

Traduction automatique et terminologie automatique

Jules Dansereau

Volume 23, numéro 2, juin 1978

URI : <https://id.erudit.org/iderudit/002456ar>

DOI : <https://doi.org/10.7202/002456ar>

[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

Dansereau, J. (1978). Traduction automatique et terminologie automatique. *Meta*, 23(2), 132–140. <https://doi.org/10.7202/002456ar>

Traduction automatique et terminologie automatique*

1. INTRODUCTION

Plusieurs personnes se demandent depuis quelques années pourquoi la traduction automatique ne s'est jamais servi des banques de terminologie qui existent. À première vue, il semble assez naturel de penser qu'une banque de terminologie bilingue soit utilisée aussi comme dictionnaire pour un système de traduction automatique, d'où une certaine économie pour les bailleurs de fonds.

Toutefois, on se rend vite compte que cette idée, bien que fort intéressante, ne résiste pas à un examen plus minutieux de la question. Dans l'état actuel des choses, il semble difficile de concevoir qu'un système de traduction automatique puisse consulter directement, par ordinateur, une banque de terminologie. Ceci ne veut pas dire qu'une telle banque ne peut pas être utile à ceux qui mettent au point un système de traduction automatique ; ces gens consulteront une banque comme n'importe quel autre client (le traducteur humain, par exemple). La terminologie est en effet un préalable à l'établissement d'un système de traduction automatique. La seule relation qui puisse exister entre une banque de terminologie et un système de traduction automatique en est un de « fournisseur » à « client », par l'intermédiaire de la personne qui rédige le dictionnaire de traduction automatique. Cette affirmation est peut-être un peu forte, mais c'est quand même elle qui guidera la suite de cet exposé.

Avant d'aller plus loin, il importe de préciser ceci : par *système de traduction automatique*, nous entendrons un système qui cherche à reconnaître, par des procédés linguistiques, la structure d'une phrase, de façon à parvenir à une traduction correcte de cette phrase. Nous excluons donc toutes les tentatives, utiles ou non, de traduction automatique mot à mot. Par *banque de terminologie*, nous entendrons uniquement son utilisation pour des fins de traduction. Toute autre utilisation ne sera pas discutée dans le présent exposé.

* Je désire signaler que les propos présentés dans cet exposé ne sont pas uniquement le fruit d'une recherche personnelle ; ils sont le résultat des recherches menées par le Groupe de recherches pour la traduction automatique de l'Université de Montréal (T.A.U.M.). Exposé présenté au 6^e Colloque international de terminologie organisé par la Régie de la langue française du Québec (Pointe-au-Pic, du 2 au 7 octobre 1977).

2. DIFFÉRENCE ENTRE LES DEUX OUTILS

La cause fondamentale de l'incompatibilité entre une banque de terminologie et un système de traduction automatique réside dans la nature même de ces deux outils. En réunissant toute la documentation terminologique dans un énorme fichier, dont la mise à jour est facile, une banque de terminologie cherche à rendre rapidement accessible à un traducteur toute l'information nécessaire à la traduction *d'une* unité de signification dans une phrase. Un système de traduction automatique cherche plutôt à traduire toute la phrase.

a) *Démarche du traducteur humain*

Pour bien illustrer cette différence, examinons la situation d'un traducteur humain qui consulte une banque de terminologie. Celui-ci a lu une phrase (ou un texte) et a repéré une ou plusieurs unités de signification qu'il ne connaît pas ou qu'il connaît mais dont il ignore l'équivalent dans la langue cible. Mais pour repérer ces unités de signification, le traducteur a dû, consciemment ou non, faire l'analyse de la phrase, c'est-à-dire établir les rapports existant entre les différents mots de cette phrase. C'est de cette façon seulement que le traducteur a pu identifier l'unité de signification qui lui pose des problèmes. Sa connaissance de la langue source lui a permis de faire l'analyse de la phrase ; et cette connaissance fonctionne très souvent de façon spontanée, voire inconsciente.

b) *Démarche d'un système de traduction automatique*

D'un autre côté, lorsqu'une phrase est donnée à un ordinateur pour être traduite, celui-ci doit d'abord faire l'analyse de la phrase, tout comme le traducteur humain, avant de tenter quelque traduction que ce soit. Le linguiste qui met au point un système de traduction automatique doit rendre sa connaissance de la langue le plus explicite possible, de manière à pouvoir la formaliser par la suite, et ainsi rendre l'ordinateur apte à établir les rapports entre les mots d'une phrase. Ceci implique donc deux opérations :

- 1) il faut d'abord donner pour chaque mot toutes les possibilités de regroupement de ce mot avec d'autres,
- 2) et ensuite donner des règles, dites syntaxiques, qui permettront de décider quel regroupement il faut choisir dans tel ou tel contexte.

Et c'est seulement après cette seconde opération, si tout a bien fonctionné, que l'ordinateur aura identifié les diverses unités de signification contenues dans une phrase.

La première opération, nous l'appelons « constitution du dictionnaire d'analyse de la langue source ». Dans un système de traduction automatique tel que défini précédemment, il y a généralement deux dictionnaires : le dictionnaire d'analyse et le dictionnaire de traduction. La division peut être uniquement conceptuelle, mais elle est souvent réalisée physiquement.

À l'aide d'exemples, nous allons voir qu'une banque de terminologie est incompatible, même théoriquement, avec un dictionnaire d'analyse d'un système de traduction automatique. Examinons la fiche de SUPPLY RESERVOIR, choisie parmi les fiches qui m'ont été fournies par la Banque de terminologie du Québec :

- lorsqu'un traducteur consulte cette fiche, il a déjà analysé la phrase et sait que SUPPLY et RESERVOIR forment ensemble un groupe nominal. Sur la fiche de SUPPLY RESERVOIR, la catégorie indiquée est effectivement NOM ;
- un dictionnaire d'analyse dans un système de traduction automatique ne peut absolument pas présumer qu'une occurrence particulière de la suite SUPPLY RESERVOIR est un nom, puisqu'on peut trouver :

(1) SUPPLY RESERVOIR WITH OIL

On devra donc indiquer au dictionnaire d'analyse que SUPPLY peut être un nom ou un verbe et l'analyseur décidera ce qui est réalisé dans le contexte donné.

Autre exemple : on peut imaginer une fiche dont l'en-tête serait TURN ON. Voilà une unité de signification très simple pour un traducteur, mais dont la reconnaissance peut poser quelques problèmes à un analyseur automatique, car il est possible de trouver :

(2) TURN ON THE RADIO

mais aussi :

(3) TURN THE RADIO ON.

Dans un dernier exemple, c'est l'analyseur qui identifiera l'unité TURN ON à partir du fait qu'au dictionnaire d'analyse, on aura indiqué que TURN peut co-occurrencer avec ON pour former une unité et que ON, dans ce contexte, ne peut être une préposition introduisant un groupe nominal. L'analyseur devra aussi décider de l'interprétation à donner à ON dans des exemples comme :

(4) THE SWITCH TURNS ON | THE RADIO

et

(4A) THE WHEEL TURNS | ON THE AXLE.

On peut déjà se faire une idée du genre d'information qui doit être mise dans le dictionnaire d'analyse d'un système de traduction automatique : il faut donner, pour chaque mot non seulement ses différentes catégories grammaticales possibles, mais, en plus, pour chaque catégorie, il faut donner les caractéristiques syntaxiques et sémantiques de ce mot, de façon à pouvoir bien identifier par la suite, au moyen de règles syntaxiques, la portée de ce mot dans la phrase. Ceci revient à définir, pour chaque catégorie grammaticale, un ensemble de sous-classes syntaxiques ou sémantiques. Par exemple, pour la préposition anglaise DURING, on indiquera qu'elle fait partie d'une sous-classe de prépositions qui se combinent avec les noms exprimant une opération

qui dure dans le temps. Le nom **GROUND** n'aura pas cette caractéristique alors que **TEST**, comme nom, appartient à cette sous-classe de noms. Cette sous-classification nous permettra d'éviter de traduire :

(5) **DURING THE GROUND TEST STAND CLEAR OF AIRCRAFT**
par :

DURANT LE SOL VÉRIFIER LE BANC À L'ÉCART DE L'AVION
Il faut en effet prendre **STAND** comme verbe et non **TEST**, et traduire par :

DURANT L'ESSAI AU SOL SE TENIR À L'ÉCART DE L'AVION.

Voici un exemple simplifié d'entrée au dictionnaire d'analyse :

CHECK

1) si c'est un nom,

— il sera abstrait et nom d'action ;

2) si c'est un verbe,

— il peut, dans sa forme en **-ING**, modifier un nom :

(6) **CHECKING PROCEDURE**

— il peut former une unité avec la particule **OUT** ;

— son objet direct peut être un groupe nominal ou une complétive introduite par **THAT** ou une subordonnée introduite par **IF** ;

— son objet indirect, s'il en a un, est un groupe nominal introduit par les prépositions **FOR** ou **AGAINST**.

Ce n'est pas le genre d'information que l'on trouve généralement sur une fiche terminologique. On pourrait, bien sûr, l'ajouter, en prévoyant sur chaque fiche un espace assez grand pour la contenir. Nous examinerons cette possibilité plus loin, mais on peut déjà dire que si certaines fiches étaient utilisées par un dictionnaire d'analyse automatique, il faudrait aussi en rejeter un grand nombre ; entre autres, toutes celles du type **TURN ON** et **SUPPLY RESERVOIR**. Ces fiches devraient alors être marquées de façon spéciale pour que le programme de consultation les refuse. De plus, dans les fiches acceptées, on n'utiliserait que la partie qui concerne la langue source.

Après la consultation du dictionnaire d'analyse, l'analyseur établit les rapports entre les mots et donne la structure de la phrase. Le dictionnaire de traduction est alors consulté.

Le mot **CHECK** nous servira aussi d'exemple de consultation du dictionnaire de traduction. L'exemple est simplifié et ne s'applique qu'au domaine de l'aéronautique. Les règles de traduction sont le résultat d'une étude préliminaire faite par un groupe de traducteurs dirigés par M. André Petit du Bureau des traductions du gouvernement fédéral :

1) le nom **CHECK** : traduction simple pour l'instant : **VÉRIFICATION** ;

2) le verbe CHECK :

— si la préposition introduisant l'objet indirect est AGAINST, traduire CHECK par COMPARER et AGAINST par À ;

(7) CHECK COCKPIT GAUGE AGAINST TEST GAUGE →
COMPARER LA JAUGE DE L'HABITACLE À LA JAUGE
D'ESSAI

— si CHECK est accompagné de la particule OUT et que le sujet est une pièce d'équipement, traduire par SATISFAIRE AUX ESSAIS si la phrase est active et par ÊTRE SOUMIS À UN ESSAI si elle est passive ;

(8) IF TRANSMITTER CHECKS OUT IN PREVIOUS
PROCEDURE →
SI LE TRANSMETTEUR SATISFAIT AUX ESSAIS...

(9) IF COMPLETE EQUIPMENT IS BEING CHECKED OUT →
SI L'ÉQUIPEMENT COMPLET EST SOUMIS À UN ESSAI...

— si le sujet est une pièce d'équipement et qu'il n'y a pas OUT, traduire par ARRÊTER ;

(10) THE VALVE CHECKS BACK PRESSURE →
LE CLAPET ARRÊTE LE REFOULEMENT

— dans les autres cas, on traduit CHECK par VÉRIFIER, mais il reste d'autres modifications à faire dans la phrase qui dépendent du verbe CHECK ;

(11) CHECK IF VALVE IS INSTALLED BACKWARDS →
VÉRIFIER QUE LE CLAPET N'EST PAS INSTALLÉ
À L'ENVERS

La négation doit être ajoutée en français si le complément de CHECK est un défaut (BACKWARDS) ;

(12) CHECK THE CABLE GUIDE FOR PROPER INSTALLATION →
VÉRIFIER L'INSTALLATION DU GUIDE DU CÂBLE

Ici, les deux compléments ont été inversés, l'adjectif PROPER et la préposition FOR ont disparu ;

(13) CHECK RESERVOIR FOR LEAKS →
VÉRIFIER QU'IL N'Y A PAS DE FUITE AU RÉSERVOIR

On a la même structure syntaxique que dans l'exemple précédent, mais l'objet de la vérification est ici un défaut. On peut difficilement accepter comme traduction VÉRIFIER LA FUITE DU RÉSERVOIR car ceci impliquerait qu'il est normal d'y trouver une fuite.

Pour chacune des traductions possibles d'un mot, il faut aussi indiquer au dictionnaire de traduction toutes les caractéristiques morphologiques (genre, type de conjugaison, irrégularité de certains accords) et les caractéristiques syntaxiques ou sémantiques propres à la langue cible.

Même si, à première vue, une fiche terminologique se rapproche plus du dictionnaire de traduction que du dictionnaire d'analyse, on constate encore une fois que l'information nécessaire au dictionnaire de traduction n'est pas du même genre que celle généralement accumulée sur une fiche terminologique. Mais, comme pour le dictionnaire d'analyse, on pourrait aussi ajouter sur la fiche ce type d'information.

3. CONSÉQUENCES D'UNE TENTATIVE D'ADAPTATION

Il serait intéressant de voir ce que cela pourrait impliquer pour une banque de terminologie d'insérer sur ses fiches toute l'information nécessaire aux deux dictionnaires d'un système de traduction automatique.

Tout d'abord, il faudra modifier les programmes de consultation de la banque de façon à ce que l'utilisateur usuel ne soit pas gêné par toute cette codification et que ce dernier ne reçoive que l'information grammaticale qu'on retrouve normalement sur une fiche terminologique. De la même façon, il faudra sélectionner l'information destinée au système de traduction automatique, car la fiche terminologique contient beaucoup d'information non pertinente pour un tel système.

J'ai démontré tantôt que seulement certaines fiches terminologiques pourront être utilisées au dictionnaire d'analyse et qu'il faudra donc sur chaque fiche, inclure une marque pour indiquer si elle est acceptable ou non pour ce dictionnaire. Et si la fiche est acceptable, on ne doit retenir, pour l'instant, que la partie langue source. Voilà donc une deuxième complication informatique pour la gestion de ces données.

Voici une troisième complication. En général, la fiche terminologique ne contient qu'une seule acception du mot. Si un mot a plusieurs acceptions, selon les domaines d'emploi par exemple, il y aura plusieurs fiches distinctes. Or, il se peut très bien que, peu importe l'acception, les caractéristiques de ce mot en langue source soient les mêmes ; par exemple, une VALVE, peu importe s'il s'agit d'un robinet, d'un clapet ou d'une soupape, est toujours un objet concret, un ensemble composé de plusieurs morceaux ayant une fonction de contrôle. L'analyseur, pour ce mot, n'aura probablement pas à tenir compte du domaine d'emploi. Donc, si on inclut l'information sur toutes les fiches pour ce mot, il faudra imaginer un moyen de ne la retenir qu'une seule fois. Et, si on code l'information sur une seule des fiches, il faudra décider et indiquer laquelle.

Supposons que ces problèmes soient résolus et poursuivons la discussion.

Il est possible qu'une fiche terminologique qui n'a pu être utilisée dans le dictionnaire d'analyse puisse l'être dans le dictionnaire de traduction. C'est peut-être le cas de fiches comme SUPPLY RESERVOIR, car l'unité de signification aura été reconnue par l'analyseur. Alors il faudra insérer une marque différente sur ces fiches, marque signifiant « non utilisable à l'analyse mais utilisable à la traduction ».

Théoriquement, il semble qu'une fiche qui a été utilisée dans le dictionnaire d'analyse pourrait l'être dans le dictionnaire de traduction. Puisque aucune expérience de ce genre n'a été tentée, il demeure difficile de l'affirmer avec certitude. Toutefois, il paraît évident qu'un bon nombre de fiches seront inutilisables par les deux dictionnaires. Par exemple, une fiche comme LEGAL ADVISER (donnée comme exemple par la Banque de terminologie du Québec) ne peut servir à l'analyse, car on ne peut présumer dans une phrase que LEGAL se rattache à ADVISER ; ainsi :

(14) CHECK IF IT IS A LEGAL ADVISER COMPANY.

Si cette phrase est prononcée par un détective, on peut présumer qu'il veut savoir si la compagnie est légale. La même fiche ne peut pas non plus servir au dictionnaire de traduction. En effet, si on considérait LEGAL ADVISER comme une seule entrée à ce dictionnaire, comme un idiome donc, il faudrait aussi considérer comme idiome :

LEGAL EXPERT
 LEGAL WRITER
 LEGAL PRACTITIONER
 LEGAL PROFESSION, etc.

Il semble donc plus judicieux de faire une entrée pour LEGAL, qui spécifiera les conditions qui font que LEGAL se traduit parfois par JURIDIQUE et parfois par LÉGAL et de traduire séparément ADVISER.

De toute façon, le système de traduction automatique devra posséder son propre dictionnaire pour les mots grammaticaux, comme les prépositions, les conjonctions et certains adverbes, et pour les mots généraux, c'est-à-dire ceux qu'on retrouve autant dans le langage courant que dans le langage scientifique et technique. Je ne crois pas qu'actuellement, une banque de terminologie ait une fiche pour la préposition IN ou le verbe TO HAVE, par exemple. Or, ces mots se retrouvent partout, dans tous les contextes et le choix d'une équivalence en langue cible pour ces mots est très souvent étroitement lié à l'environnement syntaxique et sémantique dans lequel ils se trouvent. Il est, par exemple, faux de dire que la traduction française de IN est DANS : le *Harrap's* donne une page complète, en deux colonnes, seulement pour ce mot, et je suis presque certain qu'il ne couvre pas tous les cas. Ces mots sont donc les plus difficiles à traduire pour un système de traduction automatique. Les autres mots ne posent souvent que très peu de problèmes, à partir du moment où on dispose de l'équivalent.

4. CONCLUSION

L'utilisation automatique d'une banque de terminologie par un système de traduction automatique exigerait que l'on fasse donc préalablement une sélection dans les fiches de la banque. Cette sélection serait basée sur des critères linguistiques trop souvent difficiles à préciser. De plus, la traduction automatique de bonne qualité ne peut pas encore prétendre à l'universalité. Pour l'instant,

si on veut faire de la traduction automatique qui soit rentable et dont les résultats soient utilisables, il faut resreindre le système à un seul champ d'activité humaine ou à quelques champs dans lesquels la syntaxe de la langue est très apparentée. Ceci implique que les critères de sélection des fiches risqueraient d'être très étroitement liés à un seul système de traduction automatique et n'auraient alors aucune valeur de généralité.

Cette sélection, difficile à définir et n'ayant de valeur que pour un seul système, entraînerait des complications informatiques importantes dans la gestion des fichiers. Tout cela pour aller chercher dans la banque l'information d'analyse et de traduction pour les mots qui le plus souvent ne présentent presque pas de problèmes pour un système de traduction automatique.

Devant tout cela, il me semble plus simple de continuer à considérer une banque de terminologie et un système de traduction automatique comme deux outils distincts, ayant chacun leur utilité.

Ceux qui élaborent un système de traduction automatique iront chercher, comme n'importe quel autre usager d'une banque, l'information dont ils ont besoin et incluront ces données dans leurs dictionnaires d'une manière entièrement compatible avec leur système. Ainsi les maîtres d'œuvres d'une banque de terminologie n'auront pas à se préoccuper d'un client spécial qui exigerait beaucoup d'eux, peut-être même au détriment des usagers pour qui une banque de terminologie est d'abord conçue.

Ceci n'exclut évidemment pas la collaboration essentielle entre un groupe de traduction automatique et les terminologues lorsqu'il s'agit de faire une recherche sur un domaine d'activité particulier. Comme je l'ai déjà dit, l'établissement de la terminologie reste un préalable à la traduction automatique. Peut-être qu'un jour nous verrons la possibilité d'aller plus loin, surtout avec le développement de l'informatique et de la linguistique contrastive. Mais les relations entre une banque de terminologie et un système de traduction automatique ne seront jamais vraiment différentes de celles qui existent entre une banque et un traducteur humain. La traduction automatique et la banque de terminologie sont deux outils complémentaires pour le traducteur : ce que le système de traduction automatique ne pourra pas traduire, les traducteurs s'en chargeront avec l'aide d'une banque de terminologie. L'utilisation parallèle de ces deux outils complémentaires aura un résultat commun : uniformiser la terminologie.

JULES DANSEREAU

BIBLIOGRAPHIE

- BAUDOT J., BOURBEAU L., CHEVALIER M., DANSEREAU J., ISABELLE P., LEHRBERGER J., POULIN G., STEWART G., *Projet AVIATION — rapport d'étape*, TAUM, Université de Montréal, mai 1977.
- DUBUC, Robert, *Description du système TERMIUM*, Banque de terminologie de l'Université de Montréal, édition révisée, janvier 1975.

- DUPLAIN, Jacques, *Guide de rédaction de la fiche de terminologie utilisée en entreprise*, rédigé sous la direction de Réjean Kirouac et Mireille Lacasse-Bernatchez, Banque de terminologie du Québec, Régie de la langue française, juillet 1977 (document provisoire).
- RONDEAU G., GRÉGOIRE J.-F., TESSIER P., « Banques de terminologies et linguistique computationnelle », *Meta*, vol. 22, n° 3, p. 184 à 191, septembre 1977.