

ITU and UNESCO Global Symposium on Promoting the Multilingual Internet

Geneva, 9-11 May 2006

Paper Submitted
by:

Marcel Diki-Kidiri

Union Latine - Unesco

L'accès au cyberspace des les langues peu dotées

Introduction

Le cyberspace est ouvert à toutes les langues du monde car son infrastructure n'est pas soumise à l'autorité d'un pouvoir central qui déciderait de son utilisation. Il suffit, en principe, de disposer d'un ordinateur relié à un fournisseur d'accès Internet pour afficher sur la Toile des données textuelles, iconographiques ou sonores dans la langue de son choix. La mise en oeuvre de ce principe qui est un facteur fondamental de la démocratie au niveau mondial, nécessite, cependant, que soient réunies un certain nombre de conditions techniques et de ressources humaines et financières que nous allons examiner dans le présent document. En rédigeant celui-ci, nous avons voulu répondre aussi simplement que possible à la question simple de savoir comment faire pour qu'une langue peu dotée en ressources linguistiques et/ou informatiques, voire en ressources humaines, puisse arriver à trouver une place dans le cyberspace et y être active?

Il nous faut d'emblée répondre à une question préalable. Qu'est-ce qu'une langue peu dotée? Il s'agit d'une langue qui ne dispose pas suffisamment, voire pas du tout, d'un certain nombre de ressources dont sont généralement dotées les grandes langues du monde, à savoir: une orthographe stable dans un système d'écriture donné, des ouvrages de référence (grammaires, dictionnaires, ouvrages littéraires) des oeuvres de diffusion massive (presse écrite et audiovisuelle, films, chansons et musique) des ouvrages techniques et d'apprentissage (publications techniques et scientifiques, ouvrages didactiques, manuels) divers supports de communication au quotidien dans la langue (affiches, publicités, courriers, notices, mode d'emplois, etc.) et pour finir, un nombre abondant d'applications informatiques dans cette langue. Sur le plan des ressources humaines, une langue peu dotée peut devenir une langue en danger de disparition si elle n'est plus parlée que par un petit nombre de locuteurs. Il devient alors nécessaire, pour la sauver, d'augmenter le nombre de ses locuteurs en l'enseignant par tous les moyens techniques possibles.

Heureusement, une langue peu dotée n'a pas nécessairement tous ces handicaps à la fois. Elle peut être majoritaire, écrite, enseignée à l'école, et manquer cruellement de ressources informatiques ou même de ressources linguistiques en quantité et en qualité suffisantes. Il est donc plus juste de dire que les langues peu dotées le sont de manière très variée, depuis les langues en grand danger de disparition jusqu'aux langues émergentes qui possèdent déjà une bonne partie de ces ressources mais en nombre estimé insuffisant et incomplet. De nombreuses langues sur tous les continents répondent à cette définition des langues peu dotées. On peut citer en Europe le breton, l'occitan, le basque, en Amérique, la quasi totalité des langues amérindiennes, en Asie, le myanmar, pour n'en citer qu'une parmi des centaines, en Océanie, la quasi totalité des langues autochtones des îles polynésiennes, micronésiennes et mélanésiennes. En Afrique où se parlent un tiers des langues du monde, soit environ 2000 langues, les langues les mieux dotées (afrikaans, kiswahili, hausa, ...) se comptent sur les doigts d'une main et font partie des langues émergentes, donc des langues peu dotées.

Nous partons donc de l'hypothèse la plus défavorable, à savoir, une langue qui fut jadis porteuse d'une culture florissante, mais qui n'est plus parlée aujourd'hui que par une poignée de personnes âgées dans un petit village au fin fond de la brousse africaine, disons en Afrique centrale pour être aussi loin que possible de toutes les côtes. Appelons cette langue le *ndeka*. Une telle langue a pratiquement aucune chance d'accéder un jour au cyberspace, car elle risque bien de disparaître à tout jamais à la mort de son dernier locuteur, et ils ne sont plus très nombreux à la parler. Voici qu'un jeune étudiant issu de ce village trouve que des outils comme SIPI, Wikipedia et bien d'autres encore offrent une formidable opportunité de conserver la mémoire de cette langue et à travers elle celle de son village et de la culture de ses ancêtres. Ce serait même le moyen le plus sûr de faire apprendre cette langue aux jeunes de sa génération et ainsi lui redonner une nouvelle vie. Il vient nous trouver pour nous demander de le guider dans son projet. Avec cette hypothèse en tête, notre

exposé se veut didactique afin d'accompagner, pas à pas, tous ceux qui nous rejoindront à n'importe quelle étape du chemin pour porter jusque dans le cyberspace toutes les langues peu dotées quelle que soit leur fortune.

1. Élaboration de ressources linguistiques

1.1. De l'oral à l'écrit

L'objectif de cette première étape est de doter la langue d'un minimum de ressources linguistiques, à savoir: une orthographe dans un système d'écriture, une grammaire écrite, un dictionnaire et un recueil de textes aussi fourni que possible. C'est un travail de linguiste que nous allons évoquer dans les grandes lignes en visant les objectifs intermédiaires les plus indispensables.

11.1. *Recueil de textes*

La première chose à faire est d'aller trouver les locuteurs de la langue et d'enregistrer autant de textes que possibles, des récits de vie, des contes, des proverbes, des chansons, de la poésie, des légendes, des conversations, des narrations, etc. Pour faire ces enregistrements, il vaut mieux apprendre des techniques de recherche sur le terrain, et utiliser autant que possible des appareils de qualité professionnelle, afin d'obtenir la meilleure qualité sonore possible, car le son ainsi enregistré devra subir de nombreux traitements ultérieurement. Ce serait dommage qu'à la deuxième copie la qualité du son se dégrade rapidement! Dans des conditions de travail normales, les enregistrements sont immédiatement traduits avec l'assistance des locuteurs de la langue et transférés dans une base de données textuelles.

11.2. *Transcription phonétique*

En utilisant l'alphabet phonétique international (API) les linguistes peuvent transcrire très exactement n'importe quel son de n'importe quelle langue du monde. Les textes recueillis plus haut peuvent donc être ainsi transcrits. Mais il faut bien comprendre que la transcription phonétique reproduit très fidèlement chaque son tel qu'il est prononcé en un moment donné. Or dans un mot comme *papa* le premier P est légèrement plus explosif que le second parce qu'il est en début de mot et l'autre entre deux voyelles. Cette différence est habituellement négligée car elle n'entraîne aucune conséquence pour le sens du mot, mais une bonne transcription phonétique la reflètera comme ceci [p[◌]apa]¹. La transcription phonétique est indispensable pour permettre une bonne analyse des sons de la langue, mais elle n'est pas du tout la meilleure façon de l'écrire pour un usage courant.

11.3. *Analyse et notation des phonèmes*

Si la transcription phonétique est faite correctement, c'est-à-dire avec minutie et fidélité aux sons réellement prononcés, il sera alors aisé de procéder à une analyse phonologique dont l'objectif est

1 On utilise les crochets carrés pour indiquer que le mot est transcrit phonétiquement.

d'identifier les sons utiles de la langue qu'on appelle les *phonèmes*. Il s'agit de sons qui entraînent une différence de sens quand ils changent. Par exemple, en français, on dira que /p/² et /b/ sont deux phonèmes différents parce que leur différence de prononciation permet à elle seule de distinguer les mots « pain » et « bain ».

Par contre, certaines personnes prononcent le mot *roi* avec un R roulé et d'autres avec un R grasseyé. Ces deux « R » sont phonétiquement différents et transcrits respectivement [r] et [R]. Pourtant cette différence phonétique n'entraînant aucune différence de sens, ces deux sons ne représentent qu'un seul phonème que l'on conviendra de noter /r/. Un autre exemple, la suite de sons [gz] que l'on trouve dans le mot « exact » prononcé [egza] et la suite de sons [ks] du mot « extra » prononcé [ekstRa] représentent toutes deux le même phonème que l'on conviendra de noter /x/.

Ainsi, l'analyse phonologique d'une langue aboutit à établir la liste complète des phonèmes de cette langue et la manière de les noter. On peut alors utiliser cette liste de phonèmes pour écrire un texte. On parle alors d'une *notation phonologique* par opposition à la *transcription phonétique*. Dans une notation phonologique, ce sont les phonèmes, et non plus les sons, qui sont représentés. La notation sera donc bien plus économique et fournit une bonne base pour l'élaboration d'une orthographe.

11.4. Analyse et notation des tons

On peut prononcer une syllabe sur une hauteur de voix (ou *registre*) relativement grave, moyenne ou aigue comme les notes de musiques *do ré mi*. La majorité des langues africaines exploitent cette variation de la hauteur de la voix pour supporter des différences de sens avec une même combinaison de phonèmes. Ainsi, en *sängö*³, les mots [ká] « là-bas », [kā] « plaie », et [kà] « et » sont vraiment trois mots différents, et ce ne sont pas du tout des homonymes. On qualifie ces langues de *langues à tons*.

Une langue à tons peut exploiter de deux à cinq registres différents, mais la plupart des langues africaines n'en exploitent que deux ou trois. Avec deux registres, on distingue un ton haut et un ton bas. Avec trois registres, on distingue un ton bas, un ton moyen et un ton haut. Au-delà on distingue, en plus, un ton supra-haut et/ou un ton infra-bas. C'est à l'issue d'une analyse tonologique que l'on détermine si une langue est à tons ou non et, dans l'affirmative, le nombre de registres utilisés, la fréquence des tons, la meilleure façon de les noter, etc.

Dans une langue à tons, un mot est nécessairement composé au moins d'une voyelle et d'un ton, et une forme comme *ka*, évoquée plus haut, ne revêt de sens que lorsqu'on précise le ton sur lequel il doit être prononcé. Il faudrait donc logiquement envisager de noter les tons dans le système orthographique de la langue. Toutefois, toutes les langues à tons ne font pas un usage identique des tons. Certaines se servent des tons non seulement pour distinguer des mots au niveau du lexique, mais aussi des formes verbales et des modalités personnelles et aspectuelles, donc des valeurs grammaticales. Dans de telles langues les tons occupent une place bien trop importante pour être négligés. Par contre, dans les langues qui ne recourent aux tons qu'au niveau du lexique et de façon relativement limitée (par exemple, pour ne distinguer que des mots brefs où l'ambiguïté peut être plus élevée) on a souvent pris le parti de ne pas noter les tons dans l'orthographe courante. Il y a donc là une décision à prendre au cas par cas. Toutefois, il faut souligner que pour les travaux de recherche linguistique sur une langue à tons, la notation des tons est absolument indispensable, car dans ce genre de travaux, on doit prendre en compte toutes les informations sur la réalité de la langue, et les tons constituent une dimension importante et définitoire d'une langue à tons.

2 On utilise les barres obliques pour indiquer que ce qui est à l'intérieur est un phonème ou une suite de phonèmes.

3 Le *sängö* est la langue nationale de la République Centrafricaine et aussi langue officielle conjointement au français.

11.5. *Elaboration d'une orthographe*

Les études phonologiques et tonologiques aboutissent à un système de notation de la langue qui représente correctement les sons utiles de la langue, et ce, de façon aussi économique que possible. Il permet à 90% d'écrire ce qu'on prononce sans raffinement phonétique et cela suffit pour que l'on soit tenté de se contenter de la notation phonologique et de faire l'économie d'études orthographiques. En fait, celles-ci s'avèrent indispensables pour doter vraiment une langue d'un système conventionnel capable de prendre en compte l'ensemble des besoins d'expression écrite des locuteurs. En effet, à la différence de la notation phonologique qui ne reflète que les sons utiles de la langue, une orthographe bien conçue intègre aussi la notation des relations grammaticales et des idées. Considérons, par exemple, les exemples suivants:

- a) /ɛl ʃāt /
- b) elle chante
- c) elles chantent

L'exemple (a) présente la notation phonologique qui correspond aussi bien à la phrase (b) qu'à la phrase (c). Si le français était écrit uniquement avec une notation phonologique, il n'y aurait aucun moyen de distinguer ici le singulier du pluriel comme le permet la convention orthographique.

Prenons un autre exemple en sängö, une langue qui distingue un ton haut, un ton moyen et un ton bas. En notation phonologique, ces tons sont notés respectivement par un accent aigu [á], un accent plat [ā] et un accent grave [à], mais ces accents ne se trouvent pas sur les machines courantes en Centrafrique. Une orthographe est pourtant destinée à un large public, et l'on doit pouvoir écrire la langue aussi bien à la main qu'à la machine et avec l'ordinateur. C'est pourquoi il faut élaborer une convention orthographique comportant un ensemble de règles qui permettent d'écrire la langue de façon plus pratique.

Le tableau suivant montre les différentes étapes vers l'élaboration d'une orthographe optimale et stable pour une langue, notamment le sängö :

<i>Transcription phonétique</i>	<i>Notation phonologique</i>	<i>Orthographe 1 (47 %)</i>	<i>Orthographe 2 (53 %)</i>	<i>Traduction</i>
1. [sùkùlà]	/sùkùlà/	sukûla	sukûla	laver
2. [sā ⁰ gō ^ˉ]	/sāngō/	sängö	sahngo	sango (langue)
3. [ᵐbásáᵐbálá]	/ᵐbásáᵐbáláa/	mbâsâmbâlâ	mbâssambala	sept
4. [māī ⁰ gō ^ˉ]	/māīngō/	māīngö	mayhngo	évolution
5. [k ^w ā] ~ [k ⁰ à]	/kùà/	kua	kua	travail
6. [k ^w ā] ~ [k ⁰ ā]	/kūā/	küä	kwa	poil
7. [k ^w á] ~ [k ⁰ á]	/kúá/	kûâ	kwâ	mort

On peut voir que la transcription phonétique reflète la prononciation des mots telle qu'ils sont enregistrés au magnétophone. Elle est donc propre à la recherche linguistique, mais trop fine pour servir de base à une orthographe courante destinée au grand public. La notation phonologique de la deuxième colonne ne retient que les sons utiles, les phonèmes, et de ce fait, elle est plus propice à servir de base à une orthographe courante. Mais le marquage de tous les tons constitue un obstacle pour une écriture pratique et une lecture rapide. En outre, les diacritiques, ces symboles utilisés pour marquer les tons (accent aigu, accent plat et accent grave) ne sont pas disponibles sur toutes

les machines et ne peuvent être rendus correctement qu'avec un ordinateur. Ce qui n'est pas à la portée de tous les utilisateurs d'une langue peu dotée. La convention orthographique de la troisième colonne permet d'économiser 47 % de la fréquence des tons et de les écrire avec le tréma et l'accent circonflexe, deux symboles qui figurent en touche libre sur toutes les claviers français des machines à écrire, et des ordinateurs utilisés en Centrafrique. Malgré cela, cette convention laisse tout de même 53 % de la fréquence des tons à noter avec des accents. Pour réduire encore davantage le taux de fréquence, on a effectué une réforme orthographique dont on peut voir les effets en quatrième colonne. Cette nouvelle convention permet d'économiser 53 % de la fréquence des tons dans un texte. On peut dire qu'à ce stade, le sāngö (ou sahngo⁴) est enfin doté d'une orthographe vraiment optimale et stable.

11.6. Réalisation d'ouvrages de références

L'analyse des textes recueillis plus haut permet d'élaborer une grammaire complète, qui couvre la phonologie, la syntaxe, les types d'énoncés et le discours. Généralement, une grammaire de référence est écrite pour servir de base à plusieurs ouvrages plus didactiques tels que des grammaires d'apprentissage de la langue et, le cas échéant, des manuels de classe.

L'étude du vocabulaire tiré des mêmes textes, et éventuellement complété par des enquêtes ciblées, permet de rédiger un dictionnaire de référence aussi complet que possible, lequel servira de base à des lexiques ou de petits dictionnaires d'apprentissage. Les mêmes textes vont servir de contenus à des livres de lecture, des nouvelles, afin d'inciter à la création d'autres ouvrages littéraires si ce n'est à partir de la tradition orale, du moins à partir de l'observation de la vie de tous les jours dans la communauté de locuteurs.

La réalisation d'ouvrages de référence et d'ouvrages utilitaires dans la langue n'est pas une étape marginale dont on peut faire l'économie. Les langues peu dotées ont toujours besoin d'ouvrages de ce genre, même quand elles font partie des langues émergentes, et à plus forte raison quand elles sont en danger de disparition. En outre ces ouvrages sont indispensables pour renforcer, voire créer, une conscience communautaire chez les locuteurs, surtout quand il n'en reste que peu. Ils sont indispensables pour enseigner la langue aux jeunes et ainsi augmenter leur capacité à apprendre d'autres choses, et pour alphabétiser les adultes et par ce moyen leur ouvrir la voie à plus de connaissance.

1.2.. Développement terminologique

Pour que les membres de la communauté de locuteurs utilisent leur langue comme médium de communication dans le cyberspace, il est absolument indispensable que cette langue soit dotée des termes techniques nécessaires pour exprimer les réalités de ce nouvel espace. Par exemples: courrier électronique, adresse électronique, copie conforme, se connecter, brancher (les appareils), télécharger, publier, réseaux, site, page Web (page sur Toile), naviguer, etc. Rien que le vocabulaire nécessaire pour nommer le matériel informatique que l'on utilise constitue déjà un lexique conséquent qu'il faut entièrement créer surtout que dans la majorité des cas, ce matériel ne fait pas partie des objets culturels traditionnels et n'est donc ni connu ni nommé. Dès lors, les méthodes de développement terminologique devront être mis en oeuvre au sein de la communauté pour doter la langue de néologies culturellement acceptables. Le développement terminologique est une activité continue et permanente qui sera de plus en plus générée par la communauté des locuteurs elle-même au fur et à mesure qu'elle développera une culture accrue du cyberspace. Il ne constitue donc pas une « étape » sur le chemin de l'accès au cyberspace, mais une activité de consolidation qu'il est

4 Ici, le H inséré devant NG signifie que le ton moyen du A vaut aussi pour le O. C'est l'application d'une règle plus générale qui permet de ne pas recourir aux diacritiques pour informer sur les tons du mots dans certains contextes.

nécessaire de commencer à un moment donné et de poursuivre indéfiniment.

2. Élaboration des ressources informatiques

2.1. Systèmes d'écriture et orthographe

Avant d'aborder cette étape, il est utile de préciser certains concepts trop souvent mal compris ou confondus surtout par les non spécialistes, à savoir ce qu'on appelle un système d'écriture, un système orthographique (également appelée une convention orthographique ou tout simplement une orthographe) un caractère, un jeu de caractères et une police de caractères.

21.1. Les systèmes d'écriture

Considérons l'écriture latine, l'écriture arabe, l'écriture hébraïque, l'écriture chinoise, l'écriture éthiopique, l'écriture hiéroglyphique, l'écriture cunéiforme, l'écriture maya. Chacune utilise un ensemble de symboles spécifiques avec des règles d'assemblage qui lui sont propres. Ce sont tous des systèmes d'écriture différents. L'Afrique compte une bonne douzaine de systèmes d'écriture dont les plus utilisés sont l'écriture latine, l'écriture arabe, l'écriture éthiopique, l'écriture n'ko, et l'écriture tifinagh. Lorsque l'on se propose d'écrire une langue qui n'a jamais été écrite, il convient de savoir dans quel système d'écriture le faire pour que la communauté des locuteurs puisse immédiatement s'en servir. En général, on tient compte du système d'écriture le plus courant dans la région qui sert déjà à écrire d'autres langues parlées sur le même territoire.

21.2. Les systèmes orthographiques

En utilisant un même système d'écriture, on peut écrire des langues très différentes. Ainsi l'écriture latine sert à écrire la majorité des langues d'Europe de l'ouest, d'Amérique et d'Afrique. Mais chaque langue exploite différemment les ressources du système d'écriture en fonction de ses propres structures phonologiques, syntaxiques, énonciatives et sémiotiques. Les règles d'écriture qui en résultent constituent un système orthographique propre à chaque langue. Pour illustrer cela, considérons le fait de dédoubler une consonne simple dans les quatre langues suivantes :

En italien, la gémation est importante car elle peut entraîner une différence de sens. Par conséquent, le dédoublement de la consonne sert à noter la gémation et oblige à prononcer fortement la consonne dédoublée : *tutto* tout, *oggi* aujourd'hui.

En français, le redoublement d'une consonne n'a qu'une valeur étymologique ou esthétique et n'entraîne pas du tout une prononciation particulièrement forte de la consonne : *addition*, *apprécier*, *atteler*. Comparer avec *adorer*, *apercevoir*, *atelier*. Que l'on écrive ces mots avec ou sans consonnes doubles, la prononciation est la même et le sens aussi ne change pas.

En anglais, même s'il ne s'agit pas d'une règle générale, le dédoublement de la consonne après une voyelle sert souvent à sélectionner la prononciation de la voyelle en question. Ainsi, le /i/ se

prononce [ai] dans *write* écrire, *hide* cacher, mais [i] dans *written* écrit, *hidden* caché. Le dédoublement du /t/ et du /d/ ne sert pas à signaler une prononciation forte de ces consonnes comme en italien, mais tout simplement à sélectionner entre les deux prononciations possibles de la voyelle /i/ qui les précède.

En sängö (ou sahngo) l'une des règles de l'orthographe réformée utilise le dédoublement de la consonne simple pour marquer le début d'une suite tonale identique. Ainsi, au lieu d'écrire *täsëmä* rangée, où le ton moyen est marqué par un tréma sur chacune des trois syllabes du mot, on écrit *tassema*, le dédoublement du /s/ suffit à indiquer que le ton moyen de la première syllabe se reporte sur toutes les syllabes qui suivent jusqu'à la fin du mot. De même, *mbâsâmbâlâ*, sept (avec une suite de tons hauts) est réécrit *mbâssambala* en application de la même règle. Le redoublement du /s/ suffit à indiquer que le ton haut du premier /â/ reste valable pour toutes les syllabes suivantes jusqu'à la fin du mot.

On voit bien comment les règles orthographiques sont propres à chaque langue même quand elles utilisent le même système d'écriture.

2.2. Jeux et polices de caractères

On appelle *caractère* tout symbole discret et significatif faisant partie d'un système d'écriture. Ainsi, dans le système d'écriture latin, l'espace blanc qui sépare deux mots est un caractère au même titre que l'accent aigu, une lettre, une virgule ou un chiffre.

22.1. Les jeux de caractères

Une fois l'orthographe de la langue établie, on peut faire la liste exhaustive des caractères nécessaires pour écrire la langue dans les règles de cette orthographe. Cette liste constitue ce qu'on appelle un *jeu de caractères*. En le comparant avec les caractères utilisés en standard dans les ordinateurs, on peut vérifier si tous les caractères du jeu en font partie ou non. Le cas le plus facile à traiter est celui où tous les caractères du jeu font partie des caractères standard, car dans ce cas, il n'y a rien de spécial à faire avant de pouvoir utiliser un ordinateur pour écrire la langue. Mais il est bien plus fréquent de constater que tous les caractères dont on a besoin ne se trouvent pas en standard dans les jeux de caractères des ordinateurs. On parle alors de *caractères spéciaux* pour désigner ceux qui ne s'y trouvent pas. Un jeu de caractères peut comporter, par exemple, sept (7) voyelles et vingt-quatre (24) consonnes, quelques diacritiques (circonflexe, tréma, tilde), dix chiffres (de 0 à 9) des opérateurs mathématiques (+, -, %, <, >, =) des signes de ponctuation, et enfin des symboles comme l'astérisque (*) le paragraphe (§) ou encore les symboles monétaires (\$, £, €). Un autre jeu de caractères aura peut-être cinq voyelles et non sept, et trente ou quarante consonnes et non vingt-quatre, etc.

22.2. Les polices de caractères

Une police de caractères est d'abord un ensemble de définitions des formes de caractères dessinées dans un même style artistique. Les polices de caractères sont aujourd'hui des logiciels qui comprennent, certes, un stock de jeux de caractères mais aussi des règles de représentation de ces caractères à l'écran, et des fonctions informatiques pour les gérer « intelligemment ». La création d'une bonne police de caractères de qualité professionnelle nécessite des connaissances pointues et une grande compétence technique. C'est pourquoi les polices de caractères sont des produits industriels protégés par des marques déposées, tels que Times New Roman, Garamond, Arial, etc. Les polices gratuites sont rarement satisfaisantes, bien qu'il en existe d'excellentes dans le domaine des logiciels libres. Si donc vous avez besoin de créer des caractères spéciaux pour écrire votre langue, après les avoir cherchés vainement dans les polices existantes, il vous faudra vous adresser

à une personne compétente ou à une entreprise spécialisée pour vous faire fabriquer des polices avec le jeu de caractères de votre choix. Vous aurez, en plus, la garantie que ces polices respectent les normes standard internationales.

2.3. Codage des caractères

Il est important de bien comprendre comment les caractères sont codés pour être utilisés en informatique, ne fut-ce que pour pouvoir poser les bonnes questions aux spécialistes lorsqu'on souhaite faire fabriquer des polices de caractères comportant des caractères spéciaux pour sa langue. La plupart des langues africaines récemment écrites le sont avec l'alphabet latin enrichi de nombreux caractères tirés du stock de caractères définis par l'Association internationale de Phonétique (API) et adaptés par l'Institut africain international (IAI). Pour rester simple, nous allons nous limiter à ce cas de figure, sans prendre en compte d'autres systèmes d'écriture. Du reste la question de caractères spéciaux ne se pose pas pour les autres systèmes d'écriture tels que le n'ko, l'éthiopique, le tifinagh, puisque tous les caractères de ces systèmes sont normalisés et codés, ceux qui ne le seraient pas sont des variantes de caractères normalisés et non des caractères spéciaux ayant des valeurs distinctes.

23.1. Le codage sur un octet

Imaginez un fil électrique sur laquelle la seule façon de créer un événement est de couper le courant puis de le rétablir. On obtiendrait une valeur nulle (0) quand le courant est coupé et une valeur positive (1) quand le courant est à nouveau rétabli. En un point A donné, le long du fil, ces deux valeurs sont les seules possibles et permettent de répondre « Oui ou Non » à la question « le courant passe-t-il? ». On peut donc voir le point A comme une unité minimale de stockage d'informations qui ne peut prendre que deux valeurs : 0 ou 1. On représente cette unité par un chiffre binaire (*binary digit* en anglais) appelé *bit*⁵. Si avec un seul bit on peut exprimer 2 valeurs (0, 1), avec 2 bits on peut exprimer 4 valeurs (00, 01, 10, 11) et avec 3 bits, 8 valeurs (000, 001, 010, 011, 100, 101, 110, 111). La progression est donc exponentielle. Le tableau suivant montre le nombre des valeurs que l'on peut exprimer selon le nombre de bits utilisés:

Nombre de bits	1	2	3	4	5	6	7	8
Nombre de valeurs	2	4	8	16	32	64	128	256

Si, par convention, nous décidons d'attribuer une lettre différente à chacune des valeurs obtenues avec 3 bits, nous aurions par exemple ceci :

000 = a 001 = b 010 = c 011 = d 100 = e 110 = f 111 = g

Il ne nous serait pas possible de coder toutes les lettres de notre alphabet. En réalité, on n'a jamais codé les caractères sur 3 bits mais d'abord sur 7 bits, puis sur 8 bits. Les caractères codés sur 7 bits sont numérotés de 0 à 127 et constitue le jeu appelé ASCII. Ce jeu contient les 26 lettres de l'alphabet latin de base et aucune lettre accentuée. L'ajout d'un seul bit a permis de coder 128 caractères supplémentaires de 129 à 255. Cette plage a été utilisée pour loger les caractères spéciaux dont on a besoin aussi bien pour l'imprimerie et le dessin de graphiques que pour diverses langues. Les jeux de caractères ainsi obtenus contiennent bien chacun 256 caractères codés de 0 à 255, mais sont tous partiellement différents. Ils correspondent au standard ISO-8859-*n* (où *n* indique le numéro d'un jeu particulier). Ces codes sont toujours représentés dans la machine par une suite de bits. Un paquet de 8 bits est appelé un *octet*.

⁵ Mot valise composé par le début de *binary* et la fin de *digit* : *bi-* + *-it* > *bit*.

Ce sont les polices de caractères qui établissent le lien entre, d'une part, les codes (suite de 0 et de 1) représentant les caractères dans l'ordinateur, et d'autre part, les dessins correspondants (appelés *glyphes*) qu'elles affichent à l'écran. Ainsi, la police Times New Roman va lire le code 97 correspondant au « a minuscule latin » et affichera à l'écran l'un des « a » suivants : a, **a**, *a*, **a**, selon le choix de l'utilisateur. Ces quatre « a » ont en commun le style propre à la famille de la police Times New Roman caractérisé par le plein et le délié et surtout le petit ergot (appelé *sérief*) à la base de la lettre. La police Arial agira pareillement, mais avec des lettres bâton qui n'ont pas de sérifs.

23.2. Limite du codage sur un octet

Avec l'accélération à la fois de la puissance des ordinateurs, des échanges internationaux et des besoins de l'édition, on a vite atteint les limites du codage sur un octet. L'un des inconvénients majeurs de l'utilisation du codage sur un seul octet réside dans la nécessité d'utiliser un jeu de caractères spécifique pour une langue particulière qui a besoin de caractères spéciaux. Ainsi, le jeu standard ISO-8859-1 (aussi dénommé ISO-latin-1) associe au code 198 la lettre « *Latin capital lettre æ* » c'est-à-dire « Æ ». Mais cette lettre peut être inutilisée par votre langue peu dotée qui a besoin d'une autre lettre « e ouvert », soit « □ », non prévue. En utilisant les outils logiciels adéquats, l'utilisateur lambda que vous êtes a la possibilité de redéfinir le caractère codé à cet endroit afin de remplacer « Æ » par « □ » dans une police locale. En conséquence, pour lire un texte écrit dans cette langue avec cette police locale, il faut avoir ladite police sur sa machine. Or une police bien faite est un produit industriel, souvent commercialisé, et que l'on ne peut pas distribuer à tout va gratuitement. L'utilisateur occasionnel n'a pas non plus envie d'investir dans une police nouvelle chaque fois qu'il change de langue pour consulter un document! Même si on dispose d'polices gratuites, tout envoi de document (à un collègue, un éditeur, ou un imprimeur) doit obligatoirement s'accompagner d'une police locale « jetable » après usage! Ces contraintes ne sont qu'une petite partie des inconvénients des polices locales, qui rendent, par ailleurs, d'appréciables services à leurs utilisateurs lorsqu'elles sont localement partagées. De toutes les façons, avec un maximum de 256 codes possibles, le codage sur un seul octet est évidemment insuffisant pour coder les 2000 caractères de base de l'écriture chinoise, et ceci vaut aussi pour le japonais et le coréen.

ISO 8859-1 Latin-1

0	32	64	@	96	'	128	160	192	À	224	à
1	33	65	A	97	a	129	161	193	Á	225	á
2	34	66	B	98	b	130	162	194	Â	226	â
3	35	67	C	99	c	131	163	195	Ã	227	ã
4	36	68	D	100	d	132	164	196	Ä	228	ä
5	37	69	E	101	e	133	165	197	Å	229	å
6	38	70	F	102	f	134	166	198	Æ	230	æ
7	39	71	G	103	g	135	167	199	Ç	231	ç
8	40	72	H	104	h	136	168	200	È	232	è
9	41	73	I	105	i	137	169	201	É	233	é
10	42	74	J	106	j	138	170	202	Ê	234	ê
11	43	75	K	107	k	139	171	203	Ë	235	ë
12	44	76	L	108	l	140	172	204	Ì	236	ì
13	45	77	M	109	m	141	173	205	Í	237	í
14	46	78	N	110	n	142	174	206	Î	238	î
15	47	79	O	111	o	143	175	207	Ï	239	ï
16	48	80	P	112	p	144	176	208	Ð	240	ð
17	49	81	Q	113	q	145	177	209	Ñ	241	ñ
18	50	82	R	114	r	146	178	210	Ò	242	ò
19	51	83	S	115	s	147	179	211	Ó	243	ó
20	52	84	T	116	t	148	180	212	Ô	244	ô
21	53	85	U	117	u	149	181	213	Õ	245	õ
22	54	86	V	118	v	150	182	214	Ö	246	ö
23	55	87	W	119	w	151	183	215	×	247	÷
24	56	88	X	120	x	152	184	216	Ø	248	ø
25	57	89	Y	121	y	153	185	217	Ù	249	ù
26	58	90	Z	122	z	154	186	218	Ú	250	ú
27	59	91	[123	{	155	187	219	Û	251	û
28	60	92	\	124		156	188	220	Ü	252	ü
29	61	93]	125	}	157	189	221	Ý	253	ý
30	62	94	^	126	~	158	190	222	Þ	254	þ
31	63	95	_	127		159	191	223	ß	255	ÿ

Source: Progiiciels-bpi <<http://www.progiiciels-bpi.ca/tcao/apercu.html>>.

23.3. Le codage sur plusieurs octets

Pour dépasser les limitations du codage sur un octet, le consortium Unicode et l'Organisation internationale de normalisation (ISO) ont mis au point la norme ISO/IEC-10646 qui code sur deux octets un jeu de 65.536 caractères, appelé « Jeu Universel de Caractères⁶ » (JUC-Unicode ; en

⁶ Voir la page du site du consortium Unicode <<http://www.unicode.org/fr/charts/>>

anglais : *Universal Characters Set* (UCS). De quoi loger tous les caractères de tous les systèmes d'écriture du monde! C'est parfait pour les idéogrammes chinois, mais pour les caractères latins qui se contentaient d'un codage sur un octet, le codage sur deux octets est un habit un peu trop ample. Aussi, pour assurer la compatibilité avec les documents existants, plusieurs systèmes de codage ont été proposés pour ces caractères codés sur deux octets. Le plus populaire d'entre eux est le codage UTF-8.

Mais il est rare que l'on ait besoin d'assurer la conversion soi-même, car les logiciels d'application s'en chargent. Ce qu'il est important de bien comprendre est que dans Unicode, chaque caractère ne peut être défini qu'une fois, en ce sens qu'un même code numérique ne peut renvoyer qu'à une seule définition de caractère et vice versa.

2.4. Premières ressources de base

24.1. Jeu et police de caractères

La première ressource informatique à élaborer pour une langue qui n'en a pas est de définir un jeu de caractères à partir de la liste des caractères que l'on a retenus pour l'orthographe de la langue. Si le jeu de caractères dont on a besoin ne contient aucun caractère spécial, on peut utiliser immédiatement tous les logiciels courants. Par contre, si l'on a besoin de caractères spéciaux, alors, avant de se lancer dans la création d'une police particulière, il faut d'abord vérifier si on ne trouve pas son bonheur avec les polices spéciales disponibles sur Internet, tels que *Lucida Sans Unicode*, *Gentium*, *Doulos SIL*⁷, et les polices africaines des Progiiciels-bpi⁸. Si on ne trouve toujours pas l'ensemble des caractères dont on a besoin dans aucune de ces polices existantes, alors on doit vraiment définir son propre jeu de caractères. On serait alors bien inspiré de vérifier que ce jeu puisse valoir aussi pour plusieurs langues de la région afin de mieux rentabiliser les efforts tout en garantissant la possibilité de produire plus tard des textes multilingues sans avoir à changer de polices d'une langue à l'autre. Deux bons exemples de définition de jeux de caractères 8 bits (mais compatibles Unicode) sont fournis par les Progiiciels-bpi. Chacun de ces deux jeux (*afrful* et *afrlin*) couvre plusieurs langues⁹.

Une fois défini le jeu de caractères, il faut créer la police correspondante. De nos jours, les experts déconseillent vivement la création de nouvelles polices 8-bits en raison de leurs limitations évoquées plus haut. Il vaut mieux s'adresser à un spécialiste pour se faire créer une police contenant les caractères voulus et respectant la norme Unicode. Ce recours est encore plus indispensable si on utilise un système d'écriture non latin et non encore représenté dans Unicode! Le site <<http://www.freelang.com>> propose un service gratuit de création de polices à la demande de ses visiteurs avec leurs données. La police créée est en partage libre et gratuit sur le site. On trouve sur l'Internet plusieurs outils logiciels pour créer des polices locales, notamment aux adresses suivantes:

1) Des polices prêtes à l'emploi et des outils pour en créer:

<http://scripts.sil.org/TypeDesignResources>

2) Des polices prêtes à l'emploi à télécharger!

<http://scripts.sil.org/SILFontList>

3) Pour visualiser les glyphes des polices

http://scripts.sil.org/SILEncore_Glyphs

7 . Voir le site de la SIL :< <http://scripts.sil.org/SILFontList>>

8 Voir Progiiciels-bpi <<http://www.progiiciels-bpi.ca/tcao/apercu.html>>

9 On peut voir aussi le site burkinabé <<http://www.abcburkina.net/sedelan/index.htm>> pour les langues du Burkina Faso.

24.2. Claviers virtuels

Une fois que l'on dispose d'une police de caractères correspondant au jeu de caractères que l'on s'est choisi, on peut l'installer sur son ordinateur en suivant la procédure d'installation des polices prévue dans le système d'exploitation que l'on utilise. Et lorsqu'on a besoin d'utiliser ces polices dans un texte, la plupart des traitements de texte modernes ont une fonction « Insertion/Caractères spéciaux » qui vous permet d'ouvrir une fenêtre de glyphes et de sélectionner le caractère désiré pour l'insérer dans le texte au point de saisie. Cette façon de procéder est fastidieuse et ralentit considérablement la saisie d'un texte dans la langue peu dotée. Il est donc indispensable d'élaborer des combinaisons de touches plus pratiques permettant d'accéder aisément aux caractères spéciaux d'une police particulière.

Il faut remarquer que toutes les langues dominantes des pays industrialisés ont un clavier créé exprès pour chacune d'elles, voire pour certaines variétés régionales d'une même langue. Par exemple, le clavier du français canadien n'est pas le même que celui du français de France. Non seulement tous ces claviers se distinguent par la position des lettres A Z et Q W, mais encore, ils présentent une très grande liberté dans la disposition des caractères de ponctuation et naturellement des caractères spéciaux (diacritiques, caractères accentués, symboles monétaires). En Afrique, aucune des grandes langues africaines (*kiswahili*, *zulu*, *hausa*, *fulfulde*, etc.) ne dispose de claviers qui lui soient propres. Alors, lorsqu'on est une petite langue peu dotée, il ne faut pas espérer disposer de si tôt d'un clavier physique. La solution la plus simple est donc d'élaborer un clavier virtuel.

En se basant sur le clavier physique utilisé dans la région où se parle la langue peu dotée, on commence par créer une table de combinaisons de touches faciles à se rappeler. Par exemple, on décide qu'il faut frapper la suite des touches « < »+« O » (sauf les guillemets et le +) pour obtenir un « o ouvert minuscule », soit « ◊ », et pour obtenir la majuscule correspondante « ◊ » on doit frapper les touches « Maj »+« > »+« O ». Soit la table suivante:

« < »+« O » = « ◊ »

« Maj »+« > »+« O » = « ◊ »

Ces combinaisons ne sont ergonomiques que si les crochets brisés « < » et « > » sont superposés sur une même touche avec le second placé au dessus du premier. Ces crochets sont alors utilisés comme des touches mortes. Aussi, pour pouvoir les utiliser à nouveau en tant que crochets brisés, il suffit de taper la touche « barre d'espacement » après les avoir tapés. Il faut donc ajouter les lignes suivantes à la table :

« < »+« espace » = « < »

« Maj »+« > »+« espace » = « > »

Les caractères « < » et « > » peuvent alors être utilisés comme modificateurs généraux pour obtenir d'autres caractères spéciaux en minuscule et en majuscule. La chose se complique lorsque l'on doit mettre un ou deux diacritiques sur un caractère déjà modifié, par exemple un accent aigu pour indiquer un ton haut et un tilde pour indiquer une nasalité. Il faudrait alors bien vérifier que le logiciel à utiliser pour interpréter la table supporte l'enchaînement des commandes du genre: « accent aigu » + « tilde » + « modifieur » + « caractère de base » = « caractère modifié accentué nasal ». L'un des meilleurs logiciels actuels pour la création de claviers virtuels est le Keyman^(TM) de la société Travult. On trouvera plusieurs claviers virtuels réalisés avec ce logiciel pour les langues africaines à l'adresse suivante:

<http://scripts.sil.org/SILKeyboards>.

Il suffit d'installer Keyman sur son ordinateur pour qu'il installe les claviers et les polices

correspondantes. On peut installer plusieurs claviers et passer de l'un à l'autre par une simple combinaison de touches tel que « Alt+K ». Il faut simplement s'assurer que la combinaison de touches choisie ne soit pas déjà utilisée pour autre chose par un des logiciels installés sur l'ordinateur.

24.3. Logiciels de traitement de corpus assisté par ordinateur

Dès lors que l'on utilise des caractères spéciaux, il faut penser aux outils de base qui permettront d'effectuer des tris alphabétiques, des conversions de codes des caractères, d'alignement de textes, etc, qui seront tous très utiles pour la manipulation et le traitement des textes à produire dans la langue peu dotée en vue d'un affichage sur la Toile. L'ensemble d'outils le plus abouti que l'on peut directement utiliser pour les langues africaines (et d'autres encore) provient des Progiciels-bpi que nous citons ci-après :

« Ce document présente une suite de cinq progiciels de traitement de corpus assisté par ordinateur (TCAO) qui sont désignés à l'aide des noms suivants: **Alibi**, **Concorde**, **Recode**, **Ventile**, **Vocable**. Plus bas, le **Tableau 1** fournit une brève description de ces cinq progiciels de TCAO. »¹⁰

Progiciels

TABLEAU 1 - Description des progiciels de TCAO

Recode **Recode** est un progiciel de conversion automatique de jeux de caractères. Il permet de traiter et de recoder environ 175 jeux de caractères différents et une douzaine de **surfaces** de fichiers. Selon la paire de jeux de caractères indiquée dans la commande d'appel (d'un jeu de départ vers un jeu d'arrivée), Recode effectue un recodage du ou des fichiers d'entrée. Puisque chaque jeu de caractères peut être converti vers la plupart des 174 autres, plusieurs milliers de conversions différentes sont alors possibles.

Ventile **Ventile** est un progiciel de production de statistiques textuelles. Comme élément textuel, Ventile permet de compter le nombre de paragraphes, de phrases, de mots et de caractères par fichier. Comme statistiques textuelles, Ventile produit les mesures de fréquence absolue, trois mesures de tendance centrale (le mode, la médiane et la moyenne arithmétique) et cinq mesures de dispersion (le minimum, le maximum, l'écart quartile, l'écart moyen et l'écart type). Les résultats statistiques sont affichés numériquement sous la forme d'un tableau et graphiquement sous la forme d'un histogramme.

Vocable **Vocable** est un progiciel de production de listes de vocabulaire. Cet outil permet de dépouiller des textes et de fabriquer la liste des vocables actualisés dans ces textes. Ces listes peuvent être triées en ordre alphabétique habituel (de gauche à droite), en ordre alphabétique inverse (de droite à gauche) ou en ordre de fréquence décroissante. Chaque vocable peut être accompagné de la fréquence de ses occurrences et de la liste de toutes ses références textuelles pour chaque occurrence.

Concorde **Concorde** est un progiciel de production de concordances de mots en contexte. Cet outil permet de dépouiller des fichiers textuels et de produire une liste alphabétique des éléments-vedettes avec leur contexte immédiat. L'élément-vedette d'une concordance peut être basé sur les mots dans leurs contextes phrastiques ou sur les

10 D'après la source: <http://www.progiciels-bpi.ca/tcao/apercu.html>

caractères dans leurs contextes de mots. Les mots-vedettes peuvent être triés en ordre alphabétique habituel ou en ordre alphabétique inverse.

Alibi est un progiciel d'alignement bi-textuel (ou bilingue). Cet outil permet de dépouiller en parallèle deux fichiers et d'aligner automatiquement les constituants textuels provenant de cette paire de fichiers. Comme constituants textuels, Alibi peut aligner les paragraphes, les phrases ou les mots de ces deux textes.

Faut-il souligner que la réalisation de tels logiciels n'est pas à la portée du premier venu et qu'il faut vraiment utiliser les outils existants d'autant plus qu'ils sont performants et fonctionnent aussi bien sous les systèmes d'opération les plus populaires (Windows et Linux). Associés aux jeux de caractères, aux polices et aux claviers virtuels, ces logiciels constituent les premières ressources de base pour le traitement d'une langue peu dotée.

2.5. La localisation des logiciels

On entend par localisation de logiciels l'activité qui consiste à prendre un logiciel conçu dans une langue et un pays étranger et à l'adapter à la langue et à la culture d'un autre pays. Tout logiciel est composé de deux parties : le code informatique, et l'interface homme-machine. Il n'est jamais question de toucher au code sauf dans des cas exceptionnels comme la nécessité de modifier le sens de l'écriture ou l'ordre de tris alphabétiques quand ceux-ci n'étaient pas prévus d'origine pour être modifiés. Les logiciels sont de plus en plus prévus pour être localisés et par conséquent, la partie interface homme-machine est aisément identifiable et accessible soit grâce à un fichier prévu pour cela, soit en utilisant des outils appropriés.

Sont concernés par la localisation d'un logiciel, tous les menus (qu'ils soient déroulants, fixes, flottants ou contextuels) toutes les boîtes de dialogues, les notifications, les informations d'aide en ligne, les boutons et commandes de navigation, l'imagerie (logos, drapeaux, armoiries, illustrations, revêtements (*skin* en anglais) et design), la représentation aux normes locales des unités de mesure (distance, poids, volume, temps, monnaie, densité, etc.); les tutoriels et la documentation (manuels d'installation, d'utilisation, et de référence). L'objectif de la localisation est bien évidemment de permettre à l'utilisateur de travailler dans un environnement informatique qui lui est culturellement et linguistiquement familier et donc facile à maîtriser. En conséquence, sa langue devient une langue de travail des technologies de l'information et de la communication. Ainsi, la localisation des logiciels dans une langue peu dotée contribue à valoriser cette langue aux yeux des utilisateurs, notamment de ses propres locuteurs.

Lorsqu'on veut localiser des logiciels pour une langue peu dotée, à moins d'en être le propriétaire, il vaut mieux le faire avec des logiciels libres de tous droits, comme ceux distribués sous licence GPL (*General Public Licence*)¹¹. La distribution africaine de Linux nommée Ubuntu¹² offre au monde entier un environnement en ligne appelée *Rosetta*¹³ où on peut se connecter, choisir un logiciel libre et se mettre à le traduire dans sa langue. Un ensemble d'outils de gestion de projet est fourni sous le nom de *Launchpad*¹⁴ et constitue une aide précieuse pour l'utilisateur qui se lance dans la localisation d'un des logiciels. Un tel dispositif offert gratuitement à toutes les langues du monde est vraiment une initiative extrêmement rare et généreuse, une vraie aubaine pour les langues peu

11 Pour plus d'information sur les licences des logiciels libres, voir: <<http://www.gnu.org/licenses/license-list.fr.html>>

12 Voir : <<http://www.ubuntu.com>>

13 Voir: <<https://launchpad.net/rosetta>>

14 Voir: <<http://launchpad.net>>

dotées qui pourront en profiter pour accéder plus facilement au statut de langue de travail du cyberspace. Ceci dit, il faut souligner que la localisation des logiciels n'est pas une étape indispensable pour accéder au cyberspace. On peut déjà y faire accéder une langue peu dotée avant même de penser à mettre en route un projet de localisation. Dès que l'on dispose des premières ressources de base (cf. section 2.4. ci-dessus) on peut déjà commencer à élaborer des contenus culturels et linguistiques pour le cyberspace.

3. Elaboration des contenus culturels

3.1. Les contenus culturels

Les contenus culturels sont textuels, sonores et iconographiques. Tous peuvent être numérisés et portés sur la Toile ou transportés à travers le cyberspace pour être partagés.

31.1. Le texte

Outre le corpus utilisé pour décrire la langue et créer ses premiers ouvrages de référence, il convient de collecter ou de créer encore plus de textes, de tout genre, depuis les petites annonces jusqu'aux récits d'épopées légendaires, les traductions d'oeuvres littéraires ou encore la tenue de journaux. Tous ces textes vont faire l'objet de traitements informatiques divers grâce aux logiciels de traitement de corpus assisté par ordinateur (TCAO) et les résultats de ces traitements peuvent être diffusés sur la Toile autant qu'ils peuvent servir à améliorer l'étude de la langue elle-même.

31.2. Le son

Les documents sonores peuvent être constitués d'enregistrements liste de mots ou de phrases vocalisées, de chansons, de récitals, de légendes et de contes racontés, d'interviews, de reportages, mais aussi de pure musique instrumentale, de sonnailles et de bruits de la vie et de la nature témoignant de l'environnement culturel local. La numérisation de tous ces documents sonores nécessite du matériel d'enregistrement de qualité professionnelle qui peut coûter une petite fortune pour qui vient du Tiers-Monde et parle une langue peu dotée. La création musicale en particulier peut demander des appareils coûteux. Il en est de même pour les reportages naturalistes pour capter les chants des oiseaux et les cris des insectes nocturnes dans le souci d'étudier l'environnement local et l'écologie naturel. Il suffit que ces documents sur l'environnement soient accompagnés de commentaires dans la langue locale pour que celle-ci s'en trouve valorisée.

Tous les fichiers son ainsi obtenus doivent être filtrés pour être débarrassés des souffles et autres bruits parasites. Il y a ensuite un long travail à faire en studio pour découper et étalonner chaque section de l'enregistrement sonore afin de constituer des fichiers plus légers, plus facilement transportables. Aujourd'hui, il est possible de diffuser des documents sonores sur Internet et dans le cyberspace comme on le fait pour une station radio ou pour les films. Mais sans l'image ces documents sonores, si intéressants soient-ils, ressemblent à la télévision noir et blanc à l'heure de la couleur.

31.3. L'image

Les documents iconographiques regroupent les dessins, les logos, les photos, et la vidéo. Si les

logiciels de dessins sont extrêmement courants comme les appareils de photographie, les caméras de qualité sont déjà un peu chers. Toutefois, on ne peut pas s'en passer pour filmer une danse, une cérémonie, une scène de vie, une vue panoramique, un reportage, etc. Toutes ces images peuvent servir à un montage et subir des retouches pour en améliorer la qualité. Les fichiers image sont aujourd'hui directement numérisés dès leur création. Il faut néanmoins s'assurer de la très bonne qualité de l'image, car les traitements nécessitent plusieurs copies et cela peut ternir quelque peu la qualité de l'image diffusable.

Chaque fois que cela est possible, il est préférable d'associer le son, l'image et le texte dans la production de documents destinés à la diffusion sur l'Internet. Mais si une langue peu dotée n'était pas encore écrite, il est toujours possible de créer des documents sonores et iconographiques, à l'instar d'un film et/ou de photos commentées, et de les placer sur la Toile.

3.2. Préparer les voies d'accès au cyberspace.

32.1. Le site sur Toile

Le site sur Toile est la meilleure façon d'affirmer la présence d'une langue dans le cyberspace, car on peut le décliner à l'infini, depuis la page personnelle jusqu'au portail multimédia offrant de très nombreux services divers et variés. Avant de se lancer dans la création d'un site, il convient de prendre en compte les observations suivantes:

Un site écrit dans une langue peu dotée ne peut être lu que par les locuteurs de cette langue. Cela est difficile si ces locuteurs ne savent pas lire et écrire dans leur langue. Le site peut servir alors à les y encourager. Quand on vient de mettre au point une orthographe pour la langue peu dotée, il est préférable d'envisager de créer un site bilingue utilisant simultanément la langue peu dotée et une langue de plus grande diffusion parlée dans la région¹⁵. Dans ce cas, il faut absolument veiller à ce que la langue peu dotée soit présente dans les boutons et commandes de navigation et tous les messages destinés au visiteur du site afin d'éviter qu'elle ne se retrouve en position d'objet de curiosité et perde son statut de langue de travail sur la Toile.

Les boutons de navigation, les menus de commandes, les titres, les logos, les textes décoratifs, les animations et les styles de présentation du site, sont tous des éléments à préparer à l'avance, avec des logiciels spécialisés comme *Button Studio*TM. Les fichiers texte destinés à s'afficher sur la Toile sont écrits dans le langage HTML (*HyperText Markup Language*). On peut écrire une page sur Toile avec un simple traitement de texte comme *Open Office .org* qui est capable d'enregistrer directement au format HTML. Et il existe de très nombreux outils mieux adaptés à la création des sites, des plus simples et souvent gratuits comme *Netscape*TM *Composer* aux plus sophistiqués et professionnels comme *Dreamweaver*TM. Pour débiter, on trouvera de quoi s'informer et se former à l'adresse suivante:

<http://www.unice.fr/urfist/ResInternet.html>

La composition de la page du site est une affaire de goût et de culture.

32.3. Le courrier électronique et la messagerie instantanée

C'est l'outil de communication le plus utilisé de l'Internet. Mais, pour des raisons de sécurité, il est très difficile de mesurer le taux d'utilisation de telle ou telle langue dans les échanges mondiaux par courrier électronique. L'utilisation d'une langue peu dotée dans ce service passera donc inaperçue. Par contre, la localisation d'au moins un logiciel de messagerie et de messagerie instantanée

¹⁵ Exemple, le site de l'Association YSB SAHNGO : <http://sango.free.fr>

constitue une vraie contribution à l'équipement de la langue peu dotée.

32.4. La téléphonie IP et la téléphonie mobile

La téléphonie IP (*Internet Protocol*) permet de téléphoner en échangeant des fichiers son en temps réel via l'Internet. Cette technologie peut donc être utilisée librement par les locuteurs d'une langue peu dotée. Il leur suffit de télécharger un logiciel adéquat comme *Skype*TM, le plus populaire du moment, et l'installer chacun sur son ordinateur. Seule la localisation de l'interface utilisateur du logiciel peut constituer une ressource propre pour la langue peu dotée. Il en est de même pour la téléphonie mobile dont la seule façon de contribuer à l'équipement de la langue peu dotée est de localiser l'interface utilisateur des téléphones, comme cela se fait en Afrique du Sud.

32.5. Les fora et listes de diffusion

L'installation d'un forum ou d'une liste de diffusion dans un site permet à la communauté des locuteurs de la langue peu dotée de dialoguer entre eux par écrit. Sous réserve que l'environnement logiciel supporte les caractères spéciaux requis, cette technologie peut s'avérer un excellent moyen de faire vivre la langue peu dotée dans le cyberspace. En effet, sans une communauté qui anime et active les échanges dans la langue, celle-ci court le risque de n'être qu'un élément de décor inerte dans le cyberspace. Or ce n'est certainement pas là le but de tout l'effort que l'on peut faire pour amener une langue peu doter dans le cyberspace. De nombreux portails commerciaux et non commerciaux proposent gratuitement la possibilité de créer des fora communautaires. En voici quelques-unes des plus populaires :

<http://www.yahogroups.com>

<http://www.google.com>

<http://www.ubuntu.com>

<http://www.free.fr>

<http://lists.kabissa.org/mailman/options/a12n-entraide/mdkidiri%40free.fr>

4. Développer la communauté des utilisateurs

4.1. La création d'une communauté d'utilisateurs

L'existence d'une communauté d'utilisateurs instruite dans sa langue est absolument nécessaire pour faire vivre cette langue dans le cyberspace quand elle y accèdera. En effet, si on oeuvre pour amener des langues peu dotées à occuper une place dans le cyberspace, ce n'est pas du tout pour en faire des enjoliveurs, du mobilier ou encore des figurantes, mais bien pour qu'elles y vivent et y soient productives pour leurs locuteurs, lesquels sont les premiers concernés par la maîtrise et la promotion de leurs langues dans le cyberspace. Il est donc important d'enseigner à lire et à écrire dans la langue peu dotée à ses propres locuteurs, et surtout aux jeunes. L'initiation à l'utilisation de l'ordinateur et aux logiciels de bureautique et de communication devrait s'accompagner de l'enseignement de la langue peu dotée, par exemple, dans le cadre d'activités associatives, si ce n'est à l'école. Les associations locales sont généralement plus à même de mettre en place des formations de ce genre que les écoles qui dépendent lourdement du système éducatif national. Mais le plus souvent les associations locales sont peu argentées et pas du tout aidées, surtout pour des petites langues peu dotées. Il est clair que tout soutien apporté à des associations locales engagées dans le développement des langues en relation avec les TICs contribue à la promotion de la diversité linguistique et culturelle dans le cyberspace.

4.1.1. Régénération d'une communauté

Lorsqu'une langue peu dotée n'a plus que quelques locuteurs, il est absolument incontournable de consolider le nombre des locuteurs en enseignant la langue à des jeunes par tous les moyens disponibles : activités culturelles, centres de lecture et de culture, activités associatives, et si possible à travers l'école. Il s'agit en fait de recréer un environnement vital pour la langue. Plus la langue sera enseignée et valorisée, plus elle pourra servir à produire du contenu culturel que l'on pourra mettre sur un site. Il faut bien être conscient que cela demande un travail colossal qui nécessite un investissement important tant en ressources financières qu'en ressources humaines et une détermination sans faille sur le long terme. Les associations culturelles peuvent être très actives localement pour réhabiliter une langue en danger et maintenir un intérêt croissant en sa faveur, notamment chez les décideurs administratifs et politiques.

4.1.2. Consolidation d'une communauté

Dans le cas d'une langue peu dotée mais qui dispose d'une population de locuteurs non négligeable, l'éducation des gens dans leur langue est une entreprise de longue haleine qui exige également de mobiliser beaucoup de ressources humaines et financières. Dans ce cas, l'accès de la langue au cyberspace est l'un des moyens de mobiliser et dynamiser la communauté des utilisateurs de la

langue. La consolidation de cette communauté ne peut donc pas être un préalable à l'accès de la langue au cyberspace, mais une activité permanente qui concourt à faire vivre la langue dans cet espace. Ici bien plus encore que dans le cas des langues en danger de disparition qui ne jouissent pas d'un nombre aussi considérable de locuteurs, il est nécessaire d'impliquer les décideurs politiques et administratifs dans le processus de valorisation de la langue peu dotée. Sans cela, les efforts des associations ne suffiront pas à garantir durablement son essor, faute de moyens conséquents sur le long terme.

4.2. Les aspects politiques et socio-culturels

42.1. Des arguments pour convaincre

Comme il est rappelé dans les paragraphes ci-dessus, l'action d'une association pour développer une communauté d'utilisateurs d'une langue peu dotée a une limite qui est liée aux moyens financiers que peut mobiliser cette association. Il est donc très important de sensibiliser les responsables politiques à tous les niveaux sur tous les avantages qu'il y a à soutenir l'instrumentalisation des langues peu dotées et le renforcement des capacités de leurs locuteurs à s'en servir. Avec le soutien des administrations locales, il est plus facile d'envisager l'enseignement des langues à l'école, notamment dans le primaire où la langue locale est la mieux placée pour servir de langue des premières acquisitions des savoirs chez l'enfant. Un enseignement bilingue ne peut réussir même localement qu'en étant inscrit dans un cadre juridique de l'Education nationale. On sait par plusieurs expériences d'enseignement bilingue impliquant des langues africaines au Burkina, au Sénégal; au Mali, au Burundi et au Rwanda, que les enfants apprennent bien mieux dans leur langues maternelles et obtiennent de bien meilleurs résultats scolaires. A terme, ils deviennent des citoyens plus instruits. Or un homme instruit est bien plus capable d'avoir une vision globale et une compréhension approfondie des problèmes complexes du monde moderne et de contribuer à y apporter des solutions durables dans sa sphère de vie et d'activités. C'est une personne qui pourra mener à bien un projet socio-économique autant pour lui-même que pour sa famille, son village et sa région. L'éducation multilingue impliquant la langue maternelle est la clé de l'essor économique.

Outre l'avantage éducatif, une politique favorisant la diversité linguistique peut générer des emplois, notamment, des traducteurs, des rédacteurs de manuels, des éditeurs, des animateurs, etc. La diversité linguistique est donc bien un enjeu économique et le développement des langues peu dotée y participe. Lorsque la langue peu dotée concerne toute une région dans un pays, c'est toute la politique linguistique au niveau national qui est interpellée par son développement. Celui-ci devra alors s'inscrire dans un plan national d'aménagement politique mettant en oeuvre une politique linguistique nationale clairement affirmée en faveur de la diversité linguistique. En outre, une telle politique respecte les droits linguistiques qui font partie des droits fondamentaux de l'homme.

42.2. Des instruments de référence

Au-delà des autorités politiques nationales, il est recommandé de rechercher aussi des soutiens auprès d'organisations internationales qui oeuvrent pour la promotion et la valorisation de la diversité linguistique dans le monde, ce qui implique au moins une attitude bienveillante vis-à-vis des efforts que peuvent fournir des associations locales pour développer des langues peu dotées. Il faut savoir qu'au cours des quarante dernières années, les organisations internationales ont consacré de nombreuses rencontres aux langues des populations locales, souvent qualifiées de « langues nationales » par opposition aux langues européennes héritées de la colonisation. Ces rencontres ont permis l'élaboration et l'adoption de plusieurs textes, appelés « instruments internationaux » sur lesquels on peut s'appuyer pour mener des actions concrètes. Parmi celles organisées à

l'initiative des Etats et des Organisations internationales, Amadou Touré¹⁶ note pour l'Afrique, entre autres :

1. La réunion de l'UNESCO sur l'emploi des langues vernaculaires dans l'enseignement (Monographies sur l'éducation de base) Paris, 1953 ;
2. La conférence régionale de l'UNESCO sur la planification et l'organisation des programmes d'alphabetisation en Afrique – Abidjan, Mars 1964 ;
3. Le Congrès de la Société Africaine de linguistique – SLAO- Accra, 1965 ;
4. La réunion d'experts organisée par l'UNESCO pour l'unification des alphabets des langues nationales : fulfulde, hawsa, kanuri, mandingue, songhay-zarma, tamasheq. Bamako, 28 Février - 05 Mars 1966 ;
5. La Conférence intergouvernementale sur les politiques culturelles en Afrique (UNESCO-OUA), Accra – 1975 ;
6. La réunion organisée par l'Agence de Coopération Culturelle et Technique sur la promotion des langues nationales – Yaoundé, 1977 ;
7. La conférence internationale de l'ACCT sur « la recherche linguistique, l'emploi et l'enseignement des langues en Afrique : les moyens de renforcer la coopération entre les Etats »-Yaoundé, 1983 ;
8. La réunion de l'UNESCO sur les stratégies de promotion des langues Africaines – Conakry, 1984 ;
9. La réunion d'experts de l'UNESCO sur la promotion des langues africaines comme instruments de culture et d'éducation permanente – Yaoundé ;
10. La réunion d'Accra sur le Projet de Charte d'Action pour la promotion et l'utilisation des langues africaines dans l'éducation - Août 1996 ;
11. La Conférence intergouvernementale sur les politiques linguistiques en Afrique
12. (UNESCO-OUA-ACCT), Harare – 1997 ;
13. La Consultation africaine sur la création de l'Académie Africaine des Langues – Bamako – 25-27 Mai 2001 ;
14. La conférence de Bamako sur le multilinguisme dans le cyberspace - Mai 2005 ;
15. La 33^{ème} session de la Conférence Générale de l'UNESCO tenue à Paris en Octobre 2005 qui a adopté le projet de résolution initié par l'ACALAN présenté par le Bénin et relative à l'organisation de cinq colloques régionaux sur les langues transfrontalières et les langues de moindre diffusion en Afrique.

Plusieurs grandes rencontres ont permis l'élaboration d'instruments normatifs et de textes règlementaires tels que :

1. La Charte de l'OUA – 1963 ;
2. Le Manifeste culturel panafricain d'Alger -1969 ;
3. La Charte culturelle de l'Afrique – OUA, Port Louis, 1976 ;
4. Les résultats de la 1^{ère} conférence des Ministres africains de la culture de l'OUA, Port-Louis, 1986 ;
5. Le Plan d'Action de Lagos de l'OUA pour le développement économique de l'Afrique - 1980 ;
6. Le Plan d'Action linguistique pour l'Afrique - Addis Abeba, 21-25 Juillet 1986 ;
7. Le Plan Décennal de l'OUA sur les langues et les traditions orales - 1987 ;
8. Le traité d'Abuja créant la Communauté économique Africaine - 1991 ;
9. Le Plan régional de collecte des traditions orales en Afrique Australe - Harare, 1993 ;
10. Le Plan d'aménagement linguistique de l'espace francophone - 1990-2000 (ACCT 1993) ;

16 Document de travail pour les colloques de l'Académie africaine des langues, 2006.

11. Le Programme d'Action de la Décennie de l'Education en Afrique - Harare, 1999 ;
12. Le Traité créant l'Union Africaine, Lomé- 2000 ;
13. La Résolution de la 31^{ème} session de la Conférence Générale de l'UNESCO sur l'Académie Africaine des Langues – 2001 ;
14. La décision CM/Déc.613(LXXIV) du 37^{ème} Sommet des Chefs d'Etats et de Gouvernement de l'OUA créant l'Académie Africaine des Langues - Juillet 2001;
15. La Déclaration universelle sur la diversité culturelle adoptée par la 31^{ème} session de la Conférence Générale de l'UNESCO - 2001 ;
16. Les recommandations sur la promotion et l'usage du multilinguisme et l'accès universel au cyberspace adoptées par la 32^{ème} session de la Conférence Générale de l'UNESCO - Octobre 2003 ;
17. Les textes fondateurs du Sommet Mondial sur la Société de l'Information (Genève 2003 et Tunis 2005) ;
18. La décision AU/Dec.92(VI) de la sixième Session Ordinaire de la Conférence de l'Union Africaine relative à une seconde décennie de l'éducation pour l'Afrique (2006-2015) et au cadre d'action de la seconde décennie (Doc.EX.CL/224-VIII Rev.2) – Khartoum 2006 ;
19. La décision AU/Dec.94(VI) de la sixième Session Ordinaire de la Conférence de l'Union Africaine sur la Charte révisée de la renaissance culturelle africaine – Khartoum 2006 ;
20. La décision AU/DEC.95(VI) de la sixième Session Ordinaire de la Conférence de l'Union Africaine sur les statuts de l'Académie Africaine des Langues – Khartoum 2006 ;
21. La décision AU/Dec.96(VI) de la sixième Session Ordinaire de la Conférence de l'Union Africaine sur le lien entre la culture et l'éducation – Khartoum 2006 ;
22. La décision AU/Dec.98(VI) de la sixième Session Ordinaire de la Conférence de l'Union Africaine relative à la proclamation de 2006 comme l'année des langues africaines – Khartoum 2006 ;

Des institutions ont été créées, entre autres :

- l'ICA (Institut Culturel Africain) Dakar - Sénégal ;
- le CELHTO (Centre d'Etudes Linguistiques et Historiques par Tradition Orale) Niamey - Niger ;
- le CERDOTOLA (Centre Régional de Documentation sur les Traditions Orales et les Langues Africaines) Yaoundé - Cameroun ;
- l'EACROTANAL (East African Center of Research on Oral Traditional and National Languages) Zanzibar - Tanzanie ;
- le CICIBA (Centre International des Civilisations Bantou) Libreville - Gabon ;
- le BASE (Bureau Africain des Sciences de l'Education) devenu aujourd'hui IPED (Institut Panafricain de l'Education pour le Développement). Kinshasa – République Démocratique du Congo.

La plupart de ces institutions, il faut le reconnaître, sont restées en deçà des attentes, fautes de moyens financiers pour fonctionner pleinement comme il se devrait. Il n'est donc pas évident d'obtenir leur soutien pour mener une action locale de valorisation d'une langue peu dotée. Cependant, avec un peu de persévérance et beaucoup de bonne volonté, on peut arriver à s'en faire des alliés dans sa lutte, et ce n'est pas rien!. En effet, un inventaire des instruments normatifs existants et un examen des rapports des différentes rencontres organisées au cours de cette longue période, montrent que tous les pays de Afrique ont, à un moment ou un autre et à des degrés divers, pris des initiatives créant des conditions favorables à la promotion de leurs langues nationales. C'est pourquoi, il convient de tout mettre en oeuvre pour les amener s'impliquer encore davantage dans des actions concrètes en faveur de la diversité linguistique. Seule une volonté politique forte et permanente conduira à débloquent les fonds nécessaires pour organiser l'éducation multilingue, le

développement des langues et de leur usage dans toute la vie de la nation afin que soient générés les biens culturels, économiques et sociaux pour un meilleur être de tous.

5. Conclusion

Dans le souci de promouvoir et de valoriser la diversité linguistique et culturelle dans le cyberspace, il convient d'aider les langues les plus défavorisées à accéder à cet espace. Si on peut y arriver avec une petite langue orale non écrite et menacée d'extinction, on peut, à plus forte raison, le faire avec toutes les langues peu dotées en meilleure posture.

La première étape consiste à entreprendre les études linguistiques qu'il faut pour élaborer les ressources linguistiques indispensables, à savoir : une liste des phonèmes, un alphabet et une orthographe, une grammaire, un dictionnaire et un recueil de textes.

La seconde étape est consacrée aux travaux de traitement informatique de la langue afin d'identifier ou d'élaborer les ressources informatiques compatibles, à savoir : un jeu de caractères, au moins une police de caractères, un clavier virtuel, des logiciels de traitement de corpus, lesquels peuvent être utilisés aussi pour affiner l'analyse linguistique de la langue et améliorer les ressources linguistiques.

La troisième étape consiste à élaborer les ressources culturelles et à les conditionner pour les rendre partageables dans le cyberspace. Il s'agit d'enregistrer et de numériser beaucoup de données textuelles, sonores et iconographiques et de les préparer pour les afficher dans un site. Il s'agit aussi de créer les éléments constitutifs d'un site tels que les menus, les barres de navigation, les titres et autres textes de communication homme-machine. Dans certains cas, la localisation de logiciels est nécessaire pour contribuer à valoriser la langue en tant que langue de travail et à la doter de ressources informatiques complémentaires.

Enfin, il convient d'apprendre à créer un site dans la langue peu dotée, éventuellement en complémentarité avec une autre langue de plus grande diffusion. On trouve sur Internet tous les outils nécessaires pour assurer un tel apprentissage, ainsi que pour créer un forum, et localiser des logiciels libres. Avec un site, un forum, une liste de diffusion, une téléphonie IP, de la musique, des images fixes et des vidéos, la petite langue est désormais installée dans le cyberspace. Mais pour qu'elle y vive, il faut développer une communauté d'utilisateurs capable de s'en servir intensément. Aider les associations locales à développer de telles communautés, c'est contribuer à promouvoir et à valoriser la diversité des langues et des cultures dans le cyberspace

Références

Consortium Unocide : <http://www.unicode.org/fr/charts/>

Création de polices en ligne : <http://www.freelang.com>

Environnement de traduction et de localisation de logiciels libres: <https://launchpad.net/rosetta>

Liste de polices de caractères : <http://scripts.sil.org/SILFontList>

Progiciels-bpi <http://www.progiciels-bpi.ca/tcao/apercu.html>

SEDELAN (Service d'édition en langues nationales de Koudougou, Burkina Faso) :
<http://www.abcburkina.net/sedelan/index.htm>

Site bilingue sahngo-français de l'association YSB SAHNGO pour la promotion de la langue sahngo : <http://sango.free.fr>

Université de Nice: Ressources multimédia: formation à la création de site Internet:
<http://www.unice.fr/urfist/ResInternet.html>

Annexe 1 : Jeux de caractères AFRFUL¹⁷

Les jeux de caractères africains **afrful** permettent de couvrir les alphabets du bambara, de l'éwondo, du français et du fulfulde. Le jeu de caractères *AFRFUL-102-BPI_OCIL* représente un jeu de caractères, en version encodée, qui est utilisé pour l'affichage à l'écran des langues africaines suivantes: le bambara, l'éwondo, le fulfulde (ou peuhl). Les noms alias acceptables pour désigner ce jeu de caractères encodés sont *afrful102bpiocil*, *bambara*, *ewondo*, *fulfulde*, ou *bra*.

Le jeu de caractères *AFRFUL-103-BPI_OCIL* représente un jeu de caractères, en version translittérée, qui est utilisé pour l'encodage au clavier des langues africaines suivantes: le bambara, l'éwondo, le fulfulde (ou peuhl). Les noms alias acceptables pour désigner ce jeu de caractères translittérés sont *afrful103bpiocil*, *tbambara*, *tewondo*, *tfulfulde* ou *tbra*. Pour chacun des 255 caractères présentés dans la liste plus bas, voici la signification de chacune des colonnes :

Dec		indique la valeur décimale du caractère;
Oct		indique la valeur octale du caractère;
Hex		indique la valeur hexadécimale du caractère;
Mne		fournit le code mnémotechnique à deux lettres de la norme <i>RFC 1345</i> ;
UCS2		indique la valeur UCS du caractère;
Kbd		fournit la convention d'encodage au clavier de ce caractère;
AFRFUL-102-BPI_OCIL		donne le nom du caractère, si possible en français, conformément à la norme <i>ISO 10646</i> .

Dec Oct Hex Mne UCS2 Kbd AFRFUL-102-BPI_OCIL

0	000	00	NU	0000	nul nul
1	001	01	SH	0001	soh début d'en-tête
2	002	02	SX	0002	stx début de texte
3	003	03	EX	0003	etx fin de texte
4	004	04	ET	0004	eot fin de transmission
5	005	05	EQ	0005	enq demande
6	006	06	AK	0006	ack accusé de réception positif
7	007	07	BL	0007	bel sonnerie
8	010	08	BS	0008	bs espace arrière
9	011	09	HT	0009	ht tabulation horizontale
10	012	0a	LF	000A	lf interligne
11	013	0b	VT	000B	vt tabulation verticale
12	014	0c	FF	000C	ff page suivante

Dec Oct Hex Mne UCS2 Kbd AFRFUL-102-BPI_OCIL

13	015	0d	CR	000D	cr retour de chariot
14	016	0e	SO	000E	so hors code

¹⁷ Source: Progiciels-bpi <<http://www.progiciels-bpi.ca/tcao/apercu.html>>

15	017	0f	SI	000F	si en code
16	020	10	DL	0010	dle échappement transmission
17	021	11	D1	0011	dc1 commande d'appareil un
18	022	12	D2	0012	dc2 commande d'appareil deux
19	023	13	D3	0013	dc3 commande d'appareil trois
20	024	14	D4	0014	dc4 commande d'appareil quatre
21	025	15	NK	0015	nak accusé de réception négatif
22	026	16	SY	0016	syn synchronisation
23	027	17	EB	0017	etb fin de transmission de bloc
24	030	18	CN	0018	can annulation
25	031	19	EM	0019	em fin de support
26	032	1a	SB	001A	sub caractère de substitution
27	033	1b	EC	001B	esc échappement
28	034	1c	FS	001C	is4 séparateur de fichier
29	035	1d	GS	001D	is3 séparateur de groupe
30	036	1e	RS	001E	is2 séparateur d'article
31	037	1f	US	001F	is1 séparateur de sous-article
32	040	20	SP	0020	< > espace
33	041	21	!	0021	! point d'exclamation
34	042	22	"	0022	" guillemet
35	043	23	Nb	0023	# dièse-symbole numéro (croisillon)
36	044	24	DO	0024	\$ symbole dollar
37	045	25	%	0025	% symbole pour cent
38	046	26	&	0026	& perluète
39	047	27	'	0027	' apostrophe
40	050	28	(0028	(parenthèse gauche
41	051	29)	0029) parenthèse droite
42	052	2a	*	002A	* astérisque
43	053	2b	+	002B	+ signe plus
44	054	2c	,	002C	, virgule
45	055	2d	-	002D	- tiret-trait d'union, signe moins
46	056	2e	.	002E	. point
47	057	2f	/	002F	/ barre oblique
48	060	30	0	0030	0 chiffre zéro
49	061	31	1	0031	1 chiffre un
50	062	32	2	0032	2 chiffre deux
51	063	33	3	0033	3 chiffre trois
52	064	34	4	0034	4 chiffre quatre
53	065	35	5	0035	5 chiffre cinq
54	066	36	6	0036	6 chiffre six
55	067	37	7	0037	7 chiffre sept

Dec Oct Hex Mne UCS2 Kbd AFRFUL-102-BPI_OCIL

56	070	38	8	0038	8 chiffre huit
57	071	39	9	0039	9 chiffre neuf
58	072	3a	:	003A	: deux-points
59	073	3b	;	003B	; point-virgule
60	074	3c	<	003C	< signe inférieur à
61	075	3d	=	003D	= signe égal à
62	076	3e	>	003E	> signe supérieur à
63	077	3f	?	003F	? point d'interrogation
64	100	40	At	0040	@ a commercial - arobase
65	101	41	A	0041	A lettre majuscule latine A
66	102	42	B	0042	B lettre majuscule latine B
67	103	43	C	0043	C lettre majuscule latine C
68	104	44	D	0044	D lettre majuscule latine D
69	105	45	E	0045	E lettre majuscule latine E
70	106	46	F	0046	F lettre majuscule latine F
71	107	47	G	0047	G lettre majuscule latine G

72	110	48	H	0048	H	lettre majuscule latine H
73	111	49	I	0049	I	lettre majuscule latine I
74	112	4a	J	004A	J	lettre majuscule latine J
75	113	4b	K	004B	K	lettre majuscule latine K
76	114	4c	L	004C	L	lettre majuscule latine L
77	115	4d	M	004D	M	lettre majuscule latine M
78	116	4e	N	004E	N	lettre majuscule latine N
79	117	4f	O	004F	O	lettre majuscule latine O
80	120	50	P	0050	P	lettre majuscule latine P
81	121	51	Q	0051	Q	lettre majuscule latine Q
82	122	52	R	0052	R	lettre majuscule latine R
83	123	53	S	0053	S	lettre majuscule latine S
84	124	54	T	0054	T	lettre majuscule latine T
85	125	55	U	0055	U	lettre majuscule latine U
86	126	56	V	0056	V	lettre majuscule latine V
87	127	57	W	0057	W	lettre majuscule latine W
88	130	58	X	0058	X	lettre majuscule latine X
89	131	59	Y	0059	Y	lettre majuscule latine Y
90	132	5a	Z	005A	Z	lettre majuscule latine Z
91	133	5b	<	005B	[crochet gauche
92	134	5c	//	005C	\	barre oblique inversée
93	135	5d	>	005D]	crochet droit
94	136	5e	'>	005E	^	accent circonflexe
95	137	5f	_	005F	ˉ	trait bas
96	140	60	!	0060	`	accent grave
97	141	61	a	0061	a	lettre minuscule latine a
98	142	62	b	0062	b	lettre minuscule latine b

Dec Oct Hex Mne UCS2 Kbd AFRFUL-102-BPI_OCIL

99	143	63	c	0063	c	lettre minuscule latine c
100	144	64	d	0064	d	lettre minuscule latine d
101	145	65	e	0065	e	lettre minuscule latine e
102	146	66	f	0066	f	lettre minuscule latine f
103	147	67	g	0067	g	lettre minuscule latine g
104	150	68	h	0068	h	lettre minuscule latine h
105	151	69	i	0069	i	lettre minuscule latine i
106	152	6a	j	006A	j	lettre minuscule latine j
107	153	6b	k	006B	k	lettre minuscule latine k
108	154	6c	l	006C	l	lettre minuscule latine l
109	155	6d	m	006D	m	lettre minuscule latine m
110	156	6e	n	006E	n	lettre minuscule latine n
111	157	6f	o	006F	o	lettre minuscule latine o
112	160	70	p	0070	p	lettre minuscule latine p
113	161	71	q	0071	q	lettre minuscule latine q
114	162	72	r	0072	r	lettre minuscule latine r
115	163	73	s	0073	s	lettre minuscule latine s
116	164	74	t	0074	t	lettre minuscule latine t
117	165	75	u	0075	u	lettre minuscule latine u
118	166	76	v	0076	v	lettre minuscule latine v
119	167	77	w	0077	w	lettre minuscule latine w
120	170	78	x	0078	x	lettre minuscule latine x
121	171	79	y	0079	y	lettre minuscule latine y
122	172	7a	z	007A	z	lettre minuscule latine z
123	173	7b	!	007B	{	accolade gauche
124	174	7c	!!	007C		barre verticale
125	175	7d	!)	007D	}	accolade droite
126	176	7e	'?	007E	~	tilde
127	177	7f	DT	007F		del suppression
128	200	80	PA	0080		pad caractère de bourre

129	201	81	HO	0081	hop octet supérieur prédéfini
130	202	82	BH	0082	bph arrêt permis ici
131	203	83	NH	0083	nbh aucun arrêt ici
132	204	84	IN	0084	ind index
133	205	85	NL	0085	nel à la ligne
134	206	86	SA	0086	ssa début de zone sélectionnée
135	207	87	ES	0087	esa fin de zone sélectionnée
136	210	88	HS	0088	hts arrêt de tabulateur horizontal
137	211	89	HJ	0089	htj tabulateur horizontal avec justification
138	212	8a	VS	008A	vts arrêt de tabulateur vertical
139	213	8b	PD	008B	pld interligne partiel vers le bas
140	214	8c	PU	008C	plu interligne partiel vers le haut
141	215	8d	RI	008D	ri index inversé

Dec Oct Hex Mne UCS2 Kbd AFRFUL-102-BPI_OCIL

142	216	8e	S2	008E	ss2 remplacement unique deux
143	217	8f	S3	008F	ss3 remplacement unique trois
144	220	90	DC	0090	dcs chaîne de commande d'appareil
145	221	91	P1	0091	pu1 usage privé un
146	222	92	P2	0092	pu2 usage privé deux
147	223	93	TS	0093	sts mise en mode transmission
148	224	94	CC	0094	cch annulation du caractère précédent
149	225	95	MW	0095	mw message en attente
150	226	96	SG	0096	sga début de zone protégée
151	227	97	EG	0097	ega fin de zone protégée
152	230	98	SS	0098	sos début de chaîne
153	231	99	GC	0099	sgci introducteur de caractère graphique unique
154	232	9a	SC	009A	sci introducteur de caractère unique
155	233	9b	CI	009B	csi introducteur de séquence de commande
156	234	9c	ST	009C	st fin de chaîne
157	235	9d	OC	009D	osc commande de système d'exploitation
158	236	9e	PM	009E	pm message privé
159	237	9f	AC	009F	apc commande de progiciel
160	240	a0	NS	00A0	/_ espace insécable
161	241	a1	0181	B[lettre maj. latine B crosse
162	242	a2	018A	D[lettre maj. latine D crosse
163	243	a3	0194	G[lettre maj. latine Gamma
164	244	a4	0198	K[lettre maj. latine K crosse
165	245	a5	01A4	P[lettre maj. latine P crosse
166	246	a6	01AC	T[lettre maj. latine T crosse
167	247	a7	01A9	S]	lettre maj. latine ESH
168	250	a8	019D	N]	lettre maj. latine N hameçon (enj)
169	251	a9	014A	N[lettre maj. latine ENG (Sami)
170	252	aa	0189	D]	lettre maj. latine D africain (edh)
171	253	ab	<< 00AB	<<	guillemet angulaire double vers la gauche
172	254	ac	0191	F]	lettre maj. latine F hameçon
173	255	ad	-- 00AD	\-	tiret conditionnel (trait d'union virtuel)
174	256	ae	01B3	Y[lettre maj. latine Y crosse
175	257	af	0193	G[lettre maj. latine G crosse
176	260	b0	DG 00B0	_DG	symbole degré
177	261	b1	0253	b[lettre min. latine b crosse
178	262	b2	0257	d[lettre min. latine d crosse
179	263	b3	0263	g]	lettre min. latine gamma
180	264	b4	0199	k[lettre min. latine k crosse
181	265	b5	01A5	p[lettre min. latine p crosse
182	266	b6	01AD	t]	lettre min. latine t crosse
183	267	b7	0283	s]	lettre min. latine esh
184	270	b8	0272	n]	lettre min. latine n hameçon à gauche (enj)

Dec Oct Hex Mne UCS2 Kbd AFRFUL-102-BPI_OCIL

185	271	b9	014B	n[lettre min. latine eng (Sami)
186	272	ba	0256	d]	lettre min. latine d hameçon rétroflexe
187	273	bb	>> 00BB	>>	guillemet angulaire double vers la droite
188	274	bc	0192	f]	lettre min. latine f hameçon
189	275	bd	0294	?^	lettre latine coup de glotte
190	276	be	01B4	y[lettre min. latine y crosse
191	277	bf	0260	g[lettre min. latine g crosse
192	300	c0	A! 00C0	A`	lettre maj. latine A accent grave
193	301	c1	A' 00C1	A\'	lettre maj. latine A accent aigu
194	302	c2	A> 00C2	A^	lettre maj. latine A accent circonflexe
195	303	c3	A? 00C3	A~	lettre maj. latine A tilde
196	304	c4	A: 00C4	A"	lettre maj. latine A tréma
197	305	c5	018E	E<	lettre maj. latine E réfléchi (schwa)
198	306	c6	AE 00C6	AE+	lettre maj. latine AE (ligature)
199	307	c7	C, 00C7	C\,	lettre maj. latine C cédille
200	310	c8	E! 00C8	E`	lettre maj. latine E accent grave
201	311	c9	E' 00C9	E\'	lettre maj. latine E accent aigu
202	312	ca	E> 00CA	E^	lettre maj. latine E accent circonflexe
203	313	cb	E: 00CB	E"	lettre maj. latine E tréma
204	314	cc	!! 00CC	I`	lettre maj. latine I accent grave
205	315	cd	I' 00CD	I\'	lettre maj. latine I accent aigu
206	316	ce	I> 00CE	I^	lettre maj. latine I accent circonflexe
207	317	cf	I: 00CF	I"	lettre maj. latine I tréma
208	320	d0	0190	E[lettre maj. latine E ouvert
209	321	d1	N? 00D1	N~	lettre maj. latine N tilde
210	322	d2	O! 00D2	O`	lettre maj. latine O accent grave
211	323	d3	O' 00D3	O\'	lettre maj. latine O accent aigu
212	324	d4	O> 00D4	O^	lettre maj. latine O accent circonflexe
213	325	d5	O? 00D5	O~	lettre maj. latine O tilde
214	326	d6	O: 00D6	O"	lettre maj. latine O tréma
215	327	d7	0152	OE+	digramme soudé maj. latin OE (ligature OE)
216	330	d8	0186	O[lettre maj. latine O ouvert
217	331	d9	U! 00D9	U`	lettre maj. latine U accent grave
218	332	da	U' 00DA	U\'	lettre maj. latine U accent aigu
219	333	db	U> 00DB	U^	lettre maj. latine U accent circonflexe
220	334	dc	U: 00DC	U"	lettre maj. latine U tréma
221	335	dd	018E+0301	E<'	lettre maj. latine E réfléchi accent aigu
222	336	de	0186+0301	O[\'	lettre maj. latine O ouvert accent aigu
223	337	df	01B2	V[lettre maj. latine V crosse
224	340	e0	a! 00E0	a`	lettre min. latine a accent grave
225	341	e1	a' 00E1	a\'	lettre min. latine a accent aigu
226	342	e2	a> 00E2	a^	lettre min. latine a accent circonflexe
227	343	e3	a? 00E3	a~	lettre min. latine a tilde

Dec Oct Hex Mne UCS2 Kbd AFRFUL-102-BPI_OCIL

228	344	e4	a: 00E4	a"	lettre min. latine a tréma
229	345	e5	0259	e<	lettre min. latine schwa (e culbuté)
230	346	e6	ae 00E6	ae+	lettre min. latine ae
231	347	e7	c, 00E7	c\,	lettre min. latine c cédille
232	350	e8	e! 00E8	e`	lettre min. latine e accent grave
233	351	e9	e' 00E9	e\'	lettre min. latine e accent aigu
234	352	ea	e> 00EA	e^	lettre min. latine e accent circonflexe
235	353	eb	e: 00EB	e"	lettre min. latine e tréma
236	354	ec	e! 00EC	i`	lettre min. latine i accent grave
237	355	ed	i' 00ED	i\'	lettre min. latine i accent aigu
238	356	ee	i> 00EE	i^	lettre min. latine i accent circonflexe
239	357	ef	i: 00EF	i"	lettre min. latine i tréma

240	360	f0	025B	e[lettre min. latine e ouvert
241	361	f1	n? 00F1	n~	lettre min. latine n tilde
242	362	f2	o! 00F2	o`	lettre min. latine o accent grave
243	363	f3	o' 00F3	o\	lettre min. latine o accent aigu
244	364	f4	o> 00F4	o^	lettre min. latine o accent circonflexe
245	365	f5	o? 00F5	o~	lettre min. latine o tilde
246	366	f6	o: 00F6	o"	lettre min. latine o tréma
247	367	f7	0153	oe+	digramme soudé min. latin oe (ligature oe)
248	370	f8	0254	o[lettre min. latine o ouvert
249	371	f9	u! 00F9	u`	lettre min. latine u accent grave
250	372	fa	u' 00FA	u\	lettre min. latine u accent aigu
251	373	fb	u> 00FB	u^	lettre min. latine u accent circonflexe
252	374	fc	u: 00FC	u"	lettre min. latine u tréma
253	375	fd	0259+0301	e<'	lettre min. latine schwa accent aigu
254	376	fe	0254+0301	o['	lettre min. latine o ouvert accent aigu
255	377	ff	028B	v[lettre min. latine v de ronde

Annexe 2 : Jeu de caractères AFRLIN¹⁸

Les jeux de caractères africains **afrlin** permettent de couvrir les alphabets du français, du lingala, du sango et du wolof. Le jeu de caractères *AFRLIN-104-BPI_OCIL* représente un jeu de caractères, en version encodée, qui est utilisé pour l'affichage à l'écran des langues africaines suivantes: le lingala, le sango, le wolof. Les noms alias acceptables pour désigner ce jeu de caractères encodés sont *afrlin104bpiocil*, *lingala*, *sango*, *wolof*, ou *lin*.

Le jeu de caractères *AFRLIN-105-BPI_OCIL* représente un jeu de caractères, en version translittérée, qui est utilisé pour l'encodage au clavier des langues africaines suivantes: le lingala, le sango, le wolof. Les noms alias acceptables pour désigner ce jeu de caractères translittérés sont *afrlin105bpiocil*, *tlingala*, *tsango*, *twolof* ou *tlin*. Pour chacun des 255 caractères présentés dans la liste plus bas, voici la signification de chacune des colonnes :

Dec	indique la valeur décimale du caractère;
Oct	indique la valeur octale du caractère;
Hex	indique la valeur hexadécimale du caractère;
Mne	fournit le code mnémotechnique à deux lettres de la norme <i>RFC 1345</i> ;
UCS2	indique la valeur UCS du caractère;
Kbd	fournit la convention d'encodage au clavier de ce caractère;
AFRLIN-104-BPI_OCIL	donne le nom du caractère, si possible en français, conformément à la norme <i>ISO 10646</i> .

Dec	Oct	Hex	Mne	UCS2	Kbd	AFRLIN-104-BPI_OCIL
0	000	00	NU	0000	nul	nul
1	001	01	SH	0001	soh	début d'en-tête
2	002	02	SX	0002	stx	début de texte
3	003	03	EX	0003	etx	fin de texte
4	004	04	ET	0004	eot	fin de transmission
5	005	05	EQ	0005	enq	demande
6	006	06	AK	0006	ack	accusé de réception positif
7	007	07	BL	0007	bel	sonnerie
8	010	08	BS	0008	bs	espace arrière
9	011	09	HT	0009	ht	tabulation horizontale
10	012	0a	LF	000A	lf	interligne
11	013	0b	VT	000B	vt	tabulation verticale
12	014	0c	FF	000C	ff	page suivante
13	015	0d	CR	000D	cr	retour de chariot
14	016	0e	SO	000E	so	hors code
15	017	0f	SI	000F	si	en code
16	020	10	DL	0010	dle	échappement transmission
17	021	11	D1	0011	dc1	commande d'appareil un

Dec	Oct	Hex	Mne	UCS2	Kbd	AFRFUL-102-BPI_OCIL
18	022	12	D2	0012	dc2	commande d'appareil deux
19	023	13	D3	0013	dc3	commande d'appareil trois
20	024	14	D4	0014	dc4	commande d'appareil quatre

¹⁸ Source: Progiciels-bpi <<http://www.progiciels-bpi.ca/tcao/apercu.html>>

21	025	15	NK	0015	nak accusé de réception négatif
22	026	16	SY	0016	syn synchronisation
23	027	17	EB	0017	etb fin de transmission de bloc
24	030	18	CN	0018	can annulation
25	031	19	EM	0019	em fin de support
26	032	1a	SB	001A	sub caractère de substitution
27	033	1b	EC	001B	esc échappement
28	034	1c	FS	001C	is4 séparateur de fichier
29	035	1d	GS	001D	is3 séparateur de groupe
30	036	1e	RS	001E	is2 séparateur d'article
31	037	1f	US	001F	is1 séparateur de sous-article
32	040	20	SP	0020	< > espace
33	041	21	!	0021	! point d'exclamation
34	042	22	"	0022	" guillemet
35	043	23	Nb	0023	# dièse-symbole numéro (croisillon)
36	044	24	DO	0024	\$ symbole dollar
37	045	25	%	0025	% symbole pour cent
38	046	26	&	0026	& perluète
39	047	27	'	0027	' apostrophe
40	050	28	(0028	(parenthèse gauche
41	051	29)	0029) parenthèse droite
42	052	2a	*	002A	* astérisque
43	053	2b	+	002B	+ signe plus
44	054	2c	,	002C	, virgule
45	055	2d	-	002D	- tiret-trait d'union, signe moins
46	056	2e	.	002E	. point
47	057	2f	/	002F	/ barre oblique
48	060	30	0	0030	0 chiffre zéro
49	061	31	1	0031	1 chiffre un
50	062	32	2	0032	2 chiffre deux
51	063	33	3	0033	3 chiffre trois
52	064	34	4	0034	4 chiffre quatre
53	065	35	5	0035	5 chiffre cinq
54	066	36	6	0036	6 chiffre six
55	067	37	7	0037	7 chiffre sept
56	070	38	8	0038	8 chiffre huit
57	071	39	9	0039	9 chiffre neuf
58	072	3a	:	003A	: deux-points
59	073	3b	;	003B	; point-virgule
60	074	3c	<	003C	< signe inférieur à

Dec Oct Hex Mne UCS2 Kbd AFRFUL-102-BPI_OCIL

61	075	3d	=	003D	= signe égal à
62	076	3e	>	003E	> signe supérieur à
63	077	3f	?	003F	? point d'interrogation
64	100	40	At	0040	@ a commercial-arobase
65	101	41	A	0041	A lettre majuscule latine A
66	102	42	B	0042	B lettre majuscule latine B
67	103	43	C	0043	C lettre majuscule latine C
68	104	44	D	0044	D lettre majuscule latine D
69	105	45	E	0045	E lettre majuscule latine E
70	106	46	F	0046	F lettre majuscule latine F
71	107	47	G	0047	G lettre majuscule latine G
72	110	48	H	0048	H lettre majuscule latine H
73	111	49	I	0049	I lettre majuscule latine I
74	112	4a	J	004A	J lettre majuscule latine J
75	113	4b	K	004B	K lettre majuscule latine K
76	114	4c	L	004C	L lettre majuscule latine L
77	115	4d	M	004D	M lettre majuscule latine M

78	116	4e	N	004E	N	lettre majuscule latine N
79	117	4f	O	004F	O	lettre majuscule latine O
80	120	50	P	0050	P	lettre majuscule latine P
81	121	51	Q	0051	Q	lettre majuscule latine Q
82	122	52	R	0052	R	lettre majuscule latine R
83	123	53	S	0053	S	lettre majuscule latine S
84	124	54	T	0054	T	lettre majuscule latine T
85	125	55	U	0055	U	lettre majuscule latine U
86	126	56	V	0056	V	lettre majuscule latine V
87	127	57	W	0057	W	lettre majuscule latine W
88	130	58	X	0058	X	lettre majuscule latine X
89	131	59	Y	0059	Y	lettre majuscule latine Y
90	132	5a	Z	005A	Z	lettre majuscule latine Z
91	133	5b	<	005B	[crochet gauche
92	134	5c	//	005C	\	barre oblique inversée
93	135	5d	>	005D]	crochet droit
94	136	5e	'>	005E	^	accent circonflexe
95	137	5f	_	005F	_	trait bas
96	140	60	!	0060	`	accent grave
97	141	61	a	0061	a	lettre minuscule latine a
98	142	62	b	0062	b	lettre minuscule latine b
99	143	63	c	0063	c	lettre minuscule latine c
100	144	64	d	0064	d	lettre minuscule latine d
101	145	65	e	0065	e	lettre minuscule latine e
102	146	66	f	0066	f	lettre minuscule latine f
103	147	67	g	0067	g	lettre minuscule latine g

Dec Oct Hex Mne UCS2 Kbd AFRFUL-102-BPI_OCIL

104	150	68	h	0068	h	lettre minuscule latine h
105	151	69	i	0069	i	lettre minuscule latine i
106	152	6a	j	006A	j	lettre minuscule latine j
107	153	6b	k	006B	k	lettre minuscule latine k
108	154	6c	l	006C	l	lettre minuscule latine l
109	155	6d	m	006D	m	lettre minuscule latine m
110	156	6e	n	006E	n	lettre minuscule latine n
111	157	6f	o	006F	o	lettre minuscule latine o
112	160	70	p	0070	p	lettre minuscule latine p
113	161	71	q	0071	q	lettre minuscule latine q
114	162	72	r	0072	r	lettre minuscule latine r
115	163	73	s	0073	s	lettre minuscule latine s
116	164	74	t	0074	t	lettre minuscule latine t
117	165	75	u	0075	u	lettre minuscule latine u
118	166	76	v	0076	v	lettre minuscule latine v
119	167	77	w	0077	w	lettre minuscule latine w
120	170	78	x	0078	x	lettre minuscule latine x
121	171	79	y	0079	y	lettre minuscule latine y
122	172	7a	z	007A	z	lettre minuscule latine z
123	173	7b	(!	007B	{	accolade gauche
124	174	7c	!!	007C		barre verticale
125	175	7d	!)	007D	}	accolade droite
126	176	7e	'?	007E	~	tilde
127	177	7f	DT	007F		del suppression
128	200	80	PA	0080		pad caractère de bourre
129	201	81	HO	0081		hop octet supérieur prédéfini
130	202	82	BH	0082		bph arrêt permis ici
131	203	83	NH	0083		nbh aucun arrêt ici
132	204	84	IN	0084		ind index
133	205	85	NL	0085		nel à la ligne
134	206	86	SA	0086		ssa début de zone sélectionnée

135	207	87	ES	0087	esa fin de zone sélectionnée
136	210	88	HS	0088	hts arrêt de tabulateur horizontal
138	212	8a	VS	008A	vts arrêt de tabulateur vertical
139	213	8b	PD	008B	pld interligne partiel vers le bas
140	214	8c	PU	008C	plu interligne partiel vers le haut
141	215	8d	RI	008D	ri index inversé
142	216	8e	S2	008E	ss2 remplacement unique deux
143	217	8f	S3	008F	ss3 remplacement unique trois
144	220	90	DC	0090	dcs chaîne de commande d'appareil
145	221	91	P1	0091	pu1 usage privé un
146	222	92	P2	0092	pu2 usage privé deux
147	223	93	TS	0093	sts mise en mode transmission

Dec Oct Hex Mne UCS2 Kbd AFRFUL-102-BPI_OCIL

148	224	94	CC	0094	cch annulation du caractère précédent
149	225	95	MW	0095	mw message en attente
150	226	96	SG	0096	sga début de zone protégée
151	227	97	EG	0097	ega fin de zone protégée
152	230	98	SS	0098	sos début de chaîne
153	231	99	GC	0099	sgci introducteur de caractère graphique unique
154	232	9a	SC	009A	sci introducteur de caractère unique
155	233	9b	CI	009B	csi introducteur de séquence de commande
156	234	9c	ST	009C	st fin de chaîne
157	235	9d	OC	009D	osc commande de système d'exploitation
158	236	9e	PM	009E	pm message privé
159	237	9f	AC	009F	apc commande de progiciel
160	240	a0	NS	00A0	/_ espace insécable
161	241	a1	0190+0300	E[lettre maj. latine E ouvert accent grave
162	242	a2	0190+0301	E['	lettre maj. latine E ouvert accent aigu
163	243	a3	0190+0302	E[^	lettre maj. latine E ouvert accent circonflexe
164	244	a4	0190+030C	E[\v	lettre maj. latine E ouvert caron
165	245	a5	0186+0300	O[lettre maj. latine O ouvert accent grave
166	246	a6	0186+0301	O['	lettre maj. latine O ouvert accent aigu
167	247	a7	0186+0302	O[^	lettre maj. latine O ouvert accent circonflexe
168	250	a8	019D	N]	lettre maj. latine N hameçon (ENJ)
169	251	a9	014A	N[lettre maj. latine ENG (Sami)
170	252	aa	004E+0302	N^	lettre maj. latine N accent circonflexe
171	253	ab	<< 00AB	<<	guillemet gauche
172	254	ac	004E+0308	N"	lettre maj. latine N tréma
173	255	ad	-- 00AD	\-	trait d'union virtuel (tiret conditionnel)
174	256	ae	01B3	Y[lettre maj. latine Y crosse
175	257	af	0186+030C	O[\v	lettre maj. latine O ouvert caron
176	260	b0	00B0	_DG	symbole degré
177	261	b1	025B+0300	e[lettre min. latine e ouvert accent grave
178	262	b2	025B+0301	e['	lettre min. latine e ouvert accent aigu
179	263	b3	025B+0302	e[^	lettre min. latine e ouvert accent circonflexe
180	264	b4	025B+030C	e[\v	lettre min. latine e ouvert caron
181	265	b5	0254+0300	o[lettre min. latine o ouvert accent grave
182	266	b6	0254+0301	o['	lettre min. latine o ouvert accent aigu
183	267	b7	0254+0302	o[^	lettre min. latine o ouvert accent circonflexe
184	270	b8	0272	n]	lettre min. latine n hameçon à gauche (enj)
185	271	b9	014B	n[lettre min. latine eng (Sami)
186	272	ba	006E+0302	n^	lettre min. latine n circonflexe
187	273	bb	>> 00BB	>>	guillemet droit
188	274	bc	006E+0308	n"	lettre min. latine n tréma
189	275	bd	0294	?^	lettre latine coup de glotte
190	276	be	01B4	y[lettre minuscule latine y crosse

Dec Oct Hex Mne UCS2 Kbd AFRFUL-102-BPI_OCIL

191	277	bf	0254+030C	o[_v	lettre min. latine o ouvert caron
192	300	c0	A! 00C0	A`	lettre maj. latine A accent grave
193	301	c1	A' 00C1	A'	lettre maj. latine A accent aigu
194	302	c2	A> 00C2	A^	lettre maj. latine A accent circonflexe
195	303	c3	01CD	A _v	lettre maj. latine A caron
196	304	c4	A: 00C4	A"	lettre maj. latine A tréma
197	305	c5	018E	E<	lettre maj. latine E réfléchi
198	306	c6	AE 00C6	AE+	lettre maj. latine AE (ligature)
199	307	c7	C, 00C7	C\,	lettre maj. latine C cédille
200	310	c8	E! 00C8	E`	lettre maj. latine E accent grave
201	311	c9	E' 00C9	E'	lettre maj. latine E accent aigu
202	312	ca	E> 00CA	E^	lettre maj. latine E accent circonflexe
203	313	cb	E: 00CB	E"	lettre maj. latine E tréma
204	314	cc	! 00CC	I`	lettre maj. latine I accent grave
205	315	cd	I' 00CD	I'	lettre maj. latine I accent aigu
206	316	ce	I> 00CE	I^	lettre maj. latine I accent circonflexe
207	317	cf	I: 00CF	I"	lettre maj. latine I tréma
208	320	d0	0190	E[lettre maj. latine E ouvert
209	321	d1	N? 00D1	N~	lettre maj. latine N tilde
210	322	d2	O! 00D2	O`	lettre maj. latine O accent grave
211	323	d3	O' 00D3	O'	lettre maj. latine O accent aigu
212	324	d4	O> 00D4	O^	lettre maj. latine O accent circonflexe
213	325	d5	O? 01D1	O _v	lettre maj. latine O caron
214	326	d6	O: 00D6	O"	lettre maj. latine O tréma
215	327	d7	0152	OE+	digramme soudé maj. latin OE (ligature OE)
216	330	d8	0186	O[lettre maj. latine O ouvert
217	331	d9	U! 00D9	U`	lettre maj. latine U accent grave
218	332	da	U' 00DA	U'	lettre maj. latine U accent aigu
219	333	db	U> 00DB	U^	lettre maj. latine U accent circonflexe
220	334	dc	U: 00DC	U"	lettre maj. latine U tréma
221	335	dd	011A	E _v	lettre maj. latine E caron
222	336	de	01CF	I _v	lettre maj. latine I caron
223	337	df	01D3	U _v	lettre maj. latine U caron
224	340	e0	a! 00E0	a`	lettre min. latine a accent grave
225	341	e1	a' 00E1	a'	lettre min. latine a accent aigu
226	342	e2	a> 00E2	a^	lettre min. latine a accent circonflexe
227	343	e3	01CE	a _v	lettre min. latine a caron
228	344	e4	a: 00E4	a"	lettre min. latine a tréma
229	345	e5	0259	e<	lettre minuscule latine schwa (e culbuté)
230	346	e6	ae 00E6	ae+	lettre min. latine ae (ligature ae)
231	347	e7	c, 00E7	c\,	lettre min. latine c cédille
232	350	e8	e! 00E8	e`	lettre min. latine e accent grave
233	351	e9	e' 00E9	e'	lettre min. latine e accent aigu

Dec Oct Hex Mne UCS2 Kbd AFRFUL-102-BPI_OCIL

234	352	ea	e> 00EA	e^	lettre min. latine e accent circonflexe
235	353	eb	e: 00EB	e"	lettre min. latine e tréma
236	354	ec	! 00EC	i`	lettre min. latine i accent grave
237	355	ed	i' 00ED	i'	lettre min. latine i accent aigu
238	356	ee	i> 00EE	i^	lettre min. latine i accent circonflexe
239	357	ef	i: 00EF	i"	lettre min. latine i tréma
240	360	f0	025B	e[lettre min. latine e ouvert
241	361	f1	n? 00F1	n~	lettre min. latine n tilde
242	362	f2	o! 00F2	o`	lettre min. latine o accent grave
243	363	f3	o' 00F3	o'	lettre min. latine o accent aigu
244	364	f4	o> 00F4	o^	lettre min. latine o accent circonflexe
245	365	f5	01D2	o _v	lettre min. latine o caron
246	366	f6	o: 00F6	o"	lettre min. latine o tréma

247	367	f7	0153	oe+	digramme soudé min. latin oe (ligature oe)
248	370	f8	0254	o[lettre min. latine o ouvert
249	371	f9	u! 00F9	u`	lettre min. latine u accent grave
250	372	fa	u' 00FA	u'	lettre min. latine u accent aigu
251	373	fb	u> 00FB	u^	lettre min. latine u accent circonflexe
252	374	fc	u: 00FC	u"	lettre min. latine u tréma
253	375	fd	011B	e\	lettre min. latine e caron
254	376	fe	01D0	i\	lettre min. latine i caron
255	377	ff	01D4	u\	lettre min. latine u caron