

La notion de besoin terminologique et la naissance de lexiques spécialisés au dix-neuvième et au vingtième siècles

Certaines techniques marquent et fondent une époque économique en même temps que culturelle: par exemple, le chemin de fer pour le dix-neuvième siècle, l'informatique pour le vingtième.

À l'apparition d'une nouvelle technique, ses dénominations ou celles des éléments matériels qu'elle apporte, sont éminemment variantes.

Cette variation signale le travail linguistique qui se fait pendant des années, parmi les spécialistes et pendant le processus de vulgarisation, pour ranger la nouveauté dans le monde connu tel que le représente la langue stabilisée. Le besoin lexical est ainsi comblé mais remplacé par un besoin plus proprement terminologique.

1 Introduction

Nous distinguons ici deux types de besoin qui se font ressentir lors de la création d'un vocabulaire nouveau: besoin lexical et besoin terminologique. Le premier est souvent comblé par une profusion synonymique de différents ordres, que nous examinerons; le second relève plutôt de l'épuration, de la décision plus ou moins consciente de ne retenir qu'un seul terme pour dénommer un concept reconnu dans la nouvelle sphère d'activités. Pour illustrer ces mécanismes, nous choisissons deux exemples tirés du domaine français: le vocabulaire des chemins de fer au dix-neuvième siècle, et celui de l'informatique au vingtième.

2 Le chemin de fer

Il existe en français tout un vocabulaire ancien, tiré des traditions techniques, qui va être réutilisé lors de l'introduction en France des chemins de fer (Wexler 1955). Le procédé majeur de création lexicale est alors le déplacement de sens, d'abord par extension, puis par spécialisation.

Par exemple, *chemin ferré* ou *voie ferrée* existent depuis des siècles, avec le sens de «chemin rendu ferme par des pierres couleur de fer», puis pour désigner des chemins où les bandes de roulement sont en bois. *Chemin de fer* n'est donc pas un calque de

railway, contrairement à ce que dit le *Dictionnaire de l'Académie française* récemment paru, car la traduction littérale de *rail way* et *iron railway* aurait été *route à ornières* et *route à ornières de fer*. Mais il est vrai que pendant plusieurs années, l'emprunt direct *railway* concurrence sérieusement *chemin de fer* ou *voie ferrée*. Pour ce que nous appelons *gare*, on empruntera à la navigation les mots *embarcadère*, *débarcadère*, *port-sec*, et le mot *gare* lui-même. Ce que nous appelons *rail* a été nommé *guide*, *bande*, *longuerine*, *tringle*, *plaque*, *barreau*, etc., et surtout *ornière*. C'est face à une quinzaine de mots français différents qu'a été adopté le mot *rail*. Sa prononciation est encore fluctuante: [Rɛl] ou [Raj].

On hésite beaucoup sur tous ces termes. Par exemple, pour *voie ferrée* ou *chemin de fer*, les hésitations et polémiques portent sur différents points: son exactitude, car il y a des «ornières» métalliques (toute la largeur de la voie n'est pas «ferrée»), le nombre (ce que nous appelons 1 ligne = 2 voies = 4 rails), et la forme (*ornière*, *rainure*, *bande*). Dans *chemin de fer*, *fer* est d'abord un vrai sémantème (qualifié de «fondu» ou «forgé»), puis, dès lors que le fer est toujours laminé, on ne le précise plus. Un peu plus tard, quand le fer est remplacé par l'acier, l'expression *chemin de fer* est conservée sans égard au sémantisme de *fer*.

Pendant les trente premières années du dix-neuvième siècle, le fait majeur est l'hésitation terminologique. Si globalement, avec l'habitude, on passe de la description (périphrases, syntagmes descriptifs comme *bandes de fer*), à la dénomination (*rail*), cela se fait à travers mille variantes - d'ailleurs les

techniques elles-mêmes, y compris les modes d'organisation, ne se fixent que progressivement.

À propos de chacun des termes qu'il étudie, Wexler montre que la langue française proposait plusieurs solutions. L'appel à des terminologies voisines, par extension de sens puis spécialisation, est massif, les néologismes de forme peu nombreux (ex. *locomotive*), la dérivation réduite, la composition un peu moins. Les calques, pour la plupart, n'ont pas survécu. Quant aux emprunts, ils ne s'expliquent pas par une incapacité de la langue à désigner, mais par des causes complexes.

Une première série de causes est extralinguistique, externe, c'est-à-dire sociale. De toute évidence, le prestige du pays d'origine d'une technique, aux yeux des techniciens ou même du grand public, joue un rôle, d'autant plus qu'il se rencontre avec des valeurs politiques (libertés démocratiques). Si l'adoption du mot *tunnel* a évincé les périphrases (*passage souterrain*), ainsi que les autres termes français disponibles (*galerie, percée*), y compris son correspondant étymologique *tonnelle*, il est patent que la popularité et le prestige, dans l'information circulant à l'époque en France, du projet de « tunnel » sous la Tamise à Londres, y est pour quelque chose. Un autre aspect le confirme peut-être. Alors qu'avant 1830, dans la région de Saint-Étienne, seuls des mots français sont utilisés, les premiers passages souterrains nommés « tunnels » sont ceux de la région parisienne, quelques années plus tard. Cela indique peut-être un lien avec l'accès à l'information sur l'étranger (qui est vraisemblablement plus facile à Paris qu'à Saint-Étienne), mais surtout un lien avec des milieux « à la mode ». Le même raisonnement vaudrait pour la douzaine de mots empruntés à l'époque (*ballast, tender, rail...*).

Il ne s'agit pas ici de la langue anglaise en tant que telle, mais des notions qu'elle porte, par quelque

moyen que ce soit. Si l'on trouve en Grande-Bretagne des termes latins, on se les approprie également. D'où le « double emprunt » des mots latins utilisés en anglais et adoptés en français: *via, terminus*, ou de néologismes latinisants comme *viaduc* ou *locomotive*.

La surabondance des possibilités lexicales du français oblige à se demander quel est le rôle du mot emprunté. On doit alors remarquer qu'une particularité de l'emprunt est sa non-motivation: les mots français au contraire sont linguistiquement motivés. Il est remarquable qu'on ne discute pas du bien-fondé sémantique du choix d'un mot emprunté, puisque la motivation linguistique est absente, alors que les mots français suscitent des débats. Il n'est pas invraisemblable que cette absence de motivation soit un facteur adjuvant, comme si le mot non-motivé opérait une sorte d'arbitrage. Face à une terminologie mal fixée, l'emprunt bénéficie de l'avantage de n'être pas linguistiquement contestable: il ne renvoie à rien d'autre que son référent, il est plus arbitraire que des mots français, plus motivés, et plus chargés de connotations. Il est, en quelque sorte, un signe plus purement saussurien. En admettant la notion de besoin comme préalable à l'innovation lexicale, on est amené à préciser à quel besoin supplémentaire ou différent répond le mot emprunté: la fixation du lexique, et non sa création. Il faut donc bien parler de besoin et de problème **terminologiques** et non lexicaux.

3 L'informatique

Deux raisons nous autorisent à privilégier la période de la troisième génération des ordinateurs (vers 1970) dans l'histoire de la constitution du vocabulaire français de l'informatique. D'abord, il s'agit de la période de l'implantation

effective de l'informatique en Europe, époque où plusieurs catégories de professionnels (mais pas encore le grand public) sont directement concernées: fabricants (ingénieurs, techniciens, commerciaux), utilisateurs, chercheurs et enseignants. C'est aussi à partir de cette période qu'on commence à disposer d'une documentation linguistique adéquate: les premiers vocabulaires européens sont parus au début des années 1960, mais c'est surtout la thèse de Christiane Marcellesi qui permet le premier instantané de la terminologie de l'informatique à la fin de la période de l'implantation. En effet, il s'agit d'une période où l'informatique cesse d'être classée « secret défense » (la guerre de Corée, où elle a fait ses preuves dans le domaine militaire, n'était pas loin) et où les fabricants aussi commencent à franciser leurs manuels, et à se lancer dans la publicité. Ce n'est pas encore la période de « l'informatique pour tous », mais elle se prépare déjà.

D'où vient ce vocabulaire? Christiane Marcellesi est formelle: il s'agit de la « constitution d'un champ lexical nouveau à partir d'un champ lexical déjà constitué dans une langue étrangère » (1972: 15); « l'ensemble du matériel lexical de l'informatique de langue française est traduit ou adapté de l'anglo-américain. » (1972: 20.) Notre propos est de voir comment cette traduction, cette adaptation, s'est faite, ou, dans d'autres termes, comment le français a comblé ce besoin lexical et terminologique. Cette tâche n'est pas facile, car il ne s'agit pas d'un vocabulaire un et indivisible. Même à l'époque où l'informatique était l'apanage d'un nombre restreint de spécialistes, les formes que pouvaient prendre l'expression étaient déjà très variées: manuel d'utilisation, manuel pédagogique, article, texte publicitaire, conférence, présentation, conversation d'atelier... toutes avaient des caractéristiques qui leur étaient propres, et toute mention du

«discours informatique» ou «vocabulaire informatique» représente une simplification, certes nécessaire, mais réductrice.

En 1972, l'anglais avait en partie comblé son besoin terminologique et il existait déjà un vocabulaire informatique de base relativement établi, du moins entre spécialistes de langue anglaise, mais encore presque dépourvu de supports lexicographiques. Or, on aurait pu penser que le français aurait comblé son besoin terminologique grâce à des emprunts directs massifs. En réalité, Christiane Marcellesi en relève très peu, du moins dans son premier corpus. Lors d'un travail postérieur, elle fait état de davantage d'emprunts directs dans le discours des pupitreurs, indication que différents acteurs dans le même secteur disposent de différents vocabulaires employés selon les situations, et que le taux d'emprunts directs peut varier de façon importante.

Examinons brièvement les quelques emprunts directs de cette période:

background (hapax), *batch processing*, *bit*, *computer*, *firmware*, *foreground* (hapax), *hardware*, *light pen*, *manware*, *off line*, *on line*, *packages*, *software*, *remote batch*, *time sharing*.

Mis à part les deux hapax, la plupart des anglicismes étaient plus employés en 1970 que les équivalents français relevés, sauf peut-être pour *batch processing* (*traitement par lots*) et *time sharing* (*temps partagé*). Pour *hardware* et *software* les équivalents aujourd'hui courants (*matériel* et *logiciel*) n'apparaissent pas encore, et pour *on line*, *off line*, seul le dernier (*autonome*) est encore assez courant. On constate avec intérêt que sur les quinze termes relevés ici, dix ont été remplacés dans le premier arrêté de la Commission ministérielle de terminologie, de 1981. Les anglicismes, déjà relativement peu nombreux au départ, ont depuis pratiquement disparu au cours de la

période de la maturation terminologique.

Si l'on relève peu d'emprunts directs, on peut supposer qu'il existait une politique efficace de traduction. En effet, les grandes sociétés d'informatique commençaient à publier leurs propres glossaires, tel IBM:

«Nous avons aussi introduit dans ce glossaire quelques néologismes. C'est surtout lorsqu'il nous a semblé difficile ou inélégant d'adopter les termes français déjà existants. Les néologismes sont presque toujours dérivés de racines nobles, c'est-à-dire grecques ou latines, et nous n'avons pas pensé devoir reculer devant quelques audaces dans ce domaine.» En réalité, Mme Marcellesi en relève également fort peu.

En outre, différents organismes publics et parapublics se penchent déjà sur l'adaptation du français aux nouvelles technologies et à l'informatique en général. Il s'agit dans un premier temps du Comité d'étude des termes techniques français, et du Conseil international de la langue française.

La néologie se faisait donc de façon consciente, comme la citation d'IBM le suggère, mais elle se pratiquait aussi inconsciemment, grâce à une combinaison de trois facteurs: le premier, les matrices lexicogénétiques du français, ses possibilités de dérivation, de composition et de néologie sémantique; le deuxième, le modèle de l'anglais qui suggérait dans quel sens on pouvait traduire; le troisième et dernier facteur, les origines scientifiques et techniques de l'informatique (électronique, cybernétique etc), qui, malgré sa nouveauté, était aussi le résultat de la convergence de toutes sortes de techniques existantes.

Examinons d'abord, à l'aide de cette première analyse de Marcellesi, quelles ressources linguistiques propres au français étaient exploitées pour l'élaboration de ce vocabulaire

en réponse au besoin de communication dans le domaine, c'est-à-dire en réponse au besoin lexical. Sur le plan sémantique, elle parle d'«emprunts» faits à d'autres vocabulaires français, que nous préférons appeler opérations métasémiques (métonymie et métaphore). Du vocabulaire «commun», Marcellesi compte *élément*, *dispositif*, *ensemble*, *machine*, *mode*, *organe*, *séquence*, *système*, *temps*, *traitement*. On s'aperçoit qu'il s'agit souvent d'un sens technique déjà attribué à un mot du vocabulaire «commun», soit lié à la fabrication industrielle (*dispositif*, *ensemble*, *machine*, *système*, *traitement*), soit à un domaine particulier comme *balai* (électricité), *accès*, *adresse*, *bande*, *cadre*, *carte*, *couteau*, *entrée*, *feuillet...* La part du vocabulaire élaboré pour d'autres spécialités est donc importante. À part l'électricité, elle cite le rail (*aiguiller*, *aiguillage*), et le commerce (*charger*, *décharger*, *gérer*, *stock*, *stocker...*).

Cette exploitation de ressources lexicales existantes présuppose une langue déjà «standardisée», déjà équipée en vocabulaire de base industrielle et commerciale. Ce n'est pas le cas de la totalité du vocabulaire emprunté à d'autres domaines cependant: on constate un recours intensif aux métaphores anthropomorphiques, qui doivent être bien plus facilement transposables dans d'autres langues (*bras*, *tête*, *cerveau*, *mémoire*, *naître*, *lignée*, *génération*, *nourrir*, *alimenter*, ainsi que le vocabulaire des déplacements: *aller*, *envoyer*, *chercher*, *avancer*, *démarrer*, *défiler*, *entrer*, *sortir*, *descendre*, *transiter...* ainsi que celui de la conversation: *dialogue*, *dire*, *interroger*, *répondre...*).

Cette tendance ne s'est pas démentie depuis l'époque des gros ordinateurs: avec l'avènement de la micro-informatique, l'ordinateur a remplacé la machine à écrire, le calculateur, en incorporant leur vocabulaire. Il s'agit en partie de

métaphores: les presse-papiers, les corbeilles, les bibliothèques, les sauts de lignes ne sont pas à comprendre de manière littérale: on fait **comme** avec un presse-papiers... etc. La démarche sémantique de création de nouveaux sens à partir de vocabulaires existants ne s'est donc pas arrêtée avec la troisième génération des ordinateurs. On peut penser que les autres stratégies, la dérivation, la composition (surtout savante), abondamment mises à profit dans les années soixante, continuent de fournir le complément du renouveau lexical.

Avant d'examiner le rôle qu'a joué le modèle linguistique anglais dans ces stratégies, il convient de dire un mot de la constitution matérielle de l'informatique en tant que science et surtout en tant que technique. Certes, il s'agit d'une grande nouveauté, mais, comme pour les chemins de fer, la plupart de ces éléments étaient déjà connus, certains de longue date (les opérations mathématiques), d'autres depuis peu (la cybernétique). De même, la micro-informatique a beaucoup emprunté aux techniques plus anciennes: le clavier du micro-ordinateur est presque celui de la machine à écrire, et il est donc normal d'employer toujours le même vocabulaire. Encore une fois, cette réexploitation a pu se faire en français, parce qu'il existait déjà un vocabulaire de la technologie précédente.

Le recyclage de la technologie va donc de pair avec le recyclage du vocabulaire, mais en partant d'une technologie relativement avancée et d'un vocabulaire relativement évolué. La question se pose maintenant de savoir comment cela s'est fait, et nous évoquerons trois facteurs d'ordre différent. Le rôle de l'anglais dans l'adaptation du vocabulaire, celui des traducteurs et des auteurs de dictionnaires et de lexiques, et celui des instances de normalisation.

La métasémie largement exploitée par le français dans la constitution du vocabulaire de l'informatique est-elle d'origine anglaise ou de fabrication propre? La réponse doit être circonstanciée parce qu'on peut toujours arguer d'un parallélisme attribuable à des facteurs non linguistiques: les mêmes circonstances pourraient logiquement appeler le même genre de métasémie. L'influence de l'anglais était importante: d'une part, on retrouve les mêmes métasémies: les «emprunts» à la langue générale, à l'électricité ainsi que la quasi-totalité des anthropomorphismes sont les mêmes en anglais. On les retrouve d'ailleurs dans les autres langues européennes. En outre, on constate une forte similitude de forme, due à l'exploitation du fonds gréco-latin.

Cette influence n'est pas systématique cependant: certaines métaphores employées en français n'ont pas d'équivalent métaphorique en anglais (*aiguillage*), et à l'inverse, de nombreuses images anglaises ne sont pas reprises en français (*dump* ← *décharge*; *bootstrap* car faisant appel à un idiotisme, etc). Il se peut aussi que la métaphore, une fois installée dans la communauté linguistique, puisse s'exploiter de façon autonome. C'est le cas des virus informatiques: une fois établi le parallèle entre le programme parasite et le virus, toute la panoplie de comparaisons suit.

De même, au niveau de la dérivation et de la composition, y compris savante, bien plus abondantes que pour les chemins de fer, on relève une ressemblance formelle frappante entre l'anglais et le français. Dans la très grande majorité des cas il s'agit de constructions parfaitement bien formées en français, quoiqu'avec une nette augmentation de types marginaux, tels que la composition nom + nom (*langage machine*, *mémoire-tampon*), qui, comme le fait remarquer Michael Picone (1988), existaient déjà en

français, mais dont la fréquence a été accrue à cause de l'influence anglaise, entre d'autres facteurs. Les ressemblances entre le français et l'anglais, souvent facteur de fertilisation terminologique, peuvent néanmoins aboutir à des entorses morphologiques, syntaxiques et sémantiques. Comme exemple des premières, on relève dans le corpus de 1970 des formes comme *addeur* (← *adder*), depuis remplacées.

Quel est le rôle alors des professionnels de la traduction, et des auteurs de dictionnaires? La question n'a pas de réponse facile.

C. Marcellesi le minimise quelque peu, en disant qu'ils viennent bien tardivement et ne traduisent pas toujours bien: les premiers lexiques en effet étaient des traductions très littérales de l'anglais. On note néanmoins une très nette différence entre le discours anglicisé des pupitreurs de 1970 et celui de l'utilisateur de la micro-informatique vingt ans plus tard. En effet, les fabricants ont su profiter des traductions (souvent des traductions «maison») pour présenter leurs produits à un public plus large. L'anglais sert encore, mais davantage pour la connotation, comme par ailleurs dans la publicité en général.

Les organismes dont nous avons fait mention plus haut ainsi que les commissions ministérielles de terminologie ont joué un rôle également si l'on compare la première liste de termes de la CMT de l'informatique. On note, à plus de dix ans d'intervalle, que la quasi-totalité des termes proposés s'emploie régulièrement aujourd'hui. Une rapide comparaison avec l'allemand (Humbley 1988) révèle une situation bien différente. D'une part, la Commission s'est attaquée exclusivement à un nombre restreint de termes de base, posant problème et susceptible d'une large diffusion. On fait donc abstraction de tous les autres problèmes moins voyants, normalisés consciemment par

l'Afnor ou d'autres organismes similaires, ou inconsciemment par un consensus professionnel. Il semblerait qu'en allemand, cette deuxième catégorie ait connu le même sort qu'en français, mais que les termes très en vue (*computer, hardware, software...*) n'aient pas eu de révision en l'absence de tout organisme officiel.

4 Conclusion

Un danger de la notion de « besoin lexical » serait de supposer que le besoin par rapport à un objet pourrait être satisfait par un mot: de ce fait, il ne donnerait plus lieu à évolution. En réalité, la stabilité n'est jamais complète, le besoin se renouvelle sans cesse. Plus justement, on replacera la question du lexique spécialisé parmi les différents modes de référencement que sont la périphrase ou la dénomination approchée, avec leurs procédés: néologismes de sens, de forme (composition, dérivation), emprunt lexical (langue ancienne ou moderne), et leurs combinaisons. Nous avons vu que le besoin lexical est satisfait par des moyens variés et variants.

Le « terme » se définit au contraire par un idéal de biunivocité, de stabilité (non-variation), et de faible synonymie. Le terme est une norme autant qu'un mot, il a sa place dans les nomenclatures techniques, dans des usages réglementés ou réglementaires et administratifs: en tout cas, dans des types de discours spécifiques.

C'est pourquoi aussi la terminologie, comme « noyau dur » du lexique, est une partie de la langue où l'on envisage facilement une intervention consciente et volontariste; les interventions sur

d'autres aspects sont plus problématiques.

Dans les cas où l'on ne dispose pas d'un terme normalisé, mais de mots divers (y compris des emprunts), il est donc plus juste et plus précis de parler de « besoin terminologique » que de « besoin lexical ». Les deux notions ne doivent **pas être confondues**.

Une dernière question se pose concernant la « portabilité » des mécanismes que nous venons de décrire. Dans quelle mesure l'expérience française est-elle transposable à d'autres langues? Nous avons vu que l'adaptation du vocabulaire anglais des chemins de fer et de l'informatique s'est faite non sans mal en français malgré certains atouts, dont le premier est la présence d'un vocabulaire scientifique et technique déjà considérable. En l'absence de ce formidable acquis, la tâche des adaptateurs, puis des normalisateurs, aurait été bien plus difficile. Les ressemblances entre l'anglais et le français, parfois source d'interférences négatives, sont à inscrire globalement à l'actif de l'adaptation. En outre, la présence de nombreux groupes soucieux d'une francisation rapide de ce vocabulaire, à l'intérieur et à l'extérieur des sociétés d'informatique, a permis un degré d'assimilation supérieur à celui des autres langues européennes.

La cause n'est pas perdue pour les communautés linguistiques n'ayant pas cette tradition scientifique et technique. Certaines ressources linguistiques, y compris sémantiques, sont parfaitement exploitables, comme toute la gamme des métaphores anthropomorphiques que nous avons examinées. De même, l'emprunt peut être aménagé selon les traditions de la langue concernée. Comme pour l'enseignement, c'est la base qui doit être mise en place de façon prioritaire avant l'adaptation de techniques plus sophistiquées.

Bibliographie

Marcellesi (Christiane), 1972: *Approche synchronique du vocabulaire de l'informatique* (3^e génération), Thèse de troisième cycle, dactylographiée, Université de Paris X, 822 p.

Marcellesi (Christiane), 1973: « Le langage des techniciens de l'informatique: quelques aspects de leur vocabulaire écrit et oral », dans *Langue française*, 17, p. 59-71.

Humbley (John), 1987: « L'emprunt sémantique dans la terminologie de l'informatique », dans *Meta*, 32, p. 321-325.

Humbley (John), 1988: « Comment le français et l'allemand aménagent la terminologie de l'informatique », dans *Banque des Mots*, p. 85-148.

Humbley (John), « La néologie dans la planification terminologique », dans *Actes du colloque: « Éléments pour une théorie de la standardisation linguistique »*, Académie suisse des sciences humaines, à paraître.

Picone (Michael David), 1988: *De l'anglicisme et de la dynamique de la langue française*, Atelier national de reproduction des thèses, Université de Lille III, Lille. 405 p.

Prince (Violaine), 1988: « La pidginisation du français par le jargon américain de l'informatique », dans *Actes du II^e colloque du GEPE*, Strasbourg, p. 15-37.

Wexler (Peter), 1955: *La formation du vocabulaire des chemins de fer en France, 1778-1842*, Genève, Lille, Droz, 160 p.

Jean-Michel Eloy,
Délégation générale à la langue
française
et
John Humbley,
Centre de terminologie et de néologie,
Paris,
France.