

Une translittération arabe/roman pour un e-document

Mostafa BANOUNI¹, Azzeddine LAZREK² et Khalid SAMI³

¹ *Université Ibn Zohr, Faculté des Sciences
Département de Physique, B.P. S/20, Agadir, Maroc
Tél. : +212 48 22 01 09 Fax : +212 48 22 01 00*

`mbanouni@voila.fr`

² *Université Cadi Ayyad, Faculté des Sciences
Département d'Informatique, B.P. 2390, Marrakech, Maroc
Tél. : +212 44 43 46 49 Fax : +212 44 43 74 09*

`lazrek@ucam.ac.ma`

³ *Université Cadi Ayyad, Faculté des Sciences
Département de Mathématiques, B.P. 2390, Marrakech, Maroc
Tél. : +212 44 43 46 49 Fax : +212 44 43 74 09*

`k_sami@ucam.ac.ma`

Résumé : La préparation du document arabe dans un environnement roman, dans le but d'un traitement ou tout simplement dans le but de rendre la forme phonétique ou orthographique du document, fait appel à une translittération arabe/roman. Plusieurs tables de translittération ont été, et sont, proposées. Ces translittérations ne constituent pas une correspondance biunivoque de caractère arabe à caractère roman. De plus, elles utilisent certains signes qui ont soit la même signification, aussi bien pour un document arabe que pour un document roman, soit une interprétation particulière pour certains pré-processeurs. Par ailleurs, ces translittérations ne prennent pas en compte les différentes formes que peuvent prendre certains caractères utilisés en tant que symboles littéraux dans une expression mathématique arabe. Dans cette

contribution, nous présentons une translittération qui tient compte des questions sus-mentionnées.

Abstract: The preparation of Arabic documents in Roman script based languages, for the purpose of processing or producing the phonetic form, needs an Arabic/Roman transliteration. Many tables of transliterations have been or are in use. How to make a choice in the special situation where the document consists of mathematics written in a fully Arabic script, with specific symbols? which transliteration to use for texts to be treated by information processing systems? In this contribution, we present proposal for a transliteration which takes account of the above-mentioned issues.

Mots-clés : Traitement de document, e-document, Translittération, Encodage, Transcription, ASCII, Unicode.

1. Introduction

Une translittération arabe/roman, une romanisation, est une application bijective entre l'ensemble des éléments de l'alphabet arabe (lettres, signes diacritiques, chiffres, signes de ponctuation, etc.) et un ensemble de combinaisons d'éléments de l'alphabet roman. La table de conversion qui représente la translittération est destinée à être utilisée pour la *transcription* de mots, initialement *écrits en arabe*, dans un système d'écriture appartenant à un environnement qui, pour une raison quelconque, requiert l'usage de l'alphabet roman comme dans les bibliographiques, les notices, les documents officiels, etc. Il est à remarquer, en particulier, que la table de translittération n'est pas destinée à transcrire la représentation phonétique des mots de manière directe. C'est la forme *écrite* à l'aide de l'alphabet arabe, et non sa prononciation, qui constitue la donnée d'entrée.

Il y a donc lieu de distinguer entre divers concepts qui peuvent prêter à confusion alors qu'ils ne sont pas synonymes, à savoir : transcription, translittération et encodage. On parle de transcription lorsqu'il s'agit de procédés destinés à enregistrer de manière scripturale la parole (partant de la représentation phonétique des mots et non de leur représentation scripturale dans le système d'écriture). Elle peut être représentée par le même alphabet ou par l'alphabet phonétique international. (ex. le mot français *éléphant* admet comme transcription $e\ell eF\tilde{a}$).

L'encodage est la transcription d'un texte, selon les règles d'un code, en vue de sa transmission ou de son traitement (ex. le code ASCII de la lettre e est 65 en hexadécimale).

Effectivement, contrairement à la transcription, le texte d'entrée en translittération est d'abord transcrit en écriture arabe, alors que la transcription se base sur la forme phonétique du discours. La transcription d'un discours permet ainsi au lecteur de lire un texte même s'il ne connaît pas l'écriture arabe. Par contre, la translittération peut être utilisée pour faire respecter une prononciation plus précise et offrir l'opportunité de retourner au texte original [Bee98], [BS53] et [Wel75].

Une translittération peut reposer sur la correspondance phonologique ou morphologique, d'une part, et sur la correspondance orthographique, d'une autre part.

Étant donné que la prononciation de certains caractères de l'alphabet roman varie selon la langue du contexte, il est naturel qu'il y ait plusieurs translittérations. La lettre *j* (jota espagnol) serait la plus indiquée pour translittérer la lettre arabe KHAH¹ dans un environnement où l'espagnol castillan est la langue de référence. La même lettre arabe serait représentée plutôt par le doublon *kh* là où le français règne et où la lettre *j* serait plutôt utilisée pour représenter la lettre arabe JEEM. L'administration marocaine, par exemple, translittère la lettre KHAH avec *kh*. La translittération dépend donc directement de la langue de référence du contexte de préparation du texte d'entrée.

2. Classification

Les translittérations sont bien plus anciennes que les ordinateurs. Avec l'arrivée du codage standard ASCII, la plupart de ces translittérations ont dû être revues.

Quand la translittération d'un caractère donné est composée de deux signes, ces signes peuvent être combinés de deux façons différentes : verticalement ou horizontalement, c'est-à-dire, en superposition ou en succession. Ainsi, une proposition de translittération de DAD par *ḍ* ou par *.ḍ* peut être une solution pour un document qui n'est pas destiné à être traité électroniquement mais ce peut être une mauvaise translittération dans le cas contraire. La contrainte d'utiliser le code ASCII impose de revoir tout système de translittération jadis adopté pour l'écriture à la main. En effet, le système ASCII traite le symbole *ḍ*, *ḍ* avec un point en dessous, comme une concaténation de deux symboles et non comme un symbole unique comme le ferait l'utilisateur qui écrit ou celui qui lit un texte après avoir consulté la table de translittération. On ne peut pas utiliser directement ce symbole pour dénoter la lettre arabe DAD. L'idée la plus simple alors est de décomposer le symbole en une succession de signes ASCII : dans notre exemple, ceci revient à remplacer le symbole *ḍ* par *.ḍ*. Cela peut constituer une alternative acceptable pour un texte qui n'est pas destiné à subir un traitement par la suite, tel un message électronique. Un traitement du texte entraînerait beaucoup d'autres modifications. Par exemple, dans l'exemple précédent :

- le signe point, déjà utilisé pour translittérer le caractère arabe point, serait dévoyé de son utilisation originale comme signe de ponctuation ;
- le point fait appel automatiquement à l'application de certaines règles typographiques comme la gestion des espaces avant et après le point ;
- l'espace mémoire requis est plus important.

Bien que le standard propriétaire Unicode, ou la norme ISO/CEI 10646, d'encodage universelle soit en train de s'imposer peu à peu, il reste beaucoup de contextes

¹Les noms des caractères arabes sont ceux utilisés dans Unicode².

²<http://www.unicode.org>

dans lesquels, pour la transmission et le traitement de documents arabes, l'utilisation du simple et traditionnel code ASCII standard reste une nécessité pratique.

L'alphabet roman, même avec l'utilisation des lettres majuscules, n'offre pas assez de signes pour la translittération de tous les éléments utilisés dans le système alphabétique arabe. Quelques consonnes arabes ne correspondent à aucun élément phonétique des langues romanes. La lettre AIN, par exemple, a été translittérée de plusieurs façons (ex. *, ` , e , E , a , A , . . .).

Une translittération peut être caractérisée par les propriétés suivantes :

non-ambiguïté ou bi-univocité : l'application entre les éléments des deux alphabets est une bijection ;

complétude : la translittération couvre tous les éléments du système alphabétique arabe ;

compacité : le nombre de signes utilisés pour la translittération d'un élément est optimal ;

portabilité : la translittération n'utilise que les caractères ASCII standard. La translittération peut être saisie à partir d'un clavier standard. Le texte saisi est transportable. Une translittération ASCII permet de conserver la portabilité du document, ainsi saisi ou transcrit, sur tout système informatique ;

lisibilité humaine : le texte est lisible, il peut être facilement lu par un être humain. La translittération doit utiliser uniquement les lettres de l'alphabet cible pour ne pas s'éloigner trop de la forme phonétique d'origine ;

affichabilité : le texte résultat peut être affiché sur un écran ASCII ou à l'aide d'une imprimante standard ;

mnémotechnicité : la translittération est facilement mémorisable.

Généralement, ces propriétés présentent quelques antagonismes. Il faut chercher un compromis.

Les différentes translittérations en usage peuvent être classées en deux catégories :

translittérations scripturales : les translittérations traditionnellement utilisées pour l'écriture manuelle ;

translittérations électroniques : les translittérations utilisées pour des textes susceptibles d'être traités par des systèmes informatiques, ne comprenant que des caractères ASCII.

2.1. Translittérations scripturales

Ci suit, une liste non exhaustive de quelques translittérations scripturales en usage :

Orientalist Scholars : la convention internationale de l'Orientalist Scholars adopte une proposition de translittération faite à Rome (1936).

EI (Encyclopedia of Islam) : utilise une translittération basée sur les conventions normatives anglaises, les diacritiques et les consonnes voyellisées (1913-1960) (Cf. TAB. 5).

BS 4280 (British Standard 4280)³ : propose une translittération consistante mais elle n'a pas été largement utilisée (1969-1983).

Beyrouthamendé : a été adoptée par un grand nombre de pays arabes (1972).

ISO 233 (ou 46-002)⁴ : la norme ISO de translittération arabe (1984).

ISO 233-2 : la norme ISO de translittération simplifiée Partie 2 (1993).

ISO/FDIS 233-3 : la norme ISO de translittération persane Partie 3 (1999).

ALESCO (Arab League Educational, Scientific and Cultural Organization) : la translittération utilisée par l'ALESCO [Org94] (Cf. TAB. 5).

SAWS (Scientific Arabic Writing Systems) : la translittération adoptée lors d'une conférence sur les systèmes de l'écriture arabe scientifique (1995) [Org94] (Cf. TAB. 5).

2.2. Translittérations électroniques

Ci suit, une liste non exhaustive des translittérations électroniques en utilisation. Toutes ces translittérations sont présentées dans TAB. 6 :

Qalam⁵ : la translittération morphologique développée par A. Heddaya en contribution avec W. Hamdy et M. H. Sherif, (1985-1992).

ditroff/ffortid : la translittération utilisée dans le système de traitement de texte ditroff/ffortid développé par J. Srouji et D. Berry [SB92].

ArabTeX : la translittération utilisée dans l'extension, du système de traitement de document TeX, ArabTeX développée par K. Lagally [Lag92]. Cette translittération utilise certains symboles tels que ^, _ et . ayant une signification particulière en mode mathématique en TeX.

³<http://www.edesign.demon.co.uk/translit.htm>

⁴<http://www.iso.org>

⁵<http://eserver.org/langs/qalam.txt>

Ω : la translittération utilisée dans le système de traitement de document multilingue Ω développé par Y. Haralambous et J. Plaice [HP97]. Le système Ω offre à l'utilisateur la possibilité de définir sa propre translittération.

Buckwalter⁶ : la translittération de l'orthographe de l'arabe standard moderne développée par le lexicographe T. Buckwalter. Cette translittération est une véritable transcription orthographique.

ARABVISL (ARABic Visual Interactive Syntax Learning) : la translittération utilisée dans le projet de recherche VISL⁷ de l'Institut de Langage et de Communication de l'Université du Danemark (1996).

3. La translittération **TransTec**

La translittération **TransTec**, que nous proposons ici, repose sur deux analyses : une analyse phonétique et une analyse morphologique. Le but principal est d'utiliser la même translittération en mode mathématique et en mode texte dans un système de composition du document mathématique arabe dont les expressions symboliques arabes sont composées à l'aide de symboles arabes spécifiques et dont l'écriture se déroule de la droite vers la gauche. Dans la translittération proposée, les règles suivantes sont appliquées (Cf. TAB. 7) :

- conserver l'utilisation des lettres minuscules ASCII qui sont habituellement utilisées dans les translittérations ;
- utiliser les lettres majuscules, qui sont des signes supplémentaires, pour translittérer toutes les autres lettres alphabétiques. Le trait majuscule n'est plus utilisé pour son caractère typographique (les noms propres, le début d'une phrase, etc.) ;
- limiter, le plus que possible, l'utilisation aux lettres romanes du code ASCII ;
- minimiser le nombre de lettres utilisées.

3.1. *Lettres*

L'analyse phonétique peut engendrer les quatre classes de lettres suivantes, selon que leurs correspondants romans sont communément d'usage ou d'usage restreint, ou bien qu'elles s'approchent phonétiquement de certaines lettres courantes ou non. Une autre classe de lettres est déterminée par une correspondance morphologique.

3.1.1. *Phonétique communément d'usage*

Nous entendons par *communément d'usage* que la consonne est rencontrée dans la plupart des langues indo-européennes. Les lettres ALEF, BEH, TEH, JEEM, DAL,

⁶<http://www.xrce.xerox.com/research/mltt/arabic>

⁷<http://visl.hum.sdu.dk/visl/ar>

REH, ZAIN, SEEN, FEH, KAF, LAM, MEEM, NOON, HEH sont translittérées par les lettres a, b, t, j, d, r, z, s, f, k, l, m, n, h respectivement. L'équivalent roman est alors écrit en minuscule (Cf. TAB. 7).

3.1.2. Phonétique d'usage restreint

Les correspondants romans sont écrits en majuscules. Ils sont illustrés par des exemples de mots de différentes langues indo-européennes (Cf. TAB. 1).

Lettre	Nom	Translittération	Remarques
خ	KHAH	X	Xavier (espagnol), lettre χ (grec)
ش	SHEEN	C	Cinco (italien)
ص	SAD	S	Saut (français)
ض	DAD	D	Dalle (français)
ط	TAH	T	Taux (français)

Table 1: Phonétique d'usage restreint

3.1.3. Phonétique approchée

Dans cette classe, on trouve les lettres THEH, THAL et ZAH (Cf. TAB. 7). La lettre THEH, qui se prononce comme le "th" du mot anglais "tooth", peut être prononcée par le profane comme un F. La lettre THAL qui se prononce comme le "th" du mot anglais "mother" est représenté par un Z car elle est ainsi prononcée dans certains dialectes arabes orientaux. La lettre ZAH a une prononciation semblable à celle du THAL mais plus grave. Si l'on ne prête pas une grande attention lorsqu'elle est prononcée, on peut l'entendre comme un V grave.

3.1.4. Phonétique rare

Cette classe regroupe les lettres de phonétique rare ou inexistante dans les langues indo-européennes. Ce sont les gutturales HAH, AIN, GHAIN et QAF (Cf. TAB. 2). La correspondance du QAF avec une lettre minuscule q nous écarte de la règle mentionnée plus haut car la forme majuscule Q est réservée pour une autre fin.

Lettre	Nom	Translittération
ح	HAH	H
ع	AIN	E
غ	GHAIN	G
ق	QAF	q

Table 2: Phonétique rare

3.1.5. *Correspondance morphologique*

Dans cette classe, on trouve la lettre TEH MARBUTA qui possède une double prononciation selon que l'arrêt de la parole a lieu à son niveau ou pas. En fait, elle est prononcée comme un HEH en cas d'absence de voyelles dans une phrase ; elle est prononcée comme le TEH autrement. Devant cette ambiguïté et le manque de caractères romans disponibles s'approchant de ces deux prononciations, nous avons opté pour une correspondance morphologique : le squelette de TEH MARBUTA, dépourvue de points diacritiques, dans ses deux formes : isolée (ٲ) ou finale (ٲ), ressemble schématiquement au caractère Q. Ce type de correspondance peut s'appliquer également pour consolider la correspondance de la lettre AIN au caractère E.

3.2. *Voyelles*

Les voyelles sont notées avec des signes diacritiques qui permettent de savoir comment est prononcé un caractère dans un mot donné. Ces voyelles sont les suivantes :

- voyelles longues (ALEF, WAW, YEH, ALEF MAKSURA)
- voyelles courtes (FATHA, DAMMA, KASRA)
- tanwyn (FATHATAN, DAMMATAN, KASRATAN)
- SHADDA
- SUKUN

Les trois voyelles longues, ALEF, WAW et YEH, dites encore *horoof al-mad*, sont représentées respectivement par a, w et y. Le ALEF MAKSURA qui conduit au même effet que la lettre a est représenté par la lettre majuscule Y car il a un trait morphologique commun avec le YEH qui, lui, est représenté par un y minuscule.

Les voyelles courtes FATHA, DAMMA et KASRA sont représentées respectivement par e, u et i. Le tanwyn associé à ces trois voyelles est assuré par l'ajout d'un N majuscule à leur suite, pour une correspondance phonologique, ou leur dédoublement, pour une correspondance morphologique. Ces correspondances utilisent deux caractères, ce qui est utilisé dans certains styles calligraphiques arabes.

Pour la SHADDA et le SUKUN, nous avons adopté une correspondance morphologique. Ainsi, on a associé à la SHADDA un w majuscule (le w minuscule est utilisé pour désigner le WAW) et au SUKUN un o minuscule (Cf. TAB. 3).

3.3. *HAMZA, MADDA et WASLA*

L'analyse contextuelle est utilisée pour déterminer la forme des lettres (initiale, médiane, finale et isolée) (ex. م , م , م , م). Elle peut être aussi utilisée pour déterminer la présence de ligatures (ex. LAMALEF ل). Mais elle ne peut pas être utilisée pour déterminer quand est ce que le signe HAMZA combiné avec une voyelle longue, est au-dessous ou au-dessus du ALEF ou sur la ligne. En effet, cela nécessite que le texte

Lettre	Nom	Translittération
و	WAW	w
ي	YEH	y
آ	ALEF MAKSURA	Y
ـ	FATHA	e
ـ	DAMMA	u
ـ	KASRA	i
ء	FATHATAN	eN ou ee
ء	DAMMATAN	uN ou uu
ء	KASRATAN	iN ou ii
ّ	SHADDA	W
◌	SUKUN	o

Table 3: Voyelles

soit voyellisé. De plus, il n'y a pas de règles orthographiques strictes et les quelques règles en la matière varient entre le Machrek et le Maghreb arabe.

Cette correspondance utilise deux caractères au lieu d'un seul. En fait, les écritures HAMZA ON ALEF, HAMZA UNDER ALEF, HAMZA ON WAW et HAMZA ON YEH correspondent à la superposition de deux caractères : le caractère HAMZA et l'un des caractères ALEF, WAW ou YEH.

La MADDA ON ALEF est associée à la combinaison de lettres Ma. La WASLA ON ALEF est désignée par La (Cf. TAB. 4). Les lettres capitales utilisées ici ne servent que de marqueurs et n'intervient pas pour la prononciation, ce qui peut être contre-intuitif pour un lecteur habitué à l'alphabet roman.

Lettre	Nom	Translittération
ء	HAMZA ON LINE	A-
أ	HAMZA ON ALEF	Aa
إ	HAMZA UNDER ALEF	AY
أ	HAMZA ON WAW	Aw
أ	HAMZA ON YEH	Ay
آ	MADDA ON ALEF	Ma
أ	WASLA ON ALEF	La

Table 4: HAMZA, MADDA et WASLA

3.4. KASHIDA

La KASHIDA ou TATWEEL est l'allongement de certaines lettres d'une façon curviligne. La KASHIDA n'est pas une lettre en elle-même. Elle est utilisée pour

justifier un texte et en calligraphie. La KASHIDA n'est pas utilisée dans les expressions mathématiques, on peut donc utiliser un symbole pour la représenter. On a choisi le signe - pour la désigner, mais on peut aussi utiliser la lettre K pour la désigner ou même aucune lettre puisqu'elle n'est pas prononçable.

3.5. Symboles

La translittération des symboles dépend de leurs types. La translittération des signes de ponctuation arabes . : ; ! ? est , . ; ! ? respectivement. La translittération des chiffres arabo-indiens, du système de numération décimal, est donnée par les chiffres arabes tout en faisant la distinction de la forme des chiffres utilisée au Maghreb Arabe 0 1 2 3 4 5 6 7 8 9 de celle utilisée au Machreq Arabe . ١ ٢ ٣ ٤ ٥ ٦ ٧ ٨ ٩. La translittération des délimitants (ex. les parenthèses (et)) est identique. Cependant, il y a un problème d'appellation et de direction des délimitants. En effet, la parenthèse qui se trouve à gauche (respectivement à droite) d'une expression est la parenthèse ouvrante (respectivement fermante) en roman mais c'est la parenthèse fermante (respectivement ouvrante) en arabe.

Rappelons-nous qu'en mode mathématique arabe :

- il y a moins de lettres arabes, considérées comme des symboles littéraux de désignation. En effet, les lettres ne portent pas de point diacritique et elles ne sont pas voyellées. Par exemple, les trois lettres BEH, TEH et THEH (ب ت ث) donnent le seul caractère (ب) prononcé BEH ;
- l'écriture symbolique n'est pas cursive. Par exemple, l'angle ALEF, BEH et JEEM se note (أ ب ج ou ا ب ج). La forme initiale et isolée des lettres doit être translittérée différemment (أ et ا). Cela permet d'élargir la liste des symboles de désignation ;
- les lettres ne suivent pas le même ordre que celui de l'alphabet, par exemple, (ج) pour JEEM et non HAH mais (ز) pour REH et non ZAIN.

La translittération proposée pour les symboles littéraux arabes dans une expression mathématique (Cf. TAB. 7) repose sur :

- l'utilisation des lettres minuscules pour la forme finale, qui est la forme la plus utilisée en mathématique arabe ;
- l'utilisation des lettres majuscules pour la forme initiale, si son glyphe présente une distinction par rapport à celui de la forme finale.

$\$ \amrl \{ \{ \amfrac \{ \amleft \{ \amssum \limits_{b=1}^T \} \amcos c_{_b} \amright \} + \{ \amsqrt \{ 2c \} \} \{ 3T-4 \} + 9 \} \$$

$$9 + \frac{\sqrt{2} + \left(\sum_{i=1}^T \text{جتا} \right)}{4-3}$$

Lorsque les commandes sont autorisées dans une translittération, cela offre une nouvelle dimension pour translittérer les lettres. Cette approche suppose un traitement du texte saisi. Par exemple, dans le système \TeX , l'alphabet grec est saisi avec des commandes, (ex. la commande \gamma produit la lettre γ et la commande \Gamma produit la lettre Γ). Cette approche est utilisée pour translittérer la famille de caractères avec une queue ou avec leur contour (ex. \jeem pour \char"203C , \Jeem pour \char"203D , \JEEM pour \char"203E , \jjeem pour \char"203F , \JJJeem pour \char"2040 et \JJJEEM pour \char"2041) de la fonte **NasX** utilisée dans le système **RyDArab**⁸ [Laz01a], [Laz01b] (Cf. TAB. 8).

4. Conclusion

La nécessité d'adopter une translittération s'impose à l'utilisateur, des nouvelles techniques de communication, qui ne dispose pas d'un environnement arabe ou entièrement arabisé. Cette nécessité est également rencontrée dans l'échange de l'information entre des systèmes hétérogènes. Plusieurs propositions de translittérations ont été avancées afin de permettre la saisie du texte arabe en l'absence d'un environnement informatique adéquat. Ces translittérations associent à certains caractères arabes un ou deux caractères romans ou signes. L'utilisation de signes non alphabétiques est, à notre point de vue, préjudiciable pour les deux raisons suivantes :

- on pénalise ces signes qui sont utilisés soit dans l'écriture du texte arabe moyennant quelques transformations éventuelles du glyphe, soit en tant que caractères de contrôle pour certains pré-processeurs ;
- la lecture et la saisie du texte, composé à l'aide de ces signes, ne sont pas faciles.

Nous avons proposé une translittération caractérisée par l'utilisation quasi exclusive des lettres romanes sous leurs deux traits d'écriture majuscule et minuscule. Nous avons, pour cela, procédé à deux approches : phonétique et morphologique. La translittération proposée préserve au maximum la correspondance en usage dans la plupart des translittérations existantes. La translittération **TransTec** est : non ambiguë, complète, compacte (puisqu'elle utilise une correspondance lettre à lettre sauf pour les lettres ou les signes composés), portable (ASCII standard) mnémotechnique et relativement lisible.

Nous ne prétendons pas avoir proposé ici une solution définitive à la problématique de translittération. L'utilisateur d'un système informatique préfère souvent avoir la liberté de choisir, ou même de définir, la translittération qui lui semble la plus convenable. Les systèmes de traitement de l'information doivent en conséquence tenir compte de cette diversification à l'aide de convertisseurs adéquats. L'unification d'une

⁸L'extension **RyDArab**⁹ est un système de traitement de document mathématique arabe que nous avons développée à l'Université Cadi Ayyad de Marrakech. Il permet la composition d'expressions mathématiques composées de symboles spécifiques dont l'écriture se déroule de la droite vers la gauche.

⁹<http://www.ucam.ac.ma/fssm/rydarab>

table de translittération reste cependant un objectif incontournable pour faciliter la communication entre utilisateurs.

Bibliographie

- [Bee98] Kenneth R. Beesley. Xerox Arabic Morphological Analysis and Generation, Romanisation, Transcription and Transliteration. *The Document Company - Xerox*, 1998.
- [BS53] R. Blachère and J. Sauvaget. *Règles pour éditions et traductions de textes arabes*. Les Belles Lettres, Paris, 1953.
- [HP97] Yannis Haralambous and John Plaice. Multilingual Typesetting with Ω a Case Study: Arabic. In *Proceedings of the International Symposium on Multilingual Information Processing*, pages 137–154, 1997.
- [Lag92] Klaus Lagally. Arab \TeX - Typesetting Arabic with Vowels and Ligatures. *Euro \TeX '92, Prague*, 1992.
- [Laz01a] Azzeddine Lazrek. A package for typesetting Arabic mathematical formulas. *Die \TeX nische Komödie, DANTE e.V.*, 13(2/2001):54–66, 2001.
- [Laz01b] Azzeddine Lazrek. Aspects de la problématique de la confection d'une fonte pour les mathématiques arabes. *Cahiers GUTenberg*, 39–40, Le document au XXI^e siècle:51–62, 2001.
- [SB92] Johny Srouji and Daniel Berry. Arabic formatting with ditroff/ffortid. *EPODD*, 5(4):163–208, 1992.
- [Wel75] Hans Wellisch. *Transcription and Transliteration: an annotated bibliography on conversion of scripts*. Silver Spring, Maryland: Institute of Modern Languages, 1975.
- [Org94] المنظمة العربية للتربية والثقافة والعلوم، مكتب تنسيق التعريب، أعمال مؤتمر التعريب السابع لإقرار مشاريع المعاجم ونظم الكتابة العربية العلمية، الخرطوم، السودان، 1994.
(Organisation arabe de l'éducation, la culture et les sciences, Bureau de coordination de l'arabisation, *Les actes du 7^e congrès d'arabisation pour approuver les projets de lexiques et les systèmes de l'écriture arabe scientifique*, Khartoum, Soudan, 1994.)

Une translittération arabe/roman pour un e-document 135

Lettre	Nom	EI	ALESCO	SAWS
ا	ALEF	a	a	a
ب	BEH	b	b	b
ت	TEH	t	t	t
ث	THEH	<u>th</u>	<u>t</u>	<u>th</u>
ج	JEEM	<u>dj</u>	j	j
ح	HAH	<u>h</u>	<u>h</u>	<u>h</u>
خ	KHAH	<u>kh</u>	<u>k</u>	<u>kh</u>
د	DAL	d	d	d
ذ	THAL	<u>dh</u>	<u>d</u>	<u>dh</u>
ر	REH	r	r	r
ز	ZAIN	z	z	z
س	SEEN	s	s	s
ش	SHEEN	<u>sh</u>	<u>s</u>	<u>sh</u>
ص	SAD	<u>s</u>	<u>s</u>	<u>s</u>
ض	DAD	<u>d</u>	<u>d</u>	<u>d</u>
ط	TAH	<u>t</u>	<u>t</u>	<u>t</u>
ظ	ZAH	z	Ḍ	z
ع	AIN	.	.	.
غ	GHAIN	<u>gh</u>	<u>g</u>	<u>gh</u>
ف	FEH	f	f	f
ق	QAF	<u>q</u>	q	q
ك	KAF	k	k	k
ل	LAM	l	l	l
م	MEEM	m	m	m
ن	NOON	n	n	n
ه	HEH	h	h	h
ة	TEH MARBUTA		h ou t	t
و	WAW	w	w	w
ي	YEH	y	y	y
ى	ALEF MAKSURA			
ـ	FATHA	a	a	a
ـ	DAMMA	u	u	u
ـ	KASRA	i	i	i
ـ	FATHATAN	an	an	(an)
ـ	DAMMATAN	un	un	(un)
ـ	KASRATAN	in	in	(in)
ـ	SHADDA	¹	¹	¹
ـ	SUKUN	²	²	²
ـ	HAMZA ON LINE	.	.	.
أ	HAMZA ON ALEF	a	a	a ou 'a
إ	HAMZA UNDER ALEF	i	i	i ou 'i
أ	HAMZA ON WAW	u	u	u ou 'u
أ	HAMZA ON YEH			'i
أ	MADDA ON ALEF	ā	ā	ā
أ	WASLA ON ALEF			
-	KASHIDA			

Table 5: Quelques translittérations scripturales

¹La voyelle SHADDA d'une lettre est indiquée par le dédoublement de la lettre en question.

²La voyelle SUKUN est omise.

Lettre	Nom	QALAM	ditroff/ffortid	ArabTeX	Ω	Buckwalter	ARABVISL
ا	ALEF	aa	a	a ou A ou _a	A	A	aa
ب	BEH	b	b	b	b	b	b
ت	TEH	t	t	t	t	t	t
ث	THEH	th	c	.t	th	v	th
ج	JEEM	j	j	ˆg	j	j	j
ح	HAH	H	h	.h	H	H	H
خ	KHAH	kh	x	.h	kh	x	kh
د	DAL	d	d	d	d	d	d
ذ	THAL	dh	Z	.d	dh	*	dh
ر	REH	r	r	r	r	r	r
ز	ZAIN	z	z	z	z	z	z
س	SEEN	s	s	s	s	s	s
ش	SHEEN	sh	C	ˆs	sh	\$	sh
ص	SAD	S	S	.s	S	S	S
ض	DAD	D	D	.d	D	D	D
ط	TAH	T	T	.t	T	T	T
ظ	ZAH	Z	Z'	.z	Z	Z	Z
ع	AIN	'	e	'	'	E	*
غ	GHAIN	gh	R'	.g	gh	g	gh
ف	FEH	f	f	f	f	f	f
ق	QAF	q	q	q	q	q	q
ك	KAF	k	k	k	k	k	k
ل	LAM	l	l	l	l	l	l
م	MEEM	m	m	m	m	m	m
ن	NOON	n	n	n	n	n	n
ه	HEH	h	H	h	-h	h	h
ة	TEH MARBUTA	t ou h	t'	T	"h ou "t	p	-a
و	WAW	w	w	w ou U ou _u	w ou U	w	w ou uu
ي	YEH	y	y	y ou I ou _i	y	y	y ou ii
ى	ALEF MAKSURA	ae	A'	Y ou _A	I	Y	aa
ـ	FATHA	a	'	a	a	a	a
ـ	DAMMA	u	'	u	u	u	u
ـ	KASRA	i	E	i	i	i	i
ـ	FATHATAN	aN		aN	aN	F	an
ـ	DAMMATAN	uN		uN	uN	N	un
ـ	KASRATAN	iN		iN	iN	K	in
ـ	SHADDA	l	˜˜	l		˜	l
ـ	SUKUN	-	O	2	<>	o	2
ء	HAMZA ON LINE	'	'	'	—	'	'
أ	HAMZA ON ALEF		A	'a	'a	>	a
إ	HAMZA UNDER ALEF		i	'i	'i	<	i
ؤ	HAMZA ON WAW		w'	'w	'u	&	
ي	HAMZA ON YEH		Y'	'y	'y	}	
آ	MADDA ON ALEF	˜aa	a˜	'A	'A	—	˜aa
أ	WASLA ON ALEF	e	U		"A	{	
ـ	KASHIDA			-	-	-	-

Table 6: Quelques translittérations électroniques

¹La voyelle SHADDA d'une lettre est indiquée par le dédoublement de la lettre en question.²La voyelle SUKUN est omise.

Une translittération arabe/roman pour un e-document 137

Lettre	Nom	Texte	Expression mathématique					
ا	ALEF	a	ا	a				
ب	BEH	b	ب	b	ر	B	به	\BEH
ت	TEH	t						
ث	THEH	F						
ج	JEEM	j	ح	j	ح	J	جه	\JEEM
ح	HAH	H						
خ	KHAH	X						
د	DAL	d	د	d				
ذ	THAL	Z						
ر	REH	r	ر	r				
ز	ZAIN	z						
س	SEEN	s						
ش	SHEEN	C	س	c	س	C	سه	\SHEEN
ص	SAD	S	ص	s	ص	S	صه	\SAD
ض	DAD	D						
ط	TAH	T	ط	T			طه	\TAH
ظ	ZAH	V						
ع	AIN	E	ع	e	ع	E	عه	\AIN
غ	GHAIN	G						
ف	FEH	f	ف	f	و	F	فه	\FEH
ق	QAF	q	ق	q				
ك	KAF	k	ك	k			كه	\KAF
ل	LAM	l	ل	l	ل	L	له	\LAM
م	MEEM	m	م	m	م	M	مه	\MEEM
ن	NOON	n	ن	n	ن	N		
ه	HEH	h			ه	H	هه	\HEH
ة	TEH MARBUTA	Q						
و	WAW	w	و	w				
ي	YEH	y	ي	y	ي	Y	يه	\YEH
ى	ALEF MAKSURA	Y						
ـ	FATHA	e						
ـ	DAMMA	u						
ـ	KASRA	i						
ء	FATHATAN	eN ou ee						
ء	DAMMATAN	uN ou uu						
ء	KASRATAN	iN ou ii						
ـ	SHADDA	W						
ـ	SUKUN	o						
ء	HAMZA ON LINE	A-	ء	A				
أ	HAMZA ON ALEF	Aa						
إ	HAMZA UNDER ALEF	AY						
ؤ	HAMZA ON WAW	Aw						
ي	HAMZA ON YEH	Ay						
آ	MADDA ON ALEF	Ma						
ء	WASLA ON ALEF	La						
-	KASHIDA	K ou -						

Table 7: La translittération **TransTec**

Code	Commande	Glyphe	Code	Commande	Glyphe	Code	Commande	Glyphe
32	\alef	ا	64	\aalef	أ	96		
33	\beh	ب	65	\bbeh	ب	97	\BEH	ب
34	\jeem	ج	66	\jjeem	ج	98	\JEEM	ج
35	\dal	د	67	\ddal	د	99		
36	\waw	و	68	\wwaw	و	100		
37	\reh	ر	69	\rreh	ر	101		
38	\tah	ط	70	\ttah	ط	102	\TAH	ط
39	\yeh	ي	71	\yyeh	ي	103		
40	\lam	ل	72	\llam	ل	104	\LAM	ل
41	\meem	م	73	\mmeem	م	105	\MEEM	م
42	\noon	ن	74	\nnoon	ن	106		
43	\sheen	س	75	\ssheen	س	107	\SHEEN	س
44	\ain	ع	76	\aain	ع	108	\AIN	ع
45	\feh	ف	77	\ffeh	ف	109	\FEH	ف
46	\sad	ص	78	\ssad	ص	110	\SAD	ص
47	\qaf	ق	79	\qqaf	ق	111		
48	\hamza	ء	80	\hhamza	ء	112	\HEH	ء
49	\lamalef	لا	81	\llamalef	لا	113	\KAF	ك
50	\Yeh	ي	82	\YYeh	ي	114	\TTAH	ط
51	\Noon	ن	83	\NNoon	ن	115		
52	\Reh	ر	84	\RReh	ر	116		
53	\MEem	م	85	\MMEem	م	117		
54	\Beh	ب	86	\BBeh	ب	118	\BBEH	ب
55	\Jeem	ج	87	\JJeem	ج	119	\JJJEEM	ج
56	\Heh	هـ	88	\HHeh	هـ	120	\HHEH	هـ
57	\Kaf	ك	89	\KKaf	ك	121	\KKAF	ك
58	\Lam	ل	90	\LLam	ل	122	\LLAM	ل
59	\Meem	م	91	\MMeem	م	123	\MMEEM	م
60	\Sheen	س	92	\SSheen	س	124	\SSHEEN	س
61	\Ain	ع	93	\AAin	ع	125	\AAIN	ع
62	\Feh	ف	94	\FFeh	ف	126	\FFEHEH	ف
63	\Sad	ص	95	\SSad	ص	127	\SSAD	ص

Table 8: La fonte de symboles littéraux arabes **NasX**