

Banques de terminologies et linguistique computationnelle

Guy Rondeau, Jean-François Grégoire et Philippe Tessier

Volume 22, numéro 3, septembre 1977

URI : <https://id.erudit.org/iderudit/004466ar>

DOI : <https://doi.org/10.7202/004466ar>

[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

Rondeau, G., Grégoire, J.-F. & Tessier, P. (1977). Banques de terminologies et linguistique computationnelle. *Meta*, 22(3), 184–191.
<https://doi.org/10.7202/004466ar>

Banques de terminologies et linguistique computationnelle*

0.1 Les banques de terminologies sont une manifestation récente correspondant à des besoins sociaux nouveaux. La linguistique computationnelle devra résoudre, pour assurer leur développement, un certain nombre de problèmes. Nous examinerons ici brièvement l'un de ces problèmes : celui de savoir si dans leur conception actuelle les banques de terminologies répondent véritablement à leur objectif et, sinon, quelles modifications il faudrait prévoir à moyen et à long termes.

0.2 L'objectif général des banques de terminologies existantes est de fournir à une variété d'utilisateurs, quant aux données terminologiques qu'elles contiennent, des réponses précises, complètes et sans bavures, dans un minimum de temps et à un coût raisonnable. Ajoutons que ce dernier aspect, celui du coût, n'a pas encore fait véritablement l'objet des préoccupations des banques qui en sont encore au stade expérimental.

0.3 Nous utiliserons la terminologie du terminologue, telle qu'on la trouve par exemple dans les travaux de R. Dubuc, et dans laquelle un terme T est un composé comprenant une notion N et une appellation (ou étiquette) A. Dans une unité terminologique, une notion peut être complexe et une appellation peut être lexicalement simple ou multiple (par exemple, pour N = « pièce de métal servant à réunir deux pièces en les traversant par une fente commune » (Auger 1976) on a $A(L_1) = \text{« cutter »}$ et $A(L_2) = \text{« clavette transversale »} \neq \text{« clavette »}$.

1.0 Dans l'état actuel des choses, les banques de terminologies obéissent à des contraintes de deux ordres :

— le mode d'entrée des données dans la banque, et par conséquent l'accès aux données, est défini quant au contenu par des terminologues à partir d'une démarche terminologique qui n'est pas nécessairement celle d'autres usagers ;

* Le terme « terminologies » est employé ici au pluriel, contrairement à l'usage reçu, car ce que l'on met en banque, ce sont des termes ou des séries de termes (=terminologies). La terminologie elle-même est l'ensemble des principes directeurs et des règles méthodologiques qui permettent d'obtenir des terminologies. Les lignes qui suivent reprennent l'essentiel d'une communication présentée par les auteurs en juin dernier, dans le cadre de COLING 76.

— la structure des fichiers, et par conséquent le mode d'interrogation et d'extraction des données, est défini par des informaticiens à partir des devis préparés ou fortement influencés par des documentalistes (v.g. STAIRS).

Il en résulte un certain nombre de faits que nous examinerons après avoir considéré d'abord les deux distinctions suivantes :

Quant aux types d'utilisateurs de la banque ;

Quant aux types de réponses de la banque.

1.1 Quant aux types d'utilisateurs de la banque

	Objectif : établir les relations	Démarche	Type de rapport linguistique
Terminologie		$A_1 \longrightarrow N_1$ $A_2 \longrightarrow N_2$? ou $N \longrightarrow A(\text{néol.})$ (terminologique)	M ou C
Traducteur	$A(L_1) \longrightarrow A(L_2)$ pour $N(L_1) = N(L_2)$	 (du traducteur)	C
Autre	$A \longrightarrow N$ ou $N \longrightarrow A$	$A \longrightarrow N?$ (lexicologique)	M ou C

1.1.1 Pour le terminologue

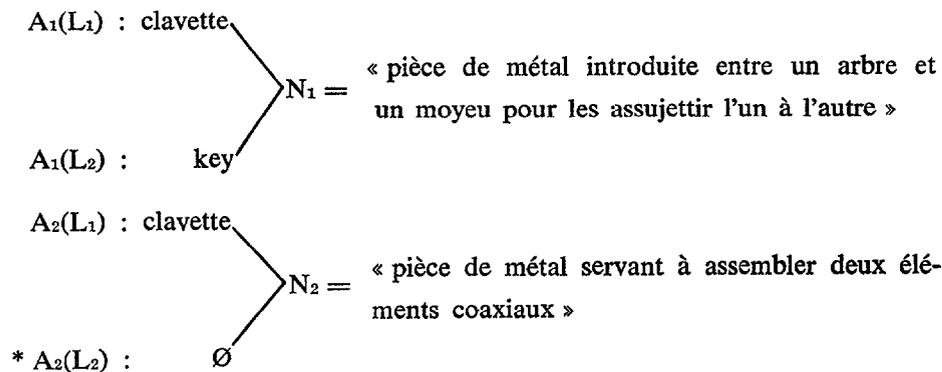
L'objectif consiste à établir des relations entre une notion N et une ou plusieurs appellations, de façon à découvrir si, à l'intérieur d'une langue donnée, il existe pour telle notion plusieurs appellations (et *vice versa* dans certains cas)

ou, de façon plus générale, quelles sont toutes les appellations qui, dans les langues naturelles, correspondent à telle notion.

Cela pose évidemment le problème de la non-correspondance des notions d'une langue à l'autre, mais nous n'aborderons pas ce problème ici.

Essentiellement, la démarche du terminologue est la suivante : étant donné d'une part l'appellation A_1 correspondant à la notion N_1 et d'autre part l'appellation A_2 , correspondant à la notion N_2 , y a-t-il correspondance entre N_1 et N_2 ? Si oui, alors A_1 et A_2 sont des synonymes au sens terminologique, comme par exemple logiciel, programmerie, software, etc. Sinon, A_1 et A_2 sont les étiquettes de deux unités terminologiques distinctes.

Cette démarche s'applique à l'intérieur d'une même langue ou sur un couple de langues. Pour revenir à notre exemple de clavette, considérons $N(L_1) = N(L_2)$:



Le cheminement de la démarche du terminologue est du type appellation \rightarrow notion \rightarrow appellation, sauf dans le cas de la création néologique, où il est du type notion \rightarrow appellation.

1.1.2 Pour le traducteur

L'objectif du traducteur est essentiellement de découvrir quelle est l'appellation de L_2 qui correspond à une appellation donnée de L_1 . La notion ne constitue pas l'objet propre de son travail, comme c'est le cas pour le terminologue ; la notion sert de fil conducteur et le traducteur a atteint son objectif dès qu'il a trouvé une étiquette correspondant de façon satisfaisante au terme donné dans la langue de départ. Contrairement au terminologue, le traducteur dans l'exercice de ses fonctions ne cherchera pas systématiquement à découvrir toutes les appellations qui, en langue d'arrivée, correspondent à la notion donnée dans la langue de départ.

La démarche du traducteur est caractérisée par le fait qu'elle s'exerce toujours sur un couple de langues (alors que le terminologue peut travailler en mode monolingue). Le cheminement de la démarche du traducteur est du type

appellation de $(L_1) \rightarrow A(L_2)$, soit par passage direct, soit par vérification dans les deux sens de la correspondance appellation-notion.

1.1.3 Autre

Quant à l'usager qui n'est ni terminologue ni traducteur, étant donné telle appellation, son objectif est soit d'obtenir la notion qui lui correspond, soit encore, étant donné telle notion, d'obtenir l'appellation correspondante.

Sa démarche est proprement d'ordre lexicologique, le terme étant entendu aussi bien dans le sens de la fabrication que de la consultation des dictionnaires. On peut citer comme exemples le lecteur d'un texte technique qui rencontre un terme inconnu, ou le rédacteur d'un manuel technique qui cherche le terme correspondant à la notion qu'il désire exprimer. La néologie de première génération est également du même type, car le concepteur d'une idée ou d'un procédé ou d'un produit doit nécessairement suivre le cheminement notion-appellation.

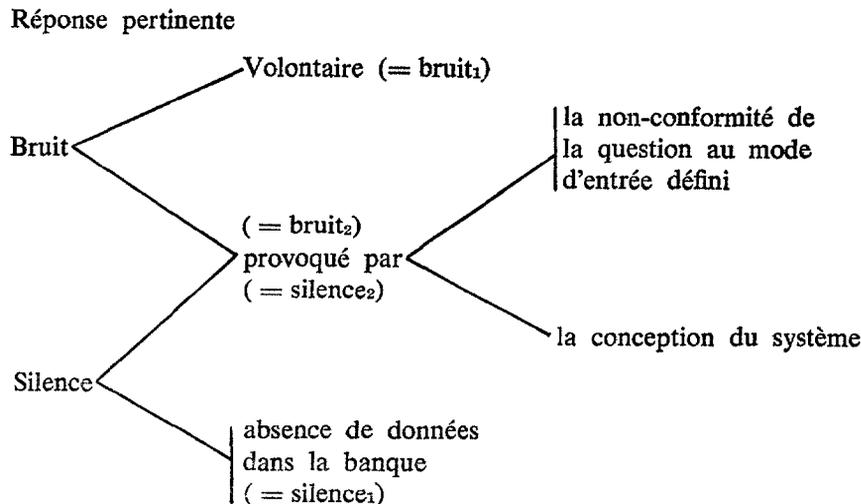
Contrairement au traducteur et comme le terminologue, l'usager de type « autre » peut travailler en mode monolingue ou bilingue.

1.2 Quant aux types de réponses de la banque

1.2.0 On en distingue trois : les réponses pertinentes, le bruit et le silence.

Par réponse pertinente, on entend celle qui correspond exactement à la question, sans fournir d'informations inutiles et sans laisser dans l'ombre un seul renseignement utile. De façon idéale par conséquent, pour fournir une réponse totalement pertinente, une banque de terminologies devrait pouvoir analyser non seulement la question, mais également celui qui la pose, afin de pouvoir tailler sur mesure des réponses en fonction du profil d'intérêts du questionneur.

1.2.1 Le tableau suivant résume les types de réponse :



1.2.2 Exemples

- Bruit₁ Q = open top car
 R = gondola
 hopper car (parfois open top car)
 etc.
- Bruit₂ Q = car
 R = flat car
 passenger car
 elevator car (ascenseur)
 fish car (app. de pêche)
 box car
 carload
 etc.
- Silence₂ Q = fishing car
 dans la banque : fish car ;
 R = Ø, si l'on n'a pas de système de troncature
- Silence₁ Q = Northern Car (gr. d'étoiles de la Gr. Ourse)
 R = Ø

Exemples de bruit causé par la non-conformité de la question au mode d'entrée défini :

Premier cas

La question n'a pas été précisée au moyen de paramètres, v.g. domaine(s).

Ex. : Q = car
 R = avalanche de fiches

Deuxième cas

La question est un problème de traduction, alors que les données sont classées en termes de problèmes terminologiques.

Ex. : Q = thread
 R = enfiler
 rentrer
 monter

Exemples de bruit causé par la conception du système :

Premier cas

Le système est conçu en fonction de l'automatique documentaire ; la première R est donc en général du type « bilan ».

Deuxième cas

Troncature non sélective

Q = car

R = cardboard, carbon paper,
cartoon, side-car, scaroid,
cart, carol, carnivore
cartridge, discarnate, carte-de-visite,
scarificator, etc.

2.0 Il en résulte les faits suivants :

2.1 Si l'interrogateur est un terminologue, la majorité sinon toutes les réponses qu'il obtient sont du type : pertinence, bruit₁ ou silence₁.

Si l'interrogateur n'est pas un terminologue, il a de fortes chances d'obtenir des réponses du type :

bruit₂, silence₁, silence₂, pertinence

En ignorant bruit₁ et silence₁, qui n'ont pas de rapport avec notre propos, on constate que bien qu'une banque de terminologies ait pour objectif de fournir des termes à une masse d'utilisateurs non terminologues, elle est conçue de façon à favoriser les terminologues au détriment des autres usagers. En d'autres termes :

Première constatation

Dans l'état actuel des choses, les banques de terminologies doivent compter sur un intermédiaire-terminologue pour assurer un maximum d'efficacité à leurs usagers.

2.2 Si l'interrogateur est un traducteur, sa démarche est caractérisée par le fait qu'il travaille toujours sur un couple de langues et qu'il tente en premier lieu d'obtenir l'appellation (L₂) correspondant à l'appellation (L₁) ; ce n'est qu'en cas d'insuccès qu'il entreprendra le cheminement : appellation (L₁) → notion → appellation (L₂). Or, dans leur forme actuelle, les banques ont été conçues par des terminologues, — essentiellement des « étiqueteurs de notions » — pour servir, entre autres usagers, des traducteurs, — essentiellement des « chercheurs de l'appellation (L₂) qui correspond à l'appellation (L₁) ».

Ainsi, dans l'exemple suivant (Dubuc 1976) :

to thread a needle (enfiler une aiguille)
to thread a heddle (rentre une lisse)
to thread a loom (monter un métier)

le problème de traduction porte sur *thread* parce qu'on a en L₂ trois significés pour un signifiant unique, mais il ne peut être résolu que grâce à la solution du problème terminologique, qui porte sur l'ensemble V + O. Or, il n'est pas sûr que le terminologue chargé d'alimenter la banque accorde la même impor-

tance aux problèmes terminologiques sous-jacents aux problèmes de traduction qu'aux problèmes purement terminologiques. En d'autres termes :

Deuxième constatation

Dans l'état actuel des choses, les banques de terminologies n'assurent pas aux traducteurs des services correspondant au maximum de leur potentiel.

2.3 Si l'interrogateur n'est ni terminologue ni traducteur (il peut être rédacteur de manuel technique, aide-terminologue qui fait de la recherche thématique, professeur de langue, fonctionnaire qui utilise le français comme nouvelle langue de travail, etc.), sa démarche est caractérisée par le fait que, comme le terminologue, il cherche l'appellation correspondant à telle notion sans pouvoir compter sur les connaissances ou l'expérience pratique de la terminologie que possèdent le terminologue et le traducteur ; ou encore, s'il s'agit, par exemple, du lecteur d'un document qui rencontre un mot nouveau, il cherche la notion correspondant à telle appellation, un peu comme s'il consultait un dictionnaire. Dans les deux cas, la démarche est encore caractérisée par le fait que l'interrogateur devrait normalement pouvoir travailler à l'intérieur d'une seule langue mais que la solution la plus simple consiste souvent à passer par le pont d'une deuxième langue plutôt que par celui de la notion. En d'autres termes :

Troisième constatation

Dans l'état actuel des choses, les banques de terminologies ne peuvent rendre que des services marginaux, et encore le plus souvent par le biais d'une langue naturelle-pivot, aux usagers qui ne sont ni terminologues ni traducteurs et qui travaillent ou désirent travailler à l'intérieur d'une même langue.

3.0 Conclusions

3.1 À la lumière de ce qui précède, on peut poser comme objectifs à long terme, pour les banques de terminologies :

3.1.1 L'élimination progressive du terminologue-intermédiaire entre les banques et leurs usagers non terminologues.

3.1.2 La capacité d'assurer aux usagers non terminologues des services au moins équivalents en qualité à ceux que reçoivent à l'heure actuelle les terminologues.

3.1.3 La capacité de travailler en mode unilingue.

3.2 Pour atteindre ces objectifs, il faudra réaliser à moyen terme un certain nombre de sous-objectifs, parmi lesquels on peut citer :

3.2.1 La conception d'une structure de fichiers qui permettent des interrogations séquentielles à niveaux multiples avec cheminement conversationnel (v.g. balayage des termes en vedettes, puis des définitions et contextes, puis d'autres champs comme celui des sources documentaires).

3.2.2 L'établissement de nœuds notionnels définis comme les ensembles de relations sémantiques qui unissent des signifiants (pour une ou plusieurs langues) autour d'un signifié. (Ce sous-objectif à lui seul est de taille et il demandera des recherches approfondies, mais il pourrait contribuer à la solution d'au moins deux problèmes : celui des fiches multilingues et celui que l'on trouve noté en 3.1.3).

3.2.3 La conception et la réalisation d'un système dans lequel coût et temps d'interrogation soient indépendants du volume de la banque. Or, avec l'approche du type « documentaire » la progression du rapport est exponentielle.

3.2.4 La conception et la réalisation d'un système dans lequel coût et temps de réponse soient en rapport inverse de la qualité de pose de la question.

Le problème étant ainsi posé, c'est de la linguistique computationnelle que devra venir la solution. Car la solution ne pourra être que le produit des efforts conjugués de linguistes (terminologues, traducteurs, etc.) et d'informaticiens.

G. RONDEAU
J.-F. GRÉGOIRE et
P. TESSIER

BIBLIOGRAPHIE

- AUGER, Pierre et collaborateurs, R.L.F., « Mémento de terminologie » in *Actes du Colloque canadien sur les fondements d'une méthodologie générale de la recherche et de la normalisation en terminologie et en documentation*, G. Rondeau, réd., D.G.T.D., Bureau des traductions, Secrétariat d'Etat, Ottawa, 1976, p. 85.
- DUBUC, Robert, « Définition et objectifs de la recherche terminologique » in *Actes du Colloque canadien sur les fondements d'une méthodologie générale de la recherche et de la normalisation en terminologie et en documentation*, G. Rondeau, réd., D.G.T.D., Bureau des traductions, Secrétariat d'Etat, Ottawa, 1976, p. 17.