

Les voies actuelles de recherche en terminologie et terminotique

Communication présentée dans le cadre de la 7^e *Université d'automne en terminologie*,
Université de Rennes II, septembre 2000

Préambule

Ce texte de vulgarisation ne vise qu'à brosser à gros traits quelques aspects de la recherche actuelle en terminologie. Destiné aux participants de l'Université d'automne en terminologie, il propose un certain nombre de liens vers les sites Internet des organismes cités et des projets de recherche brièvement décrits. Un plus grand nombre de liens pourra être trouvé dans le cadre de l'*Infoport de la terminologie*, consultable sur le site Internet du centre de recherche TERMISTI¹.

Introduction

Le fait d'être un observateur extérieur et d'avoir le temps de la réflexion place l'universitaire dans une position privilégiée pour envisager l'avenir. À l'heure actuelle, les recherches explorent des voies assez diversifiées, mais toujours liées au développement du génie logiciel. Dans le même temps se creuse un gouffre de plus en plus profond entre les progrès de l'ingénierie de la langue et la pratique professionnelle du traducteur indépendant, qui rembourse son ordinateur par mensualités, hésite à acquérir de coûteux cédéroms, s'est récemment connecté à Internet et se demande ce qu'est S.G.M.L. Un meilleur dialogue entre les chercheurs, les développeurs de programmes, les traducteurs et les consommateurs de traduction améliorerait assurément la situation.

1 Historique

Avant de voyager dans le futur, il semble à tout le moins utile de revenir sur le passé récent de la terminologie.

1.1 Eugen Wüster

Il est de bonne coutume d'attribuer, sinon « l'invention » de la terminologie, du moins sa théorisation à Eugen Wüster. Cet ingénieur germanophone, spécialiste de la machine-outil et auteur d'une thèse sur la gestion des vocabulaires spécialisés, a produit bon nombre de textes concernant la façon de bien concevoir des glossaires multilingues. Très influent dans le cadre du comité technique 37 de l'Organisation internationale de normalisation (ISO), en charge de la terminologie, il en a profondément marqué les nombreuses normes² : jusqu'à la fin des années 80, celles-ci ont clairement été marquées par ses théories.

L'approche d'Eugen Wüster est indéniablement celle d'un ingénieur qui voit dans la normalisation une solution aux problèmes de communication entre les langues. Son approche conceptuelle, éloignée du modèle saussurien, a été très longtemps boudée par les linguistes, beaucoup plus intéressés par une approche descriptive des faits de langue.

1. www.refer.org/termisti/infoport.htm

2. www.iso.ch/catf/01020.html

Cependant, il serait par trop caricatural de réduire les travaux de Wüster à cette approche normative. Son célèbre *Dictionnaire de la machine-outil*³ offre un modèle de dictionnaire de traduction spécialisé, dont l'approche est, paradoxalement, beaucoup plus descriptive que normative. Si on prend la peine de les lire, on se doit de reconnaître que les travaux⁴ de Wüster ont au moins eu le mérite de remettre en cause une pratique du traducteur qui consiste à engranger, sans grandes vérifications, de simples listes d'équivalents disposés sur deux colonnes.

1.2 La terminologie, instrument d'aménagement linguistique

Depuis la fin des années 60, époque où les travaux de Wüster ont commencé à être davantage connus, la terminologie est aussi un instrument de politique linguistique dans un certain nombre de pays. Au Québec ou en Catalogne, la production et la diffusion d'équivalents aux mots anglais ou castillans est un enjeu culturel fondamental. Au Canada, en Suisse ou dans l'Union européenne, l'activité terminologique est la garantie d'un État (ou d'une entité supra-étatique) respectueux du multilinguisme.

Les bases de données terminologiques produites dans ce contexte légal ont généralement été constituées par accumulation. Ainsi, *Eurodicautom*, une base d'autant plus incontournable qu'elle est consultable gratuitement sur Internet⁵, a surtout été conçue - dans un premier temps - par le déversement de données issues de dictionnaires déjà parus⁶. Dans ce cadre, l'essentiel du travail terminologique a longtemps consisté à d'éviter les contradictions, effacer les doublons et ajouter de nouvelles langues.

Ces grands chantiers terminologiques ont vu le jour avec les premiers développements de l'informatique au sein de l'entreprise et de l'administration. Leur modèle de données demeure, aujourd'hui encore, profondément marqué par les systèmes documentaires qui ont présidé à leur création. Ceci handicape lourdement leurs possibilités d'évolution.

1.3 La révolution de la micro-informatique

Lorsque la micro-informatique connaît son véritable essor, vers la moitié des années 80, les bases de données offrent à l'honnête traducteur la possibilité de consulter rapidement les fiches terminologiques qu'il a - dans le meilleur des cas - stockées dans de nombreuses boîtes à chaussures au fil des années. En effet, à chaque fois qu'il rencontrait un problème de traduction, il faisait une fiche avec le terme en langue source, le terme en langue cible ; s'il se montrait méticuleux, il veillait à classer chaque fiche en fonction du domaine : la fraise du cultivateur (*strawberry*) n'est pas celle du dentiste (*drill*), ni celle du veau (*calf's caul*), ni celle de Henri IV (*ruff*).

3. Wüster (E.), *Dictionnaire multilingue de la machine-outil. Notions fondamentales, définies et illustrées, présentées dans l'ordre systématique et l'ordre alphabétique. Volume de base anglais-français*, London, 1968, Technical Press.

4. On trouvera une bibliographie des traductions françaises connues, à l'adresse www.refer.org/termisti/theoweb9.htm#W

5. www2.echo.lu/eurod/

6. Au sujet de l'évolution d'*Eurodicautom*, on lira l'excellent article de Goffin (R.), « EURODICAUTOM, la banque de données terminologiques multilingues de la Commission européennes, 1973-1997 », dans *Terminologie et traduction*, Luxembourg, Office des publications officielles des Communautés européennes, vol. 2.1997, p. 29-73

L'avènement de l'ordinateur personnel et la facilité offerte par le traitement de texte n'ont toutefois pas permis de libérer du temps pour saisir toutes ces fiches ou surtout pour se demander si la structure adoptée sur le papier permettait de gérer correctement l'équivalence à l'écran. Après avoir dépensé beaucoup d'argent pour acheter un excellent logiciel comme *Multiterm*, bien des traducteurs n'ont pas trouvé le temps d'y engranger une information valide. Dans le pire des cas, ils ont même gaspillé beaucoup d'énergie à y encoder des données selon une structure qui en rendait la consultation très aléatoire...

Une décennie à peine s'est écoulée depuis l'époque où *WordPerfect 4*, *Word 3* et *DBase* tournaient sur des 8086 dotés de 512 ko de Ram. Aujourd'hui, la puissance des P.C. a considérablement augmenté et permet d'envisager de réaliser chez soi des choses qui étaient jusqu'alors réservées aux gros systèmes. Force est de reconnaître que la pratique terminologique n'a pas connu un semblable progrès, alors même que les recherches en linguistique appliquée ont proposé de nouveaux modèles de gestion immédiatement utilisables.

Dès le début des années 90, beaucoup de chercheurs ont compris – ou plutôt cru comprendre – que l'intérêt des États et des institutions internationales pour la gestion informatique des terminologies constituait une source intéressante de financement. Dans les facultés, beaucoup de linguistes, de toutes les spécialités, se sont intéressés à la terminologie. Si certains se sont contentés de « surfer » sur les budgets, d'autres, au contraire, ont heureusement mis un point d'honneur à faire progresser la discipline, au risque de creuser davantage l'écart entre la situation du traducteur professionnel et les possibilités offertes par l'ingénierie linguistique.

2 Les grands chantiers

Il semble difficile de faire un large tour d'horizon en peu de mots. Nous nous cantonnerons donc à évoquer les projets récemment développés dans le cadre de la francophonie et de l'Union européenne.

2.1 Francophonie : le Réseau international de néologie et de terminologie

Le RINT, ou Réseau international de néologie et de terminologie, regroupe un grand nombre de pays francophones, représentés aussi bien par des organismes en charge de la politique linguistique que par des universités. Outre une revue, intitulée *Terminologies nouvelles*, le RINT tient à jour plusieurs inventaires utiles pour le traducteur et le terminologue. Tant l'inventaire des travaux de terminologie que celui des outils terminologiques sont consultables au départ du site Internet du réseau⁷.

Le RINT finance également des développements de nature informatique. Il a ainsi contribué au développement d'*Adepto-Nomino*, un excellent logiciel capable d'extraire des « candidats-termes » dans un corpus textuel rédigé en français et d'immédiatement transférer dans une base de fiches d'attestation les contextes jugés pertinents. Aussi étrange que cela puisse paraître, un tel logiciel ne semble pas disponible pour d'autres langues que le français, notamment l'anglais. Il convient, en outre, de mentionner que ce logiciel est aujourd'hui intégré dans un véritable extracteur de connaissances : *Nomino v2*, développé par l'Université du Québec à Montréal⁸ (cf. 2.2.1 et 3.2).

7. www.rint.org

8. www.ling.uqam.ca/nomino

Depuis plusieurs années, le RINT tente également de construire une base d'attestation des néologismes qui puisse être enrichie et consultée sur Internet. Ce projet est nommé *Balnéo*, comme « boîte aux lettres néologique »⁹. Son objectif est de collationner sur un site commun les néologismes de la langue française à mesure qu'ils apparaissent. L'idée est séduisante, car il semble, en soi, très utile pour un traducteur de pouvoir vérifier s'il peut traduire tel nouveau concept déjà désigné en anglais, en allemand ou en russe par tel mot qu'il invente en français ou qu'il croit avoir déjà rencontré dans l'usage. Aujourd'hui que le serveur est opérationnel¹⁰, il convient toutefois de s'interroger sur la volonté véritable des organismes membres du RINT - parfois chargés d'une « tutelle » sur la langue - d'alimenter un projet qui montre la langue française dans toute sa diversité et sa vitalité.

2.2 Les programmes européens (D.G. XIII)

L'idée d'une Europe politique constitue un défi linguistique majeur, puisque l'Union européenne entend respecter et promouvoir la diversité des langues. Il n'est pas possible de retracer ici l'historique de l'ensemble des programmes européens consacrés à cette problématique. Nous nous focaliserons plutôt sur deux récents programmes de la direction générale XIII, en charge l'ingénierie de la langue, qui accordent une attention particulière au problème des vocabulaires spécialisés.

2.2.1 TELEMATICS APPLICATIONS

Le premier de ces deux vastes programmes s'appelle *Telematics Applications*¹¹. Il concerne les applications télématiques pour l'administration, le transport, les soins de santé, l'environnement et l'ingénierie linguistique. Les projets qui portent plus spécifiquement sur la terminologie et le lexique au sein de ce programme sont en voie d'achèvement. Ils ont pour noms INTERVAL et EUROWORDNET.

Le projet INTERVAL répond à un appel d'offres portant sur la validation des données terminologiques. En effet, un projet préliminaire, du nom de POINTER, a montré que les bases de données terminologiques posent des problèmes permanents de nettoyage, de correction, d'enrichissement... Leur nature même montre qu'il serait utile de mettre au point des méthodologies pratiques de validation qui puissent notamment être secondée par l'outil informatique. Le site d'INTERVAL¹² présente déjà les premiers résultats de cette recherche.

Le projet EUROWORDNET¹³ ne concerne pas la terminologie à proprement parler. À son origine, on trouve *Wordnet*, un dictionnaire électronique anglais conçu sous la forme d'un réseau conceptuel. *Wordnet* a été élaboré dans la continuité des travaux de Quillian¹⁴ par l'équipe du professeur Miller de la Princeton University, un des maîtres à penser de la science cognitive et de la

9. Cet acronyme a été conçu au début d'Internet, lorsqu'on pensait encore devoir utiliser la messagerie électronique pour alimenter la base, qui est gérée sur un serveur de l'Université de Rennes II.

10. www.uhb.fr/Langues/Craie/balneo/balneo.pl

11. www2.echo.lu/langeng/fr/index.html

12. www.mcs.surrey.ac.uk/Research/CS/AI/new_interval/

13. www.let.uva.nl/~ewn/

14. Quillian (M.R.), 1967 : « Word Concepts : a Theory and Simulation of some Basic Semantic Capabilities », dans *Behavioral Science*, vol. 12, n° 5, p. 410-443.

psycholinguistique. Dans ce dictionnaire monolingue, consultable sur Internet¹⁵, on ne trouve pas tant des définitions que des liens de sens entre les concepts (essentiellement les relations « est un type de » et « est une partie de »).

EUROWORDNET a les mêmes ambitions, mais dans un cadre multilingue, puisqu'il s'agit de transformer *Wordnet* en un dictionnaire trilingue (anglais-espagnol-néerlandais). S'il est relativement aisé de bâtir un réseau conceptuel dans une langue, il est beaucoup moins évident de le faire dans plusieurs. Toute personne qui a suivi un cours de linguistique générale a appris que d'une langue à l'autre, on ne subdivise pas la réalité de la même manière. Ce problème est d'ailleurs fondamental dans le cadre de la traduction en général et de la conception de bases de données terminologiques multilingues en particulier. La solution de Wüster, et de l'École de Vienne à sa suite, est de tenter de normaliser les concepts à l'échelle internationale. Il sera donc très intéressant de vérifier si l'extension de *Wordnet* à deux autres langues a été résolue ou non par le biais d'une ontologie.

Aujourd'hui, les travaux de divers centres de recherche sont orientés vers la constitution de ce que l'on nomme des « bases de connaissances terminologiques ». Il s'agit d'arriver à un outil très puissant qui permette d'extraire des connaissances d'un texte et de produire une information terminologique solide. Cette approche résolument conceptuelle est également fondée sur l'exploitation de liens entre les concepts. Si l'on arrive aujourd'hui à programmer des outils capables d'extraire des concepts et des liens conceptuels au départ d'un texte (voir notamment le site de *Nomino v2* à l'Université du Québec à Montréal¹⁶ et celui du projet *Ikarus* à l'Université d'Ottawa¹⁷), l'étape suivante sera de vérifier comment des informations conceptuelles extraites de corpus textuels dans des langues différentes peuvent être exploitées pour établir des équivalences (cf. 3.2).

2.2.2 MULTILINGUAL INFORMATION SOCIETY

Le programme MLIS¹⁸ (*Multilingual Information Society*) a pour principaux objectifs d'encourager la construction d'une infrastructure européenne en matière de ressources langagières, d'encourager le développement des industries de la langue et de promouvoir l'utilisation de l'ingénierie linguistique dans le secteur public.

2.2.2.1 Commercialisation de données

La terminologie occupe une place importante dans le cadre de ce programme. Un premier appel à propositions a débouché sur une série de projets visant à tenter de commercialiser des données terminologiques ou lexicales sur Internet. La plupart des consortiums ont commencé leurs travaux au cours du premier semestre de l'année 1998.

Les responsables de la plupart de ces projets ont choisi de diffuser des données existant déjà dans des dictionnaires imprimés sur papier et se sont retrouvés confrontés à la nécessité de les convertir de manière à ce qu'on puisse les exploiter à l'aide d'une interface d'interrogation. En effet, la très grande majorité des dictionnaires disponibles sur le marché sont encore rédigés en fonction

15. www.cogsci.princeton.edu/~wn/

16. Cf. note 8.

17. www.csi.uottawa.ca/~kavanagh/Ikarus/IkarusInfo.html

18. www2.echo.lu/mlis/

d'un mode de présentation hérité de la tradition lexicographique et destiné à faciliter la consultation sur un support écrit. Force est de constater qu'en cette fin de XX^e siècle, de nombreux auteurs de dictionnaires de traduction spécialisés organisent la macrostructure et la microstructure de leurs ouvrages sans songer un instant à leur exploitation en base de données. Quand bien même le contenu est disponible sur un support électronique, tel un fichier *Word*, il nécessite ensuite un long et fastidieux travail manuel de conversion pour être déversé dans une base de données. On peut espérer que ces projets donneront la mesure économique du surcoût engendré par ce travail de balisage et de conversion de dictionnaires conçus « à l'ancienne ».

L'une des principales tâches pour les responsables de ces projets est donc d'abord de clarifier ces données à l'aide du format de balisage S.G.M.L. de manière à ce qu'on puisse identifier clairement chaque information : terme, synonyme, définition, etc. Dans la mesure du possible, on essaye d'utiliser des balises et une structure qui soient communes au plus grand nombre, en sorte que les données puissent être échangées et facilement adaptées aux exigences de telle ou telle interface d'interrogation. Plusieurs consortiums ont ainsi choisi de recourir au format-pivot GENETER¹⁹, conçu par André Le Meur à l'Université de Rennes II, et qui a été proposé à l'ISO comme norme internationale d'échange de données terminologiques. GENETER suit une structure hiérarchique qui vise à décrire successivement le concept, puis la langue, puis le terme. Cette structure, déjà décrite par Wüster comme étant de loin la plus apte à permettre de résoudre des problèmes d'équivalence, devrait être reconnue comme une évidence impérieuse. Le paradoxe est qu'un produit phare comme *Multiterm*, utilisé par un grand nombre de bureaux de traduction, n'impose pas cette structure minimale, en sorte que rien – en dehors du savoir-faire du terminologue – ne garantit la conception d'un glossaire multilingue réellement fiable.

Lire le descriptif du format-pivot GENETER est assez édifiant. On se rend compte qu'il s'agit d'une véritable « usine à gaz » destinée à tenir compte de toutes les hérésies terminologiques observables à l'intérieur de bases de données et de dictionnaires parfois réputés. On ne peut qu'émettre le souhait que l'ensemble de ces projets débouchera sur une volonté commune de concevoir un logiciel de production terminologique qui, tout en offrant une très large liberté d'utilisation, garantisse le respect fondamental de la hiérarchie concept-langue-terme.

Une autre grande difficulté pour les responsables de ces projets est de concevoir une exploitation commerciale des données. Internet fourmille de petits sites offrant un accès gratuit à des données terminologiques. Si leur qualité est souvent inégale, il ne faut pas oublier qu'*Eurodicautom* est gratuitement accessible sur la toile, et ce avec l'appui du même programme MLIS. Tout nouveau site commercial devra donc impérativement présenter des données d'une qualité certaine, couvrant des domaines clairement identifiables, pour pouvoir espérer atteindre un équilibre commercial. En outre, il faut se rappeler que s'il ne travaille pas dans une grande entreprise, une grande organisation, une administration ou une université, l'internaute doit payer son accès au réseau téléphonique : il ne prendra la peine de se connecter et d'ouvrir son porte-monnaie que s'il a la certitude de trouver l'information dont il a impérieusement besoin.

Finalement se pose le problème de la protection des sources : les pirates de l'Internet ne doivent pas pouvoir s'introduire sur le site de l'éditeur pour s'emparer de la totalité des données mises en vente.

19 www.uhb.fr/geneter

2.2.2.2 « Forums » terminologiques

Le deuxième appel à propositions lancé par MLIS à propos de la terminologie consiste en la création de « forums ». Nous mettons le mot entre guillemets, car il fait d'abord songer aux forums de discussion entre traducteurs ou à des listes de distribution comme celle de l'association T.L.S.²⁰ ou celle de l'Union latine²¹. Il s'agit plutôt, ici, de promouvoir la conception de sites Internet qui offriraient des possibilités de concertation aux rédacteurs de dictionnaires multilingues. Si certains grands organismes, tels l'Organisation internationale du travail²² ou l'Union internationale des télécommunications²³ ont déjà mis leurs bases de données sur Internet, d'autres pourraient être tentés de rejoindre le mouvement, surtout si on leur montre qu'un site de rédaction « virtuel » permet un gain de qualité et de productivité. Logiquement, ces projets intéressent d'abord des organisations qui ont une position de leaders ou de normalisateurs dans leur domaine et qui ont des données terminologiques rédigées par des équipes dispersées dans le monde.

S'il est vrai que l'on trouve déjà sur Internet de très belles pages de liens consacrées à la terminologie²⁴, il semble temps pour les terminologues et les terminographes de passer à l'étape ultérieure et d'engranger immédiatement des données dans une base commune, comme cela se passe déjà en néologie dans le cadre du projet *Balnéo* (cf. 2.1). Jusqu'à présent, les équipes de rédaction internationales doivent se réunir une ou deux fois par an pour organiser leur travail, évoquer les problèmes de chevauchement culturels, résoudre les désaccords, etc. Une interface commune sur Internet permettrait d'échanger des données provisoires, d'utiliser des outils de dépouillement, de consulter de la documentation en ligne, d'interroger des experts du domaine, de prendre en compte les commentaires souvent pertinents des utilisateurs et de distribuer une version de la base de données sans cesse remise à jour. Un forum devrait pouvoir également fournir des produits dérivés d'aide à la traduction, comme des microglossaires, des textes alignés ou encore des mémoires de traduction, à l'exemple de ce que l'on trouve déjà sur le site de l'Organisation mondiale de la santé²⁵.

3 Les grandes tendances

3.1 La voie pragmatique

Beaucoup d'efforts de recherche concernent la traduction assistée par ordinateur (T.A.O.), une approche pragmatique, qui demeure assurément la plus profitable à court terme. Les sociétés de développement d'outils de T.A.O. les plus réputées sont celles qui proposent une véritable « boîte à outils » du traducteur, utilisable sur un ordinateur personnel fonctionnant sous les différentes versions de *Windows* et interfaçable avec les principaux traitements de texte. Cette boîte à outils intègre le plus souvent :

20. www.uhb.fr/listes/tlsfrm@uhb.fr/

21. termilat@francenet.fr

22. *Iloterm* : ilis.ilo.org/ilis/ilisterm/ilintrte.html

23. *Termite* : www.itu.int/ITU-Databases/Termite/index.html

24. On les découvrira au départ de l'*Infoport de la terminologie* : www.refer.org/termisti/infoport.htm.

25. www.who.int/pll/cat/tao_resources.html

- un logiciel de gestion terminologique ;
- un gestionnaire de mémoire de traduction ;
- un aligneur de corpus capable d'alimenter la mémoire de traduction ;
- voire, un outil de traduction automatique susceptible de servir épisodiquement.

Pour l'heure, dans les bureaux de traduction, on assiste à de réels investissements dans de semblables postes de T.A.O. À défaut d'études chiffrées, des échos convergents donnent à penser que ce type de poste de travail n'est réellement efficace que s'il est géré correctement dans le cadre de domaines de traduction marqués par une forte redondance lexico-syntaxique et donnant lieu à une production textuelle conséquente²⁶. Force est, par ailleurs, de constater que ces postes de travail, essentiellement dédiés au traducteur n'intègrent généralement pas de fonctionnalités destinées à faciliter le travail du terminologue, preuve que les deux activités sont considérées comme deux marchés distincts. On ne trouve guère de véritable « poste de travail du terminologue » commercialisé, tout au plus une myriade de logiciels, que les rares organismes disposant d'un véritable service de terminologie tentent d'interfacer.

Un certain nombre de contributions publiées dans le numéro spécial de *Terminologies nouvelles* consacré aux banques de terminologie²⁷ présentent de semblables logiciels plus ou moins spécialement dédiés aux terminologues, qu'il s'agisse d'un véritable poste intégré, comme *Latter*, utilisé au Secrétariat d'État du Canada pour enrichir *Termium*²⁸, ou de logiciels de recherche de candidat-termes, tels *Tact*, *Sato*, *Termplus*, *Lexter* ou *Adepto-Nomino*.

3.2 La révolution de l'intelligence artificielle

Il reste que toujours confrontés à l'urgence, la plupart des terminologues travaillant dans des bureaux de traduction se bornent à effectuer un travail de compilation critique des dictionnaires disponibles. Ce réflexe est des plus naturels chez des personnes qui ont généralement une formation initiale de traducteurs. Force est de constater que les outils d'exploitation de corpus sont surtout utilisés par des universitaires qui se sont intéressés aux vocabulaires de spécialité par l'intermédiaire de la lexicologie et de la linguistique de corpus.

Ce sont sans conteste les mêmes qui explorent actuellement la piste, déjà évoquée au travers d'EUROWORDNET (cf. 2.2.1), des bases de connaissances terminologiques et qui passent plus de temps à dialoguer avec les spécialistes de l'intelligence artificielle qu'avec les traducteurs. On en arrive à produire aujourd'hui des logiciels qui, comme *Nomino v2*, non seulement extraient des termes, mais arrivent aussi à identifier les concepts désignés et à isoler leurs relations sémantiques. Ainsi, le logiciel est capable d'extraire des connaissances d'un texte consacré, par exemple, aux oiseaux de la corne de l'Afrique, et de préciser à quelle famille appartient tel oiseau, dans quel environnement il niche, de quoi il se nourrit, etc.

26. Heidi (U.), « La mise en place du Translator's Workbench (TWB) : concurrence avec SYSTRAN et élément humain », dans *Terminologie et traduction*, Luxembourg, Office des publications officielles des Communautés européennes, vol. 1.98, p. 102-116.

27. Rousseau (L.-J.), édit., *Banques de terminologie. Actes de la table ronde (Québec, 18- et 19 janvier 1996)*, dans *Terminologies nouvelles*, Bruxelles, Agence de la francophonie et Communauté française de Belgique, juin et décembre 1996, n° 15.

28. www.bureaudelatraduction.gc.ca/term-1f.htm

On en arrive à une nouvelle forme de dictionnaire spécialisé où la définition voire l'équivalence pourraient être générées par la machine. Et qui permettrait à tout moment de retourner vers des textes spécialisés pour obtenir des compléments d'information ou des exemples de fonctionnement phraséologique.

L'idée s'impose progressivement qu'il convient de changer les contenus et la présentation des dictionnaires parce que l'informatique permet à présent de s'affranchir totalement des impératifs de présentation rendus nécessaire par la lecture sur papier. On peut désormais parler de « fiche virtuelle » en ce sens qu'un traducteur chargé de la traduction du manuel de maintenance d'un réacteur d'Airbus pourra aisément savoir où se situe telle pièce, à quoi elle sert, quels sont ses composants, quelles sont les expressions figées dans lesquelles intervient sa désignation en français et comment la désignent les ingénieurs espagnols... Toute cette information sera recherchée en temps réel dans la base de connaissances. En dehors des cas litigieux ou des domaines du savoir qui ne se prêtent guère à l'extraction de connaissances, il ne sera donc plus nécessaire d'accéder à une fiche terminologique conforme au modèle conceptuel enseigné dans les bonnes institutions universitaires et conçue au fil d'une démarche qui va du texte vers les dictionnaires en passant par l'interrogation des spécialistes. Ce sera d'autant plus profitable que cette démarche, rarement mise en pratique dans la vie professionnelle parce que trop fastidieuse, est précisément celle qui est mise en œuvre dans les bases de connaissances.

3.3 Les progrès de la traduction automatique

En dépit des croyances, la traduction automatique fait d'énormes progrès pour certains couples de langues. On se doit aussi de constater qu'une traduction brute suffit souvent à satisfaire un lecteur qui est plus intéressé par l'information technique ou scientifique que par la qualité du texte. Ce n'est assurément pas un hasard si les principaux moteurs de recherche utilisés sur Internet proposent une traduction automatique, même perfectible. Par ailleurs, Les regroupements que l'on observe ces derniers mois dans le secteur privé entre les bureaux de traduction et les entreprises de génie linguistique suggèrent l'existence d'une volonté économique forte de progresser dans la direction de la traduction automatique, de la reconnaissance vocale et de la synthèse de la parole.

Il n'est pas sot de penser que demain, on aura autant besoin de traducteurs capables de surveiller les outils de traduction automatique que de traducteurs traditionnels. Dans un tel cadre, on peut s'interroger sur le devenir des fiches terminologiques. Le traitement automatique des langues naturelles n'a pas recours au même type d'information que la traduction «manuelle ». Il exige la conception de ce que l'on nomme des «lexiques-grammaires », fournissant des informations très précises sur le comportement grammatical et sémantique du mot dans la phrase. La question se pose donc de l'avenir, à très long terme, des fiches terminologiques telles que nous les concevons actuellement pour une armée de traducteurs. La problématique de leur adaptation éventuelle ouvre des perspectives de recherche très intéressantes, surtout si l'on tient également compte des avancées des bases de connaissances terminologiques.

Une bonne manière de se construire une idée de la problématique actuelle consiste à consulter le numéro de *Terminologie et traduction*²⁹ récemment consacré à la gestion de la terminologie, de la traduction assistée et de la traduction automatique au sein de la Commission européenne.

3.4 L'impératif de la formation

L'ensemble de ces perspectives soulève finalement la question de la formation des acteurs professionnels. Le métier de terminologue existe-t-il réellement et continuera-t-il d'exister ? quel terminologue former ? à quel niveau ? pour quoi faire ? Le numéro 17 de la revue *Terminologies nouvelles*³⁰ aborde certains aspects de ce vaste champ d'interrogation.

Il semble en tout cas certain que la pratique évoluera considérablement au fil des années et que comme dans beaucoup d'autres domaines professionnels, les acteurs de terrain devront être capables d'adapter en permanence leur profil aux exigences du moment. Le premier impératif semble donc d'éviter de tomber dans le piège d'une formation uniquement dédiée aux outils du moment et de plutôt veiller à acquérir une bonne culture générale tant en matière de langue que de linguistique appliquée.

Marc Van Campenhoudt
Centre de recherche Termisti
Institut supérieur de traducteurs et interprètes
Bruxelles

29. *Terminologie et traduction*, Luxembourg, Office des publications officielles des Communautés européennes, vol. 1.98.

30. Van Campenhoudt (M.), édit. *Terminologie et formation*, dans *Terminologies nouvelles*, Bruxelles, Agence de la francophonie et Communauté française de Belgique, décembre 1997, n° 17.