جامعة نايف العربية للعلوم الأمنية

Naif Arab University For Security Sciences

حول حوسبة رومنة أسماء الأعلام العرب

أ.د. أحمد شرف الدين أحمد

٢٠٠٣م

حول حوسبة رومنة
أسماء الأعلام العرب

أ.د. أحمد شرف الدين أحمد

حول حوسبة رومنة أسماء الأعلام العرب

المخلص

يقصد «برومنة » أسماء الأعلام العرب كتابة هذه الأسماء بالحروف الرومانية . و نظراً لتفاوت و اختلاف طرق الرومنة المستخدمة فقد جرت محاولات لتقييسها . و بالفعل فقد أصدرت مواصفة قياسية عالمية لذلك . إلا أنه نادراً ما يُجْمِع المهتمون بالرومنة على أسلوب موحد فى ذلك . و بالرغم من أن الرومنة تتم حالياً بصورة يدوية فى الغالب إلا أنه توجد بعض النظم التي تحوسب عملية الرومنة . فى هذا البحث نستعرض طرق الرومنة و كيفية حوسبتها كما نقدم نظام الرومنة و الذي سبق للباحث أن اقترحه من قبل مع مقترح لناظم رومنة محوسب .

المقدمة

تنوع لغات البشر آية من آيات الله ، قال تعالى {$ّمُنَ اَيّاتُهُ خّلَقِ يسّمّوّاتُ $ّالاّرَضُ$ّاخَتُلافِ $ّلَسُنّتُكَمَ$ّلَوّانُكَمَإنّ فُي ذّلُكّ لآيّاتُ لَلّعّالُمُينّ<ڑ٢٢ڑ>} (سورة الروم) . هذا و قد انفرد حفص عن عاصم بقراءتها بكسر اللام و الباقون من القراء بفتحها ولذلك فقد جمعت القراءتان العلم و العمومية . و تقدر عدد اللغات الحية حالياً بالآلاف (منها أكثر من ٣٠٠ لغة فى الهند وحدها) . و أشرف هذه اللغات و أعلاها قدراً هي لغة القرآن الكريم ـ كلام الله ـ آي اللغة العربية .

وبصفة عامة فإن أي لغة طبيعية (أي ليست لغة لآلة) يمكن تقسيمها إلى خمسة مستويات :

أ ـ المستوى الصوتي phonology

ب ـ مستوى شكل وتركيب الكلمات (الصرف) morphology

ج ـ مستوى تركيب الجمل (النحو) syntax

د ـ مستوى المعنى (الدلالي) semantics

هـ ـ المستوى البرجماتى (الاستعمالى)pragmatic

و تتفاوت اللغات الطبيعية فى كل جانب من هذه الجوانب اختلافاً بيناً .

و من أجل التواصل بين الشعوب فإنه من الضروري أن تتم عمليات الترجمة بين هذه اللغات . على إنه يحدث فى بعض الأحيان أن لا يمكن ترجمة بعض الكلمات أو العبارات من لغة إلى أخرى . و هذا قد يكون راجعاً لأسباب متعلقة باللغة ذاتها أو لأسباب ثقافية أو اجتماعية . و لعل هذا أظهر ما يكون فى أسماء الأعلام . و فى هذه الحالة يتم التعبير عن الكلمة الأصلية (فى لغة المصدر source) بأحرف مكافئة فى اللغة الهدف (target) و هو ما يعرف بالتعبير عن كلمات لغة بأحرف لغة أخرى أو Transliteration . فإذا ما كانت أحرف اللغة الهدف هي الحروف «الرومانية» فإنه أحيانا يطلق على هذه العملية Romanization أو الرومنة . و من الضروري أن نلفت الانتباه إلى أن هناك نوعا من عدم الاتفاق حول مدلول هذين المصطلحين و أعني : Transliteration و Romanization (١) . و بالرغم من أن كلمة Transliteration ـ من الناحية اللغوية البحتة ـ لا تشترط التماثل أو حتى التقارب الصوتي بين منطوق الكلمة الأصلية و الكلمة المكتوبة فى اللغة الأخرى إلا أن استخدام هذا المصطلح فعليا قد أضاف هذا الشرط كأمر واقع في كثير من الأحيان عند الكلام على اللغات الطبيعية . و جدير بالذكر أن عملية كتابة لغة بأحرف لغة أخرى لها عمومية

أكبر من كونها تستخدم فى اللغات الطبيعية فهي تستخدم فى اللغات الرسمية Formal Languages و فى بعض علوم الحاسب الأخرى مثل كتابة المترجمات و ترميز البيانات (٢ ـ ٣).

لقد بدأت عملية الرومنة منذ قرون عديدة حينما اتصلت الحضارتان الإسلامية العربية و الأوربية خصوصاً على أيدي المستشرقين. و بالرغم من هذه البداية المبكرة فإن معظم نظم الرومنة كانت اجتهادية وتختلف من شخص لآخر كما سنبين لاحقا. و فى الوقت الراهن – و خاصة بعد أحداث ١١ سبتمبر ٢٠٠١ الشهيرة فى الولايات المتحدة الأمريكية – فقد ظهرت أبعاد أمنية لهذا الموضوع. فعلى سبيل المثال فقد وجدت السلطات الأمنية فى تلك البلاد صعوبة فى تتبع الوثائق التي تحتوى على اسم «أسامة بن لادن» باللغة الإنجليزية نظراً لاختلاف طرق كتابتها من شخص لآخر أو حتى لنفس الشخص (٤) «انظر موقع بوابة العرب Arab Gateway : Arabic Words and†the Roman Alphabet .

فى هذا البحث سنتعرض باختصار لنظم الرومنة و مشاكلها مع وضع الشروط الأساسية التي نرى ضرورة توافرها لنظم الرومنة المثالية. ثم نتعرض لمحاولات الحوسبة التي جرت فى هذا الشأن و نعرض لنموذج مقترح للرومنة و ذلك كله مع التركيز على رومنة أسماء الأعلام العربية بمعنى كيف يكتب اسم العلم العربي بحروف رومانية. .

نظم الرومنة

بدأت المحاولات الأولى للرومنة بواسطة المستشرقين في المقام الأول و كانت مبنية على اجتهاداتهم الشخصية فى كيفية كتابة الكلمة العربية كما

تنطق (حسبما سمعوها) بالأحرف اللاتينية . وقد شاب هذا الأمر عدة مشاكل منها أن الكلمة العربية الواحدة قد تنطق بعدة أساليب تبعاً لاختلاف المكان أو الأصل حتى لو راعى كل منهم النطق الفصيح لها قدر الإمكان . أضف إلى ذلك أن عملية الرومنة تتأثر بلغة المستشرق الأصلية فمن كانت لغته الأصلية الإنجليزية قد يستخدم الحرف a للدلالة على الفتحة ، sh للدلالة على حرف ش بينما قد يستخدم نظيره الألماني الحرف e للدلالة على الفتحة ، sch أو ch للدلالة على حرف ش و هكذا . . . و لذلك فإن كتابة اسم العَلَم العربي بالحروف الإنجليزية يمكن أن تتم بطرق عديدة . وحالياً يقوم اللغويون العرب و المستشرقون بعملية الكتابة هذه بطريقة يدوية، في أغلب الأحيان ، و هذا ينتج عنه ما يلي :

١ ـ عدم الاتساق فى عملية الرومنة فى داخل النص الواحد . و على سبيل المثال فإن لورانس العرب فى كتابه المعروف كان يكتب نفس الاسم العربي بعدة طرق مختلفة فعلى سبيل المثال فإن كلمة جَدَّة (اسم المدينة المشهورة بالمملكة العربية السعودية) كتبها أحيانا Jeddah و أحيانا أخرى Jidda . كما أنه كتب اسم عبد المعين بست طرق مختلفة (٤) .

٢ ـ عدم الاتساق فى عملية الرومنة بين النصوص المختلفة سواء أكانت لنفس الشخص أو لمجموعة مختلفة من الأشخاص . و هذا من باب أولى بطبيعة الحال . و قد استقصيت كتابة اسم شرف الدين مرومنا فوجدته يكتب بالصور التالية :

Sharaf Eldin, Sharaf El Din, Sharaf ElDin, Sharaf-Eldin, Sharaf-El-Din, Sharafeldin, Sharafel Din, Sharafelddin, Sharafelldin, Sharafudin, Sharafuddin, Sharaf Aldine, Sharaf Al Din, Sharaf Al Dine, Sharaf-Al-Din, Sharaf El deen, ...

٣ـ عدم ضمان استرجاع الكلمة العربية من الكلمة «المرومنة» كحالتها الأصلية (أي العملية العكسية للرومنة)

٤ ـ البطء فى «الرومنة » .

٥ ـ احتمال الخطأ البشرى نتيجة للعمل اليدوي أكبر منه مما لو كان هذا الأمر يتم حاسوبياً .

و لعل الحل المنطقي لهذا الأمر هو وضع نظام موحد ملزم لعملية الرومنة أي وضع مواصفة قياسية لذلك . و قد تعجب أيها القارئ الكريم حينما تعلم أن هناك عدة مواصفات مقترحة و بعضها متداول بين الباحثين إلا أن أيا منها لم يتحقق لها الانتشار الكافي . فعلى سبيل المثال يوجد حالياً على الساحة عدة طرق للرومنة و من أهمها (٥ ـ ٩) :

١ ـ مكتبة الكونجرس الأمريكية (L.C) و لعلها الأكثر شيوعاً .

٢ ـ المواصفة البريطانية للرومنة BS4280 (BSI) .

٣ـ المواصفة الخاصة بدائرة المعارف الإسلامية .

٤ ـ المواصفة الخاصة بالأيزو ISO .

٥ ـ «المواصفة» الخاصة بالمجلة الدولية لدراسات الشرق الأوسط (IJMES) .

٦ ـ «المواصفة» الخاصة بمعهد الدراسات الإسلامية بجامعة مكجيل بكندا .

و يبين الملحق أ «المواصفة» الأخيرة .

و ربما يدعو للدهشة أن كثيراً من الباحثين لا يستخدمون أيا منها بل أن بعضهم ربما يبتدع طريقته الخاصة للرومنة كما في (١٠) .

وفي الوقت الحالي و نتيجة لانتشار الإنترنت و الهواتف المحمولة

يستخدم كثيراً من الناس أسلوباً طريفاً للرومنة وذلك في الاتصال على شبكة الإنترنت (سواء فى البريد الإلكتروني أو على شبكة المعلومات أو في غرف المحادثة) أو فى الرسائل القصيرة على الهاتف المحمول (أو الجوال) فعلى سبيل المثال الهمزة وتمثل بالرقم ٢ ، ح و يمثلها ٧ ، ع و يمثلها ٣ ، خ و يمثلها الرمزين ٧٬ أما باقي الحروف فأنها تنطق كما تكتب و على كل فإننا لن نتعرض لهذا الأسلوب فى هذا البحث .

و من عيوب نظم الرومنة السابقة أنها تحتاج لرموز خاصة إضافة للحروف العادية الأبجدية . و هذا بدوره يؤدي لصعوبة الحوسبة نظرا لعدم وجود هذه الرموز على لوحة المفاتيح بشكل مباشر (مثل وضع نقاط أعلى حرف غير منقوط أصلا) . هذا إضافة إلى أن بعض الحروف العربية يتم رومنتها بأكثر من حرف واحد لاتيني مما قد يوقع في اللبس . هذا علاوة على صعوبة تذكر هذه الرموز بصفة عامة . و يلاحظ أيضا أن العملية العكسية لاسم مرومن قد لا تنتج الاسم العربي الأصلي . ونظراً لعدم اتباع طريقة واحدة لرومنة الأسماء ، بل إنه حتى بافتراض وجود هذه الطريقة فإنه لا ضمانة لاستعمالها بواسطة عموم الناس أو بواسطة موظفي الدولة (الجوازات مثلاً) ، فإن أي اسم معين يمكن كتابته بعدة طرق مختلفة فمثلاً شرف الدين ـ كما سبق و أن أوضحنا ـ يمكن كتابتها بست عشرة طريقة على الأقل ، بينما كلمة سليمان أمكن رصد ٤٠ طريقة مختلفة لكتابتها (١١) . و هذا بدوره يؤدي لمشاكل مختلفة خاصة عند استرجاع الأسماء المرومنة من قاعدة بيانات . و بالرغم من أن المواصفة البريطانية BS2480 تعتبر متسقة إلى حد معقول إلا أنها لا تستخدم على نطاق واسع . و فى البحث (١٢) يوجد مقارنة بين هذه النظم يمكن الرجوع إليها . و بالرغم من وجود مواصفة قياسية للصوتيات (١٣) International Phonetic Alphabet إلا

أنها معقدة بدرجة لا تلائم إلا المختصين فى مجال اللغويات فقط . و قد يبدو أن البديل الطبيعي لذلك هو التعويض عن كل حرف عربي بنظيره الصوتي الإنجليزي (الروماني) إلا أن هذا غير متيسر نظراً لأن معظم الحروف العربية لا يمكن التعويض عنها بحرف واحد و ذلك مثل حرف ش مثلاً كما أن بعض الحروف ليس لها نظير صوتي فى لغة الهدف .

متطلبات نظم الرومنة الجيدة

يمكن لنا أن نضع الخصائص الأساسية التالية لمتطلبات ضرورية لنظم الرومنة

١ ـ التناظر ١ ـ ١ بين كل حرف فى لغة المصدر (العربية) ، لغة الهدف (الإنجليزية) و هذا الشرط يصعب الالتزام به تماماً وذلك لما يلي:

أ ـ الحروف العربية تحتاج إلى علامات التشكيل كي يمكن ضبطها وتكتب هذه العلامات أعلى أو أسفل الحرف الأصلي . و بالتالي فإن التناظر يمكن الالتزام به آلياً (حيث أن التشكيل يعتبر حرفاً مستقلاً) و لا يمكن الالتزام به يدوياً (حيث أن الحرف المشكل هو حرف واحد فقط) .

ب ـ فى معظم الأحوال لن يمكن نطق اسم العلم صحيحاً إلا للشخص المتمرس فقط .

٢ ـ إمكانية استرجاع الحرف العربي الذي سبق رومنته إلى حالته الأصلية بدون لبس .

٣ ـ الشمولية بمعنى أن كافة الرموز و الأصوات اللغوية فى لغة المصدر يكون لها مقابل فى لغة الهدف .

ولبيان مدى أهمية التناظر ١ـ١ بين الأحرف فى لغتي المصدر و الهدف فأنني أورد مثالاً بسيطاً وباستخدام مواصفة معهد الدراسات الإسلامية و التي سبق الإشارة إليها . إن كلمة مثل شما [وهى من أسماء العائلات المشهورة فى مصر حالياً ، واسم قرية فى محافظة المنوفية كذلك] تكتب حسب هذه المواصفة shma و كلمة سهما تكتب : shma

أيضاً حيث أنه فى هذه المواصفة يستخدم حرف s للدلالة على حرف س ، حرف h للدلالة على حرف هـ ، و الحرفين معا sh للحرف ش أما حرف m فيستخدم للدلالة على حرف م . و هذا يسبب اللبس فى عملية الاسترجاع (عكس الرومنة و لنسمها التعريب) وأيضاً فى عملية الرومنة ذاتها يجعل من الصعوبة النطق مباشرة بالكلمة المرومنة صحيحا .

حوسبة الرومنة

بدأت محاولات حوسبة عملية الرومنة ببعض الأبحاث الأكاديمية ولعل أول هذه الأبحاث و التي أثمرت بعد ذلك و التي أمكن رصدها هو الخاص برسالة الدكتوراه من جامعة جورج تاون للسيد بول روكنك (١٤) (وهو مهتم باللغة العربية و يكنى ب «أبو سامي») . و في رسالته و ما تلاها من أبحاث متممة قام بدراسة مشكلة رومنة الأسماء العربية و استخدم المنطق المشوش كأسلوب للبحث في قواعد البيانات متعددة اللغات . و قد انضم إلى إحدى شركات البرمجيات الشهيرة في هذا المجال و هي شركة أبتك (١٥) (Apptek) حيث قامت الشركة بتطوير بعض البرمجيات والأدوات اللازمة (و التي يمكن بناء تطبيقات أخرى عليها) مثل NameFinder ـ و الذي يمكن استخدامه للبحث عن الأسماء في قواعد

البيانات متعددة اللغات. و من ضمن الأدوات (Case Tools) الخاصة بالشركة Diacritizer و الذي يمكنه تشكيل الكلمات العربية و كذا Transliteration/ Romanization Tool و التي تستخدم لاستخراج الأسماء المرومنة. و تعتمد جميع هذه الأدوات على بناء قاعدة بيانات كبيرة و متزايدة إضافة للمنطق المرتبط بعملية البحث. هذا و كانت مجموعة من الباحثين في جامعة انديانا قد سبقت د. روكنك و طورت نظاما للرومنة ضمن الأعمال اللغوية المحوسبة للغة العربية (١٦) و التي أسفرت عن نظام كلام Qalam إلا أن هذا المشروع لم يستمر على ما يبدو كما أن الموقع الخاص به ضمن الجامعة لم يعد موجودا.

و في عام ١٩٩٤ قام أربابي و آخرون من شركة IBM (١٧) باستخدام تقنية الشبكات العصبية لفلترة الأسماء غير الموثوق بها و من ثم يتم إرسال الأسماء الأخرى الموثوق بها إلى قاعدة معرفة لرومنتها. وقد تبنت شركة زيروكس مشروعاً ضخماً لمعالجة اللغة العربية أنتج فى عام ١٩٩٦ محلل صرفي للغة العربية (١٨ ـ ١٩) و بالرغم من أن هذا المشروع ليس معنياً بصفة أساسية بعملية الرومنة إلا أن ما تم تطويره من محلل صرفي يمكن أن يشكل أساساً لبناء نظام رومنة على مستوى عال من الكفاءة.

و لعل من أهم و أفضل الأعمال التي تمت بعد ذلك هو ما قام به مجموعة من الباحثين في جامعة جنوب كاليفورنيا (٢٠ ـ ٢٢) و كلها مبنية على الرومنة مع التقارب أو التماثل الصوتي باستخدام بعض القواعد والتي تضمن الوصول لرومنة معقولة و إن كانت قد تفشل أحيانا.

و يلاحظ في سائر الأعمال التي سبق الإشارة إليها المميزات التالية :

١ ـ تقوم بالرومنة مع مراعاة التماثل أو القرب الصوتي بين الاسم المرومن و الاسم العربي .

٢ ـ تقوم بالتعرف ـ و بنسبة متفاوتة من النجاح ـ على الأسماء التي رومنت بطرق مختلفة .

٣ ـ معظمها لا يحتاج لتشكيل الاسم العربي مسبقا .

٤ ـ يقوم معظمها ببناء قاعدة بيانات للأسماء

وعلى الجانب الآخر فإن هناك بعض الملاحظات عليها كما يلي :

١ ـ أنها لا تلتزم بالتناظر ١ ـ ١ بين الحرف العربي و الحرف الانجليزي .

٢ ـ أنها لا تشمل كافة الحروف و الحركات الموجودة و المستخدمة في اللغة العربية المعاصرة . على سبيل المثال بعضها لا يعرف همزة الوصل و لا يفرق بين التاء المربوطة و الهاء المربوطة و هكذا . أيضا هناك حروف تكتب خاصة بالقرآن الكريم و لا تتناولها هذه النظم . كما أن بعض الحروف ، و إن لم تكن عربية أصلا ، إلا أنها أصبحت شائعة الاستخدام حاليا مثل الفاء ذات ثلاث نقاط و التي تنطق مثل حرف V

٣ ـ أنها لا تراعي الاختلافات البينية في اللغة العربية المعاصرة بين البلدان العربية المختلفة .

٤ ـ أنها صعبة الكتابة باستخدام لوحة المفاتيح .

٥ ـ أنها قد تفشل أحيانا في التعرف على الاسم المرومن .

النظام المقترح

سبق و أن اقترح الباحث نظاماً للرومنة فى (١٢) و يمتاز هذا النظام بتحقيقه للشروط الأساسية المطلوب توافرها فى نظم الرومنة علاوة على إضافة بعض الرموز الأخرى للكتابة الصوتية لبعض أحكام التجويد فى القرآن الكريم . و سنعيد هنا الجداول الأساسية المقترحة للرومنة (ملحق ب) . هذا و يعتبر بناء البرنامج الذي يقوم بعملية الرومنة فى هذه الحالة بسيط للغاية حيث لا يتعدى النظر فى جدول الرموز و استبدال كل رمز عربي بما يقابله .

و بالرغم من المزايا الواضحة لهذا النظام فإنه به بعض أوجه القصور لعل أهمها صعوبة التعلم كما أن النطق بكلمة مرومنة بدون مران قد لا يعطي صوتا مماثلا للصوت الأصلي للكلمة . على ذلك فمثلاً كلمة سَلَكَ (الفعل) تكتب : salaka و التي يمكن نطقها بصورة مناسبة للأصل العربي أما كلمة طالب (بدون تشكيل) فتكتب : – TUlb

وتكتب كلمة طَالبٌ Tauleboo و التي ستنطق بصورة مخالفة للأصل العربي . ولعل هذا الأَمر هو العيب الأساسي فى النظام المقترح . أما إذا أردنا استخدام أحد نظم الترميز الأخرى مثل LC و التي تتميز بسهولة النطق (ولكنها لا تستوفى المتطلبات الأساسية سالفة الذكر) فإن استخدام أحد أساليب الذكاء الاصطناعي (مثل الشبكات العصبية) و/ أو قواعد البيانات الكبيرة يصبح أمرا لامناص منه .

و فى هذه الطريقة تتم عملية الرومنة بغض النظر عن التشكيل . و بالتالي فإن عملية الرومنة الآلية تكون مباشرة و مبنية على النظر في جدول التحويل

مباشرة . يبين الشكل رقم (١) المعمارية العامة للنظام المقترح و ذلك بصورة إجمالية و سوف نشرح هنا كل جزء من هذه المعمارية على النحو التالي :

١ ـ يتم إدخال اسم العلم إما مشكلا (إذا توافر ذلك) أو غير مشكل (و هو الغالب) .

٢ ـ إذا كان الاسم مشكلا فإنه يذهب مباشرة إلى نظام الرومنة أما إذا كان غير مشكل فإنه يدخل على نظام التشكيل — و ذلك فقط إذا ما أردنا استخدام نظام للرومنة غير النظام المقترح أما في النظام المقترح فإنه لا يهم كون الاسم مشكلا أو غير مشكل .

٣ـ في نظام التشكيل يتم تشكيل الحروف القابلة للتشكيل . و نظراً لأن عملية التشكيل لا تعطى جواباً واحداً فى معظم الأحيان فإنه من الضروري أن يرتب هذا النظام مخرجاته ـ إذا تعددت ـ بحسب نسبة احتمالها مع مراعاة أن استخدام أحد أساليب الذكاء الاصطناعي (مثل الشبكات العصبية) و/ أو قواعد البيانات الكبيرة جدا يصبح لامناص منه .

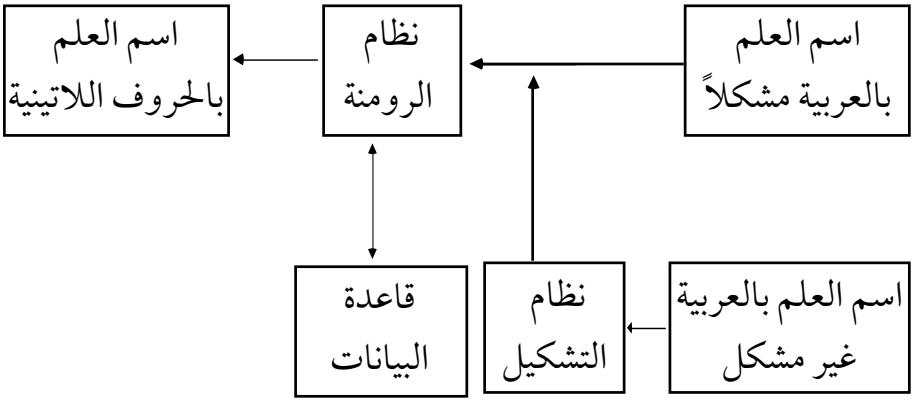
و لعل هذا هو العيب الأساسي فى النظام المقترح . أما مميزاته فهيب الالتزام بالتناظر التام ١ـ١ بين الكلمات العربية و المرومنة . إضافة لذلك فإن تكوين قاعدة بيانات بالأسماء العربية و رومنتها يحقق فائدتين أساسيتين و هما :

I ـ الاتساق في عملية الرومنة و السرعة بحيث لا نحتاج بعد ذلك لإجراء عملية التحويل إلا للكلمات الجديدة .

II ـ استخدام قاعدة البيانات هذه كقاموس إضافي عند التدقيق الإملائي .

الخاتمـــــــة

حاولنا في هذا البحث المختصر إلقاء بعض الضوء على مسألة الرومنة للأسماء العربية . و قد رأينا أن نظم الرومنة متعددة و أن أيا منها لا يستوفي كافة الشروط المتوقعة من نظم الرومنة . إذ أن بعضها يحاول المحافظة على التشابه أو التماثل الصوتي على حساب بعض الأمور الأخرى مثل التناظرية بين الحروف العربية و النظير الروماني . و لعله من المناسب أن نبين أن هذه المشاكل ليست قاصرة على رومنة الأسماء العربية فحسب بل إن معظم اللغات الطبيعية تشترك في هذا (٢٣) . وقد بينا باختصار مميزات و نواقص هذه النظم المختلفة . كما استعرضنا الأبحاث و بعض المنتجات الخاصة بذلك بصورة إجمالية . و مع ازدياد نظم الترجمة الآلية فقد تم تطوير بعض النظم التي تحوسب عملية الرومنة . و قد قدمنا نظاما مقترحا للرومنة و الذي نعتقد أنه مناسب للحوسبة بصورة عالية .



الشكل رقم (١) معمارية النظام المقترح

المراجع

المراجـــع

1.http://www.cis.upenn.edu/~cis639/arabic/info/romanization.html Contains details of the Xerox Arabic Morphology project.

2.Mitloehner, J., “Porting APL programs via ASCII-transliteration“, ACM APL Quote Quad, Vol. 23, no. 1, July 1992, pp. 148-155

3.http://www.informatik.uni-stuttgart.de/ifi/bs/publications/rep9719/ 1.htmOne of the publications of Prof. Dr. Klaus Lagally, a dominant scholar in Arabization.

4.http://www.al-bab.com/arab/language/roman1.htm#TRANSCRIPTION A good discussion about the “standards“ used for Romanization of the Arabic text.

5. http://lcweb.loc.gov/catdir/cpso/romanization/arabic.pdf The LC Romanization table.

6. British Standard Institution, Transliteration of Arabic characters, BS 4280, 1968 (with its amendment).

7. International Standards ISO 233-2, Transliteration of Arabic characters into Latin characters, 1993.

8. Lance Anderson and World wide Fonts, LaserTransliterator for Windows, 1997.

9. http://www.arts.mcgill.ca/programs/islamic/info/font.html Contains the Arabic Transliteration “standard“ devised by the institute of Islamic studies, McGill University, Canada.

10. Hanna, S. and N. Greis, “Writing Arabic - A linguistic approach from sounds to script“, E.J. Brill, Leiden, 1972.

11. Arbabi, M., S. Fischthal, V. Cheng and E. Bort, “Algorithms for Arabic name transliteration“, IBM J. of research and development, Vol. 38, no. 2, Mar., 1994, pp. 183-193.

12. Sharaf Eldin, A., “Computer-Assisted Arabic Transliteration“, 7th.. ICAIA, 1999.

13. http://www2.arts.gla.ac.uk/IPA/ipachart.html Contains the International Phonetic Alphabet.

14. Roochnik, P., “Computer-based solutions to certain linguistic problems arising from the Romanization of Arabic names“, Ph.D. dissertation, Georgetown university, Washington DC, 1993.

15. http://www.apptek.com/ The web site of the Application Technology Inc. A software house that has some products in multi-language computing. Dr. Roochnik is affiliated with them.

16. http://www.cs.indiana.edu/hyplan/dmulholl/info/Qalam.txt This site is no more on the web. It was there some years ago reporting the “Qalam“ system.

17. Arbabi, M., S. Fischthal, V. Cheng and E. Bort, “Algorithms for Arabic name transliteration“, IBM J. of research and development, Vol. 38, no. 2, Mar., 1994, pp. 183-193.

18. Beesley, K. R., “Computer Analysis of Arabic Morphology: A two-level Approach with Detours“, In Comrie,

B. and Eid, M., editors, perspectives on Arabic Linguistics III: papers from the third Annual Symposium on Arabic Linguistics , pages 155 _172,†john Benjamins, Amsterdam, 1991.

19. Beesley, K. R. , “Arabic Finite-state Morphological Analysis and Generation“, In COLING-96 proceedings, volume 1 , pages 89-94 Copenhagen, Center for Sprogteknologi. The 16th International Conference on Computational Linguistics, 1996.
20. Knight, K. and J. Graehl, “Machine Transliteration“, Proceedings of the 35th. Annual Meeting of the Association for Computational Linguistics, pp. 128-135, 1997.
21. Stalls, B. and K. Knight, “Translating Names and Technical Terms in Arabic Text“, Proceedings of the COLING/ACL Workshop on Computational Approaches to Semitic Languages, 1998.
22. Al-Onaizan, K. Knight, “Transliteration of Names in Arabic Text“, obtained via http://www.isi.edu/natural-language/projects/rewrite/yaser-acl02-WS.ps
23. Ridley, M.J., “A code for bibliographic records transliterated from Greek“, literary and Linguistic Computing, Vol. 7, pp. 27-29.

الملحق (أ)

جدول الرومنة لمعهد الدراسات الإسلامية التابع لجامعة ماكجيل

f	=	ف	z	=	ز	b	=	ب
q	=	ق	s	=	س	t	=	ت
k	=	ن	sh	=	ش	th	=	ث
l	=	ل	s{	=	ص	j	=	ج
m	=	م	d{	=	ض	h{	=	ح
n	=	ن	t{	=	ط	kh	=	خ
h	=	هـ	z{	=	ظ	d	=	د
w	=	و	‘	=	ع	dh	=	ذ
y	=	ي	gh	=	غ	r	=	ر

الملحق (ب)

النظام المقترح للرومنة جدول رقم ١ ـ الصوائت

Sequence	Arabic Character	Transliterated into
1	ء	'
2	ا	U
3	آ	A
4	ب	b
5	ت	t
6	ث	B
7	ج	j
8	ح	H
9	خ	K
10	د	d
11	ذ	c
12	ر	r
13	ز	z
14	س	s
15	ش	C
16	ص	S
17	ض	D
18	ط	ST
19	ظ	Z
20	ع	W
21	غ	G

Sequence	Arabic	Transliterated into
22	ف	f
23	ق	Q
24	ك	k
25	ل	l
26	م	m
27	ن	n
28	هـ	h
29	ة	g
30	و	w
31	ي	y

الجدول رقم (٢)

Sequence	Arabic Character	Transliterated into
1	ـَ fat-ha	a
2	ـُ Dhamma	o
3	ـِ Kasra	e
4	ـْ Sokon	-
5	تنوين	doubling the vowel
6	ـّ Shadda	doubling the consonant
7	ئ	Y‘
8	ي Alef Maqsora	Y
9	ؤ	w‘