

Penser les nouvelles fonctionnalités du thesaurus

Cet article tente de suivre l'évolution dynamique du paradigme linguistique en sciences de l'information à travers les applications de la terminologie. De l'indexation à la recherche d'information multilingue, quels rôles peut jouer le thesaurus au sein des systèmes d'information. La présentation suivra en quelque sorte l'évolution de l'apport des applications de la terminologie aux systèmes d'information. L'accent sera particulièrement mis sur le nouveau rôle que le thesaurus pourrait jouer et sur la nécessité de l'adapter à la nouvelle écriture sur écran. J'examinerai aussi la possibilité de son basculement fonctionnel, vers un usage d'interface pour l'accès à l'information.

Termes-clés :
thesaurus ; réseau sémantique ; accès au contenu ; recherche d'information ; recherche d'information textuelle multilingue ; dispositif de navigation sur le Web.

(1) Les approches conceptuelles caractéristiques de la terminologie ont été à la base du courant de l'acquisition de connaissances, car l'acquisition des connaissances propres à un domaine constitue le fondement même de la terminologie et la distingue des approches lexicographiques. Plusieurs travaux dans le domaine de l'Intelligence artificielle (IA), et du Traitement automatique du langage naturel (TAL), ont été entrepris dans ce cadre (cf. Ahmed 1993, Meyer 1993, Séguela *et al.* 1997; Skuce *et al.* 1997, entre autres).

Introduction

Sous la poussée des technologies et des recherches théoriques multidisciplinaires, les méthodes de traitement et de représentation des textes subissent actuellement de grands bouleversements. Bien que, dans les systèmes d'information, plusieurs fonctions soient encore remplies par l'être humain, de plus en plus de travaux de recherche aboutissent à la mise au point de techniques ou d'heuristiques prenant en charge l'analyse des textes, la constitution de lexiques, celle de thesaurus, l'indexation et l'interrogation de bases de données textuelles, fondées pour certaines sur des approches linguistiques. C'est le cas de plusieurs systèmes actuellement disponibles sur le marché : *Spirit*, *Nomino* pour ne citer que deux systèmes travaillant sur le français. La qualité du résultat dépend alors des capacités du système à traiter la langue dans laquelle est formulée la requête. Les approches linguistiques sont, par ailleurs, de plus en plus utilisées dans des applications du type *Knowledge management* (un processus d'organisation et de mise à disposition d'une information pouvant être contenue dans une messagerie, un système de fichiers, une base de données ou un site Web).

Au sein du paradigme linguistique en sciences de l'information, la terminologie figure au premier rang. Elle intervient dans diverses applications telles que l'indexation (manuelle et automatique), la construction de

vocabulaires contrôlés, les interfaces d'interrogation de bases de données textuelles monolingues et multilingues, la traduction assistée par ordinateur (les outils d'aide à la traduction étant de plus en plus intégrés à des systèmes d'information), le résumé automatique, etc.

Ses liens avec les systèmes d'information et de documentation ont été mis en évidence dans divers travaux (Wüster 1968, Dahlberg 1978 et 1981, Rondeau 1983, Sager 1990). Des travaux plus récents ont souligné l'importance de la terminologie dans les systèmes d'information et de documentation en tant qu'outil d'« extraction de concepts » construits dans les documents par le biais de termes qui servent de vecteurs ou de clés d'accès au contenu. Par exemple, Cabré (1998 : 87-88) écrit :

« (...) des trois questions clés en terminologie par rapport à la pensée et aux connaissances (l'acquisition des connaissances, la formation des concepts et l'organisation des concepts), la troisième est celle qui occupe davantage et l'attention et le travail des terminologues [...]. Cette question a fait l'objet d'analyse de la part de spécialistes des domaines techno-scientifiques qui s'efforcent d'établir la structure conceptuelle de leur discipline ».

Plusieurs outils d'accès à l'information prennent spécifiquement en charge cette dimension conceptuelle ⁽¹⁾. C'est le cas notamment des thesaurus, qui sont fondés sur la prise en compte des relations conceptuelles et sémantiques (représentées sous la forme d'un réseau) organisant un domaine de connaissance. À chaque domaine de

connaissance, on fait correspondre une représentation plus ou moins spécifique. Certains domaines se prêtent ainsi à des classifications hiérarchiques (classifications zoologiques, botaniques, etc.), tandis que d'autres se caractérisent par des relations partitives (méronymie/holonymie) tels que le corps humain/animal, la machine-outil ou les entités géographiques et organisationnelles. Les relations associatives quant à elles caractérisent le domaine des sciences humaines et sociales.

Dans cet article, j'examinerai l'apport de la terminologie à l'accès au contenu informationnel (§ 1.) et comment ses applications peuvent améliorer les approches systèmes. Je prendrai comme exemple le thesaurus⁽²⁾ (§ 2.) qui se situe dans les approches systèmes. Même si ces approches présentent des limites (cf. Temmerman, Cabré ici même), il est possible d'apporter une aide à l'utilisateur au moyen des réseaux sémantiques qui caractérisent le thesaurus. J'examinerai donc la possibilité d'adapter cet outil à la nouvelle écriture sur écran et l'intérêt de son basculement fonctionnel (§ 3.), vers un usage d'interface pour l'accès à l'information.

1 Place de la terminologie et de ses applications au sein des paradigmes d'accès au contenu des documents

Pour situer l'apport de la terminologie aux modèles d'accès au contenu de documents, je donnerai un bref aperçu des deux principaux paradigmes de recherche d'information (cf. Ihadjadène 1999 pour une étude critique et détaillée, Kolmayer 1997). Deux paradigmes sont actuellement dominants : le

paradigme système et le paradigme usager.

1.1 Le paradigme système

Le paradigme système constitue, depuis plus de trente ans, le paradigme dominant en accès à l'information. D'une façon générale, ses partisans considèrent que ce sont les fonctions de traitement de l'information, notamment celles de l'appariement entre les requêtes et les descriptions des documents, qui constituent le cœur du système.

L'idée sous-jacente à ces systèmes est que le degré d'appariement entre les termes de la base et ceux de la requête de l'utilisateur permet d'indiquer la pertinence des documents retrouvés. L'objectif essentiel consiste à améliorer les performances de recherche selon deux mesures : le bruit et la précision. Des études tentent d'améliorer les procédés d'appariement, au travers de divers modèles reposant sur une dichotomie bien connue : l'appariement de formes (les approches booléenne exacte et booléenne étendue, le *clustering*, etc.) et l'appariement du sens (reformulation, paraphrasage, inférence, etc.) qui sont des approches basées sur la linguistique et/ou l'intelligence artificielle. (cf. Dachelet 1990, Mustafa el Hadi *et al.* 1993).

Le système met en correspondance les termes fournis par l'utilisateur et ceux du système (terme d'indexation). Cette opération est réalisée via un appariement syntaxique et sémantique. Une fois cette transformation réalisée (le passage du langage naturel au langage de représentation du système), on aboutit à une expression booléenne mettant en relation des termes d'indexation qui permet alors la sélection des documents. Ces derniers sont ordonnés selon la valeur de la mesure de correspondance avec la requête.

Dans ce type d'approche, on permet à l'utilisateur de formuler sa requête en langage naturel ; on lui évite de connaître et de maîtriser l'usage des opérateurs booléens.

Malgré les limites des approches de l'appariement de forme (cf. Dachelet 1990, Kolmayer 1997, Ihadjadène 1999, entre autres, et § 1.2.), ces approches sont cependant largement utilisées pour des raisons bien connues :

- L'investissement réalisé dans ces systèmes est si considérable que les modifier ne serait pas économiquement viable ;
- Les techniques alternatives n'ont pas été testées dans des environnements en grandeur réelle ;
- Les résultats obtenus par des techniques alternatives ne sont pas à l'heure actuelle suffisamment supérieurs, pour justifier les changements.

1.2 Le paradigme usager

Quant au paradigme usager, ses tenants considèrent que le besoin d'information n'est pas toujours défini et fixe, mais au contraire évolutif. Pendant la recherche surviennent parfois de nouveaux besoins d'information au gré des différentes informations rencontrées (Le Coadic 1998). Ces approches montrent les limites des approches systèmes et rappellent la pertinence des critiques de l'«indexation lexicale» formulées par Amar (1997, 1999 : 4) et les avantages d'une «indexation discursive» qui laisserait ouvert à l'utilisateur d'un système tous les parcours interprétatifs (cf. Amar ici même).

(2) «Un ensemble contrôlé de termes utilisés pour l'indexation de l'information contenue dans une base de données, et utilisés également pour la recherche de l'information dans cette même base de données» (Hudon, 1995).

On considère ici l'utilisateur, son «comportement», son intention comme un élément central (en complément de l'importance de l'appariement) du processus de recherche d'information. Les approches usager témoignent de l'arrivée de chercheurs issus d'autres disciplines telles que la psychologie, les sciences éducatives, les sciences sociales, la communication ou les sciences cognitives.

Contrairement au paradigme système, il n'existe pas une seule approche usager mais plusieurs qu'on peut rattacher à deux grandes catégories générales: les approches cognitives et les approches dynamiques. D'un point de vue méthodologique, ces approches s'inspirent souvent des méthodes ethnographiques et socio-cognitives (*cf.* Kolmayer 1997; Ihadjadène 1999).

2 Intérêt du thesaurus et avantages de sa transformation en dispositif de navigation sur le Web

Avec les méthodes d'indexation automatique et la possibilité d'accès au texte intégral, l'indexation à l'aide de langages documentaires est en perte de vitesse. Ce phénomène devient plus perceptible avec l'Internet. Amar (*cf.* ici même) examine l'évolution de la pratique d'indexation et en distingue deux types qui correspondent chacun à un contexte documentaire donné. On peut, en effet, être d'accord avec Amar (*cf.* ici même) sur l'inadéquation des langages documentaires pour l'indexation.

Cependant, malgré ces critiques, je proposerai ici d'assigner un nouveau rôle aux langages documentaires, tels que le thesaurus. Cet outil représente en effet un savoir existant, partagé par les spécialistes d'un domaine. Et l'interprétation des textes peut faire

appel à ces savoirs. Il importe dans ce cadre d'examiner comment l'on peut conserver ce type d'outils, tout en les adaptant aux nouvelles exigences et à l'amélioration des approches systèmes d'une part, et en les affinant vers une assistance à l'utilisateur d'autre part.

Je présente ci-dessous les principaux avantages de cet outil.

2.1 Structure sémantique du thesaurus et assistance à l'utilisateur

La disposition des relations permet de mieux comprendre et de mieux représenter les aspects notionnels des termes. Les concepts s'organisent ainsi dans différents types de configurations et explicitent les relations notionnelles qui caractérisent un domaine de connaissances. Cette structure sémantique pourra aider l'utilisateur dans sa tâche de recherche de l'information textuelle.

2.2 Amélioration de l'accès au contenu d'une base

La structure sémantique du thesaurus permet, par ailleurs, d'améliorer l'accès au contenu d'une base de données, de réduire le bruit et le silence, d'améliorer le taux de précision et, de ce fait, permet à l'utilisateur de mieux trouver ce qu'il cherche. À la différence d'un index, chaque descripteur du thesaurus est le noyau d'un mini-réseau autour duquel se développe un ensemble de relations logico-sémantiques: générique-spécifique, générique-partitive et spécifique-partitive, relation d'équivalence, relation d'association. Cette relation lie des termes appartenant à des catégories différentes mais où plusieurs cas de figure rappellent la forte implication de l'un des deux termes choisis pour l'indexation chaque fois que l'autre est consulté: discipline ou branche d'un

domaine liés aux objets et/ou aux phénomènes étudiés (sylviculture/forêt); un procédé ou une opération et son agent ou son instrument (régulation de température/thermostat), etc.

2.3 Utilisation des relations sémantiques du thesaurus comme heuristiques dans les systèmes automatiques pour le filtrage de l'information

La requête de l'utilisateur peut être reformulée à l'aide d'heuristiques prédéfinies fondées sur la structure du thesaurus utilisé (*cf.* T. GID, *Spirit*, Fluhr 1992, Fluhr *et al.* 1999). L'objectif de cette reformulation est de produire, à partir de la question de l'utilisateur, la totalité des reformulations équivalentes que l'on pourrait trouver dans les documents. On «gère» ainsi la diversité d'expression linguistique des différentes notions.

Les informations utilisées par le module de reformulation sont de nature syntaxique ou sémantique. Pour la sémantique, on utilise le type de relation existant entre les termes.

Le processus de reformulation peut aussi s'appliquer lorsque la base contient des documents rédigés dans une autre langue que celle de la question.

Les relations peuvent avoir une origine morphologique (termes dérivés de la même base, dérivation morphologique appelée dans le domaine des sciences de l'information «stemmatisation»), ou sémantique. Par exemple la synonymie (SYN), l'appartenance (générique-spécifique (G-S)) d'inclusion partitive (GP-SP), associatives (A), etc.

[i] Exemple de *stemmatisation*:
indexeur <—> indexation

[ii] Exemples de *relations thesaurales*:
index <—A—> descripteur
véhicule <—G—S—> voiture

Les relations peuvent être absolues ou déclenchables après vérification d'un contexte. Elles peuvent être conditionnées par le résultat de l'analyse syntaxique, en particulier par l'étiquetage grammatical :

fin (adj.) —SYN—> étroit
 fin (nom) —A—> arrêt
 poste (nom masculin) —SYN—> fonction
 poste (nom féminin) —SYN—> P.T.T. (3)

2.4 Utilisation du thesaurus dans les environnements de recherche d'information textuelle multilingue (RITM)

La RITM est une méthode de recherche de textes écrits dans des langues différentes de celle qui a été utilisée pour rédiger la requête. Bien qu'il soit possible de traduire tous les documents dans la langue utilisée pour rédiger la requête, pour des collections de grande taille, l'approche la plus économique pour construire un système RITM consiste à traduire la requête dans toutes les langues des collections sur lesquelles porte la recherche au moment de la recherche (cf. Cowie *et al.* 1997, Fluhr *et al.* 1999). Pour dépasser les limitations des dictionnaires bilingues génériques, des lexiques et des thesaurus spécifiques, construits à partir de vocabulaires contrôlés, ont été utilisés pour rechercher des informations textuelles (cf. Cowie *et al.* 1997, Hamann 1998). Hudon (1997: 84) discute l'importante potentielle de l'usage des thesaurus dans les environnements bilingues/multilingues en signalant leur usage dans les systèmes de transfert de connaissances, pratique courante au Canada et dans les pays de l'Union européenne.

3 Adaptations nécessaires du thesaurus

Certaines adaptations sont cependant nécessaires :

3.1 Enrichissement des relations logico-sémantiques dans les thesaurus

La richesse d'un thesaurus se mesure à celle du réseau sémantique issu de sa représentation conceptuelle. Beaucoup de thesaurus souffrent de la pauvreté des relations sémantiques utilisées : on n'y trouve généralement que les relations de base génériques/spécifiques (fausses dans certains cas), les relations d'équivalence et les relations associatives. Quant à la relation partitive, elle est quasiment absente et souvent remplacée par la relation hiérarchique. Les relations sémantiques qui caractérisent les travaux terminologiques peuvent en partie servir de modèle pour la construction d'un tel outil. Ces relations sont détaillées dans les travaux suivants : Wüster (1968), Felber (1984), Cruse (1986), Van Campenhoudt (voir notamment le site de Termisti). Dans ce courant classique de la terminologie, le réseau sémantique de *WordNet* (4) peut inspirer l'élaboration des relations sémantiques des thesaurus. Bien que cette base de données soit considérée comme un produit hybride, car elle réunit certaines caractéristiques des bases de données lexicales et d'autres tirées de certaines bases de données terminologiques enrichies par des relations sémantiques, elle pourrait cependant servir de standard. La typologie des relations correspond, en effet, à celle qui est présente aussi bien dans les bases de données terminologiques que dans les thesaurus de langue (comme celui de Roget) ou dans les thesaurus documentaires.

Autrement dit, nous avons l'exemple d'un modèle complet de relations sémantiques associées à un thesaurus dynamique.

Les approches cognitivistes et socioterminologiques auront par ailleurs un rôle important dans l'enrichissement des relations sémantiques (cf. Temmermann, Diki-Kidiri ici même). Ces approches permettraient de préciser et de développer le type de relations sémantiques nécessaires à la transformation de thesaurus en dispositifs de navigation. Si ces dispositifs visent essentiellement à assister l'utilisateur et améliorer les approches systèmes, la prise en compte des aspects cognitifs, culturels et sociologiques est fondamentale. Les approches classiques et innovantes pourront alors être complémentaires.

3.2 Mises à jour permanentes

Le thesaurus est un outil qui devrait idéalement être en construction permanente, puisqu'il représente le corpus sémantique d'un domaine en évolution. Les relations établies ne devraient pas être figées une fois pour toutes, les spécialités évoluant rapidement.

3.3 Transformation de thesaurus en dispositif de navigation sur le Web

Un dispositif de navigation est un dispositif d'organisation du contenu, qui peut prendre la forme d'un réseau conceptuel de type thesaurus, et qui

(3) Pour plus de détails sur la reformulation, cf. Fluhr (1992), Fluhr *et al.* (1999), *Spirit, Manuel d'utilisation*, T. GID (Groupe Technologies).

(4) www.cogsci.princeton.edu/~wn/

propose à l'utilisateur de parcourir les nœuds du réseau en allant d'un concept général à un concept particulier et vice versa. Concevoir les thesaurus de cette manière permet alors d'adapter les modes d'accès aux besoins des utilisateurs.

Ce nouveau type de fonctionnement est d'ores et déjà à l'œuvre dans *Multites Thesaurus construction* ⁽⁵⁾, qui est relié, dans sa version d'évaluation en ligne, aux moteurs de recherches tels que Altavista et Yahoo.

La mise en œuvre informatique de ces dispositifs de navigation demande alors de comparer l'architecture des thesaurus et la structure sous-jacente aux sites Web et de déterminer le type de lien logico-sémantique le plus apte à être utilisé comme lien hypertextuel pouvant pointer sur des sites Internet liés aux domaines de connaissances couverts par le thesaurus/dispositif de navigation.

On peut ainsi imaginer que dans un avenir proche des thesaurus conçus manuellement ⁽⁶⁾ ou à l'aide de listes de termes extraits automatiquement pourront être transformés en dispositif de navigation.

Conclusion et perspectives

Quelle que soit l'application, il est devenu nécessaire de revoir les normes de construction de thesaurus en tenant compte des points suivants :

- Mise à jour des normes selon les nouvelles exigences de l'écriture sur écran ;
- Développer les relations logico-sémantiques issues des approches classiques et faire en sorte qu'elle s'adaptent aux diverses formes de représentations ;
- Enrichir la réflexion en intégrant les travaux terminologiques basés sur des approches de la sémantique cognitive (cf. ici même Temmerman) ;

– Concevoir les langages d'indexation (terminologies et thesaurus) en donnant une plus grande place à l'utilisateur, ce qui suppose (1) que ce soient des systèmes qui tiennent compte de différents types de connaissances ; (2) d'envisager l'emploi de « niveaux de langue » différents et de parcours distincts d'accès à l'information.

Compte tenu de l'incessante augmentation du volume de l'information et de l'accroissement du nombre des utilisateurs des systèmes de recherche d'information, le besoin d'adapter les réponses aux particularités des utilisateurs s'impose. De ce fait, l'efficacité d'un système de recherche d'information ne dépend pas seulement de l'exactitude des réponses mais également de l'adéquation des réponses aux particularités des utilisateurs.

*Widad Mustafa el Hadi,
Équipe Ingénierie linguistique et
systèmes d'information,
Université Charles De Gaulle-Lille III,
Lille.*

Bibliographie

Ahmad (K.) 1993 : « Terminology and Knowledge acquisition: A Text-Based approach », *TKE' 93*, pp. 56-70.

Amar (M.) 1997 : *Les fondements théoriques de l'indexation, une approche linguistique*, Thèse de doctorat de l'université Lyon II.

Amar (M.) 1999 : « L'indexation : un choix de mots ou un choix de textes ? Vers une " indexation discursive " sur l'Internet », communication présentée lors des Deuxièmes journées du Chapitre français de l'ISKO, Lyon, 21-22 octobre 1999.

Cabré (M.-T.) 1998 : *La terminologie, Théorie méthode et applications*, Armand Colin & les Presses de l'Université d'Ottawa (U. Linguistique).

Cowie (J.), Davis (M.) et Zajac (R.) 1997 : « Text Evaluation Conference TREC - Un point de vue multilingue », dans *Actes des 1^{res} JST Francil*, Avignon, France, 15-16 avril 1997, pp. 1-5.

Cruse (D. A.) 1986 : *Lexical Semantics*, Cambridge, Cambridge University Press.

Dachelet (R.) 1990 : « État de l'art de l'informatique documentaire », Inria, Rapport de recherche n° 1201, Paris, avril 1990, publié dans *Le Document électronique*, Cours Inria dirigé par Christian Bornes, 11-15 juin 1990, Châtelaillon (Charente Maritime), pp. 107-132.

Dahlberg (I.) 1978 : « Conceptual definitions », *International Classification*, vol. 5, n° 3, pp. 142-150.

Dahlberg (I.) 1981 : « Conceptual Definitions for Interconcept », *International Classification*, vol. 8, n° 1, pp. 16-22.

David (S.) 1993 : *Les unités polylexicales: éléments de description et reconnaissance automatique*, Thèse de doctorat en linguistique théorique et formelle, Paris, Université de Paris VII.

Dumas (L.) Plante (A.) et Plante (P.) 1997 : *Analyseur Linguistique de Nomino, Version 1*, Montréal, UQAM.

Fellbaum (Ch.) 1998 : *WordNet, An Electronic Lexical Database*, MIT Press.

(5) www.concentric.net/~Multites/thesaurus/.

(6) Dans le cadre de la Maîtrise SID (Systèmes informationnels et documentaires, Université Lille III) les étudiants construisent des thesaurus dont ils choisissent eux-mêmes les thèmes. Nous disposons ainsi de thesaurus qui couvrent 70 domaines environ de l'activité humaine. Ils vont de la *Dégustation du vin à l'Euro* en passant par le *Diabole et la sorcellerie au moyen âge*. Un projet en cours de réalisation vise à transformer certains de ces thesaurus en dispositifs de navigation.

- Felber (H.) 1984: *Manuel de terminologie*, 1984, Paris, Unesco.
- Fluhr (Ch.) 1992: «Le Traitement du langage naturel dans la recherche d'information documentaire», dans *Interfaces intelligentes dans l'information scientifique et technique*, Inria dirigé par Christian Bornes, 18-22 mai 1992, Klingthal (Bas-Rhin), pp. 103-130.
- Fluhr (Ch.), Schmit (D.), Andrieux (C.), Ortet (Ph.) et Combet (V.) 1999: «Cross-lingual Interrogation for Multilingual Catalogs», présenté pour Le *Third European Conference on Research and Advanced Technology for Digital Libraries*, Paris, France, 22-24 septembre 1999.
- Harmann (D.) 1998: «The Text Retrieval Conferences (TREC) and the Cross-Language Track», in *Proceedings of the First International Conference on Language Resources and Evaluation*, Grenada, Espagne, 28-30 mai 1998, pp. 517-522.
- Hudon (M.) 1995: *Le thesaurus: conception, élaboration, gestion*, Montréal, Asted.
- Hudon (M.) 1997: «Multilingual Thesaurus construction: Integrating the Views of Different Cultures in One Gateway to Knowledge and Concepts», *Knowledge Organization*, vol. 24, n° 2, pp. 84-91.
- Ihadjadene (M.), 1999: *La recherche et la navigation dans un SRI grand public: le cas des Hypercatalogues*, Thèse de doctorat, Université de Lyon I.
- Kolmayer (E.) 1997: «Organisation des connaissances et expertise chez les usagers», dans Maniez et Mustafa el Hadi, *Organisation des connaissances en vue de leur intégration dans les systèmes de représentation et de recherche d'information*, Coll. Travaux & Recherche, Édition du conseil scientifique de l'université Charles-De-Gaulle-Lille III.
- Le Coadic (Y. F.) 1998: *Le besoin d'information*, ADBS édition, Paris 1998.
- Meyer (I.) et Bowker (L.) 1992: «Cogniterm: An Experiment in Building a Knowledge-Based Term Bank», in *Proceedings of the Fifth Euralex International Congress* (Euralex '92), Tampere, Finlande, Tampereen Yliopisto, pp. 159-172.
- Mustafa el Hadi (W.) 1993: «État de l'art de l'extraction de concepts à partir du langage naturel», Rapport SGDN/Credo, Université Lille III.
- Mustafa el Hadi (W.) 1998: «Automatic Term Recognition & Extraction Tools: Examining the New Interfaces and their Effective Communication Role in LSP Discourse», dans Mustafa el Hadi (W.), Maniez (J.) et S. Pollitt (édit.), 1998: *Structures & Relations in Knowledge Organization*, Proceedings of the 5th International ISKO Conference, Lille 25-29 Août, Würzburg, Ergon Verlag, pp. 204-212.
- Oard (D.) et Diekema (A.) 1999: Cross-Language Information Retrieval, in *Annual Review of Information Science and Technology*, vol. 33, pp. 223-256.
- Rondeau (G.) 1983: *Introduction à la terminologie*, 2^e édition, Chicoutimi, Gaëtan Morin.
- Sagar (J.-C.) 1990: *A Practical Course in Terminology Processing*, Amsterdam/Philadelphia, John Benjamins Publishing Company.
- Séguéla (P.) et Aussenac-Gilles (N.) 1997: «Un modèle de base de connaissances terminologique», dans *TIA-97, Actes des deuxièmes rencontres «Terminologie et intelligence artificielle*», Université Toulouse Le Mirail, 3-4 avril 1997.
- Skuce (D.), Mayer (I.) et Kavanagh (J.) 1997: «Bases textuelles et bases de connaissances sur le Web: le système Ikarus», *Ves Journées Scientifiques du Réseau thématique LTT*, Tunis 25-27 septembre 1997.
- T. GID (Groupe Technologies) 1995: *Spirit, Manuel d'utilisation*.
- Van Campenhoudt (M.) 2000: *Cours de terminologie*, site de Termisti: www.refer.org/termisti/termisti.htm.
- Wüster (E.) 1968: *Dictionnaire multilingue de la machine outil*, Londres, Technical Press.