

Les outils de la terminotique:

Typologie des logiciels d'aide à la terminologie et/ou d'automatisation de la chaîne de travail en terminographie

L'automatisation de la chaîne de travail en terminographie fait appel à l'utilisation de nombreux logiciels, ce texte propose une typologie des logiciels applicables à la terminotique et décrit les nouveaux environnements informatiques d'aide à la terminographie tels que le propose le poste de travail du terminologue. Les analyseurs textuels qui offrent aux terminologues des moyens performants d'exploration et d'utilisation du texte technico-scientifique sont ici présentés plus en détail en regard des nombreux services qu'ils peuvent rendre dans le traitement de la chaîne de travail en terminographie.

Termes-clés:
traitement des langues naturelles;
technologies de traitement de
l'information écrite;
automatisation lexicale; terminotique;
terminologie assistée par ordinateur.

La première rencontre de la terminologie avec l'informatique ne date pas d'hier, il faut la situer au début des années 1960 avec l'avènement de projets comme Dicautom en 1963 (Dictionnaire automatisé de la Communauté européenne du charbon et de l'acier, Luxembourg) qui deviendra Eurodicautom quelque dix ans plus tard (1973, Banque de données terminologiques de la Commission des communautés européennes). Quelques années plus tard, d'autres grands projets de banques de terminologie voient le jour également, pour ne citer que les systèmes BTM, Banque de terminologie de l'université de Montréal, qui donnera naissance au début des années 1970 à Termium (Banque de terminologie du Gouvernement du Canada), BTQ (Banque de terminologie du Québec - de l'Office de la langue française) et Lexis (*Bundessprachenamt*, RFA). Cette rencontre est également une première pour l'époque parce qu'elle aura permis d'utiliser des données textuelles pour nourrir les systèmes de gestion de bases de données initialement développés pour traiter des valeurs numériques.

«La lexicographie et la terminographie sont à compter parmi ces applications qui ont bénéficié du traitement automatique des données que permettent les ordinateurs. La rencontre de la terminologie appliquée et de l'informatique a créé un champ nouveau de travail qu'on a dénommé *terminotique*. Fait curieux, la terminologie a été l'une des premières disciplines de la linguistique à avoir utilisé activement l'électronique pour diffuser ces

données avec les banques de terminologie, mais elle a tardé à tirer profit de l'informatique pour d'autres types de traitement qui concernent en particulier la chaîne de traitement des recherches terminologiques et terminographiques.» (Auger *et al.*, 1991: 121.)

L'avènement de la micro-informatique à compter des années 1980 et la large diffusion sociale qui s'ensuivit, de même que les progrès du génie logiciel d'orientation textuelle auront permis de dépasser cette seule fonction de stockage et de gestion de données terminologiques des banques de termes. Ce n'est qu'à compter du milieu des années 1980 qu'on pourra parler de *terminotique*, nouvelle étiquette recouvrant les divers processus d'automatisation terminographique.

On peut définir la terminotique de deux façons différentes. Dans la première, très générale, elle est «le domaine qui vise l'automatisation des tâches du terminologue» (P. A.). Dans la seconde, la terminotique est entendue comme «l'ensemble des techniques et des outils qui contribuent à automatiser la chaîne de production terminographique» (P. A.). Cette seconde définition, plus explicite, doit ici être complétée en précisant que les techniques dont il est fait état dans cette définition sont celles de l'informatique textuelle et de l'informatique linguistique et que les outils, quant à eux, sont ceux tirés des nouvelles technologies de l'information (NTI). Dans le cas d'une automatisation partielle, on parlera plutôt de *terminologie assistée par ordinateur* (sur le modèle de *traduction assistée par ordinateur*). Le présent article va proposer une typologie des outils (ou des aides)

logiciels organisée selon le déroulement de la chaîne de travail du terminologue. Cette typologie va inclure les outils propres à la terminologie, comme ceux de la terminologie assistée par ordinateur.

Lorsque l'aide fournie au terminologue est systématique et que les outils logiciels sont implantés sur un poste informatique individuel et qu'ils fonctionnent le plus souvent sous un même logiciel intégrateur de type *Windows* (Microsoft Corporation), on parle alors de «poste de travail du terminologue»⁽¹⁾. Ainsi, le Bureau de la traduction du

(1) L'expression *Poste de travail du terminologue* a été créée par le Bureau de la traduction du Gouvernement canadien, un des premiers services publics au monde à avoir conduit des recherches dans ce domaine.

(2) Élisabeth Blanchon du CTN (CNRS, Inalf) a proposé une typologie des logiciels d'aide à la terminographie qu'elle classe en six catégories: logiciels de terminologie, logiciels documentaires, dictionnaires électroniques, projets de recherche, et logiciels de terminologie «maison» (Blanchon 1991: 168). Le lecteur trouvera dans cette typologie la liste de nombreux logiciels d'aide à la terminographie, particulièrement de la première catégorie «logiciels de terminologie».

(3) Pour des raisons évidentes, nous avons dû restreindre le nombre de logiciels cités en exemple. Il va sans dire que la liste présentée dans cette partie est limitative et qu'il existe de nombreux produits comparables ou substituables à ceux cités ici. Cette liste restreinte n'est pas basée sur des considérations d'ordre qualitatif. Ainsi, par exemple, il existe sur le marché de nombreux concordanciers aussi valables les uns que les autres qui peuvent être utilisés avec profit par le terminologue.

Gouvernement canadien a mis au point un tel poste. Le poste en développement est distinct du projet «poste de travail du traducteur (PTT)» lequel comporte toutefois des logiciels d'assistance terminologique (gestion de glossaires, recherche de termes, dictionnaires électroniques, etc.). Ce dernier poste a été mis au point par le Centre d'innovation en technologie de l'information (Citi, Montréal), anciennement Centre canadien de recherche en informatisation du travail (CCRIT).

Le travail du terminologue, en simplifiant un peu, consiste à amasser et sélectionner des données [1], à traiter ces données selon les méthodes de la terminographie [2] et enfin à extraire des données formalisées pour être diffusées dans des répertoires terminologiques (dictionnaires, lexiques, vocabulaires, etc.) [3] (Auger *et al.* 1991). Ces trois grandes divisions du travail terminologique vont nous fournir la première articulation (niveau) de la typologie que nous allons présenter ici. Le deuxième niveau, quant à lui, sera constitué d'un survol des étapes possibles d'automatisation du traitement de la chaîne de travail terminographique. L'ensemble présenté est fondé sur la chaîne de travail en usage en terminographie. On peut ajouter ici qu'aujourd'hui les pratiques terminographiques sont relativement normalisées au plan international, ce qui justifie la typologie présentée ici.

1 Des logiciels pour chaque tâche du terminologue⁽²⁾

Le traitement informatique de la chaîne de travail en terminographie nécessite l'utilisation de plusieurs types de logiciels dédiés au traitement de l'information textuelle⁽³⁾.

Plusieurs d'entre eux sont des produits génériques disponibles dans le commerce, d'autres sont soit des produits spécialisés dédiés au traitement linguistique ou encore des produits spécifiquement développés pour le terminologue - terminographe. Ils permettent de colliger les données recueillies par le terminologue [1], de les traiter [2] et de les éditer [3]. Leur utilisation est liée au degré d'automatisation désiré par le terminologue - terminographe.

1.1 Les logiciels de collection de données

L'automatisation des divers processus de la chaîne de travail en terminographie repose sur la disponibilité de données sous forme numérisée, il est donc important que la saisie informatique des données supportant la démarche du terminologue se situe tout au début du processus, d'où la nécessité de disposer d'outils de saisie adéquats.

1.1.1 Collection de données sur le terrain (sources orales)

Les nouvelles orientations tracées par la socioterminologie vont amener de plus en plus les terminologues à travailler sur le terrain en se préoccupant des divers niveaux d'usage qui constituent les langues de spécialité et les différents discours socioprofessionnels. En outre, l'implantation et le développement des langues nationales dans les pays du Sud en se fondant sur une description préalable vient renforcer cette option.

- Logiciels d'enregistrements de données: type mini-SGBD. Ce type de logiciel permet de saisir rapidement des données sur le terrain sous forme ou non de fiches et de les réexporter vers d'autres systèmes plus élaborés pour un traitement ultérieur. p. ex. *Shoobox* v.12a, SIL, Texas), Celar.

- Logiciels de traitement de texte (principalement les éditeurs de type Ascii). Ils permettent la saisie directe, la manipulation et la mise en forme des données recueillies afin d'en permettre ultérieurement le traitement.

1.1.2 Traitement de données écrites (Roc)

La terminotique s'appuyant essentiellement sur la disponibilité d'un corpus numérisé, le terminographe doit dans bien des cas procéder à la saisie optique des textes à analyser ou au reformatage de textes déjà disponibles sous divers formats.

- Logiciels de Roc (reconnaissance optique de caractères): p. ex. *Omnipage Pro*. Ce type de logiciel permet de numériser des textes imprimés pour les rendre utilisables par un ordinateur.

Leur utilité n'est plus à démontrer dans la perspective de l'utilisation par le terminologue de larges corpus pour fonder son travail.

- Logiciels de transfert de données numérisées de format x à format y. D'ores et déjà, les grands traitements de textes commerciaux permettent directement un grand nombre de transferts interformats (p. ex. *WordPerfect* à *Word*) et même intersystèmes (p. ex. de Mac à PC).

1.1.3 Téléchargement de données déjà numérisées (télématique)

De plus en plus les grands réseaux informatiques (*Internet*, par exemple) diffusent des corpus de textes déjà numérisés qui sont alors rendus disponibles aux utilisateurs qui n'ont qu'à les télécharger depuis leur poste de travail. Ces bases de données textuelles servent à de nombreux utilisateurs d'horizons très différents: spécialistes techniciens ou scientifiques, chercheurs des sciences humaines, documentalistes, etc.; elles peuvent constituer une source «économique» de textes prêts à être traités pour le linguiste ou le terminologue.

- Logiciels standards de communication: p. ex. *ProCom*, *Telix*, *Kermit*. Ce type de logiciel permet de communiquer à distance avec des serveurs externes et d'importer de l'information directement réutilisable en interrogeant des bases de données.

- Logiciels de communication de type Fax: p. ex. *WinFax Pro3*. Ce type de logiciel permet même de numériser à la réception des textes imprimés expédiés à partir de télécopieurs conventionnels.

- Logiciels de «capture d'écran» p. ex. *Snipper*. Ce type de logiciel permet de charger sélectivement des données apparaissant sur un écran d'ordinateur et de les stocker dans des fichiers électroniques.

- Logiciels d'interrogation plein-texte pour les Doc (disques optiques compacts), qui permettent d'extraire des données sur disque optique compact de type texte et de les stocker dans des fichiers électroniques; p. ex. *La Presse électronique*, *Le Robert électronique*.

1.2 Les logiciels de traitement des données

La liste des nombreux logiciels énumérés dans cette partie illustre la complexité des processus de manipulation de l'information textuelle mis en œuvre dans la perspective d'une automatisation de la chaîne du traitement terminographique. Dans les faits, le degré d'automatisation des procédures terminographiques varie grandement d'un projet à l'autre. Au minimum, les données terminologiques sont stockées sur un support informatique (traitement de texte ou SGBD) pour en permettre l'interrogation. L'automatisation totale, à l'opposé, met en œuvre un traitement beaucoup plus poussé de l'information textuelle. Dans ce cas, la terminologie est extraite automatiquement de textes en format

électronique, puis traitée (sélectionnée, enrichie, stockée et éditée) par l'ordinateur jusqu'à l'obtention du produit final.

- Logiciels de conversion de textes sous et vers divers formats. P. ex. de *WordPerfect 5.1* au format Ascii ou SGML⁽⁴⁾.

- Logiciels de correction automatique ou assistée de textes:

corrections orthographique et grammaticale principalement. P. ex. correcteurs des grands progiciels de traitement de textes (*WordPerfect*, *Word - Grammatik*) ou correcteurs dédiés comme *Hugo* (Logidisque, Montréal), *Correcteur 101* (Machina Sapiens, Montréal).

- Logiciels de statistique permettant l'obtention de données quantitatives concernant les textes analysés (corpus). P. ex. *Cart* pour Dos (Systat, Illinois, U.S.A.) et *Hyperbase* (sous Mac *Hypercard*, E. Brunet, Un. de Nice).

- Logiciels de segmentation de textes* en unités lexicales simples et/ou complexes (dépouillement terminologique) en modes automatique ou assisté.

[* Les analyseurs de type plein-texte comme *Sato* (Centre d'Ato, Uqam), *Tact* (Un. de Toronto), *WordCruncher* (Johnston & Company, Utah) comportent nécessairement cette fonction].

P. ex. segmenteurs comme *Dat* (segmenteur de textes en unités simples, J. Ladouceur, Un. Laval),

(4) Des groupes spécialisés comme TEI (*Text Encoding Initiative*) et Iso travaillent à l'élaboration de normes pour l'encodage des données structurées ou textuelles de façon à être facilement exportées d'un système à un autre (cf. respectivement *Base for terminological data* de TEI et *Mater* d'Iso, ainsi que Iso 8879 SGML). On peut encore citer la norme *MicroMater*, dérivée de *Mater*, mise au point par le groupe TRG à Brigham Young University (Provo, Utah), conjointement avec l'*American Translators Association*.

Syrex (segmentation assistée de textes en unités complexes, J. Ladouceur, Un. Laval), *NP-Extractor* (D. Bourigault, France) ou *Termino* (segmentation automatique de textes en unités simples et complexes, OLF - Centre d'Ato).

- Logiciels de concordances* (concordanciers permettant l'extraction automatique de contextes).

[* Les analyseurs de type plein-texte comme *Sato* (Centre d'Ato, Uqam), *Tact* (Un. de Toronto), *WordCruncher* (Johnston & Company, Utah) comportent le plus souvent cette fonction].

- Logiciels d'analyse de collocations* permettant d'identifier automatiquement les groupes répétitifs de lexèmes (syntagmes terminologiques, candidats au TC, phraséologismes).

[* Un analyseur comme *Tact* comporte cette fonction qui peut être effectuée sur un ou plusieurs lexèmes et même sur l'ensemble d'un texte].

(5) Il existe aujourd'hui un grand nombre de ces logiciels de type SGBD dédiés au stockage de données terminologiques. Ces logiciels permettent le plus souvent au terminologue et au traducteur de créer des « glossaires personnels » adaptés à leurs besoins et de les interroger directement à partir d'un traitement de texte (par exemple, *Termex*, un des plus connus). Dans ses formes plus évoluées, ce type de logiciel permet de gérer de puissantes banques de terminologie sur micro-ordinateur PC ou Macintosh (p. ex. *Multiterm* pour PC fonctionnant sous l'intégrateur Windows). Certains logiciels de gestion terminologique permettent en outre l'exportation de données terminologiques « maison » vers une banque de terminologie, par exemple *Latter*, une composante du poste de terminologue mis au point au Secrétariat d'État du Canada qui permet d'exporter facilement des données vers *Terminium*.

- Logiciels de lemmatisation (ou lemmatiseurs)* qui permettent de rassembler sous une forme de base l'ensemble des occurrences d'un même mot / terme.

[* Cette fonction est rarement effectuée par les analyseurs textuels conventionnels, certains comme *Sato* permettent d'attribuer automatiquement des propriétés grammaticales ici - à chaque item lexical grâce à l'application de dictionnaires préalablement construits].

- Éditeurs de textes (de type Ascii ou autres) ou interfaces spécifiques pour sélectionner et trier l'information (cf. fichiers de concordances, listes de termes etc.).

- Logiciels d'interface texte - SGBD (de types relationnel - standards ou gestionnaires de glossaires personnels -, SQL, textuel et hypertextuel) pour le stockage en format structuré de l'information sélectionnée⁽⁵⁾. P. ex. *Termex* et *Dimap-2* (*Dictionary Maintenance Utilities*, v. 2.0, CL Research, Gaithersburg, Maryland).

- Logiciels d'analyse notionnelle (comparateurs avancés de textes) permettant soit une représentation arborescente des notions présentées dans un texte ou encore la mise en parallèle de bases de données textuelles en LA et en LD sur la base d'une analyse lexicale thématique (fusion de textes et analyse de contenu). P. ex. *Fusion* (J. Ladouceur, Un. Laval, Québec) et *WordNet* (*An on-line English Lexical Reference System*, Princeton University).

1.3 Logiciels d'extraction d'informations terminographiques

Les produits modernes basés sur l'interrogation en langue naturelle permettent à l'utilisateur un accès facilité aux bases de données structurées ou plein-texte.

- Logiciels d'interrogation de bases de données (structurées, textuelles,

hypertextuelles) en langage procédural ou en langue naturelle permettant la production directe de dictionnaires électroniques. P. ex. *Naturel* qui permet la fouille en langue naturelle de textes en formats divers (traitement de textes de tous formats, SGBD, hypertextes) et maximise l'extraction de l'information en format SGBD.

- Logiciels d'édition permettant l'édition automatique de dictionnaires papiers.

2 L'informatisation de la chaîne de travail en terminographie

La terminographie, nous l'avons vu, comme de nombreuses disciplines qui reposent sur le traitement de l'information écrite, a évolué naturellement dans le sens d'un rapprochement constant avec l'informatique et les moyens nouveaux que celle-ci offre. Ce domaine d'application qui privilégie le mot (ou le terme) sous ses aspects formels repose sur un traitement à la fois extensif et intensif de l'information textuelle; extensif parce qu'il traite du texte sous toutes ses formes, prenons pour exemple la diversité des informations nécessaires à l'accomplissement du travail du terminologue et les tâches de traitement qui s'y rattachent; intensif parce qu'il exige désormais le traitement d'un corpus documentaire considérable mettant en œuvre des tâches lourdes, invariablement répétitives et souvent peu adaptées au fonctionnement de l'humain. Ici l'informatique permet de se rapprocher sensiblement d'un idéal méthodologique d'exhaustivité et de rapidité difficilement concevable pour un acteur humain, mais tout à fait réalisable compte tenu des vertus de l'automate.

Qu'il s'agisse de:

- dépouillement automatique et assisté de textes;
- constitution automatique de bases de données lexicographiques et terminologiques;
- élaboration automatique de dictionnaires généraux ou spécialisés;
- préparation de dictionnaires électroniques, généraux ou spécialisés pour assister les systèmes de TAO;
- création de postes de travail automatisés et intégrés pour le lexicographe, le terminologue ou le traducteur;
- élaboration de thésaurus documentaires multilingues en langue naturelle;
- d'analyse de contenu de textes (extraction et hiérarchisation notionnelles);
- correction automatique et assistée de textes ou de bases de données;
- génération automatique de textes, des processus d'automatisation et de manipulation de l'information textuelle doivent avoir été mis en œuvre préalablement ou concurremment à l'élaboration de ces outils. Pour ne donner qu'un exemple, toute l'informatique d'orientation documentaire contemporaine repose en bonne partie sur la disponibilité de dictionnaires terminologiques électroniques robustes qui doivent être confectionnés rapidement au fur et à mesure des besoins ponctuels qui se présentent.

En considérant, cette fois, la chaîne de travail terminographique et en examinant de près le processus méthodologique lié à la confection d'un dictionnaire terminologique du début des travaux jusqu'à sa mise en marché, on peut distinguer quatre grandes catégories de tâches:

1) Préparation du travail terminologique: il s'agit ici d'accomplir des tâches préliminaires qui sont avant tout d'ordre documentaire (exploration du

domaine, recherche de documentation, sélection et enregistrement des sources);

2) Travail terminographique proprement dit: l'accomplissement de tâches proprement lexicographiques ou terminographiques liées au traitement d'écrits généraux et spécialisés (mise en forme et traitement d'un corpus de textes, dépouillement, établissement de la nomenclature, extraction de données terminologiques et paraterminologiques, rédaction de définitions, établissement des équivalences entre deux langues etc.);

3) Le stockage et le traitement des données recueillies (structuration de bases de données textuelles enrichies, relationnelles etc.);

4) L'édition et la diffusion d'un produit final (dictionnaire terminologique) accessible sous diverses formes.

Tous les projets terminographiques comportent ces quatre grandes phases, mais là s'arrêtent les similitudes; dans la réalité, la différenciation des besoins et des approches (traductionnelle, terminologique traditionnelle, aménagiste, socioterminologique, textuelle discursive, cognitiviste) a fait en sorte que les façons de faire des individus et des groupes varient considérablement et qu'il est illusoire de songer à l'élaboration d'outils terminologiques génériques et universels.

Toutefois, malgré ces divergences, il n'est pas vain de tenter d'isoler les opérations fondamentales que met en œuvre l'approche terminographique systématique fondée sur l'analyse de textes de LSP, courant particulièrement fertile ces dernières années et qui permet d'utiliser au maximum les possibilités de l'ordinateur et des logiciels qui les régissent. Dans cette approche transparaît clairement la notion d'*atelier* ou de *poste de travail*

terminographiques où sont conduites autour d'un automate d'analyse textuelle (analyseur plein-texte) diverses stratégies d'extraction d'informations terminologiques et paraterminologiques, comme le découpage du texte en ses constituants formels morphologiques et syntaxiques (« mots », termes simples, termes complexes, collocations, phraséologismes, contextes), sémantico-cognitifs (repérage notionnel, hiérarchisation notionnelle, synonymie, polysémie, procédés définitoires, attribution des anaphores) et pragmatiques (formulations spécifiques aux textes de LSP, stratification lexicale et niveaux terminologiques). Ces descriptions peuvent, en outre, être qualifiées par l'attribution de diverses propriétés aux items formels du texte (enrichissement des textes par divers marquages: grammatical, notionnel, etc.) et quantifiées (utilisation de mesures terminométriques: fréquence absolue et relative, répartition, dispersion, etc.).

Dans un second temps seront engagées des stratégies de sélection de l'information extraite (cf. les données terminologiques directement extractibles: entrées (vedettes), variantes et synonymes, contextes, références (sources), puis celles inférées du texte: définitions, notes sur la notion ou sur l'usage). En effet, la puissance des analyses-machine et l'exhaustivité qu'elle permet, produit, au contraire des analyses manuelles, une prolifération de l'information qu'il convient de gérer de façon particulière pour en maximiser les bénéfices. Une telle stratégie, par exemple, permettra de tirer les meilleurs contextes du corpus, d'exploiter les procédés définitoires utilisés dans les textes pour la rédaction de définitions, d'utiliser les nombreux contextes « délaissés » pour la rédaction de notes linguistiques, encyclopédiques et

notionnelles. Bref, utiliser au maximum l'information extraite des différentes analyses et n'utiliser que cette information (objectivité des divers recensements). Dans l'optique de la prise en compte des niveaux socioterminologiques des différents types de discours spécialisés (cf. typologie des textes et constitution du corpus), une foule de renseignements utiles pourront être tirés des analyses (p. ex. phénomènes de simplification, définition des argots professionnels, etc.).

En une troisième étape, toute l'information colligée et triée est stockée sous forme de base de données électronique sur un support quelconque (traitement de texte, SGBD, hypertexte) pour constituer le dictionnaire terminologique final, accessible « en ligne » (format plein-texte ou structuré) ou en format papier. L'avenir de ces dictionnaires électroniques réside très certainement

dans le format plein-texte enrichi, plus dynamique que le format structuré (fiches traditionnelles des SGBD).

Un pareil programme ne va pas sans poser bien des difficultés, les postes actuels de terminographie demeurant essentiellement une réunion de logiciels divers qui fonctionnent sous un intégrateur de type *Windows*. Les logiciels ne sont pas toujours compatibles les uns avec les autres (routines de transfert d'information à élaborer pour lier chacune des phases), ils présentent - les analyseurs textuels au premier chef - des difficultés énormes d'exploitation pour l'utilisateur moyen, etc. La plus grande difficulté consiste à mettre au point un protocole de travail très strict où la planification des opérations est bien effectuée, de même que l'intégration des liens entre les différentes étapes en terme de transfert des données, de façon à assurer une gestion efficace et sécuritaire de l'information.

3 Conclusion

En conclusion à cette brève revue des méthodes de travail de la terminotique et des logiciels utilisables pour automatiser la chaîne de travail en terminographie, on peut dire qu'il est aujourd'hui possible pour le terminologue de travailler dans cette voie, sans posséder de compétence particulière autre qu'être devenu un bon utilisateur du micro-ordinateur pour se protéger des aléas les plus courants en informatique, soit les pertes de données dues à de mauvaises manipulations ou au mal fonctionnement du matériel⁽⁶⁾. Un autre problème engendré cette fois par l'utilisation des outils de manipulation du texte en est un de volume de l'information générée et partant d'une gestion efficace d'une masse d'information forcément volumineuse et parfois même

écrasante pour un manipulateur humain. Ces problèmes résolus, le terminologue va obtenir les bénéfices attendus pour ses efforts: économie de temps appréciable, résultats de travail plus cohérents (ou systématiques), exhaustivité de ses recensements, etc. Le bénéfice le plus important vient du fait que l'automatisation oblige le terminologue à travailler avec des textes et que ce faisant, il va fonder ses travaux sur la seule exploitation des textes analysés, ce qui laisse peu ou pas de place pour les jugements fondés sur des approximations subjectives.

Quant à l'avenir, on peut affirmer sans risque de se tromper que les travaux terminographiques conduits en mode automatisé et qui seront de plus en plus nombreux, seront d'une meilleure qualité parce qu'ils participeront d'une analyse et d'une description plus fine des langues de spécialité et qu'ils seront issus d'une mise à plat étendue des différents types de discours spéciaux.

*Pierre Auger,
Département de langues et linguistique,
Université Laval,
Québec.*

Références

- [1] Auger (P.), L'Homme (M.-C.) et Drouin (P.), 1991: « Automatisation des procédures de travail en terminographie », dans *Méta*, vol. 36 (1), p. 121-128.
- [2] Blanchon (E.), 1991: *Évaluation de logiciels terminologiques*, Rapport pour le ministère de la Recherche et de l'Espace, Paris, CTN (CNRS, URA 1576).
- [3] TKE'93, 1993: *Terminology and Knowledge Engineering: proceedings Third International Congress on Terminology and Knowledge Engineering* (Cologne, RFA, 25 - 27 august 1993), Infoterm, Frankfurt/M, Indeks Verlag, 472 p.

(6) Il existe aujourd'hui autour des groupes importants de terminologues des structures de réflexion portant sur l'automatisation de la chaîne de travail en terminographie et ses outils. On pourrait citer ici, entre autres, l'Association pour la terminologie et le transfert des connaissances (GTW) et son groupe de travail pour l'élaboration d'une méthodologie de l'évaluation de logiciels de terminologie et la Réunion interorganisations sur la terminologie et la traduction assistée par ordinateur (Jiamcatt) qui se préoccupe de ces questions. Il faut citer également les importants colloques TKE'92 et TKE'93 organisés sous l'égide d'Infoterm sous le titre *Terminology and Knowledge Engineering* qui portent sur les derniers développements en terminotique, particulièrement l'approche cognitive.

Bibliographie

- Auger (Pierre), 1988: «Le travail du terminologue amélioré et simplifié», dans *Circuit*, numéro consacré aux industries de la langue, n° 22, p. 12-13.
- Auger (Pierre), 1989: «Informatique et terminologie: revue des technologies nouvelles», dans *Méta*, vol. 34 (3), Montréal, p. 450-456.
- Auger (Pierre), 1989: «La terminotique et les industries de la langue», dans *Méta*, vol. 34 (3), Montréal, p. 450-456.
- Auger (P.), L'Homme (M.-C.) et Drouin (P.), 1991: «Automatisation des procédures de travail en terminographie», dans *Méta*, vol. 36, n° 1, p. 121-127.
- Auger (Pierre), 1991: «Terminographie et lexicographie assistée par ordinateur: État de la question et prospective», *Les industries de la langue, Perspectives des années 1990*, Montréal, p. 669-680.
- Auger (Pierre), 1992: «La terminotique ou la terminologie assistée par ordinateur: les courants de la recherche au Québec», dans *Actes du symposium international «Terminologie et documentation dans la communication spécialisée»*, Secrétariat d'État du Canada - Infoterm, Ottawa, p. 276-287.
- Auger (Pierre), 1993: «La terminologie selon une approche textuelle: une représentation plus adéquate des vocabulaires de spécialité», dans *Actes du XV^e Congrès international des linguistes*, vol. I, Québec, P.U.L., p. 1990-1993.
- Butler (Christopher) *et al.*, 1992: *Computers and written texts*, Oxford (UK) et Cambridge (USA), Basil Blackwell Ltd., 306 p. (Applied Language Studies).
- Schaetzen (Caroline de), 1987: «S.g.b.d. et terminologie», dans *Le linguiste (De Taalkundige)*, vol. 33, n° 3.
- Schaetzen (Caroline de), 1987: «Terminologie en langues africaines et gestionnaires de données textuelles», dans *Le langage et l'homme*, vol. 22, fasc. 2, p. 172-174.
- Schaetzen (Caroline de) et Meert (O.), 1987: «Un outil d'aide à la création et à la gestion de bases de données terminologiques pour les langues africaines», dans *Le langage et l'homme*, vol. 22, fasc. 2, p. 157-165.
- Lancashire (Ian), 1991: *The Humanities Computing Yearbook 1989-1990 (A Comprehensive Guide to Software and Resources)*, Clarendon Press, Oxford, 701 p.
- Office de la langue française, 1988: «Terminologie & Informatique», dans *Terminogramme*, Québec, n° 46, 32 p.
- Perron (Jean), 1989: «Termino: un système de dépouillement terminologique», dans *Terminogramme*, n° 54, p. 3-9.
- Secrétariat d'État du Canada, 1991: *L'Actualité Terminologique (terminotique / terminotics)*, Ottawa, vol. 24 (2), 27 p.
- TKE'93, 1993: *Terminology and Knowledge Engineering: proceedings Third International Congress on Terminology and Knowledge Engineering* (Cologne, RFA, 25 - 27 august 1993), Infoterm, Frankfurt/M, Indeks Verlag, 472 p.