

Théorie et Pratique de la Formation Professionnelle de Lexicographes : Treize Ans d'Expérience Lilloise

Pierre Corbin

Université Charles de Gaulle - Lille III / SILEX - UMR 8528 du CNRS

B.P. 149

59653 VILLENEUVE D'ASCQ Cedex

FRANCE

pierre.corbin@libertysurf.fr

Resumé

Cette contribution relate l'expérience (histoire, principes, méthodes) de treize ans de formation professionnelle de lexicographes à l'Université de Lille III, d'abord dans le cadre du "Diplôme Européen de Lexicographie", puis dans celui du D.E.S.S. "Lexicographie, Terminographie et Traitement Automatique des Corpus".

1. Former des lexicographes dans un contexte éditorial en mutation

Le présent article saisit l'opportunité offerte par EURALEX 2004 pour présenter aux spécialistes de tous pays les options et les méthodes de l'expérience originale de formation de lexicographes qui se poursuit à l'Université de Lille III depuis treize ans : bien repérée en France, en particulier par les relations suivies qu'elle entretient avec les professionnels de l'édition dictionnaire¹ et des industries de la langue, mais sans pour autant que l'articulation de l'ensemble de ses spécificités soit nécessairement appréhendée par ceux-ci, elle bénéficie d'un écho international tangible, notamment via le recrutement de ses étudiants, mais néanmoins circonscrit de façon imprécise et variable selon les pays, que cette contribution pourra permettre d'amplifier, notamment auprès de la communauté anglophone.

L'intérêt de relater l'expérience lilloise tient à son originalité, qui réside d'abord dans son existence même, les instances stables de formation de lexicographes n'étant sans doute pas légion dans le monde, et ensuite dans ses modalités, qui peuvent se caractériser comme des successions d'équilibres entre des principes retravaillés en permanence et des contingences de diverses natures tenant aussi bien aux conditions matérielles de nos enseignements qu'aux évolutions technologiques et aux fluctuations de l'édition et du marché dictionnaires. Le présent compte rendu, qui fait suite à des états intermédiaires partiels², ne prétend pas proposer un modèle qui serait transposable dans d'autres contextes, mais seulement, par un témoignage circonscrit sur une pratique consolidée de formation professionnelle associant enseignement et recherche, nourrir la réflexion de ceux qui, ailleurs, empruntent d'autres voies comme de ceux qui aspireraient à développer à leur tour des cursus comparables.

Le thème choisi pour cet exposé est en outre représentatif de l'activité métalexigraphique institutionnalisée en France. Sans être très nombreuse, la communauté de ceux qui mènent des recherches dans ce domaine existe, plus ou moins disséminée, mais il n'y a que deux centres qui lui soient suffisamment dédiés pour lui donner un rayonnement

inscrit dans la durée, à des titres d'ailleurs différents et complémentaires : celui de Cergy-Pontoise, historique et culturel, créé il y a plus de dix ans par Jean Pruvost, et celui de Lille III, que j'anime depuis une vingtaine d'années, et qui est axé sur la lexicographie contemporaine envisagée sous l'angle de ses pratiques, des incidences du marché sur son évolution, de son outillage linguistique et informatique et de la formation de ses praticiens, ce dont j'entends rendre compte ici en tant qu'initiateur, pivot et mémoire de notre cursus professionnalisant, dans lequel je suis le seul à avoir été formateur de façon ininterrompue depuis sa création³.

1.1. Repères historiques

L'émergence à Lille III d'une formation professionnelle de lexicographes, qui a été rendue possible par la conjonction d'une circonstance historique déterminée par une dynamique internationale de l'intérêt pour les "dictionnaires" et de l'expérience d'une équipe de recherche en linguistique lexicale, le laboratoire SILEX, dont une partie des investissements s'étaient déployés dans le domaine de la métalexicographie, s'est produite dans le contexte très particulier d'une mutation technologique mondiale (l'informatisation) de l'activité lexicographique et d'une modification, en France, des conditions économiques de celle-ci.

1.1.1. Genèse et histoire de la formation professionnelle lilloise. Nous avons une dizaine d'années d'expérience de la recherche et de l'enseignement de la métalexicographie, par le double canal de l'activité de l'équipe SILEX à Lille III et du cursus de linguistique française que j'avais construit à l'Université de Valenciennes à partir de 1981, quand la formation s'est mise en place, au niveau bac + 5, au début des années 90, dans l'élan international d'institutionnalisation de la discipline qui fit instaurer un "Diplôme Européen de Lexicographie" (D.E.L.) simultanément à Exeter et à Amsterdam, et un peu plus tard à Barcelone.

Dans sa première phase, jusqu'en 1999, la formation, qui accueillait alors en moyenne de cinq à dix étudiants par an, fut sanctionnée par un diplôme d'université. Le volume horaire, sans être réduit, ne permettait pas de tout faire et les moyens humains favorisaient certaines orientations : la lexicographie, monolingue et bilingue, avait la part centrale, la terminographie était très marginale, l'informatique donnait lieu à une initiation de qualité mais insuffisamment développée, et la formation profitait des recherches avancées du laboratoire en linguistique, notamment en morphologie dérivationnelle et en sémantique lexicale. Si les étudiants effectuaient un stage d'au moins deux mois dans une entreprise ou une institution spécialisée, les professionnels étaient peu représentés dans l'équipe enseignante.

Depuis 1999, la formation s'inscrit dans le cadre d'un D.E.S.S. (Diplôme d'Études Supérieures Spécialisées) habilité nationalement, qui accueille désormais de dix à quinze étudiants chaque année et qui dispose de moyens horaires nettement plus substantiels, ce qui a permis un rééquilibrage des contenus, un élargissement de la plate-forme d'enseignement que traduisent ses deux intitulés successifs ("Lexicographie et Terminographie", puis, depuis 2002, "Lexicographie, Terminographie et Traitement Automatique des Corpus" (LTTAC)) et l'intégration significative de professionnels représentant divers horizons des industries de la langue dans l'équipe de formateurs, la familiarisation des étudiants avec les réalités du terrain étant également renforcée par l'extension à trois mois de la durée minimale du stage.

1.1.2. *Évolution concomitante de l'activité lexicographique française.* Notre formation s'est développée dans un contexte technologique et économique contrasté. Si l'informatisation de nos entreprises lexicographiques, qui ne s'est pas faite à marche forcée, a été un élément plutôt favorable pour nos étudiants, dont certains, à date récente, ont aussi pu trouver à s'intégrer dans des instances publiques ou privées plus spécifiquement axées sur le traitement automatique du langage, la traduction de la conjoncture économique générale dans ces deux secteurs ne créait pas les conditions les plus favorables que nous aurions pu souhaiter. À la tendance lourde de la précarisation de l'emploi, qui donne le pas au temporaire sur la stabilité, se sont ajoutées, pour les industries de la langue et la terminographie, la fragilité des start-up et le déficit de création d'emplois publics et, pour la lexicographie, la reconfiguration à épisodes des structures éditoriales, qui avait débuté vers 1985 par l'absorption de Larousse et Robert par le Groupe de la Cité et qui perdure aujourd'hui avec l'éclatement imposé par les instances européennes de l'ex-Vivendi Universal racheté par Lagardère, dont on attend encore l'épilogue. Un effet ironique de la gestion rentabiliste qui a diminué le nombre des emplois stables et, de ce fait, réduit la pratique de formation de lexicographes *in situ* semble avoir été de créer les conditions d'un relais de l'Université pour assurer celle-ci tout en induisant simultanément la possibilité pour les entreprises d'exploiter la qualification de nos diplômés et même de nos stagiaires, aptes à s'intégrer dans le processus d'informatisation en cours, pour réduire leurs coûts de production : une analyse purement économiste de cette concomitance pourrait probablement conclure, sans caricature, que notre formation professionnelle, dont la mise en œuvre a eu pour moteur l'intérêt spéculatif ordinaire d'universitaires pour une discipline, s'est trouvée en adéquation avec les intérêts financiers de grands groupes industriels.

Sur le plan proprement éditorial, la toile de fond française de notre formation professionnelle était caractérisée par la restriction de la lexicographie générale à trois éditeurs, entre lesquels certains créneaux déterminés fixaient la concurrence (les "dictionnaires de langue" scolaires, de poche et de milieu de gamme ainsi que certains bilingues pour les trois, le "dictionnaire encyclopédique" monovolume de référence pour Larousse et Hachette), alors que d'autres restaient des territoires spécifiques (le haut de gamme des "dictionnaires de langue" pour Robert, le pré- et parascolaire pour Larousse, le bilingue latin-français pour Hachette). La conjoncture et l'échec commercial de produits originaux incitaient davantage au conformisme qu'à l'innovation⁴, même si, à date récente, la créativité a refait surface dans certains secteurs. L'informatique investissait le stockage des données, la gestion des catalogues, la codification de la rédaction et les supports de consultation, mais pas la constitution des ressources : le français n'ayant pas la même diffusion internationale que l'anglais, nos éditeurs n'ont pas connu l'aiguillon des "learners' dictionaries" qui a lancé leurs homologues britanniques dans la course aux corpus électroniques, et il est significatif que les rares ouvrages commercialisés en France à avoir affiché l'utilisation de cette technologie comme une plus-value soient des répertoires bilingues anglais-français coproduits avec des éditeurs britanniques à partir des ressources de ces derniers (Oxford pour Hachette, Collins pour Robert).

Hors du cadre de la lexicographie générale commerciale, par contre, le paysage était plus ouvert, qu'il s'agisse de lexicographie spécialisée ou d'agrément (une mosaïque d'éditeurs), de réalisations universitaires⁵ de référence conduites en France (ex-INaLF) ou

dans l'espace francophone limitrophe (Neuchâtel, Louvain-la-Neuve, Leuven) concernant des parlers régionaux ou le français langue étrangère, ou de rétroconversion de répertoires anciens en ligne ou sur CD-ROM (Champion, Redon).

1.2. Constantes

Si notre pratique didactique a connu et connaîtra encore des évolutions significatives sous l'effet conjugué des modifications du contexte professionnel et de la progression de notre réflexion, trois principes stables gouvernent nos options de formation depuis les origines :

1.2.1. La définition d'un profil professionnel. Cinq ordres de connaissances constituent les atouts maîtres de qui aspire aujourd'hui à faire carrière en lexicographie et dans des secteurs connexes des industries de la langue : (i) une maîtrise accomplie de sa langue maternelle et d'au moins une autre langue, assortie d'une intuition métalinguistique fine ; (ii) un bagage solide en linguistique ; (iii) une connaissance théorique et pratique aboutie des méthodes et techniques lexicographiques et terminographiques ; (iv) de l'aisance dans le maniement des outils informatiques les plus adaptés aux manipulations de tous ordres effectuaibles sur des matériaux langagiers ; (v) une vue globale des métiers et de la vie de l'édition, dans leurs dimensions techniques et économiques.

Ce portrait idéal⁶, à laquelle aucune de nos filières universitaires ne prépare intégralement et dont notre diplôme ne peut pas prendre en charge toutes les composantes, guide nos arbitrages dans l'articulation du recrutement de nos étudiants et du choix de nos contenus pédagogiques en fonction de nos moyens : la formation est ouverte à une large pluralité de cursus d'origine (sciences du langage, langues vivantes, français langue étrangère, lettres modernes ou classiques, philosophie, auxquelles s'ajoutent les parcours des étudiants étrangers francisants dans leurs pays d'origine), que nous panachons au mieux au sein de chaque promotion afin de personnaliser les profils professionnels individuels, notamment en fonction de la variété des langues pratiquées, pour lesquelles nous n'assumons aucun enseignement. L'implication du caractère nécessairement composite de chaque promotion est l'impossibilité de présupposer un niveau de connaissances partagées dans les matières sur lesquelles nous concentrons nos efforts (cf. *supra* § 1.1.1.) et dans lesquelles nous devons mener nos étudiants le plus loin possible en un temps réduit, ce qui fait de leur motivation, telle qu'elle s'exprime dans les dossiers de candidature et les entretiens que nous avons avec les candidats, un critère de sélection important.

1.2.2. Une équipe de formateurs soudée. L'organisation de notre équipe de formateurs est originale par rapport à celle d'autres formations de même niveau, qui souvent juxtaposent une mosaïque d'interventions ponctuelles sans réelle articulation. Nous privilégions la cohésion et la solidité : un noyau lillois coordonné assure en coprésence l'essentiel des enseignements fondamentaux en lexicographie, en linguistique et en informatique, et des intervenants extérieurs, universitaires et professionnels du privé, avec lesquels nous travaillons en coopération, nous font profiter plus ponctuellement de leur expertise spécifique⁷. Cette manière de procéder, dont le coût énergétique est élevé pour les responsables de la formation qui sont la cheville ouvrière de ce dispositif, a pour effets bénéfiques irremplaçables de dynamiser les enseignements par leur fertilisation réciproque et de faciliter l'assimilation par les étudiants de l'interconnexion des différentes disciplines

auxquelles ils sont formés, ce qui est particulièrement précieux au regard des contraintes de progression rapide évoquées précédemment.

1.2.3. L'articulation de la formation et de la recherche. La constante la plus typique de notre pratique lilloise de formation de lexicographes est l'articulation entre formation et recherche, qui se traduit dans un effort permanent pour modéliser l'activité lexicographique en prenant du recul par rapport aux différentes manières de faire et en proposer une construction théorique suffisamment distanciée et ouverte à une pluralité de conceptions et de conditions d'exercice possibles pour que la formation de nos diplômés ne circoncrive pas leur qualification à la capacité de s'acquitter de missions définies à court terme, mais les appareille durablement, par l'adaptabilité, pour l'exercice d'une gamme évolutive de métiers aux facettes nombreuses et changeantes, et laisse ouverte l'orientation vers la recherche. Cette spécificité de notre approche est l'objet du développement qui va suivre.

2. Options de formation

La présentation de nos options de formation constitue à la fois un arrêt sur image et une idéalisation, en ce qu'elle fixe en un instantané une compilation articulée de l'expérience accumulée depuis treize ans qui n'est pas intégralement transmissible sous cette forme et dans cet ordre aux étudiants d'une seule promotion et qui ne peut que connaître de nouvelles évolutions au fil des sessions à venir.

2.1. L'objet de la formation

Nous formons à l'élaboration de répertoires métalinguistiques.

2.1.1. Implications. De cette formulation de la circonscription de l'objet auquel s'applique notre formation découlent deux ordres d'implications théoriques et pratiques :

(i) Incrire dans l'ensemble des répertoires les objets à l'élaboration desquels nous préparons revient à spécifier ceux-ci, sous le seul angle de leur textualité et abstraction faite de leur contenu, comme appartenant à une classe de productions discursives⁸ ayant en partage d'être des collections relativement homogènes de modules élémentaires couplant une adresse et des données, assorties d'un principe d'ordre. Cette option a deux conséquences intéressantes :

– Du point de vue théorique, elle neutralise, par la dissociation des discours et de leur matérialisation (à la différence, par exemple, de définitions des "dictionnaires" comme livres⁹ ou comme ouvrages¹⁰), les variations de support (papier ou électronique), de conditionnement (plusieurs versions sur un même support) et d'autonomie (publication indépendante ou en dépendance d'autres discours) qui peuvent advenir à un même discours, tout en mettant mieux l'accent sur le principe de consultabilité que ne le font les caractérisations des mêmes productions en termes de listes ou d'ensembles d'unités linguistiques¹¹.

– Du point de vue pratique, cette inclusion des productions discursives sur lesquelles se focalise notre formation dans l'ensemble plus vaste des répertoires élargit les perspectives professionnelles de nos étudiants, le partage de propriétés structurelles par tous ses sous-types et, parmi ceux-ci, la complexité spécifique des répertoires métalinguistiques permettant à ceux qui en ont acquis une bonne maîtrise ainsi que celle des balisages aujourd'hui utilisés pour les structurer (cf. *infra* § 2.2.5., 2.3.2. (i) et 2.3.3. (i)) de se trouver dans des conditions

favorables, moyennant une familiarisation avec des documentations particulières, pour rédiger des annuaires, des guides, des catalogues, etc.

(ii) Spécifier un ensemble de productions discursives comme constituant la classe des répertoires métalinguistiques revient simultanément à sous-catégoriser les répertoires du point de vue de leur contenu référentiel et les discours métalinguistiques du point de vue de leur organisation textuelle, en pointant comme caractéristique identitaire la division du propos en un nombre de fragments proportionné à celui des unités linguistiques qu'il paraît pertinent de sélectionner. Cette circonscription de l'intersection de deux classes de productions discursives, respectivement caractérisées par leur structure et par leur contenu, qui définit pour notre formation un objet sans ambiguïté, a trois conséquences, assumées comme des avantages en raison des clarifications qu'elles apportent sur des terrains propices à la confusion :

– On peut mesurer des écarts quantitatifs et éventuellement qualitatifs, mais il n'y a pas à poser de différence d'essence, entre, par exemple, de grands recueils très documentés dévolus à la description extensive d'un idiome et les minces répertoires qui, dans quelque publication que ce soit, apportent des informations métalinguistiques sur des items jugés importants pour la compréhension de son contenu : ce ne sont que jalons dans un continuum ressortissant au domaine couvert par notre formation. Corrélativement, (a) il n'y a pas de bénéfice opératoire à tenter de refonder théoriquement des distinctions conceptuelles pour les termes *dictionnaire*, *lexique*, *vocabulaire* et *glossaire*, dont l'usage brouillé observable en tant que dénominations de répertoires métalinguistiques est la résultante de leur histoire et s'enracine en partie dans l'origine de leurs emplois en latin médiéval puis en français¹² ; (b) la distinction domaniale et (peut-être) conceptuelle admise entre lexicographie et terminographie, que répercute l'intitulé même de notre diplôme pour la clarté de l'affichage de l'aire de formation qu'il vise à couvrir, n'apparaît pas comme corrélée à une différence de nature entre les répertoires ressortissant à l'une ou l'autre des deux pratiques¹³, que le terme *lexicographie*, spécifié respectivement par *générale* et *spécialisée*, peut suffire à subsumer.

– Toute production discursive organisée en répertoire mais de contenu non homogène est à analyser comme un amalgame de plus d'un répertoire de types différents, résultant de l'interclassement soit d'articles (par exemple entre un répertoire métalinguistique et un répertoire culturel, ou entre l'un de ceux-ci et son index intégré dans la même structure d'adressage, etc.), soit de données au sein d'articles (par exemple de données métalinguistiques et de données culturelles, etc.). Cette approche typologique de l'hétérogénéité constitutive de nombreuses productions discursives organisées en répertoires évite les écueils de l'usage mou ordinaire du terme *dictionnaire* pour référer à tout et n'importe quoi (y compris des non-répertoires) sans pour autant impliquer de s'engager, à l'instar de Rey-Debove (1971 : 19), dans une redéfinition axiomatique, qui, même si elle déjoue les risques d'aporie, court celui de ne pas peser suffisamment face à l'usage incontrôlé commun, y compris chez les professionnels et les spécialistes, ce dont on peut trouver un indice dans l'autre tentative de clarification conceptuelle et terminologique que constitue la dénomination *dictionnaire de langue*, dont l'existence même, dans le couple contrastif qu'elle forme avec *dictionnaire encyclopédique*, constitue une dénégation de la capacité de *dictionnaire* seul à remplir de façon stable la même fonction dénomminative sans pour autant abolir pleinement sa signification restreinte originelle, ce qui participe au flou

ambiant : il ne paraît, en pratique, ni moins scientifique ni plus inconfortable de qualifier avec la précision typologique souhaitable les répertoires que l'on veut discriminer quand l'exigence de rigueur s'impose, et, quand elle se fait moins pressante, de s'accommoder des flottements de l'usage banal de *dictionnaire*, un mot monolexématique de sens vague pouvant mieux servir, en certaines occasions, les besoins d'une expression fluide qu'une périphrase adéquate.

– Dans une typologie rigoureuse des productions discursives, *métalinguistique*, qui spécifie un contenu référentiel, n'appartient pas au même paradigme qu'*encyclopédique*, qui envisage tout contenu, quel qu'il soit, sous l'angle de la conjonction d'extensivité et de fiabilité du traitement qui en est proposé : chacun de ces deux termes représente une valeur d'une variable, binaire pour la première et continue pour la seconde, et les deux variables peuvent se combiner dans une classification croisée. Selon l'ampleur de sa couverture empirique et la qualité de ses développements, une production discursive métalinguistique, qu'elle se présente ou non sous les espèces d'un répertoire, pourra donc, au même titre qu'une production discursive non métalinguistique, être ou ne pas être encyclopédique¹⁴. Partant, le couplage reçu, à la fois fédérateur et oppositionnel, qui associe "dictionnaires" et "encyclopédies", aussi bien dans le catalogage de la librairie que dans les cours d'initiation pour étudiants en sciences du langage, pour référer d'une part aux répertoires totalement ou principalement métalinguistiques, d'autre part à des sommes de connaissances non métalinguistiques qui acceptent sans l'imposer l'organisation discursive du répertoire apparaît comme un artefact conceptuel sans portée théorique ni pratique pour notre formation et comme un obstacle épistémologique pour penser avec rigueur le typage des productions discursives impliquées.

2.1.2. *Modalités de mise en œuvre*. Dans les enseignements de notre D.E.S.S., qui articulent pratiques rédactionnelles et analyses textuelles, les options de principe qui viennent d'être exposées se répercutent dans l'importance particulière qui est accordée à la discrimination des types de discours qui peuvent se rencontrer dans une même production discursive, ce qui passe par l'exploration et le décryptage sémiotique d'une gamme ouverte de répertoires amalgamant un discours métalinguistique à d'autres variétés de discours. Observateurs vigilants de l'évolution du marché lexicographique français¹⁵, nous portons en conséquence une attention particulière aux hybridations nouvelles qui peuvent y apparaître¹⁶ et, plus généralement, à toutes les incidences des conditions de production des répertoires à composante métalinguistique significative sur leur conception et leur réalisation.

Initialement centrée sur la lexicographie générale, notre formation a fait, au fil des années, une place progressivement plus importante à la lexicographie de spécialité, au prorata du renforcement de notre potentiel en matière de traitement automatique des corpus (cf. *supra* § 1.1.1.), et dans ce cadre nous faisons explorer à nos étudiants des lexiques mêlant terminologie et jargon qui sont assez peu pris en compte par la lexicographie générale comme par la terminographie et qui présentent la caractéristique d'être à la fois spécialisés, en tant qu'ils sont liés à des investissements dans des activités spécifiques, et largement diffusés, du fait que celles-ci mobilisent de nombreux locuteurs à un titre ou à un autre : nous avons ainsi effectué des préenquêtes sur les lexiques du tiercé et du rugby, et commencé à construire un corpus d'oral transcrit concernant celui du football (cf. *infra* § 2.3.2.).

2.2. Une modélisation des pratiques lexicographiques

Notre formation constitue le déploiement d'une modélisation des pratiques lexicographiques.

2.2.1. *Motifs*. Construite progressivement dès les débuts du D.E.L., amplifiée et approfondie année après année au fil des promotions d'étudiants, cette modélisation toujours en évolution résulte des effets conjugués d'un choix et d'une contrainte :

– À la fois en raison d'une option assumée de théorisation et pour prendre en compte la diversité des situations concrètes très variables auxquelles nos étudiants de diverses nationalités sont susceptibles d'être confrontés concernant leurs idiomes respectifs, nous nous efforçons de couvrir la gamme la plus large de pratiques lexicographiques, de celles qui s'appuient sur les technologies les plus pointues à celles qui s'exercent encore dans une plus ou moins grande précarité¹⁷, sans indexation sur une manière de faire particulière.

– Nous ne sommes pas en prise directe avec les différentes pratiques effectives des lexicographes : notre connaissance consolidée des milieux de la lexicographie et du traitement automatique du langage, nos interactions avec nos partenaires professionnels de divers bords et les informations que nous apportent les rapports de stages de nos étudiants dans les limites du respect de la confidentialité nous permettent de construire une représentation approchée de diverses façons de procéder, mais nous ne disposons ni de l'ensemble des ressources exploitables ici ou là, ni de tous les logiciels utilisés par les uns ou les autres (certains pouvant d'ailleurs être développés *in situ*) et, n'étant pas engagés dans un processus de production de répertoires métalinguistiques soumis à des contraintes de productivité, nous n'avons pas développé de routines de travail et notre gestion du temps est tout autre que celle des instances éditoriales. Nous ne pouvons donc pas prétendre reproduire en laboratoire l'effectivité des pratiques rédactionnelles, qu'il revient aux stages de faire découvrir à nos étudiants, mais seulement simuler certaines d'entre elles. Par contre, l'interaction de nos positionnements de chercheurs et de notre position d'observateurs privilégiés de l'activité lexicographique nous met en situation favorable pour modéliser celle-ci dans sa globalité et sa diversité.

2.2.2. *Principes*. La modélisation des pratiques lexicographiques conçoit celles-ci comme l'exécution de programmes rédactionnels élaborés pour mettre en œuvre des projets éditoriaux. Partant, elle vise à articuler une typologie des projets (§ 2.3.1.) avec une représentation des programmes qui prend la forme d'une décomposition de l'activité d'élaboration de répertoires métalinguistiques en strates articulées ordonnées selon la logique de leur succession (§ 2.3.2. à 2.3.4.), qui n'est pas nécessairement toujours respectée, voire même pas intégralement perceptible, dans les pratiques effectives (dont certaines peuvent donner lieu à des télescopes entre différentes phases), ni aisément décryptable en toutes circonstances dans les productions discursives qui en découlent, mais qui est impliquée par ce que celles-ci donnent à lire. Imprégnée d'une familiarité au long cours avec une pluralité de répertoires métalinguistiques variés, mais visant à dépasser le niveau des observables pour atteindre celui des principes d'engendrement qui leur sont sous-jacents¹⁸, cette modélisation peut être caractérisée comme empirico-déductive.

2.2.3. *Articulation*. Le modèle des programmes rédactionnels articule trois strates :

– *une strate métalinguistique* (§ 2.3.2.), où s'opère le rassemblement des données (matériaux et options de traitement métalinguistique) retenues comme contenu pertinent pour un répertoire particulier qui est supposé être la concrétisation d'un projet défini ;

– *une strate discursive* (§ 2.3.3.), où s’effectue la mise en discours des données métalinguistiques sélectionnées en fonction du programme retenu comme adéquat pour la réalisation du projet, dont le protocole rédactionnel a vocation à fournir la codification, un même ensemble de données pouvant faire l’objet d’un nombre indéfini de traitements discursifs ;

– *une strate de posttraitement* (§ 2.3.4.), où sont mises en œuvre les options retenues pour l’affichage du discours élaboré dans la strate précédente, qui sont susceptibles de remplir plusieurs fonctions, une même production discursive pouvant faire l’objet d’un nombre indéfini de traitements typographiques, et où peuvent aussi être effectuées informatiquement des réorganisations structurelles et des insertions automatiques de séquences textuelles à caractère systématique.

2.2.4. *Corrélat*. Le bénéfice majeur de cette représentation stratifiée du processus d’élaboration des répertoires métalinguistiques est la discrimination conceptuelle entre les classes d’informations sur les unités linguistiques dont leur texte est porteur, qui ressortissent à la strate métalinguistique, et les composants et dispositifs structurant les répertoires, qui relèvent de la strate discursive. Les premières et les seconds ne s’identifient pas les uns aux autres – une définition n’est pas un sens d’un item mais une formulation de ce sens parmi d’autres possibles, une contextualisation d’un item (un “exemple”) n’est pas une propriété combinatoire, flexionnelle ou autre de cet item mais la possible monstration d’une réalisation particulière de ces propriétés parmi d’autres possibles, etc. –, et leurs relations ne sont pas bijectives – une même information pouvant être traitée discursivement par plus d’un composant ou dispositif d’article, et un même composant d’article pouvant véhiculer plusieurs informations de types différents –, ce dont la production courante fournit de multiples témoignages. Cette distinction cruciale¹⁹ pour appréhender la complexité sémiotique des répertoires métalinguistiques, maîtriser la palette des ressources activables quand on est rédacteur, et simplement savoir interpréter sans contresens et avec le maximum de profit ce qu’on lit quand on est utilisateur n’est ordinairement pas mise en œuvre par les métalxicographes²⁰, même ceux qui, comme Hausmann et Wiegand (1989), poussent à un degré de finesse élevé l’analyse structurelle des “dictionnaires”²¹.

2.2.5. *Modalités de mise en œuvre*. Nous trouvons dans l’initiation de nos étudiants aux formalismes linguistiques et à l’outillage informatique qu’utilise aujourd’hui la lexicographie la plus technologisée un adjuvant efficace pour les sensibiliser à la dissociation des trois strates du modèle : les outils de constitution de corpus informatisés, choisis selon les types visés de collecte (cf. § *infra* 2.3.2. (i)) et d’investigation au sein de ceux-ci, permettent d’afficher des données discursives brutes (le “real language” cher à la lexicographie britannique), de les calibrer, et de les trier selon différents critères, mais également, d’automatiser plus encore leur analyse, en exploitant des règles linguistiques et/ou des statistiques, au moyen de lemmatiseurs et d’étiqueteurs morphosyntaxiques puis de scripts écrits en fonction d’objectifs particuliers ou de concordanciers manipulant des motifs de recherche exprimés soit par des expressions régulières, soit par des automates et des transducteurs, pour en fixer les patrons combinatoires, des bases de données pouvant être impliquées à tout moment pour stocker ces matériaux discursifs éventuellement enrichis ; les systèmes de gestion de bases de données relationnelles en langage SQL et les éditeurs de documents structurés en SGML ou en XML, déjà impliqués dans le stockage, sont également

exploitables pour la structuration et la rédaction des répertoires ; en fonction des interfaces de saisie mobilisées et d'éventuelles conversions de bases SQL en XML, les états des bases ou, préférentiellement, les feuilles de style (CSS ou XSL) donnent leur habillage typographique aux répertoires rédigés et structurés, dans lesquels des restructurations et des postinsertions peuvent avoir été effectuées en langage XSLT.

2.3. Un état du modèle

L'état actuel du modèle évolutif des pratiques lexicographiques comporte des éléments de typologisation des projets éditoriaux et une représentation des tâches propres à chacune des trois strates constitutives des programmes rédactionnels permettant la réalisation des projets (cf. *supra* § 2.2.3.).

2.3.1. Le projet. Le projet est l'élément déterminant d'une réalisation lexicographique, à la fois moteur de sa mise en route, guide de son programme et repère pour son analyse, et ce qui fonde, dans son principe, la différence entre le travail des lexicographes et celui des linguistes : les uns comme les autres élaborent des constructions métalinguistiques à partir de données langagières observées, mais, prototypiquement, alors que celles des seconds, qui peuvent se déployer librement dans l'espace de la science, ont pour critère d'évaluation leur adéquation aux données dont elles sont supposées construire une représentation, celles des premiers, déterminées par le projet éditorial qui définit leurs finalités propres, s'évaluent en fonction de leur capacité à rendre à leurs utilisateurs destinataires les services impliqués par ce projet, ce qui a comme corrélat, pour la théorie de la lexicographie, la nécessité de penser la relation complexe de cette pratique à la linguistique avec nuance et sans *a priori* réducteurs²². Curieusement, les typologies de "dictionnaires", même si elles prennent en compte des fonctionnalités des répertoires métalinguistiques (monolingues vs bilingues, généraux vs spécialisés, etc.), ne semblent pas spécifiquement axées sur la classification des projets et paraissent faire davantage de place à des descriptions immanentes (présence ou non de telle rubrique ou classe d'information). Le constat de ce manque, qui semble regrettable au regard de l'éclairage qu'un tel outil serait susceptible d'apporter sur les raisons d'être des répertoires métalinguistiques et des redistributions qu'il pourrait faire opérer dans leurs classements, m'a incité, à date relativement récente, à amorcer une typologie des projets, organisée comme une structure de traits non hiérarchique, et dont je livre ici le bâti, élaboré pour l'instant à partir de trois paramètres :

(i) *les mobiles*. La dynamique de mise en œuvre d'un projet de répertoire métalinguistique peut obéir à deux mobiles compatibles mais dissociables :

– *un mobile intrinsèque*, d'ordre intellectuel, correspondant au désir de faire exister un répertoire donné pour répondre à des besoins de consultation supposés pour une classe d'usagers hypothétiques ;

– *un mobile extrinsèque*, d'ordre commercial, reposant sur le désir de responsables d'entreprises éditoriales de réaliser des profits par la vente du répertoire envisagé, dont des études de marché peuvent servir à évaluer la plausibilité de réussite.

Ces deux mobiles, qui convergent, à des degrés divers, dans beaucoup de réalisations des maisons d'édition spécialisées des sociétés d'économie libérale, déterminent aussi, indépendamment l'un de l'autre, la genèse de différents répertoires, qu'il s'agisse de projets intellectuels sans perspectives commerciales comme le *Trésor de la langue française* ou de

projets commerciaux sans substrat intellectuel comme le *Dictionnaire de français de Larousse*²³.

(ii) *les destinations*. Par-delà l'émiettement des fonctionnalités particulières qui donnent à la diversité des répertoires métalinguistiques sa raison d'être, et dont le typage fin, éclairant en tout état de cause et déjà exploré par les typologies existantes, peut prendre appui sur leurs discours paratextuels (titres, couvertures, préfaces, catalogues...), contrôlés si nécessaire par l'analyse de leur texte (en cas d'insuffisance du paratexte ou de distorsions entre les fonctionnalités effectives d'un répertoire et celles qui sont affichées, par exemple pour certains bilingues présentés comme à la fois bidirectionnels et bifonctionnels), les finalités des productions lexicographiques semblent pouvoir être ramenées à deux grandes classes conceptuellement distinctes mais susceptibles d'être conjointes selon des équilibres variables :

– proposer des données foncièrement destinées à être réinvesties dans des pratiques linguistiques, qu'elles soient orientées vers la compréhension ou vers la production d'énoncés ;

– proposer des données dont la portée métalinguistique intrinsèque soit exclusive ou de nature à concurrencer le réinvestissement dans des pratiques linguistiques, qu'il s'agisse de diffuser des informations culturelles (étymologies, datations, citations anthologiques...) éventuellement réinvestissables dans des pratiques académiques (*Dictionnaire historique de la langue française*), ou de constituer des trésors de connaissances testimoniaux et patrimoniaux ou plus distractifs sur un état de langue et/ou une classe de discours particuliers (*Trésor de la langue française*, certains "dictionnaires d'auteur"), ou encore de viser à décrire extensivement le lexique d'un idiome selon les principes d'une théorie linguistique, avec le double objectif de valider celle-ci par l'expérimentation et de renouveler la connaissance du lexique par son application systématique à de larges ensembles lexicaux (volumes parus du *Dictionnaire explicatif et combinatoire du français contemporain* d'Igor Mel'čuk, maquettes du projet de *Dictionnaire dérivationnel du français* de Danielle Corbin²⁴).

(iii) *les destinataires*. Les destinataires des répertoires métalinguistiques, dont le ciblage est déterminant dans la gestion des projets en lexicographie commerciale, ne semblent pas avoir inspiré de typologies au même titre que ceux-là, alors même que leur affichage dans les paratextes de ces derniers fournit des matériaux de premier intérêt pour une socio-histoire de l'édition lexicographique. Les recherches concernant le domaine français actuel qui ont été engagées dans cette direction avec nos étudiants n'ont pas atteint leur maturité mais permettent déjà d'inventorier différents paramètres récurrents et de prendre la mesure de la diversité des catégorisations qui sont activées²⁵ :

– *l'âge*, qui peut s'exprimer par la mention d'années (« dès 7 ans », « 9/12 ans ») ou de classes d'âge aux frontières floues (de « tout petits » à « adultes ») transposant dans certains cas des échelles sportives (de « benjamin » à « senior ») ;

– *un niveau d'expertise* dans une échelle de progression, qui peut être (a) celle des degrés scolaires, exprimée par des noms de types d'établissements (« maternelle », « collège »), de cycles d'enseignement (« cycle 3 ») ou de classes (« le dictionnaire CP, CE »), (b) celle du niveau d'apprentissage d'une langue, dénommé génériquement (« débutants ») ou spécifié de façon plus précise (« enfants qui ne savent pas encore lire »), (c) celle des statuts d'appre-

nants (« élèves », « apprenants »), (d) celle de compétences linguistiques spécifiées (« ceux qui connaissent les langues anciennes ») ou (e) celle de la maîtrise de l'usage des dictionnaires (« si l'enfant ne savait pas consulter son dictionnaire ») ;

– *un statut d'usagers*, qu'il spécifie ceux-ci en tant que locuteurs (« un locuteur non francophone ») ou qu'intervenants dans des processus didactiques (« parents », « maîtres »), ou qu'il ne les caractérise pas (« le lecteur », « tous les publics ») ;

– *une attente*, qui peut être linguistique (« tous ceux qui souhaitent avoir de la langue l'usage le plus juste »), dictionnaire (« tous ceux qui veulent avoir sous la main un dictionnaire vraiment pratique ») ou culturelle (« toutes celles et tous ceux qui désirent apprendre »).

2.3.2. *La strate métalinguistique*. Dans la modélisation du programme qui organise la mise en œuvre d'un projet de répertoire métalinguistique, la distinction d'une strate métalinguistique et d'une strate discursive se succédant dans cet ordre est une construction théorique par rapport à la réalité de pratiques rédactionnelles communes dans lesquelles elles s'interpénètrent, les rédacteurs ayant pour tâche d'insérer des données linguistiques dans les subdivisions d'une grille qui préformate une matrice textuelle et s'affiche dans un masque de saisie. Et c'est cette intrication même du traitement métalinguistique et de sa mise en discours qui rend nécessaire leur discrimination conceptuelle, qu'il s'agisse (a) de sensibiliser nos étudiants au fait qu'une même donnée linguistique est sujette à recevoir une pluralité de mises en discours (le trait sémantique 'massif', exprimé linguistiquement dans le *Dictionnaire du français* (« LE LAIT » s.v. **lait**)²⁶, l'est métalinguistiquement dans un autre dictionnaire monolingue pour apprenants allophones avancés, britannique celui-ci, le *Collins COBUILD English Dictionary for Advanced Learners* (« N-UNCOUNT » s.v. **milk**)²⁷) ou (b) de restaurer, dans l'analyse interprétative du texte des répertoires, les messages métalinguistiques sous-jacents aux options discursives retenues, qui ne s'affichent pas toujours de façon explicite (dans le *Robert junior illustré*, les contextualisations à sujet féminin « *Flora s'est blessée à la main* » (s.v. **blessé** 1.) et « *Elle s'est cassé la jambe en faisant du ski* » (s.v. **casser** 2.), parmi d'autres informations dont elles sont porteuses, sont l'unique moyen sélectionné pour transmettre celle concernant les variations de l'accord du participe passé des verbes pronominaux en fonction de la construction de ceux-ci).

Trois étapes logiques s'enchaînent dans la strate métalinguistique :

(i) *l'accès à des données langagières* exploitables pour donner corps aux projets, qui implique la mobilisation de ressources documentaires et la mise en œuvre de stratégies de collecte et de traitement dans des conditions qui ne sont pas foncièrement différentes pour les lexicographes et pour les linguistes :

– Les ressources peuvent être soit discursives (énoncés fixés (écrit, oral enregistré) ou volatils (oral saisi au vol)), soit métalinguistiques (préexistantes (répertoires, grammaires, travaux de linguistique...) ou expérientielles et suscitées (recours du lexicographe à sa compétence de locuteur ou aux témoignages d'informateurs)). L'importance relative de l'utilisation (indépendante ou combinée) de ces types de ressources dans les pratiques lexicographiques est sujette à de grandes variations en fonction de leur disponibilité respective, de leur commodité d'utilisation, de la valeur symbolique qui leur est reconnue (poids de l'écrit, notamment "littéraire", dans la culture française), de la nature des projets, des conditions économiques et technologiques du travail des lexicographes, de la qualité de

leur sens linguistique et du degré de confiance accordé à chacun des moyens qui peuvent être mis en œuvre.

– La collecte de données peut s’effectuer à partir de ressources associées anarchiquement à raison de leur accessibilité ou sélectionnées méthodiquement en fonction d’un projet ou d’un ensemble de projets, et exploiter soit sélectivement soit compilatoirement celles qui sont utilisées, les deux valeurs de chaque paramètre étant combinables avec celles de l’autre pour dessiner un continuum de stratégies, toutes effectives, qui s’échelonne de l’anarchique avec sélection (ressources aléatoires donnant lieu à des prélèvements) au méthodique avec compilation (corpus structurés informatisés), en passant par l’anarchique avec compilation (le web ou tel(s) CD-ROM de presse comme corpus) et le méthodique avec sélection (ressources ciblées donnant lieu à des prélèvements).

– Le traitement des données, en fonction des options de collecte, se distribue, au gré de la diffusion des évolutions technologiques, entre les classiques fiches dédiées aux relevés sélectifs, relayées désormais par des bases de données, et, pour les procédures les plus technicisées, des interactions complexes entre dispositifs de stockage et outils d’investigation.

Dans la formation de nos étudiants, nous nous efforçons de couvrir le plus largement possible cet éventail de pratiques, des moins outillées aux plus technologiques : les premières parce qu’elles sont loin d’être abolies de toute activité lexicographique, y compris chez nos éditeurs les plus notoires, où, par exemple, s’effectue toujours une veille néologique par dépouillement sélectif de ressources (notamment des périodiques), et parce qu’elles exercent le sens linguistique et le savoir-faire lexicographique ; les secondes parce qu’elles familiarisent avec la maîtrise d’outils informatiques utiles professionnellement et qu’elles permettent de prendre une première mesure des apports mais aussi des limites de cette instrumentation qui investit progressivement la lexicographie et les autres disciplines linguistiques²⁸ À l’écart des partis pris et de discours idéologisés qui donnent une tonalité nouvelle, quelques décennies plus tard, à un vieux débat sur les corpus²⁹, nous lançons chaque année nos étudiants dans des réalisations collectives originales mettant en œuvre un segment ou un autre de la gamme des possibles : de 2000 à 2002, ce fut l’élaboration d’une base de données relationnelle leur permettant de stocker des relevés sélectifs de “néologismes” effectués “manuellement” dans le magazine *Elle* ; en 2001-2002 l’élaboration de bases de données lexicales exploitant, dans divers domaines de spécialités, des corpus constitués de pages web ; en 2002-2003, celle d’un corpus XMLisé, conforme aux principes de la TEI, de transcriptions de retransmissions radiodiffusées de matchs de football, pour tester nos hypothèses sur l’intérêt de corpus d’oral à haut rendement³⁰ ; enfin, en 2003-2004, la création d’une DTD pour la rédaction, dans un éditeur XML, d’un répertoire des termes techniques du D.E.S.S. prenant comme ressource documentaire les enseignements dispensés par les formateurs, puis la conversion de ce répertoire en base de données pour susciter la réflexion sur les différences de logique de structuration entre ces deux outils de stockage codifié.

(ii) *la sélection des données* (items et propriétés de ceux-ci) à faire figurer dans un répertoire particulier, qui est le produit de plusieurs niveaux de filtrage articulés :

– celui de la collecte, qui circonscrit les données accessibles en fonction des paramètres qui viennent d’être évoqués, auxquelles sont appliqués les autres filtres ;

– celui de l'appareillage conceptuel et technique utilisé pour interpréter les données brutes de la collecte, en isolant des items et en identifiant des propriétés de ceux-ci : différentes procédures d'investigation ne produisent pas la même liste de candidats termes à partir d'un même corpus ; toutes les méthodes de segmentation morphologique n'isolent pas les mêmes éléments (par exemple *simul-* dans *simulacre*, comme le fait le *Robert méthodique*, bien que la finale *-acre* ne soit pas un morphème) ; toutes les analyses syntaxiques ne débouchent pas sur les mêmes patrons (par exemple la distinction de deux constructions différentes de *décider de* selon la nature nominale (*décider du sort de quelqu'un*) ou infinitive (*décider de venir*) du complément introduit par *de*, qu'opère le *Trésor de la langue française*) ; etc. ;

– celui du projet, qui détermine une configuration caractérisée par des principes quantitatifs (un ordre de grandeur) et qualitatifs (choix d'items, de types d'informations et, au sein de ceux-ci, choix d'informations pour chaque item), dont la mise en œuvre est susceptible d'activer une pluralité de critères : ceux de l'analyse linguistique (par exemple par la sélection ou l'éviction de classes d'items), ceux de l'appréciation des usages (par la fréquence, mesurée ou intuitive, la variation dans ses multiples dimensions (spatiale, temporelle, registrale, sociolectale, domaniale...), le statut institutionnel (normes, recommandations officielles), l'attestation dans certains dictionnaires³¹), des facteurs culturels (incidence de la diffusion des *realia* (tartiflette, DVD...) sur la sélection lexicale), des positions éthiques (options concernant les usages offensants), des obstacles juridiques (questions de propriété lexicale), des choix par rapport à des habitudes conceptuelles (présence ou non de noms propres dans une nomenclature "de langue") ou à des échelles de structuration et d'homogénéité (du plus aléatoire au plus cohérent³²), d'autres encore sans doute.

(iii) *le traitement métalinguistique des données sélectionnées*, qui est fonction du projet et des savoirs linguistiques accessibles aux lexicographes. Si les répertoires métalinguistiques expérimentaux destinés à un public averti qui appliquent un modèle théorique à une certaine échelle (par exemple le *Dictionnaire explicatif et combinatoire*, cf. *supra* § 2.3.1.), qui sont l'exception, ont vocation à déployer tout l'appareil conceptuel de celui-ci, l'articulation des besoins et des moyens est à penser en d'autres termes pour l'essentiel de la production :

– Les besoins en outils conceptuels plus affinés que ceux des connaissances métalinguistiques ambiantes (c'est-à-dire celles véhiculées par les enseignements scolaires de masse et les matériels didactiques qui les accompagnent et les soutiennent) varient d'une part en fonction du profilage des répertoires (tendanciellement, plus ceux-ci sont l'expression de projets correspondant à des ambitions modestes, moins une plus-value métalinguistique a de pertinence), d'autre part en fonction de la nature des unités linguistiques à traiter (la technicité du traitement des unités du lexique constitue une gradation, en haut de laquelle figure en particulier celui des unités à sens instructionnel (mots "grammaticaux", affixes))³³.

– Concernant les moyens, il convient de distinguer ce qui est souhaitable pour la qualification des lexicographes et ce qu'il est opportun d'activer dans l'élaboration de répertoires métalinguistiques particuliers. Dans le débat concernant la première, où certains, notamment dans le domaine anglophone, ont formulé des réticences marquées quant à l'intérêt pour les lexicographes d'avoir une expérience de linguistes³⁴, nous considérons au contraire, au diapason de ceux qui ont écrit les grandes pages de la lexicographie et de la

métaléxicographie françaises des dernières décennies, qu'une telle formation, pour les outils qu'elle fournit et les postures d'observation et d'analyse qu'elle installe, ne saurait jamais être trop poussée³⁵ et nous déplorons que les contraintes horaires et l'hétérogénéité des parcours antérieurs de nos étudiants nous limitent à une sensibilisation méthodologique, que nous axons prioritairement sur la morphologie constructionnelle, les structures argumentales, la sémantique lexicale et les formalismes. Par contre, la pluralité des modèles linguistiques, leur inégale valeur, et le caractère par essence mouvant de la recherche scientifique, peut-être amplifié par la relative jeunesse de la discipline, suggèrent la prudence en matière d'application de théories linguistiques en lexicographie, même quand elle n'induit pas de difficultés de métalangage (cf. *infra* § 2.3.3. (ii)) : si les analyses de la combinatoire verbale développées sous la direction de Maurice Gross au L.A.D.L.³⁶ ont trouvé à s'intégrer harmonieusement et discrètement dans la matrice classique des articles du *Grand dictionnaire encyclopédique Larousse*, le traitement des mots construits par des transformations de phrases selon le modèle initial de la grammaire générative transformationnelle dans le *Dictionnaire du français langue étrangère* ou d'après les principes distributionnels d'Eugene Nida dans le *Robert méthodique* rencontre trop de problèmes d'adéquation descriptive pour susciter une totale adhésion.

2.3.3. *La strate discursive*. Dans la strate discursive, le traitement métalinguistique reçoit la mise en forme définie comme adéquate pour un projet donné, qui peut activer différents vecteurs (en plus du texte, l'image fixe et, sur support électronique, l'image mobile et le son) et qui met en œuvre des opérations de structuration et de rédaction :

(i) *la structuration des répertoires métalinguistiques*. Un répertoire métalinguistique est une architecture de structures plus ou moins complexe, dont le principe de complexification, fondé sur des combinaisons d'optionnalisations, d'itérations, de mutualisations, d'effacements, de déplacements et d'enchaînements opérables à différents niveaux, génère des écarts entre la structuration effective et ce qui s'en donne à voir aux utilisateurs, dont il est possible de rendre compte au moyen de dispositifs descriptifs comparables à ceux de grammaires formalisées, qui agiraient sur des éléments dont les balises pourraient intégrer des attributs. La structure fondamentale théorique est l'unité de traitement, constituée d'une occurrence de la forme d'adressage choisie pour l'item à une des valeurs duquel il s'agit d'associer des données, et d'une séquence des composants dévolus dans le programme du répertoire à des fonctionnalités spécifiques, eux-mêmes susceptibles d'être décomposés en éléments et soumis à des contraintes de position relative, d'optionnalité et d'itérabilité. L'assemblage de plusieurs unités de traitement se déploie dans les deux dimensions de la hiérarchie et de la linéarité :

– elles peuvent être regroupées en strates éventuellement distribuées sur plusieurs niveaux hiérarchisés matérialisables selon différents procédés, celles du niveau le plus élevé étant elles-mêmes regroupables en articles, à leur tour susceptibles d'être rassemblés en blocs d'articles qui constituent, avec les articles non regroupés, les unités de consultation du répertoire, dont les modalités d'organisation conditionnent les modes d'utilisation de celui-ci ;

– les unités constitutives de chaque niveau hiérarchique sont enchaînées linéairement, leur ordre de succession pouvant être déterminé par des propriétés des données traitées, et donc être porteur d'informations, ou ne relever que d'options d'affichage conventionnelles (ce qui

vaut aussi, au niveau supérieur, pour l'organisation des articles et/ou blocs d'articles en une structure de consultation).

Un exemple illustrera, de façon simplifiée, l'ensemble de ce mécanisme. Soit le dispositif textuel associé à l'adresse **multiple** du *Larousse pratique* :

« **multiple** adj. (du lat. *multiplex*, qui a beaucoup de plis). 1 Qui se produit de nombreuses fois : *Ce magnétoscope a été réparé à de multiples reprises* (syn. nombreux). 2 Qui est composé de plusieurs parties : *Une prise multiple* (= sur laquelle on peut brancher plusieurs appareils). 3 Qui se présente sous des aspects nombreux et variés : *Nous vous proposons de multiples solutions*. ☐ *Grossesse multiple*, donnant naissance à deux enfants ou plus. ◆ n.m. Nombre entier qui contient un autre nombre entier plusieurs fois exactement : *12 est un multiple de 4*. ☐ *Plus petit commun multiple* ou *P.P.C.M.*, le plus petit des multiples communs à plusieurs nombres. »

Il peut être engendré à partir du patron structurel suivant (suffisant pour l'exemple présenté mais qui ne reflète pas la complexité globale du répertoire)³⁷ :

[composant d'adressage]₁ [composant de catégorisation?]₂ [composant étymologique?]₃ [composant définitionnel]₄ [composant de contextualisation?]₅ [composant de glose de contextualisation?]₆ [composant synonymique?]₇

Ce patron génère sept unités de traitement (transcrites ici sans leur codification typographique, qui intervient dans une étape ultérieure, cf. *infra* § 2.3.4.)³⁸ :

(a) [**multiple**]₁ [adj.]₂ [du lat. multiplex, qui a beaucoup de plis]₃ [qui se produit de nombreuses fois]₄ [ce magnétoscope a été réparé à de multiples reprises]₅ [nombreux]₇

(b) [**multiple**]₁ [adj.]₂ [du lat. multiplex, qui a beaucoup de plis]₃ [qui est composé de plusieurs parties]₄ [une prise multiple]₅ [sur laquelle on peut brancher plusieurs appareils]₆

(c) [**multiple**]₁ [adj.]₂ [du lat. multiplex, qui a beaucoup de plis]₃ [qui se présente sous des aspects nombreux et variés]₄ [nous vous proposons de multiples solutions]₅

(d) [grossesse multiple]₁ [donnant naissance à deux enfants ou plus]₄

(e) [**multiple**]₁ [n.m.]₂ [nombre entier qui contient un autre nombre entier plusieurs fois exactement]₄ [12 est un multiple de 4]₅

(f) [plus petit commun multiple]₁ [le plus petit des multiples communs à plusieurs nombres]₄

(g) [P.P.C.M.]₁ [le plus petit des multiples communs à plusieurs nombres]₄

Ces unités de traitement sont articulées en quatre sous-articles regroupant ce qui a trait respectivement à l'adjectif *multiple* (a-c), au nom *grossesse multiple* (d), au nom *multiple* (e) et aux noms *plus petit commun multiple* et *P.P.C.M.* (f-g). Les sous-articles (a-c) et (d) d'une part, (e) et (f-g) de l'autre sont à leur tour regroupés en deux articles, afférents respectivement aux items impliquant l'adjectif *multiple* (a-d) et le nom *multiple* (e-g). Enfin, ces deux articles sont regroupés dans un bloc unique (a-g) réunissant tout ce qui concerne *multiple* quelle que soit sa catégorie, qui constitue une unité de consultation autonome dont le repère alphabétique est l'adresse du premier de ses articles. L'ensemble de ces opérations d'assemblage s'accompagne d'une mutualisation des segments redondants dans les positions adéquates (avant le composant 4 de (a) pour les composants 1, 2 et 3 de (b) et (c), avant le composant 2 de (a) pour le composant 1 de (e), après le composant 1 de (g) pour le composant 4 de (f)), ce qui donne au bloc d'articles la structure de surface suivante :

- bloc d'articles afférent à un ensemble d'items ayant en commun l'élément *multiple*
 - article afférent à un ensemble d'items ayant en commun l'adjectif *multiple*
 - sous-article afférent à l'adjectif *multiple*
 - zone de mutualisation de (a), (b) et (c) : ([composant d'adressage]₁ [composant de catégorisation?]₂ [composant étymologique?]₃)_(a-c)
 - zone de traitement spécifique de (a), (b) et (c) : ([composant définitionnel]₄ [composant de contextualisation?]₅ [composant synonymique?]₇)_(a) ([composant définitionnel]₄ [composant de contextualisation?]₅ [composant de glose de contextualisation?]₆)_(b) ([composant définitionnel]₄ [composant de contextualisation?]₅)_(c)
 - sous-article afférent à *grossesse multiple*
 - zone de traitement spécifique de (d) : ([composant d'adressage]₁ [composant définitionnel]₄)_(d)
 - article afférent à un ensemble d'items ayant en commun le nom *multiple*
 - sous-article afférent au nom *multiple*
 - zone de traitement spécifique de (e) : ([composant définitionnel]₄ [composant de contextualisation?]₅)_(e)
 - sous-article afférent à *plus petit commun multiple* et *P.P.C.M.*
 - zone de traitement spécifique de (f) et (g) : [composant d'adressage]_(f) [composant d'adressage]_(g)
 - zone de mutualisation de (f) et (g) : [composant définitionnel]_(f-g)

Il reste alors à la première, c'est-à-dire n'exploitant que des ressources documentaires (cf. *supra* § 2.3.2. (i)) sans utiliser de matrice textuelle d'options, pour partie prises en charge par les protocoles rédactionnels qui ont vocation à homogénéiser l'écriture des répertoires, et dépendant pour le reste de l'inspiration des lexicographes :

- décisions quant aux formes à retenir comme adresses : accumulées ou sélectionnées, développées ou abrégées, représentant des occurrences actualisées d'items ou subsumant leur paradigme, isomorphes à des formes d'items ou postiches ou de pure commodité (un échantillon d'adresses utilisées pour accueillir le traitement de *au fur et à mesure* (Ø + de + que) ne donne qu'un aperçu de la gamme des possibles : **au fur et à mesure** dans le *Robert Micro*, **fur et à mesure (au)** dans le *Petit Larousse illustré*, **fur et à mesure** dans le *Grand dictionnaire Larousse-Chambers*, **fur** dans le *Dictionnaire Hachette langue française*, etc.) ;
- décisions quant au métalangage : en utiliser ou le contourner (cf. *supra* § 2.3.2. (iii)) et, dans le premier cas, le sélectionner (en fonction d'hypothèses sur celui qui est accessible au lectorat visé) et retenir un principe d'affichage (*in extenso* ou sous forme abrégée) ;
- décisions quant au style définitionnel : définitions analytiques classiques (**fatigué** : « Dont l'activité est diminuée à la suite d'un effort excessif » (*Dictionnaire du français*, 1.)) ou expérientielles (**tired** : « If you are tired, you feel that you want to rest or sleep » (*Collins COBUILD*, 1), recourant à un lexique restreint ou assumant une complexité ;
- décisions quant aux contextualisations : citer (des références "autorisées" ou des extraits bruts de corpus), forger, adapter³⁹ ;
- etc..

Par ailleurs, la non-bijectivité des relations entre classes d'informations et composants d'articles (cf. *supra* § 2.2.4.), qui permet que certaines informations puissent être prises en

charge par plus d'un composant et que certains composants puissent servir de support à des informations de type différent, donne un certain jeu au processus de rédaction, qui peut, le cas échéant, être exploité avec une relative systématisme pour véhiculer différents contenus en surinvestissant certains composants : les contextualisations, notamment, donnent la possibilité d'exhiber des formes fléchies, des constructions, des cooccurrents, mais aussi des informations référentielles, voire de faire circuler des messages idéologisés.

La complexité des opérations qui s'effectuent dans la strate discursive, l'importance de ce qui s'y joue pour la consultabilité et l'interprétabilité des répertoires par leurs utilisateurs, les aspects originaux de notre approche également⁴⁰, nous amènent à mettre un accent particulier, dans le D.E.S.S. lillois, sur ce segment capital de l'activité lexicographique⁴¹ et à pousser la formation à un niveau de technicité élevé, en utilisant pour ce faire toutes les ressources offertes par les systèmes de gestion de bases de données et par les langages de structuration de documents, tant pour la création de répertoires originaux (à titre d'exercices⁴² ou dans le cadre du développement de projets professionnels) que comme outils d'analyse de répertoires existants permettant de typer avec finesse leurs composants et leurs contenus⁴³.

2.3.4. La strate de posttraitement. La strate de posttraitement prend comme entrée le produit de sortie de la strate discursive pour lui appliquer une mise en forme particulière, maintenant de plus en plus communément gérée informatiquement par une feuille de style CSS ou XSL complémentaire de la DTD, mais aussi pour y insérer de façon automatisée, via le langage XSLT, des éléments textuels récurrents, notamment ceux qui matérialisent la structuration discursive (signes alphanumériques ou diacritiques, certaines ponctuations⁴⁴) : dans le bloc d'articles **multiple** du *Larousse pratique* (cf. *supra* § 2.3.3. (i)) pourraient ainsi être postinsérés le losange, les carrés et les numériques qui en marquent la structuration, les identificateurs (« SYN. » et « = ») et les joncteurs (« ou ») de composants, et les ponctuations (points, virgules et deux points) et parenthèses qui en indiquent les frontières.

Même si, dans le cadre de rétroconversions de répertoires métalinguistiques non conçus initialement en fonction de l'informatique, le balisage du texte en éléments peut prendre appui sur les codifications typographiques, il n'en reste pas moins que, dans leur principe, celles-ci sont dissociées de la structure discursive, qu'elles peuvent donner à voir dans une gamme d'autant plus ouverte de mises en formes que les textes lexicographiques présentent à cet égard une complexité élevée. Pour autant, la typographie n'est pas strictement cantonnée dans la matérialisation de la structuration effectuée dans la strate discursive et peut dans certains cas se trouver directement porteuse de signification : par exemple quand des variations formelles au sein d'une adresse affichent une distorsion entre la séquence prise en compte au titre de l'item au sujet duquel l'article apporte des informations et celle qui est pertinente pour l'ordonnancement de ce dernier dans la macrostructure du répertoire : ainsi en va-t-il, dans le *Robert pour tous*, de l'adresse « **AU FUR ET À MESURE** », dans laquelle les capitales identifient l'item linguistique pertinent, alors que les caractères gras de corps plus grand délimitent la séquence prise en compte pour l'insertion de l'article à sa place alphabétique ; ou encore quand, comme dans le *Larousse pratique*, le corps de l'adresse du premier article d'un bloc (par exemple **multinational, e, aux**), plus grand que celui des autres lorsqu'elles sont affichées (en l'occurrence **multinationale**), signale cette adresse comme valant aussi pour le bloc dans son ensemble,

bien que toutes les données qui lui sont associées ne concernent que l'item auquel est consacré le premier article.

Il est donc important, à l'instar des formations aux métiers de l'édition, d'exercer nos étudiants à la perception du vaste ensemble de procédés de typographie et de mise en page qui peut être activé dans les réalisations lexicographiques et de les sensibiliser au rôle déterminant du posttraitement non seulement pour traduire en codes distincts les constants changements de composants et, à l'intérieur de ceux-ci, d'éléments qui sont consubstantiels aux textes lexicographiques, mais pour le faire en rendant leur affichage fonctionnel et confortable pour les usagers en dépit de la contrainte d'avoir à contenir de grandes masses textuelles dans un espace physique limité, ce qui implique d'associer la dimension esthétique, mais aussi l'astuce, à l'approche technique : le succès d'un répertoire peut en dépendre.

3. Conclusion : témoigner de ce qui fut, se projeter dans ce qui sera

Tout comme il y a plusieurs façons de faire de la lexicographie, on peut imaginer plusieurs démarches pour former des lexicographes. Il ne s'agissait ici que de porter témoignage sur une expérience originale par sa rareté, sa longévité et surtout la posture de ses animateurs, à la fois humbles et ambitieux, conscients de leurs limites, définies par ce qui ne peut s'apprendre que sur le terrain avec leurs partenaires professionnels, sensibles aussi aux incertitudes induites par les mutations du contexte éditorial, mais toujours soucieux d'innover, de se renouveler, de s'adapter, modifiant fortement leurs supports et leurs contenus d'une année sur l'autre, et qui, d'un cursus professionnalisant initialement sanctionné par un simple diplôme d'université, et aujourd'hui habilité nationalement et en cours d'intégration dans un master, ont fait un laboratoire permanent de recherche à la fois sur un domaine d'activité – la lexicographie, la terminographie et le traitement automatique des corpus – et sur la formation de ses acteurs. Il valait la peine, je crois, de laisser, même sous une forme très ramassée, une trace d'un état de cette aventure intellectuelle et humaine qui se poursuit et porte déjà en germe ses évolutions à venir, et que j'étais le seul à pouvoir ressaisir dans sa globalité.

Dédicace et remerciements

Je dédie ce texte à la mémoire de Danielle Corbin et à Nathalie Gasiglia, qui ont chacune marqué de leur empreinte une des deux phases de la formation lilloise de lexicographes, le D.E.L., plus linguistique, pour la première, et le D.E.S.S., plus informatique, pour la seconde, que je remercie en outre, ainsi que François Corbin, pour avoir bien voulu m'aider à affiner ma réflexion et me faire part de leurs observations critiques.

Notes

1. Des étudiants issus de notre centre de formation ont collaboré à de nombreux "dictionnaires" généraux de la dernière décennie publiés en France et quelquefois à l'étranger.
2. Cf. Cabré Castellví (1994) et P. Corbin (1995a ; 2002).
3. Ce statut particulier génère d'ailleurs un flottement dans le positionnement énonciatif du présent exposé, qui rend compte d'une expérience collective (cf. *infra* n. 7), mais menée dans un cadre construit à trois (d'abord avec Danielle Corbin puis avec Nathalie Gasiglia) et sous-tendue par la

vision personnelle présentée ici : il en résulte quelques emplois de *nous* à géométrie variable, en hommage à tous les artisans de la formation, bien que les conceptions que je développe n'engagent que moi.

4. Cf. P. Corbin (1998, § 2.).

5. J'y inclus celles développées au C.N.R.S.

6. Inspiré de celui, humaniste et ambitieux, de Rey (1986 : 97-99).

7. Sont intervenus ou interviennent encore dans la formation, régulièrement ou occasionnellement : Martyn Back, Henri Béjoint, Anne-Marie Berthonneau, Jean Binon, Sandrine Boulon, Didier Bourigault, Ralph Brockmeier, Estelle Bultez, M. Teresa Cabré, Laurent Catach, Danielle Corbin, François Corbin, Pierre Corbin, Pierre Cotte, Georgette Dal, Sophie David, Trea De Lange, Marie-Jeanne Derouin, Thierry Fontenelle, Bénédicte Gaillard, Nathalie Gasiglia, Michel Glatigny, Maurice Gross, Valerie Grundy, Ulrich Heid, Bruno Hemonnet, Marta Krol, Dominique Le Fur, Sylvain Loiseau, Claude de Loupy, Philippe Monnier, Raphaëlle Mourey, Widad Mustafa el Hadi, Hans Paulussen, Thierry Poibeau, Henri Portine, Pollet Samvellian, Thierry Selva, Jean Senellart, Max Silberstein, Agnès Tutin, Danièle Van de Velde, Marleen Van Peteghem, Serge Verlinde, Jean Véronis, Chantal Wionet, Franck Wisniewski.

8. Cf. Dubois (1970 : 35), Rey (1982 : 9).

9. Par exemple Landau (1989 : 5).

10. Par exemple Rey-Debove (1971 : 27).

11. Par exemple *The New Encyclopaedia Britannica*, art. 'Encyclopaedias and Dictionaries'.

12. Cf. Weijers (1991 : 43-52).

13. Cette position croise, par exemple, celle de Van Campenhoudt (2000 ; 2001 : 182).

14. L'encyclopédisme des "dictionnaires de langue" peut se lire en filigrane dans Rey (1982 : 8-9), mais sans que l'auteur esquisse une révision critique des dénominations en vigueur.

15. Cf. P. Corbin (1991 ; 1998).

16. À cet égard, la prolifération des ouvrages pré- et parascolaires qui associent, dans des combinaisons librement renouvelées, des dispositifs dictionnaires, imagiers, abécédaires, narratifs et autres offre un terrain d'investigation particulièrement intéressant (cf. P. Corbin, 2001).

17. Ce dont les journées de lexicographie bilingue organisées périodiquement par l'INALCO apportent des témoignages récurrents.

18. Notre modélisation renoue avec l'esprit du « modèle génétique théorique conçu pour décrire toute activité lexicographique » de Rey (1970 : 53) et partage avec celui-ci une dimension algorithmique, mais elle en diffère dans son organisation comme dans son objet, en ce qu'elle s'attache aux processus d'engendrement des répertoires métalinguistiques et non à la typologie de ces derniers.

19. Déjà formulée dans P. Corbin (1989, § 2.2.3.).

20. Exception faite pour Atkins (1996, § 1.1.), dans une perspective plus pragmatique que théorisante, et sans préjuger de ce que seront les options définitives de la révision en cours de la norme ISO 1951 ('Présentation / représentation des entrées dans les dictionnaires').

21. Cf. P. Corbin (à paraître a).

22. Cf. P. Corbin (1995a ; 2002).

23. Rapidement extrait, en 1997, du *Petit dictionnaire français* de 1996 dans un contexte de concurrence conflictuelle avec Maxilivres, qui commercialisait depuis 1995 un très comparable *Dictionnaire de la langue française* (cf. P. Corbin, 1998, § 2. et n. 41).

24. Cf. D. Corbin (1990 ; 1991 ; 1993 ; 1997) et D. et P. Corbin (1991).

25. Les exemples donnés n'ont que le statut d'échantillons illustratifs.

26. Cf. 'Avant-propos', p. X.

27. Cf. 'Grammar', p. xxix.

28. Sur le développement des linguistiques de corpus, cf. Habert, Nazarenko et Salem (1997) ; sur leur outillage, cf. Habert, Fabre et Issac (1998) et Bowker et Pearson (2002)
29. Cf. P. Corbin (1980).
30. Cf. Gasiglia (à paraître a ; à paraître b ; à paraître c) et P. Corbin (à paraître b).
31. Critère assumé de complémentation de la nomenclature du *Trésor de la langue française*, cf. Imbs (1971 : XXVI-XXVII).
32. Une rétroaction d'options de traitement sur les décisions de sélection pour créer certaines cohérences pouvant aboutir à des artifices (cf. D. et P. Corbin, 1989, § 2.3.2.2.).
33. Je condense ici à l'extrême des points développés dans P. Corbin (1995a, § 2.2.2.) et (2002, § 2.).
34. Cf. P. Corbin (1995a, § 1.3.).
35. *Id.*, § 2.2.1.
36. Laboratoire d'Automatique Documentaire et Linguistique.
37. Les crochets encadrent les composants, les numéros souscrits codent leur position de base dans la séquence et le point d'interrogation indique leur caractère optionnel.
38. Le soulignage code les valeurs de composants qui, étant en relation de un à plusieurs avec celles d'autres composants, induisent des mutualisations dans l'assemblage des unités de traitement. Les identificateurs (« SYN. », « = »), joncteurs (« ou ») et délimitateurs (parenthèses, ponctuations) de composants, dont l'insertion relève du posttraitement (cf. *infra* § 3.2.4.), ne sont pas figurés ici.
39. Cf. P. Corbin (1995b) pour l'analyse de la production de contextualisations "naturelles" à partir d'extraits de textes littéraires dans le *Micro Robert*.
40. Le fait de suivre une logique propre et de construire un modèle de façon autonome ne nous empêche pas de devoir beaucoup au texte fondateur de Hausmann et Wiegand (1989) en matière d'analyse structurelle des articles de répertoires métalinguistiques (cf. *supra* § 2.2.4.), et en particulier à la distinction décisive entre ce qui, au cœur de l'article, ressortit à la « structure d'adressage » et ce qui y relève de la « microstructure » proprement dite.
41. Déjà au cœur du « programme de recherches métalexigraphiques » de P. Corbin (1989, § 2.).
42. En particulier, en 2002-2003, un dictionnaire des actions de jeu du football articulé à une ontologie et exploitant le corpus constitué (cf. *supra* § 2.3.2. (i)), ainsi que la construction d'une base de données tyant les images du *Robert Junior illustré* en connexion avec les recherches de François Corbin (2002 ; à paraître a ; à paraître b) sur l'iconographie dictionnaire, et en 2003-2004 le dictionnaire des termes de notre D.E.S.S. déjà évoqué (*ibid.*).
43. On en trouve une première expression, appliquée au *Robert & Collins Junior bilingue*, dans Corbin et Gasiglia (2004).
44. Cf. Corbin & Gasiglia (2004, § 2.2.2.3.2.).

Références

A. Répertoires métalinguistiques

- Collins COBUILD English Dictionary for Advanced Learners*. 2001. (3^e édition.) Glasgow: HarperCollins Publishers.
- Dictionnaire de français*. 1997. Paris: Larousse.
- Dictionnaire de la langue française*. 1995. Paris: Éditions de la Connaissance.
- Dictionnaire du français*. 1999. Paris: Dictionnaires Le Robert.
- Dictionnaire du français langue étrangère. Niveau 1 et Niveau 2*. 1978-1979. Paris: Librairie Larousse, 2 vol.
- Dictionnaire Hachette langue française*. 2000. Paris: Hachette.
- Dictionnaire historique de la langue française*. 1992. Paris: Dictionnaires Le Robert, 2 vol.
- Grand dictionnaire encyclopédique Larousse*. 1982-1985. Paris: Librairie Larousse, 15 vol.
- Grand dictionnaire Larousse Chambers anglais-français / français-anglais*. 1999. Paris: Larousse.
- Larousse pratique. Dictionnaire du français au quotidien*. 2003. Paris: Larousse.

- Le Micro-Robert. Dictionnaire du français primordial.* 1971. (1^e édition.) Paris: S.N.L. - Le Robert.
Le Petit Larousse illustré. 2003. (édition 2004.) Paris: Larousse.
Le Robert & Collins Junior bilingue. Dictionnaire français-anglais / anglais-français. 2002. Paris: Dictionnaires Le Robert.
Le Robert Junior illustré. 1999. (nouvelle édition entièrement revue et enrichie.) Paris: Dictionnaires Le Robert.
Le Robert méthodique. Dictionnaire méthodique du français actuel. (1^e édition.) 1982. Paris: Le Robert.
Le Robert Micro. Dictionnaire d'apprentissage de la langue française. 1998. (nouvelle édition.) Paris: Dictionnaires Le Robert.
Le Robert pour tous. 1994. Paris: Dictionnaires Le Robert.
Mel'čuk, I. et al. 1984-1988-1992-1999. *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques I, II, III et IV.* Montréal: Les Presses de l'Université de Montréal, 4 vol.
Petit dictionnaire français. 1996. Paris: Larousse.
Trésor de la langue française. Dictionnaire de la langue du XIX^e et du XX^e siècle (1789-1960). 1971-1994. Paris: Éditions du Centre National de la Recherche Scientifique (t. 1-10) / Gallimard (t. 11-16).

B. Autres références

- Atkins, B. T. S. 1996. 'Bilingual dictionaries: past, present and future' in *EURALEX'96 Proceedings*. Göteborg: Göteborg University, 515-546.
Bowker, L. et Pearson, J. 2002. *Working with Specialized Language. A practical guide to using corpora.* London and New York: Routledge.
Cabré Castellví, M. T. 1994. 'Entrevista a Danielle Corbin i Pierre Corbin, directors del Diploma Europeu de Lexicografia'. *Caplletra. Revista internacional de filologia* 17: 337-348.
Corbin, D. 1990. 'Homonymie structurelle et définition des mots construits. Vers un "dictionnaire dérivationnel"' in Centre d'études du lexique (éd.), *La définition*. Paris: Larousse, 175-192.
Corbin, D. 1991. 'La morphologie lexicale: bilan et perspectives'. *Travaux de linguistique* 23: 33-56.
Corbin, D. 1993. 'Morphologie et lexicographie: la représentation du sens dans le *Dictionnaire dérivationnel du français*' in A. Hulk, F. Meika & J. Schrotten (éds.), *Du lexique à la morphologie: du côté de chez Zwaan*. Amsterdam: Rodopi, 63-86.
Corbin, D. 1997. 'La représentation d'une "famille" de mots dans le *Dictionnaire dérivationnel du français* et ses corrélats théoriques, méthodologiques et descriptifs'. *Recherches linguistiques de Vincennes* 26: 5-37 + Supplément: I-VIII.
Corbin, D. et Corbin, P. 1989. 'Sélection et description des dérivés et des composés dans le dictionnaire monolingue' in F. J. Hausmann et al. (Hrsg.), 937-946.
Corbin, D. et Corbin, P. 1991. 'Vers le *Dictionnaire dérivationnel du français*'. *Lexique* 10: 147-161.
Corbin, F. 2002. 'Image-outil et image-spectacle: sur la dualité de l'iconographie du *Petit Larousse illustré* (millésimes 2000 et 2001)' in J. Pruvost et M. Guilpain-Giraud (dir.), *Pierre Larousse, du Grand Dictionnaire au Petit Larousse*. Paris: Honoré Champion Éditeur, 242-290.
Corbin, F. à paraître a. 'L'utilisation dictionnaire des images, du support imprimé au support électronique: l'exemple du *Robert Junior*' in Actes de la journée des dictionnaires 2000.
Corbin, F. à paraître b. 'Enquête sur un couple mal assorti: le *Robert Junior* imprimé de 1999 et son pendant électronique, vus à travers leurs iconographies'.
Corbin, P. 1980. 'De la production des données en linguistique introspective' in A.-M. Dessaux-Berthonneau (éd.), *Théories linguistiques et traditions grammaticales*. Université de Lille III: Presses Universitaires de Lille, 121-179.
Corbin, P. 1989. 'Lire les dictionnaires. Pour la constitution en champ de recherches de la lecture critique des textes lexicographiques français contemporains'. *Le français dans le monde - Recherches et applications*, numéro spécial: 31-41.
Corbin, P. 1991. 'Le maquis lexicographique. Aperçus sur l'activité lexicographique monolingue française à la fin du XX^e siècle'. *Le français aujourd'hui* 94: 6-26.
Corbin, P. 1995a. 'L'articulation entre la pratique lexicographique, la métalxicographie et la linguistique dans la formation et dans l'activité des lexicographes' in J. Pruvost (éd.), *Les*

- dictionnaires de langue. *Méthodes et contenus. La journée des dictionnaires 1994*. Université de Cergy-Pontoise: Centre de Recherche Texte / Histoire, 81-112.
- Corbin, P. 1995b. 'Le monde étrange des dictionnaires (8). Du *Petit Robert* (1967) au *Micro Robert* (1971): le recyclage des citations'. *Lexique* 12-13: 125-145.
- Corbin, P. 1998. 'La lexicographie française est-elle en panne?' in *Cicle de conferències 96-97. Lèxic, corpus i diccionaris*. Barcelona: Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra, 83-112.
- Corbin, P. 2001. 'Des imagiers aux dictionnaires (1). Cadrage d'un champ de recherche' in J. Pruvost (éd.), *Les dictionnaires de langue française. Dictionnaires d'apprentissage. Dictionnaires spécialisés de la langue. Dictionnaires de spécialité*. Paris: Honoré Champion Éditeur, 15-66.
- Corbin, P. 2002. 'Lexicographie et linguistique: une articulation difficile. L'exemple du domaine français' in F. Melka et M. C. Augusto (eds.), *De la lexicologie à la lexicographie / From Lexicology to Lexicography*. Utrecht: Utrecht Institute of Linguistics OTS, 8-37.
- Corbin, P. à paraître a. 'Lire les dictionnaires (3). Classes d'informations et composants d'articles: une approche relationnelle de l'analyse des articles'.
- Corbin, P. à paraître b. 'Des occurrences discursives aux contextualisations dictionnaires' (titre provisoire) in Actes du colloque "Entre définition et citation: l'exemple" (Klingenberg am Main, 25-27 juin 2004). Tübingen: Max Niemeyer Verlag.
- Corbin, P. et Gasiglia, N. 2004. 'Lire les dictionnaires (2). Une DTD pour le *Robert & Collins Junior bilingue*: spéculations métalexicographiques'. *Cahiers de lexicologie* 84: 135-194.
- Dubois, J. 1970. 'Dictionnaire et discours didactique'. *Langages* 19: 35-47.
- Gasiglia, N. à paraître a. 'Stratégie de constitution de corpus oraux transcrits: arguments pour un corpus plurithématique à haut rendement' in Actes des journées de linguistique de corpus de Lorient 2002.
- Gasiglia, N. à paraître b. 'Stratégie de consultation de corpus oraux transcrits: pistes méthodologiques pour l'exploration d'un corpus thématique à haut rendement' in Actes des journées de linguistique de corpus de Lorient 2003.
- Gasiglia, N. à paraître c. 'Faire coopérer deux concordanciers-analyseurs pour optimiser les extractions en corpus'. *Revue française de linguistique appliquée* 2004.1.
- Habert, B., Fabre, C. et Issac, F. 1998. *De l'écrit au numérique. Constituer, normaliser et exploiter les corpus électroniques*. Paris: InterÉditions.
- Habert, B., Nazarenko, A. et Salem, A. 1997. *Les linguistiques de corpus*. Paris: Armand Colin.
- Hausmann, F. J., Reichmann, O., Wiegand, H. E. et Zgusta, L. (Hrsg.) 1989. *Wörterbücher. Ein internationales Handbuch zur Lexicographie*. Berlin et New York: Walter de Gruyter, t. 1.
- Hausmann, F. J. et Wiegand, H. E. 1989. 'Component parts and structures of general monolingual dictionaries: a survey' in F. J. Hausmann et al. (Hrsg.), 328-360.
- Imbs, P. 1971. 'Préface' in *Trésor de la langue française*, IX-XLVII.
- Landau, S. I. 1989. *Dictionaries. The Art and Craft of Lexicography*. (2^e édition.) Cambridge: Cambridge University Press.
- Rey, A. 1970. 'Typologie génétique des dictionnaires'. *Langages* 19: 48-68.
- Rey, A. 1982. *Encyclopédies et dictionnaires*. Paris: Presses Universitaires de France.
- Rey, A. 1986. 'Training lexicographers: some problems' in R. Ilson (ed.), *Lexicography. An emerging international profession*. Manchester: Manchester University Press, 93-100.
- Rey-Debove, J. 1971. *Étude linguistique et sémiotique des dictionnaires français contemporains*. The Hague et Paris: Mouton.
- The New Encyclopaedia Britannica*. 1987. (15^e édition.). Chicago: Encyclopaedia Britannica Inc., 31 vol.
- Van Campenhoudt, M. 2000. 'De la lexicographie spécialisée à la terminographie: vers un "métadictionnaire"' in H. Béjoint et P. Thoiron (dir.), *Le sens en terminologie*. Lyon: Presses Universitaires de Lyon, 127-152.
- Van Campenhoudt, M. 2001. 'Pour une approche sémantique du terme et de ses équivalents'. *International Journal of Lexicography* 14.3: 181-209.
- Weijers, O. 1991. *Dictionnaires et répertoires au moyen âge. Une étude du vocabulaire*. Turnhout: Brepols.