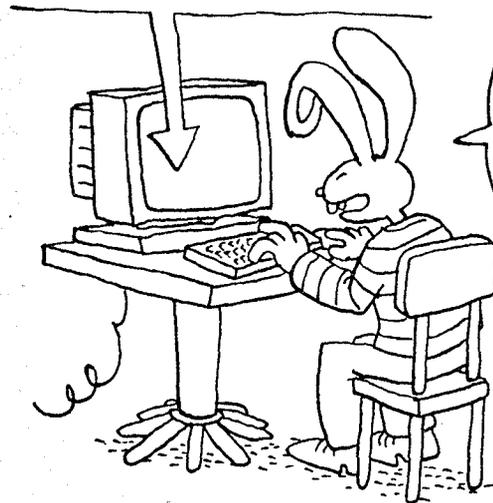


C'est compliqué d'écrire
avec ce système sans accent
et puis ça n'est pas très
élégant ni très style.



GRMML...
CE COURRIEL
DE PROTESTATION
RISQUE DE NE PAS
ÊTRE TRÈS LISIBLE.

comme informatique, industries de la langue et Internet

par Jean-Claude Corbeil

Je me souviens très bien de mon premier contact avec l'ordinateur. C'était au début des années cinquante, au Centre de calcul de l'université de Montréal, car, à cette époque, on croyait vraiment que l'ordinateur était destiné surtout, pour ne pas dire uniquement, aux calculs mathématiques. Pourtant, j'étais inscrit en linguistique et notre directeur, M. Jean-Paul Vinay, nous entraînait dans les premiers essais de traitement informatique de la langue, embryon des travaux en vue de la traduction automatique (on était très confiant à l'époque !) de l'anglais vers le français.

C'était une immense machine, une *bécane* dirait-on aujourd'hui, qui occupait à elle seule toute une grande salle, une machine sensible à la chaleur et aux courants d'air, qui ronronnait à condition que ses exigences climatiques soient respectées, une vraie plante de serre qu'il fallait aller visiter en petit groupe, pour ne pas perturber son climat !

Des grands prêtres officiaient auprès d'elle, les informaticiens. Eux seuls avaient les connaissances qui leur conféraient le droit et le pouvoir de la servir, l'intuition pour analyser ses caprices et l'amener à poursuivre ses réponses aux demandes de ses suppliants. Car, capricieuse comme

une princesse, elle s'arrêtait souvent, sans raison apparente, dans un silence absolu, laissant en plan la lente élaboration de la réponse à la question qui lui avait été posée, comportement plus angoissant encore pour les pauvres humains que les réponses ambigües de la Pythie de Delphes. Des officiants, les programmeurs, lui apportaient le menu dont elle se nourrissait et qui lui donnait vie, des centaines de cartes perforées avec précision et placées dans un ordre implacable, la moindre erreur provoquant la colère et la paralysie de la déesse ordinateur. Une armée d'assistants secondaient les programmeurs et cuisinaient, sous leur surveillance, le détail des menus, conçus ligne à ligne, une consigne rigoureusement univoque par ligne, consigne qui était ensuite transcrite, par les servants des machines à perforer, sous la forme d'une succession de vides et de pleins, sur des cartes digestibles par l'ordinateur. Si tout lui était agréable, l'ordinateur acceptait de s'exécuter et de donner suite aux suppliques des chercheurs.

Aujourd'hui, presque cinquante ans plus tard, je suis devant mon portable, j'écris ce texte, dans des conditions encore plus faciles que si j'utilisais un crayon et une feuille de papier : j'efface, je corrige, je déplace des segments de texte, je mets en réserve une phrase, au cas où je voudrais la reprendre, je donne libre cours sans hésitation aux méandres de l'inspiration et je suis le maître absolu de mon texte, contenu et présentation. À tout moment, je peux imprimer ce que j'ai écrit, pour me relire, pour profiter de ce regard à distance qu'accorde à l'œil et à l'esprit le texte imprimé. Je n'ai aucune conscience de ce qui rend possible cette écriture, aucune connaissance des détails de la programmation grâce à laquelle le langage informatique assure le traitement de ma langue, le français, dans le respect intégral de toutes ses caractéristiques. La programmation me parvient maintenant sous la forme d'un logiciel, ici un logiciel de traitement de texte, lui-même inséré dans le programme de mon ordinateur, dont il ne me reste plus

qu'à maîtriser les modes d'emploi. L'ordinateur est maintenant devenu convivial, mot clé de la micro-informatique contemporaine qui renvoie à la fois à une grande facilité d'utilisation et à la possibilité pour chaque utilisateur de personnaliser la configuration de son système selon ses besoins et ses habitudes de travail. De plus, les ordinateurs sont plus puissants qu'autrefois et peuvent contenir et traiter simultanément plusieurs applications différentes du fait qu'ils disposent d'une grande capacité de mémoire. Nous sommes très loin maintenant des modes d'utilisation de l'ordinateur des débuts de l'informatique, alors qu'à cette époque nous dépendions tous des programmeurs pour la moindre opération.

La révolution informatique a eu lieu. Mais il n'est pas certain que la révolution des mentalités se soit produite en même temps et au même rythme, soit selon les tranches d'âge, soit selon les pays. L'image de l'informatique est souvent demeurée celle d'autrefois, un univers de spécialistes, trop complexe pour le commun des mortels, d'un abord exigeant une information et une formation techniques difficiles à acquérir et à maîtriser. Cette image n'est plus vraie, elle est même la caricature de la réalité. Mais, pour le découvrir et s'en rendre compte, il faut y mettre du temps, le temps de se familiariser avec les commandes d'un micro-ordinateur et avec celles des logiciels d'application. Mais pourquoi perdre du temps, réplique-t-on, à s'initier à l'informatique alors que chacun possède déjà des habitudes de travail qui le servent bien. Pourquoi changer ? Question fondamentale, universelle dans ce monde en mutation rapide. Réponse banale : ou bien le changement est volontaire, ou bien il est imposé de l'extérieur, par l'évolution de l'environnement de travail et de vie. Question vitale également, parce que l'informatique s'est rapidement répandue et qu'elle a pénétré tous les segments d'activité et, de plus en plus, la vie privée elle-même.

La révolution informatique s'est produite d'abord et surtout aux États-Unis, donc en langue anglaise, et sous tous

les aspects de ce nouvel univers devenu rapidement un domaine commercial extrêmement lucratif : conception, fabrication et commercialisation des micro-ordinateurs, conception et commercialisation des logiciels, grande variété de manuels informatiques disponibles pour tous les produits, à bas prix, début de la standardisation des produits par l'adoption des premières normes techniques, conception et généralisation d'un réseau de communication entre les organismes de l'armée américaine d'abord, entre les centres de recherche et les universités ensuite et dont le succès a été tel qu'il s'est transformé en un réseau ouvert à tous, le réseau des réseaux d'aujourd'hui, accessible par Internet. Silicon Valley est le symbole de cette révolution, le lieu le plus mythique de cet univers, l'endroit où les millionnaires de l'informatique se comptent par centaines.

Depuis lors, toutes les langues sont à la remorque de l'anglais et cherchent à s'assurer une présence en informatique. Selon les données d'une étude du Conseil de la langue française du Québec (1998), le plus grand nombre et la plus grande variété des logiciels sont actuellement en langue anglaise : en 1994, il y avait plus de 700 logiciels de traitement de la langue anglaise, contre 425 pour le français, 350 pour l'allemand, moins de 300 pour l'espagnol et l'italien. Les utilisateurs de langue française estiment, en conséquence, que les applications en français sont trop peu nombreuses, qu'elles ne couvrent pas toute la gamme des besoins, qu'elles sont souvent des traductions-adaptations de logiciels américains, en retard par rapport aux dernières versions du produit original. Cet état de choses a généré un très net préjugé défavorable à l'égard des produits en langue française, véritable handicap au développement d'une industrie et d'un marché du logiciel en langue française.

La manière d'évaluer l'omniprésence de l'anglais en informatique est de considérer le contenu du Web sous l'angle de la langue : 80 % de tous les sites, toutes catégo-

ries confondues, sont en anglais, contre 2 % en français (données de 1998). À très brève échéance, la présence de l'anglais devrait diminuer à plus ou moins 40 % lorsque des sites seront créés dans les langues des divers pays, au fur et à mesure qu'ils se brancheront sur le réseau. Car la forte présence de l'anglais tient au fait que les pays anglophones ont été les premiers à fréquenter le réseau : au début de 1997, environ 66 % des hôtes d'Internet provenaient des États-Unis, du Canada anglais, d'Australie ou de Grande-Bretagne. La concurrence informatique entre les langues s'ajoute donc à toutes les autres manifestations de cette concurrence dans un monde de plus en plus ouvert mais qui se rétrécit chaque jour un peu plus. Une conviction commune s'est imposée peu à peu : les langues qui ne seront pas dotées d'un traitement informatisé ne pourront plus desservir correctement et complètement leurs propres locuteurs ; elles perdront petit à petit de leur utilité et de leur prestige auprès d'eux et leur pouvoir d'attraction sur le plan international diminuera.

La langue française, dans cette concurrence, possède des atouts sur lesquels elle peut s'appuyer pour améliorer sa position.

La langue française, tout comme l'anglais ou l'espagnol, est une langue de grande diffusion qui donne accès à un marché composé de millions de consommateurs potentiels de produits informatiques. Ce marché est en pleine expansion dans les pays industrialisés de langue française, au Québec et en Europe (France, Communauté française de Belgique, Suisse, Luxembourg). Il sera encore plus grand quand les conditions faciliteront la pénétration de l'informatique dans les pays de la francophonie, pays en émergence ou en développement. Encore faut-il, pour que ce marché se maintienne et s'étende, que les produits en langue française soient de haute qualité, qu'ils couvrent toute la gamme des besoins ; sinon, les produits en langue anglaise continueront de bénéficier de leur excellente réputation et, en conséquence, domineront encore le marché.

Les concepteurs-producteurs de système informatique sont très sensibles à cet aspect des choses. Leur objectif est qu'il y ait au moins un ordinateur par famille pour pouvoir ainsi, par la suite, vendre les produits et services qui exigent la possession et l'utilisation d'un appareil. Cet objectif ne peut être atteint, et ils en sont parfaitement conscients, que si les ordinateurs fonctionnent dans la langue de la famille, en conséquence dans la langue ou les langues du pays. Il faut donc s'attendre à une profonde modification de la stratégie marketing des grandes sociétés d'informatique. Un virage s'amorce qui modifiera la manière de procéder à la programmation des systèmes. Par exemple, Microsoft serait sur le point de séparer langue et programmation. La programmation se ferait en langage machine et comporterait des points d'ancrage où insérer un texte. Les textes seraient disponibles dans les différentes langues des marchés visés. Au moment de produire une version d'un produit pour un marché linguistique donné, le texte dans cette langue serait appelé par la programmation et se placerait à l'endroit voulu dans l'interface du produit. La synthèse programmation-texte serait ainsi réalisée pour chaque langue sans qu'il soit nécessaire de retoucher à la programmation. C'est ce qu'on appelle la localisation. Dans cette perspective, il serait même possible de prendre en compte la variation linguistique d'une même langue, par exemple l'espagnol, l'anglais et le français, toutes langues dont les usages nationaux comportent de légères différences qui ne gênent en rien l'intercompréhension, mais qui sont pertinentes en traitement automatique de la langue.

Le traitement informatique du français écrit est aujourd'hui possible. Les quelques problèmes qui se posent découlent de la variation lexicale du français, c'est-à-dire de la nécessité d'insérer dans certains logiciels des mots ou des sens qui font partie de la norme de la langue d'une communauté linguistique donnée, par exemple dans les correcteurs orthographiques. Il faut savoir qu'un correcteur

orthographique fonctionne par comparaison d'un lexique présélectionné avec le lexique du texte à examiner. Si le texte contient des mots qui ne figurent pas dans le dictionnaire du correcteur, un message apparaît qui demande à l'auteur de valider l'existence du mot. Plus il y en a, plus le temps de correction s'allonge. On comprend donc que les concepteurs de ces logiciels ont tendance à les adapter à la plus grande clientèle possible, en puisant dans les travaux de description lexicale du français des diverses régions ou pays de la francophonie. Ils poursuivent, pour ainsi dire, le même rêve d'un dictionnaire universel du français. Le traitement du français oral est nettement plus problématique, quoique, sans doute, dans le même état que celui de l'anglais. La difficulté est la même : comment faire qu'un ordinateur puisse éliminer les caractéristiques individuelles de la voix (timbre, prosodie, prononciation, vitesse d'élocution) pour atteindre la langue elle-même dans sa généralité, ce que tout locuteur fait naturellement en passant d'un interlocuteur à l'autre.

Le problème est ailleurs que dans le traitement informatique de la langue française. Globalement, le défi consiste à rendre disponible en langue française une panoplie aussi étendue d'outils informatiques qu'en langue anglaise pour bloquer le glissement des francophones vers l'informatique en anglais et pour consolider le marché en langue française. Relever rapidement et efficacement ce défi est urgent et vital, au moment où l'usage de l'informatique gagne la vie privée après avoir envahi la vie des entreprises et des administrations.

Le marché de l'informatique en français a pris de l'ampleur ces dernières années dans les entreprises d'abord, puis dans les milieux de la recherche industrielle ou universitaire, quoique, dans ce dernier cas, la situation varie beaucoup, enfin dans les milieux professionnels. La progression est beaucoup plus lente dans la vie privée. Au Canada, par exemple, les ménages de langue française ont nettement moins tendance à s'équiper d'un ordinateur ou à

se brancher sur Internet que les ménages de langue anglaise. Si on considère l'ensemble de la francophonie, le marché potentiel en langue française est encore plus considérable, mais les obstacles sont nombreux, économiques, à l'évidence, mais aussi techniques : stabilité de l'alimentation électrique et téléphonique, cout et entretien du matériel, support technique déficient, enseignement de l'informatique mal assuré. Ce marché des pays de la francophonie n'est pas aussi captif que celui des pays de langue maternelle française : les utilisateurs ont le choix entre les produits de langue française ou anglaise et iront vers ce qui est le plus performant et le moins couteux.

Les produits informatiques se partagent en différentes catégories où la présence du français varie considérablement.

Les micro-ordinateurs sont livrés avec un programme en langue française, que ce soit sur Macintosh ou sur la plateforme PC, sauf, parfois, pour les machines bas de gamme.

En général, les produits grand public, de consommation de masse, existent en français, au même niveau de qualité que les produits en langue anglaise : traitements de texte, chiffriers, correcteurs orthographiques, bases de données. Le courrier électronique peut être assimilé à une application grand public, puisque c'est le plus utilisé des services Internet. Il n'est pas toujours évident d'y faire usage du français : l'emploi des accents est impossible dans les adresses à cause des conventions à la base d'Internet ; dans le texte des messages, on peut accentuer normalement les lettres, quoique certains serveurs n'admettent pas les accents et affichent des signes cabalistiques en lieu et place des lettres accentuées. Le courrier électronique a pénétré dans l'usage sous son appellation américaine *e-mail*, que l'on voit partout lorsque sont indiquées les coordonnées d'une personne ou d'une entreprise. On tente aujourd'hui, en France, de lui donner une apparence française à partir de la prononciation du terme anglais, sous la forme de *mél*, avec un accent aigu incongru et impossible devant *-l*. Les

Québécois, à leur habitude, ont créé un néologisme, *courriel*, qui a l'avantage d'évoquer *courrier*, son proche parent sémantique, et d'entrer dans la série des termes informatiques en *-el* : *logiciel*, *ludiciel*, *didacticiel*, etc.

La situation du français dans les produits spécialisés (logiciels d'infographie, de comptabilité, de montage, etc.) est nettement plus difficile. Ils sont créés pour remplir des tâches d'un niveau élevé de complexité et destinés à des utilisateurs qui sont davantage sensibles à la performance d'un logiciel pour les aider à accomplir leurs travaux qu'à la langue dans laquelle il fonctionne. D'autre part, le marché des logiciels spécialisés est restreint et les utilisateurs potentiels sont plus nombreux en langue anglaise qu'en tout autre langue. On observe donc qu'ils sont le plus souvent, au départ, en langue anglaise et qu'ils sont ensuite traduits en français quand ils rejoignent un plus grand marché, lorsqu'en définitive ils se vulgarisent. C'est ce qui se passe actuellement pour les logiciels d'infographie et de mise en pages assistée, depuis le moment où l'édition s'est informatisée. Par contre, il arrive aussi que la version française ne soit pas à jour par rapport à la version américaine, qui évolue plus rapidement. D'où le préjugé que les versions françaises sont toujours en retard et donc de moindre qualité.

L'ordinateur sert de plus en plus de support au fonctionnement de produits qui ne sont pas des produits informatiques au sens strict. C'est tout l'univers des cédéroms et des jeux vidéo, mais surtout d'Internet. C'est à la fois du domaine de l'informatique et de celui de l'édition, conception et commercialisation des produits sur support CD et du domaine de la création de sites sur Internet. Il y a encore quelques années, la majeure partie de ces produits, de même que la très grande majorité des sites Internet, étaient en langue anglaise. Aujourd'hui, le nombre des produits ou des sites en langue française augmente rapidement, soit par traduction dans le cas des jeux vidéo, soit par édition de produits originaux sur cédérom, soit par création de sites

en langue française. On peut donc considérer l'avenir avec plus d'optimisme.

L'édition de didacticiels est difficile en langue française pour au moins deux raisons principales. D'une part, le marché est relativement restreint à cause de sa fragmentation entre systèmes pédagogiques et programmes d'études différents d'un pays à l'autre. Et comme un didacticiel est destiné à s'insérer dans l'enseignement comme matériel pédagogique d'appoint, il est difficile de définir un contenu qui soit adapté à toutes les situations d'enseignement, surtout au niveau du primaire et du secondaire. D'autre part, l'informatique a peu pénétré dans les classes et on ne sait pas encore très bien quel rôle elle pourrait y jouer. Autant on est convaincu qu'il est important d'initier très tôt les enfants à l'usage de l'ordinateur, autant on ne sait trop comment insérer l'ordinateur dans l'activité normale de la classe. Ici, tout est encore pour ainsi dire à l'état expérimental. Par contre, la concurrence de l'anglais est, dans ce domaine, minime puisque l'interface d'un didacticiel doit impérativement se présenter dans la langue maternelle des élèves, ou dans la langue de fonctionnement du système scolaire.

L'outil le plus stratégique de nos jours est Internet, le web, le réseau des réseaux, la toile (qu'il nous faut tisser en français !) qui a ouvert toutes grandes les portes aux nouvelles technologies de la communication et de l'information, les NTCI et, depuis peu, au commerce électronique qui se développe rapidement.

Pour avoir accès au réseau, il faut posséder un ordinateur, première condition essentielle.

Internet est à la fois un mode de communication (courrier électronique et groupes de discussion) et un mode d'information (mise en ligne de renseignements et accès à ces renseignements). La communication s'établit par ligne téléphonique ou par câble entre l'utilisateur et un intermédiaire, le serveur, qui lui-même assure par les mêmes moyens le contact avec le réseau. D'où une seconde condi-

tion d'accès à Internet : il faut être branché, c'est-à-dire s'abonner à un serveur. La manière dont est tarifée la communication (que ce soit le coût du téléphone, du câble ou du serveur) facilite ou freine l'accès à Internet, selon que le tarif est forfaitaire ou à la minute.

L'information est mise à la disposition des usagers sous forme de sites accessibles par une adresse électronique, tout comme dans la poste traditionnelle. On accède à cette information soit en composant l'adresse du site approprié, soit en repérant les renseignements au moyen de logiciels d'exploration, les fameux moteurs de recherche, qui balaisent le contenu du web avec des mots clés permettant de cerner la question, sorte « d'hameçons notionnels » susceptibles d'attraper les documents qui pourraient correspondre au sujet délimité par les mots clés.

Voilà, en quelques mots, toute la problématique d'Internet : un mode d'accès (ordinateur + ligne téléphonique ou câble + serveur), un contenu (quantité et qualité de l'information disponible en français) et un mode de recherche (la performance des moteurs de recherche en langue française).

N'importe qui peut créer un site sur Internet et y mettre ce qu'il veut, la liberté est totale, l'encadrement juridique encore très difficile. En principe, toutes les langues sont admises, sans restriction, sauf dans les adresses. La place de l'anglais diminue au même rythme que sont créés des sites dans d'autres langues. La place du français augmente rapidement, mais aussi celle des autres langues. Ce plurilinguisme d'Internet devient de plus en plus un obstacle : que faire face à une information qui nous parvient dans une langue qui nous est inconnue ? Les logiciels de décodage linguistique de l'information (appelés abusivement logiciels de traduction automatique) se développent peu à peu. Ils permettent au moins de juger si un document est intéressant et vaut la peine d'être correctement traduit. Pour

langue française, ce type d'aide est disponible et progresse en qualité.

Les propriétaires de sites peuvent être des individus, des ministères, des administrations municipales, des organismes de toute orientation et de tout statut, des entreprises, des associations, des comités d'action pour n'importe quelle cause ou évènement, des syndicats, etc. Il y a donc de tout sur Internet, le pire et le meilleur : des documents gouvernementaux, des renseignements touristiques, des catalogues (bibliothèques, maisons d'édition, commerces), des normes techniques et commerciales, des journaux et revues, de la pornographie, de la littérature haineuse, des films, de la musique, tout, absolument tout. La qualité de l'information disponible sur le web est donc extrêmement variable, hors contrôle, laissée à l'appréciation de qui la consulte. C'est sûrement l'aspect le plus troublant d'Internet. Son utilisation exige un sens critique très aigu, qui suppose que chacun possède suffisamment de connaissances pour apprécier l'information transmise et la transformer en connaissance car, pour devenir connaissance, l'information doit s'intégrer dans un ensemble. L'école a et aura de plus en plus la responsabilité d'initier les enfants à cette opération intellectuelle.

Le développement actuel d'Internet, dans toutes les directions à la fois, est si rapide et si anarchique que personne ne peut prédire ce qu'il en adviendra, pas même les grands maîtres et gestionnaires du système. Une seule chose est certaine : Internet sera de plus en plus présent dans la vie collective et dans la vie privée de chacun.

L'ordinateur, comme outil pour accomplir une grande variété de tâches, comme moyen d'accès à l'information, comme lieu du commerce électronique, comme mode de communication est là pour rester et il sera omniprésent dans la vie du XXI^e siècle.

On ne peut tourner le dos à cette révolution, pas plus que, naguère, il a été possible de tourner le dos à l'écriture

et à l'imprimerie. Impossible aussi d'admettre que le français, notre langue commune, ne soit pas informatisée et que les produits, tous les produits que nous manipulons, ne soient pas disponibles en langue française. Tout le monde est responsable de l'informatique en français, les gouvernements pour favoriser le développement d'une industrie et d'un marché en langue française, chacun d'entre nous comme consommateur, pour exiger des produits de qualité en langue française et pour faire nombre, car la taille du marché de langue française est probablement la condition la plus essentielle du développement de l'informatique en français.