

TIBOR ÓRSI

Les structures étymologiques du vocabulaire de l'informatique

Tout au long du XX^e siècle, la langue anglaise (par là nous entendons aussi l'américain) a constitué la source la plus importante des emprunts lexicaux du français. L'afflux massif des anglicismes s'est accéléré avec la naissance et le développement de l'informatique. La menace de la prolifération des anglicismes ne cesse de provoquer des réactions officielles violentes. Déjà en 1955, le mot *franglais* est composé par dérision à l'aide des syllabes initiale de *français* et finale de *anglais* pour désigner l'emploi abusif des anglicismes. Une opération systématique de désanglicisation est entreprise au début des années 1970. Quinze *commissions ministérielles de terminologie* (C.M.T.) sont constituées. Elles formulent des propositions officielles pour désangliciser les vocabulaires spécialisés des différentes branches de la science et de la vie publique. Les propositions sont publiées dans le *Journal officiel* à intervalles réguliers. Appelées euphémiquement « recommandations », elles n'en sont pas moins des règles strictes que les institutions publiques, la presse mais aussi les dictionnaires bilingues et monolingues sont tenus d'accepter, d'appliquer et de propager.

Dans la présente communication, nous focaliserons notre approche sur les aspects étymologiques de l'élimination des anglicismes. Nous procéderons à l'inventaire des techniques mises en œuvre au cours de la désanglicisation en examinant des extraits des recommandations officielles pour la francisation du vocabulaire de l'informatique comme elles figurent dans le *Journal officiel*. Le numéro du 16 mars 1999 a été consacré au vocabulaire de l'informatique. Par la suite les dictionnaires courants reprennent ces recommandations. *Le Petit Robert*, dont nous disposons des éditions régulièrement mises à jour, se prête bien au suivi des changements du vocabulaire.

Commençons par le nom de la science. Mentionnons ici que, selon le *Trésor de la Langue Française informatisé*, le terme *informatique* attesté pour la

première fois en 1962 a été inventé par Ph. Dreyfus, d'après Gilbert (1971). L'*Oxford English Dictionary* retrace la source immédiate du terme anglais correspondant au terme russe *informatika* attesté en 1966, dont la traduction anglaise apparaît en 1967. Une autre création lexicale française qui a connu un grand succès est l'objet fétiche de la science : *l'ordinateur*. Cet ancien mot français s'emploie en informatique depuis 1956, introduit à la demande de IBM France. Le terme *ordinateur*, contre toute attente, a évincé l'anglicisme *computer* et son adaptation *computeur*. *Ordinateur* entre dans les mots composés : *micro-ordinateur* s'emploie depuis 1971, abrégé familièrement en *micro* (1974) ; *mini-ordinateur* apparaît vers 1974. *Micro-ordinateur* est sérieusement concurrencé par *P. C.*, apparu vers 1982. Il s'agit du sigle anglais *Personal Computer* « ordinateur personnel », introduit à l'origine par la marque IBM.

Passons maintenant en revue des méthodes employées pour franciser la terminologie de l'informatique.

1) Une grande partie du vocabulaire scientifique est commune aux deux langues. Les terminaisons sont conformes aux règles morphologiques de chacune de ces langues. Il s'agit essentiellement de mots savants :

compatible/compatible, to compile/compiler, to connect/connecter à, error/erreur (f.), external/externe, to format/formater, hypertext/hypertexte (m.) icon/icône (m.) [mais *icône (f.)* « peinture religieuse »], *memory/mémoire (f.), menu/menu (m.), monitor/moniteur, multimedia/multimédia (m.), to select/sélectionner, text/texte (m.), virtual memory/mémoire virtuelle (f.), virus/virus (m.), etc.*

2) Certains mots savants ont pénétré l'anglais directement du latin, sans passer par le français. Évidemment, les termes français correspondants ne présentent aucune parenté étymologique avec les mots anglais. Les termes anglais appartenant à cette catégorie ont été remplacés par leurs traductions. C'est le cas de *to delete/effacer = supprimer, data/données (f. pl.)*.

3) Remplacement des termes anglais d'origine ou des mots anciennement passés en anglais

Beaucoup de termes anglais sont d'origine native, c'est-à-dire qu'ils sont issus de mots du fonds germanique commun. Nous n'avons trouvé aucun mot de ce type qui ait été épargné. Ainsi, *hardware* (attesté depuis 1515) a été remplacé par *matériel*, disponible en français. *Hardware*, employé pour la première fois en anglais dans le domaine de l'informatique en 1947, est attesté en français depuis 1965. La recommandation officielle *matériel* vise à remplacer le terme étranger. Le terme *software* est un néologisme en anglais (1960) sur le modèle de *hardware*. Le terme français *logiciel* (1965), formé sur *logique* d'après *matériel*, est un néologisme en français aussi.

Le terme anglais *disk*, *disc* et son homologue français *disque* dérivent indépendamment du grec par l'intermédiaire du latin. Le mot s'emploie en informatique à partir de 1947 : *floppy disk* (1972), *hard disk* (1975). Les calques français *disque souple* et *disque dur* apparaissent très vite. Vers 1975 *disquette* remplace *disque souple* et *disque* sera réservé au *disque dur*. L'addition du suffixe diminutif *-ette* sert à éloigner la forme française de l'original anglais. L'évolution des techniques a produit *disque compact* (v. 1982), calque de l'anglais *compact disc* (1979), lui-même fréquent en français y compris sous ses initiales *CD*.

4) Remplacement des anciens emprunts faits au français

Beaucoup de mots anglais, parmi eux certains termes informatiques sont en effet d'origine française. Qui croirait que le terme anglais *file* soit apparenté aux mots français *fil* et *file* ? La diphtongue [ai] confère au mot un caractère anglais. *File* apparaît en anglais au XVI^e siècle, l'usage du mot en informatique date de 1954. L'équivalent français *fichier* s'est spécialisé lui aussi de l'article de bureau à la désignation du fichier informatique. L'identité graphique (*fi-*), et aussi phonique [f] facilite l'association des deux termes.

5) Coexistence de doublets

Le terme anglais *CD-ROM* (1983) est francisé en *céderom* sans réussir toutefois à éliminer la forme anglaise. Pour citer un exemple, le manuel d'utilisation du *Petit Robert électronique* utilise uniquement la forme *CD-ROM* en tant que sigle (autrement dit chaque lettre est écrite en majuscule), alors que celui du dictionnaire *Le Littré* alterne le sigle *CD-ROM* et l'acronyme *céderom*. *CD-ROM* semble plus prestigieux, par rapport à la forme *céderom* qui paraît plutôt familière. Le sigle en français est invariable alors que la recommandation officielle *disque optique compact* et sa forme abrégée *doc* ainsi que l'acronyme *céderom* s'intègrent parfaitement à la morphologie française. Remarquons également qu'en anglais les sigles peuvent prendre la marque du pluriel : *PCs*, *CD-ROMs*.

Le mot anglais *bug* « erreur d'un logiciel entraînant des anomalies de fonctionnement » a été francisé en *bogue* par un arrêté ministériel de 1983. Le sens du mot anglais s'est développé de « bestiole nuisible » en « défaut » ; il s'agit d'une métonymie de la cause. D'après une citation titrée de l'*Oxford English Dictionary* sous l'entrée *BUG* n. et datant de 1889, un *bug* aurait causé la défaillance du phonographe de l'inventeur Edison. Le mot anglais paraît avoir été rapproché du mot français le plus proche en forme. Le mot *bogue* recouvre trois homonymes : 1. « enveloppe piquante de la châtaigne » ; 2. « anneau de fer fixé au manche des gros marteaux de forge » ; 3. « espèce de poisson » auxquels s'ajoute la forme francisée de *bug*. Comme les trois mots qui existaient originellement désignent des réalités ou des créatures assez rares et que seul le troisième est masculin, l'élargissement de la série homonymique n'empêche pas la compréhension. J'ai testé la fréquence des deux formes *bogue* et *bug*, par une recherche sur la Toile. Pour l'équivalent de l'expression *millenium bug* le moteur de recherche *altavista.fr* a produit 15800 résultats pour *le bug de l'an 2000*, alors que *le bogue de l'an 2000* n'en a produit que 2560, soit le sixième de la fréquence de l'anglicisme. Après une vingtaine d'années de coexistence, la forme recommandée n'a pas pu s'imposer. Le verbe anglais formé sur *bug* est *debug* « supprimer les erreurs de programmation ». La forme verbale française *déboguer* (1983) semble plus courante que *débuger*, sans doute en raison de la plus grande facilité à l'intégrer dans le système phonétique et morphologique.

6) Créations françaises originales

Prenons le cas du symbole typographique @ commercial. Ce symbole s'emploie depuis le XVI^e siècle dans les registres marchands où il désigne l'unité de mesure italienne *anfora* « amphore » ainsi que l'unité de mesure espagnole et portugaise *arroba*. En anglais, le @ commercial apparaît au XVIII^e siècle pour désigner le nombre des articles par rapport au prix. Depuis l'invention des e-mails, ce signe typographique sert de séparateur dans le libellé des adresses électroniques entre le nom de l'utilisateur (*user*) et celui du domaine (*domain*). Ce symbole aurait été choisi par Ray Tomlinson, le concepteur-projeteur du système des courriers électroniques. Par rapport à la forme anglaise *at sign* [ætsain], le correspondant français a plus de corps et plusieurs variantes : *arobase*, *arrobese*, *arobas*, *arrobas*. L'origine de ce mot est incertaine. L'emploi moderne attesté depuis 1995 dérive peut-être de la locution *a rond bas* (*de casse*). Il s'agit d'un terme technique d'imprimerie. *Casse* est emprunté à l'italien *cassa* « caisse » et désigne une boîte ou un meuble contenant les caractères typographiques. *Bas de casse* est la partie inférieure du meuble proche de la main du typographe et contenant les caractères les plus courants, *a* y compris.

Le terme anglais *e-mail* est un autre anglicisme dont la francisation mérite d'être examinée. Le mot anglais *mail* est un emprunt à l'ancien français *male* « sacoche, malle » issu du francique **malha* (d'où vient aussi le mot hongrois *málha*). *Mail* est attesté en anglais au sens de « courrier » en 1654, *e-mail* apparaît en 1982. Il s'agit de l'abréviation de *electronic mail*. L'expression apparaît en français comme anglicisme en 1994, calqué ensuite en *courrier électronique*. La contraction de cette dernière expression a produit l'équivalent possible *courriel*, utilisé d'abord au Québec et en Belgique. À l'écrit, l'administration recommande, devant une adresse électronique, l'abréviation *mél.* qui ne doit pas être employée comme substantif. Le choix de *mél.* a été sans doute facilité par la ressemblance de la forme prononcée à celle de l'original anglais tout en permettant une interprétation française.

7) Motivation phonétique

Le choix du terme français est motivé par son association à la forme prononcée de l'anglicisme. Nous avons déjà mentionné à ce propos les termes *mél.* et *fichier*. L'expression anglaise *on-line* apparaît en 1950. Depuis 1989 on utilise la forme francisée *en ligne*. On peut faire remonter le mot anglais *line* et le mot français *ligne* au même étymon latin : *linea*. *On-line* s'emploie uniquement dans un contexte informatique de même que son équivalent français. La préposition *en* [ã] correspond étymologiquement à *in* en anglais. La prononciation de l'anglais *on* [on] est plus proche de celle de *en*.

8) Motivation sémantique

Bit (1948) est un mot-valise combinant *bi-*, première syllabe de *binary* « binaire » et *-t*, dernière lettre de *digit* « chiffre ». Le terme désigne le chiffre binaire et l'unité élémentaire d'information. Au-delà de cette décomposition, le néologisme révèle en même temps une forte motivation sémantique : en anglais *bit* représente des homonymes qui signifient « petite quantité, morceau » ; « mèche de perceuse » et « mors de cheval ». Chacun des sens énumérés dérive d'un ancien verbe germanique *bitan* « mordre ». Le sens anglais le plus courant est « petite quantité » qui s'harmonise parfaitement avec le signifié du terme technique. *Bit* en tant qu'emprunt technique français est immotivé. En anglais, l'ensemble de huit *bits* constituant une unité complète d'information est *byte* (1964). C'est un mot-valise irrégulier forgé de *bit* et *bite* « petit bout, morceau ». La recommandation officielle est *octet* qui semble emprunté à l'anglais *octet*, attesté en 1919 comme terme de physique nucléaire. Dans les années 1960, il est passé en informatique par réemprunt. *Octet* est composé de l'élément latin *oct-* « huit » et du suffixe *-et*. Il s'agit donc d'un mot motivé. Pour indiquer la multiplication, dans le système scientifique international, on ajoute des préfixes d'origine grecque : *giga-*, *méga-*, etc.

9) La vitalité du français

Nous avons déjà abordé *matériel* et *logiciel*. Ce dernier illustre la naissance du nouveau suffixe *-iel* qui apparaît dans d'autres créations françaises : *progiciel* (1962) formé de *pro(duit)* et *(lo)giciel* qui se substitue à *package* ;

didacticiel (1979) « logiciel à fonction pédagogique », qui peut être rendu en anglais seulement par une longue paraphrase : *educational software programme* ; *ludiciel* (1980) « logiciel destiné à des jeux » formé de *ludi(que)* et (*logi*)*ciel* qui correspond à *computer game*. Dans chacun des exemples ci-dessus, le suffixe natif français a été ajouté à un radical savant créant ainsi des formations hybrides. Un autre suffixe de création récente est *-ique* qui sert à désigner toutes les activités s'effectuant au moyen de l'informatique : *bureautique, monétique*, etc.

La langue française a bien intégré certains emprunts. *Internet* (1973) est de formation hybride. *Inter* « entre, parmi » est une préposition latine qui s'ajoute au mot natif anglais *net* abrégé de *network* « réseau ». Le mot *Internet* s'est répandu en français vers 1995. Immédiatement après, un nouveau mot a été créé en français pour désigner « l'utilisateur de l'Internet » : *internaute*. L'élément *-naute* du grec *nautês* « navigateur » apparaît dans une dizaine de mots composés en français dont le plus ancien est *aéronaute* « personne qui pratique la navigation aérienne ». En 1794, il s'agissait des frères Montgolfier. La création la plus récente est *entreprenaute* (1999). Ce mot-valise est forgé sur *entrepren(eur)* et (*inter*)*naute* et désigne un « créateur d'entreprise sur Internet ».

Les termes que nous venons d'étudier ne représentent qu'un fragment de la terminologie de l'informatique. Les catégories sous lesquelles nous avons rangé les termes se recoupent partiellement. Assez souvent, un nouveau terme apparaît dans sa forme étrangère comme *driver* (1972), ensuite sa forme est adaptée en *driveur* avant d'être officiellement remplacée par *pilote*. Certaines adaptations persistent. C'est le cas de *spool* adapté en *spoule*. *Modem* (1958) a été forgé en anglais sur *mo(dulator)* et *de m(odulator)*, mais peut s'analyser également en français comme *mo(dulateur)* et *dé m(odulateur)* grâce au vocabulaire scientifique commun. Comme les arrêtés terminologiques sont pris avec une certaine régularité, la période d'incertitude des anglicismes dans le domaine de l'informatique se réduit.

La tendance générale qui s'accroît est de débarrasser la langue française des anglicismes dans tous les domaines. Si *pipeline* prononcé d'abord [païplain], ensuite [piplin] est rejeté comme étranger au français, *oléoduc* est-il plus français ? L'anglicisme criard est remplacé par un « gallicisme » qui n'est

pas un mot indigène du tout. Le fait que le terme *oléoduc* se compose d'éléments savants sur le modèle d'*aqueduc* en fait-il un mot d'ordre supérieur ?

L'expression *photographie digitale* ne pourrait-elle pas faire partie du vocabulaire français ? Le mot *digital* dérive du substantif *digit*, qui signifie « doigt », mais aussi « chiffre », primitivement « nombre inférieur à dix que l'on peut compter sur les doigts ». En terme d'informatique, *digital* s'applique à ce qui « emploie des nombres exprimés directement en chiffres dans un système décimal, binaire ou autre ». En anglais moderne, *digital* ne s'emploie plus dans le sens de « relatif aux doigts ». *Digitale* n. f. en tant que terme de botanique est attesté depuis 1545. L'adjectif savant correspondant au mot *doigt* s'emploie depuis 1732 dans des expressions aussi courantes qu'*empreinte digitale*. Comme terme de numérotation, *digital* est un emprunt à l'anglais et ce fait scelle son destin : il sera supplanté par *numérique*, mot non moins autochtone.

Face à l'énorme pression linguistique étrangère, les réactions officielles – tout en prétendant défendre le français contre l'anglo-américain – ne l'appauvrissent-elles pas à la fois en refusant des compromis acceptables ? Les éléments communs du vocabulaire quotidien ainsi que le vocabulaire savant disponible dans les deux langues pourraient servir de terrain d'entente.

Il paraît évident que l'attitude officielle, hostile à l'égard des emprunts récents à l'anglais et à l'américain, résulte d'un sentiment de menace contre la langue et la culture françaises. L'attitude officielle consiste à se débarrasser des anglicismes à tout prix. Reste à examiner si les utilisateurs de la langue se conforment aux recommandations officielles. La lecture des revues comme *L'Express* ou *Paris Match* ainsi que des recherches effectuées sur Internet suggèrent que de nombreuses recommandations demeurent lettre morte.

TIBOR ÓRSI

École supérieure Eszterházy Károly, Eger
Courriel : orsitibor@t-online.hu

Bibliographie

- Le Littré. Dictionnaire de la langue française. Version CD-ROM*, Marsanne, Redon.
- Le Petit Robert sur CD-ROM. Version 2.2*, Paris, Dictionnaires Le Robert/Sejer, 2004.
- Lenoble-Pinson, M, *Anglicismes et substituts français*, Paris – Louvain-la-Neuve, Duculot, 1991.
- Oxford English Dictionary. Second Edition on CD-ROM. Version 3.1*, Oxford, Oxford University Press, 2004.
- Dictionnaire historique de la langue française*, REY, Alain (sous la dir. de), Paris, Dictionnaires Le Robert, 1992.
- Trésor de la Langue Française informatisé (TLFi)*, Paris, CRNS Éditions, 2004.
- Journal officiel du 16 mars 1999. Vocabulaire de l'informatique et de l'internet*,
<<http://www.culture.gouv.fr/culture/dglf/coeter/16-03-99-internet-listes.html>>