INTERNATIONAL JOURNAL OF ENGLISH LANGUAGE, LITERATURE AND TRANSLATION STUDIES (IJELR)
A QUARTERLY, INDEXED, REFEREED AND PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
http://www.ijelr.in (Impact Factor : 5.9745 (ICI)

RESEARCH ARTICLE

Vol. 4. Issue.4., 2017 (Oct-Dec)

INTERNATIONAL STANDARD SERIAL NUMBER INDIA
2395-2628(Print):2349-9451(online)

THE COMPLICATED TRANSCRIPTION OF ARABIC SCRIPT INTO ENGLISH

SADDIK GOHAR
Professor of English Literature, Translation and Criticism & Chair of the English Literature Department, UAE University

ABSTRACT

Transliteration is the representation of a word or a phrase in the closest corresponding letters or characters of a different alphabet or language so that the pronunciation is as close to the original word or phrase. Within the theoretical framework of contemporary scholarship on the issue of transliteration, this paper aims to construct an appropriate transliteration dynamics initially through an investigation of the problems encountered in the process of transliterating Arabic proper names into English. In the argument of the paper, two major reasons are identified as the central obstacles confronting translators in the context of transliteration. The first reason is integrated in Arabic language idiosyncrasies such as omission of diacritics, variability in Arabic dialects – due to geographical differences, variations between Classical Arabic, Modern Standard Arabic (MSA) and Modern Conversational Arabic (MCA) etc. The second major reason is rooted in distinctions between Source Language (SL) and Target Language (TL) structures particularly morphological and heuristic variations in addition to differences in alphabetical systems which result into serious consequences integral to the process of transliteration.

1.1 Literature Review: An Introduction

The process of transliterating Arabic script into the Roman alphabet involves major difficulties and complications. In an era where electronic information retrieval for security purposes becomes integral to intelligence organizations on the national, regional and global paradigms, experts, in the Arab world, should develop a standard internationally agreed system for transforming Arabic text into Roman alphabet. Variations of the transliteration of Arabic names leads to serious problems because names are different from one document to another. Further, the inconsistency of Arabic pronunciation across the Arab territories extending from the Arabian Gulf to the Atlantic Ocean and the variations of Arabic vowels from one district to another create problems on different levels. No serious attempt was made to standardize the spelling of proper names though successful attempts were made standardize the spelling of geographical locations on maps. In this context, the Arabic transliteration system, adopted by the United Nations in 1972 after the recommendation of the Arab Experts Conference in Beirut (1971) was implemented in some Arab countries but it was not accepted in North African Arab countries which used a French-oriented transliteration dynamics. In Egypt, there is a local transliteration system which excludes both English and French orthographies. But it is relevant to point out that some modifications are recently integrated into all the transliteration systems in the Arab world like the use of (dh) to represent (ظ) instead of (Z) and the consistent use of Arabic long vowels فتحة ضمة كسرة(aah'oo' ee). However, transliteration problems continue to manifest themselves

everywhere. The omission of diacritics in written material published in newspapers, magazines and popular press lead to major problems because readers are not able to fill in the spaces and provide their own diacritical marks. Moreover, the Arabic glottal soundالهمزة (alhamzah) which does not exist in English is considered as one of the chronic problems encountered by transliteration experts and mechanical retrieval applications - for example when a word starts with همزة (hamzah), in pronunciation, it is not included in transliteration.

Arabic–English cross language retrieval dictionary-based transliteration or Machine Translation (MT) do not work properly with languages with two different alphabets (like Arabic and English). Problems involving variations in heuristics and linguistic knowledge of both languages complicate the process of transliteration. The irregularity of Arabic proper names, which have more than one correct spelling according to context and due to variations in diacritics as in - محمدٌ، محمداً، محمدٍ - problematizes the transliteration process because the existence of multiple Arabic spellings requires the generation of multiple alternative spelling in English. Other potential complications result from variations in short vowels in Arabic and English orthographies . Unlike Arabic, English has many short vowels that appear in transliteration. There are more differences in the phonetic inventories of the two languages. For example, Arabic has no (P) sound but it has two different (t's) - ط ، ت - while English has no (ظ, ض, غ, ع)sounds. Moreover, there are Arabic names like فاطمة or شيخة that end with هاء(h) sound but transliterated in two different ways as in "Fatima" and "Fatimah". Besides, it is not easy for a non-native Arabic reader to identify the difference between the letter / sound هاء (h) as in هدى "Hoda" and حاء (h) as in " Hamid " due to similarity in transliteration symbols and because English does not have the ح - حاء - sound. Furthermore, the dark (L) is transcribed in English as one sound/letter but in Arabic it is pronounced as a double sound as in (Allah).

Regardless of these difficulties and complications a great deal of research on Arabic-related transliteration was pursued to fulfill the purpose of Machine Translation and Electronic Information Retrieval. For example, Alanzi (2004) discusses algorithmic software that is able to produce standard transliteration/Romanization of Arabic alphabet name presentation. Arbabi and Fischtal (1994) developed a hybrid neutral network and knowledge -based system to generate multiple English spellings for Arabic personal names. Further, Knight and Graehl (1992) developed a five–stage statistical model to perform backward transliteration, that is, recover the original English name from its transliteration into Japanese Katakana. Stalls and Knight (1998) adapted this model for backward transliteration from Arabic to English of English names. These systems are very complicated involving a great deal of human design simply because they tackle the issue of backward transliteration, not only forward transliteration.

Moreover, Darwish, et al (2001) provided a hand-crafted English to Arabic transliteration system. Each English letter was mapped to the closest Arabic letter or letters. These mappings were designed manually and most English letters were given a single Arabic equivalent but few had more than one. Abdul-Jaleel and Larkey (2003) developed a dynamics similar to Darwish's system and one version of their system takes into account the phonetic context. They also attempt to evaluate the performance of their transliteration whereas Darwish did not present any evaluation of his transliteration system. Furthermore, Koehn et al (2007) developed a dynamics called Moses which is a phrase- based Statistical Machine Translation System (SMTS). Throughout this system, transliteration is performed on a separate level while language models are constructed and combined during decoding to find out the most likely results. Stalls and Knight (1998) argue that results are not satisfactory particularly when attempts are made to model transliteration as a combination of both grapheme and phoneme level transformation.

Furthermore, there are several internationally recognized transliteration systems for Arabic- in addition to other local standards – like American Library Association – Library of Congress, ALA–LC and the transliteration system of the International Journal of Middle East Studies, (IJMFS) as well as Encyclopedia of Islam. These systems vary in their degree of precession and consistent use of diacritics. For example, there are many Arabic consonants that have no equivalent in the Latin alphabet. Systems like ALA-LC and Encyclopedia of Islam use pairs of Latin letters to designate a single Arabic consonant in the cases of ث ، ذ ، خ ، ظ ، غ ، ش. However, it is relevant to argue that using two letters to designate a single sound is a less precise method of transliteration. Moreover, the International Phonetics Association (IPA) has a comprehensive

transliteration / transcription system but it is not easily consumed by normal readers because it is not a keyboard friendly system.

2.1 Proper Names

Proper names constitute names of individuals, places, organizations and other items including acronyms - nouns that are typically capitalized in English. Proper names may exist in the source and target cultures particularly when the concerned languages have similar alphabets. Nevertheless, problems still arise when proper names – even in languages with similar alphabets - are transformed / transcribed from one language to another. For example the name "Karl" appears in different forms like "Carl, Karl, Carlos" in European languages with similar orthographic systems. Peter New mark (1988) argues that proper names/nouns should be transformed accurately in order to preserve nationality and the specificity of the source language (SL). Proper names, transliterated into target languages (TL) unfortunately have different pronunciations due to variations in stress, phonological replacement and substitution of vowels. Therefore, proper names should be rendered in the closest sounding letters of a different target alphabet taking into consideration the morphology and phonology of the original names. This process requires a solid transliteration dynamics which takes into consideration variations in language structures particularly when languages have different orthographic systems.

The identification of Arabic names involves several problems due to variations in the transliteration (Romanization) of Arabic script as well as the elusive structure of Arabic names which does not conform to western data entries and filling systems. For example, there is a class of Arabic compound given names beginning with the word عبد - servant or slave- followed by the word الله "Allah "- God - like عبد الله "Abd Allah". In some cases, the initial word عبد "Abd" is followed by the definite article ال written in three different forms according to variations in pronunciations "al- el- ul" followed by a word which is a descriptive name of الله - God- like فتاح قيوم ، صمد ، رزاق، - "Fattah, Qayyum, Samad, Razzaq " respectively. The preceding name combinations are written in Arabic as follows : (عبد الفتاح - عبد القيوم - عبد الصمد -عبد الرزاق) respectively. These names are sometimes transliterated as single words like "Abdulfattah" but they also appear with spaces and/or hyphens in between as in "Abd al Qayyum" or "Abd el Samad" or " Abdul-Razzaq". It is obvious that the names prefixed with the lexical item عبد"Abd" are problematic in transliteration because of variations in spelling such as in عبد الرحمن which could be transliterated in different forms "Abdulrahman, Abd al Rahman. etc". Despite the spaces between the parts of the name, it is impossible to separate or invert these parts and components.

Due to variations in spelling, these compound names like their counterparts which start with أبو "Abu" are problematic in transliteration. There is a class of compound names which start with "Abu" meaning "father of " as in أبو جهاد أبا جهاد ، ، أبي جهاد" Abu Jihad – Aba Jihad – Abi Jihad ". Due to the rules of Arabic syntax, these names are more problematic in transliteration than other compound names because the prefix "Abu" appears in different forms in Arabic as أبا -أبو – أبي according to its position in the sentence as clarified in the preceding example. In other words, the three prefixes have the same meaning but they occur in different forms in accordance with different grammatical contexts.

In addition to transliteration complications resulting from the use of patronymic names such as أبو "Abu", more difficulties are caused as a result of the frequent use of the word إبن "Ibn" which means " son of " as in " Ibn Batuta " which means the "son of Batutah". The historical and medieval lexical item إبن "Ibn" also appears in different forms as بن "Bin" – بن تميم " Bin Tamim " or as ولد "weld" which also means "son of" and is used in Arab countries located in North Africa like ولد المياح " son of Almayyah". As a result of variations of spellings, these compound proper names are transliterated in different forms which lead to ambiguity and confusion.

Nevertheless, when divergent spellings are widespread, the solution lies in double-posting as in words like قرآن - Koran or Quran and أسامة - Osama or Usama. Further, due to variations in spelling of names, sorting problems occur because Arabic names which start with (al) are problematic. Standards for sorting such particularly the criteria adopted by the Library of the Congress attempt to find solutions with the definite article (al) and the Arabic diacritical marks. According to the standard sorting systems, the initial definite article (al) should be eliminated and diacritics as well as Arabic letters indicated by apostrophe like الهمزة

(Alhamzah) are disregarded whereas a hyphen is considered “to divide a word for word- by -word sorting (Hedden, 2007 : 11). Nevertheless, the existence of different practices of sorting which deal with Arabic language leads to variations in transliteration as well.

Furthermore, Arabic names have peculiar structures and distinctive orthographic and pronunciation characteristics which cause difficulties during the process of transliteration, formulation of name searches and identification. Some of these difficulties are rooted in the basic historical components of Arabic names including given names, patronymic names like Ibn, Bin or Bint as well as names starting with كنية" kunya" like "Abu or Umm" in addition to الألقاب "alalqab" or titles that refer to names of professions, geographical locations and tribal affiliations. Moreover, Arabic language includes a number of auxiliary name elements or affixes and prefixes used in the formation of proper names like عبد"Abd" or الدين"Aldeen" in سيف الدين"Seif Aldeen ". The transliteration of these elements is inconsistent as follows : "Abdul, Abdel, Abdal or Aldeen, Addin, Eldin". Further, these elements may be attached to the name or modify it as in the name عبد الصمد "Abdusamad " which appears in different forms like "Abdal Samad, Abd el Samad, Abdul Samad, Abd-alsamad … etc". Theselexical items like " al, el, Abu, ibn, bin and bint " are fragments which can not constitute complete names. Alien to western alphabets, these suffixes and prefixes constitute a source of segmentation variations and spelling inconsistencies when transliterated into English orthography. Due to these auxiliary elements, Arabic proper names have different transliterations. This phenomenon inevitably leads to dramatic consequences which cause embarrassment and confusion on the part ofnon-Arabic readers.

In this context it is obvious that identifying proper names in Arabic script is a difficult process because names do not start with capital letters and thus they cannot be marked in the text. Further, there are different types of proper names in Arabic such as names given at birth likeعلى “Ali” or names including the prefix أبو “Abu” which means "the father of "and is used in compound names like أبو عادل “Abu Adel” or أبو زينب “Abu Zeinab ”. Moreover, there are other compound names which include the prefix إبن “Ibn” which means "the son of" as in إبن زياد “Ibn Zeyad ” which means “the son of Zeyad” in addition to name combinations including a name and an adjective such as مسيلمة الكذاب "Mosaylema Alkathab" which means Mosaylema " the liar". There are other name combinations based on one’s personal profession like يوسف النجار“Yusuf Alnajjar” – "Yusuf the carpenter" – as well as names based on one’s religion like أبو لؤلؤة المجوسى"Abu Luluah Almajusi" –"Abu Luluah the Magi"- or مينا النصرانى"Mena Annasrani" which means "Mena the Christian". Furthermore , there arename combinations based on one’s region or tribe like سمير الاسكندرانى Samir Aleskandarani – "Samir of Alexandria" or سالم الجهنى Salem Aljuhani "Salem who descends from the tribe of Juhaynah). In a related context, it is explicit that Arabic geographic names are not easily identified in Arabic alphabet whereas historical names are accurate in transliteration. Nevertheless, the transliteration of names pertaining to the Islamic faith changed over time, out of respect to the religion’s followers. Therefore, “Moslem” becomes “Muslim” and “Koran” becomes “Quran”.

Apparently, multiple transliterations of Arabic names make it difficult for governments and security apparatuses to track persons with criminal records and potential terrorists. The multiplicity of transliterations of Arabic names is due to a variety of reasons including the inconsistency of manual transliteration systems particularly facilities designed long time ago before the era of computers and digital retrieval. These systems depend on vague phonetic transcriptions of words like “Koran” instead of قرآن "Quran", “Mecca” instead ofمكة"Makkah", "Medina” instead of مدينة"Madinah", “Jidda” instead ofجدهJeddah". Confusion in transliteration results from variations between written and spoken Arabic texts as well as differences in Arabic pronunciation due to geographical variations. For example, the Arabic name ذكي "Thaki" is pronounced as زكى "Zaki" in Egyptian dialect whereas the name ثابت"Thabet" is pronounced سابت "Sabet" in Egyptian dialect. On this basis, it is obvious that transliteration, based on local dialects, is misleading and ambiguous. Moreover, Arabic names have consonants that are different from their English equivalent sounds. For example the glottal sound الهمزة "Ahamzah" in Arabic has no equivalent in the Roman script. Obviously, the transliteration of Arabic sounds into English is a complicated process because Arabic, unlike English, has nine consonant sounds which do not exist in English language likeظ,غ,ع,ضetc. While the Arabic alphabet has twenty eight letters , English has twenty six letters, however, the English alphabet has several repeated sounds. For example the letters K,

Q, C all denote the same sound "K". In other words, the phoneme "K" is represented in English by three graphemes "K, Q, C". This phenomenon does not exist in Arabic language.

3.1 Transliteration

Difficulties in transliteration of Arabic names into English or standard phonetic Roman representation is also due to variations between Modern Standard Arabic, Modern written Arabic and Modern Conversational Arabic. Modern standard Arabic is commonly used among linguists and the computational researchers community although there is often little agreement on its definition. MSA is nobody's native or local tongue though we have an existing writing and reading MSA community. Apparently, MSA is mainly the language of written discourse used in formal communication both written and oral – with a well defined range of stylistic registers. MWA constitutes the written variety of MSA whereas MCA involves the spoken variety of MSA.

In the Arab world, there are different transliteration methods including Anglophone, Francophone and local systems that produce texts with significant spelling variations and inconsistencies in the same Arabic name. In terms of definition, transliteration is the transcription or Romanization of a word, a phrase or a text written in one writing system into another writing system. Personal names, locations and organizations are among a variety of items subjected to the process of transliterations.

There are several factors explaining the divergent methods of Romanization and transliteration from Arabic. According to Heather Hedden (2007), transliteration or Romanization refers to any rendering of words in non-Latin writing systems into languages using the Latin alphabet. Further, transliteration refers more specifically to a precise system of mapping one writing system to another, often letter by letter, so that the transliterated word becomes similar to its equivalent in the original language in terms of spelling and pronunciation. In Hedden's view, there are numerous ways to both Romanize and transliterate Arabic and although transliteration standards exist, they are not applied consistently.

Furthermore, transliteration is a two-fold process "forward / backward "formulating a representation of lexical items " words, phrases, etc"- in one language- using the alphabet of another language. Forward transliteration involves the transformation of source language (SL) into target language(TL) approximation whereas backward transliteration - the reverse process- is the reproduction of target language approximations backward into their original source language. Transliteration, in its double form, particularly backward transliteration, is a complicated process due to lack of direct correspondence between the phonetic systems of different language. As a whole, the transliteration mechanism is controlled by factors integralto the relationship between (SL) and (TL) as well as the education and experience of the person involved in the transliteration procedures. Single lexical items may have more one potential transliteration, however, variations often occur in the transliterations of proper names because they are considered as culture-specific items which constitute a high productive word class with many unknown words (Abdul-Jaleel and Larkey, 2003; Vigra and Khudanpur, 2003).

Variations in the transliteration of proper names may be due, as well, to the fact that the owners of these names may take certain liberties exploiting the spelling and pronunciation of their names to fulfill aesthetic or dubious purposes. In this context, transliteration of proper names poses challenge and consequently, solutions for such a problem may be functional in a variety of fields including tourism, security, immigration, border control, intelligence networks and counter-terrorism activities. On this basis, it is relevant to argue that the transliteration of proper names(PN) is a significant process because they constitute an extremely important component in cross language information retrieval (CLIR). Further, personal names (anthroponyms) constitute a part of the proper names and form a specific group within the vocabulary of a language. Therefore, personal names are subjected to the phonological, syntactic, semantic and orthographic rules of the language. Moreover, sources of proper names vary in quality particularly for language pairs, with different alphabets, like Arabic and English in which there is a wide variation in how Arabic names are rendered into English.

In a related context, Alanzi (2004 : 165) argues that huge amounts of non-standard Arabic database of Romanized / transliterated names exist that are in use in many private and government agencies as in passport name holders database, phone directories and geographic name database. The processing of Romanized / transliterated Arabic names is associated with several complications because transliteration is not

based on standard Arabic names. The transliterated names, generated and deployed in private and governmental agencies are problematic because such databases are inconsistent (Alanzi:166). Obviously, the inconsistent representation of transliterated Arabic names leads to the paralysis of many serious uses of these databases.

Recent studies argue that any transliteration system for Arabic has to make a number of decisions dependent on its intended field of application. The most famous applications in this connection are the two areas of Machine Translation (MT) and Cross- Lingual Information Retrieval (CLIR). Nevertheless, many problems hinder the developments of a Machine Translation module as a result of differences between vowels and consonants in Arabic and English. For example, there is no equivalent in English for some Arabic consonants like (ع),(غ), (خ), (ض) and (ح) whereas Arabic does not have English consonants like (P).

Obviously, the dissimilarity in the alphabetical systems leads to ambiguities and complications in the process of transliteration. Although the glottal stop in Arabic الهمزة exists in some spoken variety of English – replacing the "t" sound in "butter" for example- English does not have a character for this sound. Further, the same sound الهمزهcreates several problems in the transliteration of names like دعاء, هناء, وفاء, بهاء, نداء, ألاء, علاء.Moreover, the data contained in unvocalized Arabic writing is not sufficient to give non- Arab readers adequate and accurate information for correct pronunciation, and subsequently, transliteration becomes confused and fragmented. Moreover, diacritics " فتحة fatha, ضمة damma, كسرةkasra, شدة shaddah,سكونsukun " and other short vowels are considered as one of the major sources of ambiguity when dealing with Arabic proper nouns transliteration. Even native speakers of Arabic are completely embarrassed when tackling diacritics particularly when they are omitted from Arabic script.

Discussing the automatic transliteration of proper names / nouns from Arabic to English, Mehdi Kashani, Fred Popowich and Fatiha Sadat argue that the transliteration task becomes more challenging when the language pair, under consideration, uses different orthographies like Arabic and English (cited in Nasim 2009 : 37). The transliteration from Arabic to languages with different alphabets is always a source of recurrent embarrassment and confusion according to the study of Alonaizan and Knight (2002). Using Buckwalter Encoding System, a well known web search, on Google for texts on the Libyan president معمر القذافى "Muammar Alqathafi" reveals a tremendous number of variations in the spelling of the name.On this basis, an adequate transliteration dynamics should be developed in order to reproduce names as they appear in the source language. Transliteration should be conducted according to the rules, spelling, phonology, and morphology of the source language (SL) notthe target language (TL). For example several Arabic names are transliterated according to the pronunciation of non-Arab speakers like “Ahmed” instead of أحمد "Ahmad" or “Mecca” instead of مكة "Makkah" or “Emir” instead of أمير "Amir". Conventional proper names like “Muhammad” are transliterated according to non-native pronunciation rules.

There is no doubt that major problems stem from the fact that many transliterations are conducted according to the spelling of the word rather than its pronunciation. Even automatic transliteration systems and software fail to provide accurate and reliable results. For instance, Zhao et al (2007) performed an experiment utilizing the Moses SMT system to measure the accuracy of transliterating Arabic text into English via electronic apparatus. Unfortunately, the highest rate of forward translation accuracy – conducted through Moses SMT- was 43% only. Koehn et al (2007) claims that Moses SMT, a statistical Machine Translation System which allows individuals to automatically train transliteration models for any language, is a collection of translated texts "parallel corpus". Furthermore , machine readable dictionaries are not available for all language pairs and they do not contain names or technical terms or acronyms.

In other words, Language Information Retrieval Systems (LIRS) are not able to identify words which are proper names. Besides, their transliterations may not reflect current usage of the names. Therefore, transliteration through electronic or mechanical systems, like المسبارAlmisbar, عجيبAjeeb, ترجم Tarjem and others, leadsto misrepresentation of Arabic names. Due to the multiple spellings of the samename in Arabic texts, electronic devices with limited database fail to provide correct and accurate transliteration for Arabic names transformed into the Roman script. Nevertheless, there were attempts to design and develop adequate transliteration apparatuses to transform Arabic texts into English. For example, Stalls and Knight (1998) presented Arabic - to - English back-transliteration system based on the source-channel framework.

The transliteration process is based on a generative model of how an English name is transliterated into other languages.

Further, Arbabi and Fischtal (1994) developed a hybrid algorithm aiming to automate the process of manual transliteration of Arabic names, which is slow, laborious, error-prone and time consuming. The system uses neutral networks and a knowledge-based apparatus to vowelize Arabic words and proper names. One of the obstacles that confronts this system is rooted in the process of Arabic names transliteration into Roman alphabet. Arabic, like other Semitic languages, poses an additional challenge because Arabic words are written without short vowels (diacritics). On this basis, Arbabi and Fischtal's system includes a vowelization process in which the appropriate short vowels are inserted into the unvowelized input names. The network filters out unreliable names - whose pronunciation is not identified- passing the reliable names on to the knowledge-based system for transliteration or Romanization. This approach, developed at IBM Federal Systems Company, is applicable to a wide variety of purposes including visa and document processing in different destinations and across the borders.

For a variety of reasons, Arabic names are not correctly transliterated into English. Problems occur when names are transliterated rather than transliterated through electronic / mechanical systems. For example the word مصر is written "msr" in the Buckwalter Encoding System while it is pronounced "mesr". The variation between written / transliterated and pronounced versions is misleading for the non-Arab reader. Further, the omission of diacritics in Arabic texts is problematic particularly if lexical items are processed through an electronic apparatus. For example, the name "مجدي" is translated in Buckwalter as "mjdy" instead of "majdy" and "قطر", an Arabic country, is translated in Buckwalter as "Qtr" instead of "Qatar". Variations of transliteration constitute one of the major problems rooted in Arabic proper names. An Arabic name can be written in English with many different spellings. For example, the proper name صديق –the given name of the paper's author- is written only in one way in Arabic but in English the name appears in more than 25different forms like "Sadik, Saddiq, Saddeque, Saddeek, Sedeek..etc". Moreover, Arbabi and Fischtal (1994) point out that while the Arabic proper name سليمانhasonly one form in Arabic, it is written in English, in as many ways as forty different ways, such as "Solayman, Seleuman, Solomen, Suleiman and Sylayman …etc." In a similar context, Brian Whitaker (2002) demonstrates that the major problem with names in Arabic–English Information Retrieval is the great variability in spelling. Whitaker identifies thirty two different English spellings for the name of the Libyan president معمر القذافي"Muammar Alqathafi " and four spellings for القاعدة Al-Qaeda.

It is well-known that Arabic text is written from right to left using diacritics which denote short vowels and germinated consonants. Non-diacritization poses a major problem during the transliteration of Arabic texts because the readers have to substitute for the missing diacritics. Explicitly, variations in the use of diacritical marks, which are integral to the identification of Arabic names, lead to differences in transliteration. The omission of diacritics from hand written and pronounced Arabic texts leads to drastic complications in the process of transliteration. Other problems, peculiar to Arabic script, spring from variations in spelling, dialects and syntax.

Besides, the Arabic language is spoken in a vast geographic area from the Atlantic Ocean to the Persian Gulf encompassing more than 250 million people and a majority population in twenty countries. Though the written script is consistent, in all countries, significant dialectical variations occur. Further, differences between classical Arabic pronunciation and local dialects bring about different transliterations depending on the origin of the Romanized version. For example, names of places and individuals are pronounced/transliterated in different ways. Some common given names like "علي أو فاطمة"- Ali or Fatima - are pronounced/transliterated according to fixed standards while others are transliterated in terms of local dialects. A classical Arabic name like قذافي – Qathafi- which stems from the Arabic root قَذف - Qathafa - is pronounced/transliterated as "Gaddafi" in Libyan local dialect. Thus, proper names introduced into English through the press and media are misleading because they tend to assume a more popular transliteration: جدافي"Gaddafi" instead of قذافي "Qathafi".

Besides, in classical Arabic a name like "Jamal" - جمال - is transliterated as "Gamal" according to Egyptian local dialect. Arabic names including the letter - "ثاء ث " - like ثروت"Tharwat" or كوثر"Kawthar" are

transliteratedin the light of Egyptian dialect as سروت"Sarwat" and كوسر "Kawsar", thus the ثاء sound is transformed into "سين" which radically alters the meaning of the original name.Arabic names - including "ق"letters / sounds - like قابوس "Qaboos" and القرنى "Alqarni" are pronounced/transliterated as جابوس "Gaboos" and الجرنى "Algarni " in the GCC countries. Apparently the transformation of the letter / sound ق"Q" - قاف- into ج " G " - جيم - in the preceding examples in addition to other variations in dialects bring about different transliterations which is misleading on the part of non-Arab readers.

4.1 The Definite Article ال - "al".

One of the major problems in the transliteration of Arabic names lies in the definite article ال - (al) which appears in Arabic family names. Besides, many Arabic names of people, places and organizations include the definite article as part of the name. The definite articleالaccording to the Chicago manual of style is written in lowercase- unless it comes at the beginning of a sentence and is joined to the following word by a hyphen. As a definite article, it appears in different forms, a hyphenated form (-al-) or non-hyphenated form (al) or as (el) in some dialects in Egypt and Gaza strip. Since (al) is a fossilized part of Arabic proper names, there is controversy about whether it should be integrated into the name or dealt with as a separate entity, a hyphenated lexical item. Differences in opinion lead to variations in spelling which consequently bring about complications in the process of transliteration. People who believe that (al) should not be hyphenated and should be integrated into the names argue that (al) is a bound morpheme and Arabic is a language that blends morphemes.

The phonetic variations ofthe definite articleال (al) in Arabic script and pronunciation problematizes the issue of transliteration. In the GCC countries, for example, the definite articleال is pronounced (al) as in family/tribal names like النهيان"Alnaheyan" or القاسمى"Alqasemi", however, in the Egyptian dialect it is pronounced إل (el) as in family names like المنوفى "Elmunofi" or المصرى "Elmasri" in addition to names of geographical locations and cities like شرم الشيخ "Sharm el Sheikh" and المنيا "Elmenya".More complications germinate from the fact that the definite articleال(al) does not appear in name sorting. For example the name محمود العربى " Mahmud Alarabi" is sorted under عربى محمود"Arabi Mahmud". However, when the ال (al) falls into the middle of the name it is taken into consideration in the process of sorting. The Arabic definite article الis not ignored either when it becomes part of famous names like الغزالى "Alghazali"or المنفلوطى "Almanfalouti". Further, since the definite articleال contains the Arabic letter ل- اللام-(L) it becomes more complicated because- اللام- (L) in Arabic is omitted from pronunciation when followed by certain letters. In Arabic this ل (L) sound is called اللام الشمسية "L- ashamsiyya" using the word الشمس–ashams- which means (sun) as an example whereل (L) sound is eliminated during pronunciation though it appears in the Arabic script . Another example could be found in the family name of the ex-Egyptian president أنور السادات which contains the ل (L) sound in the script. Nevertheless, the word السادات is pronounced as "assadat", thus the sound / لاللام الشمسية is omitted though it appears in the written form of the name.Moreover when the Arabic ل- اللام- (L) sound appears in both written script and pronunciation it is called اللام القمرية " L- alqamariyya" since the sound ل- اللام- (L) appears in both script and pronunciation in the word القمر - alqamar- which means (moon) in Arabic. To conclude, when the ل- اللام- (L) sound appears in script and is eliminated in the pronunciation of words like الرياض"arriyad", it is called اللام الشمسية"L- ashamsiyya" . However, when the ل- اللام- (L) sound appears in script and is not eliminated from pronunciation as in words like البصرة"Albasra" or القاهرة "alqahera" – Cairo - it is identified as اللام القمرية" "L-alqamariyya".

5.1 The Colonial Legacy

Due to the impact of the colonial legacy and the interference of colonial languages with Arabic language, a variety of different transliteration systems were developed inArab countriesin North Africa and the Middle East. Further, divergent spellings and pronunciations are related to the dialect of origin or inconsistent applications of the language. For example transliteration systems in North African countries and Lebanon were influenced by the language of the French colonizers while the transliteration systems in the other Arab countries came under the impact of the English language. In Arabic speaking countries like Tunisia, Algeria and Morocco as well as Lebanon, where the French colonial legacy has left its impact on native language, the transliteration systems were influenced by the way the French colonizers reproduce and articulate Arabic language. In other words the transliteration systems were undoubtedly influenced by the ramifications and

consequences mandating from the interference of the colonial language in Arabic phonology. Thus, Arabic names like بشير ، وسيم which used to be transliterated as " Bashir, Wasim " respectively, are Romanized as " Bachir, Oussim" corresponding more closely to French pronunciation.

6.1 Arabic-English Phonetic Systems

Variations in the phonetic systems of Arabic and English leads to complications in the process of transliteration. For example Arabic has two types of vowels, according to Hayaat Alyaqout (2005): tense vowels and lax vowels. Tense vowels can be combined to create new sets of vowels including lax vowels. Further, lax vowels are mild versions of the three major tense vowels – "aah, oo, ee ". These lax vowels حركات- diacritical marks - are as follows: فتحة, ضمة, كسرة, plus some other representations. In other words, lax vowels are represented by diacritical marks placed on the targeted letters. Though the diacritics are often omitted and are not written explicitly in all texts, the reader is expected to assume the right diacritics. Since lax vowels have the power to impact tense vowels formulating new pronunciations of these vowels, they complicate the process of transliteration. In addition to the ambiguity of Arabic short vowels – a, i, u – and the difficulty of long vowels- aah, oo ee-which appear in written Arabic discourse, complicated compound letter sequence (like ظ as in " أبو ظبى" -AbuDhabi- or خ as in خمار "khemar" which means - veil- or ث and ذ as in أثاث - furniture - and ذيل- tail - respectively) are also problematic in terms of transliteration.

7.1 Recommendations

Explicitly, current transliteration systems, to some extent, produce unreliable and non-adequate results because most of their database does not contain correct Arabic names pronunciation. Therefore, the lack of accuracy in the electronic and non-electronic transliteration mechanisms and facilities introduced by language specialists and lexicographers and the dramatic ramifications of the use of disqualified transliteration apparatuses, on world safety and international security, underline the necessity of developing a new transliteration dynamics, with a comprehensive database able to retrieve Arabic proper names and related material in a correct and transparent manner.

Works Cited

Abdul-Jaleel, N. and Larkey, L. (2003). Statistical transliteration for English-Arabic Cross Language Information Retrieval. *Proceedings of CIKM*. 139-146.

Alanzi, Fawaz, S. (2004). Producing Algorithmically Standard Romanization of Arabic Names Using Hints from Non-Standards. *International Journal of Computer Processing of Oriental Languages*, 17 (3): 165-180.

Alonaizan, Y. and Knight, K. (2002). Machine Transliteration of Names in Arabic Text. *Proceedings of the ACL-02 Workshop on Computational Approaches to Semitic Languages*. 1-13.

Alyaqout, Hayat (2005). 1-Mag Scheme for Transcribing Arabic Phonemes. 1-MAQ, www.1-mag.org

Arbabi, M, Fischtal, S.M. (1994). Algorithms for Arabic Name Transliteration. *IBM Journal of Research & Development*, 38 (2): 1-11.

Darwish, K, Doermann, D., Jones, R., Oard, D., and Rautiainen, M. (2001). TREC-10 Experiments at Maryland: CLIR and Video. *In TREC 2001. Gaithersburg: NIST*, 549-561.

Hedden, Heather (2007). Arabic Names. *Centerpiece to the Index*, 25 (3): C9-C15.

Koehn, P., Och, F. and Marcu, D. (2007). Statistical Phrase-Based Translation. *Proceedings of the 2007 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology*, 1: 48-54.

Knight, K. and Graehl, J. (1997). Machine Transliteration. *Computational Linguistics*, 24 (4): 599-612.

Nasim, T. (2009) . *Transliteration : An Introduction*. Cairo : Wataniya Press.

Newmark, Peter (1988). *A Textbook of Translation*. London: Prentice Hall.

Stalls, B. and Knight, K. (1998). Translating Names and Technical Terms in Arabic Text. *Proceedings of the COLING/ACL Workshop on Computational Approaches to Semitic Languages*. 34-41.

Virga, P. and Khudanpur, S. (2003). Transliteration of Proper Names in Cross-Language Applications. *Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. 365-366.

Whitaker, B. (2002). Lost in Translation. *Mirrored fromwww.guardian.co.uk/elsewhere/ journalist/story/0.7792.730805.00.html*

Zhao, B., Bach, N., Lane, I. and Vogel, S. (2007). A log-Linear Block Transliteration Model based on Bi-Stream HMMs. *Proceedings of NAACL HLT*. 364-371.