

Rameau

Bilan, perspectives

Michel Mingam

Bibliothèque nationale de France
michel.mingam@bnf.fr

Vingt-cinq ans après sa naissance, Rameau est devenu le langage d'indexation le plus largement utilisé par les bibliothèques françaises: cet anniversaire peut être l'occasion de tenter un bilan et de risquer quelques perspectives, étant entendu que l'accent sera mis ici sur le langage proprement dit, les autres aspects de la question (techniques, institutionnels) n'étant abordés qu'incidemment.

Mais aborder à nouveaux frais la question du langage Rameau suppose, au préalable, qu'on admette quelques postulats.

La recherche matière conserve son utilité: bien entendu, pour l'importante catégorie des documents primaires sur support matériel dont l'accès - indirect - passe notamment par l'intermédiaire de catalogues (documents imprimés, documents audiovisuels sur support, documents iconographiques, etc.); mais sans doute également pour les documents de forme numérique, s'il est vrai, comme le pensent certains théoriciens, que l'hypothèse d'un document qui serait directement et entièrement autodescriptif participe d'une illusion.

Pour faciliter cette recherche, le recours à un langage d'indexation contrôlé reste pertinent, même dans un contexte informatique, en raison de certaines difficultés que la recherche par mots du titre ou la recherche « plein texte », malgré son indéniable intérêt, ne permettent pas de résoudre: documents non textuels (sons, images); documents rédigés dans des langues étrangères; titres ne comportant pas de mots significatifs; enfin, et peut-être surtout, ambiguïté inhérente au langage naturel, laquelle impose que soient réduits autant que possible les cas d'homonymie (« vol »: aéronautique ou qualifié?), sauf à engendrer du bruit en mêlant des informations hétérogènes, et les

Titulaire d'une maîtrise de lettres, **Michel Mingam** est responsable du Centre national Rameau au sein du Service de coopération bibliographique de l'Agence bibliographique nationale à la Bibliothèque nationale de France. Il a exercé auparavant en bibliothèque universitaire (Besançon et Paris VIII).

cas de synonymie (« tiers-monde » ou « pays en développement »?), sauf à engendrer du silence en dispersant un même type d'information entre des mots ou des expressions différents.

La coopération entre bibliothèques au sein de réseaux informatisés débouchant sur l'échange de données et l'élaboration de catalogues collectifs implique l'adoption d'un langage d'indexation commun: ce que Rameau est devenu, dans une large mesure, à l'échelle nationale.

Rappels

Le langage Rameau marque à la fois, paradoxalement, une rupture mais aussi une continuité par rapport aux usages antérieurs: d'où l'intérêt de revenir brièvement sur les circonstances de sa naissance, dont les conséquences se font encore sentir.

L'indexation matière

Jusqu'aux années 1970-1980, les bibliothèques alimentent leurs catalogues alphabétiques de matières sur fiches en appliquant des principes codifiés, en 1986, par la norme Afnor NF Z 44-070 : *Indexation analytique par matière*.

Cette norme définit un ensemble de règles d'indexation: utilisation du langage naturel sous une forme normalisée (usage du français, règle du singulier, etc.); description analytique des documents épousant au plus près leur niveau de précision; recours à un système de renvois d'exclusion (« voir ») pour corriger les synonymies et d'orientation (« voir aussi ») pour pallier la dispersion propre au catalogue analytique; usage de l'ordre

alphabétique pour classer les vedettes dans les catalogues, à la manière d'un index ou d'un dictionnaire, car c'est un ordre universellement connu des utilisateurs, dont la souplesse facilite en outre la mise à jour des catalogues par l'intercalation de nouvelles fiches; conséquence du classement alphabétique, adoption d'un langage précoordonné selon une syntaxe précisant l'ordre des subdivisions (de sujet, géographique, chronologique, de forme) après la vedette.

On remarquera que ces règles ont un caractère formel dans la mesure où elles indiquent comment établir une vedette matière, en particulier sur un plan syntaxique, mais non quelle vedette choisir: la norme ne concerne pas le contenu du vocabulaire lui-même. Dès lors qu'elles respectent les règles prescrites, les bibliothèques sont donc libres de créer leurs vedettes comme elles l'entendent, en les adaptant au degré de spécialisation de leurs fonds ainsi qu'aux intérêts et au niveau culturel de leurs publics, selon qu'il s'agit, par exemple, d'établissements scientifiques ou de lecture publique. Cette liberté encadrée par une norme ne pose aucun problème tant que les bibliothèques, fonctionnant isolément, développent des catalogues indépendants les uns des autres.

Les conséquences de l'informatisation

Dans les années 1970-1980, l'informatisation progressive des catalogues vient remettre en cause ce système.

D'une part, c'est l'occasion de reconsidérer les usages. Ainsi la Bibliothèque nationale choisit-elle alors d'ouvrir un nouveau catalogue matière sur des bases renouvelées en s'aidant désormais d'une liste d'autorité pour établir son fichier d'autorité matière, et non plus seulement de la norme NF Z 44-070, afin d'obtenir un langage d'indexation mieux contrôlé.

D'autre part, l'ambition longtemps caressée d'établir des catalogues collectifs mettant en commun les res-

sources des bibliothèques (catalogage partagé, récupération de notices) devient enfin réalisable: perspective qui intéresse tout particulièrement les bibliothèques universitaires.

Mais, pour être en mesure de coopérer pleinement, encore faut-il que les bibliothèques utilisent le même langage: l'établissement de catalogues collectifs suppose donc, au préalable, une exigence renforcée de normalisation. Marc Chauveinc l'exprime fortement dès 1982: la normalisation, écrit-il, est « *une condition sine qua non qu'il faut énoncer encore et toujours: sans harmonisation, donc sans normalisation, il ne peut y avoir de catalogue collectif; les bibliothèques désireuses de partager leurs moyens doivent s'astreindre à respecter une seule règle de catalogage, bonne ou mauvaise, peu importe, pourvu que ce soit la même et que toutes parlent le même langage et décrivent de la même façon le même livre*¹ ». ».

Cette exigence accrue de normalisation s'applique naturellement aussi à l'indexation matière, car « *c'est s'arrêter au milieu du gué que de partager le seul travail de signalement, en excluant l'indexation du champ de la coopération. Mais ceci suppose que les partenaires d'un réseau pratiquent le même langage d'indexation et non des patois plus ou moins proches les uns des autres*²... ». Or, pour y parvenir, l'une des solutions les plus évidentes consiste à étendre le champ de la normalisation au contenu de l'indexation lui-même, c'est-à-dire au choix des vedettes: en d'autres termes, les bibliothèques devront désormais consentir à perdre leur liberté de création des vedettes au profit d'une liste d'autorité matière commune, de type encyclopédique, où elles viendront puiser les mêmes vedettes pour désigner les mêmes contenus.

1. Marc Chauveinc, *Le réseau bibliographique informatisé et l'accès au document*, Éditions d'organisation, 1982.

2. Étienne Hustache, Annie Meyer et Christine Stotzenbach, « Salam Lamech », *BBF*, n° 3, 1986.

Le choix de Rameau

C'est dans ce contexte d'informatisation des bibliothèques que, à la suite de la Bibliothèque publique d'information, la Bibliothèque nationale en 1980, puis la DBMIST (Direction des bibliothèques, des musées et de l'information scientifique et technique) au nom des bibliothèques universitaires en 1983, choisissent d'adopter une liste d'autorité matière unique : appelée un temps Liste d'autorité-matière encyclopédique, collective et hiérarchisée (Lamech), elle constitue la matrice de ce qui va devenir le Répertoire d'autorité-matière encyclopédique et alphabétique unifié, c'est-à-dire Rameau, et s'inspire étroitement de la liste de la Bibliothèque de l'Université Laval de Québec (Répertoire de vedettes matières Laval ou RVM Laval), elle-même démarcation francophone de celle de la Bibliothèque du Congrès de Washington (Library of Congress Subject Headings ou LCSH).

Ce choix peut être facilement expliqué : pour quoi, en effet, prendre la peine, le temps, le risque aussi, de recommencer un travail déjà réalisé ailleurs, alors même que ce dernier bénéficie à la fois du label prestigieux de la Bibliothèque du Congrès et d'une version francophone - sans compter les perspectives de coopération non plus seulement nationales, mais internationales, ainsi ouvertes, et l'espoir de récupérer, aux moindres frais, les indexations des catalogues américains du type OCLC?

Tel était bien l'avis de la DBMIST : « Il importait à la fois d'aller vite et de réussir : d'où la nécessité de partir d'un langage documentaire existant et qui disposait de bases de départ solides. L'ensemble documentaire lié à la Bibliothèque du Congrès présentait des garanties sérieuses sous ces deux points de vue, puisque la première édition des LCSH est de 1914 et les versions françaises de la Bibliothèque de l'Université Laval remontent à la fin des années quarante³. »

Logique, ce choix n'en comportait pas moins des aspects discutables.

En premier lieu, adapter à la langue et aux réalités de notre pays un outil d'origine anglo-saxonne, même passé à travers un filtre québécois, ne va pas nécessairement de soi : il n'est donc pas surprenant que le vocabulaire Rameau ait porté - surtout au début, mais c'est encore souvent le cas - l'empreinte de cette source nord-américaine, dans sa forme mais aussi dans son contenu.

La Bibliothèque nationale, en 1980, au moment du choix initial de ce qui allait devenir Rameau, ne méconnaissait pas le problème, ainsi qu'en témoigne un article paru en 1983⁴, lequel souligne trois types de problèmes rencontrés alors : difficultés résultant des différences entre l'anglais et le français, tant au niveau du lexique qu'à celui de la syntaxe ; traductions des LCSH par Laval ne pouvant être reprises telles quelles, qu'elles soient impropres, différentes de l'usage attesté ou ambiguës ; enfin, particularités culturelles résultant de l'origine anglo-saxonne du répertoire.

Ces problèmes étaient d'autant moins évitables que le répertoire Laval, pour sa part, se donne pour mission de suivre le plus fidèlement possible (même si cela n'exclut pas des adaptations) la liste de la Bibliothèque du Congrès, dans un contexte canadien de bilinguisme officiel et en relation avec une politique délibérée de traduction automatique des indexations américaines : d'où un nombre notable de vedettes et de subdivisions Laval se présentant comme autant de décalques purs et simples de l'américain, sans compter quelques canadianismes (au demeurant parfaitement justifiés dans un répertoire québécois) et autres formes éloignées de l'usage courant, tel du moins qu'il apparaît dans la documentation française.

3. Étienne Hustache, *op. cit.*

4. Suzanne Jouguelet, « Un "réseau" de vedettes matières : le répertoire de l'Université Laval adopté par la Bibliothèque nationale », *BBF*, n° 4, 1983.

On pourrait également rattacher à cette origine anglo-saxonne le parti pris de présenter les vedettes au pluriel, sauf quand l'usage l'interdit, alors même que la norme française avait toujours préconisé l'emploi du singulier : ce dont témoignent encore aujourd'hui aussi bien la pratique de l'Institut de l'information scientifique et technique (Inist), dans le domaine savant (bases de données Francis et Pascal), que la liste proposée par Martine Blanc-Montmayeur et Françoise Danset⁵, dans celui de la lecture publique.

En second lieu, le choix initial s'avère d'une certaine manière paradoxal : la liste américaine date en effet du début du XX^e siècle ; c'est dire qu'il s'agit d'un langage d'indexation bien antérieur à la mutation informatique, prévu au départ pour un catalogue sur fiches, simplement transféré ensuite sur le support électronique, appliquant dès lors naturellement des principes proches de ceux de la norme NF Z 44-070, y compris le caractère précoordonné du langage (alors même que la recherche postcoordonnée semble *a priori* mieux adaptée à l'environnement informatique). En démarquant étroitement ce modèle, Laval, puis Rameau, ne pouvaient ensuite qu'en reproduire les limites.

Au moment de l'adoption de Rameau, la relative distorsion entre le caractère du langage d'indexation choisi et l'outil informatique n'avait d'ailleurs pas échappé aux responsables de l'époque puisque, du côté de la DBMIST, il était prévu de compléter ce langage, conçu d'abord comme un outil destiné aux bibliothécaires, par un puissant logiciel de gestion et de consultation appelé Alexis, développé par la société Erli, lequel était supposé capable, par sa souplesse et son ingéniosité, d'en pallier les inconvénients pour le commun des utilisateurs. Tant il est vrai qu'on estime à l'époque que « *fondamen-*

5. Martine Blanc-Montmayeur et Françoise Danset, *Choix de vedettes matières à l'intention des bibliothèques*, Éd. du Cercle de la librairie, 2002.

talement Lamech se veut un outil d'aide à l'indexation destiné à des professionnels, un outil destiné à améliorer la cohérence et la pertinence de l'indexation documentaire, mais il n'est pas sa propre fin. On peut dire que Lamech constitue une condition nécessaire, mais non suffisante, pour améliorer les possibilités d'accès du lecteur : les logiciels d'aide à l'interrogation adaptés aux différents publics devant constituer l'autre versant du service⁶ ».

En troisième lieu, il convient de garder à l'esprit que les LCSH n'ont pas été conçus, au départ, comme une liste d'autorité indépendante de tout catalogue particulier et donc susceptible d'être utilisée dans un grand nombre de bibliothèques, mais bien comme le fichier d'autorité particulier d'un catalogue particulier. Il en résulte naturellement que les LCSH reflètent et entérinent nombre d'usages propres à la Bibliothèque du Congrès et à sa politique d'indexation, lesquels se sont trouvés ensuite abusivement présentés comme devant faire partie intégrante du langage d'indexation lui-même.

Rameau aujourd'hui

Quoi qu'il en soit de ces ambiguïtés initiales, il n'en reste pas moins que Rameau s'est imposé peu à peu, des années 1980 à aujourd'hui, comme le langage d'indexation national commun à la plupart des bibliothèques françaises : fruit d'un travail collectif longuement poursuivi, la Liste d'autorité Rameau constitue ainsi désormais un réservoir terminologique très important, en constant accroissement, dont la richesse ne saurait être sous-estimée, et qui a su gagner peu à peu une relative autonomie par rapport aux réservoirs sources dont il procède.

Parallèlement, une infrastructure institutionnelle et technique a été progressivement mise en place sur

le plan national : Centre national Rameau, relevant de la BnF ; réseau des utilisateurs et partenaires régi par une convention, avec ses instances représentatives, Comité d'orientation et Comité opérationnel ; produits et outils Rameau (Liste d'autorité, *Guide d'indexation Rameau*, *Journal des créations et des modifications*, site web Rameau, Fichier national des propositions Rameau) ; aides à la formation (réseau de formateurs, supports pédagogiques). Des coopérations internationales ont également été nouées.

Il va de soi que l'élargissement de l'audience de Rameau s'est accompagné de nuances dans l'utilisation que pouvaient en faire les différentes catégories d'établissements concernés :

- à la BnF, Rameau joue ainsi le rôle de fichier d'autorité matière du catalogue, ce pourquoi il avait été adopté à l'origine (une coordination de l'indexation matière fédérant les services de catalogage de l'établissement) ;

- dans le réseau des bibliothèques universitaires, sous la responsabilité de l'Agence bibliographique de l'enseignement supérieur (Abes), Rameau est utilisé comme langage d'indexation commun permettant de partager les données au sein du catalogue collectif du Sudoc, deuxième grand réservoir catalogographique, avec celui de la BnF, à utiliser Rameau ;
- enfin, les bibliothèques de lecture publique et les organismes divers (y compris privés, tel Électre) recourent à cet outil dans les limites de leurs besoins et catalogues particuliers - les bibliothèques de lecture publique ayant toutefois commencé à s'organiser avec le concours de la Direction du livre et de la lecture.

À mesure qu'il prenait de l'ampleur, le développement de Rameau a également fait naître, ou exacerbé, un certain nombre de contradictions tenant à son histoire, à son rôle national, à sa richesse même. À celles déjà évoquées, ajoutons qu'il s'agit :

- d'une liste encyclopédique couvrant l'ensemble des activités et des connaissances, mais fonctionnant, du

fait des liens établis entre les vedettes, à la manière d'un gigantesque et paradoxal thésaurus, alors que ce type d'outil est en principe réservé aux vocabulaires spécialisés ;

- d'un langage documentaire visant légitimement la plus grande cohérence possible, mais formant en même temps une œuvre collective sans cesse enrichie de nouveaux termes, dont la prolifération même constitue un défi permanent ;

- d'un langage d'indexation réservé à ce titre, en principe, au seul signallement de contenus de type documentaire, mais cependant utilisé également pour indiquer la forme de documents écrits, sonores ou audiovisuels (les vedettes-matières jouant alors parfois le rôle de titres de forme) ;

- d'un langage susceptible d'être utilisé dans tous les types d'établissements, mais conciliant de ce fait difficilement les besoins parfois contradictoires des différentes catégories d'utilisateurs (bibliothèques de lecture publique/bibliothèques savantes ; bibliothèques généralistes/bibliothèques spécialisées).

Compte tenu de la place centrale acquise par Rameau sur le plan national, et malgré les difficultés qui viennent d'être évoquées, il ne nous semble guère raisonnable de revenir aujourd'hui sur un choix initial qu'il serait désormais bien coûteux de remettre en cause. En revanche, il n'est pas interdit de se demander comment continuer à faire évoluer cet outil, en partant des contraintes mais aussi des possibilités qu'il recèle, de telle sorte qu'il réponde de mieux en mieux aux besoins concrets des utilisateurs. Tel est l'objectif auquel voudraient contribuer les propositions qui suivent concernant le langage Rameau.

Analyses et propositions

Sur le modèle des LCSH, le langage Rameau présente une structure en trois niveaux qui fait sa richesse mais aussi sa complexité :

6. Étienne Hustache, *op. cit.*

- au niveau terminologique (= termes retenus, appelés « vedettes » + termes exclus ou rejetés), Rameau est un langage contrôlé, notamment quant à la forme du vocabulaire, à la synonymie et à l'homonymie: l'objectif est de parvenir à un langage homogène et univoque (où 1 vedette = 1 concept et 1 concept = 1 vedette), tout en multipliant les points d'accès aux termes retenus à partir de termes exclus;
- au niveau sémantique (= relations entre termes génériques, spécifiques et associés), Rameau est un langage hiérarchisé à la manière d'un thésaurus: l'objectif est de permettre une navigation entre les termes retenus afin d'élargir (termes génériques), d'affiner (termes spécifiques) ou de réorienter (termes associés) la recherche;
- au niveau syntaxique (= vedettes + subdivisions), Rameau est un langage précoordonné obéissant à des règles de construction précises: l'objectif est de permettre, à côté de la recherche par mots, une recherche par feuilletage d'index, utile pour décliner grâce aux subdivisions les différentes facettes sous lesquelles le sujet de la recherche peut être envisagé.

Concrètement, cette relative complexité oblige l'indexeur qui travaille avec Rameau à disposer des deux types d'outils suivants: d'une part, le « vocabulaire » fourni par le réservoir terminologique contrôlé et hiérarchisé que constitue la Liste d'autorité Rameau; d'autre part, la « grammaire » rassemblant les règles de syntaxe qui permettent, à partir de ce vocabulaire, de construire les vedettes-matières au moment de l'indexation - règles que l'on trouve disséminées dans les notices de la Liste d'autorité ou rassemblées dans le *Guide d'indexation Rameau*.

Nous posons comme hypothèse de départ que, contrairement à ce qui est communément admis, ces trois niveaux ne sont pas d'égale importance, mais que, tout au contraire, on doit reconnaître la prééminence du niveau terminologique dans la mesure où ce dernier, à la limite,

peut se suffire à lui-même et dans la mesure où, à l'inverse, les deux autres niveaux ne peuvent exister sans lui: un vocabulaire contrôlé, qui ne serait ni hiérarchisé ni précoordonné, est en effet parfaitement concevable; en revanche, hiérarchisation et précoordination impliquent nécessairement l'existence préalable d'un vocabulaire contrôlé.

Le niveau terminologique constitue donc bien le noyau fondamental et minimal que Rameau partage avec tout langage documentaire. C'est lui, en effet, qui permet à l'utilisateur (qu'il s'agisse de l'indexeur ou de l'utilisateur final) d'accéder directement à l'information souhaitée. C'est également lui qui commande l'accès aux deux autres niveaux du langage Rameau: en effet, si l'utilisateur n'a pu accéder à la vedette retenue (soit parce qu'elle était mal choisie, soit faute de termes exclus en nombre suffisant), peu lui importe, par exemple, avec quelles subdivisions elle est construite (niveau syntaxique).

Or, le fait que le langage Rameau fonctionne également à la manière d'un thésaurus (relations du niveau sémantique) et qu'il soit précoordonné (constructions du niveau syntaxique) a conduit jusqu'ici à aligner le niveau terminologique sur les deux autres niveaux, voire à le subordonner à eux.

Au niveau sémantique, la possibilité de suivre les liens entre vedettes est ainsi parfois prétexte à restreindre le nombre de termes exclus sous les termes retenus.

Au niveau syntaxique (héritage des règles formelles précédentes), les subdivisions, qui devraient être de simples outils secondaires de tri, permettant de préciser certains aspects d'une vedette, sont souvent conçues au contraire comme les termes d'une combinatoire où elles interviendraient, au même titre que la tête de vedette, pour exprimer, sous forme de séquences articulées, des notions pour lesquelles il peut exister, par ailleurs, des mots ou des expressions attestées dans le vocabulaire courant ou spécialisé.

Enfin, le niveau syntaxique a tendance à mobiliser une part essentielle du temps, de l'énergie et de la réflexion collective, trop souvent au détriment de la question du vocabulaire: comme si l'essentiel demeurerait toujours la définition de règles formelles d'écriture et non la gestion terminologique d'une liste d'autorité; comme si, en un mot, nous étions encore à l'époque des catalogues sur fiches.

Nous proposons donc d'inverser les priorités: il n'est certes pas question d'abandonner les niveaux sémantique et syntaxique, mais il s'agit simplement d'accorder désormais la primauté au niveau terminologique, dont les exigences devront commander l'ensemble.

Niveau terminologique

Au niveau terminologique, la vedette Rameau correspond au « descripteur », tel qu'il est traditionnellement défini à la fois par la norme et par les principales études sur les langages documentaires, à savoir: « *mot ou groupe de mot choisis parmi un ensemble de termes équivalents pour représenter sans ambiguïté une notion apparaissant dans un document*⁷ ».

Or, en nous inspirant des travaux de Muriel Amar - même si ces derniers s'inscrivent dans une tout autre perspective⁸ -, nous proposons au préalable de redéfinir et d'élargir la notion de descripteur.

On dira que la fonction du descripteur n'est pas tant de « représenter » ou d'« exprimer » le contenu d'un document que de désigner ce contenu, un peu à la manière d'un « poteau indicateur » ou d'une « étiquette » permettant de « ranger » des informations données pour les retrouver ensuite, de telle sorte que,

7. Norme NF Z 44-070 : *Indexation analytique par matière*, Afnor, 1986.

8. Muriel Amar, *Les fondements théoriques de l'indexation: une approche linguistique*, ADBS éditions, 2000.

bien orienté, l'utilisateur final puisse accéder au contenu recherché à partir du ou des termes de la requête qu'il aura formulée : autrement dit, il ne s'agit pas tant de décrire un contenu que d'y donner accès.

Dès lors, tout terme permettant de désigner ce contenu, et donc d'y donner accès, doit être considéré comme un descripteur : non seulement la vedette, mais encore l'ensemble des termes exclus, car tous contribuent également à décliner, et à épuiser (dans l'idéal), l'ensemble des possibilités terminologiques existant, dans une langue donnée, pour désigner un contenu informationnel donné.

Dans cette perspective, sera appelé « descripteur » l'ensemble, solidaire, indissociable, que constitue ce qu'on pourrait appeler la « grappe terminologique » (vedette + termes exclus), même si, dans certains cas, cette grappe se réduit au seul terme retenu, dès lors qu'il n'en existe pas de variante terminologique.

Nous postulons que c'est cette grappe terminologique qui forme le noyau dur de tout langage documentaire, donc de Rameau, car c'est elle qui, en multipliant et en contrôlant les accès matière, doit permettre au maximum de requêtes de déboucher sur une réponse pertinente, et sur celle-là seulement, à partir des termes mêmes dans lesquels ces requêtes auront été formulées, et que la grappe terminologique aura su anticiper : tant il nous semble qu'un bon langage d'indexation doit être le plus transparent possible pour l'utilisateur final (y compris l'indexeur, dont le travail devrait s'en trouver facilité et accéléré d'autant).

Dès lors, pour être pleinement satisfaisant au niveau terminologique, Rameau devra proposer un vocabulaire pertinent, univoque, riche en termes exclus et cohérent.

Un vocabulaire pertinent

En premier lieu, le vocabulaire doit être pertinent, c'est-à-dire exact, autrement dit précis et sans erreur

dans le choix des termes ; conforme aux usages majoritairement attestés dans les sources en langue française : ce sont donc ces dernières qui constituent la garantie documentaire du choix des termes retenus (étant entendu que les usages peuvent évoluer et nécessiter des actualisations du vocabulaire) ; respectueux de la langue française telle qu'elle est correctement pratiquée dans notre pays (les formes « hétérodoxes » rencontrées dans la documentation devenant, pour leur part, des termes exclus).

Cela implique que ni le choix des LCSH, ni même celui du RVM Laval ne sauraient constituer à eux seuls cette garantie : tout au contraire, ce sont les sources consultées par ailleurs qui doivent faire foi en dernière analyse. On ne saurait donc trop souligner l'importance capitale des sources, lesquelles doivent être diversifiées autant que nécessaire, puis confrontées entre elles, ainsi, bien entendu, qu'avec les LCSH et le répertoire Laval.

Un vocabulaire univoque

L'une des raisons d'être de tout langage documentaire est de corriger la polysémie inhérente au langage naturel afin de prévenir le bruit susceptible de parasiter les réponses : les vedettes doivent donc également offrir, d'emblée, un sens clair et univoque, de telle sorte que l'utilisateur comprenne immédiatement, et sans la moindre hésitation, ce dont il est question, d'une part, et qu'il obtienne l'information pertinente pour lui et celle-là seulement, d'autre part.

Dans ce but, plusieurs procédés sont mis en œuvre dans la Liste d'autorité Rameau : soit l'ajout d'une précision entre parenthèses, dite « qualificatif », permettant de distinguer entre deux homonymes : par exemple, *Vol* et *Vol (droit)* ; soit l'ajout d'un adjectif qualificatif : par exemple, *Analyse mathématique* et *Analyse numérique* ; soit, enfin, l'utilisation, pour un même terme, du singulier et du pluriel : par exemple, *Cinéma* et *Cinemas*.

On peut toutefois se demander s'il ne faudrait pas aller plus loin en ce sens, dans la mesure où certains de ces procédés laissent encore subsister une part d'équivoque, du moins à première lecture, dans l'énoncé des vedettes. Plus largement, une relecture critique de certains intitulés de vedettes ne serait sans doute pas superflue de ce point de vue.

Un vocabulaire riche en termes exclus

Mais même pertinentes et univoques, les vedettes, sauf rares exceptions, ne peuvent trouver leur pleine efficacité qu'accompagnées de termes exclus, au sein de ce qu'on a nommé la « grappe terminologique ». L'absence ou l'insuffisance de termes exclus a en effet des conséquences aisément prévisibles : la recherche documentaire s'en trouve compliquée, voire parfois impossible, l'utilisateur devant deviner le ou les termes qui ont été choisis comme vedettes ; si cet utilisateur est un indexeur, il risque d'en conclure à bon droit que la notion n'est pas entrée dans la Liste d'autorité Rameau et de proposer, en conséquence, la création d'une nouvelle vedette susceptible de venir doubler avec celle qu'il n'aura pu trouver, faute de terme exclu.

Si l'on se place du point de vue de l'utilisateur, quel qu'il soit, il apparaît donc que l'importance des termes exclus égale bien celle de la vedette, au point qu'on pourrait concevoir, à la limite, une grappe terminologique sans terme pivot, c'est-à-dire sans vedette, laquelle n'est, au bout du compte, qu'un accès à l'information parmi d'autres, dont la mise en valeur se justifie surtout par des raisons pratiques de gestion et d'affichage des données. C'est dire aussi que la richesse, donc l'efficacité, d'un langage documentaire se mesurera, tout autant qu'au nombre des termes retenus qui le constituent, à la proportion qu'il offre entre le nombre de ces derniers et celui des termes exclus : là réside même, sans doute, sa principale

« valeur ajoutée » (avec le caractère univoque du vocabulaire) par rapport à la recherche plein texte.

Remarquons par ailleurs que l'utilisation de termes exclus peut également être un moyen commode de résoudre la difficulté à laquelle se heurte le langage Rameau lorsqu'il prétend répondre, en même temps, aux besoins parfois contradictoires des différents types de bibliothèques auxquels il s'adresse. Quelle forme en effet choisir pour une vedette donnée, dans ce contexte, lorsqu'il existe à la fois une forme usuelle et une forme savante? Le dilemme se résout dès lors que la forme qui n'aura pas été retenue devient un terme exclu permettant d'accéder à l'information au même titre que la vedette.

C'est dire aussi que toute bibliothèque qui adopte Rameau (comme ce serait d'ailleurs le cas pour tout autre langage d'indexation contrôlé) doit impérativement proposer à ses lecteurs, en guise d'outil de recherche minimal, l'ensemble des vedettes correspondant à son fonds documentaire accompagnées des termes exclus qui leur sont associés.

Le besoin de disposer de termes exclus provient, en premier lieu, de l'ambiguïté et de la richesse du langage naturel: l'idéal - visé, mais jamais atteint - est de parvenir à prévoir ainsi toutes les formes de requêtes possibles, en déclinant l'ensemble des manières constatées d'exprimer un contenu donné. Il va de soi que ce travail de prospection et de veille terminologiques ne peut que bénéficier des suggestions que font les indexeurs, par l'intermédiaire du Fichier national de propositions Rameau, pour ajouter des termes exclus dans la Liste d'autorité, dans la mesure où ils sont directement confrontés aux usages de la documentation.

Le besoin de termes exclus résulte également, dans une moindre proportion, de la nature du langage documentaire lui-même et, plus précisément, de la possibilité qu'offre Rameau d'une interrogation par feuilletage d'index et par mots, sous

réserve de disposer d'un logiciel adéquat. Pour être pleinement efficace, la grappe terminologique devrait en effet tenir compte systématiquement de cette double possibilité, même au risque d'alourdir quelque peu les index, tant il nous semble que priorité doit être donnée à la multiplication des accès. Ainsi, par exemple, est-il utile de compléter certaines vedettes composées (locutions, constructions syntaxiques) avec un terme exclu inversant l'ordre de leurs éléments afin d'obtenir une double entrée d'index (*Théorème de Gödel / Gödel, théorème de*).

Il résulte de ce qui précède que la forme des termes exclus doit pouvoir être diversifiée sans restriction ni autocensure, autant qu'il est nécessaire pour faciliter l'accès à l'information pertinente. Refuser cette souplesse - par exemple, n'accepter de construction dans les termes exclus qu'avec les subdivisions du langage Rameau - reviendrait à sacrifier la richesse du niveau terminologique aux règles du niveau syntaxique, autrement dit ce qui est le plus immédiatement important du point de vue de la recherche concrète à ce qui ne l'est que secondairement.

En conclusion, si l'on admet la perspective qui vient d'être tracée, tout mot ou toute expression attestée, dès lors que lui correspond une documentation justifiant un signalement analytique, devrait donc systématiquement se trouver dans la Liste d'autorité Rameau sous l'une des trois formes suivantes: vedette; terme exclu; renvoi général d'équivalence orientant vers un ensemble de vedettes ou de constructions qui remplacent la forme rejetée.

Un vocabulaire cohérent

Le vocabulaire doit enfin être cohérent dans sa structure, c'est-à-dire sans lacunes flagrantes ni doublons.

Il peut sembler paradoxal d'en appeler à la complétude et à la cohérence d'un vocabulaire dont la vocation est encyclopédique et qui s'en-

richit en permanence en fonction des besoins rencontrés à l'indexation. La question n'est cependant pas illégitime si l'on estime que la masse critique aujourd'hui atteinte par le langage Rameau peut, au contraire, autoriser certaines évaluations rétrospectives. D'autant qu'il est possible de distinguer, parmi les vedettes Rameau, deux catégories susceptibles d'une appréciation différente: d'une part, les vedettes qui fonctionnent selon des réseaux conceptuels plus ou moins complexes et étendus; d'autre part, les vedettes qui constituent des séries ouvertes d'« objets » (animaux, plantes, minéraux, langues, groupes ethniques, etc.).

Bien entendu, il n'est intéressant de repérer rétrospectivement d'éventuelles lacunes que dans la première catégorie de vedettes que l'on vient de distinguer, dans la mesure même où celles-ci ne constituent pas des séries ouvertes mais, au contraire, fonctionnent en systèmes dont la cohérence peut être évaluée.

En revanche, il est impératif de traquer les doublons terminologiques dans toutes les catégories de vedettes, qu'elles relèvent de réseaux conceptuels ou de séries ouvertes, car ces vedettes synonymes, ou de sens trop voisin, conduisent à des tranches d'indexation parallèles qui dispersent l'information et engendrent du silence. Ce travail est d'autant plus nécessaire que les répertoires sources de départ sont loin d'être exempts de ces défauts; que l'accroissement permanent du vocabulaire Rameau constitue un facteur de risque aggravant; enfin, qu'une conception trop rigide du caractère analytique de l'indexation peut conduire à certaines dérives: l'usage veut en effet que l'on indexe au niveau le plus précis; or la dispersion de l'information est proportionnelle à sa précision; donc, plus une vedette sera précise, plus elle aura tendance à disperser l'information, au risque d'aboutir, parfois, à des indexations parallèles.

Niveaux sémantique et syntaxique

Affirmer la prééminence du niveau terminologique – dans la mesure où il commande les points d'accès directs à l'information – ne signifie pas, pour autant, abandonner les niveaux sémantique et syntaxique. Simplement, le travail qui pourra leur être consacré devra être envisagé à partir des exigences du niveau terminologique, et non l'inverse.

Niveau sémantique

Par rapport aux possibilités limitées des anciens catalogues matière sur fiches (lesquels contenaient déjà des renvois d'orientations du type « voir aussi »), le support informatique a permis de multiplier les relations entre les vedettes dans les listes d'autorité. C'est ainsi que, dans la Liste d'autorité Rameau, il est possible de suivre trois types de liens entre vedettes, permettant :

- soit d'élargir la recherche (liens génériques);
- soit de la préciser (liens spécifiques);
- soit de la réorienter (liens associatifs).

Le remarquable outil de navigation au sein du vocabulaire Rameau que constitue de la sorte le niveau sémantique intéresse aussi bien l'utilisateur final, qui peut ainsi ajuster sa requête, que l'indexeur qui y trouve une aide précieuse pour choisir le terme approprié et le bon niveau d'indexation.

L'intérêt indéniable du niveau sémantique ne saurait toutefois en masquer les difficultés. En effet, le réseau complexe des multiples liens entre vedettes rapproche la Liste d'autorité Rameau du fonctionnement d'un thésaurus, mais d'un thésaurus qui serait doublement paradoxal : d'une part, il s'agit en effet d'un thésaurus encyclopédique, alors même que ce type d'outil est en principe réservé aux vocabulaires spécialisés; d'autre part, il s'agit d'un vocabulaire ouvert

et évolutif, enrichi en permanence, là où un thésaurus est en principe plutôt stable et fermé.

Double paradoxe qui explique la difficulté à laquelle se heurte la gestion de cet énorme réseau ou échec niveau sémantique et la propension qui en résulte à s'en remettre passivement aux choix des LCSH, pourtant loin d'être eux-mêmes toujours cohérents.

Dans ces conditions, une maîtrise réfléchie du niveau sémantique passe sans doute par un juste équilibre entre, d'une part, la préservation de la richesse des liens entre vedettes, et, d'autre part, la nécessité de faire des choix et de définir des règles afin d'éviter que cette richesse ne devienne prolifération anarchique.

De ce point de vue, il est souhaitable de lier étroitement le travail sur le niveau sémantique et celui consacré au niveau terminologique, tant ces deux niveaux apparaissent solidaires en dernière analyse : le niveau sémantique du langage Rameau ne peut en effet qu'apporter une aide précieuse à la gestion et à l'amélioration de la cohérence du niveau terminologique, grâce au réseau serré qu'il tisse entre les vedettes; en retour, une meilleure appréciation de la cohérence conceptuelle d'un domaine ne pourra qu'aider à rationaliser le réseau des liens qui en est l'expression.

Niveau syntaxique

Par son niveau syntaxique, Rameau est un langage précoordonné obéissant à des règles de construction précises : la vedette peut ainsi être suivie de subdivisions de sujet, géographiques, chronologiques et de forme.

La précoordination, bien que conçue à l'origine pour des catalogues sur fiches, conserve une certaine utilité dans un contexte informatique. Lors de la consultation de gros réservoirs bibliographiques comme les catalogues de bibliothèques nationales ou les catalogues collectifs, la précoordination permet, en effet, d'atténuer le bruit documentaire en ordonnant

les réponses sous forme d'index qui déclinent les différentes facettes sous lesquelles peut être envisagée la notion recherchée : tri particulièrement bienvenu lorsque la notion fait l'objet d'une abondante littérature, car il permet de préciser la requête et évite les blocages du type « trop de réponses : affinez votre question » que provoque la recherche par mots dans certaines applications informatiques.

Il n'en reste pas moins que le caractère précoordonné de Rameau suscite des difficultés récurrentes dans sa mise en œuvre.

L'idéal, en effet, aurait été de disposer, d'une part, d'une liste de subdivisions affranchies, chronologiques, de sujet et de forme, en nombre restreint et d'un emploi très large (les subdivisions géographiques, pour leur part, ne pouvant être qu'en nombre quasi illimité); d'autre part, de quelques règles de syntaxe simples d'utilisation, avec le moins possible d'exceptions.

Or, c'est tout le contraire qui se produit : des subdivisions affranchies en très (et sans doute trop) grand nombre, dont l'emploi, pour beaucoup d'entre elles, est limité à certaines catégories de têtes de vedette, ce qui complique singulièrement les choses; conséquemment, des règles syntaxiques dont la complexité contraste avec la relative simplicité de la norme qui régissait l'indexation à l'époque des catalogues sur fiches.

Il en résulte un ensemble de conséquences fâcheuses :

- le travail – et la formation – des indexeurs s'en trouve passablement compliqué;
- le grand nombre de vedettes matières pouvant être construites à l'indexation à partir des subdivisions limite d'autant le nombre de termes exclus susceptibles de renvoyer à ces constructions dans la Liste d'autorité et les catalogues (conséquence sans doute inévitable, mais qui n'en appauvrit pas moins notablement l'accès à l'information);
- il devient parfois possible, du fait du grand nombre des subdivisions (certaines d'emploi général et

d'autres d'emploi particulier), d'indexer un même contenu de plusieurs manières concurrentes, également justes du point de vue de la règle, ce qui a pour effet de créer des tranches parallèles d'indexation qui dispersent l'information; cette richesse syntaxique peut même, dans certains cas, se substituer lors de l'indexation à un terme ou à une expression existant par ailleurs, qui dès lors ne figureront pas dans la Liste d'autorité (on aura, dans ce cas, décrit correctement un contenu grâce à la syntaxe, mais sans y donner accès à partir des ressources terminologiques pourtant en usage).

En fait, le problème ne tient donc pas tant à ce que Rameau soit un langage précoordonné, mais à ce qu'il le soit avec un luxe excessif de raffinement. Cette situation provient de ce que les subdivisions semblent davantage avoir été comprises, au départ, comme les éléments, les « briques », d'une combinatoire permettant d'exprimer toutes les nuances d'un contenu documentaire donné sous forme de séquences syntaxiques, plutôt que comme de simples outils secondaires de tri déclinant les principales facettes sous lesquelles peut être abordé un sujet: ce qui conduit à conférer aux subdivisions une importance égale aux têtes de vedettes constituant le vocabulaire, alors qu'elles ne devraient en être que de simples compléments.

Au demeurant, cette conception va de pair avec une tendance à préférer, pour indexer les sujets complexes, une seule vedette matière construite, elle-même complexe, plutôt que la décomposition du sujet en plusieurs vedettes simples – ce qui pourtant irait aussi vite, et peut-être plus vite, à l'indexation, sans compromettre en rien les recherches, puisque seraient ainsi multipliées dans l'index les clés d'accès au sujet concerné, tandis que la recherche par mots permettrait toujours de combiner *a posteriori* les notions au moment de l'interrogation. En d'autres termes, tout se passe comme s'il s'agissait de décrire un contenu dans toutes ses nuances plu-

tôt que d'y donner accès à partir de plusieurs « poteaux indicateurs ».

Au fil des années, le Centre national Rameau a toutefois cherché à simplifier en partie cet héritage, en réponse aux besoins exprimés par les membres du réseau national, au point que le niveau syntaxique a fini par mobiliser l'essentiel de sa réflexion et a eu tendance, de ce fait, à prendre le pas sur les autres niveaux du langage Rameau: on peut ainsi remarquer que le travail intense préparant chaque édition du *Guide d'indexation Rameau* tourne surtout autour du problème de la syntaxe et des subdivisions, et que, par ailleurs, l'épaisseur même de ce guide témoigne, à l'instar d'ailleurs du manuel des LCSH, de la singulière complexité du niveau syntaxique auquel il est principalement consacré. Cela a du moins permis au Centre national Rameau de développer une grande expertise dans ce domaine.

Il reste que la réforme du niveau syntaxique n'a pas été menée à son terme, pour deux raisons au moins: d'une part, malgré une volonté déterminée de simplification de la syntaxe (au risque, parfois, d'en accentuer le caractère de combinatoire), la philosophie sous-jacente n'a jamais vraiment été remise en cause, sans doute sous l'influence prégnante du modèle que constituent les LCSH; d'autre part, est venu s'ajouter progressivement un obstacle technique majeur tenant à l'augmentation, au fil des années, de la masse documentaire indexée selon ces principes, laquelle interdit désormais, du fait de son ampleur, de revenir sur certains choix initiaux: simplifier les règles de la syntaxe, supprimer des subdivisions, implique, en effet, dans la plupart des cas, de devoir corriger ensuite rétrospectivement des centaines, voire des milliers d'indexations dans les catalogues, sauf à laisser s'y introduire des incohérences, et ces corrections, la plupart du temps, doivent être menées manuellement, faute de programmes informatiques adéquats.

C'est en partant de ce constat qu'il convient, dès lors, de préciser la part

de travail qu'il reste souhaitable de consacrer à la syntaxe.

Pour ce qui est du travail de simplification des règles de la syntaxe proprement dite (limitation des exceptions à l'ordre normal des subdivisions, suppression des subdivisions inutiles et, par-dessus tout, libéralisation et assouplissement des règles d'emploi des subdivisions à chaque fois que c'est possible), il peut être continué autant que nécessaire, mais en tenant compte de la contrainte technique qui vient d'être évoquée: le volume des corrections induites dans les catalogues doit demeurer raisonnable, ou alors pouvoir faire l'objet de programmes automatisés; dans le cas contraire, sauf éventuelles rares exceptions, on devra se résigner à ajourner la réforme. Un bilan devrait permettre de mesurer ce qui est encore possible de réaliser en ce domaine, c'est-à-dire, sans doute, malheureusement assez peu.

En revanche, le travail sur les subdivisions, considéré cette fois du strict point de vue de la terminologie, doit d'autant plus être poursuivi qu'il a pour mérite de n'entraîner, le plus souvent, que des corrections manuelles limitées dans les catalogues:

- les subdivisions devraient ainsi faire l'objet d'un examen terminologique systématique au même titre que les autres vedettes, et en relation avec elles;
- un repérage systématique des subdivisions auxquelles ne correspondrait pas une tête de vedette, alors même que la notion est attestée dans la documentation, serait également à mener, afin de créer lesdites têtes de vedette lorsque c'est possible; de même faudrait-il veiller à ce que des termes ou des expressions dûment attestés dans la documentation ne se trouvent pas exclus de la Liste d'autorité, au prétexte qu'il serait possible d'exprimer un contenu analogue par une construction syntaxique à l'indexation;
- enfin, les règles d'emploi des subdivisions devraient être assouplies sans timidité en fonction des exi-

gences du niveau terminologique: certains quasi-doublons (par exemple, entre l'activité et la catégorie de personne s'y rapportant: *Lancer du poids/Lanceurs de poids*) semblent en effet se justifier avant tout par le fait qu'ils permettent l'emploi de subdivisions applicables à des catégories de vedettes différentes; or, cela revient à soumettre la terminologie à la syntaxe, alors même que c'est le contraire qui est souhaitable.

Au-delà, une simplification plus générale des règles syntaxiques pourrait être envisagée, en s'inspirant du système d'indexation allemand SWD (Schlagwortnormdatei), dont la pré-coordination a pour particularité de laisser une large initiative aux indexeurs. Dans cette perspective, l'ensemble des subdivisions seraient rassemblées en un nombre très restreint de listes et les restrictions actuelles à leur emploi en fonction de leurs domaines d'application supprimées aussi largement qu'il est possible, ce qui reviendrait à laisser cet emploi à l'appréciation des indexeurs. Une telle solution – qui supposerait une étude préalable approfondie – présenterait l'avantage d'assouplir sensiblement l'utilisation de Rameau, sans pour autant exiger de corrections rétrospectives dans les catalogues puisque la structure de la syntaxe ne s'en trouverait pas bouleversée.

Une nouvelle étape

La perspective qui vient d'être esquissée à propos du langage Rameau – et qui consiste à donner le pas à la dimension de réservoir terminologique contrôlé et hiérarchisé du langage sur celle de combinatoire syntaxique – ouvre la possibilité d'une nouvelle étape dans le travail mené, en relation avec les utilisateurs du réseau, par le Centre national Rameau.

En schématisant beaucoup, on peut en effet considérer qu'une première phase a été marquée par le travail considérable d'« accumulation primitive » du vocabulaire, à partir

des réservoirs américain et québécois; une deuxième phase a ensuite permis la mise en place d'un environnement organisationnel (réseau national) et technique (site web, Fichier national des propositions Rameau, Liste d'autorité en ligne, etc.) performant; appuyée sur ce double acquis, une troisième phase peut maintenant être envisagée sous la forme d'un travail rétrospectif, méthodique et progressif de révision qualitative du vocabulaire, centré sur la Liste d'autorité, à partir des principes que nous avons avancés.

Une telle approche, procédant par disciplines, domaines, sous-domaines, champs sémantiques et séries particulières, nous semble en effet être seule capable, par son caractère systématique, de permettre de contrôler véritablement la cohérence du vocabulaire (notamment dans le cas des vedettes formant système) et d'homogénéiser la forme des notices d'autorité (notamment dans le cas des vedettes formant des séries ouvertes d'objets); elle est en outre possible, pour peu qu'un temps suffisant y soit consacré, ainsi que l'a prouvé un essai « expérimental ».

Ce que nous proposons en l'espèce revient, en fait, à construire des « ontologies » (au sens d'organisations structurées de la connaissance) par domaines et sous-domaines, au sein d'une liste d'autorités qui resterait commune, dont le caractère homogène serait préservé, et qui finirait par constituer elle-même une manière d'ontologie encyclopédique, en raison du réseau des liens sémantiques établi entre les vedettes (ou plutôt les grappes terminologiques): si l'on met de côté les problèmes liés à la syntaxe, le langage Rameau, en tant qu'il est à la fois contrôlé et hiérarchisé, autorise en effet une telle perspective, pour peu que l'on veuille s'y consacrer d'une manière systématique.

Il va de soi qu'un tel travail suppose la plus parfaite autonomie intellectuelle à l'égard des répertoires sources (LCSH et RVM Laval). Il n'est

certes pas question de se priver de leur apport, précieux et souvent incontournable, ni d'envisager de rompre notre lien avec eux en oubliant notre dette (d'autant qu'une coopération fructueuse, reconnue par une convention, unit le RVM Laval et Rameau). Il nous semble en revanche indispensable, si l'on veut pouvoir faire évoluer Rameau et le mettre en mesure de répondre toujours mieux à sa vocation de langage d'indexation national, de ne plus le considérer (y compris inconsciemment) comme une simple transposition aménagée des deux répertoires sources, mais bien comme un répertoire parfaitement autonome, même s'il reste, par la force des choses, en partie l'homologue des deux autres. Ceci nous paraît d'autant plus nécessaire qu'il ne faut jamais perdre de vue que l'indexation matière, dans la mesure où elle recourt à la langue commune (et non, par exemple, aux indices d'une classification), ne saurait en ignorer la dimension locale toujours particulière et se plier, de ce fait, à une uniformisation internationale trop rigide.

Le programme qui vient d'être esquissé s'efforce de marier ambition et réalisme: s'inscrivant dans la continuité, refusant de céder à la tentation de la *tabula rasa*, il cherche simplement à améliorer ce qui peut l'être autant qu'il est possible, sans rupture, à moyens constants, en prenant acte de la place centrale acquise par un langage d'indexation devenu réalité nationale (voire internationale), et en prenant appui sur l'acquis d'un outil dont les pesanteurs sont certaines (en raison principalement de la complexité de la syntaxe) mais dont les potentialités (terminologiques et sémantiques) nous semblent loin d'avoir été entièrement épuisées, y compris dans le contexte des évolutions techniques en cours, si l'on veut bien admettre que ces dernières ne sauraient totalement évacuer la question des contenus intellectuels.

Juin 2005