Document generated on 07/03/2025 2:08 a.m.

Documentation et bibliothèques

Multilinguisme et langages documentaires : le projet MACS en contexte européen

Multilingualism and Index Languages: The MACS Project in Europe

Multilinguismo y lenguajes documentales: el proyecto macs en el contexto europeo

Patrice Landry

Volume 52, Number 2, April–June 2006

Les langages documentaires

URI: https://id.erudit.org/iderudit/1030015ar
DOI: https://doi.org/10.7202/1030015ar

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

Landry, P. (2006). Multilinguisme et langages documentaires : le projet MACS en contexte européen. *Documentation et bibliothèques*, *52*(2), 121–129. https://doi.org/10.7202/1030015ar

Article abstract

Providing subject access using several languages in catalogues has become an important challenge for many national libraries in Europe. Given that the Web enables anyone to search any catalogue on-line, subject access using a single language becomes, for many national libraries in Europe, an important obstacle to efficient bibliographic retrieval. Aware of this obstacle and wishing to take advantage of the wealth of networks and the availability of several index languages, four national libraries established, in 1998, the MACS project (Multilingual Access to Subjects) with the help of the Conférence des bibliothèques nationales européennes (CENL). This article outlines the strategy used by the MACS project to develop a multilingual subject access using concordances (mapping) to the following subject headings: RAMEAU (Bibliothèque nationale de France), LCSH (British Library) and SWD (Deutsche National-Bibliothek and Bibliothèque nationale suisse). This article explains how the approach was developed and the foundations upon which links were created amongst the different index languages.

Tous droits réservés © Association pour l'avancement des sciences et des techniques de la documentation (ASTED), 2006

This document is protected by copyright law. Use of the services of Érudit (including reproduction) is subject to its terms and conditions, which can be viewed online.

https://apropos.erudit.org/en/users/policy-on-use/

This article is disseminated and preserved by Érudit.

Érudit is a non-profit inter-university consortium of the Université de Montréal, Université Laval, and the Université du Québec à Montréal. Its mission is to promote and disseminate research.

https://www.erudit.org/en/

DOCUMENTATION BIBLIOTHÈQUES

Multilinguisme et langages documentaires : le projet MACS en contexte européen

PATRICE LANDRY
Bibliothèque nationale suisse
Berne, Suisse
patrice.landry@slb.admin.ch

RÉSUMÉ | ABSTRACTS | RESUMEN

Fournir un accès par sujet multilingue à leurs catalogues est devenu pour plusieurs bibliothèques nationales européennes un enjeu important. Avec le développement du Web, qui permet à n'importe qui de faire une recherche dans n'importe quel catalogue en ligne, l'accès thématique assuré par un langage documentaire monolingue dans la plupart des bibliothèques nationales européennes constitue un obstacle important à une recherche documentaire efficace. Conscientes de cette lacune et voulant profiter de la richesse des réseaux et de la disponibilité de plusieurs langages documentaires, quatre bibliothèques nationales ont lancé le projet MACS (Multilingual Access to Subjects) sous l'égide de la CENL (Conférence des bibliothèques nationales européennes) en 1998. Cet article décrit la stratégie du projet MACS pour développer, dans un premier temps, un accès par sujet multilingue grâce à l'établissement de concordances (mapping) entre les vedettes des langages documentaires suivants : RAMEAU (Bibliothèque nationale de France), LCSH (British Library) et SWD (Deutsche National-Bibliothek et Bibliothèque nationale suisse). Nous expliquerons comment cette approche a été développée et sur quelles bases les liens entre les vedettes des différents langages documentaires sont créés.

Multilingualism and Index Languages : The MACS Project in Europe

Providing subject access using several languages in catalogues has become an important challenge for many national libraries in Europe. Given that the Web enables anyone to search any catalogue on-line, subject access using a single language becomes, for many national libraries in Europe, an important obstacle to efficient bibliographic retrieval. Aware of this obstacle and wishing to take advantage of the wealth of networks and the availability of several index languages, four national libraries established, in 1998, the MACS project (Multilingual Access to Subjects) with the help of the Conférence des bibliothèques nationales européennes (CENL). This article outlines the strategy used by the MACS project to develop a multilingual subject access using concordances (mapping) to the following subject headings : RAMEAU (Bibliothèque nationale de France), LCSH (British Library) and SWD (Deutsche National-Bibliothek and Bibliothèque nationale suisse). This article explains how the approach was developed and the foundations upon which links were created amongst the different index languages.

Multilinguismo y lenguajes documentales : el proyecto macs en el contexto europeo

Brindar un acceso por tema, multilingüe, a sus catálogos se ha convertido en una postura importante para algunas de las bibliotecas nacionales europeas. Con el desarrollo de la Web, que permite a cualquiera hacer una búsqueda en cualquier catálogo en línea, el acceso temático asegurado por un lenguaje documental monolingüe en la mayoría de bibliotecas nacionales europeas constituye un obstáculo importante para una investigación documental eficaz. Conscientes de este vacío y con el fin de aprovechar la riqueza de las redes y la disponibilidad de algunos lenguajes documentales, cuatro bibliotecas nacionales han lanzado el proyecto MACS (Multilingual Access to Subjects) con el auspicio de la CENL (Conferencia de Bibliotecas Nacionales Europeas) en 1998. Este artículo describe la estrategia del proyecto MACS para desarrollar, en un primer momento, un acceso por tema multilingüe, gracias al establecimiento de concordancias (mapping) entre los encabezados de los siguientes lenguajes documentales : RAMEAU (Biblioteca Nacional de Francia), LCSH (British Library) y SWD (Deutsche National-Bibliothek y Biblioteca Nacional Suiza). Explicaremos cómo ha sido desarrollado este enfoque y sobre qué bases se crean los vínculos entre los encabezados de los diferentes lenguajes documentales.

L'ACCÈS FACILITÉ AUX CATALOGUES des bibliothèques sur le Web a grandement amélioré les possibilités de recherche de l'information. Que le chercheur accède à un catalogue en ligne ou à un portail donnant accès à un réseau virtuel de plusieurs catalogues de bibliothèques, il est quasiment assuré de trouver relativement facilement les informations qu'il cherche, à condition d'effectuer une recherche précise ; une recherche sur le portail du *Karlsruher Virtueller Katalog* (KVK, <http://www.ubka.uni-karlsruhe.de/kvk/kvk/kvk_fr.html>) sur l'écrivain québécois Michel Tremblay, par exemple, permet en quelques secondes de localiser des documents dans plusieurs centaines de bibliothèques. Par contre, le même lecteur (ou usager) qui effectue une recherche thématique dans le KVK, ou dans tous les autres réseaux de catalogues en ligne, est confronté à un environnement hétérogène dans lequel il devra faire face à plusieurs langages d'indexation et à différentes langues naturelles. Devant une telle situation, il n'a que peu de possibilités : se référer à ses connaissances de base de quelques langues naturelles ou à des dictionnaires bilingues. De plus, il devra effectuer plusieurs recherches distinctes dans les catalogues auxquels il veut accéder.

Depuis une quinzaine d'années, soit depuis les débuts de l'expansion du Web, de grands travaux ont été entrepris pour permettre aux lecteurs d'effectuer avec succès des recherches thématiques simultanées dans plusieurs bases de données ou catalogues par un guichet unique, ou selon l'approche du *one-stop*

On définit l'interopérabilité comme la capacité de plusieurs systèmes à communiquer entre eux sans ambiguïté et à échanger de l'information sans difficulté.

searching. L'objectif est de permettre à l'utilisateur d'accéder à des ressources provenant de plusieurs systèmes hétérogènes par le moyen d'une interrogation dans sa propre langue. Différentes approches ont été proposées pour assurer l'interopérabilité de différents langages documentaires de type vocabulaires contrôlés (thésaurus, listes de vedettes-matières), schémas de classification et mots-clés. La problématique que nous abordons dans cet article s'inscrit dans le contexte de ces travaux et nous mettrons l'accent sur le projet MACS (*Multilingual Access to Subjects*[1]) comme solution, au moins partielle, au problème de l'accès thématique multilingue.

Le projet MACS a été lancé sous l'égide de la Conférence des bibliothèques nationales européennes (CENL[2]) par quatre bibliothèques nationales: la Bibliothèque nationale suisse (BN), qui pilote actuellement le projet, la Bibliothèque nationale de France (BNF), la *Deutsche National-Bibliothek* (DNB)[3] et la *British Library* (BL). Le projet avait pour objectif de trouver une solution au problème de l'accès par sujet aux catalogues en ligne dans un environnement multilingue. Cet article présente le contexte dans lequel cette tentative de trouver une solution a été entreprise et décrit les facteurs qui ont mené à l'élaboration des principes de base et de l'approche méthodologique du projet MACS.

L'interopérabilité

Afin de bien comprendre pourquoi quatre bibliothèques nationales européennes ont décidé, en 1997, de concerter leurs efforts pour la recherche d'une solution multilingue, il nous semble important d'expliquer la notion d'interopérabilité et son rapport avec la problématique de l'accès par sujet multilingue. De façon générale, on définit l'interopérabilité comme la capacité de plusieurs systèmes à communiquer entre eux sans ambiguïté et à échanger de l'information sans difficulté. Dans le contexte de systèmes d'information bibliographiques, l'interopérabilité passe nécessairement par la mise en compatibilité de plusieurs langages documentaires utilisés pour fournir un accès aux collections.

M.L. Zeng et L.M. Chan (2004) dressent une typologie des différentes initiatives d'interopérabilité menées depuis une quarantaine d'années. Elles font tout d'abord la distinction entre les travaux accomplis dans deux périodes différentes: celle du développement des bases de données en ligne, principalement entre 1960 et 1990, et celle de l'Internet, de 1990 à aujourd'hui. Durant la première période, plusieurs travaux ont été faits pour rendre compatibles différents thésaurus par l'établissement de concordances simples entre les termes. Les auteures font notamment référence aux travaux ayant mené au développement de l'UMLS (*Unified Medical Language System*[4]), du *ROOT Thesaurus* et de l'*Integrated Multilingual Thesaurus* (UNESCO), en les présentant comme des exemples réussis d'intégration de plusieurs thésaurus, rendus possibles par le développement des systèmes de gestion de bibliothèques et du format de communication MARC. Aux États-Unis, plusieurs bibliothèques spécialisées ont développé un accès multilingue au moyen des *Library of Congress Subject Headings* (LCSH) et du *Medical Subject Headings* (MeSH), alors qu'au Canada l'accès bilingue aux collections est assuré par les *Canadian Subject Headings* (LCSH/CSH) et par le Répertoire de vedettes-matière de l'Université Laval (RVM).

La deuxième période a été fortement marquée par l'impact de l'Internet. Le développement du Web a grandement favorisé la mise en réseau d'un énorme volume de ressources documentaires et le développement de ressources numériques. Plusieurs initiatives, développées pour faciliter la recherche thématique dans les bases de données et les catalogues de bibliothèques, ont profité des progrès technologiques pour mieux répondre à de nouveaux besoins d'interopérabilité. La création de nouveaux environnements Web et de réseaux virtuels a accentué non seulement la diversité des ressources mises en commun, mais également celle des langages d'indexation utilisés pour décrire et accéder au contenu de ces ressources.

M. Doerr (2001) ainsi que M.L. Zeng et L.M. Chan (2004) dressent une liste de projets et d'initiatives assez variés ayant en commun une volonté d'atteindre l'interopérabilité fonctionnelle de divers types de langages documentaires, soit par l'établissement de concordances, par un mécanisme connu sous le nom de commutation (*switching*), ou par l'établissement d'une passerelle entre les différents langages. Parmi les projets qui touchent à l'interopérabilité de différents types de langages documentaires (schémas de classification, thésaurus et listes de vedettes-matière), on cite souvent en exemples les projets WebDewey[5],

1. Voir dans les sources consultées différents articles publiés sur le projet MACS. On peut consulter également le site MACS à: <https://macs.cenl.org>.
2. <http://www.cenl.org> (consulté sur Internet le 9 mai 2006).
3. La bibliothèque nationale allemande a changé de nom au début du mois d'avril 2006, passant de *Die Deutsche Bibliothek* à *Deutsche National-Bibliothek*.
4. <http://www.nlm.nih.gov/research/umls/> (consulté sur Internet le 9 mai 2006).
5. <http://www.oclc.org/dewey/versions/webdewey/> (consulté sur Internet le 9 mai 2006)

Carmen[6], UMLS, *High Level Thesaurus Project* (HILT[7]) et *UC Berkeley DARPA Unfamiliar Metadata Project*[8]. D'autres projets visent plutôt à établir des concordances entre langages documentaires de même type, par exemple, MACS (listes de vedettes-matières)[9], Merimee (thésaurus)[10], Renardus[11] et SAB/DDC[12] (schémas de classification).

Les méthodes privilégiées, au cours des 40 dernières années, pour améliorer et développer l'interopérabilité thématique continuent d'être utilisées et plusieurs de celles-ci sont à la base d'outils documentaires très développés. La traduction ou l'adaptation d'un langage documentaire existant, la dérivation d'un langage vers un autre (par exemple, le projet FAST[13]) et l'intégration de plusieurs listes de vedettes-matière ou de thésaurus en une seule liste ou un seul thésaurus (par exemple, TermSciences[14]) sont des méthodes éprouvées qui assurent un certain niveau d'interopérabilité thématique. D'autres méthodes développées depuis une dizaine d'années utilisent des applications informatiques pour établir des liens entre plusieurs types de langages documentaires, soit par traduction automatique (par exemple, Dandelon[15]), par utilisation d'un langage intermédiaire qui permet de naviguer parmi des termes équivalents dans plusieurs langues (par exemple, Renardus) ou par une méthode de cartographie (*mapping*) qui établit des concordances entre des termes équivalents dans différents vocabulaires contrôlés (par exemple, entre LCSH et la Classification décimale Dewey[16], entre le *Laborline thesaurus* et LCSH [voir Chaplan, 1995], entre le *Art and Architecture Thesaurus* [AAT] et divers thésaurus et listes de vedettes-matière couvrant le même domaine).

Contexte et travaux préliminaires

Le choix d'une méthodologie pour assurer l'interopérabilité thématique dans un contexte multilingue dépend de plusieurs facteurs. On note parmi ceux-ci la nature même des langages documentaires et les possibilités d'établir des concordances entre eux, l'environnement documentaire et le potentiel de coopération entre plusieurs partenaires de langues différentes, la situation technique et les moyens informatiques disponibles, les sources de financement public, les intérêts politiques et, enfin, les perspectives de viabilité de la mise sur pied d'un projet à long terme, qui pourrait aboutir à la création et au développement d'une solution.

En 1997, les représentants de sept bibliothèques nationales européennes, qui s'étaient réunis pour discuter de la problématique de l'accès par sujet, pouvaient compter sur une situation quasi idéale pour élaborer une solution au problème de l'accès thématique multilingue en Europe. Un groupe de travail avait en effet été mis sur pied dans le cadre du programme CoBRA+, partiellement financé par la Commission européenne, pour discuter d'accès multilingue dans le contexte des bibliothèques nationales. Ce groupe de travail pouvait compter sur le support institutionnel de chacune des bibliothèques nationales participantes dans sa recherche d'une solution à la problématique du multilinguisme. Mais très vite, le groupe a été confronté à une question délicate : quelle méthodologie fallait-il choisir afin d'assurer l'interopérabilité la plus large possible entre les catalogues des bibliothèques nationales en Europe, tout en respectant l'intégrité et le cheminement propre de chaque langage documentaire déjà utilisé ? Pour bien comprendre la nature de cette problématique et la solution proposée, il faut se rapporter à la situation en Europe durant les années 1980 et 1990. Au cours de ces deux décennies, il y a eu certes quelques efforts faits pour trouver une solution aux problèmes posés par le multilinguisme dans l'accès à l'information, mais les véritables déclencheurs furent le développement technologique et les nouvelles possibilités de coopération en Europe sous l'égide de la Commission européenne. Au cours de la même période, le développement de nouveaux langages documentaires et en particulier des listes de vedettes-matière a également ouvert de nouvelles pistes de recherche et de développement.

Au début des années 1990, les catalogues des bibliothèques d'Europe n'étaient généralement accessibles que dans une seule langue naturelle. À l'époque, il existait pourtant deux approches pour l'accès multilingue, développées dans les années 1980 en Belgique et en Suisse, pays où se posait depuis longtemps le problème de l'accès multilingue. La Bibliothèque Royale de Bruxelles (BRB) avait choisi d'indexer ses documents en utilisant le langage LCSH et de traduire les vedettes-matière en flamand et en français. Les vedettes multilingues étaient codées dans les notices bibliographiques et chaque vedette était saisie dans une notice d'autorité. En Suisse, ce sont les bibliothèques de l'École polytechnique fédérale de Lausanne (EPFL) et l'*Eidgenössische Technische Hochschule* (ETH) de Zurich qui ont collaboré pour créer un accès multilingue, utilisant la Classification décimale universelle (CDU) comme base de leur système d'accès trilingue. Chaque indice de classification avait une vedette

6. <http://www.mathematik.uni-osnabrueck.de/projects/carmen/index.en.shtml> (consulté sur Internet le 9 mai 2006)
7. <http://hilt.cdlr.strath.ac.uk/> (consulté sur Internet le 9 mai 2006)
8. <http://metadata.sims.berkeley.edu/GrantSupported/unfamiliar.html> (consulté sur Internet le 9 mai 2006)
9. <https://ilmacs.uvt.nl/pub/> (consulté sur Internet le 9 mai 2006)
10. <http://www.culture.gouv.fr/culture/inventai/presenta/bddinv.htm> (consulté sur Internet le 11 mai 2006)
11. <http://www.renardus.org/about_us/project_archive.html> (consulté sur Internet le 9 mai 2006)
12. <http://www.kb.se/ku/del_3_underlag.htm> (consulté sur Internet le 9 mai 2006)
13. <http://www.oclc.org/research/projects/fast/default.htm> (consulté sur Internet le 9 mai 2006)
14. <http://termsciences.inist.fr/> (consulté sur Internet le 9 mai 2006)
15. <http://www.dandelon.com> (consulté sur Internet le 9 mai 2006)
16. <http://www.oclc.org/dewey/updates/numbers/> (consulté sur Internet le 9 mai 2006)

correspondante qui, à l'origine, n'était qu'en allemand, et c'est en 1982 que les bibliothèques de Zurich et de Lausanne ont lancé un projet de traduction des termes allemands vers l'anglais et le français. Il est important de noter que les approches développées à Bruxelles et à Zurich n'ont pas été adoptées par d'autres pays européens à la recherche de moyens d'accès multilingue. Cela s'explique en partie par le fait que chaque bibliothèque, ou réseau de bibliothèques, avait déjà constitué des bases de données bibliographiques considérables à l'aide d'un système d'indexation donné, et qu'il était pratiquement impossible de l'abandonner pour adopter l'un ou l'autre des systèmes développés en Belgique ou en Suisse.

Les progrès technologiques et l'automatisation des catalogues ont permis de reconsidérer l'accès multilingue dans un nouveau contexte. Au cours de la conférence *Library Automation and Networking*, tenue à Bruxelles en 1990, plusieurs présentations portaient sur les infrastructures des réseaux créés dans les années 1980 et sur leur rôle dans la redéfinition des potentialités de coopération. C'est à partir de ce moment que l'on a pris conscience du fait que la technologie pouvait augmenter et élargir les possibilités de coopération qui s'étaient surtout développées jusque-là au sein des services de catalogage et de prêts inter-bibliothèques. Un des buts de la conférence était de promouvoir la coopération internationale et de favoriser la diffusion de l'information en Europe. P. Goossens (1991) et G. Clavel-Merrin (1991) ont décrit les expériences suisse et belge dans le domaine de l'accès par sujet multilingue et ont décrit l'étendue du travail à accomplir pour arriver à trouver une solution européenne. L'une des principales tâches consistait à définir une politique concernant les bases de données et le développement de réseaux de bibliothèques, qui permettrait de réduire au minimum le nombre de systèmes utilisés (et en particulier le nombre de langages d'indexation) et de faire en sorte que l'accès thématique multilingue devienne un objectif national.

Les programmes de recherche et de développement mis en place par la Commission européenne, au début des années 1990, ont été à la source du développement d'initiatives de coopération qui ont abouti à des résultats concrets. Ces initiatives ont également contribué à encourager les bibliothèques à s'unir autour de projets d'intérêts communs. Un de ces programmes, le programme de télématique pour les bibliothèques du troisième et quatrième programme cadre de la Commission européenne (1990-1998), a encouragé plusieurs bibliothèques nationales à lancer des initiatives de coopération sur le partage des ressources bibliographiques. La *Deutsche National-Bibliothek* proposa notamment, en 1995, la création du projet MULIS, qui avait pour but de créer un thésaurus multilingue regroupant cinq fichiers d'autorités et systèmes de classification. Quoique ce projet n'ait pas été retenu par la Commission européenne, l'initiative a eu le mérite de placer la problématique de l'interopérabilité des catalogues en ligne sur la liste des priorités de plusieurs bibliothèques nationales.

La proposition de la *Deutsche National-Bibliothek* avait également mis en évidence l'existence de langages documentaires européens très développés. À ce moment, la majorité des pays européens utilisaient et maintenaient des langages documentaires, thésaurus ou listes de vedettes-matière conformes aux normes bibliothéconomiques nationales et internationales. Plusieurs de ces langages avaient été développés au début des années 1980 et avaient généré, sur une période de 15 ans, des fichiers d'autorités de bonne qualité qui contenaient des vedettes couvrant tous les domaines du savoir. Les bibliothèques nationales représentées dans le groupe de travail utilisaient des listes d'autorités qui étaient généralement utilisées dans les bibliothèques de leur pays respectif: la *Schlagwortnormadatei/Regeln für den Schlagwortkatalog* (SWD/RSWK) développée en Allemagne au début des années 1980 et introduite en 1986 pour l'indexation en langue allemande, le Répertoire d'autorité-matière encyclopédique et alphabétique (RAMEAU) créé en 1980 à partir du RVM de l'Université Laval et utilisé dans les bibliothèque françaises et dans quelques pays francophones, et la LCSH, utilisée en Grande-Bretagne et adoptée par la *British Library* en 1995.

La méthodologie MACS

C'est dans ce contexte que la Bibliothèque nationale suisse, la Bibliothèque nationale de France, la *British Library* et la *Deutsche National-Bibliothek* ont accepté de relever le défi de définir un concept d'accès multilingue et d'entreprendre une étude pour déterminer s'il était faisable d'établir des liens entre différentes vedettes-matières. Pour ces bibliothèques, les conditions semblaient réunies pour évaluer la pertinence de rendre interopérables les langages d'indexation utilisés. La décision de s'engager dans une étude de faisabilité reposait sur les constats suivants:

- ▷ chaque langage d'indexation était consigné dans un fichier d'autorités important qui respectait les normes bibliographiques en vigueur dans chaque langue;
- ▷ chacun de ces langages continuait à être développé par la bibliothèque nationale concernée;
- ▷ plusieurs millions de documents avaient été indexés à l'aide de ces langages documentaires;
- ▷ dans chacune des bibliothèques, les utilisateurs connaissaient bien ces langages d'indexation.

Le projet MACS a donc été créé avec l'objectif de fournir des moyens d'accès multilingue aux bases de données des bibliothèques en utilisant des termes

FIGURE 1 : Interface de recherche

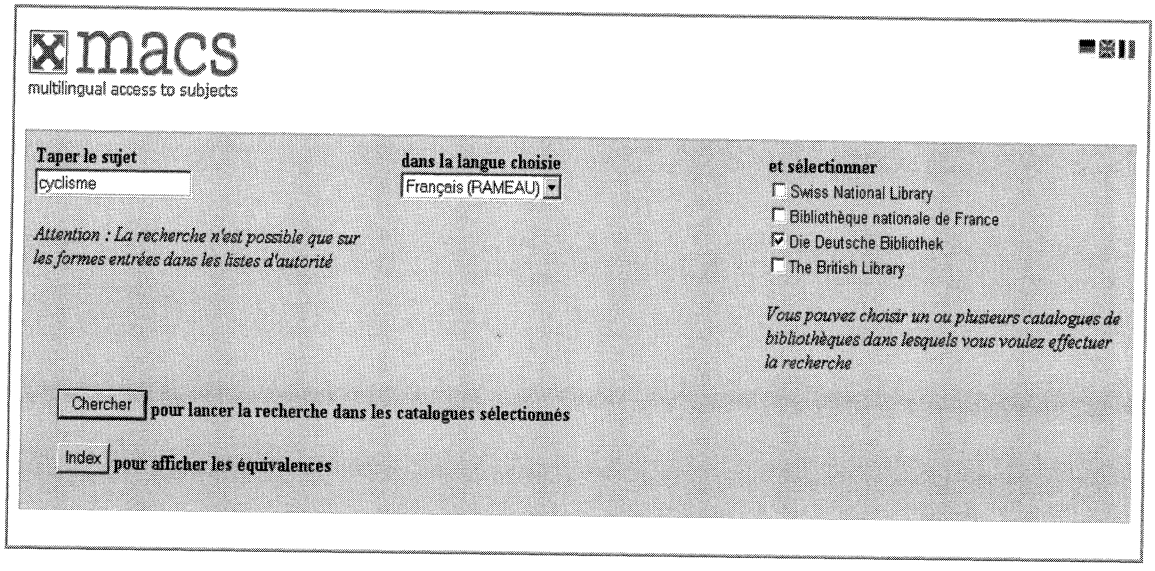
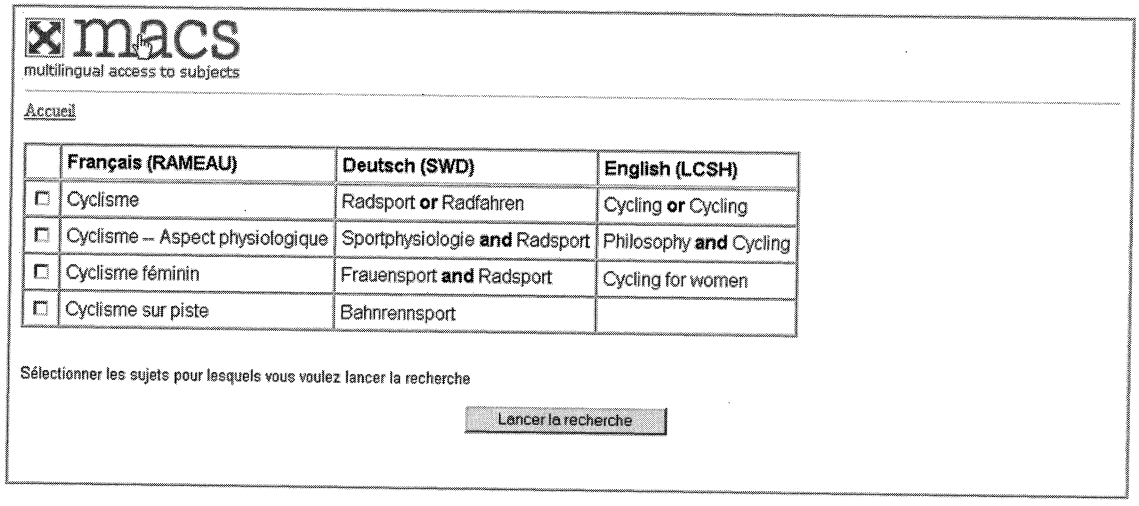


FIGURE 2 : Recherche dans l'index des équivalences



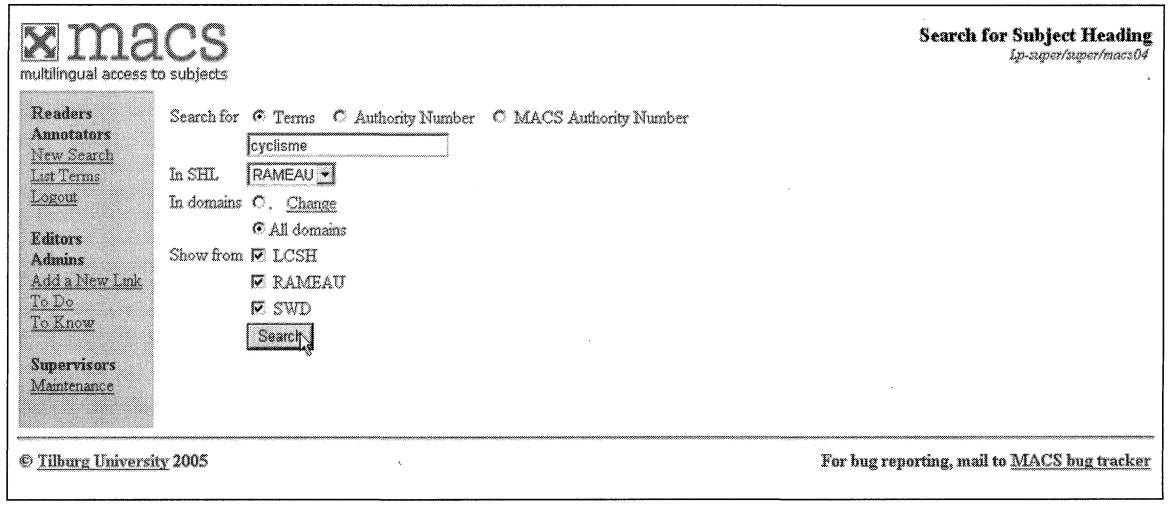
équivalents puisés dans les différents catalogues matières. L'approche qui serait privilégiée devait permettre un accès très large aux diverses bases de données par l'intermédiaire des langages d'indexation utilisés le plus fréquemment dans les bibliothèques d'Europe.

Sur le plan pratique, le projet MACS avait pour but de permettre à un usager de repérer tous les documents pertinents indexés par les bibliothèques associées au projet à partir d'une seule recherche faite dans la langue de son choix : le français, l'anglais ou l'allemand. Cette recherche multilingue deviendrait possible grâce aux équivalences établies entre les trois langages d'indexation utilisés par les quatre bibliothèques nationales : SWD/RSWK pour l'allemand, RAMEAU pour le français et LCSH pour l'anglais.

L'étude de faisabilité, publiée en 1999, a d'abord permis de définir quelques principes de base pour le développement d'un système de correspondances entre vedettes. La méthode de concordance ou de *mapping* était la seule capable de respecter les principes fondamentaux du projet. L'approche MACS repose avant tout sur le principe d'égalité des langues et des langages d'indexation. Il ne peut y avoir de hiérarchie entre les langues et c'est pour cette raison que la méthodologie qui utilise une langue pivot ou une langue intermédiaire (utilisée par Renardus, par exemple) n'a pas été privilégiée. De plus, chaque langage d'indexation est géré localement dans les fichiers d'autorités nationaux et est donc autonome ; seule la tête de vedette est utilisée dans le système de concordances.

Le projet s'appuie sur l'hypothèse que le chercheur aura préalablement recherché le terme accepté dans le fichier d'autorité du langage d'indexation qu'il utilisera pour faire sa recherche. L'approche MACS privilégie l'établissement de véritables équivalences entre les langages d'indexation, plutôt que la traduction automatique de requêtes ou de vedettes. Les correspondances sont conçues comme des grappes concep-

FIGURE 3 : Interface de la base de gestion de liens MACS



tuelles. Celles-ci sont considérées comme des ensembles individuels et non pas comme une sorte de méga-thésaurus multilingue. Enfin, la validation des liens est toujours basée sur le résultat de recherches effectuées dans les catalogues des bibliothèques. En résumé, MACS a développé une approche reposant sur l'égalité et l'autonomie des langues, sur l'établissement d'équivalences fixes considérées comme différentes grappes conceptuelles et validées par la cohérence des résultats.

L'étude de faisabilité a également permis de valider la méthodologie qui permet d'établir la concordance entre vedettes. La méthode pour relier les vedettes entre elles a été développée avec précision en utilisant les vedettes dans deux domaines : « sports athlétiques » et « théâtre ». Pour chaque domaine, trois listes monolingues de vedettes ont été établies ; ces trois listes ont ensuite été liées aux niveaux terminologique (vedette), sémantique (notice) et syntaxique (application). La méthode a été utilisée pour établir des équivalences parfaites, partielles ou complexes[17] et dans la majorité des cas, des équivalences trilingues ont pu être établies.

Les résultats de l'étude de faisabilité ayant reçu un accueil positif, le projet MACS a été établi formellement comme un projet de la CENL par les quatre bibliothèques nationales, dans le but de faire réaliser une application pour la gestion des liens et pour la recherche à partir des liens. Ces travaux, effectués de 1999 à 2001, ont mené à la création d'un prototype basé sur une application Web, qui permettait aux bibliothèques d'alimenter la base de liens d'une façon virtuelle et décentralisée, et aux utilisateurs, d'effectuer des recherches. Pour poursuivre l'analyse et raffiner l'approche méthodologique, l'équipe MACS a saisi dans cette base environ 1 000 liens trilingues et chaque bibliothèque partenaire a chargé 15 000 notices bibliographiques dans lesquelles étaient utilisées les vedettes ainsi liées. Il est alors devenu possible de valider les liens en utilisant les résultats de la recherche dans le prototype. Un accès public au prototype a été créé en 2001 et environ 500 personnes ont consulté cette base et effectué des recherches[18].

L'interface de recherche, développée pour le prototype, permet de mieux comprendre comment les liens créés assurent une recherche multilingue au niveau des têtes de vedettes. Cette interface, qui utilise le profil Zthes Z39.50, a été testée avec succès dans deux des catalogues des bibliothèques participantes, mais devra être améliorée avant d'être installée dans les OPACs des bibliothèques. La figure 1 illustre cette interface, en utilisant l'exemple d'un chercheur québécois qui désire faire une recherche sur le terme « cyclisme » dans le catalogue de la *Deutsche National-Bibliothek*.

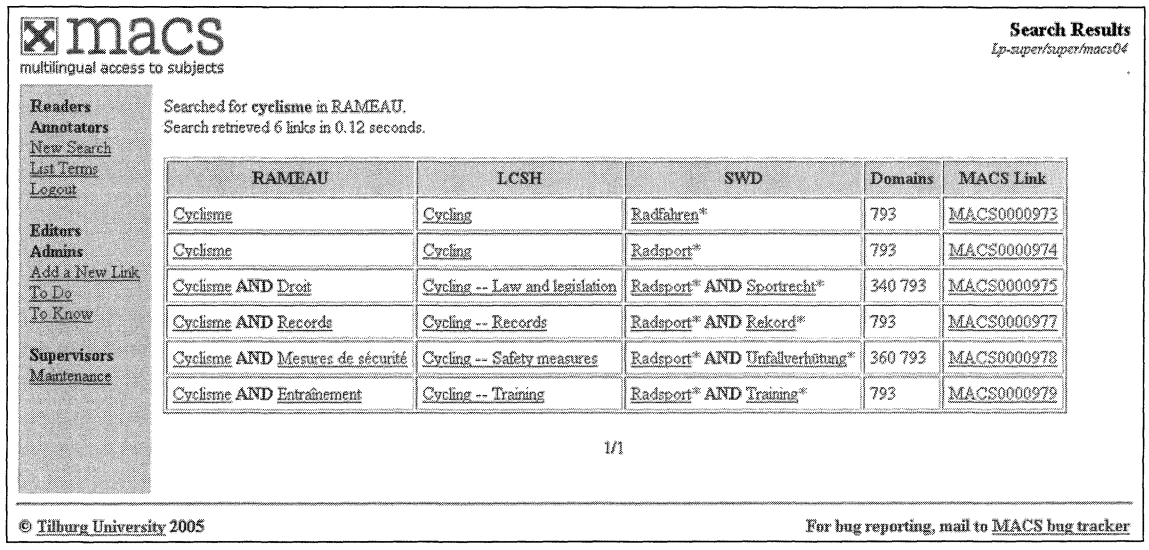
L'interface permet d'effectuer une recherche directe dans les catalogues sélectionnés ou d'utiliser un index donnant accès aux équivalences. La figure 2 illustre les équivalences de « cyclisme » repérées dans l'index.

Une base opérationnelle de gestion des liens MACS a été établie dès 2002. Le développement de cette nouvelle application tenait compte des réflexions faites sur l'alimentation de la base de liens par les différents partenaires. Une approche pragmatique a été élaborée, fondée sur une répartition efficace des tâches sur une base linguistique, et permettant à chacun de commenter la pertinence des liens. Un modèle de gestion du travail de type fédératif a été élaboré pour préciser le partage des responsabilités

17. Voir l'exemple de la vedette « cyclisme » dans la figure 4, qui illustre les équivalences complexes (1:X) pour les liens entre une vedette d'une langue liée à deux vedettes plus spécifiques dans une autre langue, et les équivalences (1:1+1) dans lesquelles deux vedettes doivent être combinées (« AND ») pour représenter le même concept exprimé par une seule vedette dans une autre langue.

18. Le prototype est accessible à partir du site MACS à : <https://macs.cenl.org>.

FIGURE 4 : Résultats de recherche (*New Search*) de liens contenant le terme « cyclisme »



macs
multilingual access to subjects

Search Results
Lp-super/super/macs04

Readers
Annotators
New Search
List Terms
Logout

Editors
Admins
Add a New Link
To Do
To Know

Supervisors
Maintenance

Searched for **cyclisme** in RAMEAU.
Search retrieved 6 links in 0.12 seconds.

RAMEAU	LCSH	SWD	Domains	MACS Link
Cyclisme	Cycling	Radfahren*	793	MACS0000973
Cyclisme	Cycling	Radsport*	793	MACS0000974
Cyclisme **AND** Droit	Cycling -- Law and legislation	Radsport* **AND** Sportrecht*	340 793	MACS0000975
Cyclisme **AND** Records	Cycling -- Records	Radsport* **AND** Rekord*	793	MACS0000977
Cyclisme **AND** Mesures de sécurité	Cycling -- Safety measures	Radsport* **AND** Unfallverhütung*	360 793	MACS0000978
Cyclisme **AND** Entraînement	Cycling -- Training	Radsport* **AND** Training*	793	MACS0000979

1/1

© Tilburg University 2005 — For bug reporting, mail to MACS bug tracker

entre les partenaires, tout en respectant les principes mêmes du projet :

- ▷ autonomie de chaque langage d'indexation ;
- ▷ responsabilité limitée dans l'élaboration du lien, chaque partenaire étant seul autorisé à ajouter une vedette appartenant à son propre langage d'indexation ;
- ▷ possibilités de propositions pour les autres parties du lien ;
- ▷ possibilités de commenter toutes les équivalences à travers une fonction « commentaires ».

La base de gestion de liens permet à chaque partenaire du projet de travailler de façon autonome et selon les possibilités de ses ressources humaines. Chaque bibliothèque partenaire est libre de traduire les interfaces de la base (présentement en anglais seulement) et de gérer les droits et permissions qu'elle accorde aux collaborateurs qui travaillent dans le projet.

Les figures 3 et 4 illustrent de façon concrète la base de gestion des liens. Chaque terme et chaque lien peuvent être cherchés par les fonctions « *New Search* » ou par « *List Terms* ».

Chaque lien est géré par un numéro de lien (*MACS Link*) et possède un ou deux indices de domaines qui permettent de repérer les liens par grandes classes thématiques. À partir de cet écran, un indexeur peut créer un nouveau lien (*Add a New Link*) ou utiliser les fonctions « *To Do* » ou « *To Know* » pour voir les tâches qu'il doit effectuer dans ses domaines de travail.

Malgré la mise en place de cette nouvelle base de gestion des liens MACS favorisant l'efficacité du travail, les partenaires ont vite compris qu'il serait trop long et très onéreux de constituer un nombre suffisant de liens pouvant assurer un accès multilingue minimal dans la majorité des catalogues si l'on n'utilisait qu'une méthode purement intellectuelle. Des solutions pragmatiques ont été proposées qui, sans pour autant remettre en question les principes de base MACS, ont permis d'accélérer l'alimentation de la base de liens. Tout d'abord, il y avait la possibilité de réutiliser les équivalences RAMEAU-LCSH enregistrées dans les notices d'autorités RAMEAU. Dans un premier temps, 30 000 de ces équivalences sur les 70 000 présentes dans le ficher d'autorités ont été extraites et chargées dans la base de liens MACS. Ensuite, la Bibliothèque nationale suisse a entrepris une expérience avec l'École de traduction et d'interprétation de l'Université de Genève pour évaluer une méthode de proposition de liens au niveau terminologique. Ces liens devaient être validés par les indexeurs de la Bibliothèque nationale suisse. Quoique que cette méthode ait donné des résultats plutôt mitigés, elle sera reprise sous une nouvelle forme en Allemagne, à l'Université de Cologne plus précisément, pour appuyer les efforts de la *Deutsche National-Bibliothek* dans ses travaux.

État des travaux

Les travaux entrepris pour la mise en place du système de liens MACS se sont terminés à la fin de 2004 et ont été officiellement validés par les partenaires MACS en janvier 2005. Se posait alors la question des suites à apporter à l'expérience. MACS était confronté, comme beaucoup de projets impliquant plusieurs partenaires, à la difficulté de passer du stade de projet à la phase d'exploitation. Le projet avait atteint ses principaux objectifs de mettre en place les moyens qui permettaient à un utilisateur d'effectuer des recherches thématiques multilingues dans la langue de son choix, et de créer une infrastructure qui assurait la création et la gestion de liens dans plusieurs langues selon une organisation de travail décentra-

lisée et flexible. Ne restait à effectuer que la mise en opération.

À l'automne 2005, la Bibliothèque nationale suisse a pris l'initiative de proposer une nouvelle structure de projet MACS dans le but de faire passer progressivement les tâches dans l'opérationnel. Ce plan de projet comportait trois objectifs: 1) intégrer les tâches de création de liens dans les processus de travail de chaque institution concernée, 2) développer une interface de recherche qui permettrait d'utiliser les liens pour la recherche, et 3) mettre en place une structure administrative permanente pour coordonner les travaux d'établissement des liens. Planifiée sur une période de deux ans, la réalisation des tâches proposées doit permettre aux bibliothèques d'évaluer l'investissement nécessaire pour rendre MACS opérationnel. Sur la plan de la production des liens, l'accent est mis sur l'ajout des vedettes SWD dans les liens contenant les vedettes RAMEAU et LCSH. Cette tâche sera menée conjointement par la Bibliothèque nationale suisse et la *Deutsche National-Bibliothek* avec le support de l'Université de Cologne. La Bibliothèque nationale suisse entreprendra une étude avec la Bibliothèque nationale de l'Italie (Florence) pour étudier la faisabilité de l'intégration du nouveau langage d'indexation italien. Pour l'interface de recherche, le projet doit évaluer la pertinence d'un développement local ou d'un développement qui se ferait dans le cadre du portail de recherche de *The European Library* (TEL[19]).

La poursuite des travaux, notamment avec la *Deutsche National-Bibliothek*, sera l'occasion de mettre en place une méthodologie MACS plus conséquente, qui devrait assurer une plus grande cohérence et une constance dans la création des liens. Le projet entreprendra ces travaux en se référant à deux nouvelles publications normatives qui présentent des recommandations sur l'établissement de liens entre plusieurs langages documentaires. Le document ANSI/NISOZ39.19-2005 *Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies* (2005), notamment le chapitre 10 «*Interoperability*», et le *British Standard BS 8723 Part 4 (DRAFT) Interoperability between vocabularies* (2006) seront les sources principales pour mieux définir la portée de certains principes MACS, en particulier celui de l'égalité des langues, de la notion de langue source et du traitement des vedettes-matière pré-coordonnées.

La publication de ces nouvelles normes ainsi que les travaux récents du groupe de travail IFLA, qui a mis en consultation le document *Guidelines for Multilingual Thesauri* (2005), montrent que l'interopérabilité entre langages documentaires est maintenant une problématique bien ancrée dans le monde bibliothéconomique. Ces normes et ces lignes directrices assurent un fondement théorique aux projets et aux initiatives qui visent à améliorer la recherche thématique monolingue et multilingue dans les réseaux. Elles ont surtout le mérite de cautionner les efforts qui ont été entrepris depuis une quarantaine d'années pour améliorer la recherche par sujet dans les catalogues et les bases de données. ◉

19. <http://www.theeuropeanlibrary.org> (consulté sur Internet le 9 mai 2006).

Sources consultées

ANSI/NISOZ39.19-2005 *Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies.* Bethesda, Maryland: NISO Press.

British Standard BS 8723 Part 4 (DRAFT) Interoperability between vocabularies. 2006 (document en consultation restreinte).

Chaplan, M.A. 1995. Mapping Laborline Thesaurus Terms to Library of Congress Subject Headings Implications for Vocabulary Switching. *Library Quarterly* 65 (1): 39-61.

Clavel-Merrin, G. 1991. Multilingual subject access in Switzerland. In *Library Automation and Networking: New Tools for a New Identity.* Herman Liebaers and Marc Walckiers dir. München, Germany: K.G. Saur, 215-228.

_____. 1999. The Need for co-operation in creating and maintaining multilingual subject authority files. In *65th IFLA Council and General Conference, Bangkok, Thailand, August 20-28 1999,* Meeting 155. <http://www.ifla.org/IV/ifla65/papers/080-155e.htm> (consulté sur Internet le 9 mai 2006).

Cordeiro, M.I. 2003. From library authority control to network authoritative metadata sources. In *Subject retrieval in a networked environment: Papers presented at an IFLA Satellite Meeting sponsored by the IFLA Section on Classification and Indexing & IFLA Section on Information Technology, OCLC, Dublin, Ohio, USA, 14-16 August 2001.* Dublin, OH: OCLC, 131-139.

Doerr, M. 2001. Semantic problems of thesaurus mapping. *Journal of Digital Information* 1 (8) <http://jodi.ecs.soton.ac.uk/Articles/vo1/io8/Doerr/> (consulté sur Internet le 9 mai 2006).

Freyre, E. et M. Naudi. 2002. MACS délie les langues. *Archimag* 157: 28-30.

_____.2003. MACS: Subject access across languages and networks. In *Subject retrieval in a networked environment: Papers presented at an IFLA Satellite Meeting sponsored by the IFLA Section on Classification and Indexing & IFLA Section on Information Technology, OCLC, Dublin, Ohio, USA, 14-16 August 2001.* Dublin, OH: OCLC, 3-10.

Goossens, P. 1991. Multilingual bibliographic access via subject in Europe. In *Library automation and networking: New tools for a new identity.* Herman Liebaers and Marc Walckiers dir. München, Germany: K.G. Saur, 206-214.

Heiner-Freiling, M. 1995. Subject indexing in the nineties: the situation in Germany. In *Subject indexing: Principles and practices in the 90's: Proceedings of the IFLA Satellite Meeting held in Lisbon, Portugal, 17-18 August 1993, and sponsored by the IFLA Section on Classification and Indexing and the Instituto da Biblioteca Nacional e do Livro, Libon, Portugal.* Robert P. Holley, Dorothy McGarry, Donna Duncan et Elaine Svenonius, dir. München, Germany: K.G Saur, 81-93.

Hudon, M. 1997. Multilingual thesaurus construction: Integrating the views of different cultures in one gateway to knowledge and concepts. *Knowledge Organization* 24 (2): 84-91.

IFLA Working Group on Guidelines for Multilingual Thesauri, Classification and Indexing Section. 2005. *Guidelines for*

multilingual thesauri. (document en consultation) <http://www.ifla.org/VII/s29/pubs/Draft-multilingualthesauri.pdf> (consulté sur Internet le 9 mai 2206).

Jouguelet, S. 1995. Evolution of subject indexing practice in France. In *Subject indexing: Principles and practices in the 90's: Proceedings of the IFLA Satellite Meeting held in Lisbon, Portugal, 17-18 August 1993, and sponsored by the IFLA Section on Classification and Indexing and the Instituto da Biblioteca Nacional e do Livro, Libon, Portugal.* München, Germany: K.G. Saur, 64-80.

Landry, P. 2001. The MACS Project: Multilingual Access to Subjects (LCSH, RAMEAU, SWD). *International Cataloguing and Bibliographic Control* 30 (3): 46-49.

_____. 2004. Multilingual subject access: The Linking approach of MACS. *Cataloging & Classification Quarterly* 37 (3-4): 177-191.

Lopes, M.I. et J. Beall. 1999. Principles underlying subject heading languages (SHLS). *Working Group on Principles Underlying Subject Heading Languages, IFLA Section on Classification and Indexing.* München, Germany: K.G. Saur.

MacEwan, A. 2000. Crossing language barriers in Europe: Linking LCSH to other subject heading languages. *Cataloging & Classification Quarterly* 29 (1-2): 199-207.

McIlwaine, I. 1995. Subject control: The British viewpoint. In *Subject indexing: Principles and practices in the 90's: Proceedings of the IFLA Satellite Meeting held in Lisbon, Portugal, 17-18 August 1993, and sponsored by the IFLA Section on Classification and Indexing and the Instituto da Biblioteca Nacional e do Livro, Libon, Portugal.* München, Germany: K.G. Saur, 166-180.

Zeng, M.L. et L.M. Chan. 2004. Trends and issues in establishing interoperability among knowledge organization systems. *Journal of the American Society for Information Science and Technology* 55 (5): 377-395.