

Fertilisation terminologique ou insémination terminologique artificielle ?

Christine Portelance

Volume 32, numéro 3, septembre 1987

La fertilisation terminologique dans les langues romanes

URI : <https://id.erudit.org/iderudit/003860ar>

DOI : <https://doi.org/10.7202/003860ar>

[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

Portelance, C. (1987). Fertilisation terminologique ou insémination terminologique artificielle ? *Meta*, 32(3), 356–360.
<https://doi.org/10.7202/003860ar>

FERTILISATION TERMINOLOGIQUE OU INSÉMINATION TERMINOLOGIQUE ARTIFICIELLE ?

CHRISTINE PORTELANCE
Université de Montréal, Montréal, Canada

Parler de fertilisation terminologique, c'est, à notre avis, aborder le problème des lacunes terminologiques, de la dépendance terminologique d'une langue par rapport à une autre, d'un point de vue positif.

Comme on le sait, il y a moins de 10% des publications technoscientifiques en français et plus de 55% en anglais : aussi, le dépouillement de textes, en vue de la confection d'un lexique français-anglais/anglais-français, est inégal. La nomenclature, dans ce cas, aura tendance à être déterminée en fonction de la langue dominante. Les trous à combler en français risquent fort d'être nombreux : on aura alors recours à la néologie.

On peut poser le problème des lacunes terminologiques d'une autre manière : le traducteur, la traductrice, aux prises avec un problème terminologique, ne pourra compter que sur elle-même devant la mutité des ouvrages terminographiques. Une solution ne sera envisagée qu'après une analyse consciente ou semi-consciente de la situation de communication en jeu ou plus exactement du contexte situationnel. Plusieurs scénarios sont alors possibles : la composition, l'emprunt et même la modulation qui peut servir de voie d'évitement. La décision sera individuelle, c'est-à-dire qu'elle sera prise par une personne, pour un texte donné ; ce qui ne veut pas dire que ce type de décision soit nécessairement sans effet sur la langue. Par exemple, le suffixe *isme* est passé dans la langue anglaise via l'emprunt de mots comme *cartésianisme* que les traductions anglaises du 17^e siècle avaient conservé à la lettre.

Il peut donc y avoir une différence entre le traitement d'une notion à l'intérieur d'une nomenclature et le traitement de cette même notion en situation discursive.

Nous aborderons la fertilisation terminologique du point de vue des nomenclatures. On s'entend généralement sur le caractère artificiel des terminologies spécialisées, caractère dû justement à l'absence de contextes situationnels. Les nomenclatures sont constituées à plus de 80% de substantifs et elles constituent une liste de dénominations correspondant aux notions ou objets de connaissance d'un domaine d'activité particulier.

Si on étudie une liste de notions, on s'apercevra de l'existence de rapports entre les notions : c'est ainsi que l'on peut parler de notion générique et notion spécifique, notion subordonnée ou coordonnée. Malheureusement, il semble impossible de concevoir, dans tous les domaines, des nomenclatures de type hiérarchique telles que les nomenclatures de botanique ou de chimie du 19^e siècle. Si un système de notions ne peut être envisagé sous l'angle de l'ordonnement des notions, comme on le fait, par exemple, dans les systèmes de classification documentaires, les rapports entre les notions n'en existent pas moins. Nous aimerions maintenant aborder ces rapports d'un point de vue formel, sous l'angle des dénominations elles-mêmes.

Examinons premièrement quelques statistiques lexicales tirées de différents glossaires, statistiques obtenues dans le cadre d'un projet de terminotique visant à l'élaboration d'une minibanque terminologique :

374 termes formés de 776 mots dont 424 mots différents

1 128 termes formés de 2 875 mots dont 1 010 mots différents

2 129 termes formés de 5 246 mots dont 1 667 mots différents

3 291 termes formés de 9 243 mots dont 2 232 mots différents

Dans un même lexique, si on examine le pourcentage des mots différents par rapport à l'ensemble des mots et ce, par tranche de termes, on s'apercevra que le taux de répétition des mots augmente considérablement avec le nombre de termes :

500 termes : 48,7% de mots différents

1 005 termes : 46,5% de mots différents

2 000 termes : 39,7% de mots différents

4 010 termes : 30,4% de mots différents

8 024 termes : 23,2% de mots différents

10 372 termes : 21,0% de mots différents

Le rapport entre l'étendue textuelle et le vocabulaire peut être comparé à celui du texte continu. Un texte de 15 000 mots contient environ 10% de mots différents, une liste de termes de même taille, 20%. À première vue, ces résultats paraissent surprenants, ils sont dus en fait à la définition de *mot différent* qui n'est pas la même dans les deux cas : pour le texte, il est question de vocables lemmatisés, pour la nomenclature, de graphies différentes. Dans le cas d'une définition commune, le vocabulaire des termes se saturerait plus rapidement que celui du texte, mais il y a quand même une similarité des phénomènes (Baudot 1986). Toutefois, ce n'est pas la comparaison qui nous intéresse, mais tout simplement l'existence même du phénomène. En effet, ces statistiques lexicales sur la saturation du vocabulaire illustrent concrètement l'hypothèse de Benveniste

les termes de base à l'aide desquels se constitue une nomenclature tendent à devenir explicites et à constituer à leur tour des combinaisons explicites au moyen d'unités lexicales identifiables séparément et organisées selon des modèles syntaxiques.

Une des caractéristiques des nomenclatures techniques est donc le potentiel combinatoire des unités qui les composent, et une façon d'aborder l'étude de ce potentiel réside en une description morphologique des nomenclatures, ce que nous tentons de faire par l'étude des matrices terminologiques.

Nous avons laissé de côté les monoterme, la dérivation et la composition gigogne, pour nous intéresser plutôt aux multitermes et en particulier aux composés à plus de deux substantifs.

Dans un corpus formé par les termes français de trois lexiques, dans les domaines de l'informatique, les télécommunications et le transport (5 129 termes) près de 37% des termes sont des composés à plus d'un substantif et la répartition est la suivante :

binominaux : 1 534

trinominaux : 299

quadrinominaux : 38

quinominaux : 7

Total : 1 878

Les matrices terminogéniques des binominaux

La matrice théorique globale est la suivante :

[N + (adj) + (prép) + (art) + N + (adj) + (adj) + adj]]

La matrice [N + prép + N] est de loin la plus prolifique (1 085 occurrences) et semble le modèle binominal par excellence. Comme on le sait, elle existe sans aucun article, mais aussi avec un article défini, indéfini, ou contracté. Notre corpus nous donne 13 autres matrices, de [N + N] *mémoire tampon* à [N + prép + N + adj + adj] *méthode d'accès séquentiel indexé hiérarchique*. Ces 13 matrices engendrent 449 occurrences.

Les déterminations sont de trois types : (+ adj) (+ prép + N), et (+N). Les déterminations adjectives sont courantes dans la langue générale, mais, dans les listes de termes, la détermination (prép + N) est la plus fréquente. Cette situation n'est qu'une plus grande utilisation d'un mécanisme bien présent dans la langue générale puisque de nombreuses expressions de ce type sont perçues comme des locutions adjectives et considérées comme telles dans les ouvrages lexicographiques : *à chaud, à froid, à cheval, de pointe*, etc. Grévisse lui-même considère que les compléments déterminatifs sont des locutions adjectives en puissance.

La détermination (+ N) est moins fréquente, mais sa présence est significative. Certains ne voient là qu'une influence de l'anglais ; ce réflexe s'explique puisque la matrice terminogénique (N + N) est la plus prolifique en anglais. Nous croyons que ce « calque » n'est possible que parce que ce modèle existe effectivement en français, et cela depuis le 13^e siècle. Il a d'ailleurs donné de nombreux composés dans la langue générale : *hôtel-Dieu, chou-fleur, betterave, timbre-poste*, etc. Ce modèle est intéressant puisqu'il est susceptible d'engendrer des séries :

mémoire tampon, produit tampon, fluide tampon, etc.

protection mémoire, impression mémoire, image mémoire, etc.

langage machine, passage machine, etc.

On peut même observer une concurrence entre des termes désignant la même notion, comme par exemple : *adresse de réseau* et *adresse réseau*, trouvés dans deux lexiques différents. Ceci laisse supposer l'existence de certaines conditions permettant la suppression de la préposition. Les prépositions les plus fréquentes, à mais surtout *de*, peuvent exprimer presque tous les rapports de subordination : destination, origine, partie d'un tout, manière, forme, etc. ; aussi, on peut penser que leur présence dans certains cas n'est que syntaxique et en conformité avec le modèle (N + prép + N). En considérant le vocabulaire terminologique d'un point de vue évolutif, on peut même envisager le composé (N + N) comme une étape préliminaire à la composition gigogne.

Darmesteter, il y a déjà longtemps, a émis l'hypothèse, à propos des composés (N + N), d'une fusion de sens par ellipse avec divers degrés de complication. Il entend par ellipse une absorption mutuelle de sèmes, dont certains sont privilégiés au détriment d'autres sèmes. Nous croyons cette thèse toujours d'actualité, mais il semble que personne n'ait été jusqu'à ce jour capable de démontrer entièrement le mécanisme de ce phénomène sémantique. Nous estimons que la productivité du vocabulaire terminologique est, entre autres, tributaire de ce phénomène et que pour cette raison il mériterait des études plus approfondies.

Nous aimerions maintenant aborder la question des composés lourds, composés à plus de deux substantifs. La dynamique d'une nomenclature dépend, comme dans les systèmes de classification documentaires, de sa capacité d'insertion de nouveaux termes, et la présence des composés lourds dans les nomenclatures peut en quelque sorte être considérée comme un épiphénomène. Voici la matrice théorique globale des trinominaux :

[(adj) + N + (adj) + (prép) + (art) + (X) + (conj) + N + (adj) + (conj) + (prép) + (art) + (adj) + (X) + N + (adj) + (X) + (X)]

dont les seuls éléments stables sont les substantifs. Cette matrice compte 17 constituants et engendre dans notre corpus 48 matrices différentes qui à leur tour produisent 344 ter-

mes. Il existe également, en très petit nombre (45 dans notre corpus), des quadriminaux et des quinominaux.

Ces composés lourds permettent l'existence de séries de syntagmes à noyau commun, comme par exemple :

- *méthode d'accès direct hiérarchique*
- *méthode d'accès avec file d'attente*
- *méthode d'accès avec file d'attente en télétraitement*
- *méthode d'accès séquentiel hiérarchique*
- *méthode d'accès séquentiel indexé hiérarchique*

En théorie, la longueur des composés lourds est infinie. C'est pourquoi nous croyons qu'à l'intérieur du cadre d'une nomenclature, c'est la capacité de la mémoire de compréhension plutôt que celle de la mémoire de rétention qui agit comme facteur de délimitation des dénominations. Avec une telle capacité générative, toute notion élaborée dans une discipline pourra être nommée et distinguée des autres notions par le jeu des références et des oppositions selon le même principe qui permet à un système linguistique d'engendrer un nombre infini de phrases.

On peut imaginer le phénomène des composés lourds comme nouveau. Pourtant Louis Renou dans un article sur l'évolution des composés nominaux en sanscrit constate que :

... les énoncés se font normalement par voie de composés, non seulement les énumérations de mots (si longues soient-elles), mais l'articulation même des règles. La seule limite est la clarté de l'énoncé...

Et un peu plus loin :

La prose classique proprement dite, tant technique (...) que littéraire, fait un usage illimité des composés. Mais ce qui, dans les textes littéraires, est surtout un ornement (...), est en revanche, dans la prose technique, un élément pour ainsi dire organique du langage.

C'est d'ailleurs à Renou que nous avons emprunté l'expression *composés lourds*, mais en lui donnant notre propre définition.

Plus près de nous, un dictionnaire technique de 1906 recensait de nombreux syntagmes dont *appareil enregistreur à encre*, *arbre de commande de transmission*, *lampe à arc en vase clos*. La longueur des composés importe peu si l'énoncé est clair, et il ne faut pas oublier que l'usage répandu d'une expression a souvent pour effet l'abrègement de celle-ci. On pense par exemple à *cinématographe* et *vélocipède*. À ce titre, il serait intéressant d'avoir des données sur les stratégies terminologiques de discours qui pourraient être obtenues par l'étude du comportement des termes en situation discursive. Nous comptons entreprendre prochainement une telle étude, mais sur une très petite échelle, en espérant un jour parvenir à mettre sur pied une banque de données textuelles.

Ce que nous avons tenté d'illustrer, tant du point de vue quantitatif que du point de vue qualitatif, c'est la dynamique du vocabulaire technoscientifique en opposition avec l'image traditionnelle des nomenclatures que l'on a souvent tendance à considérer comme une liste d'éléments figés. La statique apparente des nomenclatures est due aux caractères artificiels inhérents à l'entreprise terminologique elle-même. Or, si la mise en place de nomenclatures bien établies ne peut remplacer les différentes stratégies du discours, le potentiel combinatoire des éléments constitutifs des nomenclatures peut alimenter de telles stratégies. Nous croyons que pour obtenir une fertilisation terminologique, il faut utiliser la dynamique des nomenclatures comme un mécanisme d'insémination terminologique artificielle.

Nous estimons également que la compréhension de la dynamique des nomenclatures reliée à celle des stratégies terminologiques de discours pourrait contribuer à donner un nouveau souffle à la terminographie. En effet, des acquis théoriques dans ce domaine faciliteraient sûrement les échanges entre les différentes communautés linguistiques et permettraient une plus grande cohésion des pratiques terminologiques.

RÉFÉRENCES BIBLIOGRAPHIQUES

BAUDOT, J. (1986) : *Rapport final : BATEM*, publication interne.

BENVENISTE, E. (1974) : *Problèmes de linguistique générale*, Paris, Gallimard.

RENOU, L. (1952) : « Sur l'évolution des composés nominaux en sanskrit », *Bulletin de la Société linguistique de Paris*, vol. 52, pp. 96-116.