

Chapitre 12

Autres écritures

Ce chapitre décrit une série d'écritures n'ayant pas leur place au sein d'une des catégories présentées dans les autres chapitres :

- éthiopien ;
- chérokî ;
- syllabaires autochtones canadiens ;
- mongol ;
- déséret.

Bien que les écritures éthiopienne et mongole soient des adaptations d'écritures sémitiques, on ne peut les classer aujourd'hui parmi les écritures du Moyen-Orient. Le chérokî et les syllabaires autochtones canadiens s'inspirent de la sténographie et des lettres latines, mais leur utilisation et leur développement appartiennent au monde amérindien.

12.1 Éthiopien

Éthiopien : U+1200 – U+137F

Le syllabaire éthiopien s'utilise d'abord pour transcrire le guèze, une langue sémitique africaine. Le guèze, langue officielle de l'empire d'Axoum, disparaît entre le X^e et le XII^e siècle en tant que langue parlée ; on ne le retrouve de nos jours que dans des contextes liturgiques. Le syllabaire éthiopien s'emploie aujourd'hui pour écrire plusieurs langues vivantes originaires de l'Afrique centrale orientale, parmi lesquelles l'amharique, le tigré et l'oromo.

Éthiopien de base et étendu. Les caractères éthiopiens codés dans ce bloc correspondent au jeu de base utilisé d'ordinaire pour écrire les langues les plus importantes. Comme pour toute écriture productive, les formes de l'éthiopien de base changent occasionnellement pour produire de nouveaux caractères destinés à d'autres langues. Les recherches visant à identifier ces nouveaux caractères se poursuivent, bien que nombre de ces formes soient rares et généralement peu attestées dans la tradition typographique des langues en question.

Principes de codage. On représente traditionnellement le syllabaire éthiopien comme une matrice à deux dimensions : les voyelles et les consonnes. Unicode respecte cette structure : l'espace de codage U+1200...U+1357 s'interprète comme une matrice de 43 consonnes fois 8 voyelles, formant ainsi 344 syllabes possibles. La plupart de ces combinaisons consonnes-voyelles correspondent à des syllabes éthiopiennes écrites ; certaines, toutefois, ne s'utilisent jamais ce qui explique que certaines cellules de la matrice sont vides.

Variantes de glyphe. On peut représenter une syllabe éthiopienne donnée à l'aide de différents œils, analogues aux variantes de glyphes des bas de casse latins « a » ou « g » qui ne coexistent pas au sein d'une même police de caractères. De la sorte, un glyphe particulier, présent dans le tableau de codes pour chaque position dans la matrice, n'est qu'une représentation possible de la syllabe codée.

Sous-séries labialisées. Quelques consonnes éthiopiennes ont une forme labialisée (« en W ») ; on leur a attribué des séries particulières de consonnes dans la matrice des syllabes et ce bien que seul un sous-ensemble des formes voyellées soit réalisé. Chacune de ces séries dérivées suit immédiatement la série de la consonne principale correspondante. Les séries de voyelles incluent à la fois le « A » et « WA », il se peut donc que deux cellules différentes représentent la syllabe « consonne + W + A ». Exemple :

U+1247 = Q + WA : version inusitée de QWA
 U+124B = QW + A : ϕ SYLLABE ÉTHIOPIENNE QWA

Les deux syllabes sont alors équivalentes, on n'a codé que la forme labialisée de la consonne et non la forme « consonne + WA ». Le *Tableau 12-1* énumère les six cas particuliers.

Tableau 12-1. Formes labialisées en ~WA

Forme en ~WA	Codé au numéro	Numéro de code inusité
ϕ QWA	U+124B	1247
ϕ QHWA	U+125B	1257
ϕ XWA	U+128B	1287
ϕ KWA	U+12B3	12AF
ϕ KXWA	U+12C3	12BF
ϕ GWA	U+1313	130F

En outre, à l'intérieur même des sous-séries de formes labialisées, on considère parfois les formes de la sixième voyelle (« ~E ») comme des formes de la seconde voyelle (« ~OU »). Par exemple :

U+1249 = QW + OU : version inusitée de QWE

U+124D = QW + E : ቆ SYLLABE ÉTHIOPIENNE QWE

Les deux syllabes sont alors pratiquement équivalentes et on n'a codé que la forme en « ~E » et non celle en « ~OU ». Les six cas particuliers sont énumérés au *Tableau 12-2*.

Tableau 12-2. Formes labialisées en ~WE

Forme en ~WA		Codé au numéro	Numéro de code inusité
ቆ	QWE	U+124D	1249
ቆ	QHWE	U+125D	1259
ኧ	XWE	U+128D	1289
ኧ	KWE	U+12B5	12B1
ኧ	KXWE	U+12C5	12C1
ኧ	GWE	U+1315	1311

Saisie. L'écriture éthiopienne contenant plus de 300 caractères, les unités de saisie au clavier doivent former un plus petit jeu d'entités, d'ordinaire 43+8 codes interprétés comme les coordonnées de la matrice de syllabes. Ces codes de saisie étant avant tout des entités transitoires transformées en caractères syllabiques avant leur stockage dans un texte, Unicode ne prévoit pas de codes de saisie particuliers.

Ordre de codage et tri. Les numéros des consonnes Unicode suivent l'ordre alphabétique traditionnel. Il peut différer de l'ordre de tri précis utilisé pour une des langues se servant de cette écriture, car plusieurs langues considèrent certaines paires ou triplets de syllabes comme équivalents lors la première passe de tri. Un dictionnaire amharique peut, par exemple, contenir une section qui débute par trois syllabes semblables à *H* :

U+1200 ሀ SYLLABE ÉTHIOPIENNE HÄ

U+1210 ከ SYLLABE ÉTHIOPIENNE HHÄ

U+1280 ኧ SYLLABE ÉTHIOPIENNE XÄ

Ainsi l'ordre de codage n'induit-il aucun ordre de tri pour aucune langue précise utilisant cette écriture.

Séparateurs de mots. U+1361 : ESPACE ÉTHIOPIENNE sert traditionnellement à séparer les mots. De nos jours, l'espace mot (U+0020 ESPACE) tend à le supplanter.

Signes diacritiques. D'ordinaire, l'écriture éthiopienne n'utilise pas de diacritique, on en retrouve toutefois à l'occasion dans les textes universitaires ou pédagogiques. Les caractères U+0308 ◌̈ DIACRITIQUE TRÉMA U+030E ◌̈̈ DIACRITIQUE DOUBLE LIGNE VERTICALE EN CHEF, notamment, servent à faire ressortir une lettre ou à indiquer une gémation (redoublement d'une consonne).

Chiffres. Le système numéral éthiopien n'utilise pas le zéro et n'emploie pas de notation positionnelle. Un nombre s'écrit grâce à une puissance de 100, chacune précédée d'un coefficient (de 2 à 99). Pour chacun des termes de cette série, la puissance 100ⁿ s'écrit à l'aide de *n* caractères **CENT** (transcrits à l'aide d'un digramme quand *n* = 2). On indique le coefficient à l'aide d'un chiffre représentant les dizaines et d'un autre représentant les unités, l'un de ces deux chiffres peut être absent si sa valeur est zéro.

On représentera donc le nombre 2 345 de la façon suivante :

$$\begin{aligned}
 2\ 345 &= (20+3) * 100^1 + (40+5) * 100^0 \\
 &= 20\ 3\ 100\ 40\ 5 \\
 &= \text{VINGT TROIS CENT QUARANTE CINQ} \\
 &= \tilde{\text{ᚐ}}\ \tilde{\text{ᚑ}}\ \tilde{\text{ᚒ}}\ \tilde{\text{ᚓ}}\ \tilde{\text{ᚔ}} \\
 &= \text{U+1373 U+136B U+137B U+1375 U+136D}
 \end{aligned}$$

Une langue transcrite à l'aide de l'écriture éthiopienne peut représenter le nombre 1 000 à l'aide d'un mot, comme la syllabe amharique U+123A ᚒ SYLLABE ÉTHIOPIENNE CHI. Un nombre comme 2 345 peut donc aussi s'écrire comme il se prononce. Le cas amharique est alors analogue à celui du français :

$$\begin{aligned}
 2\ 345 &= \text{DEUX mille TROIS CENT QUARANTE CINQ} \\
 &= \tilde{\text{ᚐ}}\ \tilde{\text{ᚒ}}\ \tilde{\text{ᚑ}}\ \tilde{\text{ᚒ}}\ \tilde{\text{ᚓ}}\ \tilde{\text{ᚔ}} \\
 &= \text{U+136A U+123A U+136B U+137B U+1375 U+136D}
 \end{aligned}$$

12.2 Chérokî

Chérokî : U+13A0 – U+13FF

L'écriture chérokîe s'utilise pour écrire la langue chérokîe. Le chérokî¹ est membre de la famille des langues iroquoiennes (certains linguistes le placent plutôt parmi le groupe natchez-muskoguen). Il est parent des langues parlées par les Goyogouins, les Tsonnontouans, les Onontagués, les Hurons-Wyandots, les Toscaroras, les Onneiouts et les Agniers². Cette parenté est cependant éloignée car les Chérokis quittèrent la région des Grands Lacs de l'Amérique septentrionale il y a 3 000 ans pour s'établir en Caroline du Nord, au Tennessee et en Géorgie. Le chérokî est la langue maternelle de quelque 20 000 personnes, même si pour la plupart d'entre eux elle n'est plus leur langue principale. Chérokî se dit « tsalagi » en chérokî, ce mot désigne à la fois la langue et le peuple.

C'est à Sequoyah qu'on attribue l'invention du syllabaire chérokî entre 1815 et 1821. Il contenait à l'origine 6 voyelles et 17 consonnes. Sequoyah s'efforça de créer un alphabet original, mais les premières lettres furent par la suite modifiées pour en faciliter l'impression. Le docteur Samuel Worcester créa la première police de caractères chérokîe. Utilisant les polices à sa disposition, il emprunta un certain nombre de lettres latines et les attribua à des syllabes chérokîes. À cette époque, la lettre chérokîe « HV » disparut et le syllabaire chérokî fut de la sorte réduit au nombre actuel de 85 lettres. Les presses du docteur Worcester imprimèrent 13 980 000 pages de texte en langues autochtones, la plupart en chérokî.

Tons. Chaque syllabe chérokîe peut se prononcer à l'aide d'un parmi quatre accents de hauteur (niveaux de ton). Ce ton peut passer à un ou deux autres tons au sein de la même syllabe. Cependant, ces changements de ton influent rarement sur le sens des mots. Les tons ne sont pas indiqués.

Casse et orthographe. L'écriture chérokîe est *unicamérale*, il se peut cependant que, pour faire ressortir certaines parties d'un texte, une lettre soit composée dans un corps plus grand que les autres. L'orthographe chérokîe n'est pas normalisée : chacun écrit les mots comme il les entend.

Chiffres. Malgré l'invention par Sequoyah d'un système numéral, il ne fut jamais adopté et Unicode ne le code pas. Les nations chérokîes utilisent les chiffres européens. Les Chérokîs distinguent nettement les ordinaux des chiffres cardinaux. Pour exprimer une suite de chiffres, le chérokî utilise des ordinaux. Ainsi, le chérokî numérote-t-il les chapitres d'un livre *Chapitre premier, Chapitre deuxième* et non *Chapitre un, Chapitre deux*...

Rendu et saisie. Le chérokî s'écrit de gauche à droite et ne nécessite pas de caractères combinatoires. Il existe plusieurs conventions de saisie pour le chérokî : certaines utilisent des touches mortes et des translittérations latines ; d'autres emploient des moyens mnémotechniques phonétiques liés aux lettres latines présentes sur les touches de clavier ; d'autres encore proposent des systèmes ergonomiques fondés sur la fréquence des syllabes dans la langue chérokîe.

Ponctuation. Le chérokî utilise la ponctuation latine ordinaire.

Normes. Il n'existe pas d'autre norme de codage pour le chérokî.

¹ Il s'agit du terme ISO/CEI 10646 pour cette langue qu'on écrit aujourd'hui souvent à l'anglaise *Cherokee Chéraqui est une des ancienne graphie française.*

² En Europe francophone, on emploie à présent les noms anglais de ces tribus, à savoir : Cayuga, Seneca, Onondaga, Wyandot-Huron, Tuscorora, Oneida et Mohawk.

12.3 Syllabaires autochtones canadiens

Syllabaires autochtones canadiens : U+1400 – U+167F

Ce bloc regroupe et unifie les caractères de divers syllabaires autochtones canadiens, l'apparence des caractères constituant l'élément unificateur. James Evans inventa ces syllabaires pour transcrire les langues algonquiennes vers la fin des années 1830. D'autres collectivités et groupes linguistiques adoptèrent par la suite cette écriture et, par la même occasion, ses principes de base décrits ci-dessous. Des nations amérindiennes, membres des familles linguistiques algonquiennes, inuktitutes (esquimaudes) et athabascanes, comptent parmi les principaux utilisateurs de cette écriture. Les organismes gouvernementaux, les entreprises, les écoles et les médias utilisent cette écriture.

Agencement. Les principes structurels prônés, en 1994, par le rapport du CASEC³, essentiellement un codage de glyphes, guident l'organisation de ce répertoire. La structure canonique de chaque série de caractères est constituée d'une consonne déclinée sous cinq variantes. D'ordinaire, lorsqu'on combine la consonne avec la voyelle /e/ la forme résultante pointe vers le bas ; avec la voyelle /i/ vers le haut ; la voyelle /o/ vers la droite et la voyelle /a/ vers la gauche. En fin de syllabe et en l'absence d'une voyelle subséquente, la consonne se compose dans un corps inférieur, en exposant. Exemple :

∨	∧	>	<	◀
PÉ	PI	PO	PA	P

L'utilisation des voyelles peut également varier. Ainsi, selon l'usage inuktitut, la syllabe U+1450 ◡ SYLLABE CANADIENNE TO se transcrit-elle « TOU » en lettres latines et non « TO » ; la structure du syllabaire demeure cependant identique, quelle que soit la langue.

Disposition. La disposition des signes respecte l'ordre algonquien (pointe vers le bas, le haut, la droite, la gauche) comme l'illustre l'exemple précédent. Les formes variantes se trouvent triées au sein de chaque série de caractères. Les variantes algonquiennes précèdent les variantes inuktitutes, elles-mêmes suivies des variantes athabascanes. Cette disposition est non seulement commode, mais elle respecte aussi la diffusion historique des syllabaires Evans ; cependant, elle n'implique aucune hiérarchie. Certains glyphes ne se présentent pas dans l'ordre habituel (bas/haut/droite/gauche) ; c'est le cas, par exemple, à partir de U+146B ◡ SYLLABE CANADIENNE KÉ. La forme du glyphe de base explique cette exception à la règle structurelle ; ceci n'a toutefois pas d'incidence sur la convention.

La longueur et la labialisation de voyelle modifient les séries de caractères par l'ajout de signes diacritiques (exemple : U+143E ◡ SYLLABE CANADIENNE POUÍ). On considère ces caractères modifiés comme des syllabes à part entière. On ne les décompose pas en caractère de base et signes diacritiques. Certaines familles de langues admettent des conventions différentes de placement de diacritique. Par cohérence, souci de simplicité et pour permettre la transcription de plusieurs langues amérindiennes dans un même texte, on attribue une valeur de code différente à chacune de ces variantes.

³ Comité canadien de codage des écritures syllabiques autochtones.

12.4 Mongol

Mongol : U+1800 – U+18AF

Les peuples mongols sont les représentants les plus importants du groupe culturo-linguistique altaïque, nom dérivé du massif de l'Altaï en Asie centrale. Par le passé, ces peuples dominèrent la haute Asie et même au-delà, régnant de la Baltique à la mer du Japon. Des langues voisines se parlent de la Finlande et de la Hongrie, en passant par la Turquie et l'Asie centrale, jusqu'en Corée et au Japon. Aujourd'hui, les peuples mongols se trouvent représentés politiquement en Mongolie même (l'ancienne République populaire de Mongolie, également appelée Mongolie extérieure) et en Mongolie intérieure (ou de son nom officiel chinois : la Région autonome de Mongolie intérieure). Des populations mongoles vivent aussi dans d'autres régions de Chine.

L'écriture mongole traditionnelle, codée dans ce bloc, s'utilise couramment depuis l'époque de Gengis Khan (début du XIII^e siècle). En République populaire de Mongolie, dans les années 1940, le système d'écriture cyrillique remplaça cette écriture traditionnelle. Toutefois, une loi de 1992 la rétablit officiellement. En pratique, l'écriture mongole traditionnelle et le cyrillique coexistent en Mongolie et dans les textes érudits occidentaux. Il n'existe pas de translittération directe entre ces deux écritures. Le tableau des noms de caractères mongols fournit une correspondance partielle, mais elle n'est pas garantie être bijective. Le bloc cyrillique d'Unicode contient tous les caractères cyrilliques nécessaires à l'écriture du mongol.

L'écriture mongole traditionnelle sert à écrire une langue littéraire mongole classique qui diffère des dialectes mongols parlés. S'inspirant du mongol, le mandchou, le sibé et le kalmouk (ou oïrat) créèrent chacun leur écriture. Pour rendre correctement les textes classiques bouddhistes, on rajouta au mongol classique, au todo et au mandchou les lettres *ali gali* (également appelées galik) qui permettent de transcrire le tibétain et le sanscrit.

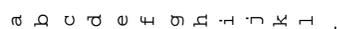
Directionnalité. Vers le XII^e siècle, les Mongols optèrent pour l'écriture ouïgoure, dérivée de l'araméen, une écriture sémitique lue de droite à gauche. À un certain moment, on transforma l'écriture comme si la page entière avait été tournée de 90 degrés dans le sens inverse des aiguilles d'une montre, de sorte qu'aujourd'hui le mongol s'écrit de haut en bas, en colonnes parallèles se succédant de gauche à droite. Le mongol possède une directionnalité unique parmi les écritures vivantes.

Dans un contexte moderne, le mongol s'écrit fréquemment en compagnie d'autres écritures ayant une autre directionnalité. En principe, chaque écriture devrait retenir sa propre orientation mais, quand cela n'est pas possible, une des écritures adopte la directionnalité de l'autre.

Si on adopte une orientation horizontale, les colonnes mongoles subissent une rotation supplémentaire de 90 degrés dans le sens inverse des aiguilles d'une montre, de sorte que les lettres se lient de la gauche vers la droite et que les colonnes se transforment en lignes à position équivalente (la première colonne devient la première ligne et ainsi de suite). Si on regarde de côté, l'ordre habituel des colonnes mongoles semblera inversé, mais cette orientation reste lisible pour de courts textes.⁴ Ce type de mise en page n'engendre pas d'effets bidirectionnels, car tout le texte est horizontal et se lit de gauche à droite.

⁴ En d'autres termes, si on désire insérer au milieu d'une phrase française le texte mongol

a e i
b f j
c g k
d h l

ce texte apparaîtra, si la largeur de ligne est suffisante, sous la forme de 

Normes. Ce n'est qu'après des années de recherches méticuleuses que des experts venus de Mongolie, de Chine et d'Occident élaborèrent le codage du mongol présenté ici. Ce groupe d'experts est sur le point de publier un document nommé *User's convention for System Implementation of the International Standard on Mongolian Encoding* qui précise les règles détaillées de formage de l'écriture mongole. La description ci-dessous ne fait que résumer l'ensemble de ces règles complexes de séquençement et de contextualisation.

Principes de codage. La relation entre la langue et l'écriture mongoles ressemble à certains égards à celle du français : une même lettre peut représenter différents sons, différentes lettres peuvent représenter le même son et la graphie des mots est parfois purement étymologique⁵ sans rapport avec la prononciation moderne. De surcroît, l'écriture mongole ressemble quelque peu à l'arabe : les lettres se joignent en cursives et adoptent des formes contextuelles calculées selon un ensemble de règles, une lettre pouvant revêtir jusqu'à dix formes différentes.

En dépit de ces variations, le principe de codage de base du mongol demeure simple, le même principe régit le français et l'arabe : un alphabet de base dont les lettres sont bien identifiées par les utilisateurs, mais dont les valeurs phonétiques ou les formes correspondantes le sont moins. Les éléments codés correspondent donc simplement aux lettres des alphabets traditionnels mongol, todo, sibé et mandchou, à leurs chiffres et à leurs signes de ponctuation.

Ponctuation. Les signes de ponctuation propres au Mongol comprennent U+1800 ◡ BIRGA MONGOL et U+1805 ◢ QUATRE-POINTS MONGOL qui indiquent, respectivement, le début et la fin d'une unité de texte (paragraphe ou article). Dans les textes mongols todo, on indique la reprise d'un mot coupé à l'aide d'un U+1806 ◣ TRAIT D'UNION VIRTUEL MONGOL TODO au début de la seconde ligne. Ce trait d'union s'emploie donc comme un U+00AD - TRAIT D'UNION VIRTUEL à la différence que le signe mongol se place au début de la ligne et non à la fin. Le bloc de *Ponctuation générale* contient certains symboles de ponctuation utilisés en mongol ; il s'agit, entre autres, de U+2048 ?! POINT D'INTERROGATION-EXCLAMATION et U+2049 !? POINT D'EXCLAMATION-INTERROGATION. Ces caractères servent à représenter ces deux signes côte à côte dans un texte vertical. Le mongol emploie le U+0020 ESPACE normal pour séparer les mots. D'autres types d'espace existent également, notamment, comme on le verra plus loin, pour insérer des coupures visuelles fines à l'intérieur des mots.

Œils. Le mongol est une écriture cursive, la forme de chaque lettre dépend d'ordinaire de sa position dans un mot. Il existe des formes pour le début (formes initiales), le milieu (formes médiales) et la fin d'un mot (formes finales). Les voyelles ont une forme isolée supplémentaire quand elles sont encadrées de blancs. Ces formes contextuelles ne sont pas toutes différentes : dans certains cas, un caractère conserve la même forme dans différentes positions (ainsi les formes initiale et médiale de U+182A ◥ LETTRE MONGOLE BA sont identiques) ; dans d'autres cas, deux caractères différents ont la même apparence (par exemple, la forme isolée de U+1824 ◧ LETTRE MONGOLE OU est la même que celle de U+1823 ◦ LETTRE MONGOLE O).

En plus des formes positionnelles, plusieurs lettres possèdent des variantes libres. L'orthographe et la grammaire, plutôt que la position du caractère au sein du mot, fixent la forme de la lettre. Le mongol utilise certaines ligatures, particulièrement celles formées par une voyelle précédée d'une consonne « arquée », c'est-à-dire une consonne telle que U+182A ◥ LETTRE MONGOLE BA à laquelle manque la hampe finale.

Le glyphe représenté dans les tableaux de caractères correspond généralement à la forme isolée pour les voyelles, à la forme initiale pour les consonnes et, si plusieurs variantes

⁵ En français, elle est même parfois faussement étymologique comme le « d » dans *poids*.

existent, à la forme la plus fréquente. À l'occasion, pour éviter que différentes lettres soient illustrées par le même glyphe, on a choisi d'autres formes.

Caractères de formage. Dans les cas où la suite contextuelle de lettres de base ne permet pas au moteur de rendu de choisir le glyphe approprié pour une lettre donnée. Unicode prévoit des caractères de formage supplémentaires afin de préciser l'œil à afficher. Sept caractères peuvent être insérés dans une suite normale de texte mongol codée pour sélectionner le glyphe rendu :

- U+180B SÉLECTEUR DE VARIATION LIBRE MONGOL UN
- U+180C SÉLECTEUR DE VARIATION LIBRE MONGOL DEUX
- U+180D SÉLECTEUR DE VARIATION LIBRE MONGOL TROIS
- U+180E SÉPARATEUR DE VOYELLES MONGOL
- U+200C ANTILIANT SANS CHASSE
- U+200D LIANT SANS CHASSE
- U+202F ESPACE INSÉCABLE ÉTROITE

Si on excepte U+202F ESPACE INSÉCABLE ÉTROITE et U+180E SÉPARATEUR DE VOYELLES MONGOL, ces caractères sont normalement invisibles. Leur seule raison d'être est de guider le processus de rendu lors de la sélection du glyphe approprié servant à représenter, dans un contexte donné, la lettre mongole de base. Un rapport détaillé des règles de sélection de formes est actuellement en cours d'élaboration. Le résultat produit par ces caractères de formage est résumé ci-dessous.

U+200C ANTILIANT SANS CHASSE et U+200D LIANT SANS CHASSE jouent, en mongol, le même rôle que dans les autres écritures (voir *Chapitre 14, Zones spéciales et caractères de formatage*). Ces caractères produisent le même effet de contextualisation sur les lettres avoisinantes que les lettres liantes et antiliantes habituelles à la différence qu'ils n'ont pas de chasse et sont donc invisibles.

U+202F ESPACE INSÉCABLE ÉTROITE : le mongol est une langue agglutinante, beaucoup de mots se forment donc par la juxtaposition après le radical d'un mot d'un ou plusieurs affixes distincts pour exprimer les rapports grammaticaux. Le radical et ses affixes constituent un seul mot, mais des espaces intra-mots séparent légèrement chacun des affixes, brisant l'intégrité visuelle du mot. U+202F ESPACE INSÉCABLE ÉTROITE représente ce blanc interne. Frontière de morphème, il joue également le rôle de sélecteur de variantes et peut modifier la forme contextuelle des lettres voisines.

Comme en mongol on utilise U+0020 ESPACE entre les mots, il est aussi permis d'employer U+00A0 ESPACE INSÉCABLE régulière afin d'éviter que deux mots ne soient séparés par une coupure de ligne. U+00A0 ESPACE INSÉCABLE n'a pas la même fonction que U+202F ESPACE INSÉCABLE ÉTROITE.

U+180E SÉPARATEUR DE VOYELLES MONGOL est une espace fine intra-mot qu'on ne retrouve qu'avant les voyelles finales U+1820 ᠠ LETTRE MONGOLE A et U+1821 ᠡ LETTRE MONGOLE È. Cette espace permet de préciser la forme précise du caractère précédent, de sélectionner une variante de forme pour ces voyelles et de produire un blanc étroit interne. Si, par erreur, cette espace apparaît avant un caractère autre que « A » ou « È », elle se comporte à la manière d'une U+202F ESPACE INSÉCABLE ÉTROITE.

Les caractères SÉLECTEUR DE VARIATION LIBRE MONGOL (UN, DEUX, TROIS) servent à distinguer les différentes formes d'une même lettre dans un même contexte, c'est-à-dire quand plus d'une forme de rendu est envisageable et qu'il faille en choisir une manuellement car celle-ci ne peut être sélectionnée par un algorithme. Les sélecteurs de variation libre suivent le caractère de base qu'ils modifient. N'étant pas productifs, ils peuvent être ignorés lorsqu'ils ne précèdent

pas immédiatement un des caractères de base répertoriés. Dans le cadre du processus de normalisation du mongol, des tableaux définissant le comportement précis des sélecteurs de forme sont à l'étude. Une fois disponibles, les résultats seront publiés sous la forme d'un rapport technique d'Unicode. Une utilisation conforme de ces caractères commande un respect scrupuleux de ces tableaux de sélection de forme.

12.5 Déséret

Déséret : U+13A0 – U+13FF

Le déséret est un alphabet phonémique (ou phonématique) inventé pour écrire l'anglais. Il fut conçu dans les années 1850 à l'Université de Déséret, devenue depuis lors l'Université de l'Utah. Son emploi fut favorisé par l'Église de Jésus-Christ des Saints des derniers jours, plus communément connue sous le nom de l'Église mormone, sous la présidence de Brigham Young (1801-1877). Le nom *déséret* provient d'un mot du Livre de Mormon qui signifierait « abeille », symbole utilisé par l'Église mormone pour représenter l'activité coopérative.

On retrouve parmi les concepteurs de l'alphabet déséret George Watt secrétaire de Brigham Young ; il avait été formé en dactylographie. Il est possible que, sous l'influence de Watt, l'alphabet phonotypique anglais de 1847 conçu par sir Isaac Pitman ait servi de modèle à l'alphabet déséret.

L'Église commanda deux polices différentes et publia quatre livres écrits en déséret. Le *Deseret News*, journal publié par l'Église, publia également quelques passages des écritures. Il existe, en outre, quelques documents historiques, des journaux intimes et quelques autres textes manuscrits écrits à l'aide de ce système d'écriture. Le déséret ne fut que rarement utilisé, par exemple en numismatique ou sur les panneaux officiels. Une pierre tombale de Cedar City (Utah) porte des inscriptions en déséret. Cette écriture ne connut toutefois jamais un grand succès et l'Église cessa d'encourager son utilisation après 1869. Aujourd'hui, l'alphabet déséret est principalement l'objet d'intérêt des historiens et des passionnés.

Structures. Cette écriture comprend trente-huit lettres. Cet alphabet est *bicaméral*, cependant les majuscules ne diffèrent des minuscules que par la force de leurs œils (leurs tailles donc) et non par la forme de ceux-ci. Les lettres sont regroupées en classes phonématiques : les lettres qui représentent des sons similaires se côtoient. Les consonnes forment notamment des paires voisées/sourdes.

Tri. Il n'existe pas de documents historiques en déséret qui soient triés, on suppose que le tri aurait été effectué selon l'ordre canonique dans lequel les lettres de l'alphabet apparaissent.

Conventions typographiques. Le déséret s'écrit de gauche à droite. La ponctuation, la mise en majuscules et les chiffres sont identiques à ceux de l'anglais. Tous les mots se prononcent de façon phonémique (phonématique) à l'exception de mots courts dont la prononciation correspond au nom de la lettre même. Ainsi, écrit-on parfois *Y* (la lettre THEE, U+1041C) pour représenter les mots *the* ou *thee*. Cette lettre se prononce habituellement comme un /*ö*/.

Normes. Il n'existe pas de norme préexistante régissant le codage du déséret.