La recherche documentaire par vedettes-matière dans une bibliothèque de collège
The access to information through subject headings in a college library
La investigación bibliográfica por materias en una biblioteca de un colegio universitario

Julien Marquis

Volume 29, Number 4, October–December 1983

URI: <https://id.erudit.org/iderudit/1053609ar>

DOI: <https://doi.org/10.7202/1053609ar>

[See table of contents](#)

Publisher(s)

Association pour l'avancement des sciences et des techniques de la documentation (ASTED)

ISSN

0315-2340 (print)

2291-8949 (digital)

[Explore this journal](#)

Cite this article

Marquis, J. (1983). La recherche documentaire par vedettes-matière dans une bibliothèque de collège. *Documentation et bibliothèques*, 29(4), 147–154.
<https://doi.org/10.7202/1053609ar>

Article abstract

After stating the main difficulties of the search for information through subject headings, the author reports on a survey made among students using the subject heading catalog in a college library. The percentage of agreement between terms used by students and subject headings is very high, yet the main problem resides at the level of specificity. An analysis of the percentage of recall terms and of the precision of the subject headings reveals also a high efficiency. Some propositions are offered to improve the students' performance and the effectiveness of subject headings.

La recherche documentaire par vedettes-matière dans une bibliothèque de collège

Julien Marquis*

Bibliothèque de Québec

Après avoir fait état des principales difficultés de la recherche documentaire par vedettes-matière, l'auteur présente les résultats d'une enquête effectuée auprès d'étudiants qui consultaient le catalogue sujets d'une bibliothèque de collège. Le taux de concordance entre les termes utilisés par les étudiants et les vedettes-matière est très élevé, mais la principale difficulté reste l'atteinte du niveau de spécificité requis. Une analyse du taux de rappel et du taux de précision des vedettes-matière révèle également un degré d'efficacité élevé. Quelques recommandations sont faites pour améliorer les performances des étudiants et l'efficacité des vedettes-matière.

The access to information through subject headings in a college library

After stating the main difficulties of the search for information through subject headings, the author reports on a survey made among students using the subject heading catalog in a college library. The percentage of agreement between terms used by students and subject headings is very high, yet the main problem resides at the level of specificity. An analysis of the percentage of recall terms and of the precision of the subject headings reveals also a high efficiency. Some propositions are offered to improve the students' performance and the effectiveness of subject headings.

La investigación bibliográfica por materias en una biblioteca de un colegio universitario

Después de plantear las dificultades fundamentales de la investigación bibliográfica por materias, el autor presenta los resultados de una encuesta efectuada entre estudiantes que consultaban el catálogo por materias en la biblioteca de un colegio universitario. El índice de concordancia entre los términos utilizados por los estudiantes y los de las entradas del catálogo por materias es muy alto, pero la dificultad principal sigue siendo el conseguir el nivel de especificidad requerido. Un análisis de los índices de poder de evocación y de los de precisión de las entradas del catálogo de materias señala también un alto nivel de eficacia. Se hacen algunas sugerencias de cara a mejorar la actuación de los estudiantes y la eficacia de las entradas del catálogo por materias.

Dans les bibliothèques et les centres de documentation, on utilise les langages documentaires pour représenter les documents et permettre leur repérage. Ces langages sont une condensation du contenu des documents et ils jouent le rôle de lien entre le langage des documents et celui des questions. Une bonne définition serait celle-ci :

Ensemble de termes normalisés avec éventuellement leurs relations mutuelles; cet ensemble est utilisé pour représenter (c'est-à-dire indexer) le contenu de documents quels qu'ils soient (écrits, visuels ou audiovisuels) aux fins de recherches ultérieures, et pour exprimer les questions¹.

Il y a trois genres de langages documentaires selon qu'ils sont basés sur les mots, sur les concepts ou sur les sujets. L'indexation à l'aide des mots donne les unitermes, celle qui utilise les concepts, les descripteurs ou mots-clés et l'indexation des sujets donne les vedettes-matière.

Ce dernier type de langage documentaire est utilisé dans la grande majorité des bibliothèques nord-américaines pour le repérage par sujets. Toutefois, son taux d'efficacité ne fait pas l'unanimité, comme l'ont démontré de nombreuses recherches.

Il nous semble intéressant de faire connaître les résultats d'une recherche que nous avons

* L'auteur est chef des services techniques à la Bibliothèque de Québec.

1 « Informatique documentaire: langage documentaire ou langage naturel? ». *Documentaliste*, vol.11 no 1 (mars 1974), 48.

menée à la bibliothèque du Cégep de Rimouski. Dans un premier temps, voyons quels sont les problèmes liés à l'utilisation des vedettes-matière.

Le problème

Divers reproches sont communément adressés aux vedettes-matière.

Il y a d'abord le manque de principes à la base de leur développement. On a toujours défini le catalogue sujets comme un instrument pragmatique répondant à des besoins immédiats. Il n'y a pas eu à l'origine de grandes recherches théoriques visant à l'établissement des principes de base, de généralisation de faits et d'expériences destinées à guider le travail dans le domaine.

Le premier code fut formulé par Cutter en 1876². Il fut longtemps le code de base régissant à la fois la constitution des catalogues auteurs-titres et des catalogues sujets. Puis il y eut la publication des règles de la Bibliothèque apostolique vaticane, d'abord en italien, puis en anglais en 1948 et en français en 1950³. Essentiellement les mêmes principes étaient exprimés mais avec davantage de règles pour couvrir les cas spécifiques.

C'est cependant Haykin qui, en publiant *Subject Headings: a Practical Guide*⁴ en 1951, offrit ce qui se rapproche le plus d'un code. C'était pourtant une description détaillée des pratiques particulières de la Library of Congress. Plus récemment, Lois Mai Chan⁵ en réexamina les principes de base et en décrivit les pratiques courantes. Mais il se défendit bien d'avoir voulu compiler des règles.

L'absence de base théorique a conduit à des incohérences et à des irrégularités dans le développement des listes de vedettes-matière.

La structure des vedettes-matière et sa très grande complexité est également un sujet de reproche. Il y a sept formes possibles de vedettes: Les mots simples (BOTANIQUE), les mots avec adjectifs (BOTANIQUE MÉDICALE), les mots avec adjectifs inversés (DIGESTIF, APPAREIL), les expressions (COÛT ET NIVEAU DE LA VIE), les expressions inversées (APTITUDE, INFLUENCE DE L'ÂGE SUR L'), les vedettes composées (RELIGION ET SOCIOLOGIE) et, enfin, les vedettes complexes (CRIMES ET DÉLITS POLITIQUES). Certaines de ces formes ne posent aucun problème. Par contre, lorsque l'accès est modifié, comme dans le cas des inversions, il y a risque de tâtonnement.

La raison invoquée pour créer les inversions est le regroupement de sujets semblables en faisant ressortir le mot le plus significatif d'une expression. Les inversions sont surtout utilisées pour les vedettes de langue anglaise, car la structure de la langue française se prête peu à cet exercice.

La terminologie des langages documentaires revêt également une importance primordiale. Pour qu'un système d'information soit efficace, les termes utilisés doivent assurer la cohérence de la représentation du contenu des documents lors de l'indexation et lors de la recherche. L'usage courant, soit l'emploi des termes les plus susceptibles d'être utilisés, demeure le principe de base du choix de la terminologie.

La variation du sens des mots et la désuétude du vocabulaire créent des problèmes de terminologie. Les nombreux développements dans certains domaines font apparaître des mots nouveaux qui ne sont pas toujours intégrés rapidement dans les listes de vedettes-matière. On continue ainsi à indexer des documents récents avec des termes dépassés ou trop généraux. La publication de mises à jour régulières des répertoires en usage atténue ce problème mais ne peut l'éliminer complètement.

C'est surtout la notion de spécificité qui, appliquée aux vedettes-matière, pose le plus gros problème. La règle élaborée par Cutter se lit ainsi: «Enter a work under its subject heading, not under the heading of a class which includes that subject»⁶. La vedette-matière devrait, par conséquent, être la représentation la plus fidèle possible du sujet d'un ouvrage.

Il faut par contre distinguer une vedette dont le sens est large, c'est-à-dire qui englobe un ensemble de concepts, d'une vedette générale manquant de spécificité. De la même façon il faut différencier une vedette spécifique d'une vedette englobant un nombre restreint de concepts. Par exemple, la vedette ZOOLOGIE est considérée comme une vedette large et la vedette CHAT comme une vedette plus étroite. La spécificité de chacun de ces termes dépend du contexte. La vedette ZOOLOGIE est trop générale si on l'applique à des ouvrages sur les chats. D'un autre côté, appliquée à des ouvrages sur la zoologie, la vedette ZOOLOGIE est aussi spécifique que la vedette CHAT attribuée à des ouvrages sur les chats.

Les usagers du catalogue sujets éprouvent beaucoup de difficultés à atteindre le niveau de

2. Charles A. Cutter, *Rules for a printed Dictionary Catalog*, Washington, D.C., Government Printing Office, 1876, 89p.

3. Bibliothèque apostolique vaticane, *Règles pour le catalogue des imprimés*, éd. française, Cité du Vatican, 1950, 402p.

4. David J. Haykin, *Subject Headings, a Practical Guide*, Washington, D.C., Government Printing Office, 1951, 140p.

5. Lois Mai Chan, *Library of Congress Subject Headings: Principles and Application*, Littleton, Colorado, Libraries Unlimited, 1978, 347p. (Research Studies in Library Science, 15).

6. Charles A. Cutter, *Rules for a Dictionary Catalog*, 2d ed., Washington, D.C., Bureau of Education, 1889, p. 66.

spécificité requis dans leurs recherches. Les études de Bates et de Knapp l'illustrent bien.

Marcia Jeannes Bates a mené une expérience au cours de laquelle des étudiants devaient trouver le sujet d'un livre de façon à vérifier la concordance de ce sujet avec les vedettes-matière du catalogue. En bref, «The rule of specific entry was found to cause a great deal of trouble. Students failed to get the same specificity level as used by the library about half the time»⁷.

Patricia Knapp, de son côté, a découvert que 31% des étudiants de niveau collégial utilisaient un terme plus général que la vedette-matière du catalogue et que 15,7% utilisaient un terme trop spécifique. Elle admet que la logique de la règle de la spécificité est inattaquable mais que son application devrait être mieux justifiée:

It follows that if the rule of specific entry is essential to subject cataloging, some justification must be made for it other than simply that of service to the user. It may be that the usefulness of the principle can be demonstrated, but it is evident that the principle causes difficulty to some users⁸.

Quant aux subdivisions, elles représentent des problèmes différents selon qu'il s'agit de subdivisions de sujets, de lieux géographiques, de périodes historiques ou de formes.

Les subdivisions de sujets sont celles où un sujet est utilisé en subdivision d'une vedette comme, par exemple, ENFANTS—MALADIES. Haykin n'admettait pas cette forme:

The use of topics comprehended within a subject as subdivisions under it is to be avoided. It is contrary to the principle of specific entry, since it would, in practice, result in an alphabetic-classed catalog...⁹

Pour sa part, Chan souligne qu'il n'y a pas de cohérence dans l'utilisation de vedettes de ce type:

Again the problem is in maintaining consistency in similar headings... It appears that these classed entries have been introduced into LCSH sporadically without sharply defined guidelines which ensure consistency in form and in scope¹⁰.

Les subdivisions géographiques sont celles qui servent à limiter un sujet à une région ou à une aire géographique particulière; on a alors des vedettes comme AGRICULTURE—CANADA. Cependant, une autre forme de vedettes représente aussi l'aspect géographique d'un sujet: il s'agit d'un nom de lieu suivi d'une subdivision sujet (CANADA—CIVILISATION). Il ne semble

pas y avoir de critères précis pour l'établissement de l'une ou de l'autre forme. Un groupe de travail formé par le comité des services techniques de l'ASTED formulait récemment cette critique à l'endroit du *Répertoire de vedettes-matière* de l'Université Laval:

Comment comprendre que des vedettes comme POLITIQUE FISCALE, POLITIQUE MONÉTAIRE, qui peuvent se subdiviser de façon «Indirecte» (i.e. qu'on peut faire suivre d'une subdivision géographique) côtoient d'autres vedettes du même type comme POLITIQUE COMMERCIALE, POLITIQUE ÉCONOMIQUE, lesquelles ne s'emploient géographiquement qu'en subdivision aux noms des pays, régions, etc. et ne peuvent pas par conséquent être subdivisées de façon «Indirecte». On retrouve donc POLITIQUE FISCALE—CANADA, mais CANADA—POLITIQUE COMMERCIALE. Une ligne de conduite unique pour ce genre de vedette devrait s'imposer¹¹.

Les subdivisions de périodes historiques sont utilisées pour limiter un sujet à une tranche chronologique particulière. Le seul problème qu'elles occasionnent se pose lors de l'intercalation des fiches au catalogue étant donné la multiplicité et la variété de ces subdivisions.

Quant aux subdivisions de forme, elles expriment la forme particulière des documents traitant d'un sujet. Elles ne font pas problème.

Méthodologie

Précisons brièvement le cadre de notre enquête. La bibliothèque du Cégep de Rimouski comptait environ 80 000 documents au moment où s'est déroulée l'enquête, entre novembre 1980 et la fin de mars 1981. La population visée était la «clientèle» étudiante régulière.

Nous avons utilisé un questionnaire comme moyen de cueillette des données. Ce questionnaire était distribué aux usagers du catalogue sujets. Ils devaient indiquer le sujet de leur recherche et les termes utilisés pour trouver leur documentation. La concordance entre les termes utilisés et les vedettes-matière du catalogue était le critère de succès.

Un certain nombre de questionnaires furent par la suite retenus pour l'évaluation du taux de rappel et du taux de précision: ceux dont les sujets de recherche étaient suffisamment précis. Les deux bibliothécaires de référence refaisaient à leur tour la recherche au catalogue sujets en indiquant les vedettes-matière appropriées ainsi que les documents répertoriés. Ils en faisaient une évaluation d'après leur pertinence.

7. Marcia J. Bates, *Factors Affecting Subject Catalog Search Success*, Berkeley, University of California, 1967, p. 11.

8. Patricia B. Knapp, "The Subject Catalog in the College Library: an Investigation of Terminology", *Library Quarterly*, vol. 14, no. 3 (July 1944), 223.

9. David J. Haykin, *Subject Headings...*, p. 35-36.

10. Lois Mai Chan, "Alphabetical Arrangement and Subject Collocation in LCSH", *Library Resources and Technical Services*, vol. 21, no. 1 (Spring 1977), 165.

11. «Le Répertoire de vedettes-matière de Laval (huitième édition): analyse critique», *Documentation et bibliothèques*, vol. 25, no 3 (septembre 1979), 152.

Une recherche parallèle était aussi conduite de façon à trouver des documents appropriés par d'autres moyens que les vedettes-matière: le système de classification Dewey et les bibliographies surtout.

Résultats de l'enquête

Sur un total de 204 questionnaires distribués, 169 ont été jugés valables et retenus; de ce nombre 107 ont été remplis par les étudiants du secteur général et 62 seulement par des étudiants du secteur professionnel même s'ils représentent la majorité du collège.

Tableau 1

**Nombre de répondants
par secteurs selon leur succès**

Secteur	Succès	Échecs	Sans concordance	Total	%
Général I	40	8	1	49	28,99
Général II	44	10	4	58	34,31
Profess. I	20	6	2	28	16,56
Profess. II	21	3	3	27	15,97
Profess. III	4	1	2	7	4,14
Total	129	28	12	169	100

Le tableau 1 révèle que sur 169 étudiants, 129 ont eu du succès dans leur démarche et qu'ils ont obtenu au moins une concordance entre les termes utilisés et les vedettes-matière du catalogue, 28 étudiants n'ont utilisé aucun terme correspondant et ont donc échoué. De plus, douze recherches ont été considérées comme des échecs pour des raisons qui ne sont pas attribuables au catalogue: il s'agissait de recherches très spécifiques nécessitant l'utilisation d'autres outils documentaires, tels les encyclopédies et les périodiques. Le taux de succès fut donc de

76,33%. En d'autres mots, trois étudiants sur quatre ont réussi.

Le tableau 2 donne la répartition des réussites et des échecs par secteurs selon qu'on a utilisé un ou plusieurs termes. Soixante-dix étudiants n'ont cherché qu'à une seule vedette-matière. De ce nombre, 52 ont utilisé un terme susceptible de leur fournir la documentation désirée alors que 18 ont utilisé un terme inadéquat. Quatre-vingt-dix-neuf étudiants ont utilisé plus d'un terme: 77 ont réussi à établir une concordance et 22 ont échoué.

Un total de 359 termes ont été utilisés par les étudiants dont 203 furent identiques à la vedette-matière qu'il fallait utiliser. On employa 129 termes non concordants et 27 sans correspondance. Ces derniers étaient ceux des 12 recherches pour lesquelles le catalogue sujets n'était d'aucune utilité.

Nous demandions également aux étudiants de noter les titres des documents repérés au catalogue. Le nombre de documents indiqués varie de 0 à 23, mais la grande majorité des étudiants ont inscrit entre un et cinq documents. Un total de 474 documents ont été repérés soit une moyenne de 2,8 documents par recherche. C'est dire que les étudiants de niveau collégial utilisent peu de documents au cours de leurs travaux.

Analyse des échecs

Nous avons fait l'analyse des échecs en utilisant la méthodologie de Knapp: classement en onze catégories des termes utilisés.

Le tableau 3 donne les résultats et présente le classement des 129 termes non concordants selon ces onze catégories.

La première catégorie est celle où l'étudiant a employé un terme trop général. Sur 129 termes non concordants, 69 sont de ce type.

Tableau 2

**Succès et échecs par secteurs selon l'utilisation
de un ou plusieurs termes**

	G I	GII	Pr I	Pr II	Pr III	Total
Nombre d'étudiants	49	58	28	27	7	169
Concordance	—	—	—	—	—	—
1 terme	13	19	6	12	2	52
Plus. termes	27	25	14	9	2	77
Total	40	44	20	21	4	129
Pourcentage du succès	81,63	75,86	71,42	77,77	57,14	76,33
Aucune concordance	—	—	—	—	—	—
1 terme	3	5	7	2	1	18
Plus. termes	6	9	1	4	2	22
Total	9	14	8	6	3	40
Pourcentage des échecs	18,36	24,13	28,57	22,22	42,85	23,66

Tableau 3

**Classement des catégories de différences
pour les termes non concordants**

Catégories de différences	1° terme	2° terme	3° terme	4° terme	5° terme	6° terme	Total	%
1. Trop général	23	17	16	7	5	1	69	53,48
2. Trop spécifique	3	6	5	4	—	—	18	13,95
3. Phrases	10	2	—	—	—	—	12	9,3
4. Subdivisions non atteintes	2	5	—	—	—	—	7	5,42
5. Lieu plutôt que sujet	—	2	2	1	1	—	6	4,65
6. Forme variante	3	1	1	1	—	—	6	4,65
7. Spécifique sans subdivision	2	1	1	—	—	—	4	3,1
8. Terme populaire	3	1	—	—	—	—	4	3,1
9. Sujet, plutôt que lieu	—	1	1	—	—	—	2	1,55
10. Terme moderne	1	—	—	—	—	—	1	0,77
11. Forme de littérature	—	—	—	—	—	—	—	—
Total	47	36	26	13	6	1	129	100%

Comment expliquer cette fréquence si élevée de termes généraux? D'après nos observations, beaucoup d'étudiants abordaient le catalogue sans avoir défini le sujet de leur recherche et, par conséquent, utilisaient des termes vagues ou trop généraux. Le catalogue leur servait de moyen d'exploration de la discipline dans laquelle ils orientaient leur recherche.

Mentionnons, à titre d'exemple, l'emploi de la vedette ENTREPRISES au lieu de ENTREPRISES, PETITES pour un travail sur les petites et moyennes entreprises, l'utilisation d'un terme désignant la classe au lieu de l'espèce, MAMMIFÈRES au lieu de LOUPS.

L'utilisation de termes trop spécifiques vient en deuxième lieu. Contrairement à la catégorie précédente, il s'agit ici d'utilisation de termes trop restreints. Dix-huit termes entrent dans cette catégorie soit un pourcentage de près de 14%. On peut en donner comme exemple une recherche sur la limnologie, science qui étudie les phénomènes physiques et biologiques se rapportant aux lacs.

Le catalogue ne comportant aucun accès à la limnologie, il aurait fallu utiliser des vedettes plus générales comme LACS ou MARAIS.

Après l'utilisation de termes trop généraux et de termes trop spécifiques, l'emploi de phrases ou d'expressions est la plus importante cause d'échecs. On compte douze cas dans cette catégorie. À titre d'illustration, un étudiant s'intéressant au rôle des policiers a utilisé le terme rôle avant la vedette-matière pertinente POLICIERS.

La quatrième catégorie d'insuccès est celle où l'étudiant utilise un terme général sans parvenir à la subdivision pertinente, comme dans cette recherche sur l'éducation au Québec où la vedette ÉDUCATION est choisie mais sans que la subdivision QUÉBEC (PROVINCE) ne soit atteinte.

La cinquième catégorie est celle où la recherche a été faite au nom de lieu plutôt qu'au sujet avec subdivision géographique. Six cas entrent dans cette catégorie.

On retrouve en sixième lieu, sur un même pied que la cinquième catégorie, l'utilisation de termes ayant une forme différente de la vedette-matière. La plupart de ces échecs sont dus à l'emploi de termes au singulier plutôt qu'au pluriel, un étudiant a employé un sigle (FLQ) au lieu de la vedette FRONT DE LIBÉRATION QUÉBÉCOIS (F.L.Q.).

La septième catégorie se rapporte à l'utilisation d'un terme spécifique au lieu d'un terme général avec subdivision et la huitième se caractérise par l'emploi d'un terme plus populaire que la vedette-matière. La neuvième comprend l'utilisation d'une vedette-matière sujet plutôt que le nom de lieu avec le sujet comme subdivision.

Tableau 4

Taux de rappel et taux de précision

Nombre total des questionnaires	46
Nombre de vedettes-matière utilisées	113
Nombre de titres repérés par les vedettes-matière	585
— pertinents	456
— non pertinents	129
Nombre de titres pertinents trouvés en recherche parallèle	125
Nombre total de documents pertinents	581
Taux de rappel: $\frac{456}{581} \times 100$	78,5%
Taux de précision: $\frac{125}{585} \times 100$	78%

Dans la dixième catégorie, est incluse l'utilisation d'un terme moderne. Un seul cas a été enregistré, celui de l'utilisation du terme *bébés éprouvettes* au lieu de la vedette *INSÉMINATION ARTIFICIELLE HUMAINE*.

Aucun échec n'apparaît dans la onzième catégorie, réservée à l'utilisation de la forme de littérature au lieu d'un sujet exprimé selon cette forme.

Taux de rappel et taux de précision

Des 169 questionnaires de l'enquête, 46 furent retenus aux fins de l'étude du taux de rappel et du taux de précision. Pour ces 46 questionnaires, 113 vedettes-matière furent utilisées par les bibliothécaires. Des 46 recherches, treize ont donné un taux de rappel nul et un taux de précision de 100%, c'est-à-dire que tous les documents identifiés étaient pertinents et qu'aucun autre document pertinent ne fut repéré à l'occasion de la parallèle. Dans les 33 autres recherches, on a enregistré des échecs de rappel et de précision.

Le nombre de titres repérés grâce aux vedettes-matière totalise 585; de ce nombre 456 furent jugés pertinents alors que 129 furent rejetés. La quantité de titres pertinents identifiés par les bibliothécaires, dans le but d'établir le taux de rappel, s'élève à 125. On en arrive donc à un total de 581 titres pertinents.

L'ensemble de ces données nous permettront de calculer le taux de rappel et le taux de précision. Le taux de rappel, mesuré par la quantité de documents pertinents identifiés par les étudiants par rapport au nombre total de documents pertinents, s'établit à 78,5%. Pour sa part, le taux de précision est mesuré par le nombre de documents pertinents qu'ont signalés les étudiants par rapport au nombre total de documents identifiés et par les étudiants, et par les bibliothécaires. On obtient ainsi un taux de 78%.

Facteurs d'échecs

Avant de faire l'analyse du taux de rappel et de précision, il importe d'identifier les divers facteurs responsables des échecs. En la matière, Lancaster en identifie quatre: le langage documentaire, l'indexation, la stratégie de recherche et l'interface usager-système¹².

Les échecs attribuables au langage d'indexation peuvent être partagés en deux groupes: ceux dus au manque de spécificité des termes d'une part et ceux dus aux mauvaises relations entre les termes d'autre part.

Les erreurs imputables à l'indexation même peuvent découler des démarches de l'indexeur, lequel pourrait avoir omis des termes nécessaires, ou utilisé un terme inapproprié, ou encore de la politique gouvernant le nombre de termes à assigner.

Tableau 5

Causes des 125 échecs de rappel

Causes d'échecs	Nombre de titres non repérés	%	Nombre de recherches	Pourcentage des 27 recherches
Langage documentaire				
— manque de spécificité des termes	9	7,2	3	11,11
Total	9	7,2	3	11,11
Indexation				
— termes nécessaires omis	9	7,2	5	18,51
— manque d'exhaustivité	23	18,4	13	48,14
Total	32	25,6	18	66,66
Stratégie de recherche				
— termes pertinents non retenus	72	57,6	9	33,33
— recherche réduite	3	2,4	1	3,7
— variation dans l'exhaustivité	6	4,8	2	7,4
Total	81	64,8	12	44,44
Autres	3	2,4	2	7,4
Total	125	100	35	

12. F.W. Lancaster, "MEDLARS: Report on the Evaluation of its Operating Efficiency", *American Documentation*, vol. 20, no. 2 (April 1969), 123.

Trois genres d'erreurs de nature stratégique peuvent se produire lors de la recherche: le chercheur pourrait ne pas exploiter toutes les approches, en négligeant un terme pertinent ou un aspect de son sujet; il pourrait utiliser des termes inappropriés et, en dernier lieu, il se pourrait qu'il y ait variation dans l'exhaustivité et la spécificité de la recherche.

Un dernier facteur, l'interface usager-système, est responsable d'insuccès lorsque la demande formulée par l'utilisateur ne correspond pas à son besoin réel d'information. Aucun échec n'a, dans notre étude, été attribué à l'interface usager-système puisque les recherches ont été menées par des professionnels.

Analyse des échecs

Nous examinerons maintenant les raisons pour lesquelles 125 documents ont échappé à la vigilance des étudiants et celles qui expliquent l'identification de 129 titres jugés inutiles en utilisant les trois premiers facteurs retenus par Lancaster. On retrouvera aux tableaux 5 et 6 l'ensemble des données.

Le langage documentaire fut responsable de 7,2% des échecs de rappel et de 58,91% des échecs de précision.

Le manque de spécificité des vedettes-matière apparaît comme le facteur dominant avec 7,2% des échecs de rappel et 48,83% des échecs de précision. Ce facteur a été responsable du non-repérage de neuf titres pertinents dans trois recherches distinctes et du repérage de 63 titres non pertinents dans six recherches.

Mentionnons à titre d'exemple une question dont le sujet de recherche était le courant électrique alternatif et continu. Le taux de rappel pour

cette question fut très élevé à cause de l'utilisation de la vedette générale ÉLECTRICITÉ laquelle a permis le repérage de 30 titres pertinents sur une possibilité de 32. Cependant, le taux de précision fut très faible avec 54,5%: seulement 30 titres sur 55 furent jugés pertinents.

Les mauvaises relations entre les termes furent, quant à elles, responsables de seulement 10% des manques de précision.

Les erreurs de repérage dues à l'indexation peuvent être attribuables à l'omission d'un terme nécessaire, à l'emploi de termes inappropriés ou à la variation du niveau d'exhaustivité de l'indexation compte tenu du nombre de vedettes-matière assignées lors de l'indexation d'un document. Ces trois facteurs sont responsables de 25,6% des échecs de rappel et de 6,2% des échecs de précision.

L'omission de termes essentiels fut responsable de 7,2% des échecs de rappel. Neuf titres pertinents n'ont pas été repérés dans cinq recherches distinctes.

L'utilisation de termes inappropriés fut responsable de cinq erreurs de précision au cours de deux recherches.

La variation de la profondeur de l'indexation fut responsable de 18,4% des erreurs de rappel et de 2,23% des erreurs de précision. De fait, 23 titres pertinents n'ont pu être repérés alors que trois titres non pertinents ont été retenus.

La stratégie de recherche, quant à elle, fut responsable de 64,8% des échecs de rappel et de 34,88% des échecs de précision. La stratégie fut la cause d'oubli de 81 titres sur 125 et du repérage de 45 titres inadéquats sur un total de 129. C'est de loin le plus important facteur d'échecs de rappel.

Tableau 6

Causes des 129 échecs de précision

Causes d'échecs	Nombre de titres	%	Nombre de recherches	Pourcentage des 21 recherches
Langage documentaire				
— manque de spécificité des termes	63	48,83	6	28,57
— mauvaises relations	13	10,07	5	23,80
Total	<u>76</u>	<u>58,91</u>	<u>11</u>	<u>52,37</u>
Indexation				
— termes inappropriés	5	3,87	2	9,52
— manque d'exhaustivité	3	2,32	3	14,23
Total	<u>8</u>	<u>6,2</u>	<u>5</u>	<u>23,8</u>
Stratégie de recherche				
— termes inappropriés	16	12,4	2	9,52
— variation dans l'exhaustivité	29	22,48	12	57,14
Total	<u>45</u>	<u>34,88</u>	<u>14</u>	<u>66,66</u>
Total	129	100	30	

L'omission de termes pertinents fut le facteur dominant avec 57,6% des échecs de rappel: 72 titres pertinents ne furent pas repérés au cours de neuf recherches. Mentionnons à titre d'exemple la recherche sur le FLQ et l'omission de la vedette FRONT DE LIBÉRATION DU QUÉBEC (F.L.Q.). Ainsi, cinq titres pertinents sont passés inaperçus.

Les erreurs dues au fait que l'utilisateur n'a pas exploré tous les aspects de son sujet sont beaucoup moins nombreuses. En fait, un seul cas se retrouve dans cette catégorie. Il s'agit d'une recherche sur la violence sexuelle où un aspect important nous a semblé être négligé, celui des crimes et délits sexuels. Trois documents auraient pu être identifiés par le biais de la vedette CRIMES SEXUELS. Il faut avouer cependant que la distinction entre le fait qu'un utilisateur ait omis un terme pertinent ou ait négligé un aspect de son sujet n'est pas toujours clairement identifiable dans la réalité.

L'utilisation de termes inappropriés fut responsable de 16 échecs de précision soit 12% de l'ensemble.

Enfin, la variation de l'exhaustivité de la formulation de la recherche fut responsable de peu d'échecs de rappel mais de plusieurs échecs de précision. Vingt-neuf titres identifiés furent jugés non pertinents et, donc, responsables de 22,48% des échecs de précision.

Conclusion et recommandations

Notre recherche nous amène aux conclusions suivantes: les étudiants qui ont participé à notre enquête obtinrent un taux de succès élevé dans

leur recherche au catalogue sujets (taux de concordance de 76,33% avec la terminologie des vedettes-matière). L'analyse des échecs révèle que leur principale difficulté est l'atteinte du niveau de spécificité requis: ils utilisent, deux fois sur trois, un terme trop général ou un terme trop spécifique.

L'amélioration de l'efficacité des étudiants nous apparaît devoir se faire autant par une meilleure initiation à la méthodologie du travail intellectuel que par la connaissance des instruments de bibliothèques. Les étudiants qui avaient suffisamment précisé le sujet de leur travail avaient toutes les chances de réussir leur recherche au catalogue.

Par ailleurs, le rendement des vedettes-matière françaises de l'Université Laval se situe également à un niveau très élevé (taux de rappel et taux de précision de 78%). Ce résultat a été obtenu, répétons-le, dans un cadre méthodologique bien particulier. Le jugement de la pertinence des documents a été posé uniquement à partir des données bibliographiques de la fiche de catalogue et certaines catégories de documents, comme les ouvrages de référence, ont été écartées. Il n'est pas certain que les résultats auraient été aussi élevés si nous avions pu remédier à ces lacunes méthodologiques que, dans les circonstances, nous avons dû accepter.

L'analyse des échecs fait ressortir que le langage documentaire et la stratégie de recherche sont les principales causes d'échecs. Une meilleure connaissance des vedettes-matière par le personnel de la référence devient essentielle si l'on veut suppléer aux lacunes du langage lui-même.