



Extrait du document disponible dans le site Web  
de l'Office québécois de la langue française :

## **Caractéristiques culturelles et linguistiques du Québec** **Conventions et pratiques dans les technologies de l'information**

### **I. JEUX DE CARACTÈRES**

La représentation du système d'écriture du français nécessite plus de 134 caractères :

- les 26 lettres de l'alphabet latin, minuscules et majuscules (52 caractères);
- les 14 lettres avec signes diacritiques, minuscules et majuscules (28 caractères)  
àÀ âÂ çÇ éÉ èÈ êÊ ëË îÏ ôÔ ûÛ üÜ ùÙ ỳỲ;
- les dix chiffres arabes (10 caractères);
- les signes de ponctuation et autres symboles courants, communs à toutes les langues européennes (38 caractères);
- les ligatures, minuscules et majuscules (4 caractères) æÆ œŒ;
- les guillemets français, ouvrants et fermants (2 caractères) « »;
- quelques lettres d'origine étrangère pour certains mots du français (dont ñÑ et öÖ).

Il faut bien comprendre que ne pas utiliser les signes diacritiques sur les caractères, qu'ils soient en minuscules ou en majuscules, peut changer le sens des mots et est considéré comme une faute d'orthographe.

Exemples :

- ON SUBVENTIONNE L'INDUSTRIE DU POISSON SALE
- CHACUN A LE SOUCI DE SA TACHE QUOTIDIENNE
- UN RESTAURATEUR BLESSE ET VOLE
- MAIS ECLATE NON BEURRE MAIS CAMELISE

Dans un contexte multilingue, il faut toutefois prévoir idéalement l'utilisation d'autres caractères, soit, à tout le moins, l'ensemble des caractères communs aux langues qui utilisent l'alphabet latin affectés de leurs signes diacritiques, auxquels il faut ajouter les caractères d'usage national et international.

Au Québec, le compromis minimal qui puisse être fait, c'est de garantir l'accès au répertoire complet de l'Alphabet latin n° 1 (**ISO/CEI 8859-1**). Pour un soutien complet, y compris les ligatures œœ, de même que le caractère Ÿ et le symbole de l'euro (€), il faut utiliser au minimum le jeu de caractères **ISO/CEI 8859-15**.

On peut trouver une représentation de ces jeux de caractères en annexe de ce document et dans le site Web de l'Office québécois de la langue française aux adresses [http://www.oqlf.gouv.qc.ca/ressources/ti/ISO\\_CEI8859-1.pdf](http://www.oqlf.gouv.qc.ca/ressources/ti/ISO_CEI8859-1.pdf) et [http://www.oqlf.gouv.qc.ca/ressources/ti/ISO\\_CEI8859-15.pdf](http://www.oqlf.gouv.qc.ca/ressources/ti/ISO_CEI8859-15.pdf).

Le jeu ISO/CEI 8859-15 offre le soutien de 27 langues basées sur l'alphabet latin alors que le jeu ISO/CEI 8859-1 offre le soutien des 14 langues suivantes : albanais, allemand, anglais, catalan, danois, espagnol, finnois, français, islandais, italien, néerlandais, norvégien, portugais, suédois.

Ces recommandations valent pour les jeux de caractères à huit bits. Récemment, l'ISO a adopté le premier jeu de caractères normalisé destiné à englober l'ensemble des langues du monde : le code JUC, pour « jeu universel de caractères codés sur plusieurs octets » (**ISO/CEI 10646**). Le JUC, qui comprend plusieurs dizaines de milliers de caractères (le nombre en est encore en croissance pour inclure le chinois historique et les langues mortes), prend en compte d'ores et déjà la plupart des systèmes d'écriture en usage aujourd'hui; il pourra théoriquement être augmenté jusqu'à deux milliards de caractères (on a récemment fixé une limite pratique à un million de caractères). Le nouveau code oriente l'évolution des technologies. Ce nouveau jeu de caractères, avec lequel **ISO/CEI 8859-1** est compatible, est soutenu en mode natif par les plus récents systèmes d'exploitation. Il est également utilisé dans Internet avec l'aide du format transformé **UTF-8** (encodage à 8 bits).

Direction de la francisation



septembre 2004  
[www.oqlf.gouv.qc.ca](http://www.oqlf.gouv.qc.ca)  
[betel@oqlf.gouv.qc.ca](mailto:betel@oqlf.gouv.qc.ca)