

**L'Informatique:
l'Instrument de l'Uniformisation ou de la Préservation
de la Diversité des Langues et des Cultures?¹**

Cette question doit être traitée en deux niveaux: (1) au niveau national; et (2) au niveau international.

1. *Au niveau national*, les libertés de chaque citoyen et le rôle de la société civile devant l'Etat;
2. *Au niveau internationale*, le droit de préservation de la diversité culturelle de chaque nation devant le géant mondial qui agit comme un pôle de pouvoir d'uniformisation, soit politique si nous considérons les Etat-Unis, soit économique si nous considérons les sociétés multinationales. Ainsi, à l'âge de la société informatisée, l'inévitable *globalisation* se confond avec l'américanisation selon certains analystes. Je voudrais vous rappeler ici une phrase du Premier Ministre turc Ecevit au sommet de Davos, en Suisse, en 2000: *Nous devons humaniser la globalisation.*

Ces forces dynamiques de la société informatisée agissent certainement mutuellement à ces deux niveaux, national et international.

Le 20 décembre 1976, le président de la République Française Mr. Valéry Giscard d'Estaing écrivait à Simon Nora, l'Inspecteur Général des Finances, pour la remise d'un rapport au sujet de l'informatisation de la société, les lignes suivantes:

“Le développement des applications de l'informatique est un facteur de transformation de l'organisation économique et sociale et du mode de vie. Il convient que notre société soit en mesure, à la fois, de promouvoir et de le maîtriser, pour le mettre au service de la démocratie et du développement humain.”

Ainsi le célèbre rapport Simon Nora/Alain Minc intitulé *l'Informatisation de la Société* était remis au Président le 20 janvier 1978 et fut publié par la Documentation Française.

En voici 2 constats importants:

1. *Si la France ne trouve pas de réponse correcte à des défis graves et neufs, ses tensions intérieures lui ôteront la capacité de maîtriser son destin.*

L'informatisation de la société est au cœur de la crise. Elle peut l'aggraver, ou contribuer à la résoudre.”²

¹ La conférence donnée par le Prof. Dr. Aydın Köksal, dans la salle de l'audience de l'Ambassade de France à Ankara, Turquie, le 17 Janvier 2002.

² Simon Nora/Alain Minc, *L'informatisation de la Société*, Rapport au Président de la République Valéry Giscard d'Estaing le 20 Janvier 1978, Documentation Française, Paris 1978, p. 9.

2. *La réflexion sur l'informatique et la société renforce la conviction que l'équilibre des civilisations modernes repose sur une alchimie difficile*: le dosage entre un exercice de plus en plus vigoureux de l'Etat et une exubérance croissante de la société civile. L'informatique, pour le meilleur ou pour le pire, sera un ingrédient majeur de ce dosage (p. 5).

Moi je trouve que cette alchimie est certes plus difficile en ce qui concerne l'équilibre entre les nations et civilisations mondiales: Est-ce que le "big brother" Etats-Unis va uniformiser et maîtriser tout, pour mettre fin à toute sorte de diversité? Etats nationaux, langues, cultures nationales vont-ils tous disparaître?

En 1980, en France, "La Commission Nationale de l'Informatique et des Libertés" publiait le document intitulé *Informatique et Libertés*,³ comprenant les textes législatifs et les textes d'application. La loi du 7 janvier 1978 pose le principe suivant (Article 1^{er}):

"L'informatique doit être au service de chaque citoyen. Son développement doit s'opérer dans le cadre de la coopération internationale."

"La commission nationale de l'informatique et des libertés" est créée en tant qu'une autorité administrative *indépendante*, composée de dix-sept membres nommés pour cinq ans (Art. 8).

La loi précise aussi:

"Il est interdit de mettre ou conserver en mémoire informatisée, sauf accord exprès de l'intéressé, des données nominatives qui, directement ou indirectement, font apparaître les origines raciales ou les opinions politiques, philosophiques ou religieuses ou les appartenances syndicales des personnes (Art. 31).

"Le titulaire du droit d'accès peut obtenir communication des informations le concernant. La communication, en langage clair, doit être conforme au contenu des enregistrements (Art. 35).

Retournons-nous maintenant en Turquie:

En 1968, jeune ingénieur électronicien, ayant découvert le pouvoir de l'informatique comme un instrument efficace de la transformation de la société, je déclarais:

"Nous utiliserons l'industrie informatique comme un instrument pour le développement de la Turquie", en 1968, donc dix ans avant le rapport Nora/Minc en France.

Le 22 Avril 1971 j'établissais "l'Association de l'Informatique de Turquie", avec mes 7 collègues informaticiens de l'Université de Hacettepe; et la "Commission de l'Informatique" au sein de l'Organisation de Planification de l'Etat le suivait en 1974, dont j'étais membre jusqu'en 1989.

Comme en France, nous avons bien constaté en Turquie combien il était difficile de prévoir et de pourvoir la préservation de notre langue, de notre culture, et de notre

³ *Informatique et Libertés*, Documentation Française, Paris, 1980.

indépendance politique et la création de notre technologie des systèmes logiciels. Il a fallu créer notre terminologie de l'informatique, (grosso modo 2.500 néologismes comme *bilgisayar* (ordinateur), *yazılım* (logiciel), *bilişim* (informatique) etc.), créer des écoles d'ingénieurs informaticiens, créer des maisons pour la production industrielle des systèmes d'application de logiciel, surtout en matière des systèmes d'information de gestion et la création d'une industrie des communications et d'autres équipements électroniques.

Nous avons, certes, réussi jusqu'à présent, de créer du néant une industrie et une société informatisée turques débutantes qui pourraient à la rigueur nous conduire vers la préservation de notre langue, culture et industrie comme une diversité, à l'échelle internationale vis-à-vis de la dominance croissante Américaine ou Anglosaxone. C'est aussi bien vrai pour la France. Mais la question qui est encore devant nous, devient de plus en plus critique avec l'avènement de la *globalisation*.⁴ La globalisation ou la *mondialisation* si vous préférez, qui n'est, à mon avis, qu'une conséquence de l'informatisation de la société interconnectée grâce au réseau global d'information "Internet" d'une part et de l'effondrement de l'Union Soviétique d'autre part. Ainsi le monde est devenu *un village électronique unipolaire*.

Un amis français m'a dit récemment que l'Anglais était devenu d'ores et déjà la langue internationale de science et de commerce, acceptée par tout le monde; donc, il ne faut pas s'en faire du fait que les caractères accentués ne peuvent pas être utilisés sur Internet. Un de mes amis tures m'a dit la même chose au sujet des caractères spécifiques tures comme le *g muet*, le *ç* ou le *ş*. Mon compatriote qui fait ses études universitaires dans un établissement où l'Anglais est utilisé comme langue d'enseignement, non seulement utilise les mots anglais comme "termes scientifiques" dans le contexte de son discours en Turc, mais aussi souhaiterait d'écrire le Turc avec les caractères soit-disant *uniformisés* de l'alphabet anglais, pour utiliser plus aisément Internet. "Il en a marre, de ces caractères", m'a-t-il dit.

Je me souviens maintenant de *1984*, le roman de George Orwell (publié en 1948) où le grand patron "big brother", qui surveillait partout tous ces citoyens (the big brother is watching you!), et qui avait bien uniformisé toutes les diversités, linguistiques, culturelles et même affectives. *La nouvelle-langue* ne comportait plus que les mots choisis par l'Etat, dont les sens étaient bien définis d'une façon uniformisée.

En 1984, l'Allemagne Fédérale et l'Etat Fédéral de Berlin, en collaboration avec le l'OCDE avait organisé une conférence internationale pour le future du défi informatique, à la mémoire de George Orwell. La question était la suivante: Est-ce possible pour nous autres, européens, russes ou chinois, de préserver notre diversité vis-à-vis des deux géants de l'informatisation, les Etats-Unis⁵ et le Japon, à l'égard de l'industrie, économie, culture, langue, collaboration internationale, division de travail, etc. Le Sous-secrétaire du Ministère de l'Intérieur Français, très pessimistement, parlait des moyens limités d'un pays relativement petit comme la France. Tout le monde était pessimistes sauf l'économiste Galbraith, qui pensait que l'Europe avec les italiens qui en faisaient partie, maîtres de la *conception* ou le *design* industriel, ne pourrait jamais accepter sa disparition complète sans trouver une solution quelque soit la gravité des conditions.

⁴ Cf. Les livres de Kennedy et de Minc.

⁵ Cf. *Le Défi Américain* de J-J Servan-Schreiber.

Je pense que nos amis, Français et Turc, désespérés devant Internet se trompent, car ils pourraient utiliser le réseau Internet avec leur alphabet respectif, pourvu que eux et leurs interlocuteurs aient sur leur ordinateur, le petit logiciel de translittération alphabétique et qu'ils veuillent, bien entendu, utiliser leurs caractères à eux.

Disons tout de suite que, sauf dans les adresses, les caractères arabes, hébreux, grecs, cyrilliques (russes), kanji (chinois) ou kanji et kana (japonais) sont aussi bien utilisables que ceux français et turcs.

Nous en sommes venu ainsi aux faits linguistiques.

Depuis l'avènement de l'ordinateur en 1945, les linguistes ont toujours voulu travailler sur le *traitement automatique des langues naturelles*, l'objectif étant l'analyse et compréhension des textes, l'étude statistique de faits linguistiques y compris *la stylistique* et enfin *la traduction automatique*.

N'importe comment, à côté des données numériques, une bonne partie des données était d'ailleurs alphabétiques, comme des noms, des adresses ou des textes.

Ainsi *l'analyse par ordinateur des langues naturelles* est devenue un domaine de recherche scientifique faisant partie du domaine plus large de *l'intelligence artificielle* dont l'un des objectifs est *la traduction automatique*. Mon thèse de doctorat, achevée en 1975, qui était dans ce domaine, était intitulée *A First Approach to a Computerized Model for the Automatic Morphological Analysis of Turkish* (*Une Première Approche à un Model Informatique pour l'Analyse Automatique du Turc*).

Dans ce travail, qui était jadis considéré par un linguiste hongrois comme "presqu'impossible sur une langue agglutinante (comme le Turc, le Hongrois, le Finnois, le Mongol ou le Japonais), au moins aussi difficile que pêcher un petit poisson dans un océan", j'ai été agréablement surpris, je dirais même ébloui, par la structure régulière, mathématique du Turc: 94% des mots agglutinés du texte turc étaient morphologiquement analysés correctement.

A l'heure actuelle un certain succès a été déjà obtenu dans le domaine de la traduction automatique, sans cependant garantir la haute qualité et fiabilité de la traduction. La machine ne produit donc qu'une ébauche qui doit être corrigée et éditée par l'intelligence humaine (d'où l'appellation de *la traduction assistée par ordinateur*).

Evidemment l'analyse des *langues fusionnelles* comme le Français, l'Anglais, le Russe et l'Arabe ou des *langues isolantes* comme le Chinois présentent des difficultés ou des avantages différents à cet égard.

L'article de Mr. Henri Hudrisier, intitulé *Risques Industriels des Langues Alphabétiques Face aux Langues Idéographiques*⁶, qui date de 1987, est une bonne source en langue française. Ce texte montre *l'iconicité de l'écriture* et les difficultés d'établir les liens sémantiques entre les mots et les concepts. Ce problème nécessite la réalisation de ce qu'on appelle en grec un *thésaurus* ou un *trésor* (des concepts).

⁶ Henri Hudrisier: *Risques Industriels des Langues Alphabétiques Face aux Langues Ideographiques*, Paris, 1987 (miméo).

Les mots écrits dans nos langues alphabétiques ne sont que les *icônes* qui représentent les sons articulés de la langue parlée, qui eux-même symbolisent plus ou moins clairement quelques concepts du trésor (des concepts) de chaque langue, pour assurer la compréhension du discours. Il faut donc tout d'abord des trésors plurilingues établis dans les mémoires magnétiques des ordinateurs, et des algorithmes et d'équipement pour accès rapide à chaque concepte. Il faut bien voir que chaque concepte peut faire partie d'un concepte plus large, ou peut en être un membre, ou peut être faite d'autres conceptes ou choses utilisées comme matières premières etc... A cause de la multiplicité de ces liens sémantiques, la structure d'un trésor linguistique est très complexe. Ce qui nécessite donc le traitement parallèle utilisant des systèmes-experts ou des systèmes à base des connaissances (=knowledge-based systems), pour une traduction en-ligne ou simultanée, sur des équipement à très grandes mémoires et très rapides... C'est ce que les japonais appellent *ordinateurs de la 5^e génération*. Néanmoins, l'existence de tels équipements ne garantie pas un entendement rapide et fiable par la machine, de la langue parlée ou écrite, faut-il encore établir sémantiquement les bases des connaissances des systèmes-experts et des trésors linguistiques, concepte par concepte, indiquant par exemple que le concepte de "naître" et "mourir" ne s'appliquent qu'aux vivants (une pierre ne meurt pas), celui de "voler" (comme un oiseau) qui ne s'applique pas à un *chien* peut cependant être utilisé pour les *avions* et *autres moyens de transport aériens* et pour des *personnes* qui les utilisent. Le concepte de "voler" (comme de l'argent) pourrait-il être utilisé ou non pour le cœur d'une personne chérie? Cet exemple montre que *la base des connaissances mémorisée* et le *trésor sémantique* ne sont que des structures artificielles pour simuler ce qui existe dans la tête de chacun de nous, qui sommes des êtres intelligents et cultivés utilisant une ou plusieurs langues pour comprendre, communiquer et traduire.

Le chinois, ainsi que le Japonais et le Coréen n'utilise pas les caractères alphabétiques mais idéographiques. Ce qui correspond, d'après Hudrisier, à *une image archéologique* de la langue. Ainsi, sans représenter les sons articulés, un icône du concepte-même est écrit sur papier. Ceci veut dire que le trésor des concepts se trouve directement dans la mémoire biologique ou dans la tête de chaque personne. Ce qui est intéressant, c'est que grâce à l'éducation, ce même trésor qui couvre à la fois la définition de chaque concepte, les liens sémantiques entre les concepts et les idéogrammes, c'est-à-dire l'écriture, chaque personne apprend pas-à-pas pendant son éducation, les choses, les concepts, l'étymologie, la géographie, l'histoire, la logique, la psychologie ou *la culture* en un mot.

Le chinois est une *langue isolante*. Chaque mot est monosyllabique et ne comporte ni déclinaison ni conjugaison: il n'ya ni préfixe ni suffixe. Les mot sont représentés conceptuellement par des idéogrammes dont le nombre dépasse trente mille. Le sens du discours change selon la place du mot dans la phrase d'après le syntaxe.

Ainsi l'écriture chinoise a beaucoup aidé à l'unité de ce grand peuple, 1.230.000.000 d'habitants aujourd'hui, où plus d'une douzaine de langues sont parlées. Les personnes qui ne se comprennent pas au moyen de la communication orale, peuvent cependant facilement lire l'écriture commune. C'est comme les chiffres arabes -en fait les icônes ou idéogrammes- que nous lisons facilement sans tenir compte des mots pronocés différemment dans nos différentes langues européennes.

De la même façon un japonais ou un Coréen, dont les langues agglutinantes comme le Turc (annexion de plusieurs suffixes de dérivation, de déclinaison ou de conjugaison) sont

à même de lire un journal chinois, utilisant les sons des mots de leurs langues respectives. Parce que les Japonais utilisent les mêmes ideogrammes chinois qu'on appelle les caractères *kanji* sauf qu'ils ont adopté seulement les 2851 d'entre eux. Ils "passent" donc ceux qu'ils ne connaissent pas.

Par contre, les japonais utilisent à la fois un autre alphabet qui s'appelle *katakana* ou simplement *kana* qui contient 51 caractères syllabiques, *phonétiques* cette fois, pour écrire les suffixes qui s'agglutinent aux racines dans la langue parlée, pour la déclinaison et la conjugaison. Donc le texte japonais consiste des caractères *kanji* chinois, chacun suivis ou non des caractères *kana* japonais. Eventuellement on pourrait écrire toute la langue avec le *katakana*, comme l'écriture du Japonais avec les caractères anglais pour fournir une écriture phonétique, utile surtout pour l'usage des étrangers.

Pour le traitement informatique du Japonais, ils suffit donc de filtrer les caractères *kana* pour identifier les mots, et leurs sens comme ils sont définis dans le trésor dans la mémoire artificielle de l'ordinateur, aussi bien que dans la mémoire biologique de chaque Japonais, par l'éducation.

Pour un chinois, la lecture d'un journal japonais est possible, sans savoir du tout le Japonais, il suffit qu'il ignore les caractères *kana*, simples icônes qui représentent les suffixes, qu'il ne connaît ni dans l'écriture ni dans la langue parlée, il ne lui reste que les caractères *kanji* du chinois.

Mr. Hudrisier considère les ideogrammes *kanji* comme *les mots valises*, et montre qu'*une langue idéographique est toujours plus rapide qu'une langue alphabétique et que ceci crée(pour demain) des risques industriels aux langues alphabétiques.*

Pour conclure:

1. Il n'y a pas de raison intrinsèque pour que le web tisse des autoroutes pour l'Anglais plutôt qu'une autre langue, si nous voulons bien utiliser nos propres langues.
2. Le réseau Internet introduit, quelque soit la langue utilisée, un écrit instantané, presque plus proche de la parole que de l'écrit traditionnel. Ceci va contribuer à la création d'une société mondiale plurilingue. Chacun de nous, nous avons autant de chance pour la préservation de nos langues maternelles respectives. Il y a aussi des outils de corrections orthographiques, grammaticales ou autres. Chaque nation doit créer ces outils pour sa langue maternelle, sa culture.
3. L'instrument informatique, pour mémoriser les mots et les concepts des langues, en établissant les trésors, les bases des connaissances et l'inventaire culturels des nations, se prête aussi bien à notre service que les Américains. Pour en tirer avantage dans la création d'une société informatisée plurilingue, nous devons établir un dialogue multilatéral entre nous autres, apprendre nos langues réciproques.

Moi j'ai le plaisir de parler Français et connaître la culture française; vous, vous auriez le plaisir de parler Turc et connaître la culture turque. Si vous français, vous le refusez, les Américains refuseront de leur part, d'apprendre le Français (et le Turc bien entendu), et leurs cultures. Le "big brother" naît de telles attitudes des autres qui s'enchaînent dans les fausses voies.

4. Multitude d'informations simultanées en réseau est une chance pour être conscient de la réalité, mais aussi crée des problèmes de filtrage. Nous devons avoir conscience du pluralisme des langues, cultures et idées, pour ne pas nous y perdre.
5. Dans ce sens, la société informatisée moderne implique déjà une société plurilingue et pluriculturelle.
6. Marshall McLuhan avait vu que *le moyen était le message (the medium is the message)*. Il prétendait que "croire qu'une arme est bonne ou mauvaise suivant son utilisation, n'était que somnambulisme". L'arme est mauvaise parce que c'est un outil de guerre; le crayon est bon, parce que c'est un outil d'écriture, et donc de civilisation.

Pour finir je veux dire que l'informatique n'est pas un outil d'uniformisation, mais plutôt un outil de préservation de nos langues et cultures et d'un meilleur entendement international, de collaboration et de paix... Sauf ceux qui ne l'utilisent pas du tout!

Bilişim: Dillerin ve Kültürlerin Tekdüzeleşmesinin mi,
Çeşitliliğinin Korunmasının mı Aracı?*

Özet

Soru 2 düzeyde irdelenebilir: ulusal düzeyde, her yurttaşın ve sivil toplumun Devlet karşısındaki özgürlükleri; uluslararası düzeyde, her ulusun dil/kültür çeşitliliğini ABD ve çokuluslu ortaklıklar karşısında koruma hakkı. “Küreselleşmeyi insancillaştırmamız gerekiyor.”

Fransa Cumhurbaşkanı’na sunulan *Bilişim Toplumunun Oluşturulması* raporundan 2 saptama: (1) Fransa yeni ve ağır (teknikbilimsel) meydan okumalara doğru yanıt bulmazsa, kendi alinyazısına egemen olma yeteneğini yitirecektir; (2) Bilişim/toplum bağlamında, çağdaş uygarlıklar arasında oluşan dengenin en önemli birleşeni, iyi-kötü etkileriyle, *bilişim*’dir. Bilişimin her yurttaşın hizmetinde olması ve uluslararası işbirliği çerçevesinde gelişmesi gereklidir; kişisel verilere erişim, herkesin hakkıdır (1978).

Türkiye Bilişim Derneği (1971), bilgisayar kullanımının doğru ve bilinçli bir çizgide yaygınlaştırılması, bilişim toplumu altyapısının hazırlanması doğrultusunda bir toplumsal değişim etmeni olmuştur. Böylece 2.500 Türkçe bilişim terimi türetilmiş; bilgisayar mühendisliği öğretimi başlatılmış; kamusal/özel kuruluşlarda bilişim dizgeleri gerçekleştirilmiştir.

Yazar, *bitişken* yapıdaki Türkçenin yanı sıra, *ayrışık* yapıdaki Çincenin özellikleri üzerinde de durmakta; *kavramçizi*’lerle yazılan Çincenin bilgisayarla işlenmesindeki kolaylıkları göstererek, bilişimin, dillerin/kültürlerin tekdüzeleşmesi için değil, tersine çeşitliliğinin korunması için, dolayısıyla daha iyi bir uluslararası anlayış, işbirliği ve barış için, bir araç olarak kullanılabilceği sonucuna varmaktadır.

* Prof. Dr. Aydın Köksal’ın, Ankara’daki Fransızca konuşanlar topluluğu önünde, 17 Ocak 2002’de Fransız Büyükelçiliği’nde yaptığı konuşma.