



Rint
Réseau
international
de néologie
et de terminologie

20

Nouveaux outils pour la néologie

Afrique centrale
et de l'Est
Afrique de l'Ouest
Canada
Communauté
française de
Belgique
France
Haïti
Madagascar
Maroc
Québec
République
centrafricaine
Suisse
Tunisie
Union latine



Revue semestrielle
coéditée par l'Agence
de la francophonie
et la Communauté
française de Belgique

N° 20
décembre 1999

Terminologies nouvelles

Introduction

Nouveaux outils pour la néologie
*par Louis-Jean Rousseau
et Loïc Depecker*
Page 2

Contributions

Recherches néologiques sur Internet
par Céline Tucac et Loïc Depecker
Page 4

Cenit: Système de détection semi-automatique des néologismes
par Sorcha Roche et Lynne Bowker
Page 12

Utilisation d'une banque de textes en terminographie
par Anouk Jaccarini
Page 17

Recherche terminographique semi-automatisée en veille terminologique: expérimentation dans le domaine médical
*par Marie-Claude L'Homme,
Claudine Bodson
et Renata Stela Valente*
Page 25

Néologie traductive
*par Adrien Hermans
et Andrée Vansteelandt*
Page 37

Néologie et terminologie: activités et réflexions
par Danielle Candel
Page 44

La conversion et l'interface infinitif/impératif/nom dans la terminologie du micro-ordinateur
par Emmanuel Aito
Page 54

Néologismes, internationalismes et mondialisation
par Alessio Petralli
Page 60

En bref

Nouvelles du Rint
Page 72

Publications
Page 75

Sommaire

Nouveaux outils pour la néologie

Présentation

La société de l'information offre un monde en expansion constante caractérisé par la généralisation de l'usage des NTIC⁽¹⁾. Cette révolution de l'information bouleverse les contenus linguistiques en mettant à la disposition des utilisateurs un volume de données textuelles qui semble sans limite, ouvrant ainsi un immense champ lexicographique et terminographique, grâce aux réseaux de type Internet. Le volume de l'information rendue disponible simultanément fait peur à plus d'un, la question majeure étant : comment se retrouver dans ces flux d'informations ?

Compte tenu du potentiel extraordinaire qu'offrent ces trésors d'information et de connaissances dans tous les domaines du savoir, le Réseau international de néologie et de terminologie (Rint) a décidé de faire le point sur leur exploitation dans les travaux sur la néologie, en regard des nouveaux outils à la disposition des utilisateurs, qu'ils soient spécialistes à la recherche de l'innovation, terminologues qui assurent la collecte et le traitement des néologismes ou simples utilisateurs de terminologie. Le Rint veut ainsi tenter de répondre à de nouvelles questions théoriques et

pratiques sur la néologie et aider les terminologues à se retrouver dans le foisonnement des possibilités et des outils ainsi offerts. Il a pour cela mobilisé les chercheurs et spécialistes susceptibles d'intervenir dans le domaine de la veille néologique afin de dégager des orientations méthodologiques sur ces questions dont les plus fréquentes sont : comment retrouver un néologisme, dans quelle(s) langue(s), par quel moyen, grâce à quel instrument, etc. Ce sont là des problèmes rencontrés par toute personne qui recherche le sens d'un néologisme, une attestation fiable, un contexte pertinent, un équivalent possible, ou encore des données conceptuelles susceptibles de contribuer à la définition, etc. De multiples points de vue sont évoqués ici, de la recherche proprement dite d'une désignation néologique ou d'un nouveau concept, à la gestion des ressources linguistiques de nature néologique.

Le Rint a, comme son nom l'indique, vocation à cela. Créé en 1986 par le premier sommet des chefs d'états et de gouvernements francophones, le Réseau international de néologie et de terminologie francophone se compose de spécialistes de la néologie et de la terminologie appartenant à des organismes à vocation linguistique. Depuis sa création, cette organisation n'a cessé de se pencher sur ces

(1) Nouvelles technologies de l'information et de la communication.

Présentation

questions. Le Rint a contribué à créer une réflexion en néologie sur le français, sur les langues partenaires de la francophonie et sur leurs rapports mutuels. Il s'est efforcé de développer des méthodologies dans ce sens, de mobiliser au fil de nombreuses rencontres scientifiques, et de susciter l'élaboration de plusieurs outils pour la veille néologique. Parmi les outils les plus représentatifs, il est possible de citer Adepto-Nomino⁽²⁾, qui est un logiciel d'extraction terminologique. Plus directement liée à la néologie est la banque de données d'attestations de néologismes, Balnéo, créée par le Rint pour favoriser l'échange et la mise en commun de matériaux néologiques et dont le système est implanté à l'Université de Rennes. Balnéo est une base de concentration alimentée par des relevés de néologismes effectués aux quatre coins de la planète. Les néologismes sont intégrés à distance par Internet. Cette base de données est mise gratuitement⁽³⁾ à la disposition de tous ceux qui souhaitent l'alimenter ou en utiliser le contenu, notamment pour la mise à jour des dictionnaires et des banques de terminologie.

L'une des priorités qui se dessinent, et que l'on découvrira dans ce numéro de *Terminologies nouvelles*, est de pouvoir accéder aux nombreux textes et contextes diffusés sur Internet. Comme l'a mis en évidence le numéro 19 de *Terminologies nouvelles*, consacré au thème «terminologie et intelligence artificielle», une néologie de corpus se développe actuellement, fondée sur les immenses masses de textes circulant sur les réseaux. Ce retour à une démarche terminologique fondée sur l'analyse des textes, lieu de création, de description et de dénomination des nouveaux concepts, rend nécessaires l'utilisation de nouveaux outils et l'implantation de pratiques renouvelées pour la néologie et la terminologie, car néologie et

terminologie sont très liées sitôt que le travail s'applique à des domaines de pointe. En dépit de ces flux absolument inédits, les moteurs de recherche, dont certains sont décrits ici, autorisent de plus en plus une recherche néologique fiable et rapide. Ils permettent, pour certains, de localiser en quelques secondes un néologisme, d'en cerner le sens grâce au contexte, en fournissant éventuellement, lorsque le document est riche, l'équivalent dans une autre langue. Ce foisonnement textuel sans précédent n'est cependant pas sans risques ou sans écueils pour le terminologue. Comment choisir les textes? Quels sont les textes fiables, tant sur le plan scientifique et technique que sur le plan linguistique? Comment assurer la gestion de ces données textuelles? Voilà quelques-unes de questions qu'il faut résoudre par de nouvelles réflexions méthodologiques.

De plus, la réelle interactivité que procure Internet sous la forme notamment de forums et de listes de diffusion permet d'interroger des personnes de tous les continents en laissant espérer une réponse ou un début de réponse.

Ce sont plusieurs de ces possibilités qui sont explorées dans ce numéro de *Terminologies nouvelles*. Que les contributeurs en soient remerciés, qui ont travaillé pour ce numéro sans l'urgence et la dynamique qu'aurait pu représenter un colloque scientifique. Qu'ils aient répondu à cet appel montre cependant que le Rint a su susciter réflexions et énergies de par le monde.

Mais c'est certainement une diffusion accrue de la langue française sur des réseaux de type Internet qui permettra d'accéder à la néologie du français d'aujourd'hui, qu'il soit d'Europe, d'Amérique, d'Afrique ou d'Asie. Un grand effort dans ce sens reste à faire pour les langues partenaires de la francophonie. Le

millénaire qui commence les verra prendre pied aussi dans la société de l'information pour peu que s'affirme dans ce sens une réelle volonté politique.

*Louis-Jean Rousseau,
Secrétaire général du Réseau
international de néologie
et de terminologie,
Office de la langue française,
Québec.*

*Loïc Depecker,
Président de la Société française
de terminologie,
Université de la Sorbonne nouvelle,
Paris III-Crettal.*

(2) Anciennement: *Termino*

(3) On peut accéder à *Balnéo* à partir du site Internet du Rint, dont l'adresse est la suivante: <http://www.rint.org>.

Recherches néologiques sur Internet

Le but de cet article est de montrer, après cinq années d'utilisation active d'Internet, le profit qu'il est possible de tirer de son utilisation dans le cadre de recherche en traduction, en terminologie / néologie et en rédaction technique. Durant ces années de mise en place sur Internet de ressources informationnelles et linguistiques de par le monde, il est devenu difficile de trouver facilement et rapidement ce que l'on cherche dans l'étonnant foisonnement de l'information qu'est devenu Internet. Cet article a pour but d'aider les utilisateurs à s'orienter plus aisément dans ces dédales en leur indiquant et conseillant les pistes qui nous semblent les plus efficaces à ce jour.

Termes-clés :
néologie, Internet,
moteur de recherche, corpus,
banque de terminologie.

1 Introduction

La question de la néologie acquiert une dimension nouvelle dans le cadre de la société de l'information, en raison de la masse des néologismes et concepts qui circulent et de la difficulté à y accéder. Dans cette perspective, la néologie revêt un nouvel aspect, notamment d'un point de vue informationnel. C'est pour cette raison qu'il est nécessaire d'exploiter les outils existants pour accéder à l'information néologique et éventuellement d'en créer de nouveaux. Cet article s'attachera essentiellement à quelques-unes des possibilités qu'offre Internet aujourd'hui dans le but d'être utile aux personnes qui commencent à naviguer sur Internet.

Nous traiterons de la néologie, particulièrement scientifique et technique, en faisant cette distinction : un néologisme est constitué d'une forme linguistique (désignation) et d'un concept (unité de pensée). La question de la recherche de néologismes peut ainsi se poser au moins de deux façons :

- Pour un concept nouveau, chercher dans une langue son ou ses équivalents linguistiques sans avoir sa désignation ;

- Pour un concept nouveau et à partir d'au moins une désignation dans une langue, chercher dans d'autres langues des équivalents linguistiques.

Dans le premier cas, il s'agit d'un travail de terminologie proprement dit puisque l'on part du concept et que l'on va vers le signe linguistique dans une langue. Dans le second, il s'agit surtout d'un travail de néologie orienté vers la traduction.

2 Les outils de recherche sur Internet

Traditionnellement, le téléphone et l'annuaire qui l'accompagne constituaient le meilleur des dictionnaires : rien de mieux que d'avoir au bout du fil un spécialiste du domaine. Et sans doute, le téléphone, tout rudimentaire qu'il apparaisse, reste le meilleur des dictionnaires de néologie. Mais on peut tout de même aller plus loin : en effet, Internet et plus particulièrement les listes de diffusion ou les forums de discussion spécialisés représentent aujourd'hui le prolongement du questionnement par simple appel téléphonique. Le tout est de savoir ce que l'on cherche et qui est à l'autre bout.

Contributions

2.1 Comment faire une recherche efficace ?

Il existe aujourd'hui plusieurs dizaines de moteurs de recherche sur Internet. Une même recherche lancée sur ces outils ne donnera pas les mêmes résultats et, d'un jour à l'autre, ceux-ci pourront être différents.

On peut distinguer comme principaux types de moteurs de recherche :

- les *robots d'exploration* qui parcourent les sites serveurs Internet et cumulent des masses d'informations dans des banques de données interrogeables grâce à des moteurs de recherche tels que *Lycos*, *Alta Vista*, etc.
- les *catégoriseurs* tels que *Live Topics* sur *Alta Vista* qui limitent la recherche à l'aide de critères précis ;
- les *catalogues sujets* qui sont des répertoires élaborés par des professionnels qui visitent, évaluent et classent les ressources disponibles sur Internet ;
- les *méta-index* ou *méta-moteurs* qui permettent, à partir d'un seul formulaire, d'effectuer une recherche dans plusieurs index à la fois.

Cependant, même si ces distinctions sont utilisées, les différences entre ces divers outils sont relativement fluides.

La question de la recherche sur Internet se résume le plus souvent à un problème documentaire. Et une recherche est efficace lorsqu'elle procure uniquement les documents souhaités. Les outils de recherche généraux peuvent s'avérer frustrants. En effet, on se fait difficilement une idée claire sur la couverture de la banque dans laquelle la recherche est lancée, ou sur son indexation. De plus, le mode de fonctionnement de la recherche n'est pas toujours clairement présenté. Il convient cependant de se familiariser progressivement avec différents outils afin de mieux juger des chances de

trouver ce que l'on cherche. De même, les possibilités d'interrogation diffèrent d'un outil à l'autre, d'où le besoin, là encore, de se familiariser avec les procédures de recherche.

En ce qui concerne les langues à sélectionner, il faut savoir si le moteur de recherche est francophone ou international : en effet, pour le premier, le mot clé sera en français alors que pour le second, il faudra partir d'une autre langue, généralement l'anglais. En tout état de cause, il est difficile de rechercher un néologisme à partir du mot clé «néologie». Ainsi, une interrogation lancée le 18 mars 1999 sur le moteur de recherche *Voilà* de France Télécom ne proposait pas moins de 698 réponses. Il faut donc interroger à partir de ce que l'on cherche. Mais c'est là aussi le plus difficile puisque la méthode de la recherche varie selon la connaissance que l'on a du concept.

Certains critères sont essentiels pour faire une recherche efficace. Il vaut mieux partir d'un ou plusieurs mots clés relativement précis et non d'un terme trop générique. Il est possible de les mettre entre guillemets, lesquels indiquent au moteur de recherche que les résultats devront comporter ces termes dans l'ordre où ils ont été entrés. On peut aussi utiliser les opérateurs booléens ou logiques, c'est-à-dire AND, OR et AND NOT. AND indique au moteur de recherche que le mot situé derrière devra obligatoirement figurer dans la réponse, OR qu'il pourra y figurer et AND NOT qu'il ne devra en aucun cas y figurer. Ces opérateurs sont généralement traduits en français par ET, OU et SAUF sur les moteurs de recherche francophones tels que *Nomade*, *Yahoo France*, etc.

Par exemple, si la recherche porte sur la manière d'orthographier «Euroland», il faudra entrer dans la boîte de dialogue «euroland*». Ainsi, une recherche lancée sur *Alta Vista* le 10 avril 1999 ne donnait plus que 200 réponses, sachant que la majorité

des sites était dédiée à la parité des monnaies des pays de la zone Euro. Cependant, plusieurs sites permettent de retrouver «Euroland» dans son contexte journalistique et un site était consacré au problème «Euroland ou Eurolande?»

De la même manière, si l'on recherche des informations ou un équivalent français d'un anglicisme néologique encore peu répandu, on peut tenter d'interroger à partir de sa forme même. Par exemple, pour l'anglicisme *job rotation*, recherché sur Internet le 18 avril 1999, 13 réponses apparaissent sur *Voilà*, 8 sur *Alta Vista* et 10 sur *Excite*. L'équivalent français proposé dans plusieurs articles est «rotation des postes». La fréquence de l'unité néologique proposée, récurrente dans divers contextes de référence, donne ainsi un bon critère d'appréciation. Ce genre de critère commence à être utilisé de manière systématique, notamment en France, par les commissions de normalisation terminologique.

Il est en revanche beaucoup plus difficile de chercher une désignation à partir de sa définition supposée. Si l'on connaît la définition mais que l'on ne sait pas à quelle désignation elle correspond, la stratégie de la recherche est différente : en effet, il faudra formuler une recherche à l'aide d'opérateurs booléens et partir des caractères du concept présents dans la définition. Le plus facile, apparemment, est de chercher un néologisme dans un domaine très spécialisé. En gastro-entérologie par exemple, la *pancolite* est une atteinte inflammatoire de la totalité du colon. Les mots clés, c'est-à-dire les caractères essentiels du concept énumérés, seront «inflammation» et «colon». L'algorithme booléen sera le suivant : (colon AND inflammat*) AND colon. Ainsi, une recherche lancée sur *Alta Vista* le 2 juillet 1999 a donné 19 sites spécialisés en gastro-entérologie et plus particulièrement sur les maladies inflammatoires de

l'intestin. Certains de ces sites sont spécialisés dans le domaine de ce type d'inflammation, et dans deux d'entre eux on trouve une définition constituée de la pancolite.

2.2 Description des divers moteurs de recherche

Il y a ici au moins deux étapes à respecter :

– L'étape du formulaire : les moteurs de recherche proposent à l'utilisateur entrant sur un site un ou plusieurs formulaires. Un formulaire est un cadre de discussion (boîte de dialogue) intégré à une page de la toile. Il permet d'indiquer les critères de recherche : dans quelle langue lancer la recherche, sur combien de temps recueillir les résultats, qu'exclure à priori dans la collecte, etc. Des explications souvent très précieuses sont fournies quant à l'utilisation de ces formulaires : conseils pour restreindre la recherche, pour l'emploi des opérateurs booléens, des parenthèses, etc., ce qui permet de structurer et d'affiner la formulation de la question. Cette étape est cruciale pour lancer une recherche. En effet, tous les moteurs insistent sur les degrés de précision de la recherche envisagée et sur la bonne formulation de la question.

– L'étape de la consultation des réponses aux questions sur le site par les utilisateurs. Ce sont les FAQ (Foire Aux Questions ou *Frequently Asked Questions*), qui sont souvent d'un très grand intérêt.

Ces deux étapes essentielles montrent l'intérêt qu'il y a à expérimenter plusieurs moteurs de recherche et à apprendre à connaître leurs qualités et leurs limites. Nous nous en tiendrons ici aux outils qui paraissent les plus performants pour la recherche néologique. Si l'on veut classer les outils en grandes catégories,

on pourrait retenir la répartition suivante :

2.2.1 Les moteurs d'exploration

On peut trouver la liste des moteurs d'exploration en posant la question « outils de recherche » à un moteur de recherche francophone ou « Search engine » à un moteur anglophone. Il est également possible d'aller directement sur le site appelé « Toile du Québec » où les moteurs les plus performants sont répertoriés. Ces outils explorent des banques constituées automatiquement par des robots informatiques visitant les sites serveurs W3. Les fréquences de mise à jour varient selon les outils, c'est pourquoi un grand nombre d'adresses URL s'avèrent périmées dans les résultats des recherches. De plus, même si le bruit y est relativement élevé, le volume des données signalées est très important. Parmi les moteurs de recherche les plus immédiatement utiles pour des recherches d'ordre général, on peut citer :

• *Alta Vista* (<http://www.altavista.com>)

Ce site est l'un des plus performants, tant pour son vaste contenu que pour sa procédure de recherche. Il offre la possibilité de rechercher dans deux banques distinctes : « The Web » pour les pages des sites W3 et « Usenet » pour les forums de discussion. Il offre deux niveaux de recherche : « simple query » et « advanced query », toutes deux documentées par des pages d'aide. Ce deuxième type de recherche est destiné à des recherches spécifiques qui nécessitent l'utilisation des opérateurs booléens : ainsi, dans la boîte de dialogue « *Search* », on entrera le concept à rechercher, et dans la boîte « *Boolean Expression* », on indiquera les termes auxquels la recherche doit se limiter. On cerne ainsi la désignation à rechercher et la définition correspondante.

Alta Vista indexe le contenu entier des documents mais le bruit y

reste élevé. C'est pourquoi le catégoriseur *Live Topics*, disponible uniquement sur le site américain d'*Alta Vista*, peut être très utile : il effectue en effet une analyse statistique des mots contenus dans les documents sélectionnés par la requête et dresse la liste des vingt termes les plus représentés. Afin de restreindre les résultats, il suffit d'indiquer « require » ou « exclude » en face de chaque terme.

• *HotBot* (<http://www.HotBot.com>)

Avec 10 millions de pages W3 indexées par jour, ce moteur de recherche américain se présente comme « le robot le plus rapide du monde ». L'interface d'utilisation est particulièrement bien développée et dotée d'une fonction d'aide exceptionnelle. Il propose plusieurs modes de recherche et permet l'utilisation d'opérateurs booléens et de parenthèses. Il est également possible de limiter la recherche en ajoutant des critères secondaires, procédure similaire à *Alta Vista* (*Advanced Query*).

• *Infoseek* (<http://www2.infoseek.com>)

Il a désormais deux modes : « *ultrasmart* » et « *ultraseek* ». Le premier est un formulaire d'interrogation couplé à un répertoire. Le second n'inclut que le formulaire. Les traditionnels opérateurs booléens AND et AND NOT sont ici remplacés par les signes « + » et « - » : en précédant le mot d'un « + », on obtient le ET, le « - » donne le SAUF. Par exemple, si on entre le terme « culture », le document peut contenir le mot « culture » ; avec « +culture », le document doit contenir ce mot ; avec « -culture », le document ne doit pas contenir ce terme.

• *Lycos* (<http://www.fr.lycos.de>)

Comme la plupart des moteurs, *Lycos* affiche un formulaire simple dans lequel indiquer l'élément à rechercher. Pour une recherche avancée, il suffit de cliquer sur

«Search Options» mais la recherche booléenne est minimale et peu documentée. La réponse peut se limiter à un maximum de 40 documents.

Enfin, se crée de plus en plus des sites thématiques: on peut, pour le monde des affaires, citer:

- *Open Text's Livelink Pinstripe* (<http://pinstripe.opentext.com>)

Cet outil est mis très régulièrement à jour et met l'accent sur l'indexation du texte intégral des pages W3. La recherche booléenne est efficace, avec deux niveaux de recherche: «*Simple Search*» et «*Power Search*» pour des recherches booléennes plus complexes avec possibilité de spécifier pour chaque terme dans quel champ effectuer la recherche. Chaque mode de recherche est explicité dans des pages d'aide.

2.2.2 Les catalogues sujets pourvus d'un moteur de recherche

Ce sont des catalogues dans lesquels les ressources Internet sont classées par sujets dont certains d'entre eux sont interrogeables à l'aide d'un formulaire. Il s'agit de répertoires produits par des professionnels qui visitent, évaluent et classent les ressources de leurs domaines de compétence. Le contenu de leurs bases est moins considérable que celui amassé par un robot, mais leurs bases sont plus cohérentes et posent moins de problèmes de bruit au moment de la recherche. En effet, les informations sont organisées hiérarchiquement, de la plus pertinente à la moins pertinente.

- *Yahoo!*

Yahoo est un excellent catalogue-sujet doté d'une interface de recherche performante. Il faut savoir que ce moteur est surtout efficace si l'on recherche une information générale sur un sujet.

Yahoo est subdivisé selon les langues:

- le français: <http://www.yahoo.fr>
- l'anglais: <http://www.yahoo.com>
- l'allemand: <http://www.yahoo.de>
- l'espagnol: <http://www.yahoo.es>
- l'italien: <http://www.yahoo.it>

– ainsi que plusieurs autres *Yahoo* spécialisés par langue ou région, proposés en bas de la page d'accueil. La recherche s'effectuant par sujets, les sites trouvés sont généralement liés à la demande formulée. Si les résultats ne conviennent pas, il est possible d'étendre la recherche sur *Alta Vista* en cliquant sur le lien *Alta Vista* au bas de la page.

- *La Toile du Québec* (<http://www.toile.qc.ca/>)

Axé sur les sites québécois ou francophones, ce site propose des moteurs de recherche spécialisés. Il est efficace pour la néologie du français dans des domaines bien précis. Ces moteurs sont en réalité des catalogues multimoteurs, c'est-à-dire des moteurs de recherche qui lancent sur différents moteurs la question posée: il s'agit de *All-in-One Search*, *BotsSpot*, *CUSI at Internet Direct* et *NetSearch*. Par exemple, *NetSearch* propose d'utiliser *Excite*, *Yahoo*, *Lycos*, *Infoseek*, etc. pour effectuer la même recherche.

- *Excite* (<http://www.excite.com>)

Ce moteur de recherche francophone et international effectue des recherches non seulement par mots clés mais aussi par concepts. Dans ce cas, *Excite* recherche non seulement des sites qui correspondent exactement à la question posée mais aussi des sites qui s'y rattachent. Par exemple, si l'on recherche des informations sur les fonds de pension des retraités, *Excite* indiquera également, dans ses résultats, les sites traitant de la condition économique des personnes âgées. Dans le cas d'une recherche par mot clé, *Excite* propose d'ajouter des termes se rattachant au mot indiqué dans le formulaire: si l'on tape «terminologie», *Excite* propose, pour affiner la recherche, d'ajouter des

termes comme «traduction, termes», etc.

2.2.3 Les méta-index

Quelques outils permettent d'effectuer une recherche dans plusieurs index à la fois à partir d'un seul formulaire. Ainsi, une même requête peut être envoyée simultanément à plusieurs banques de données et le méta-index fera la collection des résultats. De par leur nature, ces systèmes peuvent s'avérer plutôt lents et ne tirent guère profit des possibilités particulières de certains outils de recherche. Ils permettent cependant d'avoir des résultats plus variés que si la recherche était lancée sur un seul moteur.

- *SavvySearch* (<http://www.savvysearch.com>)

Site franco-anglais qui peut accéder à près d'une trentaine de banques de données, il rassemble les résultats des différents moteurs de recherche en indiquant ceux qu'il juge les plus pertinents. De plus, il synthétise l'information en supprimant toutes les données redondantes. Grâce à ce méta-index, en tapant «néologie», on peut facilement et rapidement avoir accès, par exemple, aux nouveaux termes publiés au *Journal Officiel de la République Française* en allant sur le site de la Commission générale de terminologie (accessible par ailleurs sur <http://www.culture.fr>).

- *MétaCrawler* (<http://www.go2net.com/search.html>)

Aussi intéressant que le précédant, ce site de recherche anglophone se limite aux neuf outils de recherche suivants: *Excite*, *WebCrawler*, *Lycos*, *Alta Vista*, *Infoseek*, *Thunderstone*, *Yahoo*, *LookSmart* et *The Mining Co*. Le système possède une option Java permettant de contrôler l'exécution

de la recherche, ainsi qu'un mode de recherche rapide.

- *Starting Point* (<http://www.stpt.com>)
C'est un «semi-méta-index» combinant un répertoire et un formulaire d'interrogation. Il permet l'utilisation d'opérateurs logiques. Une fois les mots clés fournis au système, le système affiche une liste de catégories et d'outils de recherche. Il suffit de sélectionner l'un d'eux, les mots clés étant automatiquement transférés.

3 Les sites utiles à la néologie

3.1 Les sites de ressources en terminologie / néologie

En entrant les mots clés «glossaire, lexique, terminologie, néologie», etc. il est possible d'accéder à de nombreux documents spécialisés sur un thème, dictionnaires, glossaires, lexiques, aussi bien en français que dans d'autres langues. Les sites sont classés ici en fonction de la masse actuelle des données de chacun d'entre eux.

- *Termisti* (<http://www.refer.org/termisti>)
Ce site, créé par l'Institut supérieur des traducteurs et interprètes de Bruxelles, est actuellement l'un des plus intéressants sur les questions de terminologie et de néologie. Il propose la consultation de glossaires multilingues réalisés par les étudiants en quatrième année à l'Institut. Pour consulter la fiche du terme recherché, il suffit de cliquer sur celui-ci dans le lexique de la langue voulue. Ce site offre également une liste très intéressante de ressources terminologiques interactives tels que les bases de données consultables, les réseaux, centres et organismes de terminologie, les listes de diffusion et forums de discussions, ainsi que des

cours sur la terminologie et la rédaction spécialisée.

- *Etis* (<http://www.unilat.org/dtil/etis/index1.htm>)

Ce serveur européen de terminologie est alimenté par les institutions internationales les plus actives en matière de terminologie. Ce site permet, grâce à de nombreux liens, la consultation d'un grand nombre de ressources aussi bien en terminologie qu'en traduction. En effet, il propose des bases de données classés par pays, des dictionnaires en ligne et des sites de terminologie qui présentent des dictionnaires, lexiques et glossaires, etc.

- *Le site de l'OMS* (http://who.int/pll/cat/tao_ressources.html)
Ce site, et plus particulièrement son unité de traduction assistée par ordinateur, propose la consultation de deux glossaires sur les maladies et le paludisme. De plus, il offre un service de traduction des termes de la santé et propose des dictionnaires multilingues sur l'anatomie, le cancer, le sida, etc. Enfin, le service de terminologie technique de l'OMS propose une liste de ressources terminologiques de la santé disponibles sur Internet en français, anglais et espagnol. Plusieurs liens permettent d'accéder à des glossaires électroniques, des bases de données terminologiques et aux documentations en ligne de l'ONU et de l'Union européenne.

- *Le CRTT* (Centre de Recherches en Terminologie et Traduction de l'Université de Lyon II) (http://www.univ-lyon2.fr/langues/LEA/LEA_RCH2.htm)

Le CRTT propose des liens précieux vers des sites traitant de lexicologie, de terminologie et de traduction. Il informe également sur les nouveautés en terminologie, les colloques organisés par le CRTT et propose une liste des divers organismes de normalisation francophones et internationaux.

- *Le Crettal* (<http://www.tele3.net/terminologie/memoires.htm>)

Y est particulièrement intéressante la consultation de mémoires de terminologie réalisés par des étudiants de maîtrise LEA, option «traduction spécialisée», dans le cadre de l'enseignement de terminologie de Paris III. Ces glossaires se présentent sous la forme de fiches accessibles en cliquant sur le terme du lexique. Ce site offre aussi la possibilité de consulter les cours sur la terminologie, des dictionnaires multilingues comme celui de l'informatique, ainsi que les cours de traduction diffusés sur la chaîne télévisée française *La Cinquième*.

- *Le Lilla* (Laboratoire d'ingénierie linguistique et de linguistique appliquée de l'Université de Nice Sophia Antipolis) (<http://www.lilla2.unice.fr>)

C'est un site très original qui propose la création en ligne, de façon automatique, de néologismes correspondant à la morphologie des langues latines. Le néologue de l'an 2000 peut ainsi créer des néologismes en partant d'un préfixe, d'un suffixe ou d'une base morphologique. Le site donne aussi la définition des néologismes qu'il a recensés.

2.2 Les banques de terminologie en ligne

- Créé par le Rint (Réseau international de néologie et terminologie francophone : <http://www.rint.org>), *Balnéo* a été la première base de données d'attestations néologiques sur Internet. Elle a pour but de collecter, échanger et diffuser les néologismes à partir de relevés effectués par des équipes de recherche francophones ou travaillant sur le français.

Balnéo offre deux types de consultation : la consultation simple ou la consultation experte. Avec la première, on peut, ou consulter la

liste de tous les néologismes répertoriés dans la base de données d'attestations néologiques, ou faire une recherche à partir d'un terme préexistant. Ainsi, pour le terme «acide», *acide** indique que l'on recherche tous les néologismes qui commencent par «acide»; **acide**, que l'on recherche ceux qui contiennent ce terme; enfin, **acide*, que l'on désire consulter ceux qui se terminent par ce mot.

- *Termium* (<http://termium.pwgsc.ca/>)

La banque de données bilingue français-anglais du Gouvernement du Canada a beaucoup évolué ces derniers temps. Elle s'étend sur de nouveaux domaines et depuis peu vers d'autres langues, particulièrement l'espagnol, en raison de l'importance que prend cette langue dans la construction du grand marché américain.

- La *Base de données terminologiques* du Cilf (Centre international de la langue française : <http://www.cilf.org/bt.fr.html>)

Le Cilf propose une banque de terminologie en ligne. À partir du terme recherché, de la langue et du dictionnaire spécialisé dans lesquels on souhaite effectuer la recherche, on accède à une liste de mots; en cliquant sur l'un d'entre eux, on accède à sa fiche qui donne ses équivalents en français, anglais, allemand et espagnol, son domaine d'application et sa définition.

- *Euterpe* (Parlement Européen) (http://germany.trados.com:4712/MT_W_LOGON) et *Eurodicautom* (Commission Européenne) (<http://www2.echo.lu/eurod/>)

Constitué au sein des institutions de l'Union européenne, ces bases proposent des dictionnaires interactifs multilingues qui donnent la définition du terme et sa traduction dans la langue cible choisie au début de la requête. Ce sont des bases très riches pour toute terminologie

entrant dans les textes réglementaires de l'Union européenne.

- *Iloterm* (<http://ilis.ilo.org/ilis/ilisterm/ilintrte.html>)

Site de l'Organisation internationale du travail, *Iloterm* est une base de données terminologiques quadrilingue gérée par l'unité de terminologie et de références du service des documents officiels. Elle traite des domaines du travail et des questions sociales; les termes y figurent en anglais avec leurs équivalents français, espagnol et /ou allemand. De plus, moyennant un abonnement, on peut avoir accès au thesaurus traitant de la terminologie du travail, de l'emploi et de la formation.

- *Termite* (http://www.itu.int/ITU_Databases/Termite/index.html)

Cette base de données de l'Union internationale des télécommunications recense les termes figurant dans les glossaires publiés par cet organisme depuis 1980 et les termes plus récents ayant trait aux diverses activités de l'Union internationale des télécommunications.

- *Le Grand dictionnaire terminologique* (http://www.lgdt.cedrom_sni.qc.ca/)

Réalisé par le Gouvernement du Québec, le Grand dictionnaire terminologique est disponible sur *cédérom* et depuis peu sur Internet moyennant un abonnement. Il est composé d'une banque de terminologie de plus de trois millions de termes techniques et d'une banque documentaire sur les travaux récents effectués en terminologie et néologie.

Parmi les nouveaux à prévoir, figure le site Internet de l'Institut national de la langue française (Inalf), qui devrait être opérationnel en l'an 2000. On pourra y consulter des bases de textes, des dictionnaires et la base d'attestations néologiques *Bornéo*.

2.3 Les organismes de terminologie

- Le *Rint*, créé en 1986 lors du premier sommet de la Francophonie, regroupe des organismes francophones à vocation terminologique et a pour vocation, entre autres, d'encourager la création terminologique en langue française, de diffuser les travaux de néologie, et de répertorier les terminologies nouvelles.

Le site du *Rint* propose la consultation de sa revue *Terminologies nouvelles*, de *Balnéo*, une liste de publications et un inventaire de ressources terminologiques.

- Le site de la *DGLF* (Délégation générale de la langue française : <http://www.culture.fr/culture/dglf>) est intéressant car il propose une compilation des messages reçus sur différentes listes de diffusion. Il s'agit des propositions faites par des professionnels abonnés à ces listes pour traduire des termes étrangers ou désigner en français des notions d'apparition récente. Ces documents ont une valeur d'information et non de recommandation officielle.

- *Realiter* (*Réseau panlatin de terminologie*) (<http://www.iula.upf.es/cpt/cptfr.htm>), rassemble des personnes, institutions et organismes de pays de langue latine impliqués dans des travaux de terminologie et de néologie. Outre des informations permanentes et les travaux en cours, ce site propose des outils de travail en néologie et terminologie dans ces langues. L'un des buts est de créer sur des formants identiques des néologismes communs dans les langues latines.

- *Riterm* (<http://www.iula.upf.es/riterm/ritermes.htm>), réseau ibéro-américain de terminologie, propose un site en espagnol et portugais. Ses objectifs consistent à développer les activités de terminologie au niveau ibéro-américain, établir un réseau de

coopération entre les différents agents terminologiques pour échanger les informations nécessaires afin de renforcer les terminologies espagnoles et portugaises.

4 Les sites d'aide à la traduction

4.1 Les dictionnaires

Nous ne détaillerons pas ici le contenu de tous les dictionnaires auxquels il est possible d'accéder mais nous proposerons plusieurs adresses intéressantes pour le traducteur.

Citons tout d'abord

<http://www.facstaff.bucknell.edu/rbeard/diction.html> qui propose l'accès à un nombre impressionnant de dictionnaires en ligne en différentes langues. On peut ainsi trouver, en français, des dictionnaires généraux mais aussi des dictionnaires plus ou moins spécialisés tels qu'un glossaire informatique des termes de la commission ministérielle de terminologie informatique, des microglossaires multilingues, un dictionnaire du cheval, etc. Il est également possible d'avoir directement accès aux dictionnaires français grâce à l'adresse suivante : <http://www.facstaff.bucknell.edu/rbeard/diction3.html#french>

Très utiles également sont les «signets des traducteurs» (<http://www.wo.int/pll/cat/TranslatorBookmark.html>), qui proposent l'accès à des banques de données anglophones et francophones, à des ressources terminologiques, à des dictionnaires généraux et spécialisés, unilingues ou multilingues, à des lexiques, des glossaires et qui permettent la consultation de plusieurs bibliothèques européennes.

L'adresse suivante :

<http://fp3.com/ljm> («Le juste mot; glossaires et lexiques») présente une liste de glossaires et lexiques français

abordant tous les domaines classés par thèmes principaux puis en sous-domaines et triés alphabétiquement. Il est possible d'effectuer une recherche à partir d'un mot clé ou d'un domaine. Il existe aussi des liens qui renvoient vers des dictionnaires pour traducteurs.

Il faut également signaler que *Termisti*, évoqué plus haut, propose une liste de dictionnaires utiles à la traduction. Il est aussi possible de rechercher soi-même le dictionnaire grâce aux divers moteurs de recherche tels que *Yahoo* ou *Alta Vista*, mais en prenant soin d'indiquer les mots clés. La recherche booléenne s'avère alors la plus apte à produire les résultats désirés.

4.2 Les listes de diffusion et les forums de discussion

Les listes de diffusion permettent à un groupe de personnes de communiquer sur un thème donné par l'intermédiaire du courrier électronique. Les messages sont distribués à tous les participants par un automate. Elles sont donc utiles à la terminologie, à la néologie et à la traduction, étant donné qu'il est possible d'avoir non seulement une réponse à la question posée, mais aussi l'avis d'un spécialiste voire d'une communauté de spécialistes.

Prenons l'exemple de *Termilat* (termilat@francenet.fr), liste électronique de diffusion sur la terminologie et les industries des langues en langues néolatines. Y circulent des demandes de traduction en langues étrangères, des informations, des offres de collaboration dans les travaux terminologiques. Cette liste est soigneusement contrôlée afin que seuls les messages pertinents soient envoyés aux usagers de la liste. Elle est en expansion constante et crée une dynamique extraordinaire entre ses différents acteurs.

Citons aussi la liste de diffusion *TLS* (<http://www.uhb.fr/Langues/Craie/tls/tls.html>) destinée à diffuser des informations se rapportant à la terminologie et qui met en œuvre un système de questions-réponses sur des problèmes terminologiques et néologiques. Orientée vers la traduction vive et spécialisée, cette liste permet de répondre dans l'urgence à des questions portant sur un terme ou sur son équivalent dans une langue.

Les forums de discussion ont les mêmes objectifs que les listes de diffusion : des spécialistes répondent à une question donnée. Cependant, ce système peut s'avérer plus rapide étant donné que la discussion peut se faire en ligne plutôt que par envoi de courriel (*e-mail*). Parmi les forums de terminologie, citons Logos (<http://www.logos.it/forumgen.html>) et *sci.lang* (<http://news.sci.lang>). Ces forums de discussion sont précieux car ils permettent aux spécialistes d'un domaine de débattre d'un point. Ils offrent également l'occasion rêvée par un traducteur de s'établir un impressionnant réseau d'informateurs. Il est donc évident qu'ils peuvent aider, comme les listes de diffusion, à la néologie parce qu'ils permettent d'obtenir une réponse à la question posée et non à une autre, et de recueillir l'avis de spécialistes, voire d'une communauté large de spécialistes selon un débat contradictoire.

5 Conclusion

Le néologue chercheur ou le néologue en herbe dispose ainsi de plusieurs grands types d'outils performants pour la néologie :
– Des répertoires de néologismes en ligne inclus le plus souvent dans des dictionnaires, lexiques et banques de données ;

– Des moteurs de recherches qui orientent vers et dans des données textuelles;

– Des extracteurs de termes éventuellement néologiques, du type *Termino*;

– Des moteurs de création de néologismes du type de celui présenté par le Lilla.

Il faudra encore attendre quelque temps pour que les systèmes détectent automatiquement les néologismes récents. Mais l'idée n'est pas si vaine: il suffirait de considérer que tous les textes existants sur les réseaux forment un immense corpus d'exclusion.

Petite internographie commentée:

Les trois moteurs de recherche les plus performants pour la recherche en terminologie/néologie sont à notre avis:

- *Alta Vista*;
- *Excite*;
- *Hot Bot*.

Les trois sites de ressources en terminologie/néologie les plus importants semblent être:

- Etis;
- OMS;
- Termisti.

Les trois banques de données les plus intéressantes sont:

- *Balnéo*;
- *Le Grand dictionnaire terminologique du Québec*;
- *Termium*.

Une adresse à retenir:

<http://ahost4u.com/zak/bases/htm>:

- Toutes les principales banques de données auxquelles il est possible d'accéder par des liens;
- Liste de sites Internet spécialisés dans la traduction, outils de recherche pour la traduction, dictionnaires et sites spécialisés dans le recrutement des traducteurs.

(1) Les auteurs remercient M. Gilles Bacouël, traducteur à la société Motus (Paris), pour sa contribution à cet article.

Liste récapitulative des adresses des moteurs de recherche Internet:

Voilà	http://www.voilà.fr
Alta Vista	http://www.altavista.com
Excite	http://www.excite.com
Hot Bot	http://www.HotBot.com
Infoseek	http://www2.infoseek.com
Lycos	http://www.fr.lycos.de
Open text's Livelink Pinstripe	http://pinstripe.opentext.com
Yahoo France	http://www.yahoo.fr
Yahoo Angleterre	http://www.yahoo.com
Yahoo Allemagne	http://www.yahoo.de
Yahoo Espagne	http://www.yahoo.es
Yahoo Italie	http://www.yahoo.it
La toile du Québec	http://www.toile.qc.ca/
SavvySearch	http://www.savvysearch.com
MétaCrawler	http://www.go2net.com/search.html
Starting Point	http://www.stpt.com

Liste récapitulative des sites utiles à la néologie:

Termisti	http://www.refer.org/termisti
Etis	http://www.unilat.org/dtil/etis/index1.htm
OMS	http://who.int/pll/cat/tao_ressources.html
CRTT	http://www.univ-lyon2.fr/langues/LEA/LEA_RCH2.htm
Crettal	http://www.tele3.net/terminologie/memoires.htm
Lilla	
Rint	http://www.rint.org
Termium	http://termium1.pwgsc.ca/
Cilf	http://www.cilf.org/bt.fr.html
Euterpe	http://germany.trados.com:4712/MTW_LOGON
Iloterm	http://ilis.ilo.org/ilis/ilisterm/ilintrte.html
Termite	http://www.itu.int/ITU_Databases/Termite/index.html

Le Grand Dictionnaire terminologique

Site de la DGLF	http://www.lgdt.cedrom_sni.qc.ca/
Realiter	http://www.culture.fr/culture/dglf
Riterm	http://www.iula.upf.es/cpt/cptfr.htm
	http://www.iula.upf.es/riterm/ritermes.htm

Liste des sites d'aide à la traduction:

Site de dictionnaires	http://www.facstaff.bucknell.edu/rbeard/diction.html
Signets des traducteurs	http://www.wo.int/pll/cat/TranslatorBookmark.html
Le juste mot	http://fp3.com/ljm

Listes de diffusion et forums de discussion:

Termilat	http://termilat@francenet.fr
TLS	http://www.uhb.fr/Langues/Craie/tls/tls.html
Logos	http://www.logos.it/gorumgen.html
sci.lang	http://news:sci.lang

*Céline Tucac et Loïc Depecker,
Union latine-Université de la Sorbonne nouvelle,
Paris III-CRETTAL (1).*

Cenit: Système de détection semi-automatique des néologismes

Les langues évoluent à un rythme accéléré, donc la détection manuelle des néologismes devient encore moins pratique. Cet article décrit *Cenit*, un système de veille néologique qui vise à identifier des néologismes dans les domaines spécialisés. Le système utilise une stratégie fondée sur l'exploitation de corpus. Nous exposons les composants et les divers stades de traitement.

Termes clés:
néologismes; système de détection des néologismes; corpus; filtres.

1 Introduction

D'après Le Meur et Depecker (1995: 48): «Les langues évoluent à la fin du XX^e siècle de façon prodigieuse, sous la poussée notamment des nouvelles techniques et de la révolution informationnelle en cours». Pour le terminologue au travail, il n'est plus raisonnable de travailler sans outils, donc l'application des ordinateurs aux divers aspects du travail terminologique continue à gagner en popularité. Depuis longtemps, les ordinateurs sont employés pour le stockage et l'organisation des données. Plus récemment, les techniques d'intelligence artificielle ont été adoptées en terminologie pour créer une nouvelle génération de ressources terminologiques qu'on appelle base de connaissances terminologiques (BCT) (Miller *et al.* 1991; Condamines 1995). Les BCT permettent non seulement le stockage des données linguistiques, mais de plus ils fournissent au terminologue la possibilité de saisir et de modéliser les données notionnelles et les arborescences de domaine.

En ce qui concerne les applications plus automatisées, des chercheurs ont examiné la possibilité de l'extraction automatique de termes, et des approches et des projets différents sont récapitulés dans Kageura et Umino (1996) ainsi que dans Bourigault *et al.* (1998). De même, plusieurs chercheurs

commencent à étudier la possibilité d'informatiser la détection des néologismes, travail qui prend du temps et qui nécessite une main-d'œuvre importante si on le fait manuellement. Il existe déjà quelques systèmes créés dans le but de détecter des néologismes. Le système *Aviator* (Blackwell 1993; Collier 1993; Renouf 1993) identifie les néologismes dans la presse anglaise pour fournir aux enseignants des ressources lexicales. De même, le projet *Obneb* (Cabré et de Yzaguirre 1995) a pour but la détection des néologismes dans la presse, mais cette fois en espagnol et en catalan. L'outil *NeoloSearch* (Janicijevic et Walker 1997) identifie les néologismes français qui se trouvent sur l'Internet, tandis que le système *Cordon* (W3 *Cordon*) utilise des techniques statistiques et donc n'est pas lié à une seule langue. Dans cet article, nous décrivons *Cenit* (*Corpus-based English Neologism Identifier Tool*), un prototype mis au point à Dublin City University. *Cenit* diffère des systèmes susmentionnés car il s'adresse spécifiquement aux terminologues et donc, contrairement à l'*Obneb* par exemple, *Cenit* détecte des néologismes dans les domaines spécialisés. L'*Obneb*, par contre, ne sert pas les besoins des terminologues car, selon Cabré et de Yzaguirre (1995: 90), l'*Obneb* «exclut préalablement toute nouveauté terminologique qui appartient à un domaine spécifique de la technologie, de la science ou de l'industrie».

2 Néologismes

Avant de décrire en détail le prototype, il faut s'arrêter un bref instant sur la question : « Qu'est-ce qu'un néologisme ? » Nous sommes d'accord avec Rey (1995 : 63) qui note que l'établissement d'une définition précise de « néologisme » est une tâche assez difficile. Néanmoins, la plupart des terminologues s'accordent sur les points suivants : un néologisme est une *unité lexicale* qui a l'air *innovateur* dans la langue à l'étude (Dubuc 1985 : 111 ; Rey 1995 : 64 ; Rondeau 1984 : 122 ; Sager 1990 : 79). Mais ils ne s'accordent pas toujours sur la nature spécifique de l'unité lexicale ou le genre particulier de la nouveauté. Par exemple, Rondeau (1984 : 127) identifie trois grandes catégories de néologismes (néologisme de forme, néologisme de sens, et emprunt), tandis que Rey (1995 : 68-70) précise trois catégories différentes (néologisme formel (y compris les emprunts), néologisme sémantique et néologisme pragmatique). Dubuc (1985 : 113-4), pour sa part, précise deux catégories principales : les néologismes de formation dite indirecte (où l'on donne à un mot existant un sens nouveau) et ceux de formation directe (où l'on crée une nouvelle entité lexicale, soit de toutes pièces, soit par regroupement d'éléments existants). Finalement, Sager (1990 : 71) n'identifie que les deux catégories suivantes de néologismes : les créations complètement nouvelles et les emprunts.

Quel que soit le mode de formation, un néologisme décrit une notion ou une réalité nouvelle. La terminologie vise à rassembler les données sur le lexique d'une langue de spécialité ainsi qu'à normaliser les termes (Sager 1990 : 2). Les ressources terminologiques (glossaires, banques de termes) devraient fournir, dans la mesure du possible, une représentation de l'état actuel du

domaine à l'étude, sur le plan aussi bien linguistique que notionnel. Mais souvent, tout au début de l'introduction d'une notion, la terminologie qui la décrit n'est pas fixe. Plutôt, d'après Sager (1990 : 82), plusieurs formes coexistent. Donc, en premier lieu, il est important que le terminologue reconnaisse, recueille et enregistre toutes les formes. Au fil du temps, il peut suivre les tendances d'usage et il peut faire des recommandations en ce qui concerne la création ou la sélection de termes préférés. Du fait qu'elle représente un travail important, la procédure de détection des néologismes est longue, trop longue pour le terminologue occupé, qui a besoin d'un outil pour lui venir en aide.

3 *Cenit* : vue d'ensemble

Nous avons créé un outil qui s'appelle *Cenit* (*Corpus-based English Neologism Identifier Tool*). *Cenit* vise à identifier des néologismes dans la langue de spécialité. Dans les paragraphes suivants, nous décrivons les caractéristiques et le fonctionnement de ce système. Néanmoins, il faut dire tout d'abord que *Cenit* n'est encore qu'un prototype ; donc il présente à l'heure actuelle quelques limites. Premièrement, nous avons choisi la langue anglaise comme point de départ, mais en principe, les techniques pourraient être adaptées au français. Deuxièmement, *Cenit* n'identifie que les néologismes de forme. Il ne peut pas détecter les néologismes de sens, créés, par exemple, par extension sémantique ou par changement de catégorie grammaticale. Le fait que *Cenit* ne traite que des unités lexicales qui consistent en un seul mot, qu'elles soient des créations complètement nouvelles ou des dérivations (créées par préfixation, suffixation, etc.) constitue une dernière limite du

prototype. C'est-à-dire qu'actuellement, *Cenit* ne peut pas détecter les néologismes de forme syntagmatique, mais nous avons l'intention d'accroître la capacité du système à cet égard dans la phase suivante du projet.

Pour réaliser *Cenit*, nous avons choisi le langage de programmation *Perl* car celui-ci manipule très bien les chaînes de caractères. Le système fonctionne sous *Unix* avec l'accès par le *World Wide Web*. *Cenit* comprend deux programmes principaux : le programme pour créer les dictionnaires spécialisés, et le programme de recherche. Le système comprend un dictionnaire de base provenant du lexique *Unix* et le programme pour créer les dictionnaires spécialisés donne au terminologue la possibilité d'ajouter d'autres dictionnaires afin d'adapter *Cenit* pour traiter un domaine spécifique.

Cenit vise à traiter un fichier d'entrée (fourni par l'utilisateur) afin d'identifier les néologismes qu'il comporte. Le programme de recherche est le programme principal qui exécute les tâches suivantes. Il transforme le fichier d'entrée en liste de types (c.-à-d. une liste qui comprend chaque forme différente figurant dans le texte) qu'il fait passer par une série de filtres. Ces filtres examinent chaque forme et essaient de l'éliminer en la comparant avec les dictionnaires. Après avoir été filtrée, la liste est considérablement réduite. Les mots qui restent peuvent être des néologismes et chacun d'entre eux est exposé sur l'écran, souligné dans le contexte d'apparition, à charge pour le terminologue d'effectuer un tri.

3.1 Programme pour créer les dictionnaires spécialisés

Le dictionnaire de base est assez général et de plus, il ne contient que des formes canoniques (c.-à-d. les

formes non-dérivées et non-fléchies). Donc, nous anticipons que les terminologues aimeront construire des dictionnaires spécialisés en utilisant des corpus de textes présents à l'état naturel. À ce stade, *Cenit* prend comme données d'entrée un corpus de textes spécialisés fournit par l'utilisateur. Il transforme ce corpus en liste d'occurrences (*tokens*) et finalement en liste de types. Évidemment, plus le corpus est grand, plus le dictionnaire spécialisé sera complet et plus les résultats seront fiables. Notons que le programme n'ajoute pas au dictionnaire spécialisé les formes qui se trouvent déjà dans le dictionnaire de base. Chacun des dictionnaires comprend 52 fichiers triés: deux par lettre (aA-zZ) pour qu'on puisse accéder soit aux minuscules, soit aux majuscules.

3.2 Programme de recherche

Dans les paragraphes suivants, nous décrivons les filtres et les stades de traitement plus en détail. L'ordre dans lequel ce programme traite les données est extrêmement important. La figure 1 illustre l'ordre de l'application des filtres. Les crochets indiquent qu'un stade est facultatif.

3.2.1 Traitement du fichier d'entrée

La première étape dans le processus d'identification des néologismes est le traitement du fichier d'entrée. *Cenit* crée trois copies différentes du fichier d'entrée. Dans la première copie, tous les signes de ponctuation sont supprimés. Le système emploiera cette copie pour créer une liste d'occurrences (*tokens*) et finalement une liste de types. La deuxième copie ressemble à la première à la différence que les points (.) ne sont pas supprimés. *Cenit* emploiera la deuxième copie pour identifier les noms propres en déterminant si un nom avec une

majuscule se trouve au début ou au milieu d'une phrase. La troisième copie est une copie exacte du fichier d'entrée d'origine où tout reste intact. Cette copie sera utilisée au dernier stade quand *Cenit* exposera les néologismes candidats dans leurs contextes d'apparition.

3.2.2 Exclusion des correspondances exactes

Le premier filtre prend comme données d'entrée la liste des types provenant de la première copie du fichier d'entrée. Chaque mot est comparé avec le dictionnaire de base et avec les dictionnaires spécialisés. Seuls les mots qui ne figurent pas dans les dictionnaires avancent à la prochaine étape.

3.2.3 Exclusion par l'analyse morphologique

Car le dictionnaire de base ne contient que les formes canoniques (par exemple, *manger*), l'utilisateur peut souhaiter exclure aussi les formes dérivées ou fléchies (par exemple, *mange*, *mangeons*, *mangé*, etc.). Cependant, la dérivation est l'un des modes les plus communs de formation des néologismes, c'est la raison pour laquelle, si l'on utilise ce filtre, on risque exclure les néologismes intéressants (par exemple, les néologismes hypothétiques *mangement* ou *mangeation*). Donc, cette étape n'est pas obligatoire. Nous supposons que la plupart des formes dérivées et fléchies «connues» se trouvent dans le dictionnaire spécialisé créé par le terminologue car la source de ce dictionnaire est un corpus de textes présents à l'état naturel. Par conséquent, nous imaginons que cette option ne sera employée que par les utilisateurs qui n'ont pas de dictionnaire spécialisé.

À ce stade, les mots sont déjà traités par le filtre des

correspondances exactes, ils ne se trouvent donc pas ni dans le dictionnaire de base, ni dans le dictionnaire spécialisé. Si on lance le filtre d'analyse morphologique, chaque mot est traité pour en déterminer la racine et si la racine se trouve dans l'un des dictionnaires, *Cenit* supprime la forme dérivée ou fléchie.

3.2.4 Séparation entre les mots qui commencent par une majuscule et les mots en minuscules

Ce stade prend comme données d'entrée la liste du stade précédent. Les mots sont divisés en deux groupes: le premier groupe comprend les mots qui commencent par un caractère majuscule, et le deuxième comprend les mots qui ne commencent pas par une majuscule. Les trois stades suivants ne traitent que les mots du premier groupe.

3.2.5 Exclusion des correspondances exactes sans faire attention au type de caractère

Toutes les majuscules sont transformées temporairement en minuscules. Ensuite, chacun des mots est comparé à nouvelle reprise avec les dictionnaires. Cette seconde comparaison vise à supprimer les mots qui n'avaient de majuscule que parce qu'ils étaient placés en tête de phrase.

3.2.6 Exclusion des initialismes

La création d'un néologisme par l'acronymie est valable, surtout si le sigle se prononce comme un mot (par exemple, *Jaser*). Lorsqu'ils sont reconnus comme mots, ces acronymes s'écrivent souvent en minuscules. Néanmoins, plusieurs terminologues ne s'intéressent pas aux sigles qui ne comprennent que des initiales tout en majuscules et qui ne se prononcent pas comme un mot (par exemple, CPTIAQ). Ce type de sigle

correspond souvent au nom d'un produit, d'une société, etc. *Cenit* propose alors un stade facultatif qui vise à supprimer les initialismes tout en majuscules. Pour chacun des mots sur la liste, le système calcule le nombre total de caractères ainsi que le nombre de majuscules, et si les deux sont égaux, *Cenit* supprime le mot.

3.2.7 Exclusion des noms propres

Ce filtre se divise en deux parties. D'abord, *Cenit* compare le mot qui précède le néologisme candidat avec une «liste d'arrêts» (stoplist) qui comprend les mots comme Monsieur, M^{me}, Professeur, Dr, etc. Si le mot précédent se trouve sur la «liste d'arrêts», le néologisme candidat est supprimé. Pour ceux qui restent, ils procèdent à l'étape suivante, qui emploie la deuxième copie du fichier d'entrée (c.-à-d. la version qui ne contient aucune ponctuation sauf les points). D'abord, le système repère précisément l'emplacement dans le texte de chaque néologisme candidat. Si le candidat n'est pas précédé par un point (c.-à-d. s'il se trouve au milieu d'une phrase), il est probable que le candidat est un nom propre et, par conséquent, *Cenit* le supprime.

3.2.8 Regroupement des mots en majuscules et des mots en minuscules

Les néologismes candidats qui ont des caractères majuscules sont maintenant réunis avec les candidats en minuscules. Cette liste représente la somme de néologismes candidats qui se trouvent dans le fichier d'entrée.

3.2.9 Exposition des néologismes candidats et de leurs contextes d'apparition

Le dernier stade consiste à exposer chacun des néologismes candidats en contexte d'apparition. *Cenit* repère l'emplacement de chaque

candidat dans la troisième copie du fichier d'entrée (c.-à-d. la copie exacte du fichier d'origine). Ensuite, le système identifie les limites de la phrase en cherchant les points et les caractères majuscules. Pour finir, *Cenit* fournit une liste de néologismes candidats en contexte d'apparition, à charge pour le terminologue d'effectuer un tri. (Noter que le terminologue peut regarder les résultats à la fin de chaque stade, s'il le désire). Ensuite, le terminologue peut ajouter au dictionnaire tout mot qui fait partie de la liste mais qui n'est pas néologisme.

3.3 Extension du prototype actuel

Le prototype actuel est trop limité pour répondre à tous les besoins des terminologues. À court terme, nous prévoyons de procéder à des extensions, comprenant la possibilité d'identifier des néologismes syntagmatiques et, peut-être, la possibilité de détecter les néologismes qui sont créés par changement de catégorie grammaticale. Pour identifier les syntagmes, nous avons l'intention d'étudier les bigrammes et les trigrammes dans les corpus. Pour détecter les changements de catégorie grammaticale, nous proposons d'examiner les collocations de termes.

4 Conclusion

Nous avons affirmé, au début de cet article, que les langues évoluent au rythme accéléré. Détecter manuellement les néologismes n'est plus pratique pour les terminologues, qui doivent souvent travailler dans des délais très serrés. Les ordinateurs se sont révélés utiles pour aider les terminologues à réaliser un travail important dont la détection des néologismes fait partie. C'est une

époque bien excitante pour la terminologie et nous espérons que le travail décrit ici pourra servir de point de départ à des recherches plus approfondies dans ce domaine.

*Sorcha Roche et Lynne Bowker,
School of Applied Language and
Intercultural Studies,
Dublin City University,
Irlande.*

Remerciements

Nous remercions Danièle Tort (DCU.LS) pour la correction des épreuves.

Bibliographie

Blackwell (S.), 1993: «From Dirty Data to Clean Language», dans Aarts (J.), de Haan (P.), et Oostdijk (N.), dir., *English Language Corpora: Design, Analysis and Exploitation*, Amsterdam, Rodopi, p. 97-105.

Bourigault (D.), Jacquemin (C.), L'Homme (M.-Cl.), dir., 1998: Computerm '98, *Proceedings of the First Workshop on Computational Terminology, COLING-ACL '98*, 15 août 1998, Université de Montréal, Canada.

Collier, (A.), 1993: «Issues of Large-Scale Collocational Analysis», dans Aarts (J.), de Haan (P.), et Oostdijk (N.), dir., *English Language Corpora: Design, Analysis and Exploitation*, Amsterdam, Rodopi, p. 289-298.

Cabré (M.T.) et de Yzaguirre (L.), 1995: «Stratégie pour la détection semi-automatique des néologismes de presse», dans *Traduction, terminologie, rédaction*, vol. 8 (2), p. 89-100.

Condamines (A.), 1995: «Analyse de textes spécialisés pour le recueil de données terminologiques», dans *Terminologies nouvelles*, n° 14, p. 35-42.

Dubuc (R.), 1985: *Manuel pratique de terminologie (2^e édition)*, Québec, Linguatex.

Janicijevic (T.) et Walker (D.), 1997 : «NeoloSearch: Automatic detection of neologisms in French Internet documents», présenté au colloque ACH-ALLC'97, Kingston, Canada. (<http://www.quics.queensu.ca/achallc97/papers/a009.html> [date d'accès 03/09/98])

Kageura (K.) et Umino (B.), 1996 : «Methods of Automatic Term Recognition: A Review», dans *Terminology*, vol. 3 (2), p. 259-289.

Le Meur (A.) et Depecker (L.), 1995 : «Balnéo: un projet de réseau informatique pour la veille néologique», dans *Terminologies nouvelles* n° 14, p. 48-53.

Miller (D.), Meyer (I.) et Michaud (D.), 1991 : «Terminologie et analyse notionnelle assistée par ordinateur», dans *Actes du Colloque international sur les industries de la langue*, Tome II, Québec, Office de la langue française et STQ, p. 781-800.

Renouf (A.), 1993 : «A word in time: First findings from the investigation of dynamic text», dans Aarts (J.), de Haan (P.), et Oostdijk (N.), dir., *English Language Corpora: Design, Analysis and Exploitation*, Amsterdam, Rodopi, p. 279-288.

Rey (A.), 1995 : *Essays on Terminology* (traduit par J.C. Sager), Amsterdam, John Benjamins.

Rondeau (G.), 1984 : *Introduction à la terminologie* (2^e édition), Québec, Gaëtan Morin.

Sager (J.C.), 1990 : *A Practical Course in Terminology Processing*, Amsterdam, John Benjamins.

Sites Internet

Aviator, University of Liverpool, Royaume-Uni : <http://radar.rdues.liv.ac.uk/newwds.html> [03/09/98]

Cenit, Dublin City University, Irlande : <http://www.compapp.dcu.ie/Projects/1998/CL/sroche.cl4/> [03/09/98]

Cordon, Institut für deutsche Sprache, Allemagne : <http://www.ids-mannheim.de/telri/seminar/jv-cordon.html> [03/09/98]

Neolosearch, Queen's University, Canada : <http://www.quics.queensu.ca/achallc97/papers/a009.html> [03/09/98]

Obneb, Universitat Pompeu Fabra, Espagne : <http://www.iula.upf.es/pjobneca.htm> [03/09/98]

Figure 1: Les stades du programme de recherche CENIT

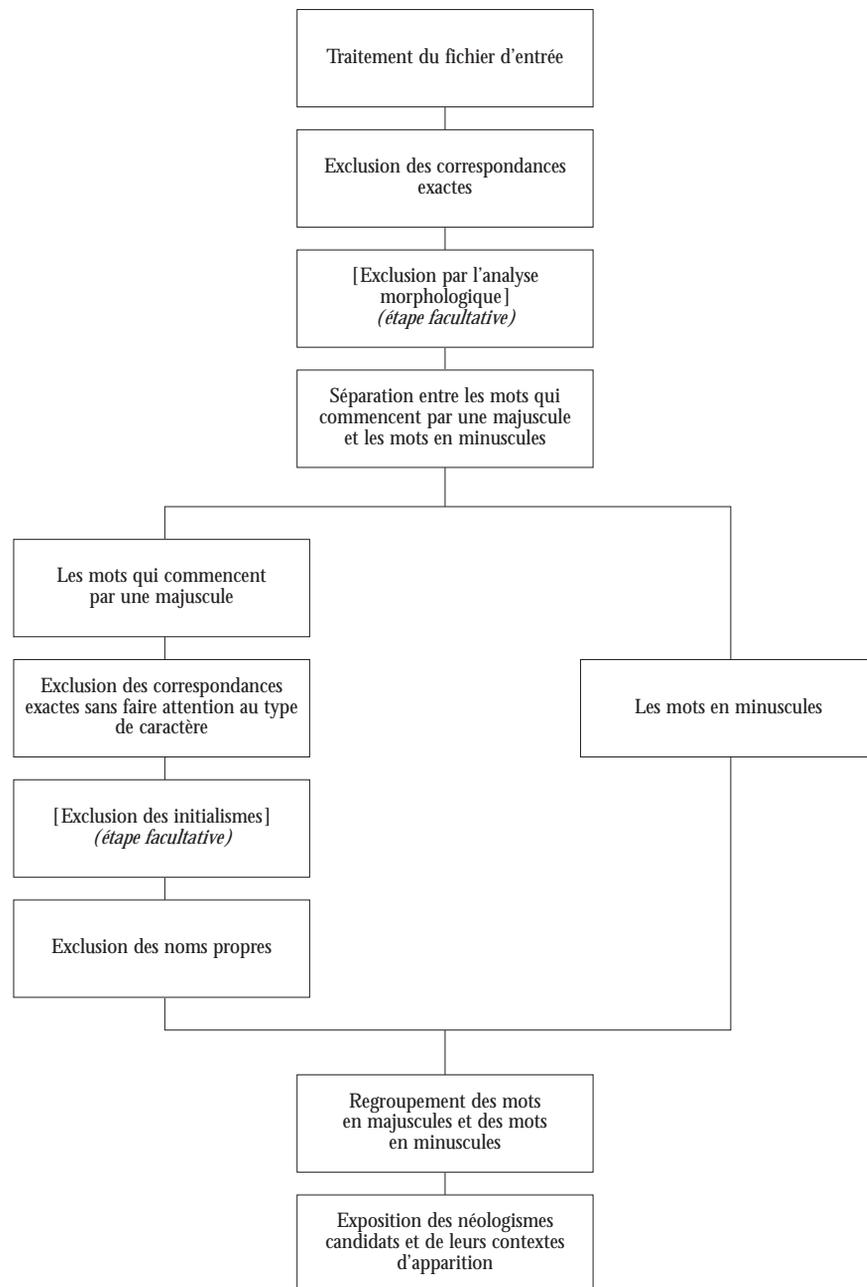


Figure 1: Les stades du programme de recherche *Cenit*

Utilisation d'une banque de textes en terminographie

L'article présente les résultats d'un projet d'expérimentation mené à l'Office de la langue française et visant à vérifier les possibilités et les modalités d'intégration d'un corpus informatisé à la chaîne de travail terminographique. Le projet se situe dans la continuité d'un projet de constitution de banques de textes déjà mené par l'Office dans le contexte de sa participation aux travaux de veille néologique du Rint. Après avoir décrit les logiciels et les méthodes de dépouillement utilisés, on formule des recommandations quant à la constitution et à l'exploitation de banques de textes en terminographie.

Termes-clés:
corpus informatisé;
gestion documentaire; terminotique;
banque de textes; logiciel.

1 Introduction

Le volume de textes disponibles sur Internet, donc en format numérique, est de plus en plus important. Cette situation présente un avantage indéniable du point de vue du terminographe; en effet, le texte écrit constitue la matière première de son travail et l'apparition de nombreux logiciels permettant le traitement informatisé de la langue lui offre la possibilité, pour peu qu'il intègre ces logiciels à ses méthodes, de traiter une quantité imposante de textes.

En 1998, dans le contexte de sa participation aux travaux de veille néologique du Rint, l'Office de la langue française a fait constituer par la Société de mathématiques appliquées (SMA) de Québec deux banques de textes, en imagerie médicale et en robotique industrielle, à partir de documents puisés à même le réseau Internet. L'objectif de l'opération était de procéder à l'analyse de ces banques de textes afin d'alimenter les systèmes de veille terminologique de la *Banque de terminologie du Québec* et de *Balnéo*⁽¹⁾. Les résultats de cette analyse ont amené l'Office à poursuivre l'expérimentation dans le domaine de la robotique industrielle.

Ce sont les résultats de cette nouvelle expérimentation que nous présentons ici: l'expérience visait à tirer d'un nouveau corpus, formé comme le premier de textes spécialisés recueillis sur Internet et constitués en banque textuelle, le plus grand

nombre possible de fiches contextuelles⁽²⁾, tout en analysant les méthodes utilisées pour le faire. On désirait ainsi évaluer plus en détail les possibilités et les modalités d'intégration d'un corpus informatisé à la chaîne de travail terminographique.

L'article précise d'abord les objectifs du projet et les logiciels utilisés, avant de s'attarder à la constitution du corpus et à la description des méthodes de dépouillement appliquées. Il présente ensuite les résultats sous forme de comparaison entre les différentes méthodes de dépouillement pour enfin formuler certaines propositions et recommandations.

2 Expérimentation

2.1 Objectifs

Notre objectif général était de vérifier comment l'utilisation d'une banque de textes informatisée pouvait s'intégrer dans la chaîne de travail d'un terminologue participant à l'alimentation d'un dictionnaire terminologique. Il s'agissait d'abord d'évaluer la pertinence de l'utilisation d'une banque textuelle comme corpus de dépouillement, par opposition, par exemple, à l'exploitation ponctuelle de chacun des textes au fur et à mesure de sa découverte. Il fallait aussi déterminer la meilleure façon d'analyser le corpus ainsi constitué, en tentant de voir comment les méthodes traditionnelles de dépouillement pouvaient être améliorées ou assistées par des outils

(1) Le système *Balnéo*, créé par le Rint, a pour objectifs la collecte, l'échange et la diffusion rapides de néologismes sur Internet: (<http://www.rint.org/cadreIndex7.htm>).

(2) Nous entendons par fiche contextuelle une fiche portant les données suivantes: terme français, contexte(s), source(s). Mentionnons également que les fiches rédigées devaient traiter de notions absentes de la *Banque de terminologie du Québec (BTQ)*.

logiciels. On désirait aussi vérifier la compatibilité et la convivialité des logiciels utilisés.

2.2 Logiciels utilisés

L'expérimentation a mené à l'utilisation d'une série de logiciels correspondant chacun à une étape du travail terminologique. Pour la recherche documentaire sur Internet, nous avons utilisé *Copernic*⁽³⁾, un moteur de recherche qui permet de conserver les archives des recherches déjà faites et qui conjugue les possibilités de plusieurs logiciels de recherche en les lançant simultanément pour ensuite faire la synthèse des résultats. Pour la gestion de la banque, nous avons utilisé l'*Atelier BDT*, logiciel mis au point par la société SMA de Québec, qui permet la constitution et la gestion de banques de textes et qui est doté d'un moteur de recherche (dans le corpus) assez performant. Pour effectuer la transition entre le format utilisé sur Internet et l'*Atelier BDT*, il est nécessaire de se servir d'un logiciel de traitement de texte; nous avons utilisé *Word*. En ce qui a trait à l'analyse textuelle, nous nous sommes servis du logiciel de dépouillement *Adepto-Nomino*. Enfin, pour la rédaction des fiches, nous avons effectué le transfert des fiches rédigées avec *Adepto-Nomino* vers *Sami*⁽⁴⁾, le logiciel d'alimentation et de mise à jour de la *Banque de terminologie du Québec (BTQ)*. Dans les cas où nous avons dépouillé les textes manuellement, les fiches ont été rédigées directement dans *Sami*.

(3) On peut télécharger une version gratuite de *Copernic* à l'adresse suivante : <http://www.copernic.com/fr/>

(4) Pour plus d'informations sur *Sami*, voir Pelletier (1996).

2.3 Constitution du corpus

Nous avons constitué, à l'aide du moteur de recherche *Copernic*, une banque d'environ 73 000 mots, dans le domaine de la robotique. Nous avons privilégié les textes rédigés en français émanant de centres de recherche universitaires ou gouvernementaux (rapports d'activités, thèses, projets de recherches, articles, etc.). En effet, ce type de texte nous paraissait offrir le plus de potentiel pour la recherche de termes nouveaux, par la crédibilité de leurs auteurs – des chercheurs de haut niveau – parce qu'ils ont été rédigés en français (et non traduits) et parce qu'ils traitent de domaines de pointe.

Nous reviendrons plus loin sur l'importance à accorder à l'étape de constitution du corpus de dépouillement. Signalons simplement que nous n'avions pas pour ambition de former un corpus se suffisant à lui-même, c'est-à-dire qui contiendrait suffisamment d'informations provenant de sources assez diversifiées pour servir à lui seul de documentation pour la rédaction des fiches terminologiques. Notre but était plutôt de constituer une banque de textes servant à trouver des termes absents de la *BTQ*, et dont la nature permettrait de recueillir des contextes très « parlants », pouvant guider le terminologue abordant ces notions pour compléter les fiches (c'est-à-dire collecter les données conceptuelles nécessaires à la rédaction d'une définition et éventuellement de notes) et lui permettre de s'engager avec confiance dans la bonne direction pour la recherche d'informations supplémentaires.

Une fois les textes trouvés, encore fallait-il les constituer en corpus, ce que nous avons fait grâce à l'*Atelier BDT*. Nous reviendrons plus loin sur les manipulations liées à cette opération.

2.4 Dépouillement

Traditionnellement, lorsqu'il s'agit d'établir une nomenclature pour un domaine spécialisé, le terminologue s'arme de son crayon et de sa patience et entend la lecture de textes récents publiés dans le domaine qui l'intéresse. Au fil de cette lecture, il se familiarise avec le domaine en question et relève ce qu'on appelle des candidats-termes, c'est-à-dire des unités simples ou syntagmatiques qui semblent se distinguer par leur caractère d'univocité et de spécificité. Le terminologue relève donc ces unités avec leur contexte et, éventuellement, les informations terminologiques qu'il déniche dans le texte et qui lui permettront plus tard de rédiger définition et notes. Cette étape, le dépouillement, gourmande de temps, devra être suivie d'une recherche exhaustive permettant de vérifier les emplois du terme, sa fréquence, ses équivalents (dans le cas d'une banque multilingue), etc. Un premier problème soulevé à l'étape du dépouillement est donc un problème de temps.

Une autre difficulté relative au dépouillement est liée à l'absence de nomenclature de départ : en effet, en veille, l'objectif est d'abord d'établir une telle nomenclature, puisqu'on est alors à la recherche de termes que l'on n'a pas encore traités, c'est-à-dire relevés, attestés, consignés, définis, etc. On ne sait donc pas exactement ce qu'on cherche, et la démarche fait appel à une bonne dose d'instinct et d'expérience; les résultats obtenus dépendent donc beaucoup de l'expérience du terminologue et de sa connaissance du domaine traité.

On a tenté, lors de l'expérimentation décrite ici, de vérifier s'il était possible de lever ces difficultés en automatisant l'étape du dépouillement des textes. Pour cela, nous avons choisi de comparer trois

approches; la comparaison s'est faite sur le plan du temps nécessaire pour produire un certain nombre de fiches. Un tableau résume cette comparaison dans la section *Analyse*.

La première approche, que nous avons appelée *dépouillement classique*, consistait à imprimer les textes et à souligner à la main les candidats-termes pour relancer ensuite ces termes dans la banque de textes afin de trouver les meilleurs contextes. Remarquons que, malgré son nom, cette démarche se distingue déjà d'une méthode plus traditionnelle puisqu'elle utilise, même si ce n'est que partiellement, les possibilités d'un logiciel comme l'*Atelier BDT* pour la recherche des contextes; pour suivre vraiment ce que nous avons décrit plus haut comme la méthode de travail habituelle du terminologue, il aurait fallu imprimer tous les textes et recueillir, en les recherchant visuellement, les meilleurs contextes pour chaque terme retenu. Cependant, notre façon de procéder a quand même donné des résultats assez probants lors de la comparaison avec les deux autres méthodes.

La deuxième approche est intermédiaire: dans ce que nous avons appelé *dépouillement semi-automatique*, il s'agissait d'établir une nomenclature d'éléments de base susceptibles de donner des termes complexes (par exemple *capteur à ultrason*) ou de racines susceptibles de fournir des termes simples et complexes (comme l'élément tronqué *holon** qui ramènera *holonomie*, *holonome*, *robot holonome*, *non-holonome*, etc.) et de lancer chacun des termes de cette nomenclature dans la banque de textes pour recueillir occurrences, cooccurrences et contextes. Il s'agit de la méthode qui correspond le mieux à l'optique dans laquelle l'*Atelier BDT* a été constitué: ce logiciel est en effet équipé d'un moteur de recherche

intéressant et performant qui permet de trouver aisément tous les contextes et cooccurrences d'un terme donné. Il suffit de lancer un terme pour trouver ses composés, des exemples de son utilisation, des contextes. Cependant, le logiciel ne fournissant pas à ce jour d'index, la principale contrainte liée à son utilisation vient du fait qu'il faut savoir quoi demander au moteur de recherche pour obtenir des résultats. Il est donc nécessaire de disposer d'une nomenclature de départ; or, comme nous l'avons déjà souligné, en veille, l'objectif est justement d'établir une telle nomenclature. Une des façons de contourner le problème est de lancer dans la banque de textes des termes de base (par exemple, pour le domaine de la robotique, des termes comme *robot*, *commande*, *manipulateur*, etc.) qu'on aura recueillis lors de lectures préliminaires. Cependant une telle méthode ne permet pas de trouver des termes composés à partir de bases auxquelles on n'aura pas songé, sinon par hasard. Si on veut agir de façon plus méthodique, il faudra donc aller plutôt vers la troisième méthode, le *dépouillement automatique* qui consiste simplement à faire analyser les textes de la banque par un logiciel de dépouillement textuel, ici *Adepte-Nomino*.

2.5 Résultats

Les résultats sont groupés selon la méthode utilisée, ce qui permettra plus loin la comparaison entre ces méthodes.

2.5.1 Dépouillement classique

Nous n'avons effectué que très peu de dépouillement selon cette méthode: elle a été utilisée surtout dans le but d'obtenir une référence, un point de comparaison avec les autres méthodes, étant donné que c'est celle qui «colle» le mieux à ce

que font traditionnellement les terminologues en situation de veille terminologique.

À partir de deux textes choisis au hasard dans la banque de textes, puis imprimés, nous avons souligné les candidats-termes, avant de constituer une fiche contextuelle pour chacun d'entre eux en les relançant dans la banque de textes pour trouver le ou les meilleurs contextes, après avoir vérifié leur absence de la BTQ. En une heure, 11 fiches contextuelles ont ainsi été rédigées.

2.5.2 Dépouillement semi-automatique

Rappelons que pour utiliser cette méthode il faut disposer d'une nomenclature de départ; nous avons donc choisi de lancer les 15 composantes des 11 termes trouvés avec la première méthode⁽⁵⁾, ce qui a permis d'isoler 62 nouvelles notions et de rédiger les fiches contextuelles correspondantes en environ 8 heures. Rappelons qu'il s'agissait ici de lancer chacun des éléments de base dans la banque de textes, de relever chacun des candidats-termes formés à partir de ces éléments, puis de relever le ou les meilleurs contextes pour chacun de ces termes avant de rédiger la fiche contextuelle correspondante dans *Sami*.

(5) Nous entendons par *composante* chacun des éléments de base d'un terme: si on a trouvé par exemple *capteur à ultrason*, on relancera *capteur* et *ultrason*.

(6) Nous avons en premier lieu tenté l'analyse de l'ensemble des textes placés dans la banque de textes avec *Tact*, logiciel mis au point à l'Université de Toronto (et dont on peut télécharger une version Web à l'adresse suivante: <http://tactweb.humanities.mcmaster.ca/>). *Tact* fournit une liste des mots trouvés dans le texte, accompagnés de leur fréquence d'utilisation; il faut, à partir de cette liste, sélectionner les termes dont on pense qu'ils pourraient donner des résultats intéressants (soit qu'il s'agisse de termes simples, ou d'éléments de base de termes complexes) et les marquer un par un avant de demander à *Tact* de produire une liste de cooccurrences ou de contextes. Pour notre banque de textes, *Tact* a produit une liste d'environ 10 000 mots, qu'il aurait fallu lire et sélectionner, sans l'aide, à cette étape du traitement, d'un contexte pour décider de la pertinence de chacun d'eux. Devant l'ampleur de la tâche, nous avons décidé de tenter immédiatement un essai avec *Adepto-Nomino*. Il faut également souligner que les manipulations préliminaires à l'analyse du texte sont peu conviviales avec *Tact*, qui fonctionne en environnement Dos, avec lequel la plupart des utilisateurs sont peu familiers.

(7) Il faut noter que la mise en forme des textes faite automatiquement par l'*Atelier BDT* nuit à l'utilisation d'*Adepto-Nomino*, à cause des différents marqueurs qui y sont introduits; pour obtenir les meilleures performances d'*Adepto-Nomino*, nous avons donc dû partir des textes dans leur format «original», c'est-à-dire sauvegardés en format *texte seulement* par *Word*.

(8) Ne sont décrites ici que les «sorties» d'*Adepto-Nomino*: pour une description plus détaillée de son fonctionnement, voir Perron (1996: 32).

(9) Le temps nécessaire à ce transfert n'a pas été comptabilisé. Dans le cadre de ce projet, il a été effectué par les services informatiques de l'Office; c'est une opération très rapide (à titre d'exemple, il faut moins d'une minute pour verser 170 fiches).

2.5.3 Dépouillement automatique⁽⁶⁾

Nous avons lancé *Adepto-Nomino*, à partir encore une fois de l'ensemble des textes placés dans la banque⁽⁷⁾. Après une analyse (qui dure entre 20 et 25 minutes pour une banque de textes telle que la nôtre), *Adepto-Nomino* produit cinq listes⁽⁸⁾: les unités complexes nominales (ucn), les unités complexes nominales additionnelles (ucna), une liste des verbes, une liste des adjectifs et une liste des noms. C'est surtout à partir de la liste des ucn et de la liste des noms que nous avons identifié des candidats-termes.

Lorsqu'on interroge n'importe lequel des éléments des listes fournies par *Adepto-Nomino*, le logiciel le replace dans son contexte (autant de fois qu'il y a d'occurrences). Il est donc possible de choisir immédiatement les contextes qui paraissent les meilleurs et de les transférer directement dans le champ approprié. En environ trois heures et demie, 89 candidats-termes ont été recueillis et leurs fiches contextuelles rédigées (après vérification dans la BTQ). Le résultat se trouvait sous la forme de fiches au format *Adepto-Nomino*, que nous avons versées dans *Sami* par la suite⁽⁹⁾.

Analyse des résultats

Le tableau 1 résume les données relatives au nombre de fiches et au temps nécessaire à leur production pour chaque méthode:

Notons d'abord que malgré son apparente incongruité (il est moins élevé que le résultat du dépouillement classique), le résultat correspondant au dépouillement semi-automatique est tout à fait normal; en effet, comme nous l'avons mentionné dans la section *Résultats*, il s'agit d'une méthode en quatre étapes:

1. Établir une nomenclature de base;
2. Lancer les éléments de cette nomenclature dans la banque de textes à l'aide du moteur de recherche de l'*Atelier BDT*;
3. Recueillir les candidats-termes complexes produits par cette recherche;
4. Pour chacun d'entre eux, rechercher le ou les meilleurs contextes.

On pourrait bien sûr recueillir les contextes (étape 4) en même temps qu'on identifie les termes complexes (étape 3), mais il semble que cette façon de faire soit plus compliquée et produise plus de «bruit»; on risque en effet de se retrouver avec des contextes inutiles qu'on devra éliminer par la suite. C'est pourquoi nous avons choisi d'établir d'abord la liste des termes complexes pour ensuite faire une recherche de contextes (cette recherche est d'ailleurs plus restreinte que la précédente, puisqu'on ne lance dans la banque de textes que le composé qui nous intéresse).

La méthode la plus performante semble être de loin, au vu des résultats reproduits dans le tableau, celle qui fait intervenir un logiciel de

Tableau 1

<i>Méthode</i>	<i>Nombre de fiches</i>	<i>Temps nécessaire</i>	<i>Nombre de fiches à l'heure</i>
Classique	11	1 heure	11
Semi-automatique	62	8 heures	7,75
Automatique	89	3 heures et demie	24,4

dépouillement comme *Adepto-Nomino*. Nous pensons qu'une telle performance s'explique par la rapidité et l'efficacité permises par l'ordinateur, combinées à la liberté laissée au terminologue; en effet, avec une telle méthode, c'est le logiciel qui propose les candidats-termes, et c'est le terminologue qui dispose de ces propositions comme il l'entend. On pourrait objecter que le terminologue, avec une telle méthode, ne lit plus les textes, et donc n'est plus aussi familier avec son domaine. À ceci nous répondrons que la performance d'un outil de dépouillement automatique de textes donne au contraire plus de temps au terminologue pour lire, et qui plus est pour lire des textes imprimés qui, de toute façon, n'auraient pas pu être analysés de façon automatique (à moins d'être d'abord numérisés, ce qui pose des problèmes différents). Nous dirions que le logiciel s'occupe des textes disponibles sous format numérique, et l'être humain des textes imprimés. De plus, c'est le terminologue qui a choisi les textes à faire analyser (et donc qui les a au moins parcourus) et c'est lui qui analyse chaque contexte proposé par *Adepto-Nomino* pour juger de sa pertinence, ce qui contribue de façon appréciable à sa familiarisation avec le domaine.

3 Propositions et recommandations

3.1 Constitution d'un corpus

La qualité et la diversité des textes choisis constituent l'assise d'un travail terminologique basé sur le corpus; elles déterminent en effet en grande partie l'utilité et la pertinence du résultat final. L'établissement de critères régissant le choix de textes destinés à faire partie d'un corpus de dépouillement nous paraît donc essentiel; la poursuite d'essais portant

sur la constitution de banques textuelles permettrait d'établir des orientations méthodologiques aptes à guider le terminologue lors de cette étape primordiale qu'est la mise sur pied de son corpus de dépouillement.

3.2 Utilité d'une banque de textes

Dans l'optique où l'on considère qu'une banque de textes doit être «rentable», c'est-à-dire que le temps qu'on a investi dans sa constitution doit avoir été utilement dépensé, il est permis de se poser des questions sur l'utilité d'une telle banque. Nous soulignerons simplement ici deux faits en réponse à ces doutes: en situation de veille terminologique, où on ne dispose pas de nomenclature de départ, l'expérimentation décrite ici fait ressortir le rapport entre le temps investi et les résultats. Pour une demi-journée passée à récolter des textes sur Internet, on peut tirer, en deux jours environ, plus de 400 fiches contextuelles⁽¹⁰⁾ de la banque ainsi constituée. L'expérimentation, bien qu'elle n'ait pas directement porté sur une situation de mise à jour du fonds terminologique, a également fait ressortir, de façon sous-jacente, le fait suivant: on a dû éliminer une certaine quantité de termes qui étaient déjà dans la BTQ; étant donné la qualité des textes stockés (consistant, rappelons-le, en grande partie en rapports de recherche universitaires), il paraît donc probable que pour ces termes la banque de textes aurait pu servir de ressource documentaire fort utile (jointe aux ressources habituelles) en situation de mise à jour. Il faudrait mener d'autres expérimentations, avec des corpus plus importants, pour vérifier cette hypothèse.

3.3 Manipulations informatiques

Il faut mentionner ici que les opérations de stockage des textes dans l'*Atelier BDT* comprennent des manipulations informatiques (création et initialisation de la banque, création d'un nouveau répertoire, puis marquage, segmentation et importation des textes, mise à jour de la page d'accueil) susceptibles de rebuter de nombreux terminologues peu familiers avec les dessous de la manipulation de fichiers, ce qui a amené certains terminologues à s'interroger sur les étapes de la constitution d'une banque de textes (Michel, 1998:10): «Qui doit faire les recherches menant à sa constitution? Les bibliothécaires sont-ils les mieux qualifiés pour ce genre de démarche? Comment en arriver à jumeler l'expertise du terminologue familiarisé avec les sites les plus intéressants et celle des documentalistes?»

En réponse à ces questions, nous proposons ici⁽¹¹⁾ un mode de formation d'une banque de textes – avec l'*Atelier BDT* – qui décharge le terminologue des tâches qui sont par trop étrangères à son travail habituel et qui, lui paraissant lourdes et malaisées, pourraient le décourager de

(10) On considère le travail accompli par un terminologue habitué à la procédure et utilisant un outil de dépouillement automatique; comme pour toute nouvelle technique, les premières expériences seront peut-être moins «rentables»...

(11) On trouvera en annexe une proposition détaillée de partage des tâches qui ne demande plus au terminologue que de sauvegarder sous le bon format le texte qu'il a trouvé en ajoutant la date de consultation et l'adresse du site.

mettre sur pied cette banque, tout en tenant pour acquis qu'il est le mieux placé pour repérer les meilleurs textes dans son domaine de recherche.

3.4 Exploitation des textes

Une autre préoccupation des terminologues portait sur la pertinence de consacrer du temps à intégrer un texte intéressant à une banque textuelle au lieu d'en extraire au moment même toutes les informations utiles. Cette question nous amène à rappeler ici le concept de *banque évolutive*: un terminologue, tant qu'il travaillerait dans le même domaine, continuerait d'enrichir sa banque et la mettrait éventuellement en commun avec ceux de ses collègues qui travaillent dans un domaine voisin du sien. Cette façon de voir est inspirée par l'observation suivante: l'expérience démontre qu'une notion ayant fait l'objet d'une recherche sur Internet «ramène» souvent plusieurs autres notions avec elle; le texte ainsi trouvé risque donc de s'avérer utile ultérieurement et il est parfois difficile de retrouver un site Internet (même s'il est marqué d'un signet – on court toujours le risque qu'il soit devenu inaccessible ou qu'il ait été déplacé) alors qu'il est possible de le conserver sur son propre disque dur, et qui plus est, avec les possibilités qu'offre le moteur de recherche d'un logiciel de gestion documentaire. De plus, le terminologue, au fur et à mesure qu'il fait connaissance avec son domaine, sait de mieux en mieux quoi chercher et où le chercher; sa banque de textes gagne donc en qualité au fil du temps, et peut devenir une source inestimable d'informations dans un domaine donné. Elle peut servir ultérieurement à des opérations de mise à jour, à familiariser un nouveau terminologue avec un domaine donné, à faire des études d'évolution de la terminologie dans une sphère

d'activité, etc. On pourrait objecter, en se référant au phénomène de saturation déjà souligné par plusieurs auteurs (voir par exemple Otman 1991b: 61), qu'on observe qu'à partir d'une certaine taille (au demeurant difficile à déterminer), un corpus de textes ne fournit plus beaucoup de terminologie nouvelle. Ce phénomène nous semble s'appliquer de façon plus marquée dans certaines situations: dans les cas de corpus en langue générale par exemple, ou dans un domaine qui connaît une évolution plutôt lente. Dans les domaines spécialisés identifiés comme des priorités par l'Office, ce phénomène ne paraît cependant pas représenter un obstacle à la constitution d'une banque évolutive, pouvant au contraire servir à indiquer que le domaine est à peu près couvert lorsque l'ajout de nouveaux textes n'apporte plus ou presque plus de terminologie nouvelle. Dans tous les cas, de nouvelles expérimentations pourraient fournir des indicateurs sur la taille optimale d'un corpus en langue spécialisée.

3.5 Convivialité et compatibilité des logiciels

Comme nous l'avons brièvement mentionné plus haut, en ce moment l'Atelier BDT et *Adepte-Nomino* ne sont pas pleinement compatibles; les textes une fois marqués par l'*Atelier* deviennent inutilisables (ou posent du moins plusieurs difficultés d'utilisation) par *Adepte-Nomino*. C'est ce problème relatif à la compatibilité entre les logiciels qui nous amène à formuler, après discussion avec les concepteurs de l'*Atelier BDT*, deux recommandations dont voici la première (qui constitue une étape vers la mise en œuvre de la situation idéale): munir l'*Atelier BDT* d'un index interactif qui permettrait d'interroger n'importe quel mot du

texte pour le replacer dans son contexte et éventuellement découvrir ses composés. Ceci éliminerait (ou du moins réduirait) le besoin d'une nomenclature de départ. Mais la solution qui nous semble la meilleure est celle-ci: mettre au point une interface de communication entre l'*Atelier BDT* et *Adepte-Nomino*, qui permettrait d'analyser directement en tout ou en partie la banque de textes constituée grâce à l'*Atelier BDT*. De plus, comme nous y avons fait allusion plus tôt, l'interface entre *Adepte-Nomino* et *Sami* existe déjà; une nouvelle interface entre l'*Atelier BDT* et *Adepte-Nomino* résulterait donc en un outil intégré constitué d'un logiciel de gestion documentaire, d'un logiciel de dépouillement et d'un logiciel d'alimentation d'une banque de terminologie, outil qui s'inscrirait aisément dans la chaîne de travail terminographique.

3.6 Proposition de chaîne de travail informatisée en situation de veille

Même s'il évolue assez rapidement, le travail du terminologue-terminographe comporte toujours les mêmes grandes étapes (Auger, 1994: 50): «1) préparation du travail terminologique [...] (exploration du domaine, recherche de documentation, sélection et enregistrement des sources); [...] 2) travail terminographique proprement dit [...] (mise en forme et traitement d'un corpus de textes, dépouillement, établissement de la nomenclature, extraction de données terminologiques et paraterminologiques, [...] etc.); 3) stockage et traitement des données recueillies [...]; 4) édition et diffusion du produit final (dictionnaire terminologique) [...]»

Ce que nous proposons ici, c'est de mieux appuyer certaines de ces

Tableau 2

Étape	Logiciel
Recherche de textes pertinents sur Internet	<i>Copernic</i>
Constitution d'une banque de textes	<i>Microsoft Word et Atelier BDT</i>
Extraction des termes	<i>Adepte-Nomino</i>
Extraction des contextes, notes, définitions, etc.	<i>Adepte-Nomino</i>
Rédaction des fiches	<i>Sami</i>

étapes à l'aide de logiciels permettant d'automatiser en partie la tâche du terminologue.

Le projet d'expérimentation mené à l'Office de la langue française concernait surtout les premières étapes du travail terminologique, soit la préparation du travail par la recherche documentaire et la sélection de sources, de même que le travail d'extraction des données terminologiques. Le tableau 2 présente sous forme schématique la chaîne de travail proposée, avec pour chaque étape un exemple de logiciel disponible⁽¹²⁾.

4 Conclusion

Il est apparu lors de cette expérimentation que l'utilisation des outils informatiques étudiés pouvait rendre plus agréable et plus efficace le travail du terminologue. À partir de logiciels existants, comme l'*Atelier BDT* et *Adepte-Nomino*, il est possible de mettre au point des outils sur mesure parfaitement adaptés au travail à accomplir, surtout en situation de veille, mais aussi pour la mise à jour, où ils s'ajouteraient aux ressources déjà disponibles. Il y a bien sûr tout un travail d'adaptation à accomplir, autant des outils aux personnes que l'inverse. Cette adaptation demande du temps et risque même de ralentir la production pendant une période donnée, où

chacun doit apprivoiser les nouveaux outils mis à sa disposition. Mais dans un monde où l'information est de plus en plus considérée comme une richesse, il est essentiel de se préoccuper de trouver des moyens de la conserver sous la meilleure forme et de la traiter le plus efficacement possible. Les terminologues sont des spécialistes du traitement de l'information sous bien des aspects; il leur revient donc de se tenir au courant des possibilités offertes par Internet et par l'informatique en général afin d'exploiter au mieux ces possibilités; ils demeureront ensuite les meilleurs juges des outils les mieux adaptés à leur travail.

*Anouk Jaccarini,
Office de la langue française,
Québec.*

(12) Il faut souligner qu'il existe de nombreux autres logiciels ayant les mêmes fonctions ou des fonctions semblables; on s'en tient ici à ceux qui ont servi au projet.

Bibliographie

Auger (Pierre), 1994: «Les outils de la terminologie: typologie des logiciels d'aide à la terminologie et/ou d'automatisation de la chaîne de travail en terminographie» dans *Terminologies nouvelles*, n° 11, p. 46-52.

Côte (Normand): «Formation aux banques virtuelles de textes – commentaires de l'équipe du tronc commun industriel», *document interne*, Office de la langue française, mars 1999.

Michel (France), 1998: «Banques virtuelles de textes scientifiques, une aide possible à la veille terminologique?» dans *Terminologies nouvelles*, n°18, p. 5-10.

Otman (Gabriel) 1991a: «Aspects de l'informatisation des activités terminologiques et traductionnelles» dans Clas (A.) et Hayssam Safar, dir., *L'environnement traductionnel: la station de travail du traducteur de l'an 2001 – Actes du colloque de Mons, 1991*, Sillery, Presses de l'Université du Québec, p. 213 à 221.

Otman (Gabriel) 1991b: «Des ambitions et des performances d'un système de dépouillement terminologique assisté par ordinateur» dans *La banque des mots*, numéro spécial 4/1991, p. 59-96.

Pelletier (Sylvie) 1996: «Le système *Sami-BTQ*: un système intégré de production et de mise à jour des données de la Banque de terminologie du Québec» dans *Terminologies nouvelles*, n°15, p. 66-76.

Perron (Jean), 1996: «*Adepte-Nomino*, un outil de veille terminologique» dans *Terminologies nouvelles*, n°15, p. 32-47.

Annexe

Proposition de partage du travail pour la constitution d'une banque de textes avec l'*Atelier BDT*⁽¹³⁾.

FRÉQUENCE	TEMPS	RESPONSABLE DE LA DOCUMENTATION	TERMINOLOGUE
lors de la création d'une nouvelle banque de textes	5 minutes	<ul style="list-style-type: none"> – sous l'<i>Atelier BDT</i>, crée une nouvelle banque – informe le terminologue du répertoire où il doit placer ses textes 	
variable (hebdomadaire, mensuelle)	une demi-journée pour environ 35 000 mots		– copie les textes qu'il trouve pertinents sous format <i>WP5.0</i> (en ajoutant l'adresse du site et la date de consultation) dans le répertoire indiqué par le responsable de la documentation
sur notice du terminologue, afin que les textes deviennent disponibles le plus rapidement possible	environ 2 minutes par texte	<ul style="list-style-type: none"> – sous l'<i>Atelier BDT</i>, marque (adresse Internet, titre et titre abrégé) et segmente les textes – importe les textes dans la banque – met à jour le fichier <i>bdt.htm</i> (page d'accueil de l'<i>Atelier BDT</i>) – vérifie la banque (sous <i>consultation</i>) 	
			– la banque devient disponible et prête à être utilisée par le terminologue

(13) Note: dans ce texte, nous nommons *responsable de la documentation* la personne (technicien ou documentaliste) qui serait chargée, entre autres, d'assister le terminologue dans la constitution d'une banque de textes.

Recherche terminographique semi-automatisée en veille terminologique : expérimentation dans le domaine médical

Le présent article fait état d'une recherche terminographique intégrant des outils informatiques (outil de dépouillement, concordancier, système de gestion de bases de données et module de création de fiches terminologiques). La recherche, qui portait principalement sur un sous-domaine de la médecine, à savoir la pharmacologie cardiovasculaire, se limitait à des concepts qui n'avaient pas fait l'objet d'un traitement préalable dans la *Banque de terminologie du Québec (BTQ)*. Nous relatons les étapes les plus importantes de la recherche et décrivons les principales difficultés rencontrées. Nous proposons également des solutions fondées sur les lacunes observées.

Termes-clés :
terminographie, terminotique,
veille terminologique,
dépouillement assisté par ordinateur.

1 Introduction

Le terminologue peut tirer profit d'un nombre croissant d'outils informatiques qu'il utilise pour accélérer ou systématiser des tâches répétitives ou fastidieuses. Nous faisons référence ici aux logiciels de dépouillement assisté par ordinateur, aux concordanciers, aux systèmes de gestion de bases de données et aux logiciels d'édition de fiches terminologiques. Ces outils sont également utilisés par d'autres professionnels, comme les traducteurs ou les lexicographes, mais le terminologue en adapte certaines fonctions à ses besoins particuliers. En effet, celui-ci applique une méthode de recherche qui rend compte d'un certain nombre de concepts rattachés à un domaine de spécialité qui a été circonscrit au préalable.

Le présent article porte sur l'utilisation de certains de ces outils dans un contexte particulier, à savoir la veille terminologique. La *veille terminologique* est une activité qui consiste à repérer des termes « nouveaux » et à colliger des renseignements sur ces termes (contextes, notes et définitions) dans le but de les décrire. Nous adoptons une définition fonctionnelle de la « nouveauté d'un terme » inspirée de Perron (1997) qui s'appuie sur l'utilisation d'un corpus d'exclusion : un terme nouveau, ou *néonyme*, est un terme qui n'a pas fait l'objet d'une description terminologique dans un répertoire donné (le corpus

d'exclusion utilisé ici est la *Banque de terminologie du Québec (BTQ)*).

La recherche a été réalisée dans le cadre d'une entente avec l'Office de la langue française, organisme gouvernemental québécois. Nous devons nous pencher sur environ 400 concepts liés au domaine de la médecine, plus précisément à la radiologie et à la pharmacologie cardiovasculaire⁽¹⁾. Les fiches terminologiques produites étaient bilingues (anglais-français), mais nous nous concentrerons sur les aspects touchant les termes français liés au domaine de la pharmacologie cardiovasculaire. Notre équipe⁽²⁾ devait procéder en outre à l'expérimentation de différents outils informatiques et déterminer s'ils présentaient un intérêt dans un contexte professionnel. Les outils testés sont *Adept-Nomino* (Perron 1996) et *Sami-BTQ* (Pelletier 1996)⁽³⁾.

Nous présentons d'abord un panorama des principales étapes de la recherche, pour ensuite nous pencher sur les difficultés liées à chacune des

(1) Ces sous-domaines de la médecine ont été choisis en fonction de l'accessibilité de textes en format électronique. Toutefois, il aurait été préférable de travailler sur un seul sous-domaine et d'enrichir le corpus de textes. Une recherche portant sur deux sous-domaines simultanément présente des difficultés notables puisque le terminologue doit se familiariser avec des notions parfois éloignées. Nous n'aborderons pas ces difficultés dans le cadre du présent article, mais elles ont été présentes durant toutes les étapes de la recherche.

(2) Nous tenons à remercier les personnes qui, outre les trois auteurs, ont participé au projet de recherche : Anne Chudobiak, Claire-Hélène Lavigne, Sophie Milliard, Jean Ntakirutimana et Margreet de Rooij.

(3) Signalons par ailleurs que nous avons intégré d'autres logiciels plus largement répandus afin de prendre en charge certaines tâches périphériques (les logiciels en question sont cités au tableau 2).

étapes franchies. En conclusion, nous suggérons quelques avenues afin d'améliorer la recherche telle que nous l'avons réalisée.

2 Principales étapes de recherche

Les méthodes de recherche terminographiques font l'objet d'un consensus depuis déjà quelques décennies. Toutefois, l'intégration de certains outils informatiques vient bouleverser en partie ces méthodes. Nous présentons au tableau 1 les différences observées entre une recherche terminologique systématique «traditionnelle» et une recherche qui a recours à des outils informatiques (cette description est largement inspirée de Auger *et al.* (1991)).

Nous nous sommes alignées sur la démarche présentée dans la troisième colonne du tableau 1, tout en tenant compte des particularités liées à la veille terminologique (il en sera question un peu plus loin). À chaque étape de la recherche, nous avons intégré un logiciel donné ou fait appel à une fonction présente dans un logiciel : le tableau 2 dresse la liste des différents outils informatiques utilisés.

La recherche, comme on peut le constater, s'appuie, en dépit de l'intégration d'outils informatiques, sur des méthodes de recherche systématique (Auger et Rousseau 1978). L'aspect prédominant de ce type de recherche est la prise en compte de l'ensemble des termes d'un domaine : le terminologue définit, ultimement, les termes les uns par rapport aux autres. Cette stratégie permet de dégager les liens existant entre les concepts du domaine (hyponymie et hyperonymie, co-hyponymie, méronymie, etc.) et ce pour parvenir à dégager les

Tableau 1 :
Comparatif des méthodes de travail traditionnelles et des méthodes de travail intégrant des outils informatiques.

	<i>Méthode traditionnelle</i>	<i>Méthode intégrant des outils informatiques</i>
Recherche documentaire : Mise en forme du corpus	Collecte de textes de nature variée portant sur un domaine de spécialité.	1. Recherche et collecte de textes en format électronique portant sur un domaine de spécialité. 2. Conversion de textes papier en format électronique (numérisation).
Repérage des unités terminologiques	Repérage de termes au moment de la lecture de textes et transcription des contextes.	1. Identification automatique de termes (logiciel d'identification automatique des unités potentiellement terminologiques). 2. Première sélection humaine des unités produites par l'identification automatique qui feront l'objet de recherches ultérieures.
Relevé de contextes	Relevé de contextes lors de la lecture des textes. Raffinement de la liste de termes devant faire l'objet de descriptions.	1. Repérage automatique de contextes pour les termes retenus et sélection humaine des meilleurs contextes. 2. Raffinement de la liste de termes devant faire l'objet de descriptions.
Analyse des contextes	Étude des contextes relevés pour la production de fiches terminologiques. Information complémentaire puisée dans des ouvrages de référence.	1. Étude des contextes relevés pour la production de fiches terminologiques. 2. Information complémentaire puisée dans des ressources électroniques (par exemple, Internet) et, si nécessaire, dans des ouvrages de référence papier.
Rédaction de fiches terminologiques	Reproduction, dans une fiche terminologique, d'une partie des informations recueillies lors du dépouillement. Ajout d'information complémentaire. Reproduction de la fiche dans une banque de données terminologiques.	1. Copie, dans une fiche terminologique, d'une partie des informations recueillies lors du dépouillement. 2. Ajout d'information complémentaire. 3. Importation de la fiche dans une banque de données terminologiques.

Tableau 2 : Outils utilisés lors de la recherche

<i>Tâche</i>	<i>Procédé ou outil informatique pour la réaliser</i>
Mise en forme du corpus	Numérisation (numériseur pleine-page), reconnaissance optique de caractères (<i>Omnipage Direct</i>) et vérification orthographique pour la correction des textes numérisés (vérificateur de <i>Word</i>). Sélection de textes à partir de sites Internet
Identification automatique de termes	Analyse linguistique (<i>Nomino</i> ⁽⁴⁾) Listes de termes gérées dans un fichier <i>Excel</i>
Relevé et analyse des contextes	Sélection de contextes proposés par le concordancier (<i>Adepte</i>) Copie des contextes sélectionnés dans <i>Adepte</i> (système de gestion de bases de données intégré dans l'environnement <i>Adepte</i>) Lecture et analyse des contextes relevés dans <i>Adepte</i> pour chacun des termes (établissement de renvois pour les termes apparentés)
Recherche d'information complémentaire	Consultation d'ouvrages de référence sur support électronique (banques de terminologie, dictionnaires sur CD-Rom) Consultation de sites Internet En dernier recours, consultation d'ouvrages de référence sur support papier
Rédaction de fiches	Copie et collage dans <i>Sami-BTQ</i>

caractéristiques qui définiront les termes dénotant ces concepts.

Nous avons dû cependant tenir compte de certaines difficultés liées à la recherche terminologique menée en contexte de veille telle qu'elle a été définie à la section 1 (rappelons que nous avons écarté des termes déjà décrits dans le corpus d'exclusion, à savoir la *BTQ*).

Dans ce contexte, il est parfois difficile d'obtenir une vue d'ensemble

(4) *Adepte* et *Nomino* sont fondus dans le même environnement de travail. Toutefois, *Nomino* fait le dépouillement et les autres tâches sont réalisées au moyen de fonctions présentes dans *Adepte*. Cette distinction ne sera pas toujours faite dans ce qui suit.

du système conceptuel d'un domaine comparable à celle donnée par la recherche systématique véritable. Cette lacune est causée par le fait qu'une partie des termes du domaine peut déjà faire l'objet d'une description terminologique, alors qu'une seconde partie reste à décrire. D'autres lacunes sont entraînées par le fait que les traitements automatiques restent imparfaits et reposent exclusivement sur un corpus qui peut lui-même présenter des imperfections (par exemple, nous avons parfois exclu des termes pour lesquels le corpus ne livrait pas suffisamment de renseignements). Pour ces raisons, il arrive que le terminologue doive composer avec des listes de termes partielles.

Pour illustrer le sens de la dernière observation, nous avons reproduit au tableau 3 une partie de la liste de termes construits à partir de la tête *abcès* (ce sont les termes trouvés dans les textes du corpus utilisé dans le cadre de la recherche).

Tableau 3 :
Liste de termes dont la tête est *abcès*

<i>Termes trouvés dans le corpus</i>	<i>Présence dans la BTQ</i>
abcès appendiculaire	
abcès bilatéral	
abcès cérébral	BTQ
abcès de Brodie	BTQ
abcès du cerveau	BTQ
abcès du foie	BTQ
abcès du poumon	BTQ
abcès du psoas	BTQ
abcès encapsulé	
abcès épidural	
abcès intra-osseux	BTQ
abcès médullaire	
abcès nécrotique	
abcès paravertébral	
abcès para-vertébral	
abcès pelvien	
abcès péri-péritreux	

On remarque que, sur un total de 17 termes, sept sont déjà traités dans la *Banque de terminologie du Québec*. Donc, en principe, ces unités ne doivent pas faire l'objet d'un traitement dans un contexte de veille terminologique. Cependant, les «trous» signalés dans les listes peuvent soulever des problèmes au moment de l'analyse des renseignements recueillis.

Par exemple, les trous posent des problèmes du point de vue de l'appréhension d'un domaine de spécialité. Nous allons commenter ces difficultés au moyen d'un certain nombre de concepts liés au concept d'«action» (effet recherché par l'administration de médicaments) qui sont reproduits à la figure 1.

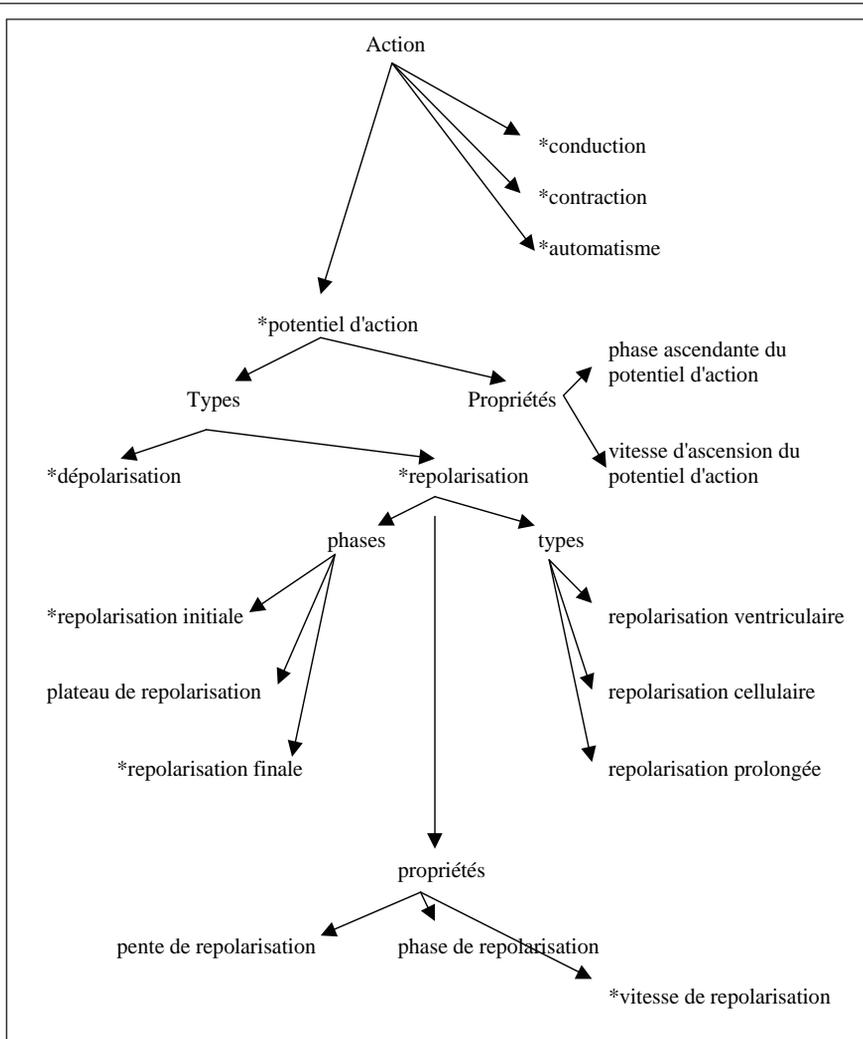


Figure 1: Concepts liés à « action »

Tableau 4: Termes construits à partir de *conduction* et de *conductibilité*

<i>Termes construits à partir de conduction</i>	<i>Termes construits à partir de conductibilité</i>
conduction tawarienne conduction cardiaque	conductibilité auriculaire conductibilité atrio-ventriculaire conductibilité cardiaque conductibilité ventriculaire conductibilité de ré-entrée

Tous les concepts présentés à la figure 1 n'ont pas fait l'objet d'un traitement pour différentes raisons (les concepts mis de côté sont signalés par l'astérisque). Nous présentons, ci-dessous, quelques causes pour lesquelles ils n'ont pas été retenus dans le cadre du projet.

Conduction: terme non relevé lors du dépouillement initial des textes français (qui accordait la priorité aux termes complexes). En revanche, certains hyponymes ont été retenus: *conduction cardiaque*, *conduction tawarienne*. De plus, certains termes faisant référence à la propriété de conduire les impulsions ont été traités: *conductibilité auriculaire*, *conductibilité atrio-ventriculaire*, *conductibilité cardiaque*, *conductibilité de ré-entrée* et *conductibilité ventriculaire*. On constate (tableau 4) le manque de parallélisme entre ces deux derniers groupes de concepts (qu'une véritable recherche systématique aurait corrigé).

Contraction: terme traité dans la *BTQ*. Ici encore, certains hyponymes ont été retenus: *contraction cardiaque*, *contraction myocardique*, *contraction spontanée*, *contraction tonique*, *contraction vasculaire*, *contraction ventriculaire*. De plus, certains termes étroitement liés à ces premières unités ont été traités dans le cadre du projet: *contractilité cardiaque*, *contractilité myocardique*, *contractilité ventriculaire*. De nouveaux parallélismes n'ont pas été saisis de façon appropriée (tableau 5).

Automatisme: terme non relevé lors du dépouillement des textes français. Certains hyponymes d'*automatisme* ont pourtant été traités: *automatisme auriculaire*, *automatisme cardiaque*, *automatisme ectopique*, *automatisme sinusal*. Un certain nombre de termes dénotant la propriété de générer des automatismes ont également été retenus: *automaticité ectopique*, *automaticité myocardique*, *automaticité*

Tableau 5 : Termes construits à partir de *contraction* et de *contractilité*

<i>Termes construits à partir de contraction</i>	<i>Termes construits à partir de contractilité</i>
contraction auriculaire contraction cardiaque contraction myocardique contraction tonique contraction vasculaire contraction ventriculaire	contractilité cardiaque contractilité myocardique contractilité ventriculaire

Tableau 6 : Termes construits à partir de *automatisme* et de *automaticité*

<i>Termes construits à partir de automatisme</i>	<i>Termes construits à partir de automaticité</i>
automatisme auriculaire automatisme cardiaque automatisme ectopique automatisme sinusal	automaticité ectopique automaticité sinusale automaticité ventriculaire

sinusale, automaticité ventriculaire.

Ici encore, des parallélismes n'ont pas été saisis de façon appropriée (tableau 6) :

Potentiel d'action: terme traité dans la *BTQ*;

Dépolarisation: terme traité dans la *BTQ*. Nous avons toutefois traité des hyponymes de ce terme :

dépolarisation diastolique, dépolarisation diastolique lente, dépolarisation spontanée, dépolarisation ventriculaire.

Repolarisation: terme traité dans la *BTQ*. Des hyponymes de ce terme ont été décrits: *repolarisation cellulaire, repolarisation en phases, repolarisation prolongée, repolarisation ventriculaire.*

Repolarisation initiale, repolarisation finale, vitesse de repolarisation: termes non relevés dans le corpus de départ.

Comme le montrent les exemples ci-dessus, seule une partie

des concepts fait l'objet d'un traitement terminographique. Cependant, tous les concepts d'un domaine doivent être appréhendés pour décrire les termes adéquatement. En contexte de veille, le terminologue doit lui-même combler les trous pour les concepts qui ne sont pas spécifiquement pris en charge dans le cadre d'un projet.

Ainsi, on peut penser, lors de la planification d'un projet de veille terminologique, qu'une description terminographique exigera moins de temps puisqu'elle ne porte que sur un nombre réduit de concepts. Toutefois, il n'en va pas tout à fait ainsi puisque les liens entre tous les concepts doivent être saisis. De plus, la systématisme, qui fait ordinairement la qualité d'une recherche terminographique, est perdue partiellement dans ce type de démarche.

3 Détail de la recherche

Nous allons maintenant nous pencher sur les étapes proprement dites de la recherche et sur certaines difficultés ponctuelles. Les sections qui suivent présentent les étapes décrites dans la troisième colonne du tableau 1. Précisons d'entrée de jeu que les différentes tâches sont décrites successivement dans le but de faciliter leur présentation. Dans les faits, cependant, les chevauchements et les aller et retour sont fréquents.

3.1 Mise en forme du corpus

Les termes à traiter devaient être extraits de textes en format électronique (qui étaient soumis, par la suite, à un logiciel de dépouillement, voir 3.2). Nous avons mis en forme un corpus de textes de pharmacologie cardiovasculaire qui comptait environ 300 000 mots⁽⁵⁾. Les textes sélectionnés avaient été numérisés lors d'un précédent projet de recherche, c'est-à-dire que les textes papier ont été convertis en format électronique.

Les textes du corpus devaient être rédigés en français, c'est-à-dire non traduits. Nous avons également pris soin de choisir des textes écrits par des auteurs différents pour constituer un corpus représentatif. À cette étape, nous n'avons pas mis l'accent sur la variété des niveaux de spécialisation des textes en tenant pour acquis qu'un échantillon suffisamment élaboré permettrait de constituer un corpus bien équilibré. Nous avons regroupé les textes accessibles en format électronique.

(5) Rappelons que nous avons travaillé sur un autre sous-domaine de la médecine, à savoir la radiologie; ces textes ont été téléchargés à partir de sites Internet.

Cependant, lors du traitement des termes, nous avons rencontré des difficultés importantes qui découlent des choix faits au départ. Il aurait été primordial, dès la première étape, de sélectionner des textes qui respectent des critères très précis.

En veille terminologique, les documents qui forment le corpus de dépouillement devraient contenir la plupart des renseignements essentiels à la rédaction d'une fiche puisque, en principe, ces renseignements ne se trouvent nulle part ailleurs (dans les ouvrages de référence, par exemple). De plus, le succès des traitements automatiques repose en très grande partie sur la qualité des textes formant le corpus de départ. Voici les principaux critères auxquels les documents devraient normalement répondre :

1. Une certaine uniformité des thèmes abordés dans un groupe de documents ;
2. Une documentation « équilibrée » : c'est-à-dire suffisamment pointue pour contenir des renseignements susceptibles d'éclairer les concepts terminologiques appartenant à un domaine relativement restreint et suffisamment accessible pour permettre au terminologue d'en tirer des définitions qui s'adressent à un vaste public ;
3. Des documents comportant un nombre d'occurrences suffisamment élevé des termes faisant l'objet d'un traitement ;
4. Un nombre élevé de textes de provenances diverses, afin d'accéder à des attestations des variantes terminologiques (synonymes, abréviations, etc.).

Ainsi, les documents intégrés dans un corpus de départ doivent faire l'objet d'une sélection rigoureuse. Cette observation reste valable peu importe le type de recherche terminologique entreprise – traditionnelle ou informatisée –, mais le problème est amplifié par le fait qu'en recherche informatisée on tente

de réunir des sources en format électronique qui ne répondent pas nécessairement à tous ces critères⁽⁶⁾. De plus, de nombreux textes présentant ces caractéristiques « idéales » (des manuels didactiques, des encyclopédies spécialisées) sont rarement en format électronique. Le terminologue peut numériser les textes idéaux, mais la numérisation devient rentable dans la mesure où le projet en est un de grande envergure.

À notre avis, la constitution d'un corpus électronique (dans une optique de recherche en contexte de veille) devrait se dérouler de la manière suivante :

1. L'identification des objectifs de la recherche (domaine, langue, nombre de concepts à décrire, mode de description des concepts) ;
2. Le repérage de documents de types variés portant sur le domaine choisi ;
3. La vérification de l'accessibilité des textes repérés en format électronique (copie de textes accessibles, conversion, si nécessaire, de textes intéressants en format électronique).

3.2 Identification automatique de termes

L'ensemble des textes sélectionnés ont été soumis au logiciel *Nomino* pour en extraire le contenu terminologique⁽⁷⁾. *Nomino* propose des listes de mots simples et de syntagmes nominaux à partir desquelles le terminologue sélectionne les unités qui feront l'objet d'un traitement terminographique.

Nous aborderons, dans les sections qui suivent, quelques difficultés soulevées par l'utilisation d'un logiciel de dépouillement. Nous décrirons d'abord les imperfections des listes proposées par le logiciel et leurs répercussions sur le travail terminographique (3.2.1). Nous nous pencherons par la suite sur deux méthodes d'utilisation des listes (3.2.2).

Dans le cadre du projet, le temps consacré à l'interprétation des listes proposées par *Nomino* a été prolongé en raison du fait qu'il fallait nous assurer que les termes retenus n'avaient pas fait l'objet d'une description préalable dans la *BTQ*. La consultation de la banque a dû être faite ponctuellement pour chacune des unités potentiellement intéressante puisqu'aucune procédure automatique n'allégeait cette tâche.

3.2.1 Listes proposées par le logiciel de dépouillement *Nomino*

Comme plusieurs autres logiciels de dépouillement, *Nomino* génère des listes de mots et de syntagmes nominaux qui sont imparfaites du point de vue de l'utilisateur : dans le cas qui nous préoccupe, les listes comportaient un grand nombre d'entrées qui n'ont pas fait l'objet d'un traitement terminographique. Le tableau 7 présente les principaux problèmes relevés dans les listes de termes complexes⁽⁸⁾.

(6) On a par ailleurs souvent tendance à exagérer l'intérêt que peuvent présenter les ressources trouvées sur Internet. Il est vrai que le *Web* offre une masse documentaire colossale mais l'expérience a montré qu'il peut se révéler difficile d'isoler, d'une part, puis de réunir, d'autre part, un nombre suffisamment élevé de documents de qualité et rédigés dans une autre langue que l'anglais (ce fait s'observe dans un certain nombre de domaines).

(7) Pour une présentation détaillée du fonctionnement du logiciel, voir Perron (1996).

(8) *Nomino* propose également des listes de mots simples. Bien que ces listes proposent également des unités terminologiques, la proportion de bruit est nettement plus élevée.

Tableau 7: Principales causes de bruit et exemples tirés des textes dépouillés⁽⁹⁾

Problèmes graphiques (fautes de typographie contenues dans le texte traités)	• <i>sur la conduction, effet indésirable</i>
Syntagme nominal non terminologique (syntagme construit à partir d'un nom mais qui n'est pas une unité terminologique)	<i>conduction au niveau des synapses des ganglions sympathiques, effet additionnel, effet bien connu; hypotension avec des variations; insuffisance préexistante</i>
Terme n'appartenant pas au domaine	<i>conduite automobile, année de l'enfant</i>
Mot mal catégorisé (un mot dans le syntagme relevé est mal analysé: verbe traité comme un nom, par exemple)	<i>contraction cardiaque cesse, hypertension artérielle comporte; insuffisance doit; résistance vasculaire pulmonaire montre</i>
Collocation plutôt qu'unité terminologique (groupe de mots ayant des affinités et qui apparaissent fréquemment dans les textes, mais qui ne sont pas des termes complexes)	<i>effet de la drogue, effet des médicaments, traitement de la migraine, traitement de l'angine de poitrine</i>
Mauvais découpage (le terme apparaissant dans le texte est soit plus long, soit plus court que le terme proposé dans la liste)	<i>hypertension à activité: hypertension à activité rénine basse; médicament à action: médicament à action alphabloquante</i>

De plus, même si les listes proposées par le logiciel de dépouillement étaient parfaites, c'est-à-dire qu'elles ne comportaient que des unités terminologiques valables compte tenu d'un projet donné (ce qui est loin d'être le cas), elles soulèvent des difficultés du point de vue de la démarche normalement adoptée par le terminologue. Ce dernier doit désormais procéder à partir d'une liste de formes linguistiques et en extraire le contenu conceptuel pertinent pour rédiger une définition et établir des liens entre concepts apparentés. Ce contenu conceptuel est moins immédiatement perceptible à partir d'une liste de

syntagmes nominaux proposés par un logiciel de dépouillement. (Nous verrons plus loin qu'une méthode permet toutefois de simuler la démarche conceptuelle.)

3.2.2 Utilisation des listes proposées par le logiciel de dépouillement Nomino

Le terminologue peut aborder les listes de syntagmes nominaux de deux manières. La première méthode consiste à procéder à un balayage systématique des entrées afin d'écartier toutes les unités ne devant pas faire l'objet d'une recherche plus poussée et de conserver les entrées qui

semblent intéressantes. Le terminologue doit d'abord s'assurer que l'unité proposée est effectivement un terme (certaines entrées sont éliminées rapidement; d'autres, au contraire, sont plus difficiles à traiter, car seule la lecture du contexte permet de prendre une décision éclairée). De plus, si la recherche est réalisée dans un contexte de veille terminologique, il doit s'assurer que le terme n'apparaît pas dans la BTQ. Nous avons testé cette technique sur une partie d'une liste de syntagmes nominaux générée par *Nomino*: nous avons éliminé environ 90 % des entrées de cette manière. Ce travail est très mécanique et peu satisfaisant, puisqu'on doit consacrer beaucoup de temps à isoler environ 10 % de termes dont l'intérêt n'est pas nécessairement assuré à cette étape.

La seconde méthode consiste à repérer, dans les listes proposées par le logiciel, des blocs de termes potentiellement intéressants sans tenir compte de l'ensemble des entrées dans les listes. Cette dernière méthode est nettement plus intéressante pour des motifs qui ressortiront clairement dans ce qui suit. Le terminologue travaille sur une série de termes apparentés sur le plan de la forme ou du sens, ou se trouvant dans un même environnement conceptuel⁽¹⁰⁾. Il utilise, comme point de départ, une série de termes construits à partir de la même tête.

(9) Les éléments entraînant du bruit varient en fonction de la stratégie d'identification adoptée par le logiciel et, parfois, de la langue traitée. De plus, la proportion de bruit dépend de l'application visée (pour de plus amples renseignements, voir Lauriston (1994); L'Homme *et al.* (1996) et Otman (1991)).

(10) Cette méthode a été suggérée dans Perron (1996).

Voici un exemple de la méthode utilisée dans les textes que nous avons dépouillés. Nous savons que les médicaments exercent un certain nombre d'«actions». Nous utilisons les termes construits à partir d'*action* comme point de départ d'un dépouillement. Nous avons reproduit dans le tableau 8 une partie de la liste des unités complexes proposées par *Adepto-Nomino* dont la tête est *action*:

Tableau 8: Liste partielle des termes dont la tête est *action*

Termes complexes
action adrénérurgique
action adrénolytique
action agoniste
action alpha
action alpha-adrénérurgique
action alpha-adrénolytique
action alpha-agoniste
action alpha-agoniste adrénérurgique
action alphabloquante
action alphabloquante accessoire
action alphabloquante non spécifique
action analgésique
action analogue
action anesthésique
action anesthésique locale
action anorexigène
action antagoniste

Un premier repérage permet d'effacer les unités complexes relevées par *Adepto-Nomino*, mais qui ne sont pas des termes. Par exemple, *action analogue* a été écarté de la liste à la suite d'une première lecture. De même, *action alpha* a été exclu, car l'examen des contextes a permis de constater qu'*alpha* était toujours placé devant un adjectif (*alpha-adrénérurgique*, *alpha-adrénolytique*, etc.).

Ensuite, un recouplement avec les termes de la *BTQ* permet d'éliminer les unités ayant déjà fait l'objet d'une description terminologique. Aucun des termes ci-dessus n'apparaissait dans la *BTQ*, mais *action coagulante*

(listé un peu plus loin) a été repéré et, par conséquent, écarté de la liste.

Les éléments qui restent peuvent en principe faire l'objet d'une recherche. Le terminologue peut, par la suite, utiliser les contextes dans lesquels ces termes apparaissent pour repérer d'autres termes liés sur le plan conceptuel. Les contextes affichés contiendront en effet d'autres termes qui seront utilisés comme point de départ pour une autre recherche. Par exemple, une recherche sur *action adrénérurgique* donne accès au contexte suivant:

La plupart des bêta bloquants sont antagonistes des actions adrénérurgiques et au niveau des récepteurs bêta 1 et au niveau des bêta 2.

Ce contexte révèle la présence d'autres termes, à savoir *bêta bloquant*, *récepteur bêta 1* et *récepteur bêta 2*. Le terminologue peut effectuer des recherches sur ces termes qui serviront eux-mêmes de pistes pour d'autres unités et ainsi de suite jusqu'à ce qu'il épuise le filon. Cette méthode a permis de repérer, dans les textes de pharmacologie, des termes associés aux familles conceptuelles suivantes:

- Noms de médicaments ou de substances biochimiques (ex.: *alcaloïde de l'ergot*; *barorécepteur aortique*; *céliprolol*);
- Propriétés ou mécanismes d'action des médicaments (ex.: *absorption digestive*; *action adrénolytique*)
- Maladies et symptômes (ex.: *angor primaire*; *bradycardie réflexe*; *déficit enzymatique*);
- Propriétés ou mécanismes d'action des parties du corps (ex. *automaticité ectopique*, *conductibilité atrio-ventriculaire*, *débit coronarien*);
- Termes anatomiques (ex.: *ganglion végétatif*; *innervation parasymphatique*).

Ainsi, la démarche qui consiste à procéder à partir de termes apparentés est moins contraignante et plus satisfaisante pour le terminologue. C'est également celle qui se rapproche

le plus des méthodes de travail traditionnelles.

3.3 Extraction et analyse des contextes

Adepto comporte un concordancier qui affiche, pour toutes les unités présentes dans les listes, les phrases dans lesquelles ces unités apparaissent. Les contextes peuvent être sélectionnés ou simplement consultés. Les contextes sélectionnés sont copiés, à la demande de l'utilisateur, dans une fiche recevant l'ensemble des contextes choisis pour un terme.

À l'étape de l'analyse des contextes, nous nous sommes aperçu que les contextes extraits des corpus ne comportaient pas l'ensemble des éléments essentiels à la rédaction de définitions. Étant donné que les corpus étaient constitués de textes très spécialisés, c'est-à-dire destinés à des spécialistes du domaine, il a fallu puiser des renseignements complémentaires dans d'autres sources plus accessibles: par exemple, des manuels pédagogiques de format papier destinés à des étudiants en médecine ou encore des sites Internet qui s'adressent à la population en général pour vulgariser certaines connaissances scientifiques. Il aurait fallu choisir, et ce dès le début, des textes de type pédagogique ou informatif grand public pour éviter un surplus de travail à cette étape.

Même si on ne peut pas toujours en tirer des définitions proprement dites, les contextes contiennent de nombreux éléments qui permettent de cerner les concepts décrits. Nous avons reproduit quelques exemples ci-dessous:

- Éléments définitoires

Absorption intestinale: les graisses, insolubles dans l'eau, sont transformées en micelles composées d'acides gras, de monoglycérides et de sels biliaires.

(Notion de transformation liée à l'absorption intestinale)

Les bêtabloquants n'ont, aux doses thérapeutiques, qu'une action adrénolytique, qui déplace le tonus neurovégétatif en faveur du parasympathique, ralentissant le nœud sinusal et freinant la conduction auriculoventriculaire. (Effet de l'action adrénolytique exercée par certaines substances)

– Relations explicites avec des concepts connexes

Certains neuroleptiques tels que l'halopéridol ou la chlorpromazine ont une action alphabloquante non spécifique. D'autres neuroleptiques tels que le sulpiride, la thioropérazine, ou le pimozide possèdent une action alphabloquante assez marquée à prédominance alpha 2 présynaptique qui pourrait expliquer leur profil «désinhibiteur», probablement lié à la libération accrue de catécholamines (fig. 34).* (Médicaments qui exercent une action alphabloquante).

Dissocier l'action de ces produits en une action antiarythmique et une action antifibrillante est plus ou moins légitime, car les deux ne sont pas sans relation. (Relation entre deux types d'actions voisines).

Le mécanisme de l'action hypotensive des hydrazinophthalazines est encore très mal connu. Certains auteurs ont supposé qu'elles inhibent directement les centres vasomoteurs. (Effets de l'action hypotensive).

Les contextes sont également utiles pour dénicher d'autres formes de renseignements :

– Attester certaines relations synonymiques établies dans les autres sources consultées (par exemple, la série synonymique suivante : *nœud atrioventriculaire, nœud auriculoventriculaire, nœud d'Aschoff-Tawara, nœud atrioventriculaire de Tawara, nœud d'Aschoff et Tawara* peut être établie à la suite d'une consultation combinée du corpus et de la documentation complémentaire) ;

– Vérifier l'exactitude des renseignements trouvés dans les autres sources consultées (par exemple, de tels renseignements pouvaient se révéler plus ou moins pertinents, être trop généraux ou porter sur une notion médicale différente) ;

– Fournir de l'information plus technique qui pouvait faire l'objet de note (ex. : durée d'une action exercée par un médicament, effets indésirables provoqués par une substance, affections traitées par un médicament, etc.) ;

– Donner le genre d'un terme, notamment d'un terme dénotant un médicament ;

– Fournir l'explicitation d'un sigle (ex. : *MIBG* est expliqué dans un contexte entre parenthèses) :

Une scintigraphie à la MIBG (méthyl iodo benzo guanidine) est réalisée en cas de suspicion de neuroblastome.

3.4 Recherche d'informations complémentaires

Comme les contextes extraits lors des premières étapes de la recherche se sont révélés insuffisants, il a fallu nous tourner vers d'autres ressources documentaires. Les renseignements recueillis étaient versés au fur à mesure dans une fiche terminologique (gérée par *Adepté*).

À cette étape, nous avons utilisé deux types de catégories de ressources documentaires, à savoir les ouvrages de référence, monographies, revues et encyclopédies (sur support informatique ou papier) et les sites Internet.

Les documents du premier groupe (dictionnaires et ouvrages spécialisés) ont aidé à produire un certain nombre de définitions. Toutefois, la recherche dans certains ouvrages requiert beaucoup de temps. Mentionnons également que les ouvrages de référence (dictionnaires, banques de terminologie) ont rarement permis de retrouver les

termes recherchés tels quels (ce que nous pouvions anticiper, puisqu'il s'agissait de néonymes). Toutefois, ils présentaient le plus souvent des renseignements d'ordre conceptuel sur des termes apparentés (hyponymes, hyponymes ou co-hyponymes).

L'Internet s'est révélé particulièrement utile parce qu'il permet d'atteindre l'information spécialisée rapidement. Il héberge un grand nombre de sites qu'on pourrait qualifier de «généralistes» comme des compilations de revues médicales, des listes de résumés scientifiques, etc. (voir Michel 1998 pour une présentation d'un certain nombre de sites médicaux).

Lorsque les recherches dans les premiers sites se révélaient infructueuses, nous tentions de localiser d'autres sites moins largement connus au moyen des moteurs de recherche (ex. : *Altavista, Copernic, Yahoo*, etc.). Les mécanismes de recherche, bien que parfois rudimentaires (la recherche se fait par simple correspondance entre chaînes de caractères), permettent d'isoler des documents web à partir d'un terme simple ou complexe. Signalons que les requêtes sur les termes complexes ont livré les résultats les plus intéressants. Par ailleurs, nos efforts ont été facilités par le fait que la recherche portait sur des termes médicaux construits à partir de racines grecques et latines. L'ambiguïté souvent générée par les recherches simples sur des chaînes de caractères a été réduite considérablement pour cette raison.

Lorsque les recherches de termes complexes se révélaient vaines, nous reprenions une nouvelle recherche à partir d'une composante d'un terme, notamment à partir de sa tête. Nous repérions fréquemment au moyen de cette méthode des contextes ou des informations sur l'hyperonyme. Fréquemment, ces contextes présentaient des informations sur le modificateur – même s'il n'était pas

mentionné explicitement – qui nous aidaient ainsi à saisir le concept dans sa totalité. Toute information pouvant aider potentiellement la description d'un concept était retenue et stockée.

Comme dans bien d'autres domaines, les sites Internet anglais sont nettement plus nombreux et riches en renseignements. Nous avons repéré ces sites de langue anglaise à partir de l'équivalent anglais d'un terme donné lorsque nous l'avions déniché dans un répertoire (pour les équivalents non relevés tels quels, nous procédions à partir des composantes d'un terme complexe). D'autres termes se sont révélés plus difficiles à traiter de cette manière, soit parce que nous n'avions pas relevé d'indices sur leur traduction, soit parce que l'équivalent anglais repéré avait d'autres acceptions non médicales. Dans les faits, ce sont les sites anglais qui nous ont permis de rédiger la plupart des définitions.

La recherche d'information complémentaire a été beaucoup plus ardue qu'on ne l'aurait pensé au départ. Cette étape a souffert des imperfections liées à la constitution du corpus.

3.5 Rédaction des fiches

Les fiches ont été saisies dans un logiciel d'édition de fiches terminologiques, *Sami-BTQ*⁽¹¹⁾, à partir des renseignements extraits des autres sources documentaires reproduites dans *Adepto-Nomino*. *Sami-BTQ* permet de conserver des versions provisoires et de les modifier tant que les versions définitives ne sont pas prêtes. De plus, ses mécanismes de recherche permettent d'atteindre des termes apparentés et la *BTQ* peut être consultée par l'utilisateur à tout moment.

(11) Pour une présentation détaillée des fonctions du logiciel, voir Pelletier 1996.

Nous avons, dans certains cas, copié les informations utiles à partir d'*Adepto* et nous les avons collées dans *Sami* (les entrées françaises et anglaises). Quant aux définitions et aux notes, elles étaient rédigées directement dans *Sami*.

Sami peut importer directement des fiches produites dans *Adepto*, ce qui évite d'avoir à reproduire les rubriques ou à les coller dans les champs prévus à cette fin. Nous n'avons pas eu recours à cette stratégie puisque nos entrées *Adepto* comportaient de nombreux renseignements dont la saisie s'est

faite sans véritable gestion de l'information dans les champs. En outre, nous trouvions plus utile de conserver tous les renseignements retenus que d'éditer les fiches de dépouillement afin de les rendre compatibles avec la structure des fiches *Sami-BTQ*. Certains renseignements apparaissant dans la fiche de dépouillement qui décrivait un premier terme ont servi à traiter d'autres termes liés sur le plan conceptuel. La figure 3 présente quelques fiches rédigées au terme de la recherche.

Figure 5: Fiches de terminologie

```

1. Conduction cardiaque
FICHER TRAVAIL
005 q010274
025 V7818
030 PUMS
035 UM
040 9812
060 médecine|
105 cardiac conduction
110 E; a
205 conduction cardiaque
210 P5; n f; F; b
215 Propagation d'une impulsion nerveuse qui prend naissance dans le nœud
auriculo-ventriculaire et qui se propage dans les artères entraînant la contraction
de ces dernières. |c|
295 SOURCE ANNÉE VOLUME NUMÉRO MOIS PAGE MENTION
a 605268 1998 12
b 603134 1998 12
c 603135 1998 12

2. Dépolarisation diastolique
FICHER TRAVAIL
005 q011024
025 V7818
030 PUMS
035 UM
040 9809
060 médecine|
105 diastolic depolarization
110 E; a
130 1- phase-four depolarization |E; a|
205 dépolarisation diastolique
210 P5; n f; F; a
215 Diminution de la polarisation de la diastole, c'est-à-dire du mouvement de
dilatation du cœur et des artères, qui alterne avec le mouvement de contraction
(systole). |b|
295 SOURCE ANNÉE VOLUME NUMÉRO MOIS PAGE MENTION
a 603134 1998 10
b 603135 1998 10

```

4 Suggestions d'améliorations à la méthode de travail

En guise de conclusion, nous rappellerons les principales étapes abordées dans le présent article en mettant l'accent sur les lacunes que nous avons observées. Nous suggérons également des éléments qui permettraient d'améliorer la méthode de travail.

Premièrement, au cours des différentes étapes de la recherche, nous nous sommes rendu compte que les textes sélectionnés n'étaient pas parfaitement adaptés à la description de 400 termes complexes du domaine médical. Nous avons donc passé beaucoup de temps à rechercher de l'information complémentaire pour mobiliser les renseignements nécessaires à la préparation des fiches terminologiques.

La recherche aurait gagné en efficacité si les objectifs avaient été bien définis au préalable et si le corpus avait été constitué de façon à les atteindre. Un corpus équilibré peut se révéler difficile à constituer dès le départ, mais ces efforts permettraient d'accroître la productivité, notamment en réduisant le temps consacré à la recherche complémentaire.

Deuxièmement, le dépouillement des textes a reposé sur un traitement automatique (par le logiciel *Nomino*) et nous avons accordé la priorité aux termes complexes. Toutefois, comme nous l'avons vu, l'acquisition de connaissances nécessaires à la description d'un terme complexe passe d'abord et fréquemment, par l'acquisition de connaissances sur un autre terme. Or, comme tous les termes du domaine ne faisaient pas nécessairement l'objet d'un traitement, notre recherche perdait en systématisme. Le travail véritablement terminologique est par excellence

systématique: un domaine de la connaissance humaine ne peut être appréhendé adéquatement que de façon systématique.

Nous croyons dès lors que la décision de réaliser ou non une description systématique doit être prise lors de la planification du travail et cela est directement lié au profil de l'utilisateur qui aura accès à cette terminologie. Ce dernier doit déterminer l'orientation que doit adopter une recherche terminographique. Le terminologue canalise et décrit les concepts d'un domaine et, de ce fait, procède à sa délimitation.

Enfin, le projet a permis de relativiser la croyance pourtant bien répandue voulant que l'utilisation d'outils informatiques accroisse considérablement la productivité. Il est vrai que certains traitements automatiques allègent nettement des tâches fastidieuses. Toutefois, ils modulent de façon sensible des méthodes de travail implantées depuis longtemps. La chaîne de travail terminographique, par exemple, devient plus séquentielle et le terminologue doit s'adapter à ces nouvelles stratégies. Le gain de productivité se fait sentir pour des projets de grande envergure et non pour de petits projets.

*Marie-Claude L'Homme,
Claudine Bodson,
Renata Stela Valente,
Département de linguistique
et de traduction,
Université de Montréal.*

Bibliographie

Auger (P.), Drouin (P.) et L'Homme (M.C.), 1991: «Un projet d'automatisation des procédures en terminographie», *Méta* 36 (1), pp. 121-127.

Auger (P.) et Rousseau (L.-J.), 1978: *Méthodologie de la recherche terminologique*. Québec: L'Éditeur officiel du Québec.

Bergenholtz (H.) et Tarp (S.), 1995: *Manual of Specialised Lexicography*, Amsterdam / Philadelphia: John Benjamins.

Boulanger, (J.C.), 1979: «Problématique d'une méthodologie d'identification des néologismes en terminologie», dans *Néologie et lexicologie. Hommage à Louis Guilbert*, coll. Langue et langage, Paris: Librairie Larousse, pp. 36-46.

Duplain (J.) et coll., 1997: *Guide d'utilisation de la fiche de terminologie*, édition mise à jour, Québec: Gouvernement du Québec.

Lauriston (A.), 1994: «Automatic Recognition of Complex Terms: The TERMINO Solution», dans *Terminology* 1(1), pp. 147-170.

L'Homme (M.C.), 1999: *Évaluation de méthodes d'informatisation de la chaîne de travail terminographique (en vue de la description de néologismes liés au domaine de la médecine). Rapport d'évaluation présenté à l'Office de la langue française* (avec la collaboration de C. Bodson, A. Chudobiak, C.H. Lavigne, S. Milliard, J. Ntakirutimana et R. S. Valente).

L'Homme, (M.C.), 1998: *Expérimentation du logiciel Adepto-Nomino en contexte de veille terminologique. Rapport d'évaluation présenté à l'Office de la langue française* (avec la collaboration de M. de Rooij).

L'Homme, (M.C.), 1996: «Definition of an Evaluation Grid for Term-extraction software», dans *Terminology* 3(2), pp. 291-312.

Meyer (I.) et Mackintosh (K.), 1996: «The Corpus from a Terminographer's Viewpoint», dans *International Journal of Corpus Linguistics* 1(2), pp. 257-285.

Michel (F.), 1998: «Banques virtuelles de textes scientifiques, une aide possible à la veille terminologique?», dans *Terminologies nouvelles* 18, pp. 5-10.

Otman (G.), 1991: «Des ambitions et des performances d'un système de dépouillement terminologique assisté par

ordinateur», dans *La banque des mots* 4 (numéro spécial), pp. 59-96.

Pelletier (S.), 1996: «Le système *Sami-BTQ*: un système intégré de production et de mise à jour des données de la Banque de terminologie du Québec», dans *Terminologies nouvelles* 15, pp. 66-76.

Pelletier (S.), 1995: *Sami. Système d'alimentation et de mise à jour de la Banque de terminologie du Québec*. Québec: Office de la langue française.

Perron (J.), 1996: «*Adepto-Nomino*: un outil de veille terminologique», dans *Terminologies nouvelles* 15, pp. 32-47.

Sager (J.C.), 1990: *A Practical Course in Terminology Processing*. Amsterdam / Philadelphia: John Benjamins.

Sinclair (J.), 1991: *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

Néologie traductive

Pour accélérer la mise à disposition des néologismes aux traducteurs et rédacteurs, le Réseau international de néologie et de terminologie a mis sur pied une action de veille néologique visant à consigner, puis à diffuser plus rapidement les termes nouveaux par le biais d'un réseau francophone d'échange de néologismes sur Internet, *Balnéo* (Boîte Aux Lettres – NÉOlogie).

Menée en Communauté française de Belgique, une recherche sur la néologie traductive s'intègre dans ce projet du Rint. Le présent article en fait la description.

Termes-clés:
néologie, traduction,
veille terminologique.

1 Introduction

Le décalage entre l'apparition des termes et leur consignation dans les dictionnaires constitue un problème pour les traducteurs, rédacteurs et spécialistes des domaines techniques, scientifiques et sociaux. Cette partie du vocabulaire spécialisé et de la langue commune fait régulièrement défaut dans les dictionnaires sur papier, mais également dans les dictionnaires et banques de données terminologiques sur support électronique (disquettes, disque optique compact et en ligne), même si ces derniers peuvent être mis à jour plus facilement et plus rapidement.

Pour accélérer la mise à disposition des néologismes aux traducteurs et rédacteurs, le Réseau international de néologie et de terminologie a mis sur pied une action de veille néologique visant à consigner, puis à diffuser plus rapidement les termes nouveaux par le biais d'un réseau francophone d'échange de néologismes sur Internet, *Balnéo* (boîte aux lettres – néologie).

Menée en Communauté française de Belgique, une recherche sur la néologie traductive s'intègre dans ce projet du Rint. Elle vise la création d'un réseau de traducteurs francophones belges qui s'engagent à signaler les néologismes qu'ils créent eux-mêmes ou qu'ils ont rencontrés au cours de leur travail. Ces

néologismes sont soumis à l'étude puis diffusés.

2 La néologie traductive

Dans les différents domaines de la vie sociale, les usagers créent des termes au fur et à mesure de leurs besoins. Le lexique s'enrichit ainsi continuellement, notamment pour dénommer ou désigner de nouvelles réalités.

2.1 Néologie primaire et néologie traductive

On peut distinguer deux types de néologie: celle où la formation d'un nouveau terme, dans une langue précise, accompagne la formation d'un nouveau concept et celle où le terme existe déjà dans une langue et où un nouveau terme est créé dans une autre langue. La situation typique dans laquelle se déroule le premier processus est la situation de travail (le laboratoire de recherche, la fabrication de nouveaux produits, etc.). Le contexte classique de la deuxième forme de néologie est la traduction. Appelons «néologie primaire» la première forme et la seconde, «néologie traductive». En effet, la néologie traductive peut être le fait, sporadique ou plus ou moins systématique, d'instances chargées de terminologie. Ces instances proposent souvent des équivalents pour des termes qui circulent déjà dans la langue d'arrivée sous forme d'emprunts, de calques ou de termes jugés mal formés.

Cependant, l'environnement le plus ancien et le plus naturel de néologie traductive, c'est-à-dire la formation et l'introduction de nouveaux termes qui ont déjà un précédent linguistique, est la traduction. C'est que les traducteurs sont les premiers à être confrontés aux nouveaux termes en LS (langue-source) et aux nouveaux concepts, pour lesquels leur métier les contraint à proposer un équivalent dans la LC (langue-cible).

Nous appelons donc *néologie* la création d'un nouveau terme par un traducteur et, pour la facilité de notre propos, *néologisme* un nouveau terme, proposé dans une traduction, tout en sachant que le débat autour du concept de néologie/néonymie est, aujourd'hui encore, plus ouvert que jamais. Il y a un contraste important entre la constatation de J. Rey-Debove en 1971 («un néologisme est un mot récemment utilisé dans les échanges et absent des corpus métalinguistiques») et l'analyse très fine de F. Cusin-Berche en 1998 (la perception du néologisme est aussi liée au sentiment linguistique du locuteur, à qui un mot «semble» nouveau, au degré de diffusion du terme dans la communauté linguistique et au fait que le néologisme peut se fonder sur la traduction d'idées non originales, mais exprimées de façon inédite, pour exprimer une certaine vision du monde).

2.2 Les traducteurs-néographes

Quoiqu'un traducteur isolé ne crée pas quotidiennement des néologismes et ne s'occupe que d'une partie de la néologie, à savoir la néologie terminologique ou dénominative, l'ensemble du monde de la traduction déploie une activité néographique variée et multiple, surtout dans les domaines où la néologie primaire est elle-même

foisonnante. Les traducteurs ont, en effet à produire dans la langue d'arrivée, un texte ayant les mêmes fonctionnalités que le texte dans la langue de départ. La valeur liée au néologisme, souvent signalée dans le microcontexte de la langue de départ, c'est-à-dire dans l'ensemble des informations véhiculées par la phrase ou le paragraphe, nécessite souvent un néologisme parallèle dans la langue d'arrivée. Une périphrase ferait disparaître la fonctionnalité du néologisme (Wijnands 1989: 2), on l'évitera donc dans la mesure du possible.

2.3 La traduction, lieu stratégique pour la néologie

Les néologismes créés par les traducteurs apparaissent immédiatement dans un contexte réel de communication, à savoir l'environnement qui favorise une propagation naturelle des nouveaux termes. L'audience du traducteur est celle du commanditaire de la traduction, qui peut être importante. Les néologismes proposés par les traducteurs circulent immédiatement dans des écrits, support qui semble, pour les termes techniques et scientifiques, la voie de pénétration principale des néologismes, une voie qui a le plus grand potentiel d'acceptabilité dans une communauté linguistique et qui assure donc d'emblée une légitimation, une valorisation et une consécration du nouveau terme.

Souvent pressé par le temps, le traducteur ne peut pas consacrer beaucoup d'énergie à la recherche du néologisme correct. Aussi, plus souvent qu'il ne le souhaiterait, il recourt à des emprunts ou à des périphrases (Hermans 1991: 3). Les traducteurs ont néanmoins conscience de leur rôle de néographes. Selon une enquête menée en Belgique, au Canada, en Suisse et au

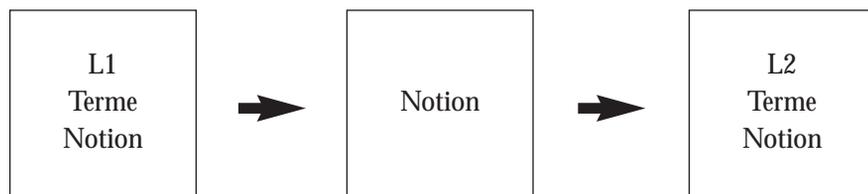
Luxembourg (Benhamida 1993), 76% des traducteurs estiment qu'ils ont un rôle à jouer dans les processus de diffusion des néologismes bien formés.

2.4 Les principes de la néologie traductive

a) Un premier principe de traduction applicable à la terminologie peut être formulé ainsi: on ne traduit pas d'une langue dans une autre. Le traducteur ne cherche pas systématiquement des équivalents pour tous les termes du texte à traduire (il ne traduit jamais mot à mot). Il crée des termes s'ils sont utiles pour la transmission correcte du message. La néologie traductive obéit en premier lieu aux principes régissant la traduction elle-même. La première obligation du traducteur n'est pas l'équivalence des termes, mais celle du message.

Le néologisme passe donc par le prisme du système notionnel avant de passer par le prisme du système de la langue. Dans le cas de la traduction par équivalence des termes, qui joue un rôle plus important dans la traduction spécialisée que dans la traduction de textes littéraires, le traducteur ne cherche pas d'emblée à traduire le terme. Il identifie la notion exprimée par le terme du texte de départ et réexprime ensuite la notion dans le texte traduit. La question qu'il se posera est donc celle-ci: quel est le terme dans la langue-cible dont le sens correspond exactement à la notion exprimée par le terme dans la langue-source? L'adéquation, c'est-à-dire la qualité qu'a un terme de bien convenir à la notion qu'il exprime, dans le contexte précis du texte à traduire, est donc l'exigence principale.

Le traducteur sait en effet qu'une réalité peut souvent être considérée et dénommée selon plusieurs facettes ou points de vue (Bedard 1986). Cette



connaissance lui permettra de créer de bons équivalents, des termes qui donnent une image aussi nette que possible du référent. La connaissance de l'univers notionnel du domaine, grâce à laquelle le traducteur appartient à la même communauté de pensée que l'auteur et le lecteur, lui permettra de reconnaître plus facilement la fonction du néologisme dans la LS.

b) Un deuxième principe, d'ordre terminologique cette fois, est le respect des traditions présidant à la création des termes dans la spécialité considérée. Chaque discipline ne possède pas seulement son système notionnel, mais également ses matrices terminogéniques, qui lui font choisir de préférence certaines lois de construction des termes. Un nouveau terme n'est pas un objet isolé mais un élément d'un système, plus ou moins structuré. Le traducteur professionnel expérimenté sait qu'en terminologie, il y a peu d'équivalents établis. Les équivalents présentés dans les dictionnaires ne sont souvent que des possibilités, auxquelles le traducteur est libre ou non de donner «vie». L'innovation est en effet contenue potentiellement dans la langue, ce sont des déterminations sociales qui provoquent son émergence à un moment donné. Cependant, le traducteur est parfois obligé de donner vie à un équivalent précis, notamment dans des contextes juridiques et officiels (normes, législations nationales ou internationales). Les traductions lexicales d'une discipline ou d'une technique incluent ce que Kocourek (1992) appelle un «entrain

onomasiologique» variable, c'est-à-dire une propension plus ou moins marquée des spécialistes à dénommer toutes les réalités abordées.

c) Le troisième principe, enfin, relève du respect de la cohérence de la langue-cible. Dans ses créations de termes, le traducteur sera conservateur et suivra les voies tracées par la langue. Les termes sont souvent motivés. Les néologismes doivent également offrir la possibilité d'engendrer des dérivés dans leur catégorie lexicale ou dans d'autres catégories lexicales.

3 Caractéristiques des néologismes de traduction

Quels caractères distinguent les néologismes primaires des néologismes créés par les traducteurs? Les hypothèses avancées ci-dessous sont basées sur le corpus des quelque 60 néologismes récoltés en 1995 par le Centre de Terminologie de Bruxelles dans les milieux de la traduction francophone belge.

3.1 Prédominance de la formation syntagmatique

Une première observation est la prédominance des syntagmes, dans notre corpus de néologismes de traduction et ce, quelle que soit la langue du terme-source. Nous émettons l'hypothèse que ce grand nombre est lié au processus même de la néologie traductive. En effet, si l'on compare les néologismes non

traductifs avec le reste du corpus, on constate que le pourcentage de syntagmes est plus élevé dans la seconde catégorie. Comparez les termes *coûtenance, déclaré, fleurage, gammiste* avec les néologismes traductifs *gestion intégrée des déchets, industrie de la récupération, recyclage énergétique*. Les premiers sont pourtant esthétiques (sonorité, connotations) «clairs», même si paradoxalement ils sont parfois cryptiques, sans être motivés à l'extrême.

a) D'une part, l'importance de la formation syntagmatique se manifeste par la comparaison du terme de la langue-source et de son équivalent français. Contrairement au français, l'anglais n'éprouve pas le besoin de préciser, dans les séquences déterminant+déterminé, les rapports logiques qui unissent ces deux éléments. Pour être plus explicite, le français utilisera par contre des prépositions, conjonctions ou subordinées relatives.

Exemples:

holding pipette → *pipette de maintien*
pipette de contention
pipette d'aspiration

(N+N)

public-key encryption → *encryptage à clé publique*
(Adj. + N + N) (N+prép.+N+Adj.)

Les syntagmes français raccourcissent néanmoins, grâce à la suppression d'éléments de liaison.

Exemples:

bulles trois trous, bulle deux trous.

Cette ellipse est attribuable à l'influence de l'anglais. L'ellipse du régissant, également utilisée dans les langues germaniques, se retrouve aussi en français, mais elle est plus rare.

Exemples:

(*bulle à*) *deux trous*
(*objets*) *encombrants ménagers*
(*déchets*) *inertes*

b) D'autre part, la prépondérance de la formation syntagmatique dans la néologie traductive s'observe également par la

comparaison avec les néologismes non traductifs, par exemple ceux du bâtiment.

Certains de ces néologismes non traductifs sont, il est vrai, des dénominations de spécialisations, dues à la profusion des méthodes, des matériaux, des techniques ainsi qu'à l'approfondissement des connaissances. Cette néologie reflète une évolution du métier, qui s'illustre parallèlement de deux façons. D'une part, on assiste à une redéfinition des tâches: une nouvelle spécialité est créée, qui, tout en regroupant des spécialités plus classiques, se concentre sur un objet spécifique (exemple: la salle de bain pour le *bainiste*, qui s'occupera autant de la plomberie que des revêtements, des boiseries éventuelles, de l'électricité, de la décoration, etc.). D'autre part, les spécialisations portent soit sur le matériau (le *métallier*), soit du lieu (le *vérandaliste*), soit de l'objet (le *dallagiste*). Les dérivés foisonnent pour ces dénominations de spécialisations, avec une préférence évidente pour le suffixe de spécialité *-iste*. Il va de soi qu'une néologie qui reflète une telle évolution ne peut pas être rendue adéquatement par des syntagmes. C'est, au contraire, la nouvelle dénomination qui crée ou renforce la création des nouveaux métiers.

Il n'en reste pas moins que beaucoup de néologismes secondaires plurilexicaux qui créent et expriment à la fois de nouvelles réalités s'expliquent sans doute par le blocage au niveau même de la notion qu'ils dénotent: un syntagme tient souvent plus de la définition et de la périphrase que de la dénomination.

En outre, la formation syntagmatique, quand elle se base sur un mot-clé auquel on ajoute une épithète, contribue *de facto* à la formation d'un schéma notionnel (générique → spécifique). L'exemple le plus frappant en est les «familles» de termes créés à partir des termes

suivants:

bulle: distinction selon l'apparence et l'utilisation (*deux trous, trois trous*)
collecte: distinction selon l'objet (*mono- ou multimatières*), le mode (*fatale, mobile*) et la forme (par *bulles, par parc à conteneurs*)
déchets: distinction notamment par l'origine (*verts, animaux, bureautiques*)
recyclage: distinction au niveau du temps (*pre- ou postconsumer*) de la matière et/ou du résultat (*recyclage énergétique*).

3.2 L'attrait du signifiant

Une deuxième caractéristique de la néologie traductive est l'attrait qu'exerce le signifiant de la langue-source. Elle revêt plusieurs formes et se rapproche du phénomène du calque.

a) Lorsque le signifiant recouvre une des caractéristiques de la notion ou du référent, par exemple sa forme (*skinpack*), le traducteur a souvent tendance à considérer le même trait pour créer un néologisme dans la langue-cible (par exemple, *pelliculage*). Ce faisant, il ne respecte pas le premier principe de la néologie traductive. Rappelons que ce principe consiste à considérer la notion au point d'oublier le signifiant en langue-source et à profiter du fait qu'une notion ou un référent peut être considéré et donc dénommé selon plusieurs facettes ou points de vue (forme, fonction, utilité, utilisation, etc.) (Bedard 1986). Voici un néologisme émanant d'une commission ministérielle de terminologie respectant ce principe de traduction. Il ne s'est pas lexicalisé:

Anglais	Notion	Français
<i>notebook</i>	→	<i>ordinateur</i>
<i>computer</i>	<i>portatif</i>	<i>plumier</i>

b) Voici un autre signe de l'attraction exercée par la langue d'emprunt. L'ordre des mots au sein des syntagmes nominaux complexes

français respecte généralement les règles d'agencement des mots en français et non celles de la langue anglaise d'origine.

Exemples:

high-band width cable connection
 → *connexion par câble à large bande*
integrated-services digital network
 → *Réseau numérique à intégration de services*
object-oriented programme/database
 → *Programmation/banque orientée objet*

Ce n'est cependant pas toujours le cas, notamment pour les syntagmes courts:

compact disk → *compact disque*

Deux modes d'attraction de l'anglais peuvent expliquer ce phénomène d'inversion déterminant-déterminé: dans les domaines comme l'informatique, l'attraction du modèle syntaxique, dont les linguistes pensaient pourtant qu'elle n'existait pas, ou alors les fausses motivations (en l'occurrence, le public traduit littéralement l'abréviation CD, qu'il connaît bien parce qu'elle figure sur tous les disques compacts qu'il achète).

c) Voici une troisième illustration de l'attrait de la langue originelle. En se transportant de l'anglais en français, un signifiant peut acquérir le statut de terme parce que le traducteur a cru le constater en anglais. Exemples de ces hapax, qui se sont lexicalisés en français, pas en anglais:

general-purpose computer
 → *ordinateur non dédié*
organizational structure
 → *structure organisationnelle*

Pour une part, ce phénomène s'explique peut-être par cette pratique de l'antéposition en chaîne de déterminants de l'anglais. Cette pratique donne une allure de substantifs plurilexématiques, donc de termes, à des expressions qui se traduiraient en fait plus élégamment par une subordonnée ou autre périphrase, en français. En ce sens, on

peut émettre l'hypothèse selon laquelle l'attrait du signifiant entraîne dans certains cas une terminologisation presque forcée du français...

3.3 De nouveaux formants internationaux

Un grand nombre de termes sont composés sur base d'*écologie*, qui devient par abrégement *éco(-)*. Cette forme abrégée est devenue un élément de composition à part entière dans le domaine de l'environnement. Ce préfixe joue un rôle comparable à celui qui provient du grec et du latin. De fait, cette particule a effectivement une origine grecque (*oikos*), mais elle s'est sémantiquement transformée, plus particulièrement ces 25 dernières années. Elle se rattache désormais au concept de l'environnement en général, et non plus de la maison, comme l'aurait exigé un point de vue strictement étymologique. Au sujet d'*éco(-)* encore, on note un flottement quant à sa graphie et à son insertion dans les composés. On l'y trouve avec et sans trait d'union.

Dans la protection de l'environnement, un autre élément fertile pour la création lexicale est l'adjectif *vert* (*écosac vert*, *verdier*). Il est relativement peu fréquent dans les termes qu'il nous a été donné de récolter, mais bien présent dans le vocabulaire de l'environnement, et ce, dans une série de langues, notamment l'allemand et l'anglais. *Eco-* et *vert* sont par ailleurs devenus synonymes, désignant tous deux la notion «respectueux de l'environnement»: *police verte*, *PC vert*, *comptabilité verte*...

3.4 Toujours des emprunts, des calques et peu de néonymes

Nous avons rencontré relativement peu d'emprunts ou de calques parmi les termes qui nous ont été soumis. De toute évidence, les domaines qui restent les plus soumis à cette pratique sont ceux de l'informatique et de la médecine.

Exemples:

workflow, drawer → *tiroir*

Nous n'avons rencontré aucun exemple de création *ex nihilo*. Ce procédé de création lexicale n'est pas fréquent dans la pratique du traducteur, et ne lui est d'ailleurs pas recommandé. Ce type de procédé est presque exclusivement réservé à la publicité (pour les dénominations commerciales, notamment). Il ressort d'ailleurs de plusieurs recherches menées dans le domaine que ce type de terme s'enracine très difficilement dans l'usage, car le locuteur le perçoit comme un «corps étranger» dans sa langue.

3.5 Des confixes hybrides

Dans les confixes, on remarque un nombre important de formations hybrides (latin-grec, classique-moderne) parmi les termes qui nous ont été soumis. L'usage ne semble donc pas éprouver de réticences vis-à-vis de ce procédé, pourtant critiqué et déconseillé par l'ensemble des linguistes et philologues. Exemples:

écophile formant moderne *éco(logie)* + formant grec *philos*
polynormé formant grec *poly* + formant moderne *norme*
domologie formant latin *domus* (*maison*) + formant grec *logos* (*discours*)

La recherche d'équivalents pour des acronymes anglais mène le traducteur à créer des termes dont on ne connaît ni le statut ni les possibilités de diffusion.

On propose par exemple *mégadvertising* pour *megatainment*, et *alicaments* comme équivalents de *medifood*. Quant à *UFO-ologist*, il est traduit par *ovniologue*.

3.6 Création lexicale liée à Internet

Les activités et termes variés proposés sur Internet ont suscité une explosion de termes ou mots nouveaux, même s'ils ne sont pas tous spécialisés. En voici quelques exemples, qui constituent un défi à la créativité du traducteur, mais qu'il doit impérativement transposer d'une manière ou d'une autre. Les termes créés sont peu stables, tant pour le fond que la forme (graphies multiples) et ils semblent hésiter entre la forme syntagmatique N+Adj (ex. *banque virtuelle* pour *web banking*) et le composé ou confixe «*cyber-* + N»: Angl. *Cyberspace*

→ Fr. *cyberespace*

Cyberocracy

→ *ybérocratie*

All. *Cyber-Zeugin*

→ *cybertémoin*

Cyber-Anwalt

→ *cyberavocat*

Cyberpolizei

→ *cyberpolice*

Vs. *Cyber-Küken*

→ *poussin cybernétique*

Cyber-Haustier

→ *animal domestique virtuel*

Angl. *cyber-car*

→ *voiture virtuelle*

Nl. *Cyberparadijs*

→ *paradis virtuel*

On le voit, il importe de mobiliser les facultés créatrices de la langue française et de ses locuteurs et de permettre aux francophones le comportement ludique dont font preuve les anglophones dans leur usage quotidien de la langue. Ces mots ou termes se stabiliseront ou disparaîtront, c'est l'usage qui

tranchera. De nombreux néologismes liés à Internet semblent ressortir exclusivement au domaine de l'écrit (média ou publications pour internautes).

Il ne suffit donc pas au traducteur, pour créer de bons néologismes, de respecter les principes de la néologie traductive rappelés plus haut, ni de connaître et d'appliquer les procédés classiques de formation lexicale en français. D'autres ressources, que mentionne Cl. Bédard (1986), sont d'excellents moteurs de création:

1. L'emprunt à des domaines techniques analogues;
2. L'extension de sens;
3. L'omission de constituants lorsque le syntagme d'origine est déjà très long.

Le traducteur dispose d'une assez grande liberté de choix, pour les moyens linguistiques réexprimant en LC, avec les mêmes fonctionnalités (Delisle 1993), le sens d'un texte en LS.

4 Conclusions

L'étude des néologismes recensés nous a permis de dégager deux caractéristiques majeures, et à notre avis mal connues, du processus même de la néologie traductive: à savoir la prédominance de la formation syntagmatique et l'attrait du signifiant de la langue-source. Pour créer de bons néologismes, le traducteur doit dès lors, plus systématiquement et plus consciemment, se dégager de la manière dont la notion est exprimée (ou le référent est dénommé) dans la langue-source. Il lui faut considérer tous les aspects et toutes les dimensions de la notion ou du référent. Il pourra ainsi utiliser l'aspect ou la dimension qui convient le mieux au discours, aux habitudes langagières et terminologiques de la langue-cible.

Cette méthode lui permettra d'avoir moins recours aux calques, mais souvent également de trouver de meilleures solutions qu'un syntagme définitoire. Il contribuerait ainsi mieux à l'enrichissement de la langue en créant des termes plus courts, plus esthétiques et réellement nouveaux.

La néologie traductive est un domaine inexploré. Les résultats d'études à son propos intéresseraient aussi bien la traductologie que la néologie en langue française. C'est que la langue française et sans doute toutes les langues du monde sont en situation de traduction, presque systématique, par rapport à l'anglais. Outre un cours de terminologie, ne faudrait-il pas dès lors ajouter au programme des spécialistes eux-mêmes, ou au moins à celui de leurs formateurs, un cours de traductologie?

*Adrien Hermans
et Andrée Vansteelandt,
Centre de terminologie de Bruxelles,
Institut Marie Haps,
Bruxelles.*

Bibliographie

- Abreu (J.M.), 1994: «L'abréviation dans le langage technique», dans *La Banque des mots*, n° 47, p. 105-115.
- Adda (R.), Bastuji (J.), Bochmann (K.) et al., 1979: *Néologie et lexicologie: hommage à Louis Guilbert*, Paris, Larousse.
- Bedard (C.), 1986: *La traduction technique. Principes et pratique*, Québec, Linguatex.
- Benhamida (L.), 1993: «Translators and Interpreters as Adopters and Agents of Diffusion of Planned Lexical Innovations: The Francophone Case», dans Ed. Wright (S.E.) and Wright (L.D.): *Scientific and Technical Translation*, Amsterdam, Benjamins.
- Boulanger (J.-C.), 1990: «La création lexicale et la modernité», dans *Le Langage et l'homme*, XXV, 4, p. 233-240.

Cusin-Berche (F.), 1998: *Le management par les mots: Étude sociolinguistique de la néologie*, Paris, L'Harmattan.

de Schaetzen (C.), 1994: «Méthodologie des attestations en néologie terminologique», dans *Le Langage et l'homme*, XXIX, 1.

Delafontaine (Christine) sous la direction de I. Hoorickx, 1994-1995: *A Terminology of Medically Assisted Procreation*, Bruxelles, Institut Libre Marie Haps.

Delisle (J.), 1993: *La traduction raisonnée*, Ottawa, Presses de l'Université d'Ottawa.

Diki-Kidiri (M.), Joly (H.), Murcia (C.), 1981: *Guide de la Néologie*, Paris, Conseil international de la langue française.

Dugas (A.), Molinier (C.), 1992: «La productivité lexicale», dans *Langue française*, n° 96, p. 3-124.

Goosse (A.), 1975: *La néologie française aujourd'hui: observations et réflexions*, Paris, Conseil international de la langue française.

Guilbert (L.), Baggioni (D.), Bastuji (J.), Dresco (P.), Fauveau (B.), Gardin, Guespin (L.), Laurian (A.-M.), Lefevre (G.), Marcellesi (C.), Mortureux, Petroff, 1974: «La néologie lexicale», dans *Langages*, n° 36, p. 3-128.

Guilbert (L.), 1975: *La créativité lexicale*, Paris, Larousse.

Hermans (A.), 1991: «Groupes de travail en terminologie», dans *Le Langage et l'homme*, XXVI, 2-3, p. 165-171.

Hermans (A.), 1991: «Quelques pratiques terminographiques des traducteurs belges», dans *Nouvelles de la FIT*, 10, 3.

Hermans (A.), 1994: «Traduction et néologie. Proposition de coopération», dans *Le Langage et l'homme*, XXIX, 3-4, p. 299-302.

Hohnhold (I.), 1990: *Übersetzungsorientierte Terminologiearbeit: Eine Grundlegung für Praktiker*, Stuttgart, Intra.

Hrsg Altner (G.) et al., 1993: *Jahrbuch Ökologie*, München, C.H. Beck.

Humbley (J.), 1993: «L'observation de la néologie terminologique: l'expérience du CTN», dans *La Banque des mots*, 5, numéro spécial, p. 65-73.

Kocourek (R.), 1982: *La langue française de la technique et de la science*, Wiesbaden, Brandstetter Verlag.

Lenoble-Pinson (M.), 1991: *Anglicismes et substituts français*, Louvain-la-Neuve, Duculot, 173 p.

Massenaux (Nadine) sous la direction de M^{me} Costa, 1992-1993: *Vokabular der Wende in den deutschsprachigen und französischen Presse, Lexeme und Termini*, Bruxelles, Institut libre Marie Haps.

Rey-Debove (J.), 1971: *Étude linguistique et sémiotique des dictionnaires français contemporains*, Mouton.

Sager (J.-C.), 1989: «Term Formation», dans *Lebende Sprachen*, XXXIV, 4.

Thiel (G.), Thome (G.), 1987: *Resolutionen: Ein empirisch entwickelter Beitrag zur Textanalyse*, Tübingen, G. Narr.

Wijnands (P.W.), 1989: «Le néologisme comme problème de traduction», dans *Lebende Sprachen*, XXXIV, 2.

Remerciements

La recherche a été encouragée par une subvention accordée par le Service de la langue française du Ministère de la culture et des affaires sociales de la Communauté française de Belgique.

La recherche dont le présent article donne quelques résultats a été réalisée par Adrien Hermans, Élisabeth Lequeue et Andrée Vansteelandt, avec la collaboration de Hélène Bertrand, de Thierry Lepage et de Pablo Sanz Moreno.

Néologie et terminologie : activités et réflexions

Pour présenter le nouveau Service de néologie et terminologie de l'Inalf, on résume les principales missions de ce service dont une tâche importante est l'assistance aux commissions spécialisées de terminologie et de néologie dans le cadre du dispositif d'enrichissement de la langue française. On propose une typologie de ses activités et on discute certains éléments méthodologiques mis en œuvre ou observés en terminologie ponctuelle.

Termes-clés : terminologie ponctuelle ; veille néologique ; description de l'usage ; étiquetage par domaines ; variation ; recommandations officielles.

L nous a longtemps semblé prématuré d'intervenir dans un numéro de revue consacré aux « nouveaux outils pour la néologie ». En effet, le SNT (Service de néologie et terminologie) créé il y a quelques mois seulement⁽¹⁾, est trop jeune pour prétendre donner aujourd'hui des renseignements généraux sur ou pour la néologie. C'est pourtant bien dans ce secteur qu'il œuvre, ayant repris, pour certaines de ses missions, une partie des activités du CTN. L'équipe, très restreinte, de permanents participant partiellement à ces travaux, a déjà, outre une longue expérience de la lexicographie, celle de la lexicographie de spécialité⁽²⁾. Elle a participé à la constitution et à l'expérimentation de plusieurs bases de données lexicales, dictionnairiques ou textuelles, dans les domaines de spécialité⁽³⁾, et traque de longue date les néologismes dans les textes⁽⁴⁾.

Après avoir décrit (en 1) les missions qui reviennent au SNT5 dans le cadre des travaux pour les commissions spécialisées de terminologie, il est nécessaire d'établir un relevé des requêtes qu'elles entraînent (en 2), qui sera suivi (en 3) de l'exposé des moyens et de la méthodologie adoptés puis de l'évocation de quelques constats et perspectives (en 4).

(1) À l'issue de deux réunions déterminantes à la Délégation générale à la langue française (DGLF) en septembre 1998, ayant eu pour but de re-diriger une partie des missions du CTN (sous la responsabilité de John Humbley) vers le SNT (sous notre responsabilité). Cela a pu se faire grâce à une convention avec la DGLF (voir le *Rapport annuel de la Commission générale de terminologie et de néologie*, DGLF, 1998, p. 24).

(2) Il s'agit de Pierrette Marchaudon, de Violette Tolédano et de nous-même, toutes trois ayant participé à la rédaction du *Tésor de la langue française (TLF)* puis, dans le cadre des travaux de l'Équipe Sciences et techniques, au *Supplément au TLF* pour les termes de spécialité et les emprunts à l'angloaméricain, et enfin, dans le cadre du projet *DFST*, à la création de bases de données lexicales dans les domaines de spécialité. On trouvera des précisions sur ce projet dans Candel 1994.

(3) Participation au choix des textes pour la base textuelle *Scitech*, exploitation de la base de dictionnaires ITC1 et 2, création de la base lexicale *Lexitech*, rédaction de vocabulaires (jusqu'en 1997).

(4) Nous élaborons notamment une étude sur la définition dans les textes de spécialité.

(5) Dans le cadre de la convention avec la DGLF citée ci-dessus en note 1.

1 Les actions du SNT dans le cadre des travaux pour les commissions de terminologie.

Trois types d'activités se dégagent : l'assistance aux commissions spécialisées de terminologie et de néologie, l'assistance à la commission générale de terminologie, caractérisées l'une et l'autre par la rédaction et la diffusion de documents en amont des séances ; et enfin, la participation aux travaux et réflexions des commissions en cours de séance, sous la forme d'avis, de propositions, de commentaires⁽⁶⁾. Ces trois activités se traduisent par la collecte d'informations préexistantes aussi bien que par des suggestions, voire des recommandations.

D'une façon générale, il s'agit de terminologie ponctuelle. Tout au plus aborde-t-on la terminologie systématique ou ses méthodes en quelques occasions, de manière ponctuelle : on peut être amené à traiter d'éléments ou de sous-ensembles d'une terminologie (*cf.* le champ de la toxicodépendance dans le domaine de la santé, avec notamment les termes *assuétude*, *addiction*, leurs synonymes et leurs équivalents). Mais il ne s'agit pas de rédiger des terminologies en tant que telles.

2 Une typologie des missions proposées au SNT dans le cadre des travaux des commissions spécialisées de terminologie.

On trouve, résumés dans le tableau 1, deux types de demandes : l'assistance à la veille néologique et terminologique (ci-dessous en 2.1) et la recherche d'un terme français

équivalent à un terme anglo-américain (en 2.2). Deux autres types de travaux pour les commissions spécialisées de terminologie et de néologie sont présentés au tableau 2 : l'élaboration ou la reformulation d'une définition (en 2.3) et l'examen ou le réexamen d'un terme ou d'une forme (en 2.4). Quelques observations générales sont présentées en 2.5.

2.1 L'assistance à la veille néologique et terminologique

La veille néologique et terminologique se fait par le repérage et la collecte de néologismes terminologiques, et concerne les termes eux-mêmes aussi bien que leurs définitions.

2.2 La recherche d'un terme français équivalent à un terme anglo-américain

Ce type d'étude regroupe trois cas de figure. Soit le terme discuté est déjà en usage en français, on en recherche un autre, qui peut être un néologisme de forme et on aménage une définition de ce terme en français. Soit le terme étudié est en train de s'implanter en français, tout comme le concept auquel il renvoie à lui-même franchi les frontières : la recherche porte alors sur un néologisme de forme et de sens et s'accompagne aussi d'une recherche de définition en français de ce terme. Soit, enfin, le terme français recherché peut lui-même être l'équivalent d'un autre équivalent français déjà en usage mais ne donnant pas satisfaction, car il est, par exemple, trop proche du terme source anglo-américain et ne s'intègre pas vraiment au moule morphosémantique du système linguistique français ; l'activité porte sur un néologisme de forme. Un

débat porte actuellement sur *leadership*, qui, d'après le *Petit Robert* 1996, est un emprunt enregistré dès 1875, avec le sens général de « fonction, position de leader » ; *leader* est enregistré dès 1929, avec le sens de « chef, porte-parole (...) », « personne qui prend la tête d'un mouvement, d'un groupe » et un renvoi à *meneur* et à *chef de file* ; *chef de file*, au sens « figuré » de « celui qui vient le premier dans une hiérarchie, qui est à la tête d'un groupe, d'une entreprise », on renvoie à *leader*. Or les spécialistes de la banque ont pour coutume d'utiliser, avec un sens particulier à leur domaine, le dérivé *chef de filat*. Deux positions sont possibles : refuser un néologisme français choquant⁽⁷⁾ pour le non-initié en raison de son mode de dérivation ou bien appuyer l'effort de francisation enregistré dans le jargon bancaire.

2.3 L'élaboration ou la reformulation d'une définition

Dans les deux premiers cas présentés au Tableau 2, l'activité est relative à un néologisme (partiel) de sens, qu'il s'agisse de la modification d'une définition pour un terme dont le sens a évolué ou de la modification d'une définition pour un terme dont le domaine d'emploi a changé. Dans le troisième cas, on cherche à formuler une définition en termes de spécialité, mais qui soit compréhensible par le plus grand

(6) Nous souhaitons mentionner l'efficacité de nos trois premières collaboratrices temporaires dans ce secteur : Gaëlle Berjonneau, Emmanuelle Bruley et Odile Raymond.

(7) Il le serait moins s'il se graphiait **chefdefilat*.

nombre. Cette préoccupation a été la nôtre dans l'ensemble de nos travaux de lexicographie traitant de mots de spécialité⁽⁸⁾. Elle s'impose aussi en terminologie, aussi fine et pointue cette terminologie soit-elle. Il faut rappeler ici le rôle du public: pour qui ces travaux de terminologie sont-ils réalisés? Quel est le public visé? Ce n'est pas seulement le public spécialisé du domaine – qui n'en a guère besoin. Ce n'est pas seulement le scientifique de haut niveau cherchant à se cultiver dans un domaine voisin du sien. C'est aussi un public plus large, journalistique, médiatique, pour qui une bonne vulgarisation s'impose: il faut «déformuler» les définitions trop spécialisées pour les «reformuler», afin que le plus grand nombre de lecteurs les comprenne. Ce public est également constitué de personnes concernées directement par ces termes dans leur travail quotidien, et qui pourtant n'ont pas toujours connaissance des notions auxquelles ils se rapportent: nous pensons au vaste ensemble professionnel constitué par le personnel de secrétariat, telles les secrétaires médicales, qui transcrivent et contribuent à rédiger les commentaires, rapports, diagnostics, ordonnances, prescriptions dictés par les médecins.

2.4 L'examen ou le réexamen d'un terme ou d'une forme

Un terme ou une forme entraînent un réexamen notamment si leur domaine d'emploi a changé. Le cercle de leur(s) domaine(s) d'appartenance s'élargit ou se rétrécit, selon que le terme spécialisé se banalise, se généralise, ou selon que, au contraire, il se spécialise davantage. Il se peut aussi simplement que le terme connaisse une «re-spécialisation», un simple transfert de domaine. Il y a lieu d'interpréter

l'activité en cause comme relative à un néologisme de forme et de sens.

L'examen de formes dont la graphie a changé permet d'enregistrer des variantes graphiques en fonction des domaines d'emploi. S'il s'agit d'un aménagement possible des graphies selon les domaines dont relèvent leurs emplois, il paraît économique de proposer une homogénéisation, allant de préférence, dans le sens d'une simplification. Nous pensons à un exemple comme la série des termes en *scan(n)*⁽⁹⁾, les différents termes terminés en *-plex(e)* (*duplex* et *multiplex* mais *multiplexe*) ou encore les variantes en *hém(o)* – / *hém(a)* –, *héma* – / *hémato* – (*hémovigilance*, *hémacrite*, *hématocrite*) ou *densi-* / *densito* – (dans *densimètre* et *densitomètre*).

Ces études permettent d'isoler des formes ou des éléments disponibles émergents, que l'on pourra proposer pour de nouvelles séries de formations: citons *-vigilance* (*hémovigilance*, *matérovigilance*, *pharmacovigilance*, *réctovigilance*, *stimulovigilance*); citons aussi l'élément *-el* en sciences et techniques spatiales, de *élément*, dans *pixel* (de l'anglais *picture element*), et, à partir de ce modèle, dans *volumel* et dans *voxel* (anglo-américain: *volume element*, *voxel*), mais qui apparaît concurrent d'une finale *-èle* dans *tachèle* (pour *tâche élémentaire*), qui correspond à l'anglo-américain *groundel*⁽¹⁰⁾.

2.5 Analyse des principaux facteurs de variation: synonymie, niveaux de spécialisation, facteur temps

Nous venons de relever quelques caractéristiques d'activités de terminologie aussi bien que de lexicographie terminologique. Dans les cas décrits en 2.2, on est en présence de séries néologiques

coexistantes et concurrentes, qui ne sont autres que des synonymes terminologiques, l'un des termes reflétant l'usage naturel, spontané, l'autre étant un terme élaboré – ou du moins recommandé – en vue de supplanter le premier. C'est la question classique des termes néologiques concurrents qui se pose, avec, comme choix d'analyse, et selon le point de vue adopté, celui de termes homonymiques, ou celui de mots polysémiques. Nous mettons en avant ici la probabilité d'une nuance apportée au sens d'un terme par la nouvelle forme servant à le désigner: plus exactement, ces nouvelles désignations véhiculent sans doute des connotations (*frimousse* est proposé comme équivalent de *smiley* (*Vocabulaire de l'Internet*, 1999) (v. Spillner 1994)⁽¹¹⁾).

Dans le cas d'une définition vulgarisatrice (2.3) il s'agit vraiment de la cohabitation de définitions, même si nous ne sommes pas là en lexicographie générale, où l'on appelle la «double définition» ce qui consiste à proposer, pour un terme de spécialité, une définition en termes spécialisés d'une part et une définition pour le grand public d'autre part. Nous devons tenir compte de niveaux de spécialisation différents dans l'art de définir: reformuler en termes simples une définition de spécialistes.

(8) Cf. la note 3 ci-dessus.

(9) Voir Candel (à paraître): § VB.

(10) Ces termes et leurs définitions ont fait l'objet de publications au *Journal Officiel* et sont actuellement en cours de révision.

(11) Les «flottements» demeurent toujours possibles (De Schaetzen 1990: 57).

Un certain nombre de cas permettent la mise au jour de sous-ensembles terminologiques devenus désuets : ils ne devraient pas trouver leur place dans un relevé néologique. Il s'agit de termes dont l'usage est révolu, le concept désigné ne correspondant plus à une réalité en usage. Ce sont des « archaïsmes néologiques », ou « néologismes archaïques », termes devenus obsolètes, en raison de l'évolution subie par leur forme ou par leur sens, d'un usage révolu, dépassé, vieilli. Toute analyse d'un terme en formation devient un élément d'analyse diachronique. L'action sur la langue consistant à recommander des termes, qui porte en principe sur des néologismes terminologiques, risque de s'appliquer à des termes connus, en usage depuis un certain temps déjà, et qui ne sont donc plus des néologismes. Les commissions spécialisées de néologie et de terminologie sont tout naturellement amenées à dépasser le strict cadre de la néologie, en cherchant à orienter ou à redresser des usages terminologiques – mais l'on peut se demander si rester dans le strict cadre de la néologie terminologique est possible.

3 Moyens spontanément mis en œuvre et méthodologie adoptée

3.1 Travail au sein du SNT

3.1.1 Recherches dans des dictionnaires et autres répertoires de mots

Des termes, ou, le plus souvent, des traductions de termes, ainsi que leurs définitions, sont recherchés, ainsi que leurs définitions, dans la quête d'équivalents français à des termes anglo-américains. On consulte à cette fin :

- a) des bases de données (*le Grand Dictionnaire Terminologique*, de l'Office de la langue française, au Québec, *Termium*, du Bureau de la traduction, à Ottawa, *Eurodicautom*, des Communautés européennes, à Luxembourg), sur Internet et / ou sur CD-Rom ;
- b) des dictionnaires spécialisés : sur CD-Rom (choix de dictionnaires de *Lexpro* – la Maison du dictionnaire, *Cilf*), sur Internet, avec ou sans définitions, domaine par domaine, en version papier⁽¹²⁾ ;
- c) des dictionnaires généraux : sur CD-Rom (*Le Petit Robert* 1996), sur Internet (plusieurs tomes du *TLF*, le tome 14 étant accessible librement), en version papier (le *TLF*, le *Dictionnaire de l'Académie*, 9^e édition) ;
- d) des dictionnaires encyclopédiques, en version papier (le *Grand Larousse universel*).

3.1.2 Recherches dans des textes

La terminologie spécifique à un domaine, objet des cas examinés ci-dessus, ne couvre pas toujours un ensemble d'usages suffisamment représentatifs du domaine étudié. Les textes ont la mémoire des mots, des usages, et de l'évolution de ces usages. Une collecte de données plus vaste s'impose alors, avec une vérification dans les textes, à une plus large échelle à travers l'interrogation de textes, sur des CD-Rom (*Encyclopædia Universalis*), des bases textuelles informatisées (*Scitech*⁽¹³⁾ de l'Inalf), ou dans divers contextes balayés grâce au moteur de recherche Altavista. Il s'agit d'une exploitation de l'usage, que l'usage soit établi ou naissant. Une recherche de la fréquence autant que de la répartition des attestations peut être rapidement révélatrice des usages à suivre ou à recommander.

3.1.3 Une démarche efficace : la recherche par synonymes

Qu'il s'agisse de la recherche ou de la vérification de termes en français, d'équivalents étrangers, ou de domaines dans lesquels un terme est utilisé, on constate que ces recherches se font le plus souvent au moyen de l'interrogation de termes qui, par le sens ou par la forme, sont proches des termes à l'étude. C'est donc une étude à partir des synonymes qui s'impose (voir Lerat 1994 : 28). En sciences et techniques spatiales, une recherche sur *stabilisation (sur) trois axes* permet d'évaluer, outre ces deux formulations elles-mêmes, des attestations de *stabilisation à trois axes, triaxiale, sur les trois axes, 3 axes*, pour deux synonymes en anglais.

3.2 Travail en collaboration

Le travail se fait d'abord et surtout avec les membres des commissions spécialisées de terminologie. Ce sont des personnalités marquantes des domaines concernés, de la sphère scientifique, industrielle, économique (le CEA, le CNES, la Banque, les hôpitaux, la pharmacie, la presse médicale, les syndicats représentatifs des professions...). On travaille en séance sur les termes, leurs définitions, leurs usages. Des propositions sont recueillies, discutées, annotées en commun et les échanges entre spécialistes des domaines concernés sont de tout premier intérêt pour les linguistes. Le rôle des personnalités du domaine est inestimable. Leur disponibilité et l'intérêt qu'elles manifestent pour la langue représentent pour les linguistes, les lexicographes,

(13) Voir le *Rapport d'étape Inalf* de juin 1997, UPR 9017, p. 21-27.

terminologues et traducteurs, un atout exceptionnel, inégalable et marquent une étape indispensable dans la réflexion sur les termes⁽¹⁴⁾.

3.3 Comparaison d'expériences différentes: lexicographie générale, lexicographie de spécialité et terminologie

Trois types de démarches sont comparées au tableau 3: la lexicographie générale (par exemple telle que nous l'avons pratiquée dans le cadre du *TLF*), la lexicographie de spécialité (par exemple telle que nous l'avons pratiquée pour le *Supplément au TLF* et pour les travaux de l'équipe sciences et technique) et la terminologie (telle que nous la pratiquons au sein du SNT dans le cadre du travail pour les commissions spécialisées de terminologie). Nous les avons caractérisées selon leurs objets d'étude, la méthode qu'elles entraînent, le mode de classement des unités étudiées et le recours à l'usage⁽¹⁵⁾. On relève de nombreux points communs entre ces trois démarches, mais surtout le même besoin d'analyse et de synthèse, ainsi que l'effort de proposition et de

conseil, mais répartis de manière différente entre ces trois expériences. Les unités à l'étude sont issues de listes spécialisées ou, plus rarement, de corpus textuels. Il faudrait systématiser le recours aux marques de domaines, aux arbres de domaines conventionnellement arrêtés mais autorisant des spécifications nouvelles. On atteste et valide les emplois et on privilégie l'usage grâce aux corpus textuels et aux recherches complémentaires dans de grandes bases de données textuelles et sur Internet. Mais, dans le traitement de termes, la question de l'usage reste largement ouverte: nous y reviendrons.

4 Constats et perspectives

4.1 Un consensus difficile

Il arrive que des spécialistes réunis au sein d'une commission ne s'accordent pas sur un terme, un domaine d'emploi ou une définition: après avoir statué, il suffit qu'ils conviennent un collègue extérieur au groupe, qui, entendant les résultats des débats, s'élève contre un terme, sa définition ou son domaine ou sous-domaine d'emploi, et tout est remis en cause. Le résultat de séances de travail en commission spécialisée peut aussi faire finalement l'objet d'un renvoi à une nouvelle série d'études et de débats, pour n'avoir pas reçu, ou pas totalement, l'agrément du Service du dictionnaire de l'Académie⁽¹⁶⁾ et de la commission générale de terminologie. En tout état de cause, il y a de fortes chances pour que les interventions ne concernent plus, alors, un véritable néologisme, mais un terme déjà ancré dans l'usage. Si le consensus est absent du début à la fin des débats, l'étude du terme est renvoyée à une date ultérieure, et peut-être abandonnée.

4.2 Une objectivité difficile

Il est plausible que l'on cherche à éviter des anglo-américanimes évitables. Mais les réactions quant à ces termes divergent d'un expert décideur à l'autre. La dénomination d'un concept familier est accueillie favorablement, fût-elle anglo-américaine. Un terme anglo-américain désignant un concept étranger à la connaissance d'un individu est naturellement moins bien accueilli. L'objectivité n'est pas partagée de la même manière selon que l'on est, ou non, familier d'un domaine de spécialité, de ses concepts et de ses désignations.

4.3 Choix conventionnels: étiquetage par domaines

Il arrive que des termes à l'étude, dans le domaine de compétence de la commission spécialisée, relèvent également d'autres domaines ou sous-domaines. Les spécialistes de la commission réunie renvoient alors ce terme à leurs collègues des autres spécialités, pour s'assurer de l'opportunité d'ajouter une autre marque de domaine à celle(s) proposée(s) par leur commission. Il n'est pas rare que la deuxième commission saisie propose alors de reformuler la définition, parallèlement à l'introduction de la nouvelle marque de domaine.

4.3.1 La terminologie des marques de domaines: une terminologie en soi

Nous sommes par ailleurs, avec les noms de domaines, dans un autre type de terminologie: c'est la terminologie des marques ou indicateurs de domaines. Or bon nombre de ces termes-marqueurs ont aussi leur place dans les dictionnaires de langue – on peut aisément consulter les longues listes de marques de domaines, en général abrégées, qui

(14) Outre les synthèses établies, ces circuits comprennent aussi des échanges avec les terminologues des autres pays francophones (Canada, Belgique, Suisse) et avec des organismes de normalisation.

(15) On se reportera avec profit à Lerat (1994).

(16) Le rôle de l'Académie est souligné dans Humbley (à paraître). Il est décrit dans le *Rapport annuel de la Commission générale de terminologie et de néologie*, 1998.

ouvrent tout dictionnaire de langue. Le *Dictionnaire de l'Académie* compte, pour le premier tome de sa neuvième édition (1994, A – Enz-), 137 marques de domaines. Le recours aux domaines est habituel en lexicographie générale, tout comme en lexicographie de spécialité et va de soi en terminologie.

Mais l'usage des marques de domaines implique toujours l'adoption de cadres conventionnels⁽¹⁷⁾. Ces marques ont la valeur de simples étiquettes et répondent à un codage. C'est, davantage encore, le cas des langages documentaires comme les thesaurus. Certains de leurs descripteurs accompagnent avec succès les descriptions linguistiques des mots de dictionnaires. Des confusions existent lorsque deux types de systèmes différents sont utilisés conjointement et que le répertoire, lexicographique ou terminologique, ne les distingue pas suffisamment. Langage administratif, institutionnel et langage usuel des initiés du domaine de spécialité peuvent ainsi être confondus. C'est ainsi que la commission générale de terminologie souhaite distinguer «ce qui relève de l'enrichissement de la langue française et ce qui relève de la terminologie administrative et juridique» (*Rapport annuel*, 1998: 11). Il faut définir les étiquettes, ou, plus exactement, déterminer les ensembles auxquels elles renvoient et les sous-ensembles qu'elles incluent éventuellement. Dans les différents types de constructions ou de recueils administratifs, on cherchera à éviter les dénominations ambiguës, ou répondant à des modes de distribution inhomogènes. Les langages documentaires n'ont pas

pour fonction première de correspondre aux termes utilisés dans le discours scientifique ou technique, mais il peut y avoir ressemblance et adéquation trompeuse. Si la néologie n'est pas, sans doute, prédominante dans les terminologies en usage ici, elle y est présente. D'où l'utilité de collaborations comme celle engagée entre la commission spécialisée de terminologie compétente pour le domaine de la santé et des groupes œuvrant à des classifications médicales, comme la Classification commune des actes des professions de santé (CCAPS), due au Pôle d'expertise et de référence national des nomenclatures de santé (PERNNS), et réalisée à partir de la Nomenclature générale des actes de professionnels (NGAP) et du Catalogue des actes médicaux (CdAM), ou la Classification alphanumérique de la documentation médicale et pharmaceutique (CANDO).

4.3.2 Points de vue terminologique ou lexicographique

Dans les démarches terminologique et lexicographique, le changement de domaine pourrait être traité comme le changement de sens. Dans le cas d'un changement de sens, s'il s'agit de modifier la définition d'un terme, que le sens de ce terme ait évolué ou qu'il y ait eu transfert ou glissement de son domaine d'emploi, on estime, en terminologie, qu'on a affaire à deux unités différentes. En lexicographie, on a là deux sens différents, ou deux emplois différents d'un même mot. Pour un changement de domaine d'appartenance, cette répartition est plus difficile à faire. D'un strict point de vue de terminologie, on admet qu'on a affaire à deux termes, relevant de deux domaines d'emploi différents, alors qu'un traitement lexicographique ferait placer sous une même entrée deux alinéas successifs,

introduits chacun par son étiquette de domaine, voire proposerait une seule définition, avec une indication de domaine supplémentaire pour une partie de la définition seulement. Ce catalogage des termes par étiquettes de domaines est lui-même modulable: ainsi, on peut aussi bien proposer en parallèle deux étiquettes de domaines, c'est-à-dire rattacher simultanément un même terme à ce deux étiquettes (par exemple: «ingénierie nucléaire» ou «santé» pour le terme *radioprotection*), ou encore proposer une hiérarchie dans le classement de ces étiquettes et recourir à une étiquette de domaine et à une étiquette de sous-domaine (par exemple «transports» par rapport à «aéronautique»). Une confusion risque de se faire jour entre les sens, les emplois particuliers de mots et les termes ayant un sens spécifique dans le cadre d'un domaine particulier d'emploi, que ce domaine relève des sciences ou des applications technologiques ou industrielles.

4.4 La veille néologique en terminologie

La «veille néologique et terminologique» a souvent pour rôle d'appuyer ou de réorienter des propositions. Mais là encore, la prudence s'impose. Ce n'est pas toujours un article de journal, ou quelques extraits de la presse du mois qui permettent de prendre une décision. On connaît les effets de mode, les usages qui se répandent aujourd'hui extrêmement vite: en 15 jours il arrive que l'on ait recueilli pour un terme autant d'attestations d'emploi qu'en 20 ans auparavant; les attestations sont donc nombreuses, mais la couverture des emplois n'est pas significative.

Cette activité est à développer, avec les spécialistes des domaines. Toutes les propositions sont à envisager: la mise en forme de

(17) Notre expérience du *Dictionnaire de sigles* a été éclairante à ce sujet.

documents déjà collectés, aussi bien que le lancement de nouvelles collectes. Bien entendu, elle existe déjà largement, et de façon inégale, conformément aux besoins des commissions spécialisées de terminologie et néologie.

On devrait tenir compte de plusieurs paramètres: (a) du nombre d'occurrences relevées par domaines de spécialité recensés, (b) du nombre d'occurrences relevées dans des répertoires de spécialité et dans des terminologies, (c) des avis des représentants des groupements professionnels, des groupes d'utilisateurs, d'usagers de la langue du domaine – souvent bien présents ou représentés dans les commissions de terminologie, (d) des types de supports textuels analysés: encyclopédies (*Encyclopædia Universalis*), Internet (*Altavista*: toutes sortes de textes), journaux, manuels, articles de vulgarisation, etc. et (e) pour un terme du vocabulaire scientifique général, du nombre de domaines ou de sous-domaines recensés.

La veille passe par une expertise que l'on appellera de première main: données rassemblées et fournies par les spécialistes eux-mêmes, documents listés, extraits de journaux fournis par les membres des commissions, ou par d'autres contributeurs. Elle profitera aussi d'une expertise «de seconde main», données provenant par exemple des mémoires de terminologie rédigés par des étudiants; à cet ensemble pourraient être agrégées les données que nous avions collectées d'ores et déjà dans le cadre de l'équipe Sciences et techniques de l'Inalf.

La veille peut par ailleurs contribuer aux travaux d'ingénierie linguistique relatifs au repérage automatique de terme.

Le SNT ne saurait être systématiquement un producteur de néologie: c'est le domaine des spécialistes. Il peut l'être

occasionnellement. Plus exactement, cette fonction serait donc d'être récepteur et organisateur, archiveur de néologie, ou du moins de collaborer à un tel ensemble d'activités.

4.5 Description de l'usage ou prescription ?

«L'usage décidera», conclue-t-on couramment à défaut de consensus. Cela revient, le plus souvent, à laisser le terme anglo-américain s'implanter. L'«usage», qui serait à corriger en «usages», serait à redéfinir, moyennant un ensemble d'enquêtes permettant de tenir compte du double critère de fréquence et de répartition des emplois.

L'usage s'impose toujours, parallèlement ou concurrence aux actions de «prescripteurs», mais, paradoxalement, c'est quand il devient gênant qu'une solution commode pour les «prescripteurs» consiste à dire «l'usage tranchera», ou «laissons faire l'usage». C'est ce qu'a tendance à faire l'Académie, «gardienne de la langue» et défenseur de l'usage, précisément. C'est ce que fait le *TLF*, qui atteste, décrit les sens, emplois et usages rencontrés. Mais ni le *Dictionnaire de l'Académie* ni le *TLF* ne prétendent être des dictionnaires de français scientifique ou technique, même si le *Supplément au TLF* (où notre équipe a choisi d'éviter les usages postérieurs à 1970 afin d'être mieux à même de juger des «usages» établis) s'en rapproche davantage, pour ce qui est de la partie spécialisée du vocabulaire pris en compte. Mais dans le cadre de l'action pour laquelle œuvrent les commissions spécialisées de terminologie et de néologie, «laisser faire l'usage» équivaut aussi bien à laisser l'anglo-américanisme s'implanter aux dépens d'un usage naissant, alternative commode, qui pourrait facilement être adoptée par les spécialistes, pour peu qu'on les y

encourage. L'exemple cité plus haut de *leadership* vs *chef de filat* fait entrevoir deux réalités: ne pas accepter de recommander officiellement le néologisme de *chef de filat* dans les textes relatifs à la Banque incite à l'angloaméricanisme, ce qui va à l'encontre de l'effort d'«obligation pour les services de l'État d'employer les équivalents français publiés à la place des termes étrangers correspondants» (*Rapport annuel de la commission générale de terminologie et de néologie*, 1998: 9). Mais il est vrai aussi que si l'usage d'ores et déjà implanté chez les professionnels du domaine est *chef de filat*, ce terme continuera sur sa lancée, même si le dispositif officiel d'enrichissement de la langue française préférerait s'abstenir de le recommander⁽¹⁸⁾.

Il y a des interférences aussi bien que de grandes dissensions entre deux types d'usages: on peut s'exprimer en langue générale courante, commune à un grand nombre de locuteurs de «la rue», ou bien user d'une langue de spécialité, pouvant atteindre un degré de spécialisation extrême, où se pose alors le problème du faible nombre de spécialistes, mais dont le langage a une portée extraordinaire du fait de ses implications en recherche, ou dans les applications industrielles, puis commerciales, etc. Avec cette diffusion possible des concepts et des termes d'abord spécialisés, le terme spécialisé rejoindra le grand public, en se banalisant.

De même que certains répertoires «généralistes» n'hésitent pas à jouer un rôle «prescriptif», comme le *Petit Robert* (1996) qui note volontiers «abus.» pour «abusif», et «abusiv» pour

(18) Voir aussi le numéro de *Terminologies nouvelles* consacré à l'«Implantation des termes officiels» (1994).

«abusivement», une forme de prescription se dégage des conseils prodigués aux spécialistes, avec des directives d'usage. C'est ce que fait le *Dictionnaire des termes officiels* en mettant en garde contre des «termes étrangers ou termes impropres à éviter ou à remplacer». C'est aussi ce que font la DGLF et la commission générale de terminologie à l'issue des travaux des commissions, s'il y a eu consensus, par les publications au *Journal Officiel* puis sous la forme de fascicules thématiques (*Vocabulaires*). On trouve ainsi, dans les fascicules de *Vocabulaires*, en notes, les précisions et recommandations suivantes: «On rencontre parfois aussi, moins bien, "organiseur"» (*Informatique*, 1998, *J.O.*; sous *agenda électronique*); «l'expression "ligne numérique à paire asymétrique" ne doit pas être utilisée» (*Internet*, 1999, *J.O.*, sous *raccordement numérique asymétrique*), «l'expression "pré-marketing" est impropre», «marketing direct" est impropre» (*Économie et des finances*, 1998, *J.O.*, respectivement sous *préplacement* et sous *vente directe*); «Il est déconseillé d'utiliser le terme "catalogue outdoor"» (*Économie et des finances*, 1998, *J.O.*, sous *catalogue du plein air*); «le terme "connectivité" est déconseillé» (*Informatique*, 1998, *J.O.*, sous *connectivité*); «le terme "inforoute", qui n'est pas recommandé» (*Internet*, 1999, *J.O.*, sous *autoroutes de l'information*); «le terme "stewarding" ne doit pas être utilisé» (*Sport*, *J.O.*, sous *stadiaire*).

C'est ce que peuvent faire aussi certaines commissions, par un avis dans la presse spécialisée sur la dénomination de concepts émergeant dans leur domaine. Une telle action est prometteuse dans le domaine médical mais il ne semble pas que les physiciens, ou les mécaniciens s'exprimant dans notre idiome bénéficient d'un média journalistique s'adressant à un public suffisamment large: il n'y a pas de revue généraliste de physiciens. *La Recherche* s'adresse

aux spécialistes de tous les domaines confondus.

Les nombreux spécialistes auxquels nous avons eu l'occasion de nous adresser ces dernières années, à l'occasion de la rédaction de définitions ou du choix de termes à faire entrer dans les dictionnaires, sont particulièrement intéressés par le processus d'entrée des mots de leurs domaines dans le dictionnaire, d'attestation de sens et d'emplois spécialisés. Aujourd'hui, lorsque nous posons des questions aux spécialistes au sujet de termes nouveaux à proposer comme alternatives «linguistiquement efficaces» aux anglo-américanisms déjà en usage, afin de corriger ou de faire agréer un usage naturel spontané, nous obtenons deux types de réponses très distinctes. Les uns sont sceptiques, car l'idée est répandue, dans la communauté scientifique, que toute action sur la langue «venue d'en haut» demeure inconnue, inutile et sans effet. Les autres sont prêts à collaborer et s'emparent même, à l'occasion, des termes que nous soumettons à leur jugement et nous disent les réutiliser aussitôt car ils leur manquaient. Il est donc bienvenu de créer ou d'aider la création et le développement de nouveaux outils méthodologiques pour la néologie, sa collecte, sa création, sa description, l'incitation à son abandon ou à son usage et finalement sa diffusion. On y arrivera en encourageant la collaboration entre les scientifiques et les linguistes, dans un commun respect.

Danielle Candel,
Service de néologie et terminologie,
Inalf,
CNRS,
Paris,
France.

Bibliographie

Candel (D.), 1994: «Vers un dictionnaire de français scientifique et technique», dans *Français scientifique et technique*, Candel (D.) éd., Didier Érudition, p. 185-201.

Candel (D.), à paraître: «Français scientifique et technique», dans *Histoire de la langue française 1945 – 2000*.

Candel (D.), Carton (J.), Marchaudon (P.) et Tolédano (V.), 1992: *Dictionnaire de sigles, domaines économiques et sociaux*, Conseil économique et social, CNRS, INaLF, La Maison du Dictionnaire, XXIV- 744 p.

De Schaetzen (C.), 1990: «Besoins et ressources belges en terminologie», dans *Terminologies nouvelles*, 3, p. 55-57.

Humbley (J.), à paraître: «Terminologie», dans *Histoire de la langue française 1945 – 2000*.

Lerat (P.), 1994: «Terminologie vs lexicographie», dans Candel (D.) éd., *Français scientifique et technique*, Didier Érudition, p. 27-36.

Rapport annuel de la commission générale de terminologie et de néologie, 1998: Premier Ministre, DGLF, 47 p.

Rapport d'étape, 1997: CNRS – INaLF, UPR 9017, 52 p.

Spillner (B.), 1994: «Terminologie et connotations», dans Candel (D.) éd., *Français scientifique et technique*, Didier Érudition, p. 53-62.

Terminologies nouvelles, 1994: *Implantation des termes officiels*, actes du séminaire (Rouen, décembre 1993), 165 p.

Tableau 1. *Principaux objets d'étude et recherches engagées.*

Principales actions engagées	Recherche d'un terme français équivalent à un terme angloaméricain			
	Assistance à la veille terminologique et néologique	déjà en usage en français	en train de s'implanter en français	synonyme d'un autre terme français déjà en usage
néologisme de forme	oui	oui	oui	oui
néologisme de sens	oui		oui (partiel)	oui (partiel)

Tableau 2. *Quelques phénomènes observés dans les principales actions de terminologie ponctuelle.*

Principales actions engagées	Élaboration ou reformulation d'une définition			Examen ou réexamen	
	pour une unité linguistique ou terminologique dont le sens a évolué	pour une unité linguistique ou terminologique dont le domaine d'emploi a changé	qui s'adresse aussi au grand public	d'un terme dont le domaine d'emploi a changé	d'une forme dont la graphie a changé
Variabilité définitionnelle			oui		
Incidence diachronique dans les usages observés	oui	oui		oui	
Incidence du domaine ou du degré de spécialisation		oui	oui	oui	oui
Variabilité formelle					oui

Tableau 3. *Trois expériences: lexicographie générale, lexicographie de spécialité, terminologie.*

Trois expériences caractéristiques à l'INaLF		Lexicographie générale	Lexicographie de spécialité	Terminologie ponctuelle
objet d'étude				
Présentation	listes de termes proposées par domaines	non	souvent	oui
	termes extraits de corpus	oui	souvent	rarement
Idiome d'origine des termes	termes français	oui	oui	souvent
	emprunts angloaméricains	rarement	souvent	souvent
	termes étrangers	non	rarement	oui
Méthode d'étude				
recherches	dans des répertoires de mots	oui	oui	oui
	dans des textes	oui	oui	souvent
	complémentaires	rarement	oui	oui
Mode de classement				
répartition des termes par domaines		souvent	oui	oui
Respect de l'usage				
constat et mise en avant de l'usage		oui	oui	?

La conversion et l'interface infinitif/impératif/nom dans la terminologie du micro-ordinateur

La formation terminologique emploie, entre autres, la conversion (dérivation impropre, hypostase) pour répondre aux besoins de dénomination et d'enrichissement lexical.

En principe, ce transfert transcategoriel aboutit, dans de nombreux cas, à l'adjectivation ou à la nominalisation.

Cette formation lexicale et terminologique révèle donc un mouvement entre les classes lexicales.

Dans la terminologie moderne du micro-ordinateur telle celle du texte intégral (continu, suivi) accompagnant le coffret du traitement de texte *Microsoft Word 6.0* (1994), ce mouvement provoque un affaiblissement de l'opposition infinitif/impératif/nom.

Cet affaiblissement permet, et en anglais et en français, la coalescence entre ces trois fonctions et ce type de formation est investi d'une capacité productrice bénéfique parce qu'économique.

Termes-clés : terminologie, conversion, dérivation, création lexicale, catégories grammaticales.

La morphologie des unités linguistiques telles qu'elles sont employées dans la langue technoscientifique nous offre l'occasion d'observer de nombreux phénomènes à la lumière des principes permettant de créer une variété d'éléments lexicaux. Les procédés que l'on observe mettent parfois à l'épreuve certains principes de création lexicale, une activité qui est souvent impétueuse sinon inattendue, d'où l'étiquette de créativité qu'on lui confère. Mais comme on le verra dans cette étude, il y a correspondance et concordance entre la langue usuelle et la terminologie selon les procédés employés; en fait, celle-ci s'enracine dans celle-là et lui sert de point de départ. Pour ce qui est de la conversion face aux catégories grammaticales que nous soulevons, il y a lieu d'examiner les effets réels et potentiels qui s'imposent.

La conversion décrit deux termes de forme semblable (et d'une origine étymologique commune) qui finissent par appartenir à deux classes lexicales différentes. Ce type de dérivation a pour synonymes l'hypostase et la dérivation impropre et même *la dérivation zéro* (en anglais *zero derivation*). Cette formation marquée par l'absence d'uffixes et par le changement de catégorie grammaticale rend possibles des constructions complexes avec de grandes implications syntaxiques comme nous le constaterons dans des syntagmes lexicalisés. On peut critiquer ce mode, mais il fournit une grande souplesse et exploite le

concept d'économie qui fait qu'un nombre restreint d'éléments formateurs permettent la création de multiples constructions terminologisées.

En principe, ce transfert transcategoriel aboutit, dans de nombreux cas, à l'adjectivation ou à la nominalisation, la signification se limitant à un seul champ sémantique. Mais ce faisant, le processus dérivationnel implique une dynamique associative, paradigmatique.

Cette formation lexicale et terminologique révèle donc un mouvement entre les classes lexicales. Il faut faire mention du fait que cette formation, loin d'être exclusive à la terminologie, est un des procédés morphologiques en grammaire, à la fois en français et en anglais. Mais dans la terminologie moderne du micro-ordinateur, ce processus provoque un affaiblissement de l'opposition infinitif/impératif/nom, observation qui découle de la prépondérance des noms et d'autres formes nominalisées en terminologie et en terminographie. Cet affaiblissement permet en français et en anglais la coalescence entre ces trois fonctions, et ce type de formation est investi d'une capacité productrice bénéfique parce qu'économique. En revanche, certains estiment que cette formation risque de perturber la grammaire dérivationnelle.

En français technoscientifique, la création lexicale occupe une place importante et les procédés qui permettent la créativité restent encore à expliciter, compte tenu de leur rapport avec certains phénomènes

linguistiques et extra-linguistiques. Mais sans conteste, ces mécanismes s'effectuent conformément à une double équation dénomination/définition, dénomination/conceptualisation.

Nous nous proposons donc d'aborder l'interface de ces catégories grammaticales au sein de la terminologie française du micro-ordinateur (donc de la micro-informatique). Pour ce faire, on renvoie aux catégories ci-dessous.

1 Participe présent (Ppr) > Nom (N)

Ce qui suit constitue l'illustration du participe présent qui aboutit au nom: *courant* (ppr>n) *alternatif*, *courant flottant*, cas de double dérivation impropre où les deux éléments du syntagme lexicalisé révèlent la dérivation impropre par excellence ppr>n; ppr>adj. dans le fonctionnement syntaxique du deuxième élément; *courant porteur* (premier élément: ppr>n; deuxième élément: n>adj. postposé); un syntagme peut même se créer à partir de cette combinaison nominale pour donner: *courant porteur commun* produisant alors une formule de type n+adj.+adj. On peut citer aussi cet élément composé: *intrants-extrants* (ppr>n: formes substantivales formées à l'occasion de l'avènement des concepts qu'elles décrivent, à savoir les données qui entrent et qui sortent (*information*, *données* pp>nfpl) du système informatique. Les *intrants*, par un processus métonymique, ont également trait aux aspects du *matériel*: *clavier*, *souris*, *lecteur optique*, *unité d'entrée* (2^e élément: n<pp) tandis que les *extrants* désignent aussi *écran*, *imprimante* (nf<vb [ppr]+e), *sortie* (pp+e>nf) *sonore*, etc.

On voit donc que ces désinences des participes nous fournissent les

suffixes formateurs: -ant(e) et -é(e). Les formes dérivées impropres qui en résultent sont utiles pour former les syntagmes lexicalisés: *imprimante* (nf<ppr [adj]) dans *imprimante sans impact*, *imprimante à impact*, *imprimante à jeu électrostatiques*, *imprimante laser* (ou le second élément est d'abord emprunt et sigle lié, intégré et non épilé anglais jouant ici le rôle du déterminant dans un syntagme lexicalisé). Dans le *Guide de l'utilisateur Microsoft Word* (1994), dorénavant le *Guide*, notre source primaire, nous avons pu relever une quantité modeste d'exemples autonomes tels *Assistants*, *exposant*. Par ailleurs, nous avons trouvé ce procédé assez productif dans les créations syntagmatiques lexicalisées comme dans les cas suivants.

2 Ppr > Adjectif (Adj) (surtout dans les syntagmes lexicalisés)

Dans *marge flottante* (nf) on remarque que *flottante* (élément régisseur à l'intérieur d'un terme complexe) est d'abord participe présent transformé en adjectif, avec la marque du féminin par l'ajout de e qui assure la définition et la cohésion syntaxique du syntagme. Il en va de même des syntagmes *menu déroulant* et *barre d'outils flottante*. D'autres exemples sont: *statistiques parlantes*, *ce qui est très expressif* ou *qui se passe de commentaire*, cette qualité recherchée des données ou des entrées en général est un des *tests de facilité*; *option croissant/décroissant*, commande qui remplace parfois les formes infinitives impératives sous certains menus et dans certaines boîtes de dialogue. On retrouve également des syntagmes du type N+Prép+N+Ppr: *point d'insertion clignotant*. Pour ce qui est du manque d'accord ou de la rupture syntaxique et de l'emploi des majuscules initiales dans certains des exemples, nous

réserve nos commentaires jusqu'au passage sur les syntagmes lexicalisés.

3 Participe passé (Pp) > Nom (N)

Illustrons ce procédé par *accusé* (pp>n, comme dans *accusé de réception*), *données* qui figurent dans *base de données*, et *porteuse* comme dans *porteuse de données*, *modulation d'amplitude sans porteuse*, *modulation à porteuse inhibée*.

Il convient de signaler la présence de ce type de dérivation impropre dans notre source primaire, à savoir le *Guide*: le *souligné*, le *double souligné* (employé parfois comme descriptif de bouton), le *barré*, les *pointillés*, le *résumé* (qui peut englober les statistiques d'un document donné), les *coordonnées* (nfpl) et le *libellé*. Pour ce dernier exemple, le *Petit Robert* (1993:1277) indique que c'est un substantif dérivé de *libeller*, c'est-à-dire *rédigé dans les formes*. Il s'agit des termes dans lesquels un acte officiel est rédigé, comme par exemple dans l'emploi spécialisé le *libellé* d'un jugement. Ce n'est que par extension sémantique que nous arrivons à *libellé* d'une demande, d'une lettre. C'est dans cette optique que *libellé* est employé dans notre source primaire. On y trouve également: *saisie* comme dans, par exemple, *saisie de texte*, *de donnée*, *de dimensions*, *de glossaire*, etc., ce dernier étant synonyme de *entrée(s)* employé souvent en petites majuscules (*ENTRÉE*) pour indiquer le descriptif de la touche. Il y a des cas où ce mode permet des syntagmes du type *touche raccourci* où le deuxième élément est un participe passé (il peut être considéré comme régressif), et *mise à jour*, *mise en forme*.

4 Pp > Adj (composition et syntagmes lexicalisés)

4.1 N+Adj où l'adjectif déterminant est un participe passé

Ce procédé est productif et répond comme dans la langue usuelle au besoin de modification des noms qui constituent les éléments principaux du syntagme: *image incorporée; espaces réservés; listes hiérarchisées (ascendantes/descendantes); répertoire partagé; dossier partagé; paramètres définis ou prédéfinis; entrée prédéfinie; séparateur prédéfini; taille prédéfinie; ombrage prédéfini; bordures prédéfinies; plage définie; texte masqué; titre intégré; cellules référencées* (plutôt dérivé du nom *référence*); *cas désactivé; ancre verrouillée; zone ombrée*; et expressions renversées telles *Word installation; fenêtre séparée; bordure hachurée; élément dessiné; ombre portée; document récupéré* (sauvegardé); *recherche poussée; polices intégrées; style intégré; texte justifié, texte non justifié (alignement de texte); marque déposée; doubles touches intégrées*. On relève ce procédé dans l'élément d'emprunt composé dans ce syntagme: *PostScript encapsulé*, marque déposée écrite avec des majuscules à l'initiale des mots constituants.

4.2 N+Adj où l'adjectif provient de la conversion du participe présent

Le cas des modificateurs à source participiale est très fréquent et comprend les participes présents: *liste déroulante* que l'on remarque dans *champ Liste déroulante; texte environnant; date courante; heure courante; titre courant; disque tournant; parenthèse fermante*, d'où

on peut s'attendre à *parenthèse ouvrante; lettres résultantes*. On rencontre moins fréquemment le format Adj+N: *gros titre; double interligne*.

L'adjectivation s'effectue bien sûr par le truchement du suffixe -able: *caractère non sécable, espaces sécable; caractère non imprimable, zone non imprimable, polices dimensionnables; ordinateur portable* devant *ordinateur portatif*.

5 N1+N2

Cette construction syntagmatique entraîne l'imprévisibilité parce qu'il est parfois difficile d'établir les rapports hiérarchiques et sémantiques entre N1 et N2 comme on le remarque dans les termes composés tandem. On essaie de déterminer les régissants et les régisseurs mais le rapport de subordination peut aller dans les deux sens. Citons par exemple *carte son* qui ne devient clair qu'après l'insertion du joncteur «de» dont la chute est à la base de l'ambiguïté ressentie; *raccourci clavier; dossier système; fichier texte* et *fichiers texte* dont le modificateur reste invariable, soulignant ainsi sa nature nominale et sa conversion incomplète; *alimentation papier* par le truchement du bac; *documents maîtres; configuration système; pilotes écran* (deuxièmes éléments invariables); *paramètres clavier* sans accord du pluriel pour N2; *service clientèle; logiciel système; espace disque; lecteur réseau; trait connecteur* (N2 passerait facilement pour adjectif comme dans *dispositif convertisseur*); *version écran; fichier source; longueur page; champs Texte* parfois sans T majuscule; *Insertion Champ; type quadrillage* (formulaire); *formule politesse; Initiales auteur; style journal; mots clés* (parfois composé *mots-clés*); *conventions clavier; Assistant Lettre*

(lettres initiales majuscules); *touche RETOUR; touche ENTREE; service clients; lettre type*.

Cette variante de dérivation impropre figure dans des syntagmes lexicalisés parfois complexes. Nous citerons ici des exemples peu complexes, réservant les plus complexes pour la rubrique des créations syntagmatiques: *vice caché; système indéterminé; paramètres définis ou prédéfinis; barres d'outils intégrées (ancrées, empilées); dossier partagé; répertoire partagé; système d'aide étendue; listes hiérarchisées; espaces réservés*. Dans ces exemples, le terme en gros, c'est-à-dire l'ensemble de la désignation, est un nom et on y remarque une prépondérance du genre masculin. Le genre du déterminant dépend de l'élément régissant du syntagme. D'autres termes syntagmatisés semblables sont: *papier carboné* (nm), *commande imbriquée* (nf), *tube ombré, réseau décentralisé, ordinateur jumelés*. On voit facilement pourquoi Darmesteter (1877: 41) qualifiait d'une manière assez simple ces formations comme celles «[...] qui ne recour[ent] pas à des suffixes». Cette nominalisation semble très commode et productive en informatique sans grands retentissements sur la nominalisation des syntagmes lexicalisés qui contiennent ces adjectifs. Les éléments en deuxième position jouant le rôle de déterminant suivent le nombre et le genre des éléments surtout nominaux qui précèdent. Le genre des syntagmes prépositionnels s'établit à partir du caractère du premier élément.

6 Adj > N

Un exemple ici est *utilitaire* (adj)>*utilitaire* (n) ce dernier étant un dispositif qui, selon le cas, permet de réaliser une opération. Ainsi, on peut avoir un *utilitaire* qui permet

d'imprimer des étiquettes facilement et rapidement. Le passage de la forme adjectivale à la forme nominale révèle l'homographie comme dans le cas des noms suivants par rapport à la catégorie grammaticale dont ils sont issus: le *périphérique*; le *graphique*; le *gras*; l'*italique*.

7 N > Adj (syntagmes lexicalisés)

Nous avons vu précédemment les noms postposés jouant le rôle d'adjectifs dans la partie sur la formation par composition lorsqu'on analysait les termes formés de deux ou plusieurs noms à la fois. On y remarquera la valeur adjectivale du nom postposé modificateur dans son rôle de complément de nom à l'intérieur des syntagmes terminologisés. Leur conversion de la catégorie nominale à la classe adjectivale dans leur fonctionnement syntaxique facilite la création de bon nombre de termes importants en terminologie de l'informatique. Dans ce cas, on pourrait parler de la composition ou de la syntagmatisation (s'il s'agit d'un syntagme) par dérivation impropre. La liste d'exemples que nous fournirons permettra de vérifier cette observation: *tabulation arrière* (n1+n2 où n2 = déterminant (adj, adv), *système anticopie* (on notera le caractère invariable de ce n2 dans son rôle d'adjectif malgré sa terminaison), *ingénieur programmeur*, *ingénieur logiciel*, *langage auteur*, *logiciel auteur*. Le *Guide* nous fournit parmi d'autres exemples: *fichiers texte*; *conventions clavier* où n2 ne s'accorde pas avec n1 du point de vue syntaxique; *Assistant Lettre* où n1 et n2 ont une lettre majuscule à l'initiale; *mode Refrappe* où n2 a une lettre initiale majuscule; *style journal*; *écran Word* où n2 = emprunt à l'anglais commençant par une lettre majuscule (marque

déposée); *Mots clés* où il y a accord entre les deux éléments; *fenêtre Word*: 2^e élément = emprunt, le 1^{er} élément représente un emploi métaphorique faisant appel à son équivalent anglais; *Police TrueType*: syntagme où le 1^{er} élément français se combine avec n2, élément d'emprunt composé postposé modificateur; *document Word*; *programme d'installation Word*; *documents maîtres*; *émulation de terminal ANSI*, *Windows NT* dont les sigles empruntés et n'ayant pas d'équivalents français dans le texte fonctionnent en tant que modificateur (rôle adjectival). Cette formation rentrerait plutôt dans la catégorie de la syntagmatisation des formants; c'est un des procédés qui permettent de contourner le besoin de créer de nouveaux éléments d'emprunt, sigles ou termes isolés, malgré l'apparence d'hybridation souvent critiquée dans les cas d'emprunt de souche anglaise. Dans la terminologie du traitement de texte qui nous préoccupe, cet emploi est privilégié à cause du souci de garder les noms déposés de plusieurs programmes appartenant au coffret, y compris les programmes supplémentaires indispensables.

8 Infinitif

• Infinitif (Inf) > Impératif (Imp)

Un autre cas frappant de dérivation impropre s'apparente (peut-être par analogie) à l'usage en sémiotique contemporaine qui recherche le sémantisme pur par son exploitation des dérivations non suffixées en suivant la formule Inf>N. C'est le cas de l'infinitif nominalisé (à source verbale infinitive) du type *l'avoir*, *le faire*, *le paraître*, *le valoir*, *le vouloir*, *le devenir*, *le pouvoir-faire*, *le savoir-faire*, *le savoir-connaître*, *le faire-connaître*, etc. Pour ce qui est de notre dépouillement, nous avons relevé les exemples qui suivent dans un texte

suivi (Armand et Bonneau: 1994) conçu d'un point de vue pédagogique et destiné à démystifier et à rendre plus accessible l'apprentissage des principaux systèmes d'exploitation et d'application ainsi que certains logiciels de traitement de texte les plus employés. Les exemples ci-dessous révèlent la transformation de l'infinitif en impératif, dérivation que l'on remarque d'ailleurs dans la langue usuelle: *Personnaliser*; *Rechercher*; *Remplacer*; *Sélectionner*; *Enregistrer sous*; *Insérer*; *Définir*; *Verrouiller*; *Déverrouiller*; *Fermer*; *Centrer Page*; *Désactiver*; *Annuler*; *Créer*.

On note dans ces infinitifs devenus impératifs que les lettres initiales sont toutes en majuscule. Ces mots de commande se trouvent dans différents menus et sont faciles à comprendre; la sémantisation de l'infinitif n'entraîne aucune complexité dans le maniement de la terminologie telle que l'emploi Armand et Bonneau. L'utilisateur non spécialiste qui cherche à apprendre cette terminologie et qui s'initie au fonctionnement de l'ordinateur ne ressent pas de grands ennuis à cet égard. Cet emploi confirme encore une fois que toute terminologie, surtout celle qui nous intéresse en l'occurrence, puise dans la langue usuelle.

À la différence de l'impératif que l'on trouve à l'intérieur du texte suivi de notre source de corpus primaire, l'impératif qui signale les commandes accessibles par le truchement des boutons ou par les options de commandes marquées sur les boutons dans les boîtes de dialogue et sur les boutons des barres d'outils, fonctionnant comme l'équivalent graphique d'un menu (qui regroupe en principe ces commandes ou options), se manifeste par le biais de l'infinitif. Cette forme infinitive est favorisée grâce à la neutralité par rapport au sujet visé. Autrement dit,

l'infinif ne renvoie à aucun pronom personnel.

À partir du *Guide* (1994) nous citerons les exemples en fonction des catégories suivantes.

- Tout d'abord, des formes infinitives simples pour signaler essentiellement des commandes: *Fermer; Enregistrer; Imprimer; Quitter* sous le menu *Fichier*. Sous *outils de vérification* nous relevons par exemple *Ignorer; Suggérer; Annuler; Ajouter; Supprimer; Modifier; Restaurer; Expliquer*, etc. Pour le lancement du programme, *Lancer; Exécuter*. Pour l'édition du document, *Rechercher; Remplacer; Atteindre; Couper; Coller; Copier; Couper-Coller*.
- La formule Inf+Prép nous donne les formes suivantes: *Enregistrer Sous* sous *Fichier; Remplacer par*.
- Pour la combinaison Inf+Pronom: *Accepter tout; Rejeter tout; Remplacer tout; Enregistrer tout; Sélectionner tout*.
- Pour la séquence Inf+Adv, on trouve *Ignorer toujours; Remplacer partout* lors de la vérification de document.
- Pour l'ordre Inf+N: *Ajouter les mots; Marquer les corrections; Fusionner réunions; Comparer version; Coller lignes; Supprimer lignes; Fermer un menu; Reproduire la mise en forme*.
- Avec un peu d'expansion, on a le format Inf+Inf+N: *Faire apparaître les corrections à l'écran* (ou à *l'impression*) une forme de construction périprastique suivie d'un nom et rendu en anglais par *Format Painter*.
- Pour la catégorie Inf+LocPrép, nous avons *Remplacer en cours de frappe*.
- En dernier lieu, pour Adv+Inf, il y a *Toujours Substituer*.

Les catégories ci-dessus indiquent la souplesse et l'économie de cette forme de dérivation neutre. Puisque l'infinif ici remplace facilement l'impératif, comme c'est d'ailleurs le cas dans la langue commune, nous craignons que la critique de cette formation lexicale ne soit plus nécessaire. Ce mode de dérivation impropre nous rappelle les mots anglais des mêmes commandes

comme *print, copy, save, replace*, etc., où la morphologie des mots, du moins la graphie, ne montre aucune différence entre l'infinif et l'impératif. On voit donc que les verbes peuvent être directement nominalisés ou adjectivés puis nominalisés; ils figurent dans les syntagmes complexes lexicalisés où les éléments qui changent ainsi de catégorie grammaticale remplissent la fonction déterminative dans l'agencement syntaxique. Ce fonctionnement syntaxique est important car on voit à la place des déterminants les verbes et noms qui fonctionnent en position postposée comme le feraient les éléments de type proprement adjectifs.

9 Conclusion

Devant ces constatations, on relèvera que Arrivé *et al.* (1986: 334) voient l'infinif comme «une forme du verbe ayant la particularité de ne connaître de marques ni de personnes, ni de nombre, ni de temps». Certains traits morphologiques confèrent à l'infinif sa forme verbale. Il est également susceptible, du point de vue syntaxique, d'assumer toutes les fonctions du nom. Il est capable de fonctionner comme l'équivalent d'un impératif dans l'expression de l'ordre, de transmettre une consigne ou une interdiction. Pour Riegel *et al.* (1994: 407), le caractère injonctif ou impératif est associé à la gamme des actes directifs où le locuteur veut obtenir un certain comportement de son interlocuteur. Nous signalerons à ce stade que dans la terminologie française de l'informatique et selon l'optique qui sous-tend la conception de cette interrogation, la présence du locuteur est presque toujours latente. Le locuteur maintient une distance entre lui-même et l'usager du système. S'il y a proximité, c'est

essentiellement entre le système informatique et celui qui s'en sert, élément qui est mis en relief à travers la serviabilité et la convivialité recherchées par les utilisateurs des dispositifs informatiques. Poursuivant leur réflexion, Riegel *et al.* (408) ajoutent que l'impératif est «doublement lacunaire» car il est limité à des personnes, et il est restreint au groupe verbal sans groupe nominal sujet exprimé. Par ailleurs, l'infinif (251; 334) est considéré comme la forme nominale du verbe. C'est là où réside son attrait en terminologie de l'informatique surtout dans la mesure où on a besoin d'établir une communication ou un dialogue virtuel entre l'utilisateur et le système informatique.

Comment cet emploi de l'infinif entrerait-il dans les dictionnaires techniques spécialisés traitant de l'informatique, par exemple? Son absence, ou son exclusion, si on lui a refusé une considération terminographique (on peut noter que les formes verbales existent dans les dictionnaires généraux de langue), est-elle attribuable à sa marque d'ambiguïté? Quoi qu'il en soit, on ne peut pas nier que l'infinif peut être employé à la place de l'impératif pour exprimer un ordre, un conseil avec la particularité graphique de lettres initiales capitales.

Au point de vue de la communication, la conversion nous rappelle les six fonctions de Jakobson (1963: 213-218; cf. Kocourek 1991: 58) en particulier la fonction conative (injonctive, appellative). L'utilisateur (ici le récepteur, le lecteur ou le destinataire) se servant du micro-ordinateur est appelé, même en l'absence apparente de l'émetteur, à exécuter des opérations pertinentes. Ce qui frappe est que l'interaction qui s'effectue entre le système informatique s'établit non à l'aide de l'impératif, mais grâce à l'infinif.

Malgré les insuffisances apparentes, après la création de ces éléments, il incombe au linguiste/terminologue d'élaborer, d'élucider le processus de création lexicale et d'établir l'interface entre les catégorisations grammaticales face aux servitudes de la grammaire de la langue concernée et de démontrer l'étendue de l'emploi en question. Jusqu'ici, les aspects portant sur les éléments verbaux ont reçu peu d'attention, au profit des nominalisations.

Il nous semble qu'il s'agit d'un emprunt sémantique intralinguistique utile pour la terminologie de l'informatique usitée ces jours-ci par un effectif considérable de personnes, d'où le concept de la sociabilisation de l'informatique. Si nous croyons à la notion que le français préfère dénommer les concepts plutôt que de les décrire, nous comprendrons pourquoi la terminologie française en général tend à s'emparer de toute ressource disponible y compris des anglicismes, malgré les mesures extralinguistiques mises en place pour endiguer les éléments inhabituels et étrangers. Il est également évident qu'il faut manier et travailler toute la gamme sémiotique d'une langue pour atteindre sa capacité dénommatrice et formatrice potentielle. Sinon, on finit par croire à son insuffisance et à son incapacité face aux nouveaux concepts, surtout des concepts technoscientifiques qui exigent une connaissance de pointe abstraite et exacte.

*Emmanuel Aito,
Département de français,
Université de Regina,
Regina,
Saskatchewan,
Canada.*

Bibliographie

St-Pierre (A.), Bonneau (A.), 1994: *La maîtrise du micro-ordinateur*, Boucherville, Éditions Vermette Inc.

Arrivé (M.) *et al.* 1986: *La grammaire d'aujourd'hui. Guide alphabétique de linguistique française*, Paris, Flammarion.

Darmesteter (A.), 1877: *De la création actuelle de mots nouveaux dans la langue française et des lois qui la règlent*, Paris, Vieweg (Genève, Slatkine Reprints, 1972).

Kocourek (R.), 1991, *La langue française de la technique et de la science*, 2^e éd., Wiesbaden, Brandstetter.

Microsoft Corporation, 1994, *Quick Results: Microsoft Word 6.0*, USA, Microsoft.

Microsoft Corporation, 1994, *User's Guide: Microsoft Word 6.0*, USA, Microsoft.

Microsoft Corporation, 1994, *Aperçu: Microsoft Word 6.0*, Ireland, Microsoft.

Microsoft Corporation, 1994, *Guide de l'utilisateur Microsoft Word 6.0*, USA, Microsoft.

Petit Robert, 1993, Rey-Debove J. et Rey Alain, dir., Paris, Le Robert.

Rieget (M.) *et al.*, 1994, *Grammaire méthodique du français*, Paris, PUF.

Néologismes, internationalismes et mondialisation

La mondialisation laisse ses traces sur toutes les langues du monde. Les nouvelles technologies permettent désormais de suivre cette évolution linguistique de façon bien plus précise que dans le passé. Alessio Petrali propose un système de veille néologique concernant cinq communautés linguistiques, et livre les premiers résultats de l'enquête sur le phénomène de la mondialisation. Cet article est une mise à jour de la communication orale présentée à Barcelone en 1997 à l'Université Pompeu Fabra.

Termes-clés:
veille néologique; corpus;
néologie; multilinguisme;
néologisme formel;
néologisme sémantique;
néologisme d'emprunt.

1 Introduction

Les langues et leurs locuteurs sont dans un état de perpétuelle confrontation: on peut dire que leur diversité est une confirmation de leur identité. Cette confrontation, qui ne manque pas de susciter beaucoup d'intérêt, peut faire l'objet d'études systématiques: traductologie, lexicographie, enseignement des langues, par exemple. La mondialisation de ces dernières années a ravivé l'intérêt pour ces confrontations interlinguistiques, et pour les études contrastives en général, en réponse à l'internationalisation croissante de la société qui crée le besoin de compétences linguistiques accrues.

La linguistique de corpus ouvre de nouvelles perspectives dans le cadre des études contrastives. Il est désormais possible d'accéder à des corpus significatifs numérisés de types variés, disponibles en particulier sur Internet et sur des supports numériques (cédéroms); le coût d'accès est variable, mais actuellement très faible, voire nul. Il existe des logiciels puissants d'interrogation textuelle, des interfaces conviviales et tout cela à des prix abordables. Cependant la constitution de corpus électroniques d'une certaine taille, choisis avec des critères cohérents et convertis dans des formats spécifiques (permettant l'interrogation par des logiciels appropriés, par exemple), est une affaire autrement complexe et qui a son prix. Cette première phase

longue et onéreuse est le tribut que le linguiste doit payer afin d'obtenir des possibilités de recherches commodées et variées encore difficilement imaginables il y a peu de temps, ainsi, par exemple, la collaboration à distance avec d'autres chercheurs.

2 Parcours de recherche

La néologie internationale est désormais en mesure de bénéficier des outils de traitement automatique et de procéder à l'exploitation de grands corpus. On mesure aujourd'hui le chemin parcouru lorsqu'on consulte les nombreux projets présentés dans le recueil de Braun-Schaeder-Volmert (1990), référence privilégiée pour les internationalismes. De plus, les corpus informatisés et les résultats de leurs analyses peuvent être transmis et rendus accessibles assez facilement et à peu de frais dans le monde entier, d'autant plus que des liaisons efficaces (Internet), des ordinateurs puissants et des mémoires de masse volumineuses sont aujourd'hui à la portée du chercheur. Ces moyens de communication modernes ont une autre caractéristique: la rapidité, qualité non négligeable lorsqu'il s'agit de pister les changements quotidiens dans des langues variées, témoignages de la néologie internationale.

Nous partons de l'hypothèse que la néologie internationale, basée essentiellement sur l'anglais, peut être diffusée quasi instantanément sur toute la planète; la télévision atteint un public de masse, les nouveaux médias une élite émergente. Les mots,

comme les images, sont donc aujourd'hui virtuellement omniprésents, et à la disposition des masses, contrairement à ce qui se passait autrefois lorsque les contacts internationaux constituaient le pré carré des personnes instruites. Il est intéressant à ce propos d'examiner comment on passe du mondial au local et vice versa. Ce qui nous a poussé vers les internationalismes était notre travail sur l'italien régional du Tessin, qui comporte de nombreux tessinismes qu'on peut appeler «d'anticipation» (Petralli 1990: 126-129). En effet, au Tessin, on relève des néologismes italiens tels que: *legge quadro* – de *loi cadre, piano direttore* – de *plan directeur, educatore*, de *educateur, riservazione* – de *réservation, compostaggio* – de *compostage, polluzione* – de *pollution, suggestione* – de *suggestion, boicotto* – de *boycott, palette* – de *palette, ordinatore* – de *ordinateur*, résultat du contact quotidien que l'italien régional du canton suisse entretient depuis longtemps avec deux grandes langues de culture européenne, le français et l'allemand⁽¹⁾. Ce qui se passe au Tessin n'est pas atypique de phénomènes analogues en Europe. On constate effectivement de nombreux parallèles dans le lexique intellectuel, la langue de la médecine ou des sciences des langues européennes. Ce qui est plus étonnant, en revanche, est le fait que la plupart des études sur les

internationalismes proviennent presque exclusivement de la linguistique allemande. La raison est sans doute l'aspiration à l'ouverture supranationale, qui a animé le monde universitaire allemand depuis la fin de la Seconde guerre mondiale.

Le projet d'un observatoire de néologie internationale s'est pourtant révélé bien plus difficile à réaliser que prévu. Il nous a fallu privilégier la néologie qualitative (les signifiants dans différentes langues d'un petit nombre de néologismes choisis et en particulier de *globalisation mondialisation*, mais aussi *multimedialità* et *interattività*, etc.) plutôt que les aspects quantitatifs, ou des listes plurilingues d'internationalismes néologiques résultant d'un dépouillement automatisé.

3 Des corpus parallèles aux corpus comparables

L'exploitation de corpus informatisés commence en 1961 avec la *Brown University Corpus of American English*, suivie vers la fin des années 60 du *LOB Corpus*, équivalent britannique de la première. Au début les corpus numérisés sont monolingues, mais rapidement on passe aux bilingues, visant une application lexicographique, traductionnelle ou pédagogique. Un des corpus bilingues le plus célèbre est la transcription des débats parlementaires canadiens, source inépuisable de textes politiques en français et en anglais. Il s'agit de corpus parallèles, ou d'une collection de textes dont les uns sont les traductions des autres. Beaucoup de corpus parallèles plurilingues sont produits de nos jours par l'Union européenne et plus récemment par la Suisse. Au début de nos recherches

nous avons utilisé certains corpus parallèles afin de tester leur pertinence pour les recherches de néologismes internationaux. Dans la recherche d'équivalents, l'avantage des corpus parallèles est indéniable sur le plan quantitatif, mais on se heurte rapidement à un inconvénient de taille. Les corpus bilingues et multilingues consistent généralement en une langue source et en une ou plusieurs langues cibles, celle(s)-ci étant le fruit d'une traduction. Il en résulte des effets de distorsion du fait de la traduction, fortement influencée par la langue source. Le lexique d'une traduction n'est donc pas représentatif de celui de la langue cible. Dans le cas des corpus comparatifs, cependant, nous obtenons des textes qui traitent du même sujet dans plusieurs langues et qui mettent en lumière des analogies de différents types, pas seulement lexicales, témoignages authentiques des grands mouvements d'hybridation culturelle et linguistique qui nous intéressent. Plus généralement, l'exploitation de corpus comparables a séduit de nombreux chercheurs et pour beaucoup d'utilisations ils représentent le *nec plus ultra* dans le cadre des corpus plurilingues.

4 Choix de secteurs et de domaines d'emplois

Après avoir testé différents corpus parallèles plurilingues, surtout dans les domaines de l'écologie, de l'économie et du droit (pour ce dernier, le recueil plurilingue de textes juridiques de l'Union européenne *Celex*, cf. <http://europa.eu.int/celex> est très utile), nés dans le contexte de l'Union européenne, nous avons décidé de nous concentrer sur des corpus comparables. Celui que nous avons constitué est tiré de sources

(1) On rappelle que le canton du Tessin, où on parle un italien régional tessinois, est l'un des vingt-six cantons suisses, qui réunit 300 000 habitants sur 2 811 kilomètres carrés. L'italien est considéré comme langue officielle, à côté du français et de l'allemand, sans oublier le romanche (voir *Terminologies nouvelles* 12 : 59).

journalistiques et comporte cinq langues (*cf.* tableau 1). Le choix de textes écrits pris dans des journaux compromet quelque peu la représentativité de notre corpus, dans la mesure où le langage journalistique représente un registre soumis à ses propres règles. Ainsi, tout en s'occupant potentiellement de tout, le langage des journaux serait plus apte à exprimer certains sujets plutôt que d'autres et à en parler d'une certaine façon (le style journalistique, qui peut d'ailleurs varier d'un journal à un autre). Pour ces raisons il faut reconnaître que certains types d'écrits sont presque totalement absents (littéraire ou poétique par exemple), alors que d'autres sont surreprésentés (politique et économique).

En tout cas, après avoir rassemblé un premier corpus, nous avons privilégié les articles qui traitent d'une façon ou d'une autre de la mondialisation. Mettant *mondialisation* en relation avec d'autres mots-clés (voir tableau 2), par exemple avec *style de vie*, *mode de vie* et leurs équivalents dans les différentes langues, avec le préfixe *cyber-* (dans les formes équivalentes en italien, allemand, anglais et espagnol), nous avons pu rassembler un nombre conséquent d'articles capables de donner la mesure de certaines tendances générales et actuelles. De telles tendances seraient, dans notre première hypothèse, à l'origine de toute une série de néologismes, qui représentent en quelque sorte certaines frontières de la nouvelle société de la communication, au moins dans l'imaginaire véhiculé par les quotidiens prestigieux cités dans le tableau 1.

5 Corpus en cinq langues

Le corpus en cinq langues (italien, français, allemand, anglais,

espagnol) comporte des articles sur ce qu'on peut appeler de façon large, «style de vie de la mondialisation et des nouveaux médias». Nous avons aussi créé un corpus plus restreint, consacré uniquement aux articles qui, toujours dans la même période (1992-1995) et dans les mêmes journaux, comportent le mot italien *globalizzazione* ou l'équivalent dans les autres langues.

Le support privilégié pour ce traitement n'est pas le papier mais bien sûr le cédérom. En plus des cédéroms, de nombreux textes journalistiques, surtout récents, peuvent être téléchargés à partir d'Internet. Cette option n'est pas une solution de facilité : le langage HTML des pages de la toile engendre des difficultés supplémentaires quand il s'agit d'obtenir un texte propre qu'on puisse interroger avec profit. Il serait également possible de s'adresser au fournisseur de bases de données qui permettent de rechercher des journaux en texte intégral, mais les coûts peuvent être relativement élevés, surtout pour des journaux à caractère économique. Après un long travail de recherche et de téléchargement de données, nous sommes parvenus à un corpus comportant aujourd'hui 1 433 000 formes⁽²⁾ pour cinq millions d'occurrences (*cf.* tableau 1).

Ce corpus choisi et équilibré selon les critères simples déjà esquissés est interrogeable à souhait avec le logiciel d'analyse textuelle très puissant DBT⁽³⁾.

6 Quelques enquêtes effectuées sur le corpus

Notre hypothèse a été confirmée par les résultats des premiers sondages, résumés ici. Après avoir constitué le corpus, nous avons dans

un premier temps voulu procéder à toute une série d'expériences qui ont donné quelques résultats plus ou moins prometteurs. En outre, il faut noter que les secteurs qui nous intéressent le plus (les internationalismes néologiques) représentent seulement une des différentes possibilités d'approfondissement qu'un corpus comme le nôtre offre au chercheur. Pour cette raison, aux recherches ponctuelles sur la diffusion d'une néologie plus ou moins internationale, il est opportun d'ajouter des réflexions à caractère sociolinguistique. Il serait donc souhaitable qu'une observation permanente à plusieurs niveaux soit conjuguée à une mise à jour continue du corpus, et à ce propos, il est nécessaire de bien définir *archive* et *collection*.

Un corpus de données textuelles peut être défini comme une «collection» (ou «archive») organisée et articulée de textes portant sur un domaine donné, dans lesquels sont représentées une ou plusieurs

(2) Un repère quantitatif peut être déduit de Volz (1996: 162) lorsqu'il dit «in order to browse the texts accumulated within a month from an average daily newspaper, a lexicographer would have to examine about 40 000 word forms, of which only 2% prove to be genuine neologisms» (*cf.* Maier-Meyer 1995: 196). En tout cas, le nombre de 800 néologismes peut paraître excessif.

(3) Pour d'autres informations, voir <http://www.ilc.pi.cnr.it/dbt/pisystem.html>. Il s'agit d'un corpus original dont l'exploitation peut, à notre avis, constituer une contribution significative aux travaux sur les corpus comparables plurilingues.

langues⁽⁴⁾. Comme nous l'avons déjà vu, préparer un corpus significatif est une opération onéreuse en termes d'énergie et de temps. En effet, à tous les problèmes de réparabilité des données textuelles, il faut ajouter la nécessité de transférer ces données dans un format utile qui permette des interrogations complexes, grâce aux logiciels d'analyse textuelle dédiés comme DBT. Naturellement, une fois constitué, un corpus de qualité permet toute une série articulée et rapide de recherches naguère difficilement imaginables. La néologie en bénéficie grandement, comme l'annonce déjà Volz (1996).

(4) Pour une définition plus « officielle », mais de langue anglaise, voir le *Glossary of Corpus Linguistics*, <<http://cgl1.bham.ac.uk/glossary.html>> « un corpus est un recueil de textes représentant la langue naturelle, choisi dans le but de caractériser un état ou une variété de langues. En linguistique informatique moderne, un corpus comporte généralement plusieurs millions de mots. La raison en est qu'il est reconnu que la créativité de la langue débouche sur une variété d'expressions tellement vaste qu'il est difficile d'isoler les récurrences qui sont les indices de la structure lexicale de la langue ». Nous décrivons ici deux types de corpus. Le premier est le corpus « témoin », qui est un recueil fini de textes, choisi en général avec beaucoup de précaution et étudié en détail. Une fois qu'un corpus témoin est établi, il ne subit aucune modification ni aucun ajout. L'autre type de corpus qui prend forme actuellement, est le corpus « moniteur », qui réexploite des textes préparés sous forme numérisée dans d'autres buts : pour des imprimeurs dans le cas de journaux, de magazines, de livres, et de plus en plus, de traitements de textes; pour la langue parlée, il s'agit de transcriptions faites dans un but juridique ou bureaucratique. »

7 A la recherche de néologismes et d'internationalismes

En se référant par exemple à Petralli (1996b: 18), nous pouvons distinguer trois catégories de néologismes : le néologisme formel (nouveau signifiant pour un nouveau signifié), le néologisme sémantique (acquisition d'un nouveau signifié pour un signifiant déjà existant avec d'autre(s) signifié(s)) et le néologisme d'emprunt (importation plus ou moins directe d'un vocable étranger avec des adaptations plus ou moins marquées). Dans la première catégorie (néologismes formels) nous comprenons le néologisme de formation : mot nouveau créé en exploitant les matériaux déjà disponibles dans le système linguistique et les recombinaisons.

7.1 Néologismes formels

Le néologisme formel est le type qui peut être le plus facilement repéré par des moyens automatisés (cf. Volz 1996, « le néologisme nouveau-né »). En effet le nouveau signifiant peut être facilement identifié en confrontant deux corpus ou deux listes de mots. Puisque le néologisme est un mot nouveau, on présuppose que la confrontation est faite entre des corpus comparables qui se réfèrent à des laps de temps divers (diverses années d'un journal, par exemple). Voir l'annexe 1, qui présente tous les mots qui commencent par a, b, c, contenus dans c92 (*Corriere della Sera* 1992) et qui ne sont pas répertoriés dans *DMI* (*Dizionario macchina dell'italiano*).

Il nous a été possible d'utiliser ce dictionnaire électronique (*DMI*) de l'Institut de linguistique computationnelle du CNR à Pise, « reflet » lexical du vocabulaire italien de la seconde moitié des années 80.

La liste des mots résultant de cette comparaison comporte 373 unités, dont une vingtaine au plus sont susceptibles de susciter l'intérêt du chasseur de néologismes italiens⁽⁵⁾. Tous les autres sont par exemple des substantifs ou des adjectifs non lemmatisés (ou des erreurs de lemmatisation), des noms propres ou des éléments de syntagmes fixes (*chief* pour *chief executive*). On se rappelle qu'il s'agit ici de 1992, et les néologismes potentiels peuvent recouvrir des phénomènes sociaux tels que le sida, des adjectifs tels que *alternatif*, *acquariano* (*dell'era dell'acquarie*), des anglicismes désormais plus ou moins intégrés dans le lexique italien comme *antitrust* et *appeal*, *chat* et *anchorman*, des représentants de suffixes internationalement productifs, comme *-izzazione* dans *attualizzazione* et *compartimentalizzazione*, des constructions « médiatiques » comme *cybersex*, des termes techniques comme *coagratrice*.

Partant des réalisations néologiques de l'italien, on peut chercher dans les corpus des autres langues les équivalents potentiels, par exemple en recherchant tous les mots qui se terminent par *-izzazione*, *-ization*, *-isierung* – *isation/ization*, *-izacion*, (*globalizzazione*, mais aussi par exemple en italien *delocalizzazione*, en français *délocalisation*, en allemand *Entterritorialisierung*, en anglais *delocalisation*, en espagnol *deslocalizacion*), ou bien les

(5) Pour une comparaison purement indicative, on peut signaler les remarques de Volz (1996: 162), qui déclare que : « pour parcourir les textes accumulés d'un mois de publication d'un quotidien moyen, un lexicographe devrait examiner 40 000 mots-formes, dont seuls 2% seraient de vrais néologismes. »

anglicismes non adaptés (*appeal*)⁽⁶⁾, ou encore les préfixes productifs comme *ciber-/cyber-*. Dans les exemples de l'annexe 1, nous nous sommes limité aux trois premières lettres de l'alphabet, mais il est évident que de telles recherches peuvent être effectuées aisément, à condition de disposer de corpus significatifs comparables, et de corpus d'exclusion comme le *DMI*. De cette façon, les corpus des différentes années peuvent être paramétrés en comparant les listes de mots, vérifiant ainsi quels sont les mots effectivement présents pendant une période donnée et absents à d'autres.

7.2 Néologismes de formation

Très souvent, en particulier dans les discours spécialisés, les néologismes sont composés en faisant appel aux matériaux linguistiques déjà disponibles en latin, en grec et de nos jours en anglais (*cf. biomouse*). Il s'agit surtout de néologismes internationaux, puisque de nombreuses langues de spécialité ont tendance à uniformiser une bonne partie de leur terminologie. Parmi les néologismes internationaux qu'on pourrait rechercher avec profit, on peut privilégier les termes faisant appel à des préfixes actuels et potentiellement très productifs. Dans le sous-corpus ITA (pour la composition de ce corpus transversal de journaux italiens, voir tableau 2)⁽⁷⁾, on présente ici (*cf. annexes 2 et 3*) toutes les occurrences des mots commençant par *post-* et *bio-*, mais

(6) Dans notre corpus transversal q94 *appeal* paraît six fois dans c94 (dont trois fois *sex appeal*), trois fois dans S94 (aucun exemple de *sex appeal*) et six fois dans t94 (dont trois fois *sex appeal*). Outre les nombreuses attestations dans i94 et n94, *appeal* paraît une fois dans la *NZZ* (en parlant de *product appeal* entre guillemets dans le texte – de la nouvelle automobile Smart) et une fois dans *TAZ 94* (mais *sex appeal*).

on aurait pu aussi bien choisir *non-*, *euro-*, etc.

Il est évident que tous les résultats ne sont pas pertinents dans la perspective de la néologie potentielle. Certaines occurrences ne sont pas *a priori* néologiques (les noms propres, par exemple), tandis que d'autres nécessitent une vérification ponctuelle dans les contextes. Pour cette raison, à des fins d'illustration, nous présentons dans l'appendice 2 tous les contextes de *bio-*, par exemple *bioclimico*⁽⁸⁾, *Bioparc*. Il convient de souligner que le logiciel DBT permet de récupérer facilement tous les types de contextes souhaités, à partir de quelques mots dans un *kwic* (*Keyword in Context*). En pratique, le logiciel DBT offre la possibilité de consulter une grande quantité d'articles de journaux de façon modulaire et cela très rapidement, ce qui ne serait pas possible en feuilletant les journaux sous forme papier.

7.3 Néologismes d'emprunt

Rechercher automatiquement des emprunts récents, passés d'autres langues en italien, est souvent facilité par la forme particulière du signifiant non adaptée ou peu adaptée. Par exemple il peut être utile de rechercher tous les mots dans le corpus ITA (*cf. tableau 2*) qui comportent au moins une des cinq lettres *j* ou *k* ou *w* ou *x* ou *y*, sauf ceux qui commencent par une majuscule, afin d'éliminer les noms propres comme New York. Une autre possibilité propre à l'italien est la recherche de tous les mots qui se terminent en *-s*, qui, dans le corpus des journaux rédigés dans cette langue, sont au nombre de 124. Parmi ceux-ci, dans le corpus des textes de 1992 à 1995, moins d'une dizaine nous ont semblé intéressants. On a ensuite effectué une recherche peu ciblée en partant d'une liste de mots tirés de Barnhart 1995, dont nous avons retenu *chat*, *down*, *ranking*, *recall*, *wired*, *splatter*, *karaoke*.

Le fait qu'on relève ces mots dans bon nombre de journaux du corpus plurilingue q94 (*cf. tableau 2*) est un bon indice de leur «internationalité».

7.4 Néologismes sémantiques

Repérer automatiquement un mot déjà existant qui acquiert un nouveau sens est évidemment une tâche difficile. Il suffit à ce sujet de rappeler le témoignage de Volz (1996: 161) quand il souligne la difficulté d'identifier «les éléments lexicaux déjà existants, soit des mots simples ou composés, dont l'usage ou le sens a été déplacé, modifié, ou étendu». Il est évident qu'il est bien plus difficile d'identifier ces mots que ceux du premier type.⁽⁹⁾ Grâce au

(7) Rappelons que le corpus de référence et les différents sous-corpus, en particulier le sous-corpus ITA comprenant quatre années (de 1992 à 1995) de trois quotidiens italiens et le sous-corpus q94 qui regroupe l'année 1994 de dix journaux dans cinq langues sont encore à un stade relativement grossier/brut qui sera affiné dans les analyses à venir. Il subsiste par exemple des problèmes de caractères accentués et de chiffres, et dans de nombreux cas il faudra examiner de plus près l'apport des contextes. Pour cette raison les matériaux présentés en annexe ont un caractère encore provisoire.

(8) *posizione geografica, oltre che per motivi orografici e *bioclimici**, *ha conservato risorse naturali di eccezionale pregio*. SV19-09-94 *Dalle Alpi al mare un tuffo 461*.

(9) Pour Volz (1996: 161) les mots du «premier type» sont des «*new-born neologisms*», des «mots isolés qui ont été créés dans une langue particulière ou bien sans modèle préexistant, comme dans le cas de la plupart des sigles, soit dans celui des créations *ex nihilo*, comme par exemple *quark* en anglais, ou par des mots ou des parties de mots empruntés, y compris les radicaux, affixes etc, à d'autres langues telles que le grec, le latin ou de nos jours l'anglais.»

logiciel DBT, il est possible d'interroger notre corpus par rapport aux mots qui l'entourent, il y a toutes les chances qu'on le relève associé à un contexte plus ou moins déterminé. Dans le cas de *globalizzazione*, les relevés dans différentes langues concernent des nuances de sens qu'il serait prématuré d'interpréter ici de façon sociolinguistique. Dans le cas d'une recherche plus ciblée de néologismes sémantiques, il est évident qu'un changement évident répété des mots du contexte signalera un possible changement de significés. Pour en donner une idée, on peut supposer que si l'on relève *raton* en espagnol (pour rendre *mouse*), on pourrait envisager la «naissance» de *topo* en italien⁽¹⁰⁾.

8 Les guillemets

Les guillemets constituent pour plusieurs raisons un signe de ponctuation important en matière de néologie⁽¹¹⁾. Nous laissons de côté pour le moment toute réflexion de type quantitatif, qui serait à mettre en relation avec le volume des différents sous-corpus, pour dire que l'emploi des guillemets pour mettre en relief un mot (et ainsi une charge néologique éventuelle) est particulièrement courant dans les journaux italiens. En particulier, l'analyse faite sur q94 met en lumière le grand nombre d'occurrences relevées dans le *Corriere della Sera* de 1994 (c94). En général, il semblerait que l'emploi de guillemets pour signaler les emprunts à l'anglais non adaptés est relativement limité (exemples tirés de c94: *future*, *hacker*, *fiction*, *correct*, *lecture*, *kidware*). Dans *Le Monde* 94 (194) de nombreux emprunts non adaptés se trouvent accompagnés de gloses métalinguistiques, d'explications ou de propositions de traduction (par exemple *feed-back* rendu par *effet de retour*, et *re-engineering* accompagné du commentaire «pour les adeptes du

jargon anglo-saxon»). Il est bien sûr possible de relever le même phénomène pour l'espagnol (V94 = *La Vanguardia* 1994). Citons comme exemple *monitoring* accompagné de la glose entre parenthèses «(en espagnol, *controlar*)». En ce qui concerne l'attitude de la presse espagnole à l'égard des anglicismes, il est instructif de consulter le *Libro de estilo* de *El País* («Ne dites pas *mass media* mais *medios de masas* ou *medios informativos*»). Entre autres anglicismes non adaptés présentés entre guillemets dans V94, citons *apartheid*, *bit*, *boom* et *lobby*, termes déjà assimilés dans plusieurs langues. Dans le *TAZ* de Berlin on relève peu d'anglicismes non assimilés signalés par des guillemets; citons à ce propos les internationalismes *cool* et *impeachment*. Parmi les préfixes véhiculés par l'anglais, citons *ciber-*/*cyber-* (allemand *Kyber-*). A propos de signes particuliers, il convient de signaler aussi l'emploi des parenthèses accompagnant tel ou tel commentaire, qui peut être significatif dans la recherche de néologismes.

Comme on le comprend bien, partant des quelques mots proposés à des fins d'exemplification (emprunts à l'anglais non assimilés entre guillemets), il n'est pas possible d'extrapoler une tendance, d'autant plus que l'emploi de ce signe de ponctuation est tributaire des traditions des différents journaux, ainsi que de celles des langues en question, mais ils sont plus qu'une question de style. Des contextes significatifs de ce point de vue, tirés du sous-corpus italien de l'appendice 6, qui comporte des syntagmes de deux mots mis entre guillemets dans le *Sole 24 Ore* 1994, comportent des calques bien prévisibles: *autostrade elettroniche*, *qualità totale*, *realtà virtuale*, *casa cablata*, etc., ainsi que des anglicismes non adaptés et spécialisés (*on line*, *client server*, *information technology*, *public company*, *convenient store* [sic], *function point*, etc.).

9 Conclusions

Cet article se propose d'illustrer un phénomène jusqu'ici peu étudié par les sciences linguistiques, mais promis à un développement croissant. La néologie internationale et les nouveaux médias qui la promeuvent sont en réalité un nouveau chapitre dans les études portant sur deux aspects de notre vie accélérée: l'élan de l'innovation linguistique (et extralinguistique) et l'intensification des contacts entre les peuples et les différentes langues. Cette tendance, favorisée par les nouveaux médias, mais aussi par la diffusion des moyens de communications (télévision, radio, fax, etc.) et de moyens de transports rapides de plus en plus à la portée des masses, donne lieu à des changements conséquents. Il s'agit de changements difficilement mesurables à moyen terme, tant leur rapidité et leur complexité rendent difficiles toutes tentatives de prévision⁽¹²⁾ allant au-

(10) En forçant un peu le trait, on peut s'attendre à des cooccurrences comme *gatto* (chat) et *formaggio* (fromage) pour l'un (**topo* pour «mouse») et *computer* (*ordinateur*) et *tastiera* (*clavier*) pour l'autre (it. *Mouse*). «Un mot qui a plusieurs sens différents aura des colloquats différents. On peut ainsi dépister automatiquement le changement de sens.» Volz (1996: 163).

(11) Pour l'italien voir Petralli 1990 ainsi que Accame (1981).

(12) On rencontre des prévisions inconsidérées, comme le fait remarquer le directeur de Medialab à Lugano, Nicholas Negro Ponte, qui voit périodiquement des scénarios révolutionnaires. Nous pouvons citer celui d'une révolution numérique qui aurait fait disparaître rapidement le papier de nos bureaux.

delà de simples propos de Café du Commerce. Les journaux représentent de toute façon un miroir significatif de l'évolution de la société, en tant que vecteurs indispensables entre les connaissances toujours plus parcellisées et le désir de la part de groupes de citoyens de plus en plus nombreux d'accéder aux informations.

Tout en observant la réserve de rigueur dans ces cas, on peut prétendre que les thèmes choisis par les grands journaux, leur «agenda setting», la langue et les mots employés constitueront une référence qu'il conviendra d'observer si l'on souhaite pister les tendances émergentes de la société contemporaine, surtout celles qui constituent le grand marché mondial (États-Unis, Asie, Europe). Mais il est vrai aussi qu'il est plus probable que dans les journaux on relève des tendances «déjà émergées» et que les jeux de la néologie soient déjà exprimés dans les revues spécialisées. Où et quand est née la «globalisation»? Quelle est l'influence du «village global»⁽¹³⁾? Rappelons toutefois que les journaux sont des vecteurs privilégiés pour transmettre au grand public le vocabulaire de nombreuses sciences (et d'autres sous-codes en général).

C'est le cas de l'économie, qui dans certains cas fait son propre marketing, mais c'est aussi le cas des sciences plus dures qui ne dédaignent pas la scène médiatique. La mondialisation est un phénomène inédit et complexe. Elle est également ambiguë, une fourre-tout où l'on range des événements significatifs et triviaux. Il arrive que pour tel nouveau concept on ne trouve pas toujours les «mots pour le dire». Les réflexions de Fanfani (1991: 12) sont significatives de ce point de vue, quand il écrit que pour l'italien «à la fin des années 80 [...] la phase de l'influence de l'anglais tire à sa fin, influence qui a commencé à l'époque

de l'expansion économique («boom») et qui était caractérisée par une pression croissante et tumultueuse d'ordre lexical surtout», tandis que de l'autre il relève «qu'on aborde une période de rapports plus mûrs et plus circonspects avec l'anglais, qui laisse entrevoir une certaine stabilisation dans le domaine lexical mais une influence plus profonde au niveau des structures de la langue». Cette (apparente) contradiction devient subitement plus compliquée quand l'auteur relève «le déclin du soi-disant «mythe américain», il s'agit probablement de son européanisation.»

Dans la pratique, le niveau superficiel des langues serait moins influencé de nos jours tandis que la structure profonde serait déjà modifiée. En somme, dans la pire des hypothèses, les Européens seraient déjà devenus bien Américains et l'exception culturelle revendiquée par les Français pour l'audiovisuel serait un paravent pour les belles âmes⁽¹⁴⁾.

Nous croyons qu'une néologie qualitative, suivant de près l'évolution de concepts cardinaux comme celui de la mondialisation/globalisation peut être utile pour mieux comprendre la portée et les possibles conséquences de l'hégémonie culturelle américaine et pourrait être à l'origine d'une nouvelle pragmatique de la communication, et non seulement en réseau. On pense à la concentration entre peu de mains des grands marchés de la communication.

Pour ces raisons nous nous proposons de suivre de près les parcours narratifs de et sur la globalisation, que de nombreux journaux européens et américains proposent avec insistance quotidiennement à leurs lecteurs. Nous avons commencé à le faire pour la première moitié des années 90, et nous avons l'intention de poursuivre pour la seconde moitié. Un modèle d'analyse sémantique efficace, appliqué à un néologisme aussi

répandu (*globalisation/mondialisation*) et à ses dérivés, serait bien révélateur. Ces perspectives méritent amplement, à notre avis, une collaboration internationale.

*Alessio Petralli,
Istituto di Scienze Sociali della
comunicazione,
Facoltà di Scienze della comunicazione,
Università della Svizzera italiana,
Lugano.*

*Traduit par J. Humbley,
Centre de terminologie et de néologie,
Université Paris XIII.*

Bibliographie

Accame (F.), 1981: «Virgolette, che passione. Interpretazioni indiscipline del comunicare implicito», *Il Caffè*, XX, juin, p. 47-59.

Barnhart (D.), 1995: *The Barnhart New Words Concordance. Supplement*, Cold Spring, New York, Lexik House Publishers.

Braun, (P.), Schaeder (B.) et Volmert (J.) (éd.), 1990: *Internationalismen. Studien zur interlingualen Lexikologie und Lexikographie*, Tübingen, Niemeyer.

(13) *War and Peace in the Global Village*, avec Quentin Fiore, date de 1969, tandis qu'en 1960, dans *Explorations in Communication: An Anthology*, d'Edmund S. Carpenter, on relève déjà *global village*.

(14) En tout cas nous pensons que la culture (et la langue) française est aujourd'hui la seule capable de résister au «leadership USA»). Comme exemple de cette résistance et de cette tension, nous pouvons citer *Le Monde diplomatique*, d'une part, et des publications comme *Wired*, de l'autre.

Fanfani (M.), 1991-1995: «Sugli anglicismi nell'italiano contemporaneo», *Lingua Nostra*, 52 (1991), p. 11-24, p. 73-89, p. 113-18; 53 (1992), p. 18-25, p. 79-86, p. 120-21; 54 (1993), p. 13-20, p. 63-71, p. 122-24; 55 (1994), p. 19-25, p. 76-77, p. 117-120; 56 (1995), p. 14-17.

Maier-Meyer, 1995: *Lexikon und automatische Lemmatisierung*. Munich: Centrum für Informations- und Sprachverarbeitung der Universität München (CIS-Bericht 95-84).

Petralli (A.), 1990: *L'italiano in un cantone. Le parole dell'italiano regionale ticinese in prospettiva sociolinguistica*, Angeli, Milan.

Petralli (A.), 1996a: *Sul nuovo articolo linguistico della Costituzione svizzera. Straordinario storia di ordinaria democrazia elvetica*, Nuova Critica, Lugano, Canobbio.
(<http://www.tinet.ch/nuovacritica/communications>)

Petralli (A.), 1996b: *Neologismi e nuovi media. Verso la «globalizzazione multimediale» della comunicazione?*, CLUEB, Bologna.

Volz (N.) 1996: «CORDON – A Joint Venture Case Study», dans: Rettig (H.) (éd.): *Proceedings of the First European Seminar: Language Resources for Language Technology*, Tihany, Hongrie, 1995. Budapest: Hungarian Academy of Sciences, p. 159-167.

Annexes 1

Ace	Anteo	Banker
Acquarellate	Anti	Bankitalia
Acquariana	Anticapitalistici	Baraka
Acquariani	Anticonsumiste	Barbour
Acquariano	Anticonsumistici	Barcellona
Ade	Antidoping	Barek
Ademon	Antieducativa	Barfogenica
Admnistration	Antioccidentali	Baring
Advanced	Antipsicanalisi	Bass
Adzubej	Antiregionalista	Batch
Aereodinamiche	Antiriciclaggio	Bates
Aficionados	Antisocialismo	Battle
Afraid	Antitrust	Battletec
After	Antoni	Bayer
Agabenghian	Antroposofi	Bbs
Agadez	Appeal	Beatnik
Age	Aramesh	Beautiful
Ager	Arcais	Beckettiani
Agers	Architet	Behnke
Agis	Arcivestivano	Belgeri
Aharon	Arcuti	Ben
Aharonot	Arden	Bencheikh
Aids	Ardenzi	Benedini
Aires	Arens	Benissmo
Airlines	Arp	Bennis
Akira	Ars	Beppe
Alcun	Art	Berdini
Alekan	Artimm	Bergmaniana
Alferj	Arts	Bergognone
Algernon	Assen	Beria
Allan	Assets	Berlanda
Allen	Associamo	Berlino
Allenby	Associated	Berlusconiani
Allilujeva	Aston	Best
Altavilla	Astra	Biancorossoverde
Alternativo	At	Biasco
Altroieri	Athuin	Bigliettini
Amd	Atma	Bilità
Americàs	Attualizzazione	Bill
Ames	Auriol	Binswanger
Amiri	Autoalimentarsi	Biomeccanico
Anchorman	Autocandidature	Biomediche
And	Autocastrarsi	Biomedico
Anders	Autoflagellarci	Biomi
André	Avalon	Biondino
Angeles	Avenue	Birobidzan
Animal	Aviv	Bisognava
Animaletto	Bad	Bisognerebbe
Animatroni	*	Biturbo
A	Baek	Black
Annus	Baffetti	Blade

Blady	Butera	Centenaro	Coffeehouse	Cumparsita
Blasi	By	Center	Cofinanziamento	Cuociuova
Blob	Ca	Centocinquantamila	Cohen	Cup
Blum	Caccerà	Centosessant'	Coinvolgimento	Cupio
Board	Cadillac	Centoventotto	Collars	Cussione
Bobbiese	Caf	Cepollaro	Collegabili	Cutugno
Bobbio	Café	Cercorsele	Collegano	Cyber
Boca	Caggiano	Cercopitecantropo	Collegare	Cybereden
Bocchiano	*	Cespe	Collegarsi	Cybermaster
Bocconiani	Cagnoni	Cgil	Collegherà	Cyberpunk
Bohémien	Cairella	Chances	Coloberti	Cyberspace
Bondismo	Calciofila	Chandler	Colpaccio	Cyberspazio
Book	California	Change	Com	Cyborg
Boomers	Calise	Chao	Comby	
Booz	Calusca	Charbonnel	Comics	
Border	Camp	Charles	Committee	
Bosniaca	Canal	Chastel	Communications	
Bosniache	Cantanapoli	Chat	Compact	
Bosniaco	Canton	Chatwin	Company	
Bot	Cap	Chief	Compartimentalizza-	
Bottigliette	Capital	Chironomidi	zione	
Boulangier	Capitan	Chord	Completmente	
Bourget	Cappiello	Chronos	Computeristi	
Bowe	Caprioliano	Cianfanelli	Computers	
Brac	Caps	Cibalgine	Comunelli	
Braccesi	Caramel	Ciber	Comunicatore	
Bradbury	Carassa	Cibersesso	Comunicatori	
Braganti	Carbo	Cibona	Concettualmente	
Braggio	Card	Cicciolina	Conchetta	
Brainstorm	Cardinal	Cidim	Conclusosi	
Braitenberg	Carey	Cilea	Configurabile	
Breakfast	Carobbi	Cina	Confindustria	
Breeden	Carroarmato	Cinefila	Confuzione	
Brett	Carroll	Cinefili	Connery	
Bridgestone	Carrà	Cinemagique	Consigliere	
Brook	Casamadre	Cinematografico	Consoles	
Brooks	Cascella	Cinesesso	Contropiano	
Brosnan	Casetti	Cinesine	Coprodotto	
Brothers	Casini	Cinesini	Corporation	
Brown	Cassiopea	Cinevideo	Corrigan	
Brownsville	Castle	Cino	Cosanostra	
Brunetta	Casuismo	Cinquantasettesimo	Coss	
Btp	Cavaradossi	Cinquecentotrenta	Course	
Bubble	Cavaturacciolo	Cinqueper cento	Cowboy	
Bud	Cbs	Cinturoni	Cranshaw	
Budyonny	Cct	Cio	Crash	
Buenos	Cd	Cipro	Credazzi	
Building	Ceccinelli	Ciàcole	Credibilità	
Bulletin	Ced	Clisura	Crem	
Buon	Cedex	Coagulatrice	Crs	
Busby	Cee	Coast	Csm	
Buscemi	Cementir	*	Cto	
Bush	Cenghialta	Coelli	Cucurrucucù	

Annexes 2

Mots comportant l'élément POST-relevés dans le corpus ITA

1) post 8	(12)
2) postal	(2)
3) postatomico	(1)
4) postazione 1	(7)
5) postazioni 1	(8)
6) postbank	(1)
7) postbellico	(1)
8) postcomunismo	(3)
9) postcomunista	(4)
10) postcomunisti	(1)
11) postdemocristiani	(1)
12) postecologica 1	(1)
13) postecologico 1	(1)
14) postel	(3)
15) postelettorale	(2)
16) posterius	(1)
17) postfascisti	(1)
18) postfazione	(3)
19) postfuturistiche	(1)
20) postindustriale	(3)
21) postindustriali	(1)
22) postinformatica	(1)
23) postlethwaite 3	(1)
24) postludio	(1)
25) postman	(2)
26) postmoderna	(7)
27) postmoderni	(1)
28) postmodernism	(1)
29) postmodernismo	(1)
30) postmoderno 1	(7)
31) postnichilista	(1)
32) postnucleare	(1)
33) postolimpico	(1)
34) postpasquale	(1)
35) postprandiale	(1)
36) postproduzione	(1)
37) postquarantenni	(1)
38) postrel	(1)
39) postribolo	(2)
40) postrisorgimentale	(1)
41) postscript	(1)
42) poststalinista	(1)
43) postunitaria	(1)
44) posturografo	(1)
45) postvendita	(1)

- DBT

E. Picchi

Mots comportant l'élément BIO-relevés dans le corpus ITA

1) bio	(4)	50) biomeccaniche	(1)
2) biochimica	(3)	51) biomeccanici	(1)
3) biochimiche	(1)	52) biomeccanico 1	(1)
4) biochimici	(1)	53) biomedica	(3)
5) biochimico	(1)	54) biomedicale	(1)
6) bioclimici	(1)	55) biomediche 1	(3)
7) biocompatibile	(1)	56) biomedicina	(2)
8) biocompatibili	(1)	57) biomedico 1	(2)
9) biocompatibilità	(1)	58) biomi 1	(1)
10) biocontrol	(1)	59) biomolecolare	(1)
11) biodegradabile	(1)	60) biomouse	(1)
12) biodegradabilità	(1)	61) bionica	(2)
13) biodiesel	(1)	62) bionici	(2)
14) biodiversità	(1)	63) bionico 1	(2)
15) bioelettronica	(1)	64) bioparco	(1)
16) bioetica	(4)	65) biopark	(1)
17) bioetiche	(1)	66) bioprò	(1)
18) bioetici	(1)	67) bioritmi	(1)
19) biofisica	(1)	68) bioritmo	(1)
20) biofisici	(1)	69) biosegnali	(1)
21) biofisico	(1)	70) biosfera	(2)
22) biogenetica	(2)	71) biosistemi 1	(3)
23) biognava	(1)	72) biotecn	(1)
24) biografa	(1)	73) biotecn	(1)
25) biografi	(2)	74) biotecnologia	(6)
26) biografia 2	(7)	75) biotecnologica	(2)
27) biografica	(1)	76) biotecnologico	(1)
28) biograficamente	(1)	77) biotecnologie	(8)
29) biografiche	(4)	78) biotiche	(1)
30) biografici	(1)	79) biotridimensionale	(1)
31) biografico	(1)	80) biottico	(1)
32) biografie	(3)	81) biove	(1)
33) biografo	(3)	- DBT	
34) biohazard	(1)		E. Picchi
35) bioimpianti	(1)		
36) bioinformatica	(1)		
37) bioinformatico	(1)		
38) bioingegneri	(1)		
39) bioingegneria	(5)		
40) biolghini	(1)		
41) biologi	(3)		
42) biologia 5	(9)		
43) biologica 1	(10)		
44) biologiche 6	(5)		
45) biologici 2	(8)		
46) biologico	(7)		
47) biologismo	(1)		
48) biologo 2	(8)		
49) biomasse	(1)		

Tableau 1

Code	Source	Année	Nombre d'occurrences	Nombre de formes
C92	«Il Corriere della Sera»	1992	98322	17209
C93	«Il Corriere della Sera»	1993	191778	26102
C94	«Il Corriere della Sera»	1994	284642	33056
C95	«Il Corriere della Sera»	1995	220055	27391
RC2	«Il Corriere della Sera»	1992	189592	25917
RC3	«Il Corriere della Sera»	1993	309938	34980
RC4	«Il Corriere della Sera»	1994	398717	39550
RC5	«Il Corriere della Sera»	1995	526884	45465
RS2	«Il Sole - 24 ore»	1992	347792	29843
RS3	«Il Sole - 24 ore»	1993	365757	30781
RS4	«Il Sole - 24 ore»	1994	513201	37617
RS5	«Il Sole - 24 ore»	1995	720270	41365
S92	«Il Sole - 24 ore»	1992	319465	27986
S93	«Il Sole - 24 ore»	1993	197467	18778
S94	«Il Sole - 24 ore»	1994	398252	29220
S95	«Il Sole - 24 ore»	1995	916863	41578
RT2	«La Stampa»	1992	143122	23335
RT3	«La Stampa»	1993	201010	29249
RT4	«La Stampa»	1994	286538	36400
RT5	«La Stampa»	1995	351209	41790
T92	«La Stampa»	1992	157905	24815
T93	«La Stampa»	1993	136518	22774
T94	«La Stampa»	1994	193824	28197
T95	«La Stampa»	1995	343754	41323
V94	«La Vanguardia»	1994	163974	20596
V95	«La Vanguardia»	1995	184401	22813
L93	«Le Monde»	1993	243681	21947
L94	«Le Monde»	1994	189015	20219
L95	«Le Monde» (01.01.95 - 30.06.96)	1995	227470	24218
R94	«Le Soir»	1994	265407	27089
R95	«Le Soir»	1995	277696	27791
M95	«MCCarthy»	1995	275966	18809
Z93	«Neue Zürcher Zeitung»	1993	110171	22803
Z94	«Neue Zürcher Zeitung»	1994	158724	30761
Z95	«Neue Zürcher Zeitung»	1995	101667	21207
B92	«TAZ»	1992	111845	23293
B93	«TAZ»	1993	84858	18942
B94	«TAZ»	1994	78590	18420
B95	«TAZ» (01.01 - 30.06)	1995	54104	13906
I92	«The Independent»	1992	461179	32372
I93	«The Independent»	1993	577624	36339
I94	«The Independent»	1994	612223	38052
I95	«The Independent»	1995	935727	46392
RI2	«The Independent»	1992	462253	32400
RI3	«The Independent»	1993	578380	36387
RI4	«The Independent»	1994	612754	38069
RI5	«The Independent»	1995	937797	46453
N93	«The N.Y. Times»	1993	66853	11049
N94	«The N.Y. Times»	1994	133958	16334
N95	«The N.Y. Times»	1995	87225	11996
			Total occurrences 15806417	Total Formes 1433378

Tableau 2

Corpus	Code	Source	Année
ITA	C92	Il Corriere della Sera	1992
ITA	C93	Il Corriere della Sera	1993
ITA	C94	Il Corriere della Sera	1994
ITA	C95	Il Corriere della Sera	1995
ITA	S92	Il Sole 24 Ore	1992
ITA	S93	Il Sole 24 Ore	1993
ITA	S94	Il Sole 24 Ore	1994
ITA	S95	Il Sole 24 Ore	1995
ITA	T92	La Stampa	1992
ITA	T93	La Stampa	1993
ITA	T94	La Stampa	1994
ITA	T95	La Stampa	1995
Q94	C94	Il Corriere della Sera	1994
Q94	S94	Il Sole 24 Ore	1994
Q94	T94	La Stampa	1994
Q94	I94	The Independent	1994
Q94	N94	The N.Y. Times	1994
Q94	L94	Le Monde	1994
Q94	R94	Le Soir	1994
Q94	Z94	Neue Zürcher Zeitung	1994
Q94	B94	TAZ	1994
Q94	V94	La Vanguardia	1994

Nouvelles du Rint

Le Comité d'orientation du Réseau international de néologie et de terminologie a tenu sa vingtième réunion à l'Institut d'études et de recherches pour l'arabisation, à Rabat, en juillet 1999. Au cours de cette réunion, le Rint, en accord avec l'Agence intergouvernementale de la Francophonie, a choisi de privilégier en 1999 les actions d'information et de diffusion.

Cette orientation est significative, car non seulement est-il important de faire connaître les réalisations du réseau, mais c'est l'une des missions fondamentales du Rint de rendre accessible toute l'information sur la néologie et la terminologie. Voici donc quelles sont les principales actions du Rint en ce sens.

Terminologies nouvelles

La revue *Terminologies nouvelles*, dont le lectorat augmente régulièrement, est maintenant diffusée en texte intégral, en format *PDF*, dans le site Internet du Rint (www.rint.org), à compter du numéro 14. Cette nouvelle forme de diffusion permettra au Rint de rejoindre un public plus large, en fonction des différents thèmes traités dans la revue.

Les thématiques traitées dans *Terminologies nouvelles* sont de plus en plus étendues et diversifiées. Le Rint

souhaite occuper tout le champ de l'aménagement linguistique et celui de l'informatisation des langues et pour ce faire, il privilégie l'ouverture et le partenariat. Ainsi, le numéro 18 portait sur « Terminotique et documentation », le numéro 19 sur le thème « Terminologie et intelligence artificielle ». Le présent numéro porte en bonne partie sur l'informatisation des travaux de veille néologique et terminologique. Les autres aspects de l'aménagement linguistique ne sont pas pour autant négligés puisque le prochain numéro de *Terminologies nouvelles* abordera la question de la terminologie dans ses rapports avec la diversité culturelle.

Afin de rendre plus accessible le contenu de *Terminologies nouvelles*, le Rint diffuse sur son site Internet une base bibliographique dans laquelle sont répertoriés tous les articles parus dans la revue depuis le premier numéro. On peut interroger cette base par auteur, par titre et par terme clé.

La revue *Terminologies nouvelles* est réalisée sous la responsabilité du module de la Communauté française de Belgique – Service de la langue française.

Inventaire des ressources terminologiques

Dictionnaires terminologiques et banques de terminologie

La connaissance de ce qui existe en terminologie est précieuse pour tous les terminologues et pour les usagers de la terminologie. Le Rint entretient et diffuse depuis de nombreuses années un *Inventaire des ressources terminologiques* constitué de plusieurs éléments. Tout d'abord, il y a l'inventaire des travaux terminologiques publiés (imprimés et sur support électronique). Dans cette partie de l'inventaire, on retrouve essentiellement des dictionnaires terminologiques de tous domaines dont l'une des langues est le français. Il en va de même pour les banques de terminologie qui ont fait l'objet d'une enquête spécifique de la part du Rint. Ces inventaires sont accompagnés de formulaires électroniques qui permettent aux auteurs, éditeurs, diffuseurs, documentalistes ou usagers d'inscrire directement de nouveaux titres.

Ressources terminologiques sur Internet

En 1999, en réponse à une demande maintes fois formulée par les usagers de son site Internet, le Rint a décidé d'ajouter un nouveau volet à son *Inventaire des ressources terminologiques*. Il s'agit des

En bref

terminologies diffusées en ligne dans l'Internet. Afin de répertorier ces ressources terminologiques parfois éphémères, Une nouvelle page web vient d'être créée à l'adresse suivante: <http://www.olf.gouv.qc.ca/ressources/termino/ressling.html#inventaire>.

À partir de cette page, on pourra consulter la base de données, qui donne accès par hyperliens aux sites terminologiques répertoriés, ou l'enrichir en ajoutant de nouvelles données à l'aide d'un formulaire électronique sur lequel on peut indiquer les données suivantes: titre, auteur, lieu d'édition, éditeur/diffuseur, public cible, tarification, type de répertoire terminologique, catégories de données, langues, domaines, mots clés, adresse URL.

L'*Inventaire des ressources terminologiques* est réalisé sous la responsabilité du module québécois – Office de la langue française.

Inventaire terminotique

Conscient de l'importance grandissante que prend l'informatisation des travaux terminologiques, le Rint a créé il y a quelques années un *Inventaire terminotique*, de façon à répertorier et à faire connaître les différents types de logiciels utilisés en terminologie. En 1998-1999, le Rint a décidé d'élaborer un nouvel inventaire terminotique de façon à refléter la grande diversité des types de logiciels. Ce nouvel inventaire est décrit dans ce numéro.

L'*Inventaire terminotique* est réalisé sous la responsabilité du module canadien — Bureau de la Traduction.

Balnéo, outil de néologie interactive dans l'Internet

Le système *Balnéo* a pour objectifs la collecte, l'échange et la diffusion rapides de matériaux terminologiques touchant plus particulièrement les

néologismes, afin d'accélérer la diffusion des termes nouveaux et afin de rendre plus facile et plus efficace la mise à jour des dictionnaires terminologiques et les banques de terminologie. Conçu de manière à simplifier la saisie, la consultation, l'importation ou l'exportation des données, en environnement *Windows*, le système *Balnéo* est ouvert gratuitement à tout organisme ou à toute personne qui souhaite participer activement aux échanges terminologiques sur Internet.

Bibliographies

Aux bases de données déjà citées s'ajoutent deux bibliographies qui sont mises à jour de façon permanente: la *Bibliographie de la néologie* et la *Bibliographie de la terminotique*. Ces deux bases de données sont également diffusées dans le site Internet du Rint.

Le système *Balnéo* et les bibliographies sont placées sous la responsabilité du module français — Délégation générale à la langue française.

L'ensemble de ces répertoires et bases de données font du site Internet du Rint un véritable infoport terminologique indispensable à tous ceux qui s'intéressent à l'enseignement et à la pratique de la terminologie, de même qu'aux usagers de la terminologie, notamment les membres des professions langagières et les écrivains.

Voici l'adresse du site Internet du Rint: <http://www.rint.org>

Louis-Jean Rousseau,
Secrétaire général du Rint.

Projets de terminotique du Rint

Le module canadien du Rint dirige le groupe de travail en terminotique du Réseau. À ce titre, il gère divers projets de terminotique réalisés dans le cadre des biennums de

l'Agence intergouvernementale de la Francophonie. Au nombre de ces projets figurent les *Grilles descriptives de logiciels de terminotique à applications spécifiques* et le *Laboratoire terminotique virtuel*.

Grilles descriptives de logiciels de terminotique à applications spécifiques

Au cours du biennum 1998-1999, M^{me} Nelida Chan, professeur de terminologie au Collège Glendon de l'Université York, à Toronto et membre du Module canadien du Rint, a rédigé un rapport portant sur des grilles descriptives de logiciels de terminotique à applications spécifiques. Ces grilles servent à évaluer les logiciels de terminotique suivants:

- logiciels de gestion de données terminologiques;
- logiciels de saisie de données terminologiques;
- concordanciers;
- logiciels de dépouillement de la terminologie;
- logiciels de production de dictionnaires;
- gestionnaires de réseaux notionnels;
- générateurs de néologismes.

M^{me} Chan a d'ailleurs présenté succinctement ces grilles dans un article paru dans la chronique «En bref» du numéro 18 de *Terminologies nouvelles* (juin 1998).

Élaborées à partir de l'analyse d'une cinquantaine de logiciels de terminotique provenant d'Amérique, d'Asie et d'Europe, les grilles répertorient, outre une dizaine d'éléments de tronc commun, un large éventail de caractéristiques générales (une soixantaine) et spécifiques (environ 900) que comportent les logiciels de terminotique à applications spécifiques. En voici quelques exemples:

- critères de gestion des données terminologiques;

- type de concordances en langues source et cible;
- mode de production de dictionnaires;
- type de réseaux et de liens notionnels;
- types et modalités de génération.

Les grilles descriptives constitueront un outil d'évaluation précieux une fois que les concepteurs de logiciels les auront remplies dans le cadre du projet d'enrichissement de l'inventaire des produits en terminotique mené par le module canadien. En effet, les langagiers, perpétuellement en quête de nouvelles avenues technologiques pour accroître leur efficacité, disposeront désormais d'une ressource unique, accessible par Internet, dans laquelle seront rassemblées l'ensemble des caractéristiques des outils terminotiques. Ces renseignements les éclaireront dans le choix d'un outil qui réponde adéquatement à leurs besoins. Les grilles sont accessibles à l'adresse URL suivante:
www.uhb.fr/langues/balneo/GRILLE_ter/accueil.htm

Laboratoire terminotique virtuel

On a vu émerger, ces dernières années, une pléthore d'outils informatiques destinés à automatiser le travail du langagier. Dans ce contexte, le comité d'orientation du Rint a confié au groupe de travail en terminotique du Réseau le mandat de fournir une expertise en terminotique aux différents modules et de déterminer dans quelle mesure les outils existants peuvent optimiser le travail du terminologue.

Aussi, dans le cadre du biennium 2000-2001, le Module canadien du Rint a-t-il présenté à la réunion du comité d'orientation du Réseau, qui s'est tenue les 28 et 29 juillet 1999 à Rabat, au Maroc, le projet intitulé *Laboratoire terminotique virtuel*.

Il s'agirait de recenser les technologies prometteuses et d'en

évaluer l'utilité pour les langagiers. Les travaux consisteraient également à poursuivre les essais de logiciels de terminotique amorcés au cours du biennium 1998-1999, à raison d'une dizaine par année, dans un cadre plus structuré, celui d'un laboratoire terminotique virtuel.

Les résultats des travaux seraient mis à la disposition de la Francophonie par le truchement du site Internet du Rint. Ils pourraient également être présentés lors de diverses manifestations en Francophonie (Journée internationale de la Francophonie, Sommet francophone, etc.).

C'est avec grand enthousiasme que le module canadien entreprendra ce projet retenu dans la programmation du Rint pour le biennium 2000-2001.

*Michèle Valiquette,
 Terminologie et normalisation,
 Bureau de la traduction,
 Travaux publics et Services
 gouvernementaux Canada.*

Publications

Discours professionnels en français

Il peut paraître pervers de rassembler en Scandinavie et de publier en Allemagne un recueil d'études sur les langues de spécialité du français. Certes, le rédacteur de l'ouvrage et initiateur du projet, Yves Gambier, est professeur de français dans un institut de traduction en Finlande et l'éditeur, Peter Lang, est plus ouvert aux études de linguistique appliquée que la plupart des maisons d'édition françaises, mais la vraie raison de cette anomalie géographique se trouve ailleurs. D'une part, le regard que porte l'étranger sur la situation en France est garant d'objectivité et, d'autre part, ce sont les étrangers qui sont en passe de prendre le relais des Français, qui délaissent, à quelques brillantes exceptions près, ce domaine de recherche.

Cette crise des études des langues de spécialité fait l'objet d'une analyse perspicace signée du rédacteur lui-même, ici l'auteur du plus long article du recueil. Il commence par esquisser la situation des LSP à l'échelle internationale, où l'anglais domine, comme le sigle en usage même dans les écrits francophones le laisse entendre. Historiquement, c'est la nécessité d'enseigner rapidement la langue des études universitaires (donc l'anglais, le français, naguère le russe) aux étudiants allophones, qui a motivé les premières études des langues de spécialité. C'est aussi l'aspect des langues de spécialité qui accuse le plus grand retard en France aujourd'hui. Traditionnellement, c'était le français, langue étrangère (Fl), qui suscitait le plus grand

nombre d'études sur les langues de spécialité du français, tradition qui s'essouffle, sauf dans le nord de l'Europe. Gambier examine ensuite la langue de spécialité par rapport à des domaines connexes, qui représentent en général une application apparentée: la terminologie, la rédaction technique, la traduction. La méthode fait appel à des critères souvent flous: le domaine, la situation de communication, l'interaction, la vulgarisation, la sociolinguistique (langue de spécialité comme instrument de légitimation professionnelle), ces dernières étant pourtant susceptibles d'évolution, donc porteuses de renouveau. Ce tour d'horizon est une raison suffisante pour justifier l'utilité de ce recueil. Mais les autres articles, la démonstration pour ainsi dire, illustrent l'enrichissement possible des langues de spécialité, lorsque l'analyse se fonde sur une approche linguistique plus complète.

Savoir ce que constituent exactement les langues de spécialité tracasse les linguistes depuis longtemps. C'est «la question des questions», comme le dit Lothar Hoffmann, cité en exergue par Finn Frandsen, qui pour sa part, tente une définition qui se libère du carcan de l'opposition simpliste entre langue générale et langue de spécialité. Pour ce chercheur, il vaut mieux examiner les entités selon des perspectives différentes et complémentaires: celle de l'épistémologie d'une part et de l'ontologie de l'autre. L'une définit les formes d'usage d'une langue, l'autre les données elles-mêmes. La nouvelle définition des langues de spécialité qu'il propose est donc asymétrique et

si elle résout en le déplaçant le problème de la langue générale, elle comporte l'inconvénient de devoir constituer un nouveau métalangage, que l'auteur n'essaie pas d'improviser.

Après tant de critiques dirigées contre l'approche qui privilégie le niveau lexical, il n'est pas étonnant de constater que la majorité des articles accuse une orientation résolument textuelle. Privilégier le textuel, et parfois l'intertextuel, n'exclut pas une multiplicité de regards, qui se révèlent tout à fait complémentaires.

Une des lectures possibles des textes de spécialité est la sémantique interprétative de François Rastier. W. Johanssen analyse une brochure présentant une entreprise à la lumière des isotopies sémantiques, dégagées dans le texte, effet de la récurrence syntagmatique d'un même sème. Cette analyse fait ressortir non seulement la cohésion du texte, dont le point de départ est différent de celui de Hassan et Halliday, par exemple, mais aussi un aperçu de la culture d'entreprise.

André Avias pour sa part fait appel à une approche prototypique inspirée de J. M. Adam pour analyser un article du *Figaro* sur le budget. Il détermine d'abord, non sans difficulté, le découpage structurel de son article en séquences textuelles, puis il choisit des prototypes. Bien entendu, cette démarche ne se limite pas aux textes spécialisés, mais elle apporte une contribution au traçage de l'argumentation dans le discours.

Le caractère argumentatif, voire polémique d'un autre genre de texte, le «Mot du PDG», sorte de préface du rapport d'activité annuel d'une entreprise, fournit le thème de

l'article de K. Flottum. L'analyse de ce genre textuel est d'ailleurs une spécialité scandinave, et l'auteur a l'avantage de pouvoir puiser dans un fonds de recherche établi. Elle fonde son analyse sur les études qu'a réalisées Ducrot sur la négation syntaxique, sur la qualité de polyphonie, surtout sur la lecture qu'en fait H. Nölke. Une analyse des différents types de négation fait ressortir l'usage dominant de l'un d'entre eux, la négation polémique, déni des «idées fausses qu'on peut avoir sur l'entreprise». Cette analyse, surtout syntaxique rappelons-le, permet non seulement de résoudre des questions posées dans les études antérieures, mais aussi de déterminer qu'il s'agit d'un type de texte argumentatif et contre-argumentatif. L'auteur apporte la preuve par ailleurs que ce discours n'est pas impersonnel, comme une lecture superficielle pourrait le laisser entendre, car les destinataires sont bien présents, mais implicites.

Ce sont les sciences cognitives qui viennent enrichir l'analyse d'un texte juridique français par L. Lundquist, dans le cas présent un jugement de la Cour de cassation. Les outils que l'auteur emploie sont empruntés à l'intelligence artificielle dans le cas du cadre (*frame*), et aux études psychologiques, dans celui de l'espace mental. Celui-là sert à situer l'analyse de l'ensemble, celui-ci structure les relations inférieures. L'auteur expose ensuite le cadre du jugement, permettant au néophyte non francophone de situer les relations structurelles, tâche malaisée surtout en comparaison de la transparence structurelle du discours juridique danois déjà familier. L'idée d'espace mental permet à son tour de bien repérer les éléments et les relations au fil du discours. Cet article est un bon exemple de perspicacité obtenue par un regard extérieur porté sur le discours juridique français, regard qui est loin d'être naïf.

A. Askelund s'inspire de la grammaire des cas pour son analyse des textes juridiques, mais c'est la traduction qui l'intéresse. Elle explique comment l'emprise du texte de départ est différente selon qu'il s'agisse de la langue maternelle ou de la langue étrangère, et on apprécie son cadre d'analyse originale et explicite de la «métamorphose» du texte traduit.

Les deux derniers articles du recueil sont à orientation lexicale. Celui de G. Dyrberg et J. Tournay renoue avec la très riche tradition danoise des dictionnaires de spécialité dont le contenu et la présentation sont constamment modifiés en fonction des besoins perçus des usagers. Le dictionnaire envisagé dans cet article est juridique d'après son contenu et asymétrique selon sa présentation. Destiné aux usagers danois, il envisage un savant mélange d'exemples et d'explications, tantôt en danois, tantôt en français. Ce sont les informations encyclopédiques qu'il convient d'inclure qui constituent le sujet de l'article, mais on passe en revue la formation de la définition, les stratégies d'exemplification de l'équivalence partielle, particulièrement bien illustrées d'ailleurs, la typologie des exemples, et la place dans le dictionnaire de ce qu'elles appellent d'après A. Kjaer, les «formules de routine». Ces dernières auraient mérité un article de plus – en quoi consistent-elles en fait ? Faut-il les présenter sous un élément de la suite, ou de façon conceptuelle, quel traitement informatique proposer pour résoudre les problèmes sans issue sur le papier.

Le dernier article du recueil, signé de P. Lederlin, traite des constructions de type N de N dans les textes économiques. L'auteur cherche des règles de production simples et maniables à proposer à ses étudiants norvégiens. Pour son cadre théorique, d'une part, il se limite aux études relativement anciennes (M. Wilmet 1986) ; d'autre part il

sous-estime l'explication de la lexie complexe, pourtant évoquée. Selon cette approche, l'étudiant apprend *appel d'offres, assiette de l'impôt* comme des unités lexicales (ou terminologiques) et non comme des locutions.

Le recueil comporte une bibliographie de 71 études portant sur les langues de spécialité françaises publiées récemment en Scandinavie.

Après avoir pris connaissance de plusieurs nouvelles approches présentées ici, on ne peut partager le pessimisme de Gambier. Certes, la plupart des auteurs peinent à définir les langues de spécialité, mais en les décrivant ils contribuent à fournir une définition par extension, plutôt que par compréhension, comme diraient les lexicographes, mais qui est en même temps plus intuitive, plus pratique, et plus consensuelle.

*Une lecture de John Humbley,
Centre de terminologie et de néologie,
Laboratoire de linguistique informatique,
Université Paris 13.*

Gambier (Yves), éd., 1998 : *Discours professionnels en français*. Francfort, Peter Lang, 233 p.

Terminologie maritime

Le parler des gens de mer est lié à l'histoire de notre civilisation et porte la trace des échanges culturels, techniques et commerciaux entre les peuples. À ce titre, il a de longue date intéressé les lexicographes, ce qui explique la riche histoire des dictionnaires de marine. Outils essentiels pour enseigner, traduire et communiquer, ces dictionnaires de marine connaissent aujourd'hui une importante métamorphose dans le cadre des industries de la langue.

La première conférence internationale *Terminologie maritime: traduire et communiquer* avait pour principal objet de réunir les spécialistes du langage des gens de

mer ou du domaine nautique concernés par l'usage, l'étude ou l'enseignement du vocabulaire maritime. Elle s'est tenue à l'Institut supérieur de traducteurs et interprètes (Isti, Bruxelles) les 15 et 16 mai 1998. Le succès fut tel qu'il a été décidé d'organiser la conférence tous les deux ans. La prochaine rencontre se déroulera à l'Université de Turku, en Finlande, en mai 2000.

Les actes du colloques, parus aux Éditions du Hazard, rassemblent les communications des différents orateurs, regroupées en six thématiques différentes :

- Histoire de la terminologie maritime ;
- Questions de normalisation ;
- Informatique et terminologie maritime ;
- Enseignement et langue de spécialité ;
- Contacts interculturels et communication ;
- Traduction spécialisée et lexicographie.

Toujours accessibles, ces textes – rédigés en français et en anglais – offrent une diversité de points de vue sur le vocabulaire de la marine en cette fin de millénaire.

Renseignements :

Centre de recherche Termisti,
34 rue Joseph Hazard,
B-1180 Bruxelles,
Belgique,
termisti@euronet.be,
tél. : +32.2.346.26.41,
téléc. : +32.2.346.21.34.

Newman (D.L.) et Van Campenhoudt (M.), éd., 1999 : *Terminologie maritime: traduire et communiquer. Actes du 1^{er} colloque international de terminologie maritime (Bruxelles, 15 et 16 mai 1998)*, Bruxelles, Éditions du Hazard.

Le sens en terminologie

Ce volume collectif rassemble onze articles consacrés aux problèmes du sens en terminologie. On y trouvera des contributions de Maria Teresa Cabré, Juan C. Sager, Monique Slodzian, Loïc Depecker, Marc Van Campenhoudt, François Gaudin, Bruno de Bessé, Ingrid Meyer et Kristen Mackintosh, Yves Gentilhomme, Claude Boisson, et de Henri Béjoint et Philippe Thoiron. Sont abordées les questions des rapports entre le terme et le mot, de la manière de définir le sens des termes (relations entre concept, signifié et référent, importance du domaine), des frontières entre langues spécialisées et langue générale. L'ensemble des contributions fait apparaître l'émergence d'un nouveau type de recherche terminologique fondée sur l'exploitation des corpus, éventuellement multilingues. Cette nouvelle terminologie se démarque de la tradition par ses méthodes et ses objectifs. Elle tend à devenir descriptive plutôt que normative et s'inscrit désormais dans le cadre d'une approche ouvertement linguistique.

Presses universitaires de Lyon
80, boulevard de la Croix-Rousse
BP 4371
F-69242 Lyon Cedex 04
France

Thoiron (Philippe) et Béjoint (Henri), dir., 1999 : *Le sens en terminologie*, Lyon, Presses universitaires de Lyon.

Je soussigné souhaite recevoir gratuitement la revue *Terminologies nouvelles*.

Nom : _____

Entreprise, organisme : _____

Fonction : _____

Adresse : _____

Ce bulletin d'abonnement est à adresser au module dont vous relevez (adresse au dos de la revue)

Descriptif bibliographique :
Depecker (Loïc)
et Rousseau (Louis-Jean), dir. :
Nouveaux outils pour la néologie,
dans *Terminologies nouvelles*,
n° 20, décembre 1999, Bruxelles,
Agence de la francophonie
et Communauté française de Belgique,
ISSN : 1015-5716.

ISSN : 1015-5716

© Tous droits de traduction
de reproduction et d'adaptation
réservés pour tous pays.

Édit. resp. :
M. Garsou, 44 boulevard Léopold II,
1080 Bruxelles, Belgique.

Numéros déjà parus

Consultables à partir du n° 14 à l'adresse www.rint.org

1, mai 1989:

Le Rint: objectifs et perspectives

2, décembre 1989:

La formation en terminologie

3, juin 1990:

Harmonisation des méthodes en terminologie (actes des séminaires de Talence et de Hull)

4, décembre 1990:

Numéro général

5, juin 1991:

Terminologie et informatique

6, décembre 1991:

Terminologie et développement I (actes du séminaire de Rabat)

7, juin 1992:

Numéro général

8, décembre 1992:

Terminologie et environnement

9, juin 1993:

Terminologie et développement II (actes du séminaire de Cotonou)

10, décembre 1993:

Phraséologie (actes du séminaire de Hull)

11, juin 1994:

Numéro général

12, décembre 1994:

Implantation des termes officiels (actes du séminaire de Rouen)

13, juin 1995:

Terminologie et entreprise

14, décembre 1995:

Numéro général

15, décembre 1996:

Banques de terminologie (actes de la table ronde de Québec)

16, juin 1997:

Enquêtes terminologiques

17, décembre 1997:

Terminologie et formation

18, juin 1998:

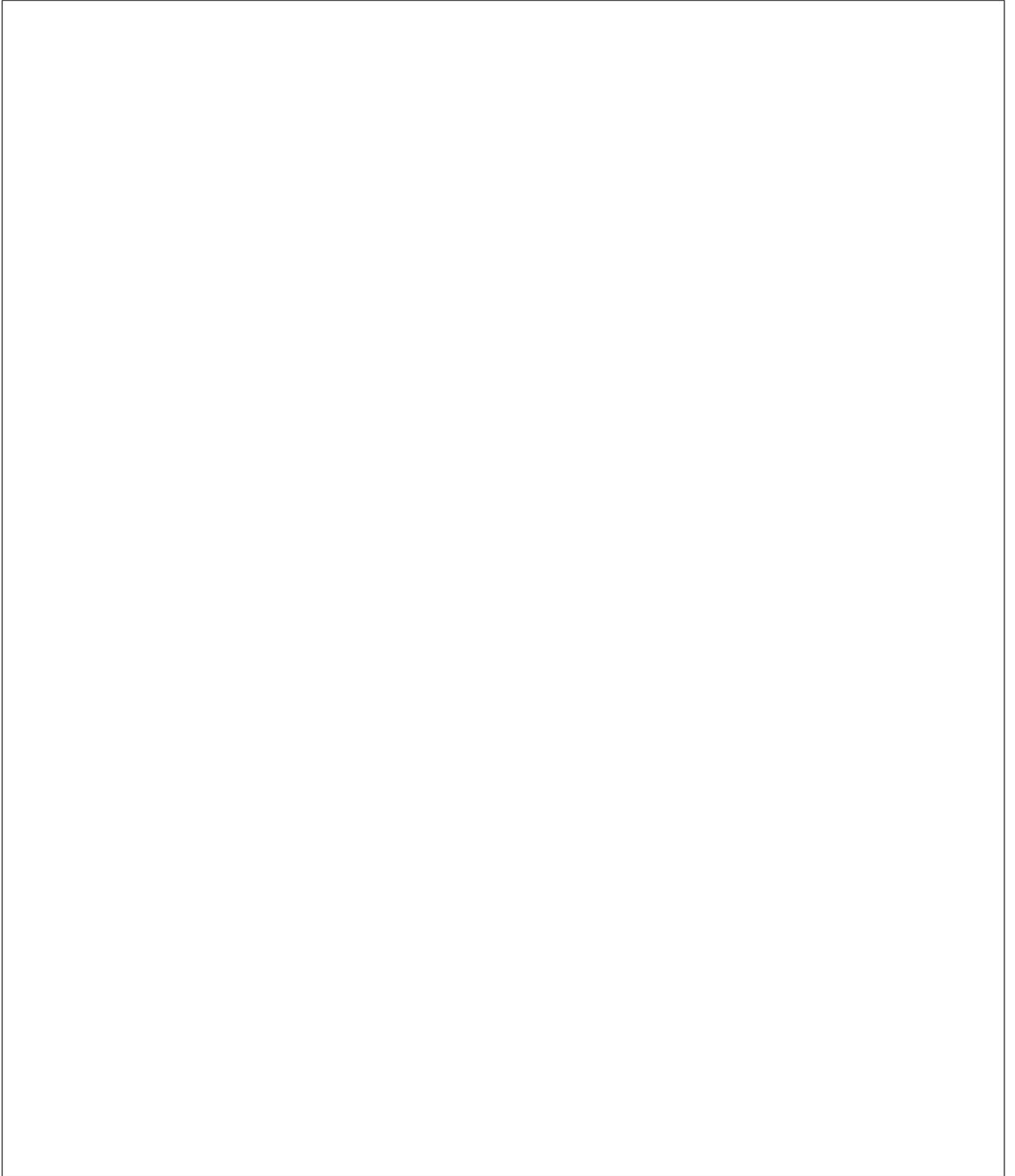
Terminotique et documentation

19, décembre 1998 - juin 1999:

Terminologie et intelligence artificielle (actes du colloque de Nantes)

20, décembre 1999:

De nouveaux outils pour la néologie



Adresses des organismes membres du Rint

Afrique centrale et de l'Est

Coordination: Centre de linguistique
théorique et appliquée
BP 4956
Kinshasa/Gombe
Zaire.

Afrique de l'Ouest

Coordination: Centre de
linguistique appliquée
Université Cheikh Anta Diop
Dakar — Fann
Sénégal.

Canada

Terminologie et Normalisation
Bureau de la traduction
Travaux publics et
Services gouvernementaux
Portage II, 3^e étage
165, rue Hôtel-de-Ville
Hull (Québec)
K1A 0S5
tél.: 1 (819) 994-5934

Communauté française de Belgique

Ministère de la
Communauté française
Service de la langue française
44, Bd Léopold II
B-1080 Bruxelles
tél.: 32 (2) 413 32 74

France

Délégation générale à la langue
française
1, rue de la Manutention
F-75116 Paris
tél.: 33 (1) 40 69 12 00

Haïti

Faculté de linguistique
Université d'État d'Haïti
38, Rue Dufort
(Quartier Bois-Verna)
Port-au-Prince
tél.: (509) 45 12 33

Madagascar

Centre des langues de l'Académie
malgache
BP 6217
Antananarivo 101.

Maroc

Institut d'études et de recherches
pour l'arabisation
BP 6216
Rabat — Instituts
tél.: 212 (7) 77 30 12

Québec

Office de la langue française
200, chemin Sainte-Foy,
Québec (Québec)
G1R 5S4
tél.: 1 (418) 643-4144

République centrafricaine

Conseil national d'aménagement
linguistique
BP 888
Bangui.

Suisse

Chancellerie fédérale suisse
Services linguistiques centraux
Section de terminologie
Gurtengasse 2-4, 4^e étage
CH 3003 Berne
tél.: 41 (31) 324 11 49

Tunisie

Innorpi
10bis, rue Ibn el Jazzar
1012 Tunis — Belvédère
tél.: 216 (1) 785 922

Modules associés

Union latine
Bureau de Paris
131, rue du Bac
F-75007 Paris
tél.: 33 (1) 45 49 60 60



Coédité par:
L'Agence de la francophonie
et la Communauté française de Belgique
(Service de la langue française
du ministère de la Communauté française
et Commissariat général
aux relations internationales)

Secrétariat du Rint:
Office de la langue française
200, chemin Sainte-Foy,
Québec (Québec)
G1R 5S4 Canada

Le Rint sur Internet:
<http://www.rint.org>