

Systemes d'information multilingues au service du traducteur

Dans le cadre de la modernisation des méthodes de travail au Service de traduction de la Commission européenne, et dans la perspective d'une ouverture encore plus large de l'Europe multilingue, un certain nombre d'outils informatiques ont été développés ou adaptés dans le but d'améliorer la gestion des problèmes terminologiques et documentaires. Les solutions proposées mettent l'accent sur la réutilisation des données dans une perspective d'intégration toujours plus poussée.

Termes-clés :
multilinguisme ; traduction ;
terminologie ; documentation ;
outils informatiques intégrés ;
Europe élargie.

Notre propos n'est pas de décrire ici un xième logiciel de gestion documentaire ou terminologique, mais plutôt de présenter un ensemble intégré de solutions aux problèmes qui se posent aux traducteurs des institutions européennes, projet principalement connu sous le nom d'Euramis (European advanced multilingual information system).

À tout seigneur, tout honneur : chacun connaît *Eurodicautom*, la base de données terminologiques multilingue gérée par la Commission des Communautés européennes, qui abrite les données de la Commission, du Conseil de l'Union, du Parlement européen, de la Cour de Justice, des autres institutions et organes de l'Union, ainsi que de multiples vocabulaires spécialisés des agences de l'ONU, de ministères, d'organismes de terminologie, ou d'associations internationales.

Historiquement, *Eurodicautom* a été conçu comme un système essentiellement interactif, destiné à être utilisé directement par les traducteurs de la Commission, ouvert au monde extérieur pour des consultations en ligne, mais permettant aussi des traitements différés, comme l'extraction de glossaires synoptiques adaptés au document à traduire, ou des glossaires thématiques, éventuellement diffusés par l'Office des publications des Communautés européennes. Cette orientation interactive explique sans doute pourquoi l'accès est aujourd'hui possible à travers diverses interfaces, donc une interface Web, donnant

accès à l'ensemble des commandes d'une façon simplifiée.

Les objectifs assignés à *Eurodicautom*, à savoir de contribuer à accroître la productivité des services, la qualité et l'homogénéité des textes, n'ont jamais été remis en cause et sont plus que jamais à l'ordre du jour pour l'ensemble des outils d'aide à la traduction et à la rédaction qu'ils convient de mettre au point, au moment où l'Europe négocie un nouvel élargissement, qui accroîtra encore sa couverture linguistique.

Parmi les difficultés rencontrées pour mettre en adéquation ces objectifs avec le développement d'une base comme *Eurodicautom* et la réalité de l'activité traduisante, un aspect doit être davantage souligné qu'il ne l'a été dans d'autres articles ou présentations. En effet, d'une part, tous les problèmes rencontrés par les traducteurs ne sont pas de nature terminologique, et, d'autre part, la meilleure des bases ne rendra aucun service si elle n'est pas consultée.

Sur ce dernier point, le manque d'intérêt suscité par la consultation par lots d'*Eurodicautom* a très tôt interpellé ses concepteurs, et deux pistes ont pu être suivies pour expliquer ce phénomène.

D'une part, le produit fourni ne correspondait pas aux besoins. En effet, à ses débuts, la consultation différée permettait de poser une série de questions et de fournir dans les quarante-huit heures une liste imprimée sur les imprimantes rapides du Centre de Calcul. Deux jours, une éternité pour de nombreux documents à traduire dans l'instant ! Et ces listes sur papier, seule présentation possible à l'époque,

manquaient de flexibilité, et obligeaient le traducteur à opérer des choix *a priori*.

L'autre aspect concernait la consultation elle-même. Pour rationaliser au maximum l'opération, il convenait de poser les questions une seule fois, pour le bénéfice de tous les traducteurs concernés, soit que la traduction d'un texte d'une certaine longueur eût été répartie entre plusieurs traducteurs d'une même langue, soit que le même texte fût traduit vers diverses langues officielles. L'expérience a vite montré que, à tort ou à raison, les traducteurs considéraient que les questions posées par un collègue ne rejoignaient pas leurs propres préoccupations, soit que les termes considérés par l'un comme étant difficiles étaient déjà connus par un autre, soit que, à l'inverse, des termes inconnus ou douteux n'avaient pas fait l'objet d'une recherche. Les tentatives opérées pour son propre compte par un traducteur isolé n'étaient guère plus satisfaisantes : trop souvent, même lorsqu'il peut, conformément aux règles de l'art, prendre le temps de lire le texte à traduire avant de se mettre à l'ouvrage, le traducteur lui-même sous-estime la difficulté ou la simplicité d'un terme. Combien de fois, malgré l'expérience, ne nous laissons-nous pas piéger lors d'un survol rapide d'un texte, et les textes les plus simples en apparence ne sont-ils pas les plus ardues à interpréter et à rendre correctement ?

Un autre facteur doit être au moins mentionné, et concerne les stratégies de recherche. En effet, conscients des environnements différents selon qu'on travaille en conversationnel ou en différé, les développeurs avaient mis au point des stratégies en principe adaptées à la recherche par lots, mais qui, faute d'avoir bénéficié de développements appropriés tenant compte de l'expérience sur le terrain, étaient nettement inférieures au niveau de

qualité auquel les recherches interactives avaient habitué les traducteurs.

Dire qu'une réponse optimale a été apportée aujourd'hui à l'ensemble des problèmes documentaires et terminologiques rencontrés par les traducteurs relèverait sans doute d'une tendance à l'exagération dont nous essayons toujours de nous garder. Mais il est un fait qu'un grand pas en avant a été franchi dans le développement et la mise à disposition d'outils et de solutions s'inscrivant de mieux en mieux, d'une façon de plus en plus intégrée, dans le cycle de la communication écrite plurilingue à la Commission.

Schématiquement, l'orientation suivie est la suivante : dès qu'un service remplit une demande (électronique) de traduction d'un document, une série d'outils de recherche sont activés, de sorte que le traducteur, lorsqu'il reçoit le texte à traduire, reçoit en fait un dossier complet, incluant essentiellement, selon les cas, une pré-traduction effectuée par un système automatisé comme *Systran*, une archive de documents électroniques traitant du même sujet, une base de fragments de textes identiques à travers des mémoires de traduction, une liste de références juridiques et administratives pertinentes, une base de données terminologiques constituée dynamiquement à partir d'*Eurodicautom* et d'autres bases de travail sélectionnées. À partir de tous ces éléments regroupés dans son poste de travail, le traducteur peut alors recomposer le texte avec une efficacité maximale à l'aide de fonctions bien adaptées, et en continuant à faire usage du sens critique qui caractérise l'exercice de sa profession. Une fiche linguistique peut également suivre le document tout au long du circuit qu'il est éventuellement appelé à parcourir à travers les différentes institutions, expliquant ainsi aux collègues ayant à connaître du texte

plus loin dans la chaîne, le cheminement l'ayant amené à choisir tel terme, telle formule, etc.

Voyons brièvement, sans entrer dans le détail des caractéristiques techniques individuelles, les principales composantes qui sont activées dans ce processus.

Le système de traduction automatique *Systran* peut être utilisé soit pour fournir une traduction, soit pour repérer des termes ou des références législatives dans un document à traduire. Dans ce dernier cas, pour les langues sources disponibles, le système analysera le texte, en extraira les termes ou les références, et les enverra à *Eurodicautom* ou *Celex* pour la suite du traitement. Pour les couples de langues suffisamment développés, le service de traduction par la machine fait intégralement partie de l'offre du Service de traduction : n'importe quel fonctionnaire peut, à travers une interface Web, obtenir une traduction «brute» en quelques minutes. La qualité obtenue suffit souvent à des fins d'information ; une fois passés par les mains expertes de traducteurs chargés d'en faire une post-édition rapide, ces documents, même s'ils n'ont pas la qualité requise à des fins de publication, conviennent parfaitement aux débats techniques des experts de différents groupes de travail. Pour améliorer les performances du système, un effort a été fait pour établir différentes catégories de textes, et pour alimenter, notamment en s'appuyant sur *Eurodicautom*, les dictionnaires de *Systran* en termes et tournures fréquemment rencontrés.

Pour le traducteur, rien n'est plus frustrant que de devoir retraduire un texte fort semblable à d'autres documents déjà passés à travers ses mains, ou de constater, en déposant sa plume ou en éteignant son ordinateur, que la traduction qu'il vient d'achever était déjà partiellement disponible dans un

autre service. Pour une administration ou une entreprise également, ce genre de situation paraît devoir être évitée dans le souci d'améliorer la gestion de l'information et diminuer les coûts. La possibilité technique de créer des mémoires de traduction est donc une innovation particulièrement importante, qui satisfait à la fois l'employé et le donneur d'ordre. Partant du principe que de nombreux textes, ou morceaux de texte, ont déjà fait l'objet d'une traduction, et constatant que le traducteur, par un effort de mémoire ou par un archivage plus ou moins judicieux de son travail, cherche à retrouver les passages pouvant lui être utiles, un effort a été fait de rassembler, d'une façon structurée, des ensembles de textes originaux et de traductions, afin de se substituer avec une efficacité plus grande à la mémoire collective des traducteurs.

Euramis repose donc essentiellement sur une très grande base de ressources linguistiques. Les textes ou fragments de textes que celle-ci contient, une fois répertoriés, validés, alignés, grâce à des outils de gestion de l'information sophistiqués, peuvent être facilement reconnus par des systèmes informatiques, proposés au traducteur au cours de la traduction, ou bien déjà être intégrés « par surimpression » en quelque sorte dans le document source, sous réserve de l'acceptation ultérieure par le traducteur. On imagine aisément les gains prodigieux qui peuvent en découler, rien qu'en temps de recherche et de frappe ! Les systèmes disponibles aujourd'hui peuvent être paramétrés pour n'accepter que des passages exactement identiques, ou tolérer un certain taux de divergence. Ils peuvent aussi adapter certaines variables du texte retrouvé dans la mémoire au nouvel environnement créé par le texte source, par exemple pour des dates, des unités de mesure, des quantités etc. Le traducteur peut constituer sa propre mémoire à partir

de documents qu'il a sélectionnés, ou puiser dans la mémoire collective et décharger dans son poste de travail les sous-ensembles qui peuvent lui être utiles.

Il n'est toutefois pas toujours possible de retrouver des fragments de texte déjà traduits, ou bien le traducteur peut avoir de bonnes raisons de mettre en doute tel ou tel équivalent d'un terme qu'il a trouvé dans une mémoire. Dans ces cas, il sera heureux de pouvoir continuer à consulter une base comme *Eurodicautom*. Toutefois, pour rationaliser et accélérer le processus, un effort est fait pour modifier les habitudes de consultation. En effet, le balayage automatique des textes, par exemple par *Systran*, permet de générer une base terminologique locale, ciblée, éventuellement enrichie par d'autres apports. Appelée dans un coin de l'écran au cours de la traduction, elle permet à tout moment au traducteur de trouver ou de vérifier la terminologie à utiliser, et de l'intégrer dans sa traduction par simple clic. On voit donc que, tant en entrée qu'en sortie, il a été tenu compte des problèmes psychologiques, techniques et organisationnels évoqués plus haut.

Les mêmes routines d'analyse sont aussi utilisées par le mécanisme de recherche dans la base législative *Celex* des références à la législation communautaire apparaissant dans les documents à traduire. Ceci permet de retrouver et d'intégrer sans effort des titres de règlements ou de directives, mais aussi de vérifier plus facilement les usages terminologiques particuliers de tel ou tel concept dans tel ou tel texte de loi.

Lorsque l'information structurée et validée a atteint ses limites, le traducteur peut souhaiter étendre sa recherche, documentaire et terminologique, à un ensemble de textes « en vrac », qu'il sélectionnera éventuellement en fonction de l'auteur, du service, du numéro, du

titre, du contenu. Il pourra alors faire appel à *SdT vista*, le serveur de documents du Service de traduction, qui contient tous les documents électroniques qui entrent au service, et toutes les versions linguistiques qui en sortent. Ce système permet aussi de rassembler des textes qui peuvent contribuer à la formation du traducteur, ou du moins qui lui permettent de se familiariser rapidement avec un domaine qu'il maîtrise moins bien ; un effort est en cours actuellement pour trouver une méthode simple pour mieux identifier automatiquement le contenu réel d'un texte. Notons enfin que le fait de disposer facilement des textes originaux est une amélioration importante. Trop souvent par le passé, les traducteurs disposaient, dans leurs centres de documentation, des textes qu'ils avaient traduits, mais ils avaient plus difficilement accès à l'ensemble des documents pertinents, et plus difficilement encore aux originaux dans leur langue que, par définition, ils n'étaient pas amenés à traduire, et qui pourtant contiennent la terminologie et la phraséologie qu'il convient de retrouver dans les traductions. Cet impressionnant fonds documentaire est également accessible à d'autres directions générales pour les parties qui les concernent, ce qui contribue à une gestion plus rationnelle de l'information dans l'institution.

Comme on le voit, le développement d'outils d'aide à la traduction s'est considérablement diversifié et amplifié au cours de ces dernières années. Il tient compte à la fois des expériences accumulées lors de l'utilisation d'outils déjà « banalisés » comme *Eurodicautom*, de l'évolution des besoins des traducteurs en fonction de la réalisation de plus en plus poussée de l'Union européenne, des possibilités nouvelles offertes par la technique, et des nécessités découlant d'une organisation du travail faisant davantage appel à la

subsidiarité. Les prochains élargissements de l'Union accéléreront encore sans nul doute ces développements et modifieront encore les méthodes de travail. La traduction du droit primaire et du droit dérivé dans les langues des pays candidats peut également bénéficier des efforts consentis, le contenu des outils évoqués ci-avant pouvant s'enrichir et s'améliorer au fur et à mesure de l'avancement du travail. En effet, il ne faut jamais perdre de vue que les outils, aussi performants soient-ils, ne sont pas une panacée, et ils ne remplissent leur rôle d'une façon optimale que si les données qu'ils utilisent sont de qualité, et si l'organisation du travail dans un cadre interinstitutionnel évolue en symbiose avec leur développement et leur installation dans les services.

*Alain Reichling,
Chef de secteur,
Commission européenne,
Luxembourg.*

Bibliographie

T&T, Terminologie et Traduction, la revue des services linguistiques des institutions européennes, Luxembourg, Office des publications officielles des Communautés européennes, vol. 1.98, ISSN 0256-7873.