

TI, le logiciel d'expérimentation notionnelle de Termisti

L'Institut supérieur de traducteurs et interprètes (Isti, Bruxelles ⁽¹⁾) a lancé un nouveau projet de recherche avec l'aide financière du Ministère de la Communauté française (Direction générale de la culture: Service de la langue française). Ce projet réunit informaticiens, enseignants et chercheurs en terminologie sous l'appellation *Termisti*. Il a pour objectif de réaliser, à travers un système de gestion micro-informatique, une description de micro-glossaires propres aux terminologies de pointe.

1 Une hypothèse de travail: luxuriance et écosystèmes

Dans une intervention récente, Daniel Blampain (1991), promoteur et coordinateur du projet *Termisti*, a montré que le terminologue était confronté à une luxuriance protéiforme: luxuriance des sciences et des techniques, luxuriance des banques de terminologie, luxuriance des méthodes, luxuriance des logiciels, etc. Face à ce constat, il justifiait l'idée d'une terminologie fondée sur «le concept d'écosystèmes qui implique non seulement l'étude d'équilibres et la partition d'une entité en systèmes relativement autonomes dont il est possible d'analyser la structure et le fonctionnement, mais aussi la mise en évidence de mécanismes de régulation».

Dans une base de données, les notions doivent, selon cette hypothèse, prendre place dans une arborescence gérée par le logiciel à la manière d'un semblable écosystème. Ceci suppose d'abord que l'on traite des micro-glossaires, c.-à-d. des ensembles de notions qui, issues d'un même domaine, assument leur autonomie par le nombre et le type des liens qui les unissent. Cela suppose ensuite que l'on identifie des règles qui permettent de gérer lesdits liens au sein de l'arborescence. Selon ce point de vue, la mise au point d'«écosystèmes terminologiques» devrait répondre aux exigences de

l'environnement informatique qu'utilise le traducteur confronté aux langages de spécialité.

L'objectif de description des micro-glossaires implique donc deux démarches parallèles et complémentaires: décrire des micro-domaines et établir les règles qui les structurent.

2 Micro-domaines et terminologies de pointe

Pour accomplir la première démarche, on a choisi de se pencher sur l'étude de terminologies nouvelles utilisées dans des secteurs clairement délimités de la recherche universitaire de pointe. Ce faisant, on espère découvrir les systèmes évoqués, partant du postulat que la terminologie du praticien répond au concept d'écosystème évoqué.

3 A la recherche de l'écosystème régulé

3.1 Comment délimiter la notion?

L'analyse des diverses approches théoriques a permis d'établir que seule une approche notionnelle permet de résoudre les problèmes traditionnels de la traduction des langages de spécialité (vide notionnel, intersection partielle, inclusion, etc.). A cet égard, il est bon de rappeler que le Secrétariat d'État d'Ottawa s'intéresse aujourd'hui – après vingt années de pratique terminologique – à un logiciel qui se réfère aux réseaux notionnels (*Code*, établi en

(1) L'Isti est un institut universitaire dépendant du Ministère de l'Éducation et de la recherche de la Communauté française. Il forme des traducteurs et des interprètes en quatre années d'études.

collaboration avec l'Université d'Ottawa).

On regrettera, par ailleurs, que l'on ait si souvent segmenté la planète des terminologies entre «descripteurs» et «prescripteurs». En effet, certaines vues simplistes aboutissent à prétendre que seule une démarche onomasiologique, dressant un inventaire préalable des notions, permet de réaliser l'approche notionnelle évoquée ci-dessus. On trouve ainsi des affirmations selon lesquelles les notions peuvent être organisées dans un système unifié une fois pour toute et apte à rendre compte de toutes les langues.

(2) La perspective ici évoquée est celle que nous tentons actuellement de développer dans une thèse de doctorat à l'Université Paris-Nord, sous la direction du professeur Pierre Lerat.

(3) Cette manière de distinguer polysémie et homonymie en fonction d'une autre langue est déjà appliquée en langue courante dans un dictionnaire alphabétique tel le *Robert*. On y considère, par exemple, qu'il y a deux mots *éponge*, car on peut identifier deux étymologies différentes: *sponga* et *sponda*. Les dictionnaires homonymiques distinguent, eux, les entrées en fonction de critères sémantiques arbitraires.

(4) On pourra nous reprocher de ne pas travailler, pour cet exemple, à partir d'une langue de spécialité. Le problème ici évoqué possède pourtant un caractère abstrait qui suscite une réflexion applicable à la terminologie: on se figurera aisément que le schéma présente la manière dont chaque langue identifie les types d'une pièce.

(5) Les couples de deux lettres ne correspondent qu'à une dénomination arbitraire de la notion à prévoir et nullement à son extension: la notion [af] ne possède que l'extension de la notion [a].

L'expérience montre que cette idée procède d'une confusion quelque peu naïve entre terminologie et science de la documentation, voire d'une forme inconsciente d'impérialisme linguistique qui consisterait à imaginer que les quelques grandes langues européennes couvrent l'ensemble des «possibles».

Même la lecture approfondie de *La République* de Platon ne nous apprendra pas comment identifier une notion. Ainsi, dans la terminologie nautique anglaise, il semble difficile d'identifier avec certitude combien de notions recouvre le terme *stay*. Celui-ci s'applique à toutes sortes d'objets (coque, machine, grément) qui occupent des fonctions relativement semblables, constat qui conduit les dictionnaires spécialisés à considérer qu'il s'agit d'une même vaste notion à laquelle ils attribuent une entrée et plusieurs acceptions. Cet exemple montre que dans un langage de spécialité donné, comme dans le langage courant, rien ne permet de distinguer la limite entre polysémie et homonymie.

3.2 Une notion «internationalement conditionnée»⁽²⁾

Notre hypothèse de travail est que la traduction rend pourtant l'approche notionnelle tout à la fois justifiée et possible. Confrontant la terminologie de deux langues, on pourra, en effet, comparer les différences de découpage notionnel d'une même réalité. L'observation des équivalences de *stay* en français permet d'établir que dans un système de traduction français-anglais apte à fonctionner dans les deux sens, il convient de prévoir six notions, quitte à créer des notions homographes dans une langue⁽³⁾.

1. *Stay* = *étai* (d'un mât)
2. *Stay* = *draille* (d'une voile)

3. *Stay* = *entretoise* ou *tirant* (d'une chaudière à vapeur)
4. *Stay* = *jambe de force* (d'un pavois)
5. *Stay* = *traverse* (d'un gouvernail)
6. *Stay* = *séjour* (dans un port)

Cela signifie que, en traduction, la notion est «internationalement conditionnée». Elle résulte de l'observation des intersections, inclusions, chevauchements, etc. comme Lyons (1970: 46-47) l'avait déjà entrevu dans son approche du problème des couleurs. Si l'on reprend ce problème (schéma ci-dessous), on conviendra qu'une base de données trilingue (langues A, B et C) devra prévoir 12 notions pour fonctionner dans les six sens de traduction envisageables (A → B, A → C, B → A, B → C, C → A, C → B).

français	rouge	orange	jaune	vert	bleu
A	a	b	c	d	e
B	f	g	h	i	j
C	p		q	r	s
	1	2 3	4	5 6	7 8
				9	10

De notre point de vue, la notion est, dans le cadre de la traduction, le résultat variable d'une comparaison du découpage de la réalité par deux langues différentes⁽⁴⁾: la notion la plus restreinte (celle qui possède la plus petite extension) dans une langue, impose son découpage aux notions plus vastes des autres langues. Le nombre et l'étendue de l'extension des notions varieront donc en fonction du couple de langues retenu. Ce principe n'est autre que celui de l'isotopie.

- A-B = 9 notions, soit: [af], [bf], [bg], [cg], [ch], [dh], [di], [dj], [ej];⁽⁵⁾
- B-C = 8 notions, soit: [fp], [gp], [gq], [hq], [iq], [ir], [jr], [js];
- A-C = 8 notions, soit: [ap], [bp], [bq], [cq], [dq], [dr], [er], [es].

A ce principe s'ajoute un autre: comme la même base peut inclure plus de deux couples de langues, le découpage doit s'effectuer en prenant en compte toutes les notions nécessaires à une traduction correcte quel que soit le couple sélectionné et le sens de la traduction. Ainsi, une base trilingue devra prévoir les extensions des notions identifiées ci-avant, soit 12 notions différentes. Chaque fois que l'on ajoutera une nouvelle langue, de nouvelles notions «internationales» devront inévitablement s'intégrer dans la base et être traduites, tandis que d'autres deviendront caduques.

- A-B-C = 12 notions, soit: [afp], [bfp], [bgp], [bgq], [cgq], [chq], [dhq], [diq], [dir], [dir], [ejr], [ejs].

Selon notre hypothèse, les relations qui unissent les notions unilingues aux notions bilingues et les notions bilingues aux notions trilingues sont des relations de types hiérarchiques qui pourraient être gérées dans un réseau notionnel informatisé. Nos travaux actuels tentent de développer des algorithmes aptes à accomplir cette tâche. Le «conditionnement international» pourrait ainsi être mis en évidence par le recours à la machine.

3.3 Le logiciel *TI*

Un logiciel prototype a été conçu qui réponde aux premiers développements théoriques de la recherche. Baptisé *TI* – initiales de *TermIsti* –, il gère des données notionnelles entre lesquelles sont établis des liens hiérarchiques et coordonnés. Nos analyses confirment, en effet, la pertinence des principes wüstériens en la matière et il nous paraît dommage qu'ils n'aient guère débouché sur une étude de leur application systématique.

Le principe fondateur de *TI* s'inscrit donc dans le prolongement

de l'École de Vienne: une notion, sa définition dans chaque langue, un nombre indéterminé de synonymes et d'équivalents et des données sur l'utilisation de chaque équivalent (notamment un contexte). L'ensemble est conçu de telle manière que chaque utilisateur s'y retrouve, sans qu'il importe que sa démarche soit dite *onomasiologique* ou *sémasiologique*, *prescriptive* ou *descriptive*. Les champs présents sur la fiche de *TI* s'inspirent de ceux d'*Eurodicautom*, banque avec laquelle des échanges sont dès lors réalisables⁽⁶⁾.

La version prototype prévoit déjà l'établissement de relations notionnelles. La conception s'inspire notamment des travaux de Lerat (1990: 84-86), lesquels ont montré l'avantage des formules prédicatives pour réaliser une telle tâche. *TI* permet au terminologue de créer à sa guise des prédications et de les utiliser pour marquer les liens unifiant les notions au sein de l'arborescence.

La prédication permet une approche beaucoup plus transparente du réseau et contribue ainsi à l'ergonomie du système. Elle est un complément idéal à la définition et au contexte, car le système permet des interrogations très puissantes:

- *Quelles sont les parties de X?*
- *Que trouve-t-on devant X?*
- *A quelle espèce appartient Y?*
- *De quel prédicat X est-il l'argument?*
- *Précisez tous les liens qui unissent Z aux autres notions de l'arborescence.*
- *Etc.*

Dès à présent, nous avons pu observer que les notions hyponymes situées tout au bas de l'arborescence par la relation espèce-genre n'étaient guère définies dans les ouvrages et dictionnaires spécialisés. Elles sont, tout au plus, énumérées dans la définition de l'hyperonyme. Du fait de leur place dans l'arborescence, ces notions hyponymes font justement l'objet de nombreuses prédications

qui suppléent adéquatement l'absence de définition.

3.3.1 Développements

Les progrès de l'intelligence artificielle nous donnent à penser que, analysant les prédicats d'un contexte, la machine pourrait proposer une assistance dans la structuration du réseau. Une approche sémantique comparant les découpages effectués dans chaque langue permettrait même d'identifier l'extension des notions internationales évoquées ci-dessus. La gestion du réseau pourrait ainsi être assistée par l'ordinateur à chaque fois que l'on intégrerait une nouvelle langue.

Cet objectif à long terme sera précédé de réalisations intermédiaires que justifie pleinement l'importance des relations notionnelles pour les langages de spécialité. Ainsi, la production de glossaires, lexiques et autres dictionnaires a tout à gagner d'une approche notionnelle. Les impossibilités de traduction sont décelées sinon résolues, les synonymes sont identifiés, les homographies sont volontairement prévues, etc.

Grâce au réseau, l'arrangement systématique des versions papiers peut être fondé sur une véritable logique et non plus sur le vague sentiment d'une logique arbitraire, évidente aux seuls yeux du terminographe-concepteur. Ainsi, des ordres de priorité dans les relations peuvent être définis pour l'impression d'un glossaire systématique. Par exemple, on demandera d'imprimer les fiches consacrées à tous les types d'une

(6) Le choix d'*Eurodicautom* s'explique par le souhait d'être compatible avec une grande banque notionnelle que consultent nos étudiants et qui a déjà servi de modèle commun à d'autres logiciels (*MC4* en France et *Lingua-PC* en Suisse, par exemple).

pièce que l'on trouve dans une machine en procédant du haut de la machine vers le bas et en fournissant toujours les composants de chacune des pièces du même type.

Les dernières études sur l'environnement informatique du traducteur (notamment Tremblay 1991) montrent tout l'intérêt d'une interface graphique dont le système multi-fenêtres permet d'unir l'ensemble des tâches. L'accès simultané au traitement de texte, aux dictionnaires et aux banques de données constitue un avantage certain. Nous pensons notamment au transfert des données entre les divers programmes (exécution d'un simple « copier-coller » grâce à la souris), à la possibilité d'ouvrir plusieurs fois la même application – pour autant que celle-ci le permette – dans des fenêtres différentes, à l'appel d'une image à partir de la notion (une fiche) ou de la notion (une fiche) à partir de l'image⁽⁷⁾. Actuellement développé sous *Dos*, *TI* pourrait être prochainement rendu utilisable dans un système qui, tel *Windows*, cumule ces avantages⁽⁸⁾.

(7) La possibilité de consulter des illustrations (plan, planches, écorchés...) présente un intérêt certain pour le traducteur. Cependant, nous ne percevons guère l'avantage qu'il aurait à visualiser un réseau notionnel graphique. A moins de pouvoir sélectionner seulement certains types de liens, on aboutit fréquemment à une toile d'araignée où le bruit dépasse l'information. De ce point de vue, la lecture de courtes prédictions nous semble beaucoup plus performante.

(8) Notons toutefois que *Windows* est gourmand en ressources matérielles et exige un investissement important auquel tout traducteur n'est pas encore prêt à consentir.

4 Conclusion

La dynamique de recherche développée par le groupe Termisti se veut déjà un écosystème dédié à la terminologie: plusieurs démarches y sont conduites de front et s'y nourrissent mutuellement. Certes, l'informatique et la sémantique lexicale jouent un rôle important dans ce projet, mais en étroite dialogue avec la description terminologique de micro-glossaires, car une modélisation pertinente suppose d'abord une connaissance adéquate.

*Marc Van Campenhoudt,
Institut supérieur de traducteurs et
interprètes de la Communauté française
de Belgique,
Bruxelles.*

Bibliographie

Blampain, Daniel; Petrusa, Philippe et Van Campenhoudt, Marc, 1991: «A la recherche d'écosystèmes terminologiques», Actes du colloque international (Aupelf - Uref) *L'environnement traductionnel* (à paraître), Université de Mons, avril 1991.

Lerat, Pierre, 1990: «L'hyponymie dans la structuration des terminologies», dans *Langage* n° 98, juin 1990, p. 79-86.

Lyons, John, 1970: *Linguistique théorique*, Paris, Larousse, 382 p.

Tremblay, Klaire, 1991: «Les projets *Logos* et *PTT*», Actes du colloque international (Aupelf - Uref) *L'environnement traductionnel* (à paraître), Université de Mons, avril 1991.