

Cet article est disponible en ligne à l'adresse :

[http://www.cairn.info/article.php?ID\\_REVUE=DN&ID\\_NUMPUBLIE=DN\\_063&ID\\_ARTICLE=DN\\_063\\_0211](http://www.cairn.info/article.php?ID_REVUE=DN&ID_NUMPUBLIE=DN_063&ID_ARTICLE=DN_063_0211)

---

## Codage informatique des langues anciennes

par Serge ROSMORDUC

| Lavoisier | *Document numérique*

2002/3-4 - Volume 6

ISSN 1279-5127 | pages 211 à 224

---

Pour citer cet article :

—Rosmorduc S., Codage informatique des langues anciennes, *Document numérique* 2002/ 3-4, Volume 6, p. 211-224.

---

Distribution électronique Cairn pour Lavoisier.

© Lavoisier. Tous droits réservés pour tous pays.

La reproduction ou représentation de cet article, notamment par photocopie, n'est autorisée que dans les limites des conditions générales d'utilisation du site ou, le cas échéant, des conditions générales de la licence souscrite par votre établissement. Toute autre reproduction ou représentation, en tout ou partie, sous quelque forme et de quelque manière que ce soit, est interdite sauf accord préalable et écrit de l'éditeur, en dehors des cas prévus par la législation en vigueur en France. Il est précisé que son stockage dans une base de données est également interdit.

---

# Codage informatique des langues anciennes

## Le cas des hiéroglyphes égyptiens

**Serge Rosmorduc**

*IUT de Montreuil  
Université de Paris 8  
140, rue de la Nouvelle France  
93100 Montreuil  
rosbord@iut.univ-paris8.fr*

---

*RÉSUMÉ. Nous présentons les problèmes liés au codage des hiéroglyphes égyptiens. Après avoir brièvement expliqué le système d'écriture hiéroglyphique, nous faisons l'historique des codages utilisés, dont la plupart proviennent des besoins de l'imprimerie. Nous insistons ensuite sur les problèmes théoriques liés à l'opération de codage elle-même, et à ses finalités, pour conclure qu'il est aujourd'hui prématuré d'inclure l'ensemble des hiéroglyphes dans ISO 10646.*

*ABSTRACT. We explain the problems met while encoding ancient Egyptian hieroglyphs. After a brief presentation of the hieroglyphic system, we give an historical account of the various encodings which have been proposed for this writing system. Next, we concentrate on the theoretical problems linked to the encoding process itself. We conclude that it's too early to include all hieroglyphic signs in ISO 10646.*

*MOTS-CLÉS : codage, égyptien ancien, hiéroglyphes, égyptologie.*

*KEYWORDS: encoding, ancient egyptian, hieroglyphs, egyptology.*

---

## 1. Introduction

L'égyptologie est une science éminemment philologique. Elle se nourrit de textes. L'informatisation a, dans ce contexte, deux mérites : elle facilite l'édition des textes, en baissant notablement le coût de l'impression ; elle permet d'autre part de constituer des bases de données lexicographiques, denrées précieuses pour l'étude d'une langue morte où les mots mal connus sont légions.

La présente contribution a pour objet de présenter les problèmes liés au codage des hiéroglyphes. Ceux-ci sont essentiellement de deux ordres. Il faut tout d'abord assigner un code à chaque signe. Mais de plus, comme la disposition des signes est complexe et parfois *signifiante*, il convient de noter leurs positions relatives. Le standard actuel dans le domaine est le *manuel de codage* [BUU 88].

Nous présenterons tout d'abord brièvement le système d'écriture égyptien. Nous ferons ensuite un historique des analyses et des représentations de celui-ci, pour enfin aborder en détail les problèmes théoriques liés à son codage informatique, dont Unicode/ISO-10646.

## 2. Description sommaire du système hiéroglyphique

Les hiéroglyphes apparaissent en Égypte vers -3100, -3200. Ils sont liés à l'embryon d'état qui se met alors en place. Un des points marquants de ce développement est sa rapidité. L'écriture a certes évolué par la suite, mais ses grands principes sont en place dès le début. Pour résumer, les hiéroglyphes égyptiens mêlent des éléments idéographiques à une majorité d'éléments phonétiques (ou plus exactement consonantiques).

Une autre caractéristique de l'Égypte antique est que durant la majeure partie de son existence, deux écritures furent utilisées en parallèle : les hiéroglyphes pour les inscriptions monumentales et sacrées, et l'hieratique, une écriture cursive dérivée des hiéroglyphes, pour la pratique quotidienne. Contrairement aux égyptologues, qui aujourd'hui apprennent d'abord à lire les hiéroglyphes, puis passent à l'hieratique (lequel constitue d'ailleurs une spécialité), les scribes égyptiens apprenaient en premier lieu la cursive, et la plupart s'en tenaient là.

### 2.1. Les signes

Les signes égyptiens ont plusieurs emplois possibles : ils peuvent être phonétiques, idéographiques, ou servir de déterminatifs. Il faut noter qu'un même signe peut avoir plusieurs emplois.

### 2.1.1. Les signes phonétiques

Ils notent 1, 2 ou 3 consonnes :  = *m* (unilitaire);  = *mn* (bilitaire);  = *bpr* (trilitaire). Ces valeurs sont généralement justifiable par un principe proche du rébus. Ainsi, le signe  qui représente une bouche, acquiert la valeur *r* parce que le mot « bouche » en égyptien est en gros prononcé « ro ». Malgré leur figurativité, ces signes, et en particulier les unilitaires, fonctionnent normalement sans référence aucune à ce qu'ils représentent. Un , pour qui a l'habitude de lire les hiéroglyphes, n'évoque plus une chouette ou un hibou. C'est simplement la manière normale d'écrire « m ».

### 2.1.2. Les idéogrammes

Il s'agit de signes dont le dessin renvoie de manière directe ou indirecte à leur sens. Ils sont souvent suivis du signe « 1 » qui les marque normalement comme idéogrammes. Exemples :  = « taureau » ;  = « bouche ».

### 2.1.3. Déterminatifs

Ce sont des classificateurs sémantiques utilisés en fin de mot. Ils permettent à la fois de segmenter le texte et de lever des ambiguïtés entre des termes qui seraient homographes sinon. Parmi les déterminatifs les plus typiques, citons , le rouleau de papyrus qui détermine les mots abstraits ; , le « bras armé » qui détermine les actions violentes et  qui détermine les mots désignant des choses petites ou mauvaises.

## 2.2. Fonctionnement du système

Un mot est typiquement écrit en utilisant des signes phonétiques + un déterminatif. Soit par exemple le mot    qui se transcrit *ʒpd*, et signifie « oiseau ». Il est écrit avec trois signes unilitaires  (pour *ʒ* « aleph »),  (pour *p*) et  (pour *d*). Ces signes sont suivis d'un déterminatif, qui est ici un dessin de canard, mais qui vaut comme déterminatif générique des oiseaux.

Le système est redondant ; un signe complexe est souvent complété par tout ou partie de sa lecture en signes unilitaires. Considérons le mot suivant :    , = *tʒšw*, (tachou) « frontière ». On peut l'analyser ainsi :  (*t*) +  (*tʒ*) +  (*š*) +  (*dét. : couper*) +  (*dét. : territoire*). On remarquera que le bilitaire  est précédé de l'écriture explicite de sa première consonne, et que le mot a plusieurs déterminatifs (selon une habitude qui tend à se développer aux époques tardives).

Les catégories que nous avons décrites sont fort utiles d'un point de vue pédagogique, mais les frontières qui les séparent sont floues : on analyse le signe  comme

un idéogramme dans le mot , *irt* « œil », mais comme un phonogramme dans le verbe , *irt* « faire ».

Terminons cet exposé du système en soulignant une évidence, sur laquelle nous reviendrons plus loin : un hiéroglyphe est un dessin... et malheureusement pour l'informaticien, tout dessin égyptien est un hiéroglyphe en puissance. Il en résulte que, plus que d'autres systèmes d'écriture, le système hiéroglyphique est *ouvert*.

### 3. Historique de la question

Le nombre des signes varie d'environ 800 au Moyen Empire à plusieurs milliers à la fin de l'histoire égyptienne. Le nombre même des signes appelle un classement. Nous avons quelques traces de classements antiques. Dans les systèmes mis au point par les égyptologues, le désir de pouvoir *imprimer* les signes, et donc de disposer d'un codage commode pour les fontes en plomb, a joué un rôle non négligeable. Enfin, l'informatisation des textes s'est située dans la droite ligne de la typographie en plomb.

#### 3.1. Classements antiques

La réflexion que les scribes égyptiens eurent sur leurs écritures est surtout attestée par la virtuosité de leur pratique ; ici, comme dans la plupart des domaines, les écrits « théoriques » sont rares.

Néanmoins nous avons quelques sources. Outre le papyrus Jumilhac, qui contient quelques remarques mythologiques concernant tel ou tel signe, les hasards de la conservation nous ont légué un papyrus tardif, provenant de Tanis, qui porte une liste de signes hiéroglyphiques, expliqués et commentés en hiératique. La liste semble être organisée partiellement selon la *forme* des signes en hiératique, ce qui est assez naturel, puisque c'était l'écriture « normale » pour un scribe égyptien.

#### 3.2. Les classements égyptologiques

Dès le départ, les égyptologues jugèrent nécessaire d'organiser les signes ([JAN 72] et [BUD 20] p. xxvi–xxx). Champollion choisit de les grouper selon l'objet représenté et non selon leurs valeurs. Ce faisant, il suivait consciemment un modèle ancien. Les Coptes, descendants des Égyptiens anciens, avaient réalisé des dictionnaires, appelés *scalae* dans la tradition érudite. Ces dictionnaires étaient organisés par thèmes, en partant du plus haut (c.-à-d. Dieu pour les coptes chrétiens) jusqu'au plus bas. La liste de signes de Champollion commence donc par les éléments célestes, puis les hommes, les parties du corps humain, les quadrupèdes, les oiseaux... pour finir par les éléments qu'il a du mal à classer : « plusieurs formes géométriques » et « images monstrueuses ». Comme Champollion le pressentait, ce système était bien plus ancien que l'égypte chrétienne et remontait à la période pharaonique. La découverte de

l'*onomasticon* d'Amenemopé, lexique thématique daté de 1000 avant notre ère, en a apporté la confirmation. Il convient cependant de préciser qu'*onomastica* et *scalae* constituent des listes lexicographiques et non des listes de signes. Nous n'avons pas d'attestation de ce principe pour classer des signes dans l'antiquité.

Le classement de Champollion fut repris, et quelque peu modifié par Lepsius en 1846, cette fois-ci dans l'intention de faciliter l'usage d'une fonte hiéroglyphique. À partir de ce moment, les besoins de l'imprimerie jouèrent un rôle déterminant dans la classification des signes.

Deux grands systèmes ont dominé le siècle, chacun associé à une fonte particulière. Le premier est celui de l'Institut français d'archéologie orientale (IFAO), mis en place vers 1907 par Chassinat, pour organiser la fonte de l'institut, qui, partant d'une base créée au XIX<sup>e</sup> siècle pour l'Imprimerie nationale par J.-J. Dubois et J.-A. Letronne, s'était considérablement étendue. Dans cette fonte, les caractères sont simplement numérotés en continu, sans groupement thématique. Un signe nouveau prend simplement le premier numéro disponible. Cette liste est marquée par le grand nombre de signes disponibles (plus de 5000) ; son développement est intimement lié aux travaux menés par de Rougé et Chassinat, et poursuivis depuis, d'édition des textes des grands temples ptolémaïques (rappelons que l'écriture de la période ptolémaïque se caractérise par le très grand nombre de signes utilisés).

L'autre grand système est celui mis au point pour la fonte dite *Gardiner*. L'égyptologue et linguiste Alan H. Gardiner est l'auteur d'une des grammaires fondamentale de l'égyptien classique. Pour les besoins de son édition, il regrettait que les fontes disponibles ne fussent pas adaptées, tant d'un point de vue typographique (les caractères étant trop grands) que d'un point de vue paléographique (leurs formes étant généralement inspirées d'exemples tardifs). Il commandita donc la réalisation d'une nouvelle fonte, dont les caractères furent dessinés par N. et N. de Garis Davies, un couple d'égyptologues et de dessinateurs de grand talent, à partir de formes trouvées dans les textes de la XVIII<sup>e</sup> dynastie. Pour l'organisation des signes, Gardiner s'inspira du système de Lepsius. La liste des signes qu'il donne en annexe de sa grammaire est donc découpée en sections : 'A' : « l'homme et ses activités », 'B' : « la femme et ses activités »... 'M' : « arbres et plantes » pour finir par 'Z' : « traits » et 'Aa' : « non classés ». Le code d'un signe est constitué par sa catégorie suivie d'un numéro d'ordre.

Le principe utilisé convient pour la plupart des signes. Néanmoins, il demande une identification de l'élément représenté par le signe. Il peut arriver que le référent d'un signe soit inconnu alors même que la valeur du signe est connue. C'est par exemple le cas du signe Aa27, , signe phonétique dont la valeur est *nd* (nedj). Parfois même, nous savons que les Égyptiens anciens eux-mêmes ont changé d'interprétation quant au référent d'un signe.

Dans d'autres cas, plusieurs signes anciens se sont fondus en un seul. , codé N18, provient de deux signes originellement distincts : l'un représente un pain ; l'autre une île. Le codage N18 le range dans la catégorie « ciel, terre, eau », mais il aurait

aussi pu trouver sa place dans la catégorie X « pains et gâteaux ». D'autres signes sont composites, et pourraient donc se ranger dans plusieurs catégories.

Parfois, les progrès de la science voudraient que l'on déplaçât un signe. Dans la majeure partie des cas, fort heureusement, il s'agit d'un signe de la catégorie Aa (indéterminé), dont le référent a été identifié. Néanmoins, pour des raisons pratiques, seule change la place des signes dans la présentation des listes. Leur code demeure.

### **3.3. Le passage à la représentation informatique**

Le coût de l'impression de textes hiéroglyphiques au moyen de caractères en plomb était prohibitif. Ce système a cependant été employé jusqu'aux années 1990, en particulier à l'IFAO, qui disposait d'un personnel compétent et d'une fonte très riche. Il n'en restait pas moins que la publication d'un ouvrage était une affaire de longue haleine. D'ailleurs, dans beaucoup de cas, les égyptologues ont préféré dessiner leurs textes plutôt que de les faire composer.

Les possibilités offertes par l'informatique dans le domaine de l'édition des textes ont assez tôt intéressé les égyptologues. On notera que l'un des premiers systèmes existant, Glyph, a été développé, sur gros systèmes, dès la fin des années 60, par Jan Buurman [BUU 76], astrophysicien (et égyptologue amateur). Les années 70 virent quelques tentatives, notamment celle de Michael Hainsworth [HAI 79], du CNRS, à l'aide d'un Apple II, pour imprimer le corpus des textes des pyramides.

Vers le milieu des années 80, les égyptologues qui utilisaient l'informatique organisèrent un groupe de travail. L'un des buts de ce groupe était de développer un standard acceptable de codage des textes hiéroglyphiques, en vue de leur impression mais aussi de la constitution de corpora électroniques. Le premier rapport de ce groupe *Informatique et Égyptologie*, édité conjointement (et artisanalement) par l'université d'Utrecht et par la Sorbonne, garde la trace des discussions et des problématiques d'alors [GRI 85, BUU 85]. À cette première table ronde, il fut décidé de développer le standard autour de la notation utilisée par le logiciel Glyph. Ce standard porte désormais le nom de *manuel de codage* [BUU 88]. Publié en 1988, ce manuel a vieilli. Il n'en reste pas moins à la base de la totalité des systèmes professionnels utilisés en égyptologie.

Michael Everson a proposé un codage ISO 10646 [EVE 99] des hiéroglyphes. Fort logiquement, il s'est appuyé sur le manuel, et surtout sur le logiciel winglyph, qui en est l'un des principaux représentants.

## **4. Le codage informatique**

### **4.1. Le manuel de codage**

Le manuel de codage comporte deux parties : d'une part, une série de constructions qui permettent de disposer des signes les uns relativement aux autres, et d'autre part une liste de codes pour les hiéroglyphes.

Les signes sont codés selon leur code dans la grammaire de Gardiner. Les signes les plus fréquents admettent généralement un second code, qui est phonétique<sup>1</sup>. Le manuel (et plus encore ses réalisations effectives) propose une extension de la liste de Gardiner, afin de permettre le codage de signes plus nombreux. La liste de Gardiner, en effet, contient 800 signes, ce qui suffit pour les textes transcrits de l'hiéroglyphique, et convient passablement pour les textes du Moyen Empire. Pour les textes plus récents, et en particulier ceux de la période ptolémaïque, c'est beaucoup trop peu. La liste est donc étendue de plusieurs milliers de codes, classés selon plus ou moins le même principe.

En plus de ces codes, un certain nombre d'opérateurs permettent de placer les signes. Les Égyptiens pouvaient écrire horizontalement ou verticalement, et de droite à gauche ou gauche à droite. Dans une orientation donnée, mettons, en ligne et de gauche à droite, les signes étaient disposés en *cadrats*, de manière à remplir l'espace le plus possible. Certains signes, comme , remplissent un cadrat à eux seuls ; mais d'autres, comme , sont trop courts ou pas assez hauts. Dans ce cas, les signes se groupent en carrés. Ainsi, le mot « ciel », *pt*, devrait-il s'écrire    (*p + t + déterminatif/idéogramme « ciel »*). En fait, il s'écrit généralement  .

Le *manuel* permet donc de superposer des signes en les séparant par ' : ', et de les grouper horizontalement en les séparant par des '-'. Comme la priorité de ' : ' est supérieure à celle de '-', notre groupe, composé des signes *p* () et *t* () et du déterminatif *pt* () devrait s'écrire (*p-t*) :*pt* en utilisant des parenthèses. Pour réduire l'usage de celles-ci, on introduit un nouvel opérateur, '\*', de même valeur que '-', mais avec une plus forte priorité, ce qui fait que notre groupe, en fin de compte, se code *p\*t* :*pt*.

#### 4.2. La proposition ISO 10646

La première proposition de codage ISO 10646 des hiéroglyphes s'appuyait sur l'état du logiciel *winglyph* et de sa liste de signes en 1999. Cette proposition a été durement critiquée [SCH 99] par l'égyptologue W. Schenkel, pionnier de l'utilisation de l'informatique en égyptologie (dès 1970), et sans doute l'un de ceux qui ont le plus travaillé sur les questions du codage des hiéroglyphes [SCH 74]. Le fond de la critique est que Schenkel estime que les listes existantes, et en particulier celle de *winglyph*, constituent un répertoire trop partiel de l'égyptien. En particulier, ces listes ont été créées pour les besoins de l'impression de textes, et non au vu d'un inventaire à visée exhaustive des signes existants dans les sources. Le niveau de description rendu possible par ces listes est très variable. Dans certains cas, des signes ont été rajoutés à la demande de l'éditeur d'un texte particulier : le rendu sera alors très proche de celui

1. Ce qui oblige à normaliser les textes pour effectuer des recherches. Néanmoins, on se rappellera que le *manuel* date de 1988, et qu'à l'époque le code était entré manuellement, sans aide à la saisie.

du texte d'origine. Dans d'autres cas, il faudra ruser, voire employer un signe erroné à la place du signe correct, faute de l'avoir dans la fonte. Dans d'autre cas enfin, des signes fréquents à certaines périodes n'apparaissent pas dans la liste de winglyph, ni donc dans la proposition ISO 10646.

On peut rajouter que cette proposition incluait les symboles nécessaires au positionnement des signes selon le Manuel de Codage ; or ceux-ci se révèlent partiellement insuffisants. Il est d'ailleurs probable à court terme que le *manuel* soit étendu ou remplacé par une DTD XML. On conçoit donc qu'il serait dommage de figer dans ISO 10646 un système appelé à disparaître à moyen terme.

Il nous semble, en ce qui concerne ISO 10646, que la première proposition de M. Everson était trop ambitieuse. La seule partie sûre et stable concerne la liste des 800 signes de Gardiner. Pour le reste des signes, il serait souhaitable d'attendre, pour aboutir à une meilleure organisation. Cela est d'autant plus vrai que la liste étendue n'est pas vraiment utile pour beaucoup d'utilisations. La liste de Gardiner, moyennant quelques additions, couvre la plupart des besoins de l'édition de textes dont l'original est en hiératique, et elle est utilisable pour les discussions grammaticales, où il n'est pas trop grave de normaliser un peu le texte. W. Schenkel estime que scientifiquement, elle est améliorable ; néanmoins, comme elle est utilisable et bien ancrée dans les pratiques, on peut considérer qu'elle a sa place dans ISO 10646.

Tout récemment, dans message publié le 18 juillet 2002 sur la liste de diffusion électronique *Ancient Near East*, C. E. Jones, bibliographe de l'*Oriental Institute* à Chicago, a demandé des lettres de soutien de spécialistes pour appuyer une nouvelle proposition de M. Everson, qui, après consultation d'un certain nombre de spécialistes, a effectivement limité la liste des hiéroglyphes codés à celle proposée par Gardiner.

## 5. De la difficulté de coder les hiéroglyphes

### 5.1. *Un système ouvert*

Le titre ci-dessus peut sembler paradoxal dans la mesure où nous parlons d'une langue morte. Néanmoins, le hasard des découvertes, la publication de nouveaux textes, voire le réexamen des textes connus peuvent à tout moment apporter de « nouveaux » signes. Certes, tout système d'écriture est peu ou prou ouvert, comme l'atteste l'apparition de l'euro sur les claviers d'ordinateurs. Néanmoins, le système hiéroglyphique est *intrinsèquement* ouvert.

Dans les textes de la pratique, les scribes ont, dès le départ, utilisé l'hiératique. Les hiéroglyphes formaient une écriture de prestige, destinée à sacraliser un texte. Cette spécialisation a permis aux hiéroglyphes de conserver leur caractère iconique ; ils étaient certes long à dessiner, mais n'étaient pas employés dans un contexte où la rapidité d'écriture était déterminante.

Le hiéroglyphe égyptien est donc resté à la fois dessin et signe d'écriture. Dans le même temps, et très rapidement, l'art égyptien s'est « hiéroglyphisé » : les représen-

tations que nous qualifierions de figuratives peuvent, dans une certaine mesure, s'interpréter comme de grands hiéroglyphes. Un exemple simple, mais typique, permettra de comprendre : les noms propres sont normalement terminés par le déterminatif qui représente un homme assis (𓂏). Cependant, ce signe est fréquemment omis, en particulier sur les parois des tombes. En effet, le nom d'un personnage y est généralement gravé devant une représentation de celui-ci. Cette représentation fonctionne à la fois comme dessin, montrant par exemple le propriétaire de la tombe en train d'inspecter ses domaines, et comme signe d'écriture, tel un déterminatif géant. Le signe fait alors système à la fois avec d'autres dessins en participant à une scène, et avec des signes d'écritures, en occupant la place normalement dévolue à un déterminatif.

Inversement, les hiéroglyphes deviennent des éléments de scène à part entière : les dieux tiennent par exemple fréquemment le signe 'nh ꜥ qui signifie « la vie », ceci pour marquer qu'ils sont dispensateur de vie. La chose est tellement habituelle que personne ne s'en étonne. Néanmoins, nous avons ici un hiéroglyphe tout ce qu'il y a de courant qui acquiert un statut d'élément iconique de plein droit dans les scènes.

Par ailleurs, une observation plus fine des scènes champêtres, par exemple, montre que les oiseaux y sont souvent figurés selon les canons des hiéroglyphes ; leur position, leur attitude est exactement celle du hiéroglyphe. Leurs pattes sont bien à plat, comme s'ils reposaient sur le sol, alors qu'ils sont perchés dans des fourrés de papyrus.

Plus encore, le texte est parfois susceptible d'une double analyse : comme texte *et* comme dessin. Les conventions standard d'écriture voudraient par exemple que le texte « adorer Rê » fût écrit ainsi : \*𓂏𓂏𓂏 ; \*𓂏𓂏 étant le verbe « adorer », écrit au moyen d'un hiéroglyphe phonétique, *dw*𓂏, et d'un déterminatif, 𓂏 ; 𓂏 étant utilisable comme déterminatif ou comme hiéroglyphe pour écrire « Rê ». Ce texte est souvent écrit : 𓂏𓂏𓂏. Les signes étant disposés de manière à former la scène même qu'ils écrivent.

Plus tard, à la période ptolémaïque, ce genre de jeux sur l'écriture fut systématique. Souvent, on l'utilisa pour superposer au texte un second discours, non plus redondant comme dans l'exemple précédent, mais complémentaire. L'idée n'était pas d'obscurcir l'écriture, mais bien plutôt de jouer sur ses capacités pour exprimer dans le même temps le plus de choses possibles. En particulier, ces jeux sont fréquents dans l'écriture des noms des dieux, l'idée étant, par le choix des signes et par leur disposition, de suggérer graphiquement une épithète ou une fonction du dieu par l'écriture de son nom. S. Sauneron, dans *les prêtres de l'ancienne Égypte* [SAU 88], p. 138, en propose un très bel exemple : l'écriture



pour le nom du dieu Ptah. Chacun des signes est phonétique (𓂏 = *p(t)*, 𓂏 = *t(ʃ)* et 𓂏 = *h(h)*), mais dans le même temps, le groupe suggère l'une des fonctions parfois assignées au dieu Ptah, qui est d'avoir séparé le ciel de la terre.

De cette proximité entre le signe et le dessin, il ressort que les frontières entre les deux catégories sont plus que perméables : un dessin, en Égypte, n'est jamais qu'un grand hiéroglyphe (D. van der Plas et R. Vergnieux ont d'ailleurs esquissé un système pour coder l'*iconographie* égyptienne au moyen des codes utilisés pour coder les hiéroglyphes [PLA 93]). Normalement, un texte est écrit pour être lisible : le répertoire des signes est donc *a priori* fixé. Il n'en demeure pas moins que la porosité des frontières entre signe et dessin fait que les scribes et dessinateurs se réservent toujours le droit de lancer des ponts entre le dessin et le texte, de créer de nouveaux signes, voire de donner de nouvelles valeurs à des signes existants.

Il peut tout simplement s'agir d'intégrer des nouveautés dans le système. Ainsi, l'arrivée du cheval en Égypte (vers 1700 av. J.-C.) entraîne la création d'un nouveau signe, . Plus généralement, les déterminatifs peuvent être plus ou moins précis ; utiliser un déterminatif nouveau et figuratif ne porte donc pas à conséquence.

Pour compliquer les choses, nous parlons d'une écriture qui a vécu sur 3000 ans. Au cours de ces trois millénaires, certains signes, originellement distincts, se sont confondus. Un même signe, inversement, a pu être décliné en plusieurs variantes, susceptibles de se spécialiser.

Pour qui code un texte dont l'original est en hiéroglyphes, le nombre de problèmes est limité. Par contre, dès que l'on aborde les textes monumentaux, on se retrouve très rapidement confronté à des signes non disponibles dans les fontes.

## **5.2. Pourquoi coder les hiéroglyphes ? entre fac simile et standardisation**

Le codage informatique tel que le sous-tend UNICODE distingue la notion de *caractère* et celle de *glyphe*. Le premier est une unité linguistique, et le second une unité graphique, les liens entre les deux étant complexes. Cette vision est adaptée à la *production* de textes. Par contre, il n'est pas sûr qu'elle le soit à l'édition scientifique de textes anciens, car la distinction entre le graphique et le linguistique ne sont pas toujours nettes.

Un des problèmes qui se pose à qui veut proposer un codage informatique des hiéroglyphes est la variété des motivations possibles. La question s'est déjà posée à propos de l'usage de hiéroglyphes imprimés. Ceux-ci sont utilisés pour deux usages : dans le corps d'un texte essentiellement imprimé en caractères latins, ou pour donner une édition « propre » d'un texte égyptien. Le premier usage est fréquent dans les grammaires et dans les articles scientifiques ; il n'a jamais posé le moindre problème théorique. L'usage de fontes dans des dictionnaires est déjà une autre affaire. On désire généralement disposer de graphies les plus fidèles possibles ; aussi la plupart des dictionnaires sont-ils entièrement manuscrits. À vrai dire, pour les dictionnaires, le problème était plus technique que théorique. Le coût de la composition typographique d'une part, la difficulté de disposer de tous les signes nécessaires pour composer l'ouvrage d'autre part, ont conduit la plupart des auteurs de dictionnaires égyptiens à reprographier un original manuscrit (on citera par exemple le monumental *Wörterbuch der aegyptischen*

*Sprache* [ERM 63] en 13 volumes !). Le récent dictionnaire de R. Hannig [HAN 95], quant à lui, bénéficie de l'informatique ; les signes manquant dans les fontes ont été simplement insérés sous forme de dessins.

Le problème est tout autre lorsque qu'il s'agit de l'édition scientifique d'un texte. En toute rigueur, la seule solution acceptable en pareil cas est le *fac simile*. Dans son ouvrage *Living in the Past*, [MAN 94] Peter der Manuelian publie ainsi, pour chacune des stèles qu'il étudie : une photographie, un dessin au trait, et, pour *faciliter la lecture*, un texte en hiéroglyphes standardisés. Seule cette démarche permet de ne pas perdre d'information. La photographie est souvent difficile à lire, car tel ou tel signe peut être quasiment invisible ; le *fac simile* au trait comporte normalement une part d'interprétation minimale, et suffit théoriquement. Néanmoins, la publication d'une version « standardisée » du texte le rend plus accessible. D'une part, le scribe antique pouvait avoir une écriture plus ou moins soignée ; d'autre part, les habitudes d'écriture d'une époque donnée peuvent déconcerter les égyptologues non spécialistes de celle-ci.

L'informatisation permet théoriquement de produire un texte « propre », tout en gardant une souplesse suffisante pour approcher des qualités du *fac simile* si nécessaire. Il est *a priori* plus facile, si un signe manque, de le dessiner pour un logiciel de composition automatique que de faire fondre un nouveau caractère dans une fonte en plomb. Gardons néanmoins à l'esprit qu'un texte informatisé est forcément standardisé. La fidélité à l'original n'est donc pas parfaite.

Cependant, cette standardisation a elle aussi son intérêt ; la publication récente des chapelles osiriennes du temple de Dendara [CAU 97] le prouve. Les chapelles osiriennes datent de la fin de l'époque ptolémaïque, vers 50 av. J.-C. À cette époque, le nombre de signes en usage était de l'ordre de plusieurs milliers, et le nombre de signes différents utilisés dans ces textes est d'environ 2000.

La publication des chapelles par S. Cauville devait au départ comporter les photographies des parois, une édition en hiéroglyphes « d'imprimerie » des textes, et une traduction. La composition du texte a été effectuée en étroite collaboration avec Hans van den Berg et Jochen Hallof, l'un auteur du logiciel *winglyph*, et l'autre auteur des fontes utilisées par ledit logiciel. Ceux-ci ont créé des signes spécialement pour l'occasion, mais le plus intéressant est qu'ils ont proposé de réaliser un index des mots attestés dans ces textes. Un tel index n'aurait pas vu le jour si les textes n'avaient été codés selon le *manuel de codage*. De plus, S. Cauville a pu publier un autre ouvrage, qui recense les signes attestés à Dendera et leurs valeur [CAU 01].

On voit là l'intérêt pratique du codage pour le scientifique. Néanmoins la question du code est loin d'être simple.

L'éditeur d'un texte égyptien souhaite normalement donner le plus de renseignements possibles sur ce texte. Dans ces renseignements, on peut distinguer plusieurs niveaux d'interprétation. Le *fac simile* est relativement neutre, et ne comporte que peu d'interprétations. Un texte en hiéroglyphes standardisés est très loin d'être aussi inno-

cent. Pour chaque signe du texte d'origine, l'éditeur a choisi tel ou tel signe dans les fontes. Il peut s'être trompé, ou avoir fait un choix malheureux.

Le statut de ce codage est donc à mi-chemin entre le niveau graphique pur et des niveaux plus interprétatifs. Les utilisations possibles du texte codé par d'autres chercheurs seront d'autant plus variées que le codage sera proche du niveau graphique. Nous sommes ici dans une situation qui peut paraître étrange à l'informaticien, dont la tendance naturelle est plutôt de coder le fond, et de laisser à l'ordinateur le soin de *produire* la forme. C'est que dans notre cas, nous disposons de la forme, et nous voulons remonter au fond. Il vaut mieux donc coder la forme que le fond, car celui-ci sera le résultat final du processus, et le reflet d'une problématique donnée. De plus, la dimension éminemment graphique des hiéroglyphes est en contradiction avec la réduction qu'impose le codage. Plus encore, on peut dire que l'égyptologue ne code jamais un *texte*, instance abstraite d'une œuvre, mais un *document matériel* précis : tel ou tel papyrus, telle stèle, etc.

Pour montrer au lecteur pourquoi il est intéressant de respecter l'aspect graphique des textes, citons Sylvie Cauville [CAU 01] :

Tout autant que dans la gravure elle-même, on discerne des “mains” dans le choix des hiéroglyphes, raison pour laquelle Émile Chassinat et Fr. Daumas ont tenu à ne pas normaliser les hiéroglyphes.

De ces mains, S. Cauville tente de tirer des informations sur l'élaboration du programme décoratif de Dendera. Or les différences observées sont *a priori* minimales ; la valeur des signes n'en est pas affectée.

Il reste que le codage complet de la forme devrait inclure tellement de détails qu'il est impraticable. Il faut faire des compromis. Un des problèmes est que les codages disponibles ne sont pas toujours cohérents [SCH 74]. Par exemple, la fonte de Gardiner propose des signes du Nouvel Empire. Pour répondre à des besoins ponctuels, certaines formes anciennes sont aussi présentes. Le signe du rouleau de papyrus (Y1 : ) , sous sa forme récente, et doublée de sa forme de l'ancien Empire (Y2 : ). On remarquera que deux codes distincts sont attribués à ces signes. D'un autre côté, la fonte Gardiner contient aussi un signe T9  doublé d'un signe T9A . T9A est aussi une variante ancienne de T9, mais ici le caractère variant est marqué explicitement dans le code. Il n'y a donc pas de cohérence complète dans le codage ; il n'y en a pas non plus dans le choix des signes : certaines variantes sont ignorées, certaines sont disponibles. Cela dépend essentiellement des besoins du commanditaire de la fonte. C'était vrai pour la fonte de Gardiner, mais cela le reste pour celle de winglyph. L'auteur des fontes de winglyph est un spécialiste de l'Égypte ptolémaïque ; le répertoire de signes disponibles dans ce domaine est donc assez étendu. D'un autre côté, certains signes fréquents et nécessaires pour coder des textes plus anciens ne sont pas fournis.

En résumé, le codage d'un texte va répondre, dans les meilleurs cas, aux préoccupations de son codeur. Il en résulte que le texte codé est *variable* pour un même texte d'origine, et que le codage est un processus complexe et *interprétatif*. Ces caractéris-

tiques étant intrinsèques à l'objet codé lui-même, il n'y a pas de solution exhaustive disponible.

### 5.3. *Le codage dans tksesh*

Nous développons un environnement coopératif pour le philologue, en particulier l'égyptologue, articulé autour d'une base de données de textes hiéroglyphiques. Le problème du codage et de l'extension des fontes s'est donc posé à nous. Nous avons finalement décidé de ne pas utiliser de fontes. Les *formes graphiques* des signes sont entrées dans la base de données, chacune a un code unique qui dépend simplement des conditions de son entrée et de son auteur. Ces glyphes sont éventuellement mis en correspondance avec des codes officiels quand ils existent. En distinguant le document d'origine, ses codages possibles par un ou plusieurs utilisateurs, le signe sur le texte d'origine, ce que le codeur en a vu, ce qu'il en a compris... nous tentons de distinguer assez de niveaux pour éviter qu'un codage trop strict n'empêche la représentation d'un phénomène scientifiquement intéressant, tout en conservant la possibilité d'effectuer des recherches dans la base sans être trop pénalisé par sa variété, à travers des informations ajoutées dans la base de donnée à propos des signes.

## 6. Conclusion

Le problème du codage présente plusieurs facettes. Il s'agit d'une part d'un acte technique, dont le but est d'obtenir une certaine portabilité des documents électroniques. Dans cette optique, le codage d'un noyau dur d'hiéroglyphes dans ISO 10646 est légitime, et peut se révéler d'une certaine utilité. Pour que ce travail-là soit fructueux, il faut dans un premier temps que sa portée soit limitée ; une légère extension de la liste de Gardiner semble l'approche la plus simple.

En revanche, pour l'usage documentaire en égyptologie, l'intérêt d'un codage unifié tel qu'ISO 10646/Unicode est limité, puisque la conception de ce qu'est un caractère va quasiment varier d'un auteur à l'autre. Dans ce domaine, il est nécessaire de se donner du temps.

Après avoir suscité beaucoup d'intérêt dans les années 1980, le problème du codage scientifique des hiéroglyphes a été malheureusement délaissé pour plusieurs raisons. La première est que la plupart des parties en présence étaient essentiellement désireuses de disposer d'un traitement de texte ; la saisie de textes codés leur apparaissait comme un produit dérivé intéressant, mais pas comme une fin. D'autre part, coder un texte prend énormément de temps, et est peu « rentable » d'un point de vue universitaire.

L'apparition de nouvelles technologies comme XML et le WEB, et surtout leur adoption par la communauté des sciences humaines, est à la source de nouveaux travaux. Comme un travail vraiment abouti sur le codage de l'égyptien ancien ne peut se concevoir sans la collaboration de spécialistes des divers domaines de l'égyptologie,

afin de prendre en compte les besoins, et non simplement de coder les fontes existantes, il est nécessaire de se donner du temps, afin que les personnes concernées se sentent impliquées et participent.

## 7. Bibliographie

- [BUD 20] Budge E. A., *An Egyptian Hieroglyphic Dictionary*, vol. I, Dover, 1920.
- [BUU 76] Buurman J., Schimmelpenninck van der Oije A. H. J., « The Composing of Hieroglyphic Texts by Means of a Computer », *Göttinger Miszellen*, vol. 19, 1976, p. 7-15.
- [BUU 85] Buurman J., « Printing of Egyptian hieroglyphs by means of a computer », *Informatique et Egyptologie*, vol. 1, 1985, p. 13-32.
- [BUU 88] Buurman J., Grimal N., Hainsworth M., Hallof J., Plas D. V. D., *Inventaire des signes hiéroglyphiques en vue de leur saisie informatique*, Mémoires de l'Académie des Inscriptions et Belles Lettres, Institut de France, Paris, 1988.
- [CAU 97] Cauville S., *Dendara, les chapelles osiriennes*, Institut Français d'archéologie orientale, 1997.
- [CAU 01] Cauville S., *Dendara, le fonds hiéroglyphique au temps de Cléopâtre*, Cybèle, 2001.
- [ERM 63] Erman A., Grapow H., Eds., *Wörterbuch der aegyptischen Sprache*, Akademie Verlag, 1926-1963.
- [EVE 99] Everson M., « Encoding Egyptian Hieroglyphs in Plane 1 of the UCS », <http://www.dkuug.dk/jtc1/sc2/wg2/doc/n1944.pdf>, 1999.
- [GRI 85] Grimal N., « Remarques sur les méthodes de classement d'une fonte hiéroglyphique automatisée », *Informatique et Egyptologie*, vol. 1, 1985, p. 57-63.
- [HAI 79] Hainsworth M., « Traitement automatique des textes en hiéroglyphique égyptien », *L'égyptologie en 1979, axes prioritaires de recherche*, CNRS, 1979.
- [HAN 95] Hannig R., *Großes Handwörterbuch Ägyptisch-Deutsch*, Philipp von Zabern, 1995.
- [JAN 72] Janssen J. M. A., « Les listes de signes hiéroglyphiques », Sauneron S., Ed., *Textes et langages de l'Égypte pharaonique*, p. 57-66, Institut français d'archéologie orientale, 1972, reprise d'un article publié dans *Chronique d'égypte*, vol. XXVII/53, 1952, p. 88-97. Historique complet des fontes hiéroglyphiques.
- [MAN 94] der Manuelian P., *Living in the Past*, Kegan Paul International, 1994.
- [PLA 93] van der Plas D., Vergniew R., « Computer-aided Research on Egyptian Stelae from the Middle Kingdom », *Informatique et Egyptologie*, vol. 9, 1993, p. 139-151.
- [SAU 88] Sauneron S., *Les prêtres de l'ancienne Égypte*, Perséa, 1988, publié d'abord aux éditions du Seuil en 1957.
- [SCH 74] Schenkel W., « Gesichtspunkte für die Neugestaltung der Hieroglyphenliste », *Göttinger Miszellen*, vol. 14, 1974, p. 31-45.
- [SCH 99] Schenkel W., « Comments on the question of encoding Egyptian Hieroglyphs in the UCS », <http://www.dkuug.dk/jtc1/sc2/wg2/docs/n2096.pdf>, 1999.