

Cahiers **GUT** *enberg*

☞ ISO-LATIN-1, NORME DE CODAGE DES
CARACTÈRES EUROPÉENS ? TROIS
CARACTÈRES FRANÇAIS EN SONT ABSENTS !

☞ Jacques ANDRÉ

Cahiers GUTenberg, n° 25 (1996), p. 65-77.

<http://cahiers.gutenberg.eu.org/fitem?id=CG_1996__25_65_0>

© Association GUTenberg, 1996, tous droits réservés.

L'accès aux articles des *Cahiers GUTenberg*

(<http://cahiers.gutenberg.eu.org/>),

implique l'accord avec les conditions générales

d'utilisation (<http://cahiers.gutenberg.eu.org/legal.html>).

Toute utilisation commerciale ou impression systématique

est constitutive d'une infraction pénale. Toute copie ou impression

de ce fichier doit contenir la présente mention de copyright.

ISO Latin-1, norme de codage des caractères européens? trois caractères français en sont absents !

Jacques ANDRÉ

*Irisa/Inria-Rennes
Campus de Beaulieu
F-35042 Rennes cedex
jandre@irisa.fr*

Résumé. Trois caractères français ne font pas partie d'ISO Latin-1 : œ, Œ et Ÿ. On montre pourquoi ils devraient y être et pourquoi ils n'y sont pas. Puis on en tire quelques leçons sur la normalisation des caractères.

Abstract. *Three French characters (œ, Œ, and Ÿ) are missing in ISO Latin-1. It is explained why they should be there and why they are not. Then, some comments are made on the standardisation of characters.*

1. Introduction

Tout le monde le sait bien, la norme ISO Latin-1 qui comprend tous les caractères de l'Europe de l'Ouest, donc de la France, ne comprend pas les caractères œ ni Œ. On sait moins souvent qu'un troisième caractère français est aussi absent, le Ÿ. En fait, cette norme n'est qu'une norme d'échange de caractères et non de glyphes (voir à ce sujet [1, 11]) et finalement on s'en sortait quand même pas trop mal, cet article par exemple étant imprimé (ou même transmis sur l'Internet par ftp) avec les bons caractères œ, Œ et même Ÿ. Mais deux phénomènes m'ont récemment fait réagir¹.

1. Diverses listes de discussion sur l'Internet sont consacrées d'une façon ou d'une autre à la typographie, par exemple `comp.font` où beaucoup de questions sont relatives aux fontes, `Uni code` au codage des caractères pour Unicode et `TYPO-L`

1. Cet article est en effet la version française d'une réponse que j'ai faite à `TYPO-L@LISTSERV.HEA.IE` le 18 novembre 1996. Des versions voisines de ce texte doivent aussi paraître dans *Cantonade* (revue du syndicat des correcteurs), *Graphê* (revue de l'APAT) et la *Revue Suisse de l'Imprimerie*.

qui tourne plutôt au niveau du code typographique anglais. Il s'est trouvé que lors du dernier trimestre 1996, beaucoup d'échanges de messages ont eu lieu sur ces trois listes à propos du codage des caractères, frisant parfois un peu la maniaquerie².

Mais ces discussions ont dégénéré vers des affirmations comme celle de Peter Flynn sur Typo-L où il affirmait « *The main point is that y-trema is a MISTAKE. According to everyone who answered, it DOES NOT EXIST, it was included by people who did not know what they were doing because of bad information.* » avec l'idée sous-jacente que puisque ces caractères ÿ et Ÿ n'existent pas, il faut les supprimer d'ISO Latin-1 et d'Unicode³.

2. Ce problème aurait pu rester mineur dans la mesure où les glyphes de ces caractères existent bien tant dans les casses en plomb que dans les tables de codage à la *Adobe Standard Encoding* ou autre ; mais il se trouve qu'aujourd'hui⁴ la transmission des informations sur l'Internet passe essentiellement (surtout lorsqu'elles sont mises en page) par HTML. Or si on peut coder en gros ce que l'on veut avec \LaTeX , il n'en est pas de même avec HTML : en effet, si l'on veut être lu par n'importe-qui, on doit employer une norme d'échange admise universellement, souvent Ascii ou, en Europe, ISO Latin-1 avec donc l'impossible d'y utiliser œ, Œ ou Ÿ, sauf à l'aide d'astuces comme celles consistant à remplacer, comme le fait par exemple *LaTeX2html*, ces caractères par une image bitmap de ces caractères, ce qui n'est jamais ni précis quant aux métriques, ni de bonne qualité. Certes, Unicode/ISO 10646 a réparé cette lacune, mais les systèmes utilisant des codes à 16 ou 32 bits sont encore très rares aujourd'hui.

2. Le problème est que sur ces listes beaucoup de gens découvrent, naïvement, leur propre culture. Des Allemands qui ignoraient tout des ligatures découvrent brusquement que l'on peut, ou l'on ne peut pas, mettre la ligature *ffl* dans le mot *Aufflage* du fait qu'il s'agit du mot composé *Auf-Flage*. Mais, en France, on a rigoureusement le même genre de phénomène lorsque les média disent « d'ici à Noël » croyant que la forme « d'ici Noël » serait erronée parce que l'on dit d'« ici à Rennes » et ne se rendent même pas compte que la faute, s'il y en avait une, serait plutôt d'employer « ici » dans un contexte temporel au lieu de « maintenant » !

3. Ou, respectivement ISO 8859-1 et ISO 10646 : voir [1, 11] pour ces diverses normes. ÿ existe dans ISO Latin-1 et Unicode, mais Ÿ n'existe pas dans ISO Latin-1 (voir tableau 1) ; heureusement, à un nouveau message de Peter Flynn à `unicode@Unicode.ORG` fin novembre, il lui fut répondu *According to the rules, a character cannot be deleted from UCS. So it [Y diaeresis] has to stay with us.*

4. Et même dès hier avec par exemple le Minitel où il est presque cocasse de voir un dictionnaire géré par le CILF (Conseil International de la Langue Française : 3615 code orthotel ou <http://www.sdv.fr/orthonet>) qui propose « oeil-de-boeuf » comme bonne orthographe de « œil-de-bœuf ».

CROÿ [Kruï]. ♦ Famille française connue depuis le **xii^e s.**, qui doit son nom au village de Crouy, dans la Somme. Elle a servi les ducs de Bourgogne, les Habsbourg, et les rois de France, et s'est divisée en de nombreuses branches (Renty, Havré, Roeux, Chimay, Solre...). ♦ **EMMANUEL DE CROÿ** (Condé, 1718 – Paris, 1784), maréchal de France. Il se distingua à Fontenoy et fortifia les ports de Dunkerque et de Calais.

FIGURE 1 – La famille Croÿ, selon le Robert [12] qui utilise donc le y tréma en petite capitale et en capitale.

2. Ÿ et ÿ sont-ils français?

2.1. Quelques attestations

Effectivement, ni Ÿ ni Ỳ ne sont attestés dans les noms communs français et aucun des principaux codes typographiques ([5, 7, 8], [10, page 102]) ne cite ces caractères; Grevisse [6, art. 79] ne liste pas le « y » parmi les voyelles pouvant avoir un tréma et Jacques Anis [3, page 90] n'en fait pas non plus état.

Mais par contre, le Ỳ apparaît dans des noms propres français tels que⁵ :

- | | |
|---------------|---|
| Aÿ | chef-lieu de canton de la Marne. |
| Rue des Cloÿs | et impasse, dans le 18 ^e arrondissement de Paris. |
| Croÿ | Il existe un hôtel, Boulevard Raspail à Paris, du nom de Croÿ. Plusieurs personnages célèbres portent ce nom ⁶ : |
- « *Albums de Croÿ* : magnifique édition intégrale de 2 500 gouaches réalisées à la demande du Duc Charles de Croÿ (1510-1612) et représentant les villes et villages des anciens Pays-Bas espagnols » (*Catalogue des Presses Universitaires du Septentrion*, Lille, 1996) ;
 - « Lorsque Charles-Quint organisa son gouvernement, il ne s'appuya que sur des Bourguignons [...]. Politiquement, il avait été l'élève d'un grand seigneur, Guillaume de Croÿ, sieur de

5. La liste donnée ci-après est grandement due à la collaboration de Patrick Andries, Thierry Fuhs, Alain La Bonté, Bernard Marti, Christophe Parent, Paul Pichaureau, Daniel Taupin et François Yergeau qui ont répondu à une question sur la liste « gut » et que je tiens à remercier ici. On m'a aussi signalé le nom de Valentin Haÿ, mais à tort car le tréma est en fait sur le u !

6. Voir figure 1 ; les divers détails donnés ici sont dus à Patrick Andries.

HAY-LES-ROSES (L') [laïləʁoz]. ♦ Ch.-l. de cant. du Val-de-Marne (arr. de Créteil). 24 378 hab. Célèbre roseraie. — Matériaux de construction.

FIGURE 2 – L'Hay-les-Roses, selon le Robert [12].

Chièvres, mais les principes inculqués par ce dernier furent mal appliqués par l'empereur. » (cité dans *Histoire de l'empire des Habsbourg*, Fayard, 1990, p. 160) ;

- *Larousse Universel* (1922) : « Croÿ (prononcer KROU-Î [circonflexe sur le I]) illustre famille française de Picardie. Un de ses membres les plus distingués, Emmanuel de Croÿ, maréchal de France, consacra une partie de sa fortune à restaurer le port de Dunkerque (1717-1787). »
- *Le Quid?* (1988) cite un noble en Picardie depuis 1207, un duc en 1598, un prince du St Empire en 1742, reconnu en France en 1768 et le 4/VII/1814.

Ghÿs serait la graphie enregistrée ainsi à l'état-civil français d'un nom dont la forme originale (d'origine flamande) était Ghij mais où la diphtongue « ij » aurait été translittérée « ÿ » du fait de la ressemblance de ces deux graphèmes lorsqu'ils sont écrits à la main.

Le Blanc de Nicolaÿ compagnie d'assurance de l'Est de la France.

L'Hay-les-Roses commune entre Paris et l'aéroport d'Orly (figure 2).

Pierre Louÿs écrivain et érudit français (1870–1925).

Moÿ nom de lieu (commune de l'Aisne notamment) et de famille.

etc. On trouverait sûrement d'autres attestations de ÿ dans des noms de lieu. Le problème est qu'il faut regarder les divers documents « à la main » et encore en ne se fiant qu'à d'anciens documents : la rue des Cloÿs est bien orthographiée dans le vieux guide Torride, mais sans ÿ dans les récents plans de Paris de Michelin ; L'Hay-les-Roses a le droit au tréma dans la tête de cette commune dans l'annuaire du téléphone des PTT (version imprimée car bien sûr la version minitel ...), mais pas dans les titres courants de ce même annuaire ; de même la rue de Cloÿs perd son tréma dans ce même annuaire, tout comme (on peut le supposer) les divers Moy ou Croy qui sont inventoriés à Paris ; même l'*Atlas des communes de France*, édité par Oberthur vers 1960, ignore ce ÿ !

2.2. D'où vient le ÿ?

On l'aura remarqué, la majorité des exemples de noms contenant des ÿ cités ci-dessus sont des noms anciens. Mais d'où vient ce caractère? Jacques Anis et Bernard Cerquiglini (respectivement professeurs à l'Université de Nanterre et de Paris 3) m'ont tout deux confirmé la chose : la voyelle « y » au Moyen-Âge est utilisée comme « i long » à des fins de lisibilité. Le tréma apparait dans ces noms comme il apparaîtrait sur un i pour casser le digramme ai ou oi. C'est à dire que tout comme « ai » se prononce⁷ « a-i » et non « ai » (comparez par exemple « haïr » et « haie »), tout comme « oi » se prononce « o-i » et non « oi » (comparez « Moïse » et « Moi »), de même (à la longueur du « i » près), « L'Haÿ-les-Roses » se prononce « L'A-i les roses » et non pas « L'Aille les roses » ni bien sûr « La Haie les roses ». Notons que si vraiment le ÿ devait ne pas exister, L'Haÿ-les-Roses pourrait au moins s'écrire « L'Haï-les-Roses » (voire « L'Haï ») plutôt que « L'Hay-les-Roses » !

De même, il faut prononcer Croÿ « Crou-y », Louÿs « Lou-ys » et Moÿ « Mo-y » et non respectivement « Croi », « Louis » ou « Moi » (et surtout pas « Croille » ou « Moille »!).

2.3. Y tréma et les codes

Ainsi que nous l'avons signalé [1], il faut distinguer les normes d'échange de caractères qui ne sont finalement que des tables de noms de caractères auxquels on associe un numéro – le nom du caractère 255 (FF en hexadécimal) en ISO Latin-1 est « LETTRE MINUSCULE y TRÉMA » – et les codages de glyphes (où par exemple Adobe associe au caractère « y dieresis » l'œil du « ÿ », ou à « oe » l'œil du caractère « œ »).

Le tableau 1 donne quelques codages (ISO Latin-1, Unicode, Adobe et le standard de Cork utilisé en T_EX : voir [1]) de ces caractères ÿ et Ÿ. Ce tableau montre une anomalie : Ÿ n'existe pas en ISO Latin-1 !

2.4. Pourquoi n'y a-t'il pas de Ÿ dans ISO Latin-1?

Il ne nous a pas été donné d'avoir de réponse sûre et concluante à ce sujet.

Une chose est certaine : la place que devrait occuper le Ÿ est occupée par le ß allemand qui n'est pourtant qu'un glyphe de « ss » (n'existant pas en capitale). En effet, tant pour les codes des lettres non-accentuées d'Ascii que pour ceux des lettres accentuées des parties spécifiques à chaque sous-ensemble d'ISO 8859-n, on a la relation

$$code_{capitale} = code_{bas-de-casse} - 32$$

7. L'emploi d'alphabets phonétiques, dont l'international API, serait sûrement plus rigoureux ici, mais moins pédagogique dans la mesure où ces alphabets ne sont connus que des seuls spécialistes!

TABLE 1 – Quelques codages de \ddot{y} et de \ddot{Y} . Une case blanche indique que le caractère n'est pas connu dans le codage correspondant.

Type	norme/standard	code du \ddot{y}	code du \ddot{Y}
Caractères	ISO Latin-1/8859-1	255 = FF ₁₆	
	Unicode/10646	00FF ₁₆	0178 ₁₆
Glyphes	Adobe StandardEncoding		
	Adobe ISO Latin1Encoding	ydiereasis 255	Ydiereasis
	Cork	\"y 184	\"Y 152

Par exemple, $code_A = 65$ et $code_a = 97$; de même $code_{\ddot{e}} = 201$; $code_{\ddot{e}} = 233$. Par contre, $code_{\mathcal{R}} = 223$; $code_{\ddot{i}} = 255$.

Il y a donc eu volonté délibérée de ne pas mettre de \ddot{Y} ; ce ne peut être un oubli. Il fallait sûrement sacrifier quelques caractères, compte-tenu du peu de places dans ces 256 caractères dont tous d'ailleurs ne sont pas imprimables. On imagine alors très bien ce qui a pu se passer : un représentant a probablement dit que puisque le \ddot{y} avait pour but d'indiquer qu'un y devait être prononcé séparément de la voyelle qui le précède, alors il ne pouvait exister de \ddot{y} en début de mot et que donc le \ddot{Y} n'existait pas.

Si tel est le cas, c'est oublier que souvent on met en capitales les noms propres et les entrées de dictionnaires (voir par exemple les figures 1 et 2), etc. On pourrait évidemment dire que, par exemple dans Le Robert, le \ddot{Y} de L'HAY-LES-ROSES n'est qu'un glyphe contextuel de \ddot{y} . Mais ce serait aussi faire preuve d'incohérence : dans ISO Latin-1, on a à la fois le \grave{u} et le \grave{U} . Or, des langues pour lesquelles cette norme a été conçue, seul le français utilise le \grave{u} (certes le \grave{u} existe aussi en catalan, en roumain, etc., mais ces langues relèvent d'une autre partie de l'ISO 8859); et en français, le \grave{u} n'existe que dans le seul mot « où »⁸), donc jamais en première lettre d'un mot. Pourquoi donc le \grave{U} existe-t-il mais pas \ddot{Y} ? Mystère ...

3. Le cas de œ et de Œ

Georges Perec a écrit un roman, *La disparition*, où pendant plus d'une centaine de pages il a réussi la gageure oulipienne de ne pas employer une seule fois la lettre e ! Pour ce faire, il a dû utiliser tout un vocabulaire (plantes, animaux, noms propres, etc.) et des situations très spéciales (dialogues où le « je » est interdit par exemple). Mais aurait-il

8. À tel point d'ailleurs que plusieurs personnes ont demandé alors pourquoi ne pas avoir carrément la touche « où » sur nos claviers ! la réponse est qu'il en faudrait alors trois : une pour « où », une pour « Où » et une pour « OÙ » !

été capable d'écrire ne serait-ce qu'une dizaine de pages racontant la vie banale d'un français moyen sans utiliser les mots suivants?

- bœuf
- chœur
- cœur
- mœurs
- nœud
- œil
- œuf
- œuvre (et chef-d'œuvre, hors-d'œuvre, manœuvre, etc.)
- sœur
- vœux
- etc.

C'est pourtant ce que arrive à toute personne voulant écrire une page Web à l'aide d'HTML en utilisant ISO Latin-1 d'où ces lettres œ et Œ ont disparu !

3.1. Œ et œ sont-elles des lettres?

Avant d'expliquer pourquoi Œ et œ ont été « oubliés » dans ISO 8859-1, posons-nous la question de savoir pourquoi ils devraient figurer dans une liste de caractères d'où sont exclues les ligatures comme « fi » puisqu'il s'agit d'un glyphe. C'est que justement Œ et œ ne sont pas des ligatures typographiques normales (même si leurs glyphes viennent de la fusion d'un o et d'un e ; mais voir figure 3) mais des ligatures linguistiques ou même des caractères orthographiques à eux tout seuls (voir [2, 9]) ; on parle parfois de bigramme ou de digramme soudé. La lettre œ a en gros trois origines en français. Elle vient d'une part de noms propres étrangers translittérés⁹, soit de noms savants¹⁰ inventés à partir de racines grecques (ou transmis via le latin, comme foetus) soit enfin de diphtongues franco-françaises sur lesquelles nous allons revenir.

Ceci n'est pas spécifique au français ; on retrouve le même genre de diphtongues dans d'autres langues :

- le néerlandais utilise les « ligatures » ij et IJ (qui n'existent pas non plus en ISO Latin-1) ;
- en breton, il n'y a pas de lettre « c » (qui s'écrit « k »), mais par contre il y a deux « ligatures » ch et c'h ;

9. Tels que Jacques Cœur, Kœchlin, Kœnig, Kœnigs, Mœbius, Œrsted, Schœlcher, etc.

10. Tels que coelacanthé, cœliaque, diœcie et monoœcie, œdème, œnologie, œsophage, œstrogène, pœcile, etc.

- en suédois, on trouve les ligatures æ et Æ d’origine latine¹¹ ; elles sont bien dans ISO Latin-1 ;
- en espagnol, jusqu’à la récente réforme (1987) de la *Real Academia Española*, les diphtongues « ll », « ch » et « rr » étaient considérés comme des lettres à elles-seules.

Ces caractères spéciaux ne sont d’ailleurs pas obligatoirement placés alphabétiquement à la place normale :

- en suédois, æ est après le z alors qu’en français on a « caducée, cæcum, cafard » ;
- en breton, l’ordre alphabétique est tel que « c’h » est après « ch » : on aura donc « chug (jus), c’hoar (sœur) » ;
- en néerlandais, la ligature ij se place selon les dictionnaires à diverses places, mais rarement après le i ;
- en revanche, en français, on a l’ordre « coercion, cœur, coexister ».

Qu’est-ce qui distingue ces ligatures « linguistiques » des typographiques ?

- Nina Catach explique dans son *Dictionnaire de l’orthographe* [4, art. cœur, nœud, œil, sœur, vœu] que tous ces mots très courts, voire monosyllabiques, ont cette graphie depuis le xvii^e siècle (Estienne, Thierry, Nicot, Furetière et l’Académie sont alors les principaux théoriciens de la langue française). Dans la quasi-totalité des cas, cette graphie répondait à deux besoins : d’une part de rappeler l’étymologie latine en « o » et, d’autre part, d’éviter l’homographie, la confusion, avec un autre mot (par exemple « meurs », « seur » et « veu », à l’époque, étaient des orthographes correctes respectivement pour « mûrs », « sûr », et « vu »).
- Cette origine se retrouve dans la prononciation de ces diphtongues. Tout comme le suédois æ se prononce « é » et non « a-e », de même le œ français ne se prononce pas comme les deux lettres « o-e ». On a même pu dire que le œ est un anti-tréma : on passe du son unique « ue » au son double « u-e » en ajoutant un tréma sur le e « uë » ; réciproquement on passe du son double « o-e » au son unique « œ » grâce à cette ligature ; selon la voyelle suivante, la prononciation diffère (comparez « œil » et « œuf »). En fait « œ » c’est un peu le ö (o-umlaut) allemand ! Nina Catach [4] fait remarquer que le tréma qui existait sur certains mots comme « moëlle » pour noter la prononciation disjointe de o et de e est tombé vers 1700 (mais « poète » s’écrivit jusqu’en 1878 « poëte »). On ne retrouve plus guère aujourd’hui ce tréma que dans des noms propres comme Joël et Noël.

11. Les caractères Å et å, qui eux aussi font partie d’ISO Latin-1, sont des inventions du début de ce siècle et remplacent les bigrammes AA et aa qui se prononcent en gros comme « o ».

TABLE 2 – Quelques codages de œ et de Œ. Une case blanche indique que le caractère n'est pas connu dans le codage correspondant.

Type	norme/standard	code du œ	code du Œ
Caractères	ISO Latin-1/8849-1		
	Unicode/10646	0152 ₁₆	0153 ₁₆
Glyphes	Adobe StandardEncoding	oe 372	OE 352
	Adobe ISO Latin1Encoding		
	Cork	\oe 245	\OE 215

- A ces raisons linguistiques, on peut ajouter des caractéristiques typographiques :
 - contrairement à « fi », on ne peut remplacer n'importe quelle séquence « oe » par « œ » : on ne peut pas écrire mœlleux ni coexister pas plus qu'on n'a le droit d'écrire noeud !
 - si l'on doit mettre une capitale à la première lettre d'un mot, on ne le fait pas aux deux premières (on n'écrit jamais FIn) ; or, on écrit bien « Œuvres de ... », à la rigueur « OEuvres de ... » mais pas « Oeuvres de » ; il en est de même en néerlandais : le *Grand Atlas Universel* (Bordas 1978, page 74) cite autour d'Amsterdam IJmuiden, IJpolder, IJmeer et IJssemeer et non Ijmuiden, Ijpolder, Ijmeer ni Ijssemeer ;
 - si l'on accepte les divisions de mots selon leur étymologie, on peut faire une césure entre o et e, jamais dans la ligature œ (par exemple, on peut diviser co-exister, mais on ne divisera jamais co-eur) ; notons bien que ce n'est pas le fait d'avoir une ligature qui interdit de diviser : en effet, en Garamond où on utilise la ligature « st », on pourra très bien couper « coexis-tence ».

3.2. Histoire d'Œ

Le tableau 2 montre la présence ou l'absence des caractères œ et Œ dans les principaux codages. Pourquoi ces deux caractères œ et Œ ne figurent-ils pas dans ISO Latin-1 ? L'histoire veut que le représentant français était absent le jour du vote et que les Islandais en ont profité pour faire passer le thorn. La vérité est un peu plus triste ; elle m'a été racontée par Alain La Bonté (Service de la prospective au Gouvernement du Québec). Il a été dès 1988 (peu après donc l'adoption d'ISO Latin-1) expert-délégué canadien officiel au groupe de travail 3 du SC2 (sur le codage de l'information)¹² :

C'est D.¹³ de Bull, qui représentait l'AFNOR au sein du GT3 en 1987... c'est au sein du GT que le travail se fait réellement. Si on attend les votes au niveau du SC, il est trop tard. Au sein du GT3, H. un francophone polyglotte a

12. Voir [11, page 17] pour le fonctionnement de l'ISO.

13. Les noms propres sont remplacés ici par des initiales {Ndlr}.

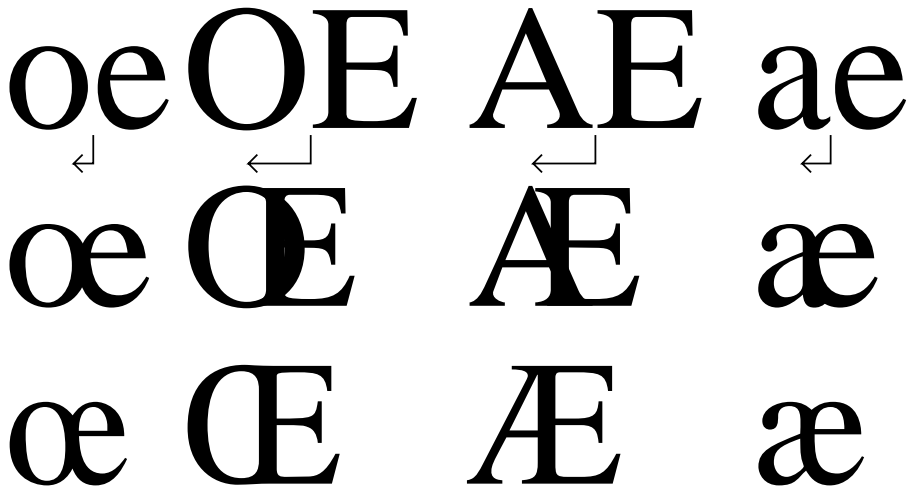


FIGURE 3 – Ces ligatures linguistiques ne peuvent être obtenues par crénage

déclaré que l'œ n'était qu'une vulgaire convention typographique... Je n'étais pas là à l'époque, je suis arrivé peu de temps après l'adoption mais tous m'ont confirmé cela.

D., sans consulter aucun Français autre que les gens de sa boîte (dont j'ai su qu'ils avaient pris une décision d'équipe uniquement parce que les imprimantes de Bull ne produisaient ni œ ni Œ), ont appuyé H. ...

Le Canada – alors représenté par Y. un anglophone qui n'avait pas d'autre argument à présenter que de dire que le Canada tenait au œ – a vivement protesté en disant que c'était un besoin. Mais c'était peine perdue, les deux francophones influents du GT3 affirmaient tous deux que cela n'était pas nécessaire et que la ligature pouvait à la rigueur être produite par un caractère de commande¹⁴ de l'ISO 6429 que le GT3 venait de produire... Plus tard (1988), j'ai produit une contribution, en français et en anglais, démontrant que l'œ servait à des fins orthographiques et ne pouvait être employé au gré de la fantaisie de tout un chacun¹⁵. Bien que H. ait alors affirmé publiquement que c'était la première démonstration intelligente qu'il voyait à cet effet, cela n'a rien changé au fait que la norme avait déjà été publiée¹⁶.

14. Cette commande revenait à faire un crénage, c'est-à-dire à coller le o et le e. Si cela passe à la rigueur pour « oe » sur un écran (pour lesquels cette norme était principalement conçue), c'est déjà moins vrai pour « Œ » et en tout cas pas du tout pour « Æ » qui est graphiquement trop éloigné de « Æ ». Voir figure 3. Peut-être est-ce pourquoi æ et Æ, eux, sont dans ISO Latin-1 ? Notons que Bernard Marti avait proposé d'utiliser la norme ISO 6937 (dérivée de CCITT T50/T51) qui aurait permis la gestion des accents flottants, mais en vain. {J.A.}

15. Ses arguments sont en gros ceux développés dans la section 3.1 ci-dessus {Ndlr}.

16. Lors du vote du SC, le représentant français, Bernard Marti du CCETT, farouche défenseur du œ, était effectivement malade ; mais de toutes façons, le CT avait déjà décidé que œ ne serait pas dans ISO Latin-1 !

Le résultat : on a \times et \div (multiplier et diviser) à la place de CE et de α , grâce à l'insistance de W. B., un Allemand, qui m'a dit déplorer que α et CE n'y soient pas (mais apparemment il est resté neutre lors du débat, ce dont je n'ai jamais pu être entièrement sûr, par contre ; comme il était le père de l'ISO 6429 sur les caractères de commande, l'argument de la 6429 devait bien lui plaire).

Voilà... La faute ultime est due à la mesquinerie d'une équipe particulière chez Bull (un membre de l'équipe m'a lui même confirmé l'histoire qui s'était passée en coulisses, je lui avais dit que je savais que D. était le responsable ; il m'a alors dit que sa petite équipe chez Bull était bien au courant et avait pris cette position sciemment, à l'insu du reste de la France et même de B. Marti)... si la France avait protesté en comité (et si, moi, le Québécois bigot avait été là, je suis bien connu pour mon entêtement en cette matière), jamais cela ne se serait fait.

3.3. Cette disparition est-elle inexorable ?

Le secrétariat pour lequel travaille Alain La Bonté est candidat pour prendre la relève du secrétariat international du SC2 dès janvier 1997, qui se situe aujourd'hui à l'ECMA, à Genève. Si les Canadiens obtiennent ce poste, le secrétariat sera toujours francophone, ce qui serait avantageux pour la francophonie. Et comme on parle d'une révision d'ISO Latin-1 (ou de la création d'un ISO Latin-n), peut-être serait-ce une façon de retrouver les caractères perdus (les Irlandais et les Danois seraient favorables au retour de α et de CE).¹⁷

4. Conclusion

Il semble que plusieurs leçons sont à tirer de cette histoire :

- comme l'a écrit lui-même Bernard Marti [11], *on constate que certaines langues régionales des pays nordiques (féorien, islandais) sont prises en compte dans ISO Latin-1 et pas les langues régionales des pays situés plus au sud (catalan, basque, ...) et que, même pour des langues nationales, certaines ligatures sont « oubliées » (α français, ij hollandais) et pas d'autres (α nordique, β allemand) ; ce résultat est une conséquence de la faiblesse du niveau de travail effectué en France sur ces sujets et marque la domination germanique en matière de science de la documentation, qui sont à la base de ces choix et en matière de normalisation associée ;*

¹⁷. Ce *Cahier* ne paraissant qu'en 1997, au moment de le mettre sous presse on apprend que c'est finalement Toshiko Kimura du ITSCJ (Japon) qui a la charge de ce secrétariat. Des bruits courent aussi que si la révision d'ISO Latin-1 il y aura, ce serait plutôt dans le sens de dire « le français y est partiellement couvert » que de nous donner accès à tous nos caractères !

- soit qu'ils n'aient pas été sollicités, soit qu'ils n'aient pas pris conscience de l'importance des normes, les typographes brillent par leur absence dans les instances de normalisation où l'on trouve donc essentiellement des « ingénieurs » ; certes, il s'agissait de normes de transmission, pas de dessins, mais d'autres normes sont en cours de préparation et concernent par exemple les claviers ou les dessins et métriques de caractères (ISO 9541) ; quand elles paraîtront, il sera trop tard pour réagir (comme c'est le cas pour DSSSL sur, en gros, les feuilles de style) ! « Mais bien sûr, s'il n'y a que des ingénieurs pour défendre la langue française ... » disent, passivement, certains typographes français ! On pourrait en revanche ajouter qu'il est quand même curieux que d'un côté on oblige les scientifiques français à organiser des colloques francophones et que de l'autre on laisse le soin aux Canadiens de défendre la francophonie !
- on est en droit aussi de se demander ce que l'on normalise : l'état d'une langue à un moment donné, ou son évolution possible ? il serait quand même dommage qu'un Louis n'ait pas le droit d'orthographier son nom d'écrivain Louÿs ni de créer un personnage du nom de Ūgaro !
- finalement, on se rend compte que si l'Académie française est responsable du devenir du vocabulaire français, que si que les divers ministères et organismes chargés de la francophonie ont nommé diverses commissions pour le suivi de l'emploi du français, aucun organisme officiel n'existe, regroupant typographes, correcteurs, linguistes, etc., qui soit en charge de suivre l'évolution de la typographie francophone ; c'est sur ces mots que Louis Guéry aussi termine l'introduction de son *Dictionnaire des règles typographiques* [7].

Remerciements Cette note doit beaucoup à des Québécois sans doute plus sensibles à la francophonie que nombre de Français, et notamment à Patrick Andries, Alain La Bonté et François Yergeau que je tiens à remercier ici. Merci aussi à de nombreuses autres personnes, notamment à celles qui ont corrigé des versions précédentes de cette note.

Bibliographie

- [1] Jacques ANDRÉ et Michel GOOSSENS, « Codage des caractères et multilinguisme : de l'Ascii à Unicode et ISO IEC/10646 », *Cahiers GUTenberg*, 20, mai 1995, p. 1–53.
- [2] Jacques ANDRÉ, « Ligatures et informatique », *Cahiers GUTenberg*, 22, septembre 1995, p. 51–86.
- [3] Jacques ANIS, *L'écriture – théorie et descriptions*, DeBoeck, Bruxelles, 1988.
- [4] Nina CATACH (sous la direction de), *Dictionnaire historique de l'orthographe française*, Larousse, 1995.

-
- [5] *Code typographique – Choix de règles à l'usage des auteurs et des professionnels du livre*, Fédération nationale du personnel d'encadrement des industries polygraphiques et de la communication, 14^e édition, 1991.
- [6] GREVISSE, *Le bon usage*, éd. Duculot, Belgique (révisions quasi-annuelles).
- [7] Louis GUÉRY, *Dictionnaire des règles typographiques*, CFPJ Éditions, Paris, 1996 (ISBN 2-85900-129-8).
- [8] *Guide du typographe romand – Règles et grammaires typographiques à l'usage des auteurs, éditeurs, compositeurs et correcteurs de la langue française*, édité par le Groupe de Lausanne de l'Association suisse des compositeurs à la machine, 5^e édition, Lausanne, 1993.
- [9] Yannis HARALAMBOUS, « Tour du monde des ligatures », *Cahiers GUTenberg*, 22, septembre 1995, p. 87–99.
- [10] *Lexique des règles typographiques en usage à l'Imprimerie nationale*, Imprimerie nationale (2, rue Paul Hervieu, 75015 Paris), 1990 (3^e édition).
- [11] Bernard MARTI et co-auteurs, *Télématique – techniques, normes, service*, Dunod, 1990.
- [12] ROBERT, *Dictionnaire universel des noms propres*, Le Robert, Paris, 1974.