Meta
Journal des traducteurs
Translators' Journal

META

L’emprunt sémantique dans la terminologie de l’informatique

John Humbley

Volume 32, Number 3, septembre 1987

La fertilisation terminologique dans les langues romanes

URI: https://id.erudit.org/iderudit/003158ar
DOI: https://doi.org/10.7202/003158ar

See table of contents

Publisher(s)

Les Presses de l'Université de Montréal

ISSN

0026-0452 (print)
1492-1421 (digital)

Explore this journal

Cite this article

Humbley, J. (1987). L’emprunt sémantique dans la terminologie de l’informatique. *Meta*, *32*(3), 321–325. https://doi.org/10.7202/003158ar

L'EMPRUNT SÉMANTIQUE DANS LA TERMINOLOGIE DE L'INFORMATIQUE

JOHN. HUMBLEY
Université de Nancy II, France

Lorsqu'on parle de la terminologie de l'informatique dans le contexte français, c'est généralement pour insister soit sur le flot d'emprunts : *slot, scratch, rack, batch, ...* soit sur les créations lexicales, réussies comme **ordinateur, logiciel** ou **progiciel**, ou passablement délaissées comme **listage, ludiciel** ou **bogue**. Ceci porterait à croire que le français de l'informatique consiste en une tension constante entre une importation massive de mots anglais et une francisation radicale.

Si on veut se convaincre du contraire, il suffirait d'examiner la composition des éléments contenus dans les listes de termes obligatoires en informatique (décrets du 22-12-1981 et 30-12-1983) pour constater que la plupart d'entre eux ne sont pas des créations intégrales, mais plutôt des aménagements sémantiques de matériel existant, aménagements qui ne bousculent pas les habitudes linguistiques des utilisateurs, ce qui explique en partie le succès de cette initiative. Cette impression est confirmée par la lecture de la presse de vulgarisation informatique, où créations ou emprunts paraissent comme des procédés presque marginaux, si on prend la peine de reconnaître les très nombreux emprunts sémantiques, qui constituent la base du vocabulaire. Une lecture attentive de la presse spécialisée nous convainc que l'emprunt sémantique d'un point de vue quantitatif enrichit davantage la terminologie que l'emprunt brut ou la création, et nous avançons quelques chiffres pour étayer cette affirmation, chiffres qui suggèrent les volumes relatifs des différents procédés d'adaptation. Il convient ensuite d'examiner les emprunts sémantiques et de se poser la question pour savoir si le type d'emprunt (brut ou sémantique, avec ou sans une mesure de création) joue un rôle dans son intégration définitive.

Nous avons deux moyens pour constituer des statistiques lexicales d'un vocabulaire donné : soit directement lorsque nous recourons à un corpus primaire, soit indirectement en se fiant aux dictionnaires spécialisés. Ici nous avons retenu les deux démarches, le caractère concret des dépouillements directs compensant et corrigeant le parti pris des lexicographes, tandis que le caractère exhaustif du dictionnaire complète l'aspect aléatoire de l'enquête sur le terrain.

Si nous prenons le dictionnaire volontariste de Pierre Morvan (*Dictionnaire de l'informatique*, Larousse, 6e édition, 1981), il apparaît que le lexicographe fait la part belle aux emprunts sémantiques et aux créations : les mots anglais identifiables comme tels se résument à quelques unités, au point où il lui faut un lexique anglais-français destiné sans doute aux lecteurs de textes spécialisés ... français ! L'enquête sur le terrain — un dépouillement d'une douzaine de magazines d'informatique de 1986 permet de redresser le tir. Ce sondage fait état de quelque 126 emprunts bruts, chiffre considérable, mais moins élevé que les 260 emprunts bruts de la presse équivalente allemande de la même époque. Il serait en outre à moduler compte tenu du nombre élevé d'hapax, mots étrangers cités plutôt qu'employés, ainsi que des mots qui ne relèvent pas de la terminologie de l'informatique.

Du point de vue de la création, nous nous heurtons au problème de la définition de ce terme, car il existe un continuum qui va de la réalisation de dérivés ou de composés virtuellement présents en français, mais non réalisés (ou peu réalisés) à une créativité plus large. Quelle que soit la définition cependant, la création lexicale ne dépasse pas une quarantaine d'éléments du *Dictionnaire Larousse*[1].

Ce sont donc les emprunts sémantiques qui constituent l'essentiel de son répertoire : quelque 800 termes, que nous tâcherons de classer dans la deuxième partie de cet exposé. Il est évident que tous ces termes ne sont pas intégrés au même point dans l'usage : *directory* s'emploie fréquemment plutôt que **descripteur de fichier**, et le dictionnaire ne donne aucune indication quant à la fréquence d'une expression plutôt qu'une autre. La très grande majorité de ces 800 termes sont cependant présents dans le dépouillement des revues, et une grande partie des absences s'expliquent par la nature partielle de l'échantillon, et aussi par le fait que la plupart de ces emprunts sont si bien intégrés que même l'œil averti ne les repère pas, et c'est seulement au moment de faire la comparaison avec les solutions trouvées dans d'autres langues qu'on s'aperçoit qu'il s'agit d'emprunts cachés.

Si l'on accepte que la majorité des termes de l'informatique sont des emprunts sémantiques, il devient nécessaire de les étudier dans le détail afin d'en comprendre le fonctionnement. Revenons aux définitions de base : un emprunt sémantique est un mot déjà existant dans la langue qui emprunte, mais qui se voit doté d'un sens nouveau, qu'on peut attribuer à l'influence d'un mot étranger. **Routine** est un mot attesté en français depuis le seizième siècle, mais lorsqu'on l'emploie avec le sens de **sous-programme**, on emprunte le sens (néologique en anglais également !) de *routine* (anglais). Certains de ces emprunts sémantiques sont considérés comme indésirables (encore un anglicisme) ; Violaine Prince[2] en a rappelé un certain nombre lors du Colloque du G.E.P. à Strasbourg en 1986, mais l'immense majorité relevée dans la presse de vulgarisation ne soulève aucune passion tout simplement parce qu'ils passent inaperçus. Ces emprunts sont invisibles parce que le nouveau sens, imposé à partir de l'anglais, est déjà très proche du sens français, et cette extension semble aller de soi. Quoi de plus normal, par exemple que d'appeler **filtre** le dispositif qui permet d'éliminer des données parasites, mais le français aurait-il trouvé cette solution si l'anglais n'employait déjà **filter** ?

Ceci est plus particulièrement le cas lorsqu'il s'agit de ce qu'on pourrait appeler des traductions transparentes, lorsque le mot anglais est identique ou très proche du mot français. Le cas d'identité sur le plan du signifiant écrit est assez courant et représente quelque 90 termes du *Dictionnaire Larousse* (exemples : **activation, bus, mode, terminal**, etc.) les mots étant des bases simples (**menu**) ou des dérivés (**transmission**). Étant très proches dans les deux langues, le nouveau sème ajouté n'étonne guère, et ce type d'emprunt minimaliste est bien accepté et souvent employé. Un cas contraire, **routine**, bien qu'assez employé, est réprouvé par le *Dictionnaire Larousse*, sans doute à cause du sens assez éloigné.

L'essentiel de la classe des traductions « transparentes » est composé de mots facilement identifiables comme les « mêmes » qu'en anglais, mais ayant subi quelques modifications orthographiques. Certains, comme **hôte** pour *host* commencent à virer vers l'opacité, mais restent suffisamment proches sur le plan de la substance pour que l'identité soit évidente. Le dictionnaire de P. Morvan[1] contient 44 emprunts sémantiques transparents composés d'un seul monème (**accès, bloc, commande**, etc.), 120 faisant entrer un ou plusieurs affixes, ceux-ci étant également transparents (**absorbeur, microlangage**, etc.), 49 autres sont des composés (**accès direct, adapteur de ligne**, etc.), composé pris dans un sens assez large et où l'ordre des éléments (**accès direct** *direct access*)

ou l'introduction de prépositions (**générateur *de* caractères**) trouble un peu la transparence : pourquoi dit-on par exemple **effet Josephson** mais le **graphe *de* Kiviat** ?

Le cas des dérivés parmi les emprunts transparents se rapproche d'un certain type de créativité lexicale, bien que confectionné sous impulsion étrangère. **Parité** est un dérivé qui, sur le plan formel, ne doit rien à l'anglais : l'emprunt reste circonscrit au niveau du signifié ; **absorbeur** est plus récent — il daterait des années cinquante[3], et on peut se demander s'il y a eu une nouvelle formation en français à partir de l'anglais *absorber*. Il existe même le cas de dérivés français qui connaissent une espèce de limbes et qui resurgissent à l'époque contemporaine (**adressage** : 1967 ; **additionneur** : 1965 ; **absorbeur** : 1975 ; **compactage** : 1967 ; **masquage** : 1967 ; **quantification** : 1967, tous relevés dans le fichier de l'INALF de Villetaneuse dans d'autres domaines que l'informatique). Les nouvelles formulations, les créations donc, sont simplement un prolongement de ce type de reprise d'élément existant virtuellement. Lorsqu'on crée **émulateur** c'est bien sûr à cause de l'anglais *emulator*, mais une création indépendante aurait été également possible grâce au paradigme existant : **émuler, émulation, émulateur** (comme **admirer, admiration, admirateur**). Cette tendance à remplir les cases vides des paradigmes peut aboutir à la création de doublons gênants (**interprète/interpréteur**, ce dernier relevé dans le *Dictionnaire Larousse*, mais absent du corpus magazine).

Les emprunts sémantiques forgés à partir d'une traduction « opaque » sont encore plus nombreux que la variété transparente que nous venons d'examiner : 70 emprunts concernant un seul monème (**appel** pour rendre *call*, **arbre** pour *tree*, **verrou** pour *lock*), plus 140 dérivés (**bourrage** pour *jam*, **délimiteur** pour *separator*, **sortie** pour *output*), et près de 250 composés, ce dernier toujours compris dans un sens large (**arrière-plan** pour *background*, **défaut de page** pour *page fault* jusqu'à **expression graphique assistée par ordinateur**).

Il existe une catégorie intermédiaire entre transparence et opacité : les dérivés ayant un radical transparent, mais un suffixe divergeant (**analogique** pour *analogue*, **branchement** pour *branch*, **adressage** pour *addressing*). Ce dernier cas est bien représenté : **-age** est employé pour rendre *-ing*, seul suffixe anglais courant n'ayant pas d'équivalent transparent en français. On a donc **(en)codage, bourrage, compactage, débottelage, (dé)multiplexage, fenêtrage, formatage, masquage, paramétrage, remplissage, routage,...** mais cette solution est loin d'être automatique et plusieurs de ces dérivés ne trouvent guère la faveur des journalistes de l'informatique. La majorité des cas, comme ceux que nous venons d'examiner, sont transparents de par leur radical, et peu visibles en tant qu'emprunts : cette sous-catégorie compte une cinquantaine de cas.

Les autres dérivés se décomposent en radical non transparent et affixe indifféremment transparent ou non (**affaiblissement** *attenuation* ; **bibliothécaire** *librarian* ; **échantillonnage** *sampling*) ; le modèle anglais ne comporte pas nécessairement un affixe (**clavier** *keyboard* ; **décalage** *shift* ; **anticipation** *look ahead*), bien au contraire, la dérivation semble relativement moins exploitée en anglais qu'en français.

Les composés représentent la catégorie la plus complexe ; même si tous les éléments ne sont pas nécessairement opaques (**analyse de données** *data analysis* ; **corps de Gallois** *Gallois field*), mais une étude exhaustive de cette partie dépasse les limites de cette étude.

À l'intérieur de cette catégorie, la plus complexe, deux tendances se dessinent. Tantôt il s'agit d'une traduction qui de toute évidence s'impose (**anneau** représente une traduction élémentaire de *ring*), une espèce de traduction transparente, même si les mots ne se ressemblent pas sur le plan du signifiant. Tantôt il s'agit d'une traduction beaucoup moins directe, différente de ce qu'on trouverait si on prenait le premier mot suggéré par un dictionnaire bilingue (**amorce** de *bootstrap* ; **bascule** de *flip-flop*, **exécu-**

tion pour *run*). Ces traductions qu'on peut appeler, d'après Vinay et Darbelnet[4], des traductions dynamiques, sont moins nombreuses qu'une traduction directe dans les cas où on rend de cette façon un seul monème (**champ** pour *field*, par exemple, **écran** pour *screen*, **fenêtre** pour *window*,...) ; les traductions dynamiques représentent une petite vingtaine de cas sur la cinquantaine que compte l'ensemble de la catégorie. Par ailleurs, on remarque que ces traductions dynamiques sont les moins facilement acceptées par les utilisateurs. Si on se réfère aux dépouillements, on constate que *string*, *array*, *bootstrap* se lisent plus souvent que leur traduction dynamique. Ceci s'explique peut-être du fait qu'une traduction dynamique, même brillante, est toujours sujette à une autre interprétation, et la présence de traductions concurrentes aboutit souvent à l'adoption définitive de ... l'emprunt brut. La catégorie comporte aussi des succès notables, tels **puce** pour *chip* (pourtant concurrencé par **pastille**), ou **débit** pour *throughput*.

La traduction simple semble aller de soi : il est évident, dira-t-on, qu'on traduit *flag* par **drapeau**, *loop* par **boucle**, *window* par **fenêtre** ; évident en français, peut-être, mais beaucoup moins si on passe à une autre langue : l'allemand, par exemple, d'après le sondage dans la presse de vulgarisation informatique analogue à celui effectué en français, retient dans ces trois cas, comme dans bien d'autres, l'emprunt brut. Il arrive aussi que des journalistes français emploient parfois l'emprunt direct (*directory* plutôt que **catalogue**, par exemple), mais de façon moins systématique que les Allemands.

Les dérivés opaques se répartissent en deux groupes à peu près égaux de traductions directes (**apprentissage** rend *learning*) et dynamiques (**accessibilité** pour *ease of use* ; **permutation** pour *swapping*). La plus forte représentation des traductions dynamiques parmi les dérivés s'explique du fait que la plupart des termes d'origine, donc anglais, comportait déjà au moins deux morphèmes, soit 80% des cas ; cette complexité incite donc à chercher des traductions dynamiques afin de mieux rendre les deux, trois ou quatre morphèmes. Encore une fois, les traductions directes sont bien mieux acceptées que les traductions dynamiques : **sauvegarde** est autrement plus fréquent dans les magazines que *saving*, mais *hard copy* se lit toujours et non **reprographe**.

Lorsqu'on étudie les traductions qui sont effectivement adoptées, on peut se demander s'il n'y a pas un autre type de transparence qui rentre en jeu. **Fenêtre** est la traduction évidente de *window*, comme chaque élève de sixième le saurait, mais *loop* ne fait pas obligatoirement partie du vocabulaire anglais de l'adulte qui fait de l'informatique : on ne relève jamais *window* dans les textes français dépouillés, mais *loop* revient constamment. De la même façon, la distance sémantique qui sépare les mots peut agir comme un frein : *string* fait partie du vocabulaire du ministère de l'Éducation nationale de la classe de 3e, mais l'idée d'appeler une suite d'éléments **ficelle** ne viendrait pas à l'esprit d'un francophone. Pierre Morvan suggère **chaîne**, qui rend à peu près l'image de *string* et qui par ailleurs s'emploie également pour une suite d'éléments.

Les francophones qui s'intéressent au vocabulaire anglais de l'informatique sont souvent frappés par les mots d'origine anglo-saxonne, si différents du français, et si difficiles à rendre en français (*latch*, *batch*,...)[5]. Mais lorsqu'on compare les emprunts bruts faits à l'anglais de l'informatique en allemand, on relève davantage de termes d'origine latine, termes dont l'origine anglaise passe inaperçue en français, grâce aux emprunts sémantiques que nous venons d'étudier. Si on pense que les mots d'origine germanique de l'anglais sont mieux assimilés dans l'allemand de l'informatique, on a peut-être raison, mais seulement dans la mesure qu'il s'agit d'emprunts bruts, et non pas de traductions.

Quelle est donc la part du vocabulaire d'origine latine de l'informatique en anglais, vocabulaire virtuellement transparent en grande partie en français ? Si l'on

compte les termes anglais donnés en équivalence des mots français dans le dictionnaire de Pierre Morvan, on arrive aux proportions suivantes :

- une centaine de termes d'origine germanique (si on ne tient compte que des monèmes : *readability* et *acknowledgement* sont donc comptabilisés comme d'origine germanique) ;
- 80 mixtes, latin et germanique (tels *call confirmation*) ;
- plus de 550 d'origine latine, soit empruntés directement au latin (*processor, preprocessor,...*), avec ou sans élément grec, soit par le truchement du français (*level, core, pattern,...*). Ceux qui remontent à l'époque anglo-normande sont souvent aussi difficiles à rendre que ceux d'origine germanique ; nous connaissons les difficultés rencontrées lorsqu'on a souhaité franciser *management*, cousin anglais de **ménagement** ; *pattern*, qui a la même origine que **patron** ; ou *display*, de **déployer**. Dans d'autres cas, assez rares toutefois, des mots empruntés au latin ont complètement disparu du français : on ne peut plus compter sur l'origine latine de *remote* pour la rendre en français. Dans la très grande majorité des cas toutefois ces mots d'origine commune servent à créer les emprunts sémantiques totalement transparents (**rejet** *reject*) ou presque (**amplificateur**). Ceci explique l'impression d'anglicisation plus poussée dans l'allemand de l'informatique : la proportion de vocabulaire latin venu de l'anglais est très importante et très visible, et les mots germaniques sont souvent plus éloignés de l'allemand contemporain que les mots anglo-normands du français.

Le français occupe donc une position à part, car la naturalisation de la terminologie de l'informatique se trouve facilitée par le fonds commun gréco-latin partagé avec l'anglais ainsi qu'une volonté d'agir, d'aménager, de créer. N'oublions pas toutefois que la création lexicale a des origines modestes, parmi lesquelles on compte l'emprunt sémantique, façon simple et rentable, surtout à une époque où les termes à rendre en français sont de plus en plus nombreux. Une connaissance des mécanismes de ce procédé nous permet d'enrichir la terminologie de façon efficace et peu voyante tout en contrôlant des glissements de sens intempestifs.

Notes

1. Pierre MORVAN (1981) : *Dictionnaire de l'informatique*, 6e éd., revue et mise à jour, Larousse, 341 p.
2. Violaine PRINCE (à paraître) : « La pidginisation du français par le jargon américain de l'informatique », *Actes du IIe colloque du Groupe d'études sur le pluralisme européen.*
3. *Petit Robert* : « 1948, de absorber ».
4. J.-P. VINAY et J. DARBELNET (1958) : *Stylistique comparée du français et de l'anglais*, Didier.
5. Georges LURQUIN (1982) : « La langue spéciale des informaticiens », AILF DOC.