

NOTE ET BILAN D'EXPÉRIENCE

L'analyse documentaire I: Les thésaurus en France

Louise Gagnon-Arguin

*L*e mouvement de normalisation dans la description et l'analyse des documents d'archives s'est amorcé au cours des années 1970 suscité par les exigences de l'informatique et les possibilités qu'elle offrait pour assurer le traitement efficace de l'information. Les efforts des archivistes dans ce sens se sont manifestés dans différents domaines et dans plusieurs pays. Ainsi se sont développées des normes de description des documents d'archives aux États-Unis (Hensen, 1989), en Grande-Bretagne (Cook and Procter, 1989) et plus récemment au Canada (*Règles pour la description des documents d'archives* (RDDA), 1990). Ces normes avaient comme objectifs, entre autres, de permettre la description des documents d'archives en vue de les rendre compatibles aux exigences des bases de données documentaires, de rendre possible un meilleur repérage de l'information et d'en assurer une plus grande diffusion auprès des chercheurs ou des usagers en général.

L'analyse des documents d'archives constitue une suite logique aux efforts de normalisation des descriptions car elle permet une meilleure connaissance de leur contenu. C'est ainsi que l'on retrouve, dans la littérature archivistique américaine, nombre d'articles qui abordent les problèmes particuliers à l'analyse et à l'indexation des documents d'archives et au développement de thésaurus qui leur sont propres. Au Canada, une recherche a aussi été menée sur le sujet. En 1988, le Comité de planification sur les normes de description créait un Groupe de travail sur l'indexation par sujet, étant conscient du lien très étroit qui existe entre la description des documents d'archives et leur indexation ou leur repérage et de la nécessité de recherche dans ce domaine. Les réalisations de ce Groupe se sont traduites par la publication, en 1992, d'un ouvrage sur l'indexation par sujet (BCA, 1992). La recommandation 9 de ce rapport soulignait «le besoin de développer des thésaurus spécialisés sur des thèmes que l'on retrouve le plus couramment dans les archives» (BCA, 1992, p. 129).

À l'occasion du Stage international des Archives nationales de France 1994, nous avons été à même de constater tout le travail accompli par les archivistes français sur le plan de la description et de l'analyse des documents d'archives, et particulièrement dans l'élaboration de thésaurus appliqués aux archives. Il nous est apparu important de faire connaître ce travail aux archivistes québécois à un moment où ils se servent de la description normalisée selon les *Règles de description des documents d'archives* (RDDA) pour réaliser l'indexation.

Notre article a donc pour objet de situer le contexte dans lequel se sont développées la recherche et les expériences sur le traitement descriptif et analytique des documents d'archives en France depuis le début des années 1980, de présenter les trois principaux thésaurus utilisés dans les centres d'archives français, d'en faire ressortir les principales caractéristiques et de tirer les conclusions utiles quant à la mise en place de tels outils de travail. C'est à partir de la littérature archivistique française et surtout d'entrevues menées auprès des responsables de ces thésaurus¹ que nous avons recueilli les informations utiles à cet article.

CONTEXTE DU DÉVELOPPEMENT DU TRAVAIL DESCRIPTIF ET ANALYTIQUE DES DOCUMENTS D'ARCHIVES EN FRANCE

Différents facteurs ont favorisé le développement du traitement descriptif et analytique des documents d'archives en France. Certains sont liés à l'augmentation de la masse documentaire à conserver ainsi qu'à la disponibilité de matériel informatique, tandis que d'autres relèvent de changements survenus dans la pratique archivistique française. Ces facteurs sont en relation étroite les uns les autres, mais ils comportent suffisamment de caractéristiques propres pour justifier leur étude particulière.

L'augmentation de la masse documentaire et disponibilité de matériel informatique

L'augmentation de la masse documentaire constitue l'élément déclencheur du renouvellement de la pratique archivistique française. En 1962, une enquête commandée par le ministre des Affaires culturelles d'alors sur l'estimation des masses de papiers existantes ainsi qu'à celle de la production pour la prochaine décennie permettait de prévoir une entrée de 800 km linéaires de documents (Pétillat et Prax, 1993, p. 307). Comme le soulignent des archivistes dans le même ouvrage: «confronté au problème de la masse des archives contemporaines, qu'aggrave la rapidité de leur prise en charge dans un système où le préarchivage incombe encore souvent aux services gérés par les archives, l'archiviste n'a plus les moyens d'attendre ou de laisser faire le temps» (Pétillat et Prax, 1993, p. 257).

C'est la situation devant laquelle se retrouvaient les archivistes français au début des années 1980. Cette situation ainsi que les causes qui vont provoquer l'avènement de nouvelles pratiques ne sont pas spécifiques à la France mais elles dérangent particulièrement dans un pays où les archives historiques avaient été l'objet des préoccupations premières des archivistes. L'augmentation de cette masse documentaire se cristallise autour de l'ouverture d'un centre de préarchivage qui deviendra, en 1986, le Centre des archives contemporaines situé à Fontainebleau. Ce Centre recueille les do-

cuments provenant des ministères, en gère la localisation, l'élimination, la conservation (officiellement depuis 1986) et la communication aux services versants ou aux chercheurs et, par là même, le repérage. Cette concentration de documents et de services allait provoquer le développement d'une nouvelle expertise.

Fort heureusement, l'informatique se développe en même temps que le phénomène de la surproduction de l'écrit. Dans les ministères, on prend en compte les besoins propres au traitement de la documentation. Par exemple, le schéma directeur des Services du premier ministre prévoit le développement informatique qui a permis la mise en place de la base de données *Varenne*. Au Centre des archives contemporaines, le système informatique a été rattaché au service informatique du ministère de la Culture.

Les logiciels documentaires de bases de données font aussi leur apparition sur le marché. Aujourd'hui, deux d'entre eux sont particulièrement utilisés. Il s'agit du logiciel *Mistral* sur lequel est développé la base de données *Varenne* et du logiciel *Texto* qui sert de structure entre autres à la base de données *Priam 3*. Des bases de données ont été développées dans différents départements. Elles servent à la fois à la description de fonds, à la gestion des archives départementales et à la gestion administrative (prêt, données statistiques demandées par la Direction des Archives de France, etc.). Parmi celles-ci, on retrouve entre autres les bases suivantes: *Gaïa*, *Ardan*, *Ariane* et *Archeia*.

L'informatique est aussi utilisée pour traiter des masses importantes d'information pour lesquelles le travail de traitement a déjà été réalisé manuellement ou encore pour traiter des séries homogènes. Aux Archives nationales de France, l'index de l'État général des fonds se retrouve dans la base de données *Égérie*². Une analyse des «documents de la série F21 (Beaux-Arts) concernant les commandes et les achats d'œuvres d'art par l'État, ainsi que leur répartition dans les musées et les édifices publics» s'inscrit dans la base de données *Arcade* (Labat-Poussin, 1984). Deux autres bases de données portent sur des dossiers ou documents en particulier. C'est ainsi que la base de données *Léonore* répertorie les dossiers de nomination des titulaires de la Légion d'honneur et que la base de données *Prof* analyse des lettres de provisions d'offices conservées dans la sous-série V (Grande Chancellerie) des Archives nationales. Il n'existe toutefois pas de bases de données pour toutes les archives déposées aux Archives nationales de France. Ce sont les divers instruments de recherche imprimés qui permettent aux chercheurs d'accéder aux fonds, séries ou documents, qu'ils veulent consulter.

L'informatique a répondu aux nouveaux besoins des archives. Cela s'exprime par le développement de deux types d'informatique: l'informatique de gestion et l'informatique documentaire. Les archivistes en ont tiré avantage en développant des systèmes pour répondre à des besoins de repérage et d'analyse.

Les changements dans la pratique archivistique

Plusieurs changements sont survenus dans la pratique archivistique française. Nous en rappellerons ici quelques-uns:

L'abandon de l'utilisation de la cotation comme méthode de classification des archives contemporaines

La circulaire AD 79-6 du 31 décembre 1979 officialise l'abandon du cadre de classification adopté pour tous les documents antérieurs au 10 juillet 1940. Désormais, ce ne sera plus la cote qui servira au rangement des documents. Les archivistes auront plutôt recours à un rangement séquentiel. «Chaque article [ou versement] reçoit un numéro d'ordre de 1 pour le premier jusqu'au nième et dernier article du versement, sans jamais utiliser de bis ou ter» (Luquet, 1992, f. 5).

Les conséquences du classement séquentiel sont principalement de deux ordres. D'une part, ce mode de classement ne permet plus l'identification des sujets puisque les documents ne sont plus placés dans des séries thématiques. D'autre part, il affecte le regroupement des documents d'un même créateur. Les archivistes doivent donc trouver une autre façon de reconstituer les fonds d'archives. L'abandon de la classification oblige donc les archivistes à se doter d'autres instruments pour repérer l'information et pour respecter la notion de fonds.

La description et l'analyse par les bordereaux de versement

Les documents d'archives versés par les ministères dans les centres de préarchivage devaient recevoir un classement ultérieur après leur arrivée. Dans un premier temps, ces versements n'étaient accompagnés que d'un bordereau sommaire de description. Deux constatations se sont vite imposées suite à l'utilisation de cette procédure. Tout d'abord, il est apparu que, compte tenu de l'importance de la masse documentaire de ces versements, il serait utopique de penser réaliser «une refonte systématique des versements au moment de leur reclassement, accompagnée d'un travail scientifique approfondi» (L'Huillier, 1982, p. 203) à cause, entre autres, de la lourdeur de l'opération (René-Bazin, 1985a, p. 87). Il fallait donc que les missions fassent «dès le départ, œuvre définitive». De plus, on avait déjà constaté que les difficultés dans la recherche sur les versements déjà réalisés étaient dues «à l'imprécision des analyses» (L'Huillier, 1982, p. 203). D'où la nécessité d'élaborer un bordereau de versement plus complet qui servirait à la production d'instruments de recherche et permettrait le repérage.

Le bordereau de versement a évolué dans le temps et constitue désormais un véritable répertoire. Il comporte une «chemise» dont la couverture décrit les aspects matériels du versement³. Une analyse plus poussée des archives sérielles ou de versements plus importants ou jugés comme tels est disposée à l'intérieur de cette «chemise». (Richefort, 1987, p. 165-175). Les principales parties du bordereau de versement relatives à la description sont la cote, le nom du service producteur et du service servant, le sigle le plus couramment utilisé pour les nommer, l'année du versement, le nombre de cartons concernés, le métrage, sa localisation dans le dépôt, les dates extrêmes des documents, des données relatives à la date d'élimination ou à celle de la révision si la première n'est pas certaine, la mention relative à la conservation permanente s'il y a lieu, la communicabilité, le mode de classement et le support. D'autres parties du bordereau contiennent des renseignements de type analytique. Il s'agit de la description de l'unité documentaire dans laquelle l'archiviste indique, en vocabulaire

libre mais en tenant compte de l'indexation qui suivra, le contenu intellectuel de l'unité documentaire⁴.

Le bordereau de versement constitue donc un «instrument de recherche définitif à vocation multiple - administrative, juridique, archivistique, documentaire, informatique - «... [il] doit, dans la forme, réunir tous les éléments utiles à ces exploitations variées» (Pétillat et Prax, 1993, p. 275).

Le développement de concepts intellectuels propres à l'analyse des documents d'archives

L'analyse des documents d'archives consiste en «un texte rassemblant tous les éléments qui identifient un document, un article ou une catégorie de documents» (Naud, 1985, p. 54). Elle repose, en France, sur un versement qui est homogène quant à son contenu et qui provient du même créateur. Il s'agit habituellement d'un ou de plusieurs dossiers. L'archiviste de la mission, qui voit à organiser logiquement le versement, détermine lui-même le niveau d'analyse.

Le raisonnement fondamental qui est à la base des éléments retenus pour l'analyse est le suivant :

- le document est produit pour franchir le temps et l'espace;
- il est porteur d'information; il a donc un **objet**;
- le document isolé n'a de sens qu'avec les autres pièces d'un dossier et en relation avec d'autres dossiers;
- la nature du document en constitue la forme et non la raison d'être (**type**);
- le document résulte d'une action qui est en relation avec la mission de l'administration concernée (**action**) (Naud et Naud, 1982).

C'est donc à partir de ce «raisonnement» que se sont structurés les différents éléments de l'analyse et que s'est imposé le besoin de thésaurus ou de liste de vocabulaire contrôlé pour réaliser de façon efficace cette analyse, d'autant plus que le traitement documentaire informatisé devenait possible dans le temps; l'informatique «permettant d'introduire en mémoire des éléments que la machine puisse réellement enregistrer et classer» (Naud, 1985, p. 53). Le rôle de l'analyse est donc de comprendre l'action et l'objet, de les hiérarchiser et de les synthétiser ou de les rassembler (Naud, 1985, p. 56). Ces notions d'**objet**, d'**action** et de **type** se retrouvent dans les trois thésaurus que nous présenterons ci-après.

Le recours aux langages et aux techniques documentaires

Dans leurs réflexions sur le traitement analytique des archives, les archivistes ont été amenés à se tourner vers les principes et méthodes de l'analyse documentaire. C'est ainsi que parallèlement à l'analyse propre aux archives et à leurs composantes, les archivistes ont exploré les instruments de travail utilisés par les documentalistes. Comme l'affirmait l'une d'entre eux: «... une réflexion sur les index thématiques ne saurait se passer d'une réflexion sur le langage et les techniques documentaires» (Vallée, 1984, p. 233).

Les thésaurus, en permettant la structuration de l'analyse et, en conséquence, l'efficacité du repérage de l'information constituaient, pour les archivistes, une réponse aux problèmes soulevés par le classement séquentiel des documents. L'introduction du *Thésaurus W* présente l'esprit de cette nouvelle approche:

La méthode de traitement des archives [avant 1980 en France], fondée sur la cotation, c'est-à-dire sur la répartition des documents dans un cadre de séries thématiques, utilisait un *langage classificatoire* (le cadre de classement des séries anciennes et modernes), techniquement à l'opposé des *langages documentaires* parmi lesquels on range les thésaurus. Un langage classificatoire obéit en effet à une logique fonctionnelle et vise à organiser le plus rationnellement possible un rangement matériel. Ces préoccupations sont étrangères à la construction d'un langage documentaire qui cherche à faciliter l'exploitation d'un stock d'informations à l'aide d'une liste de concepts jugés pertinents (France, Direction des archives de France, 1989, p. 5).

Les archivistes résistent toutefois à utiliser des thésaurus déjà créés. Dès 1992, la Banque d'information politique et d'actualités (BIPA) dispose d'un thésaurus. Les archivistes identifient les caractéristiques qui différencient les documents d'archives de l'imprimé pour justifier la création de leur propre thésaurus. Ces caractéristiques reposent sur la nature de l'information contenue dans les uns et les autres. Ainsi, ils affirment que les archives constituent une...

source documentaire non rédigée, non élaborée selon une réflexion, composée d'un ensemble de pièces, manuscrites, dactylographiées, imprimées, diverses par leur fonction et par leur forme, etc., et surtout, qui expriment une action administrative dans ses différentes formes, échelonnées dans le temps, sur un ou plusieurs objets, alors que la source documentaire des autres organismes reflète déjà une réflexion à un deuxième niveau (Playout, 1983, p. 63).

Les archivistes n'en reconnaissent pas moins les qualités propres aux thésaurus. Ils s'initieront aux normes qui régissent la constitution de tels instruments et développeront des outils d'analyse correspondant à la forme et aux sujets de l'information constituée par les administrations.

L'expérience française repose donc sur la nécessité devant laquelle les archivistes ont été placés d'organiser l'information afin de pouvoir la retrouver pour les besoins des administrations. Ils ont voulu faire «œuvre durable» devant l'impossibilité qui leur est vite apparue de ne pouvoir re-traiter les archives qui étaient conservés à Fontainebleau, et devant l'ampleur de la masse de documents recueillie dans les départements. C'est ainsi qu'ils ont adapté leurs méthodes de travail à cette nouvelle réalité. Dans un premier temps, ils ont abandonné le système de classification utilisé pour les documents jusqu'en 1940 pour adopter un classement séquentiel des versements. Ils ont ensuite élaboré des bordereaux de versement qu'ils utilisent comme instruments de recherche pour les documents contemporains. Le développement et l'utilisation de thésaurus complètent cette nouvelle réalité archivistique française.

PRÉSENTATION DES THÉSARUS

Le thésaurus est «une liste d'autorité *organisée de descripteurs et de non-descripteurs* obéissant à des règles terminologiques propres et reliés entre eux par des *relations sémantiques* (hiérarchiques, associatives, ou d'équivalence). Cette liste sert à

traduire en un langage artificiel dépourvu d'ambiguïté des notions exprimées en langage naturel» (France. Direction des archives de France, 1989, p. 5).

Trois thésaurus sont particulièrement utilisés dans les archives en France. Il s'agit, par ordre chronologique de parution, du *Vocabulaire de Varenne*, du *Thésaurus W* et du thésaurus *Priam 3*⁵. Dans les pages qui suivent, nous présenterons chacun d'eux en rappelant leur origine, en décrivant leur format et en soulignant leurs principales caractéristiques.

Le *Vocabulaire de Varenne*

Le *Vocabulaire de Varenne* (France. Mission permanente..., 1986) est un vocabulaire utilisé pour l'indexation générale des documents inventoriés dans la base de données de *Varenne*⁶. Il a été préparé par les archivistes travaillant à la Mission permanente des Archives nationales auprès des services du premier ministre de France et s'applique aux documents produits par ces services. Il sert à l'analyse des dossiers et peut même permettre l'analyse de la pièce si l'une d'entre elles revêt une importance particulière. Cette base permet donc «l'inventaire, la communication et, dans la mesure où elle leur incombe, la gestion des archives» (René-Bazin, 1985b, p. 114). La dernière édition publiée du *Vocabulaire de Varenne* date de juin 1986 mais l'édition informatique est l'objet de mise à jour continue.

Au départ, ce thésaurus a été développé et utilisé dans la base de données *Varenne* dans le but de

rendre d'abord possible et ensuite rapide la recherche, publique dans un premier temps puis privée ultérieurement, sur les documents conservés étant entendu que, de plus en plus, submergé d'informations, le public n'utilisera que ce qu'il pourra sélectionner simplement par le biais probable d'outils informatiques (René-Bazin, 1985a, p. 87).

Le *Vocabulaire de Varenne* a ceci de particulier qu'il a été «réalisé d'après le *Dictionnaire des rubriques utilisées par la division des questions écrites* de l'Assemblée nationale» (France. Mission permanente..., 1986, page couverture). Il s'inspire aussi du thésaurus de la BIPA.

Le *Vocabulaire de Varenne* se divise en deux parties: le vocabulaire lui-même et en annexe «une liste... [des] descripteurs qui ne se suffisent pas à eux-mêmes et de ce fait doivent toujours être associés à d'autres» (France. Mission permanente..., 1986, p. [i]).

Ce vocabulaire comprend plus de 800 termes placés dans un seul ordre alphabétique. Ces termes sont structurés de façon hiérarchique. C'est par un jeu de termes en majuscules que l'on peut identifier la rubrique principale ou les grandes divisions, les termes en caractères minuscules gras indiquant les subdivisions de chacune des rubriques principales. Des explications complémentaires sont fournies pour aider l'indexeur. Ainsi, des «astérisques accolés à certains descripteurs signalent que des notes d'application, consultables à la mission des archives, en précisent les conditions d'utilisation» (France. Mission permanente..., 1986, p. [i]).

A titre d'exemple, nous avons choisi le terme «logement». Il s'agit d'un terme constituant une rubrique principale. Il se subdivise en:

- accession à la propriété
- allocation de logement
- amélioration de l'habitat
- assurance construction
 - att. : ce terme se trouvera aussi sous la rubrique «Bâtiment et travaux publics»
- baux d'habitation
- construction
 - att. : ce terme se trouvera aussi sous la rubrique «Bâtiment et travaux publics»
- habitation à loyer modique
- permis de construire
 - att. : ce terme se trouvera aussi sous la rubrique «Bâtiment et travaux publics»

Le nombre de subdivisions pour chacun des termes est très variable. Ainsi, pour le terme GRÂCE ET AMNISTIE, on n'en retrouve aucune, tandis que le terme JUSTICE comprend 21 subdivisions. Il faut ajouter de plus que ces subdivisions ne se font qu'à un niveau seulement.

Le thésaurus ainsi organisé est facile de consultation puisqu'il est placé dans un seul ordre alphabétique. Par contre, il est plus difficile de visualiser la hiérarchie. Les renvois sont identifiés par des «voir». Le système d'indexation est pré-coordonné, c'est-à-dire que les termes utilisés sont déjà combinés. Ainsi, l'on retrouvera les termes «énergie», «énergie solaire», «économies d'énergie».

Le *Vocabulaire de Varenne* est utilisé à la Mission permanente des Archives nationales auprès des services du premier ministre. Ce qu'il faut surtout en retenir, c'est la particularité de son point de départ, la base de ce vocabulaire étant les besoins des usagers de l'information et de leur vocabulaire. Le *Vocabulaire de Varenne* est construit à partir des questions écrites soumises à l'Assemblée nationale à l'instar du thésaurus de la BIPA. C'est en en tenant compte que sera analysée l'information contenue dans les documents.

Une autre caractéristique, c'est que le *Vocabulaire de Varenne* s'applique à la description au dossier et à la pièce. Sous cet aspect, il peut servir à une étude utile sur la portée que peut avoir le niveau de description dans le développement d'un thésaurus.

Ce vocabulaire constitue un instrument original dont l'analyse plus approfondie pourrait permettre d'évaluer l'efficacité de telles approches. Il y a peu d'expériences de ce type dans le milieu archivistique, la plupart d'entre elles reposant plutôt sur l'analyse des documents. Il peut servir de modèles à l'élaboration d'un thésaurus pour un petit centre d'archives.

Le Thésaurus W

Le *Thésaurus W* tient son nom du cadre de classification des archives de France qui réserve la cote W aux documents contemporains, c'est-à-dire ceux qui ont été produits avant le 10 juillet 1940⁷. Commencé dans les départements, il a été ensuite

développé pour les archives départementales par le Service technique de la Direction des Archives de France. Il repose, entre autres, sur la terminologie utilisée en droit public.

Le *Thésaurus W* se divise en deux parties: un thésaurus-matière et des listes d'autorité. Le thésaurus-matière comprend des tableaux thématiques auxquels s'ajoute une liste alphabétique qui permet de retrouver les termes dans les tableaux thématiques. Les listes d'autorité portent sur des «mots-outils» qui servent à décrire «l'action» qui portera sur «l'objet» contenu dans le tableau thématique. Une autre liste comprend une «typologie documentaire» c'est-à-dire une liste de types de documents les plus couramment utilisés dans l'administration.

Le tableau thématique du Thésaurus présente 11 chapitres sous lesquels on retrouve présentés, hiérarchiquement, des rubriques, des sous-rubriques et des descripteurs «simples». Les chapitres sont:

- Administration
- Agriculture
- Communications
- Économie
- Éducation
- Équipement
- Événementiel
- Justice
- Opinion
- Société
- Temps libre

Ce thésaurus reflète «les compétences de la puissance publique... et... des questions dites «de société» (France. Direction des archives de France, 1989, p. 7). Il comprend 975 descripteurs. Il comporte en moyenne trois niveaux de description.

Ainsi, le terme «logement» fait partie de la section:

- Immobilier
 - Logement
 - Expulsion locale
 - Loyer immobilier
 - Logement de fonction
 - Logement social
 - Habitation à loyer modéré
 - Logement-foyer

L'index comporte des termes rejetés ainsi qu'un réseau de termes associés.

Ce thésaurus est un instrument de travail réglementaire pour les archives départementales. La Circulaire AD 89-3 du 31 août 1989 en rend l'utilisation obligatoire dans les départements français à compter du 1er janvier 1990.

Le *Thésaurus W* s'applique à des versements des administrations départementales, donc à des masses importantes de documents. Il reflète les activités variées des régions et des départements. Il est utilisé dans plusieurs bases de données constituées

dans les départements. Il sert aussi à compléter l'index des instruments de recherche qu'ils publient. Il est géré à la Direction des Archives de France qui en assure le contrôle et la mise à jour. Son utilisation est obligatoire dans les départements pour indexer les versements de l'administration départementale mais aussi dans la production des instruments de recherche.

L'utilisation du *Thésaurus W* semble présenter certaines difficultés d'application selon les commentaires reçus des archivistes départementaux, comme la rigidité des concepts et l'inadéquation entre les termes employés et la réalité administrative d'un département. Toutefois, si plusieurs archivistes utilisent des termes qui n'apparaissent pas dans le thésaurus, ils s'en inspirent largement. Somme toute, il s'agit d'un instrument de base qui présente l'avantage de n'être pas trop complexe, pas trop volumineux tout en suivant les règles de l'art.

Le thésaurus *Priam 3*

Le thésaurus *Priam 3*⁸ constitue le dernier thésaurus élaboré en collaboration entre la Section des missions des Archives nationales de France et le Centre des archives contemporaines de Fontainebleau. Il s'applique aux documents gérés par ce Centre et envoyés par les missions des Archives nationales établies dans les différents ministères du gouvernement français. Ce thésaurus est donc utilisé dans la base de données du même nom par le Centre et les missions⁹.

Comme c'est le cas pour les autres thésaurus, il sert à compléter la partie analyse des bordereaux de versement et s'inscrit sur la fiche des unités documentaires. Il a été établi à partir d'un échantillon représentatif de termes introduits dans le champ *objet* de la base de données *Priam 3* depuis 1984.

Priam 3 est donc le titre d'un thésaurus en même temps que le nom d'une base de données. L'introduction d'une version papier que nous avons consultée, le présente comme un microthésaurus.

Il comprend 22 rubriques :

1. Institutions
2. Administration
3. Défense nationale
4. Relations internationales
5. Justice
6. Économie
7. Finances publiques
8. Agriculture
9. Industrie
10. Environnement
11. Ressources
12. Équipement
13. Transports
14. Société
15. Travail
16. Enseignement
17. Culture

- 18. Temps libre
- 19. Sciences
- 20. Communication
- 21. Histoire
- 22. Géographie

Chacune de ces rubriques peut comporter jusqu'à sept niveaux de descriptions: un terme générique, trois termes spécifiques et des termes faisant partie d'une liste ouverte mais dont l'ajout est facultatif. Le plus couramment, les archivistes utilisent cinq niveaux. Le thésaurus *Priam 3* existe en version papier mais il prend tout son sens dans la base de données dans lequel il s'inscrit, puisqu'il permet une recherche qui situe même le niveau où se retrouve le terme indexé. Par contre, l'index du thésaurus *Priam 3* ne peut être consulté sur une copie papier mais il est disponible à tous les usagers qui disposent de la base de données.

Dans le thésaurus *Priam 3* le terme «logement» se retrouve dans le chapitre ÉQUIPEMENT. Il se présente de la façon suivante:

LOGEMENT

NA⁰. Aide publique ou privée : surindexer éventuellement

Aide de l'État

AIDE AU LOGEMENT

Accession à la propriété

Aide personnalisée au logement

EP¹¹ *Allocation-logement*

Participation des employeurs

EP *1% logement*

HABITAT

NA. Mode d'organisation du cadre de vie et ensemble des conditions d'habitation, de logement

Amélioration de l'habitat

Bidonville

TA¹² **Habitat insalubre**

Grand ensemble d'habitation

Habitat insalubre

EP *Logement insalubre*

Immeuble collectif

Maison individuelle

LOGEMENT DE FONCTION

LOGEMENT SOCIAL

EP *Logement économique*

Cité d'urgence

NA. Terme usité dans les années 1950-1960

Habitation à bon marché

EP *HBM*

Habitation à loyer modéré

EP *HLM*

Le thésaurus *Priam 3* comprend aussi des noms propres d'organismes ou d'entreprises. Pour ces noms, il n'a pas de contrôle d'autorité, mais beaucoup d'informations sont déjà enregistrées à ce titre sur les bordereaux de versement, comme les changements de noms et de responsabilités.

Priam 3 constitue le thésaurus le plus développé dans les archives de France et ses auteurs ont été particulièrement soucieux de respecter les normes de rédaction d'un thésaurus. Par son volume, il s'applique à des masses importantes de documents. La qualité de sa construction, malgré la complexité qu'impose la présence d'un grand nombre de termes, en fait un instrument de grande qualité.

La principale valeur des thésaurus développés en France repose sur leur utilisation effective. Ils n'en sont plus à l'état d'expérience puisqu'ils sont utilisés depuis plus d'une décennie. Ces thésaurus ont déjà subi l'épreuve du temps. Ils ont été corrigés¹³, révisés et restructurés pour répondre aux besoins pour lesquels ils avaient été créés. De plus, ces thésaurus sont utilisés dans plusieurs ministères et départements. Ils ont donc déjà été confrontés à différents types de recherche et testés par nombre d'archivistes au cours des ans.

QUELQUES CONSIDÉRATIONS SUR LES THÉSAURUS

Après la présentation de ces recherches et de ces expériences sur le traitement descriptif et analytique des archives en France, quelques considérations s'imposent. Elles porteront surtout sur les différences d'approche qui ont présidé à la création des thésaurus, sur les conditions particulières de leur utilisation, et sur les problèmes que pose la présence de plusieurs thésaurus pour l'analyse de documents. De plus, nous tenterons de dégager des conclusions utiles pour les archivistes québécois.

La création des thésaurus

Les éléments qui ont servi de base à l'élaboration des thésaurus utilisés dans les services d'archives français sont différents les uns des autres. Il s'agit du vocabulaire relié aux questions posées à l'Assemblée nationale pour le *Vocabulaire de Varrenne*, de celui utilisé dans la description des documents versés aux archives intermédiaires - donc basé sur les documents - pour le thésaurus *Priam 3* et celui du droit public pour le *Thésaurus W*. Les exemples fournis pour le terme «logement» démontrent bien les différences auxquelles conduisent ces trois approches. Les mêmes termes reliés au mot «logement» se retrouvent dans l'un ou l'autre thésaurus mais il n'y a pas de concordance exacte entre eux. Il n'est pas de notre propos, ici, de prendre partie pour l'une ou l'autre approche. Toutefois, nous croyons que ces expériences peuvent servir dans la préparation de nos propres thésaurus au Québec. L'analyse systématique des thésaurus français permettra de valider l'une ou l'autre approche ou tout au moins de mesurer les conséquences d'un tel choix sur le vocabulaire adopté.

L'utilisation des thésaurus

Les thésaurus que nous avons présentés sont appliqués aux archives contemporaines. Il faudra probablement développer d'autres thésaurus pour les archives

anciennes. C'est ce que semblent appréhender les archivistes français eux-mêmes. Christine Nougaret l'affirme dans *La pratique archivistique française*.

De même qu'on a fixé le vocabulaire des archives contemporaines pour l'objet, l'action et la typologie, on devra établir des thésaurus ou des vocabulaires contrôlés de mots-matière pour les autres concepts – mais aussi pour les documents plus anciens – ... (Nougaret, 1993, p. 146).

Il y a donc un autre défi à relever par rapport à ces autres documents. Toutefois, cette considération vaut-elle dans le contexte québécois? À quelle date faut-il faire remonter les «documents plus anciens»? Au Régime français? À la Confédération? L'expérience française permet, sous cet autre volet, de se poser ces questions avant d'appréhender la création de nos propres thésaurus.

Parce que les thésaurus sont développés en fonction des documents produits par les administrations contemporaines, pourront-ils répondre aux besoins des chercheurs autres que ceux des administrations? Autrement dit, pourront-ils aussi servir pour la recherche? Cette question a déjà été soulevée par Jean Favier, dès 1983, alors qu'il était directeur des Archives nationales de France et lors de la présentation des premières expériences sur le sujet. Il se disait alors «inquiet de l'aspect ponctuel que présente ce type d'interrogation qui ne lui semble pas répondre à une interrogation de type historique...». Et il conclut en «rappelant qu'il est indispensable de conserver les deux fonctions également essentielles de l'archiviste: la fonction de documentation qui est vitale à la gestion de la chose publique; la fonction historique, indispensable aux historiens» (Abrégés des débats, 1985, p. 102). Cette appréhension ne peut encore être confirmée ou niée puisque les archives indexées par les archivistes ne sont pas encore directement accessibles aux chercheurs mais la question demeure. Dans le contexte québécois, la même problématique existe. Les index préparés pour les dossiers et les documents à leur stade actif pourront-ils être utilisés lorsque les documents seront versés aux archives historiques? Pour la France comme pour le Québec, le recul dans le temps n'est pas suffisant pour permettre de répondre à cette question. Toutefois, nous avons l'avantage de pouvoir nous la poser avant même que les thésaurus soient élaborés.

Les thésaurus pourront-ils permettre l'autonomie du chercheur lors d'une démarche de recherche? Même si des chercheurs peuvent consulter actuellement des documents contenus dans les bases de données dans lesquelles sont utilisés les thésaurus, ils y ont accès par l'instrument de recherche papier constitué par les bordereaux de versement ou encore parce qu'un archiviste aura fait la recherche pour eux dans la base. Les thésaurus n'ont donc pas encore subi le test des questions des chercheurs. À ce titre, on peut supposer que le *Vocabulaire de Varenne* risque de mieux performer puisque son point de départ n'est pas le document lui-même mais les questions ou les demandes d'information que peuvent demander les députés en Chambre. La relation est donc déjà établie entre le demandeur et l'information, mais la vérification d'une telle hypothèse n'est pas encore réalisée ni en France et encore moins au Québec.

L'utilisation des notions d'objet, d'action et de type de documents

L'utilisation des notions d'objet et d'action ainsi que les types de documents comme concepts de base du repérage développés par les archivistes français méritent

d'être étudiés de près. Malgré les difficultés soulevées par les archivistes relativement à l'application des notions d'objet et d'action, il demeure qu'ils constituent des concepts intellectuellement appropriés à la réalité de l'information administrative. Ils méritent d'être étudiés et testés dans le contexte de la réalité québécoise pour conduire, sinon à leur adoption, du moins à la découverte d'approches plus appropriées.

Quant à l'utilisation de la typologie des documents comme élément d'analyse, elle est aussi une voie à explorer. Les archivistes français ont une bonne connaissance de la diplomatique. La discipline s'est développée chez eux et elle fait partie de leur formation à l'École des Chartes. L'application à l'analyse des documents dans un but de repérage et non plus seulement dans un objectif d'authentification constitue une nouvelle utilisation de la diplomatique dans un contexte contemporain. Au Québec, la typologie des documents est déjà utilisée dans l'élaboration des calendriers de conservation. Pourrait-elle aussi servir à la recherche? Constitue-t-elle une approche utilisée par les chercheurs? Répond-elle aux besoins des méthodologies utilisées? Voilà autant de questions qu'il faut se poser, toutefois, avant d'adopter ce mode d'analyse.

L'analyse, base de l'indexation

Les thésaurus produits en France servent à l'indexation de grandes séries ou dossiers déjà décrits sur les bordereaux de versement. Leur création et leur utilisation reposent sur une analyse préalable des documents. Cette démarche constitue l'approche logique à toute indexation.

Au Québec, de telles descriptions existent déjà pour les documents historiques dans les instruments de recherche. Depuis l'adoption des *Règles de description des documents d'archives* (RDDA), le contenu de ces descriptions est plus formalisé. Les règles relatives à la zone 7 «Portée et contenu» précisent même le contenu de cette description.

L'expérience française nous démontre le lien qui existe entre la description et l'analyse. Les RDDA constitueront une base sur laquelle pourra reposer une telle analyse.

Les thésaurus et les noms propres

L'index d'un instrument de recherche comporte des noms communs et des noms propres. Pour l'établir, l'archiviste utilisera deux instruments différents. Pour le choix des noms communs, il se servira de thésaurus. Ceux-ci servent à contrôler le vocabulaire dans les index et à établir les «relations sémantiques» qui y seront utilisées. Les noms propres seront tirés de la liste d'autorité constituée des noms propres dont la forme a été établie à partir des règles relatives aux noms propres.

Priam 3 comporte une liste de noms propres mais ils ne font pas partie du thésaurus lui-même. Les autres thésaurus ne comportent pas de noms propres.

Il faut retenir que la liste d'autorité est aux noms propres ce que le thésaurus est aux noms communs. L'un et l'autre sont des instruments servant à l'indexation mais ils constituent des outils de travail distincts dont les archivistes doivent se doter afin de favoriser le repérage de l'information.

Un thésaurus à la mesure de ses besoins

L'exemple français nous présente trois types de thésaurus ou vocabulaire. Le *Vocabulaire de Varenne* constitue un outil facile à utiliser par les archivistes, les indexeurs et les chercheurs. Il présente une grande souplesse de création, de mise à jour tout en permettant la hiérarchisation des termes. Cette façon de procéder, par contre, ne permet pas le développement équilibré des différentes rubriques. Toutefois, il répond aux besoins du milieu qui l'a créé. À l'autre extrême, le thésaurus *Priam*³ présente une complexité de structure et d'application. Il permet jusqu'à sept niveaux de description. Cette complexité se justifie toutefois par la masse des documents à indexer. Le *Thésaurus W* se situe entre les deux autres modèles quant à sa dimension et à sa structure. Il respecte les normes de rédaction d'un thésaurus tout en étant peu volumineux.

Les trois exemples fournis présentent des modèles de thésaurus différents. Ils ont été développés pour des besoins divers et en rapport avec la masse documentaire. Ils portent toutefois sur le même genre de documents, ceux de l'administration publique. Ces exemples nous permettent de constater qu'il peut y avoir plusieurs types de thésaurus. Le meilleur n'est-il pas celui qui répond le mieux aux besoins pour lesquels il est créé en autant qu'il respecte les règles de l'art?

L'expérience de l'élaboration et de l'utilisation de thésaurus en France peut-elle être profitable aux archivistes québécois? Nous le croyons et ce, à plusieurs titres. Il faut reconnaître que le fait que ces thésaurus soient en langue française présente un atout majeur. La littérature en archivistique fait souvent référence au thésaurus *Art and Architecture* (Peterson, 1990) comme un modèle utile aux archivistes. Toutefois, le fait qu'il soit en anglais, constitue un obstacle à son utilisation. Tout au plus, peut-on l'utiliser à des fins de référence pour la structure.

Au moment où les archivistes québécois disposent de normes de description permettant l'uniformisation des descriptions des archives, où des recherches sont proposées sur leur utilisation éventuelle dans la description de dossiers à l'actif (CREPUQ, 1994), le développement de thésaurus s'impose. D'ailleurs, le Groupe de travail sur l'indexation par sujet dans les archives (BCA, 1992) recommandait le développement de ces instruments de travail. Actuellement, quelques services d'archives commencent à normaliser l'utilisation de thèmes qui sont abordés dans les fonds. L'index d'un instrument de recherche, par exemple, comporte des noms communs et des noms propres. De plus, l'on voit apparaître différents types d'index à la fin des instruments de recherche (index des professions, index des lieux, index des thèmes, etc.). L'expérience française présente donc, à ce titre, une référence des plus intéressantes que les archivistes québécois ont intérêt à observer et à analyser.

Louise Gagnon-Arguin L'auteure est professeure agrégée à l'École de bibliothéconomie et des sciences de l'information de l'Université de Montréal.

NOTES

1. Nous tenons à remercier Mme Paule René-Bazin, responsable du Stage, qui a accepté de relire ce texte avant sa publication. Mme Bazin nous a aussi fait connaître ces expériences et mis en contact avec les principaux artisans de ces thésaurus qui ont accepté de nous rencontrer. Il s'agit de Mesdames Christine Naud, conservateur au Centre d'archives contemporaines de Fontainebleau et de Marie-Josée Corcuf, conservateur à la Direction des Archives de France, de Messieurs Jean Luquet, conservateur et responsable de la Mission des Archives nationales auprès des Services du premier ministre, de François Gasnault, chef du Service des archives au ministère de l'Économie et des Finances.
2. *Égérie* constitue l'index de l'État général des fonds des Archives nationales de France. « Cette base est l'index des quatre premiers volumes de l'*État général des fonds* des Archives nationales. Elle contient 29 149 unités documentaires et est le point de départ de toute recherche d'ordre général dans les fonds des Archives nationales » (France. Archives nationales, 1993, p.18).
3. Le versement dont il est question ici est utilisé dans le sens que les archivistes français lui donnent, à savoir le dépôt de documents dans un centre de préarchivage.
4. La division en description et analyse est de nous. Dans le milieu archivistique français, l'action de remplir ce bordereau constitue l'analyse. Comme dans le contexte québécois, les règles pour décrire et analyser sont différentes, nous avons cru bon de les distinguer dans ce travail.
5. D'autres thésaurus sont utilisés par les archivistes pour indexer les documents graphiques ou figurés (Garnier) et pour les documents architecturaux mais nous n'en ferons pas état ici.
6. La base de données *Varenne* a été constituée à partir d'un inventaire des documents versés par les archivistes de la Mission et utilisée pour la gestion et la communication des archives (Richefort, 1987, p. 170).
7. Afin de distinguer les archives contenues dans les archives départementales de celles des archives nationales, la lettre W précède le numéro du versement aux Archives nationales (W1800) tandis que, dans les archives départementales, elle le suit (1800W).
8. *Priam 1* (1977) constitue une base de données destinée à gérer les espaces au Centre de Fontainebleau, *Priam 2* (1977) sert à gérer l'information et les communications administratives entre les missions et le Centre de Fontainebleau, tandis que *Priam 3* (1984) constitue la base de données documentaires.
9. À l'exception des services du premier ministre et des services de la Présidence de la République qui utilise la base de données *Varenne*.
10. NA donne la définition exacte ou les conditions d'emploi du descripteur.
11. EP signifie «Employer pour...»
12. TA signifie «Terme associé».
13. Par exemple, le *Thesaurus W* en est à sa troisième édition qui paraîtra bientôt. Au *Vocabulaire de Varenne*, on a ajouté une table hiérarchique.

BIBLIOGRAPHIE

- «Abrégés des débats» (1985). *Actes du XXVI^e congrès national des archivistes français, Angers, 5-7 octobre 1983*, Paris, Archives nationales, 1985, p. 102-105.
- Aitchison, Jean et Alan Gilchrist (1992). *Construire un thésaurus*. Paris, ADBS, 1992, 217 p.
- Bertrand-Gastaldy, Suzanne (1986). «De quelques éléments à considérer avant de choisir un niveau d'analyse ou un langage documentaire». *Documentation et bibliothèque*, vol. 32, n^{os} 1-2 (janvier-juin 1986), p. 3-23.
- BIPA: thésaurus d'information politique, économique et sociale* (1992). 2^e éd. Paris, Direction de la documentation française, 1992, 317 p.
- Bureau canadien des archivistes (BCA) (1992). *L'indexation par sujet en archivistique: rapport du Groupe de travail sur l'indexation par sujet*. Ottawa, Bureau canadien des archivistes, Comité de planification sur les normes de description, 1992, xii, 154 p. (Publication n^o 4)
- Conférence des recteurs et des principaux des universités du Québec (CREPUQ) (1994). *La gestion des archives informatiques*. Québec, Presses de l'Université du Québec, 1994, xi, 163 p.
- Cook, Michael and Margaret Procter (1989). *A Manual of Archival Description*, 2nd ed. Aldershot, Gower, 291 p.
- France. Archives nationales (1993). *Guide du lecteur* / par Gérard Ermisse et Philippe Le TrigUILly, Paris, Les Archives, 77 p.
- France. Direction des archives de France (1989). *Thésaurus W*. 2^e éd. Paris, 1989, 249 p.
- France. Mission permanente des Archives nationales auprès des services du Premier ministre (1986). *Vocabulaire utilisé pour l'indexation générale des documents inventoriés dans la base de données Varenne*. Paris, 1986, 111 f.
- Hensen, Steven L., comp. (1989). *Archives, Personal Papers, and Manuscripts: a cataloging manual for archival repositories, historical societies, and manuscript libraries* / 2nd ed. Chicago, Society of American Archivist, 1989, ix, 196 p.
- Hudon, Michèle (1994). *Le thésaurus: conception, élaboration, gestion*. Montréal, Asted, 1994, 220 p. (Clé en main)
- Jacquard, Alex (1991). *Base de données «ARCADE»: système descriptif*. Paris, Archives nationales, 1991, 119 p. (La Documentation française)
- Labat-Poussin, Brigitte (1984). «“Arcade”: traitement par l'informatique d'un fonds des archives nationales», *Gazette des archives*, nos 125-126 (décembre 1984), p. 249-256.
- Lejeune, Anne, Isabelle Pébay et Daniel Pierre (1991). «Le projet informatique “Sirius” des archives nationales», *La Gazette des archives*, nos 152-153 (mai 1991), p. 152-153

- L'Huillier, Hervé (1982). «La rédaction des bordereaux de versement». *La Gazette des archives* n° 115 (mars 1982), p. 203-215.
- Luquet, Jean (1991). «Les archives du ministère de l'Équipement de la Grande Arche de la Défense», *La Gazette des archives*, nos 152-153 (mai 1991), p. 117-124.
- Luquet, Jean (1992). *Les archives dans les services administratifs: constitution, classement, archivage*. Paris, Archives nationales, 1992, 19 f.
- Maroteaux, Vincent (1987). «Informatique et archives contemporaines, une approche nouvelle», *La Gazette des archives*, nos 137-138 (décembre 1987), p. 144-149.
- Naud, Christine (1985). «L'analyse archivistique». *Actes du XXVI^e congrès national des archivistes français, Angers, 5-7 octobre 1983*, Paris, Archives nationales, 1985, p. 53-59.
- Naud, Gérard et Christine Naud (1982). «L'analyse des archives administratives contemporaines». *La Gazette des archives*, n° 115 (mars 1982), p. 216-245.
- Nougaret, Christine (1993). «Classement et description: des principes à la pratique», *La pratique archivistique française*, sous la direction de Jean Favier assisté de Danièle Neirinck, Paris, Archives nationales de France, 1993, p. 135-186.
- Peterson, Toni, ed. (1990). *Art and Architecture Thesaurus*. New York Oxford University Press, 1990, 3 vol.
- Pétilat, Christine et Hélène Prax (1993). «Les archives contemporaines ou l'arrivée du flux au quotidien». *La pratique archivistique française*, sous la direction de Jean Favier assisté de Danièle Neirinck, Paris, Archives nationales de France, 1993, p. 231-312.
- Playout, Arlette (1983). «Thesaurus et langage commun: bilan d'expériences». *Actes du XXVI^e congrès national des archivistes français, Angers, 5-7 octobre 1983*, Paris, Archives nationales, 1985, p. 60-82.
- Règles pour la description des documents d'archives* (RDDA) (1990) ...élaboré sous la direction du Comité de planification sur les normes de description, Ottawa, Bureau canadien des archivistes, 1990, 1 vol. (f. mobiles)
- René-Bazin, Paule (1985a). «Renouveler la recherche: l'application de VARENNE», *Actes du XXVI^e congrès national des archivistes français, Angers, 5-7 octobre 1983*, Paris, Archives nationales, 1985, p. 86-100.
- René-Bazin, Paule (1985b). «Vers une informatique archivistique». *La Gazette des archives*, n° 129 (novembre 1985), p. 114-121.
- Richefort, Isabelle (1987). «Les instruments de recherche élaborés par les missions». *La Gazette des archives*, nos 137-138 (décembre 1987), p. 165-175.
- Richefort, Isabelle et Christiane Naud (1991). «L'édition des instruments de recherche de fonds contemporains». *La Gazette des archives*, nos 152-153 (mai 1991), p. 88-99.
- Vallée, Aline (1984). «Index thématiques, tableaux méthodiques et thesaurus: luxe ou nécessité». *La Gazette des archives*, nos 125-126 (décembre 1984), p. 232-248.