

Dictionnaires informatisés et traduction

Agnes Kukulska-Hulme

Volume 34, numéro 3, septembre 1989

1. Actes du Colloque Les terminologies spécialisées : Approches quantitative et logico-sémantique et 2. Actes du Colloque Terminologie et Industries de la langue

URI : <https://id.erudit.org/iderudit/003359ar>

DOI : <https://doi.org/10.7202/003359ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Les Presses de l'Université de Montréal

ISSN

0026-0452 (imprimé)

1492-1421 (numérique)

[Découvrir la revue](#)

Citer cet article

Kukulska-Hulme, A. (1989). Dictionnaires informatisés et traduction. *Meta*, 34(3), 533–538. <https://doi.org/10.7202/003359ar>

DICTIONNAIRES INFORMATISÉS ET TRADUCTION

AGNES KUKULSKA-HULME
Modern Languages Department
Aston University, Birmingham, U.K.

L'année 1983 marque à l'Université d'Aston le début d'un programme de recherches ayant pour objectif l'exploration du potentiel de dictionnaires informatisés vis-à-vis des processus de rédaction et de traduction de textes techniques. Pour la traduction, le développement de maquettes de dictionnaires s'est fait conjointement avec une étude des rapports entre texte, processus de traduction, état des connaissances du traducteur, organisation conceptuelle du domaine de référence, exigences terminologiques et unités de traduction. Nos travaux ont produit plusieurs approches prometteuses dont je vais exposer ici les grandes lignes ; il conviendra également de les mettre en balance et de réfléchir sur des questions importantes déclenchées par les résultats.

On a pris comme point de départ le problème essentiel qui se pose au traducteur, notamment de trouver le maximum d'informations sur les unités terminologiques rencontrées dans le courant de la traduction, dans une forme «prêt-à-porter», le plus vite possible. Il lui faut savoir ce qu'une unité dénote, et comment elle fonctionne. Dans l'amalgame de données que l'on peut rassembler autour d'un terme, les données linguistiques telles que abréviations ou modèle syntactique sont entrelacées avec des connaissances spécialisées du domaine de traduction, par ex. champ d'application, relations logiques et ontologiques, processus dont un objet fait partie, et ainsi de suite. Certains éléments ont un rôle double : ainsi la définition, elle, explicite le sens du terme et donne aussi «un instantané» d'une partie de la connaissance du domaine. Pour la traduction humaine, une séparation formelle de ces aspects semble inutile ; il faut cependant que des informations de plusieurs types soient accessibles.

L'innovation en matière de dictionnaires techniques peut se traduire par des données terminologiques d'un nouveau type, par une organisation nouvelle, ou bien par une nouvelle méthode d'accès aux informations retenues dans le dictionnaire. C'est la distinction entre classification et indexation qui nous permet d'envisager des méthodes de repérage efficaces appliquées à des données amoncelées autour de termes vedettes, sans que ceux-ci soient rangés dans un ordre déterminé. Une consultation dont le point de mire n'est que l'équivalent en langue cible d'un terme spécialisé inconnu peut se faire par une simple recherche en table. C'est dire pourtant que les relations représentées dans le dictionnaire, et les relations dans le texte, sont en parfait accord — ce qui n'est pas forcément le cas. Le dictionnaire s'adresse au cas général, tandis que le texte engendre le cas spécifique. Le premier est statique ; le second est en mouvement. Et enfin, l'un n'est pas à jour par rapport à l'autre.

Alors, tout en admettant qu'une solution toute prête en forme de terme équivalent en langue cible soit très commode, il vaudrait mieux offrir également au traducteur la possibilité de fouiller, de vérifier le choix d'équivalent. N'oublions pas non plus que le traducteur a toujours besoin de compléter ses connaissances dans le domaine de référence, et qu'il lui faut des renseignements sur l'entourage des termes, c'est-à-dire sur leurs contextes, leur collocations.

Il ne fait aucun doute que la qualité des résultats obtenus à la suite d'une consultation dépend dans une large mesure de la stratégie de recherche. Ainsi, des stratégies empruntées à des dictionnaires imprimés donnent des résultats de qualité inférieure. En revanche, une recherche «libre» de mots-clés dans un dictionnaire informatisé peut réunir des fragments de connaissance dispersés dans les explications ou les citations qui accompagnent les termes. C'est un moyen de se former une image plus complète d'un objet ou d'un processus qui se présente dans le texte à traduire. On peut également choisir plusieurs mots qui représentent les concepts essentiels du domaine de traduction et les chercher dans les données associées aux termes vedettes afin de rassembler une collection de fiches dont les termes équivalents représentent le domaine restreint en langue cible. Une troisième stratégie consiste à chercher les définitions d'un dictionnaire monolingue en langue d'arrivée pour relever tous les termes qui se rapportent à une notion donnée (Fig. 1). Il faut noter aussi que dans le cas d'une base textuelle non structurée, une recherche de mots-clés fournit des exemples d'usage des termes.

MOTS-CLÉS	(langue source)
OBJET/PROCESSUS	(langue source)
CONCEPTS	(langue source)
TERMES	(langue cible)
CONCEPT	(langue cible)

Fig. 1

Il convient de rappeler ici un dilemme auquel on doit faire face dans un projet de dictionnaire électronique. Il s'agit d'un choix qui touche l'utilisateur et qui se résume en deux mots: 'libération' ou 'contrainte'. Une recherche 'libre' présuppose une formation dans les stratégies de recherche ainsi qu'une métareprésentation du contenu du dictionnaire. C'est une approche souple qui s'adapte aux méthodes de travail du traducteur. Mais elle a ses défauts: c'est lent et c'est compliqué par rapport à une consultation directe qui ne vise qu'un terme précis. Il se peut que les avantages qu'elle apporte ne soient pas suffisamment manifestes.

Quant à la «contrainte», elle est synonyme péjoratif de «conseil» ou de «guidage». Au sens positif, il peut s'agir simplement d'une interface qui aide l'utilisateur dans le système, par le moyen de profils qui s'adaptent à sa tâche et à son organisation conceptuelle des données terminologiques qui l'assiste dans la construction mentale d'un modèle du domaine de référence.

La classification conceptuelle d'un domaine spécialisé peut se faire selon plusieurs principes. Citons à titre d'exemple les principes de hiérarchie logique, d'opposition et de complémentarité, de traits, de champs, et de réseaux sémantiques. Certains domaines se prêtent également à une répartition en séries d'actions ou d'événements. On ne peut pas nier le caractère arbitraire et souvent énigmatique des catégories conceptuelles, ni le fait que la représentation soit souvent liée à une seule langue et à un seul domaine. Il y a un conflit aussi entre classification conceptuelle et classification syntactique, celle-ci correspondant à un besoin plus immédiat de la traduction. Toutefois, un dictionnaire basé sur une analyse conceptuelle a une qualité importante qui le fait recommander: c'est le fait

qu'il offre une représentation plus complète du domaine de référence. Cela devrait permettre au traducteur de combler des lacunes dans ses connaissances du domaine.

Permettez-moi de vous présenter, très brièvement, une organisation conceptuelle qui se distingue par sa capacité de satisfaire aux besoins de la traduction dont nous venons de parler. On mettra en relief les objets et les actions d'un domaine donné, ainsi que leurs qualités ou attributs. Les quatre éléments de base (Fig. 2) peuvent se combiner pour produire des syntagmes qui s'insèrent facilement dans le texte de destination. Les exemples de listes sorties du dictionnaire (Fig. 3 et 4) se rapportent au domaine de l'informatique, en particulier la terminologie de l'utilisation d'un tableur. Les listes comptent parmi leurs membres des mots et des groupes de mots d'une fréquence aussi bien élevée que basse.

objets
actions
attributs des objets
attributs des actions

objets + attributs des objets
attributs des objets + objets

objets + actions
actions + objets

actions + attributs des actions
attributs des actions + actions

Fig. 2

OBJETS

BLOC, CELLULE, COLONNE, DONNÉE, FENÊTRE,
FEUILLE, FORMAT, MESSAGE, MODE, POINTEUR, etc...

ACTIONS

ADDITIONNER, ANNULER, CADRER, COPIER, DÉPLACER,
FIXER, POINTER, PROTÉGER, SÉLECTIONNER, etc...

OBJETS + ATTRIBUTS DES OBJETS

CELLULE ACTIVE	FORMAT ACTIF
COURANTE	RÉPÉTÉ
FIXE	DE COLONNE
LIBRE	DE FEUILLE
PUITS	NUMÉRIQUE
SOURCE	NUMÉRIQUE STANDARD
VIDE	PAR DÉFAUT

ATTRIBUTS DES OBJETS + OBJETS

ACTIF(VE): CELLULE	COURANT(E): CELLULE
FORMAT	FEUILLE
FEUILLE	OBJET

OBJETS + ACTIONS

CELLULE: POINTER	POINTEUR: DÉPLACER
SAISIR	ANCRER
ÉDITER	RÉTRÉCIR
PROTÉGER	
DÉPROTÉGER	

FENÊTRE: AJUSTER

FAIRE ROULER

ACTIONS + OBJETS

POINTER: UN BLOC	SAISIR: UNE POMME
UNE CELLULE	UNE FORMULE
UNE RÉFÉRENCE	UN NOMBRE
	UN TEXTE
	UNE CELLULE
	UN TITRE

Fig. 3

ACTIONS + ATTRIBUTS DES ACTIONS

SAISIE:	MULTIPLE
CALCUL:	COMPLET
	TOTAL
	PARTIEL
RECHERCHE:	VERTICALE
RÉÉVALUATION:	HORIZONTALE
	VERTICALE
	NATURELLE
MODIFICATION:	GLOBALE

ATTRIBUTS DES ACTIONS + ACTIONS

AUTOMATIQUE:	CALCUL
VERTICAL(E):	RECHERCHE
	RÉÉVALUATION
GLOBAL(E):	MODIFICATION

etc.

Fig. 4

À l'heure actuelle, l'alimentation du dictionnaire résulte d'un dépouillement de textes qui se fait en partie à l'aide du système. Le système produit des concordances qui constituent des entrées potentielles pour le dictionnaire. Le choix d'objets est alors effectué par un spécialiste dans le domaine de référence.

À cet égard, je me permets de citer une remarque de Jean Perron relevée dans un article de *Terminogramme* : «*Le travail de dépouillement, s'il a pour objet premier de permettre au terminologue de relever les termes du domaine qu'il étudie, n'en constitue pas moins pour lui un moyen privilégié de se familiariser avec ce domaine. Aussi, nous garderons-nous bien de lui éviter tout travail de dépouillement ; il ne s'agit que de confier à l'ordinateur la part de textes que le terminologue n'a pas à lire pour parfaire sa connaissance du domaine*». (réf. 1)

Il me semble qu'il est important de souligner ce besoin de perfectionnement de la part du terminologue aussi bien que du traducteur. La connaissance du domaine, glanée d'une façon plus ou moins consciencieuse dans des textes sources, est transférée au dictionnaire et au texte cible d'une manière peu formelle. On sait très bien que la procédure de saisie des connaissances n'est pas facile à codifier ; il faut quand même essayer d'assurer son fonctionnement au moment du dépouillement de textes. Il faut ensuite que l'on puisse dégager ces mêmes connaissances des articles du dictionnaire.

Je viens de parler de stratégies de recherche en tant que moyens de retrouver des connaissances cachées dans des données pour ainsi dire «secondaires» du dictionnaire. Or, une bonne organisation conceptuelle du dictionnaire peut favoriser les procédés de rassemblement de connaissances appliqués par le traducteur. Dans l'exemple qui vient d'être présenté, ce rassemblement est facilité par la possibilité de voir d'un coup d'œil tous les objets et actions du domaine. Par ailleurs, chaque ensemble d'attributs constitue une description à partir de laquelle on peut faire des inférences sur plusieurs aspects des éléments du domaine — des aspects comme états, nombre, contenu, capacité de mouvement, disponibilité, statut, relations, disposition dans l'espace. Cette approche a pourtant des limites, surtout à cause de la polysémie des mots : ainsi, par exemple, dans le terme 'cellule active', 'active' peut signifier ou bien 'capable d'action' ou bien 'engagée dans l'action' ; c'est alors la définition qui fournira la précision essentielle à un raisonnement plus avancé. Voilà pourquoi dans ce système, la définition est toujours très importante ; autrement dit, la structure conceptuelle complète la définition, sans la remplacer.

Normalement, le traducteur fait des inférences improvisées, lorsqu'il cherche à compléter ses connaissances. Leur justesse est souvent validée lors de la révision du texte. Étant donné que la langue échappe à la logique, des erreurs sont inévitables. Dans la terminologie en question, 'cellule passive' (contraire de 'cellule active') n'existe pas ; 'cellule vide' exige un contraire plus spécifique que 'cellule pleine', et ainsi de suite.

Nous venons de parcourir deux méthodes d'accès aux dictionnaires informatisés pour la traduction de textes techniques : la recherche «libre» de mots-clés dans les explications et les exemples d'usage des termes, et la recherche guidée par un système qui présente des données terminologiques dans une optique d'unités conceptuelles appartenant à trois classes. Ce que ces deux méthodes ont de commun, c'est qu'elles offrent au traducteur la possibilité de trouver les unités terminologiques recherchées dans leur contexte linguistique usuel, et aussi un moyen de s'approprier des connaissances au domaine de traduction. Mais les résultats obtenus ne sont pas les mêmes, et plusieurs questions s'imposent : Faudrait-il choisir sa méthode d'organisation et d'accès au dictionnaire selon le domaine de traduction ? Quels seraient alors les caractéristiques du domaine qui permettraient de déterminer le choix de méthode. Nous espérons répondre à ces questions dans les mois à venir.

BIBLIOGRAPHIE

- BROWN, P. S., J. D. GOULD (1987) : «An experimental Study of People Creating Spreadsheets», *ACM Transactions on Office Information Systems* 5/3, 258-272.
- DE BEAUGRANDE, R. (1984) *Text Production : Toward a Science of Composition*, New Jersey, Ablex Publishing Corp.
- PERRON, J. (1988) : «Le dépouillement terminologique assisté par ordinateur», *Terminogramme*, n° 46, janvier 1988.
- PITRAT, J. (1985) : *Textes, ordinateurs et compréhension*, Paris, Eyrolles.
- SOWA, J. F. (1984) : *Conceptual Structures*, Reading MA, Addison-Wesley.
- VAN DIJK, T. A., W. KINTSCH (1983) : *Strategies of Discourse Comprehension*, London, Academic Press.
- VICKERY, B. C., A. VICKERY (1987) : *Information Science in Theory and Practice*, London, Butterworths.