

Kristen E. Wilson. A Guide to Copy Cataloging Arabic Materials. A Master's Paper for the M.S. in L.S. degree. April, 2005. 48 pages. Advisor: Jerry D. Saye

For most catalogers, non-Roman script languages are more difficult to catalog than those in Roman scripts, and Arabic is particularly problematic. The cataloger must have a firm grasp of the language in order to correctly supply unwritten vowels and to use the standard Arabic-English dictionary which lists words by root rather than alphabetically. This manual presents the cataloger who does not have that language knowledge with strategies for effective copy cataloging searching. Topics include the development of Arabic cataloging automation, problems of name authority, distinguishing between Arabic and other languages written in the Arabic script, and using a non-alphabetic Arabic-to-English dictionary.

Headings:

Arabic literature/Cataloging

Copy cataloging

Foreign language materials/Cataloging

Cataloging/Automation

Transliteration

Arabic language/Transliteration

A Guide to Copy Cataloging Arabic Materials

by
Kristen E. Wilson

A Master's paper submitted to the faculty
of the School of Information and Library Science
of the University of North Carolina at Chapel Hill
in partial fulfillment of the requirements
for the degree of Master of Science in
Library Science.

Chapel Hill, North Carolina

April 2005

Approved by

Jerry D. Saye

Table of Contents

Introduction	2
Definitions	2
Literature Review	
Foreign Language Cataloging	3
Library Automation and Arabic Materials	6
Cataloging and the Arabic Language	10
Language	12
Is it Arabic?	14
Anatomy of the Arabic Book and Sources for Descriptive Cataloging	18
Using the ALA-LC Romanization Table	23
The “Root” of the Problem: Using Dictionaries	25
Determining the Root of a Word	26
A Dictionary Exercise	29
Verb Patterns	31
Arabic and Muslim Names	32
Strategies for Cataloging Copy Searches	33
Conclusion	36
Works Cited	37
Reference Works About Arabic for the Cataloger	39
Appendixes	
A: Christian and Islamic Calendar Months	40
B: Answers to Dictionary Exercise	42
C: Verb Patterns and Conjugation	43
D: OCLC Connexion Record with Parallel Vernacular Fields	45
E: تم بحمد الله Tamm bi-ḥamdu lillāh	46

Introduction

This manual is designed to be a guide to copy cataloging Arabic materials for catalogers and trained paraprofessionals who do not read or speak Arabic. After familiarizing themselves with the Arabic alphabet, they should be able to use the strategies outlined below to locate appropriate records. Because of the special problems of Romanizing Arabic, in addition to the understanding necessary for classification, it is not recommended that original cataloging of Arabic materials be assigned to those without the requisite language knowledge.

Definitions and Understandings

Copy cataloging has been defined as “preparing a bibliographic record by using or adapting a bibliographic record already prepared by someone in another library or organization” (Schultz, 1995). In this guide it is assumed that users will search for preexisting copy that is an exact match (i.e., same edition), although some users may feel comfortable doing close copy, i.e., using an existing record for an older edition as a basis for cataloging a new edition. Those with more language knowledge may use this guide as a jumping-off point to do simple original cataloging, but without a certain level of language competence, it should not be considered a sufficient tool to enable production of accurate original cataloging.

When referring to cataloging in a bibliographic utility, the guide will deal exclusively with OCLC’s Connexion system. OCLC has become the dominant

bibliographic utility, and its Connexion interface will completely replace its Passport interface as of June 5, 2005. Connexion provides greater ease of searching, especially for the special problems of Arabic. Connexion version 1.3, which became available in April 2005, is the first version to provide universal access to Arabic vernacular capabilities.

For improved legibility, Arabic words and letters throughout this guide are represented at a slightly larger point size and italics are not used in the Romanization. The Arabic spelling, ALA/LC Romanization, and meaning of each word or letter is listed a way that one hopes is clear but not too tiresomely repetitive. A few of the diacritics used in ALA/LC Romanization which appear directly above or below the corresponding letter in Connexion, cannot be lined up vertically in this word processing program. In these cases the diacritic is placed after the letter that it modifies.

Literature Review

Foreign Language Cataloging

Cataloging and bibliographic control of foreign-language material has long presented a quandary to cataloging departments, and their managers have attempted a variety of solutions. Some libraries have one or more area specialists highly qualified and knowledgeable about specific countries or regions. Many members of cataloging departments, both professional and paraprofessional, have some language skills, though European languages appear to be the most common, mirroring general trends in foreign language education in the West. Students who know languages that are less common in the West are sometimes employed to do copy cataloging and/or to translate titles and otherwise assist professionals with language issues. Within consortia and through less

formal arrangements, items are sometimes exchanged with another library which has the needed language expertise. Some libraries are also resorting to outsourcing materials to vendors who can provide the needed cataloging/language expertise. Backlogging of foreign-language materials persists as well and items may remain uncataloged for years, if not decades.

In 1992, da Conturbia published the results of her survey of many aspects of ARL libraries' foreign-language material cataloging efforts. She found that a high percentage of both professional and paraprofessional cataloging staff had knowledge of non-English languages, 88.7% and 78.8%, respectively. She did not ask which languages were known by the catalogers, but libraries were asked to report which non-English languages were most strongly represented in their collections and which were projected to be future collection priorities. Spanish, German, and French were the top three responses to both questions; Arabic was chosen by only one library (likely the same one) for both questions. Clearly a fairly high percentage of catalogers with foreign-language knowledge are competent primarily in European languages.

When the need exists, libraries can sometimes create their own language competence. The University of Nevada, Reno (UNR) has the largest collection of Basque materials outside of Spain and France, but had few staff who spoke the language when it was decided to catalog 10,000 items from their backlog (Ugalde and Etcheverria, 1998). They had little hope of hiring Basque-speaking catalogers, so an innovative program was initiated with grant funding for time-limited employees. The project director, 5 staff (professionals and paraprofessionals), and student workers received two weeks of intensive Basque language training, and teachers were available as occasional resources

for later language questions. UNR's Basque Studies staff prepared guides to important terms and the bibliographic peculiarities of Basque books. Additionally they gathered resources such as bibliographies to assist with name and title authorities.

Ugalde and Etcheverria suggested that cataloging skills and the willingness to work on cracking the code of a foreign language without full fluency can be adequate substitutes when deep language knowledge is not available. In fact, bibliographic tools and informal language assistance can be used to catalog books in many languages without any language training at all. Chressanthi (2001) singlehandedly took on an important collection of translations in 21 languages, although she had knowledge of only one, Chinese. In addition to carefully studying Romanization tables and language manuals for librarians, she sought help in deciphering certain words through her university's faculty and students who possessed the requisite language knowledge. She also used the community of catalogers on AUTOCAT, a cataloger's listserv, to get help with specific questions.

Of course, in any library copy cataloging of foreign-language materials is the ideal solution in terms of time and money saved. That is, whenever such cataloging copy can be found. Faced with a book in a non-Roman script, however, a cataloger may blanch at the prospect of searching for copy. Some libraries use students who are native speakers to transliterate title and author information. This enables catalogers to search for cataloging records. If the number of works in a language does not warrant employing a student or other individual with the requisite language knowledge, however, these items tend to be backlogged at a higher rate than English language works (da Conturbia, 1992).

Library Automation and Arabic Materials

Before the age of library automation, vernacular records for non-Roman script materials were the standard. Catalog records were entered on cards and, without the technical problems that would appear in the computer age, descriptive information could be recorded in the original language, even non-Roman languages such as Chinese or Arabic. These cards were often kept in separate catalogs for individual non-Roman languages, or if the choice was made to incorporate them into the general catalog, access points were Romanized to allow for interfiling with Roman-alphabet languages (Vernon, 1991).

When computer cataloging began to develop in the 1970s, libraries in the Arab world and those in the West took different approaches to dealing with Arabic materials. Arab libraries did not adopt MARC cataloging, because MARC could not support the input and display of Arabic script. Instead they automated by “Arabizing” various Western library software systems. Major libraries of the Middle East adapted different systems to include Arabic capability (or even the same systems, in different ways). Various attempts to develop an ARABMARC system, beginning in 1977, have not come to fruition (Khurshid, 1998, 2002).

Libraries in the West, however, initiated the process of automation with the requirements of Western languages in mind. Thus they concentrated on the display of Roman script. Catalog records in non-Roman scripts were square pegs that would not fit in the round holes of early bibliographic utilities. In 1978, the Library of Congress announced that it would consider Romanization of all scripts other than Chinese, Japanese, and Korean (often referred to as CJK). Catalogers of Arabic and Hebrew

expressed an overwhelming lack of support for such a project. Both the Middle East Librarians' Association and associations of librarians of Jewish studies passed resolutions against the practice or voted against supporting it (Vernon, 1991). As time passed, however, libraries became dependent on automation and the lack of electronic access to records for Arabic and Hebrew books became less acceptable. An added attraction of Romanization of records was that it would also allow for easier processing of materials by library staff in various departments who couldn't read the non-Roman scripts. MARC cataloging of the JACKPHY languages (Japanese, Arabic, Chinese, Korean, Persian, Hebrew, and Yiddish) was initiated in 1983. Commercial applications available at the time did not support vernacular script. At this time, the Library of Congress revised its Romanization tables and began developing additional tables for less common languages. Gradually, libraries across the United States began Romanizing these records and integrating these materials into their general catalogs. By the late 1980s, most libraries had eliminated their separate card catalogs for materials in Arabic, Hebraic and Armenian scripts (Riedlmayer, 1999).

Vernon (1991) compared the problems posed by the two options for automated cataloging of Arabic and Hebrew, i.e., Romanization and retention of the vernacular. Romanization posed a theoretical problem, that of "how to represent the languages of another script in a meaningful and efficient way." Retention of the vernacular was a technical problem, that of dealing with a different script, sorting order, and direction (right-to-left as opposed to left-to-right). Advances in dealing with the theoretical problems of Romanization have not advanced significantly since the last revisions of the

ALA/LC Romanization tables, but steps have been taken to address the technical challenges that prohibited vernacular automated cataloging.

In 1991, the Research Libraries Information Network (RLIN) introduced an Arabic-script cataloging system which completed its support of the JACKPHY languages plus those in the Cyrillic script. Records could be input primarily in Arabic, “eliminat[ing] the problems of transliteration” (“LC is First,” 1991). Unfortunately, most library automated catalogs still did not have the capability to display vernacular scripts. Indeed, many do not to this day. Thus most libraries found it necessary to continue to include Romanized descriptive information.

Older Online Public Access Catalogs (OPACs), often cannot even display the diacritic marks used in Romanization. The dots, macrons, inverted hyphens and other marks that serve to distinguish letters and sounds in the original language, and thus vital meaning, may be stripped from the record, leaving the patron, perhaps a native speaker of the language, to puzzle at such an imprecise rendering of the title or term she or he is searching for (Riedlmayer, 1999). Newer OPACs are able to display diacritics, and some have even advanced to vernacular display. We are currently in a period of transition. With such differences in OPACs, the capabilities of libraries to effectively communicate information about their holdings in non-Roman languages vary widely.

In RLIN, catalogers wishing to input both Romanized and Arabic vernacular data input Romanized data in the standard MARC fields, e.g., title and statement of responsibility in 245, edition information in 250, etc. The Arabic script versions are placed in 880 fields and linked to their corresponding Romanized fields with a subfield 6 (Vernon, 1996). This system of enabling “alternative graphic representation” was also

used when OCLC (Online Computer Library Center) began its Arabic Cataloging Pilot Project in 1999. Catalogers could use the other major bibliographic utility to search, copy, and add records containing Arabic script and choose whether to export Romanized, Arabic or both scripts when exporting to their own system. In April 2001 OCLC's WorldCat contained 22,297 records with Arabic vernacular (Khurshid, 2002). By 2005 OCLC reported the number as "approximately 50,000" on the Web site for its Arabic Cataloging software, although more than 250,000 records for Arabic material are available ("Arabic Cataloging," 2004).

As of July 1, 2005, OCLC will discontinue use of its Arabic Cataloging software and its previous interface, Passport, in favor of its new Connexion Client, which has been available as an alternative for several years. Connexion 1.3, which became available in April 2005, is the first version to integrate Arabic vernacular capability and provides for parallel, linked fields. A record may have two 100 fields for author main entry, one Romanized and one vernacular, two 245s, etc. It also provides a feature allowing automated transliteration. With a couple of clicks of the mouse, the cataloger can generate a parallel Arabic field from the Romanized field. The reverse cannot be done without human input, however, due to the necessity of providing unwritten vowels. This will be further discussed in the next section.

It seems unlikely that American libraries will give up Romanization of descriptive information from non-Roman scripts in the near future, even if they enthusiastically embrace the addition of vernacular display as it is supported by bibliographic utilities and OPACs. The majority of users are most comfortable with Roman letters, and materials in non-Roman languages can be more easily processed by non-specialist library staff when

records are Romanized. Israeli libraries, which exist in a multi-script, bi-directional environment, present a different example. Research libraries in that country have significant holdings in Hebrew, Arabic, and English, and for a long time kept records for different languages in separate catalogs. Current OPACs allow this separate filing to be transparent to the user, however. An examination of the online catalog of Ben-Gurion University of the Negev revealed monolingual, monoscript records in all three languages. An exception was that Arabic and English records contained Library of Congress Subject Headings, while Hebrew records contained Hebrew Subject Headings.

Cataloging and the Arabic Language

In an article on the description of catalogers' foreign language skills, Beall (1991) stated that "[o]f the three main activities involved in cataloging – description, subject analysis, and classification – each requires a different level of language proficiency." He stated further that of the three, descriptive cataloging required the lowest degree of language knowledge because "description involves transferring specific information from an item into a bibliographic record; subject analysis involves generating new information which describes the content of a work."

While it is certainly true that it is difficult to classify or generate subject headings for materials in a language in which one is not competent, the degree of proficiency needed for these operations is exactly the opposite in Arabic. In describing an Arabic book, title and statement of responsibility, as well as edition and series information, must be Romanized for entry in the bibliographic utility.

Vernon (1996) stated that in the case of Arabic and Hebrew, Romanization is not the same as transliteration. The latter implies "simple letter-by-letter substitution." For

these languages, however, in which most vowels are not written but understood by the reader according to the grammar and context of the word, catalogers must supply vowels, and do so from a sophisticated understanding of the vocabulary and grammar. While a less knowledgeable reader may be able to get the gist of the subject of a book, accurate descriptive cataloging proves impossible for the cataloger who does not know the language well. Thus Arabic is a language which presents very specific obstacles to cataloging by persons who do not know the language.

Given Romanization in an automated environment, the problems of cataloging Arabic materials in a Western library can be divided into three major areas: language, layout, and culture. The first challenge to a cataloger when faced with an item in Arabic script is to determine which language the book is in. Among the languages that use Arabic script are Arabic, Farsi (Persian), Urdu, Ottoman Turkish (now extinct), Kurdish, and Pushto. In any of these languages, letters of the Arabic alphabet can be represented in four or more different forms, depending on their position in a word. The tradition of calligraphy, which often appears on title pages of books, can make even the order of letters in a word difficult to determine. Accurate Romanization, which will be discussed further, requires a certain level of competency in the language.

Arabic books are laid out more or less like Western books, but a higher percentage have incomplete publishing data, in the form of missing publisher, place, and date information (Khurshid, 2002). Sometimes it may be difficult to find because it is located in the back of the book, or in other unexpected places. Contemporary books usually include a publishing year in the Gregorian calendar, but many also include the year according to the Islamic calendar. The Islamic year is made up of lunar months and

is shorter than, and thus not easily convertible to, the Gregorian calendar. Arabic and Muslim names can be composed of as many as seven different elements, and their presence in authority files is spotty. Finally, the Library of Congress Subject Headings and the LC classification system reflect a Western bias when attempts are made to shoehorn Arab materials into them. The establishment of the KBP schedule for Islamic law is a welcome development, but the capability to classify and assign subject headings to certain materials is frankly inadequate (Jajko, 1993; Khurshid, 2002).

Language

The Arabic language is written with an alphabet of 28 letters plus diacritical marks that may or may not be used in printed texts, depending on the style of the book and its typography. The language is read from right to left, with the exception of numerals, which are read from left to right within the general right-to-left flow:

←————— →←—————

هذه جملة فيها الرقم 2005 و كلمات بعده. This is a sentence with the number 2005 in it and words after the number.

In North Africa, Arabic numerals (0, 1, 2, etc.) are used, while in Arab countries in West Asia and the Arabian peninsula a set of numerals adapted from an Indian numeric system are more commonly used:

10	9	8	7	6	5	4	3	2	1	0
١٠	٩	٨	٧	٦	٥	٤	٣	٢	١	٠

Books written in Arabic begin from what, to English speakers, is the back of the book – that is to say, the spine is on the right when the book is closed and face up.

When writing Arabic, most letters are connected to adjacent letters in individual words, just as in cursive writing in the Roman script languages. The different forms for

initial, medial, final, and lone letters are shown in the ALA/LC Romanization table for Arabic. Only a couple of ligatures, or specially joined forms, are required in written and printed Arabic. Whenever a lām (ل, initial ل) is followed by an alif (ا) the two appear together as لا instead of a shape like a capital U. The word meaning God, Allāh, has a special form that requires diacritic marks. Typing the word in publishing or word processing software automatically generates this form: الله.

While the Arabic alphabet can be studied and learned without too much trouble, several distinguishing characteristics make Arabic a difficult language to interpret without knowledge of the language. First, only long vowels are written in Arabic words; short vowels can be represented with diacritical marks but these are usually omitted in general texts. They are usually included in religious works such as the Qurʻān, Islam’s holy book, and collections of Hadith (sayings of the prophet Muhammad), some poetry, and in books for children (Figures 1, 2, and 3).

Second, most Arabic words are based on three-consonant roots; these roots are laid into a pattern of vowels and additional consonants to impart a specific meaning. The standard Arabic-to-English dictionary lists words by their roots rather than by

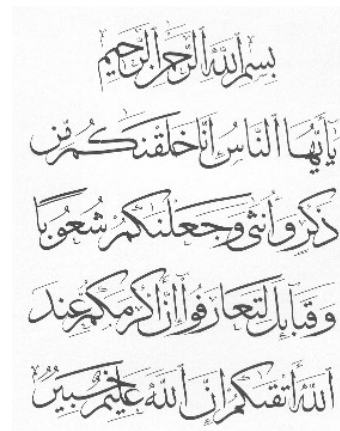


Fig. 1: A passage from the Qurʻān, Islam’s holy book, written in calligraphy with all diacritics.

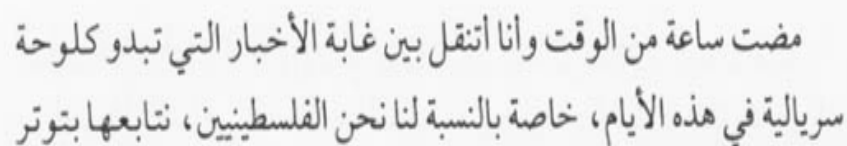


Fig. 2: Typical printed Arabic without diacritics.

their spelling, so one must be able to determine the root to look up a word. For instance,

the word انتخابات (intikhābāt, elections) is listed under the root ن خ ب or نخب (nakhaba, to select or choose). Later in the guide this subject will be addressed and tips for guessing the root of a word will be given.

Third, one must use both grammatical knowledge and context to determine which unwritten vowels are

present and thus, how the

word is pronounced and

transliterated. The word

تذكر may be Romanized

tadhakkara (he

remembered) or tadhakkur (remembering). However, only those Arabic letters

transliterated as t, dh, k, and r are present in the word. To Romanize the word one must

understand what the word is from its context and supply the appropriate vowel sounds as

well as, in the case of the k sound, the doubling.

All this sounds quite discouraging, and there is no denying that Arabic is a difficult language. However, with some tips a non-Arabic speaking cataloger should be able to locate a fair amount of cataloging copy and process many Arabic books.

Is it Arabic?

Before attempting to catalog an Arabic book, one must make sure that it is in fact written in Arabic. There are a number of other languages that are written in the Arabic alphabet at present. Others have used the alphabet in the past. One is most likely to come across books in Persian (Farsi) from Iran, Urdu from Pakistan and India, or Ottoman

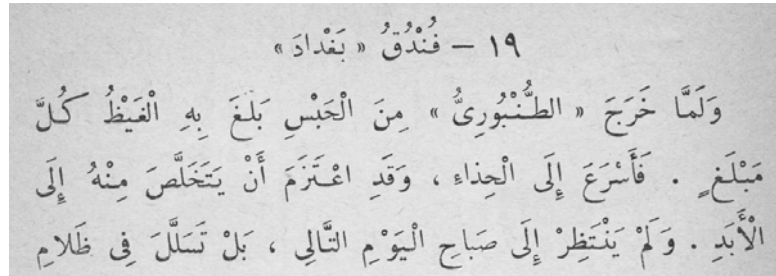


Fig. 3: Lines from a children's book, with full vocalization (diacritics for vowels and doubling).

Turkish (written in Arabic script until 1925). Below is a fuller list of languages in Arabic script and the countries where they are spoken:

Language	Countries
Arabic	Algeria, Bahrain, Central African Republic, Chad, Comoros, Djibouti, Egypt, Iran, Iraq, Israel, Jordan, Kuwait, Lebanon, Libya, (Maldives), Mauritania, Morocco, Oman, Niger, Qatar, Saudi Arabia, (Somalia), Sudan, Syria, Tunisia, United Arab Emirates, Western Sahara, Yemen
Azerbaijani (Azeri)	Azerbaijan, parts of Iran
Hausa	Nigeria, Niger, parts of Benin, Burkina Faso, Cameroon, Central African Republic, Chad, Congo, Eritrea, Germany, Ghana, Sudan and Togo <i>(Primarily written in Roman letters since the 1950s)</i>
Kashmiri	India, Pakistan
Kazakh	Kazakhstan, parts of Afghanistan, China, Iran, Mongolia, Russia, and Turkey. <i>(Written in Cyrillic and Roman script under the Soviet Union, use of Arabic script is re-emerging)</i>
Kirghiz (Kyrgyz)	Kyrgyzstan and parts of Afghanistan, China, Kazakhstan, Tajikistan, Turkey, and Uzbekistan. <i>(Written in Cyrillic and Roman script during most of the twentieth century, use of Arabic script is re-emerging)</i>
Kurdish	Parts of Iraq, Turkey, Iran, Syria, Lebanon, Armenia, Georgia, Kyrgyzstan, Azerbaijan, Kazakstan and Afghanistan <i>(Mostly written in Roman letters in Turkey and in Cyrillic in parts of the former Soviet Union)</i>
Pashto (Pushtu)	Afghanistan and parts of Pakistan
Persian (Farsi)	Iran, parts of Afghanistan or Tajikistan
Qashqai	Parts of Iran

Uighur (Uyghur)	Parts of China, Afghanistan, India, Indonesia, Kazakhstan, Kyrgyzstan, Mongolia, Pakistan, Tajikistan, and Uzbekistan
Urdu	Pakistan, parts of Afghanistan, Bangladesh, India, Nepal
Uzbek	Uzbekistan, parts of Afghanistan

One good way to determine whether a book is in one of these languages is to look for the country or city of publication; however if this is not clear or conclusive there are clues in the languages that can help the cataloger distinguish between them.

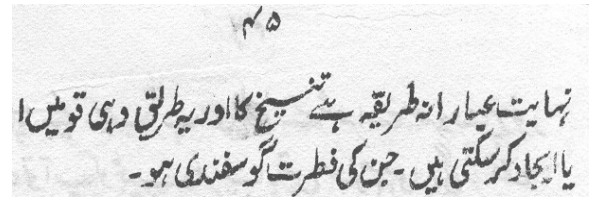


Fig. 4: Persian text in Nastaliq script. Page number in variant style (top) is 35.

First, one can look at page numbers and other instances of numerical data to look for a variation on Indic numbers used in much of the Middle East. In Iran, Pakistan, and India, digits look like those used in the eastern Mediterranean except for the numbers 4, 5, and 6 (see also the variant 3 in Figure 4):

Arabic	(found in Arabic books)	10	9	8	7	6	5	4	3	2	1	0
Indic	(found in Arabic books)	۱۰	۹	۸	۷	۶	۵	۴	۳	۲	۱	۰
Central Asian	(not found in Arabic books)	۱۰	۹	۸	۷	۶	۵	۴	۳	۲	۱	۰

Books printed in many languages using Arabic script tend to use a more elaborate looking style than do most Arabic books. Persian and Urdu are usually printed in the Nastaliq script (Figures 4 and 5).

Languages other than Arabic written in the Arabic script also include letters that have been modified to

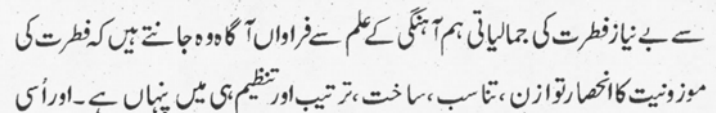


Fig. 5: Urdu text in Nastaliq script.

represent sounds not present in classical Arabic. These are generally based on an Arabic letter with a related sound, with diacritics added. For space and applicability considerations, not included in this list are letters in languages not currently written in Arabic script and letters that are assigned different sounds in other languages but look the same, which will not help anyone who does not speak the language. All of these extra-Arabic letters behave like the Arabic letterforms they are based on when written in different positions in a word.

Letter	Sound	Arabic letterform	Languages found in
گ	hard g	ك (kāf)	Kashmiri, Kurdish, Ottoman Turkish, Persian, Uighur, Urdu
پ	p	ب (bā')	Kashmiri, Kurdish, Ottoman Turkish, Persian, Uighur, Urdu
ڤ	v	ف (f ā')	Kurdish, Persian
ژ	like French j in "Jacques"	ز (zā')	Kashmiri (ts sound), Kurdish, Ottoman, Turkish (hard j sound), Urdu
ٹ	retroflex/hard t	ت (tā')	Kashmiri, Urdu
ڈ	retroflex/hard d	د (dāl)	Kashmiri, Urdu
ڑ	retroflex/hard r	ر (rā')	Kashmiri, Urdu
ټ	retroflex t	ت (tā')	Pashto
ډ	retroflex d	د (dāl)	Pashto
ړ	retroflex r	ر (rā')	Pashto
څ	ts in "lets"	ح (hā')	Pashto
چ	ch in "chin"	ح (hā')	Kashmiri, Kurdish, Persian, Ottoman, Turkish, Uighur, Urdu
ڱ	nya, ng	ك (kāf)	Ottoman Turkish, Uighur

Anatomy of the Arabic Book and Sources for Descriptive Cataloging

As mentioned earlier, Arabic reads from right to left and the component parts of Arabic books tend to be arranged in the same order as in English-language books, but of course in the opposite direction. The *Anglo-American Cataloging Rules* defines “recto” as

1. The right-hand page of a book, usually bearing an odd page number.
2. The side of the printed sheet intended to be read first.

The page intended to be read first in an Arabic book is that on the left, and it also bears odd page numbers, so we can safely define the left-hand page in an Arabic book as recto, and verso as its opposite. Half-title and title pages are usually found on recto (left-hand) pages in Arabic books. An important exception to this arrangement is the fact that tables of contents are more often placed in the back matter of Arabic books.

Title and statement of responsibility information is usually found on the title page, although occasionally there is only a half-title page. In this instance the statement of responsibility is usually taken from the cover. Arabic title pages sometimes include publisher information; less frequently they include the place of publication. Rarely do they provide a date. The title page verso tends to be an important source of information for cataloging. Indeed, more information has come to be included on it in recent years. Many contemporary books include some form of an “all rights reserved to the author” notice on this page, though Middle Eastern countries do not usually have strong copyright legislation. Rarely will one see the copyright symbol ©.

An important thing to look for on the title page recto is الطبعة الأولى (al-ṭab‘ah al-’ulá), which literally means “the first printing.” The word for printing in Arabic is not

used as narrowly as it is in Western publishing, however, and can be taken to mean publication. Library of Congress practice is to use this term as an edition statement, and to put “al- Ṭab‘ah 1.” in the 250 field. The word meaning “first” is abbreviated “1.” following AACR2 rule C.3B1.

The date of publication often appears immediately after (to the left of) this edition statement, for example:

٢٠٠٤ الطبعة الأولى (First edition 2004).

Sometimes only a probable date of publication can be determined based on its inclusion in a publisher-specific serial number for the book. For example:

٢٠٠٤/ ٨١٤٦٩٠٥

In this case the year would be recorded within square brackets in MARC Field 260, subfield c. The date is always recorded in Arabic (i.e. Western) numerals, in this example it would appear as [2004].

Sometimes the Islamic year is given in addition to, or instead of, the Gregorian or Julian year. The Gregorian year will usually be indicated with a letter mīm (م), which stands for Mīlādī, or Christian. The Islamic year will be indicated with a hā’ (هـ), which stands for Hijrī, the term used for the Islamic calendar. If only the Islamic year is given, it is recorded in MARC Field 260, subfield c, with the equivalent year or years added and enclosed within brackets. An Islamic year is shorter and thus does not match the Western calendar. Unless the month of publication is given one cannot be sure of which Western year it corresponds to. In such cases, both possible years should be recorded. For example:

\$c 1423 [2002 or 2003].

There are several calendar converters available on the World Wide Web. A good, simple one appears in the reference works about Arabic at the end of this guide. Among newly published books, it is more common for both Western and Islamic dates to be given, rather than only the Islamic date. When this occurs, the Western date is recorded in MARC Field 260, subfield c and a MARC Field 500 note can be included to note the Islamic year, for instance: “Islamic date of publication: 1423.” Month of publication may be listed and may help clarify which Gregorian calendar year is the equivalent. The eastern Mediterranean countries use different names for the months than do the North African ones, where they use names derived from the old Roman names or the French versions of the months (see Appendix A).

Occasionally, title page verso information is handily listed and labeled, as in the example in Figure 6.

واقعة الإسكندرية	اسم الكتاب
عاصم محروس عبد المطلب	اسم المؤلف
٢٠٠٤/٧٤٨٩	رقم الإيداع
I.S.B.N 977-6015-40-9	الترقيم الدولي
الأولى	الطبعة
مكتبة بلستانج المعرفة	الناشر
كفر الدوار - الحدائق - ٦٧ ش الحدائق بجوار نقابة التطبيقيين	
تليفون: ٠٤٥/٢٢٢٤٢٢٨ الإسكندرية ٠١٢٣٥٣٤٨١٤	
مطبعة الأمل - العصاره - إسكندرية	الطباعة

Fig. 6: Information from a title page verso.

The items in the right-hand column of Figure 6 say:

اسم الكتاب	ism al-kitāb	the name of the book
إسم المؤلف	ism al-mu'allif	the name of the author
رقم الإيداع	raqm al-'īdā'	the deposit number
الترقيم الدولي	al-tarqīm al-dūwalī	the international numbering [ISBN]
الطبعة	al-ṭab'ah	the edition
الناشر	al-nāshir	the publisher
الطباعة	al-ṭibā'ah	the printer

When the name of the publisher cannot be found on the title page or title page recto, it sometimes appears in a colophon at the end of the book. Note that Arab addresses often begin with the city and country rather than ending with them.

MARC Field 300 physical description information (i.e., page count, illustrations, and dimensions) can be assessed in the same way as for English-language books. It is less common to find more than one form of pagination in an Arabic book, as roman numerals are not used for preliminary matter. Occasionally, however, letters are used, and they will be in the old alphabetical order. In the modern ordering of the alphabet, shown in the ALA/LC Romanization table, letters with similar forms are grouped together. Simply count the pages and list them with an Arabic number before the pagination of the main block of text (i.e. "8, 179 p.").

Bibliographical information often appears in footnotes in nonfiction Arabic books, and bibliographies are becoming more common in modern books. When looking for bibliographies and indexes in Arabic books, the overlapping words for back matter material must be kept in mind. The word فهرس *fihris* (plural فهارس *fahāris*) is best translated as "index," but is also used to denote a table of contents. The table of contents is often called a فهرس الموضوعات *fihris al-mawḍū'āt* or "index of topics." More rarely, it is called المحتويات *al-muḥṭawayāt* or "contents." These parts of the book can

easily be distinguished from true indexes by the fact that the page numbers listed to the left of the topics are in numerical order, and correspond to the order of chapters in the book.

Bibliographies are usually listed under one, or both, of the words مصادر maṣādir and مراجع marāji‘, which mean “sources” and “references.” These pages would be listed in a 504 note. One must watch for separately titled or headed sections of references. Authors will often list non-Arabic sources in a separate section in the original language. This section is often titled مصادر اجنبية maṣādir ajnabīyah, i.e., foreign sources. Less commonly than the footnote/bibliography pattern one may find a section of endnotes, called هوامش hawāmish.

Language information is recorded in the “lang” fixed field. The code for Arabic is “ara.” When applicable, the 041 field, in which information is given about multiple languages, including translations, is used.

Below are some other terms which, even if not used in cataloging, may help you navigate through a book and determine its nature (see also Appendix E):

اطرحة (pl. اطرحات)	uṭruḥah (pl. uṭruḥāt)	dissertation
موسوعة	mawsū‘ah	encyclopedia, thesaurus, comprehensive work
مختارات	mukhtārāt	anthology, selected writings
معجم	mu‘jam	mu‘jam
الإصدار/الطبعة	al-iṣḍār/al-ṭab‘ah	edition
تأليف	ta’līf	composition (by) (may precede author’s name on title page)
مقدمة	muqaddimah	introduction
الأولى (الثاني ...)	al-faṣl al-‘ulá (al-thānī...)	the first (second...) part/chapter/section

الفصل		
الإهداء	al-ihdāʾ	the dedication
الخاتمة	al-khātimah	the conclusion, epilogue
أعمال	aʿmāl	works (i.e., previous works by the author)

Using the ALA-LC Romanization Table

The ALA-LC Romanization table and rules of application for Arabic are 10 pages in length and need to be read and re-read repeatedly by the cataloger attempting to use them. They are available in print, and on the Library of Congress' Web site at <http://www.loc.gov/catdir/cpsa/romanization/arabic.pdf>. The table shows the letters of the Arabic alphabet in their initial, medial, final, and unconnected (alone) forms. For most of the letters the different forms resemble each other and simply change slightly in order to connect to other letters in the word. This is similar to cursive writing in Roman script. All letters can be connected to the letters that precede them in a word (i.e., that appear to the right of the letter in question), but some letters do not connect to the letters that follow them (appear to their left). These are the letters ا, د, ذ, ر, ز, and و. Any letter following one of these letters will appear in its initial or final form, but without a character-sized space. In the examples below, note the forms taken by the letters *following* (to the left of) the non-connecting letters. In تورم tawarrum, the و connects to the ت before it but the ر and the م appear in their standalone forms because the و cannot connect to the ر after it and the ر itself cannot connect to the م.

تورم	tawarrum	rising
درابزین	darābazīn	banisters

يشاهد yushāhid he watches

A few possible appearances of letters do not appear in the table. The sīn (س) is very occasionally represented in printed matter as simply a long line instead of the distinctive three-point letter form. This is almost universal in handwriting but not a normal part of Arabic type fonts, so it is rare in books. There it may appear on a calligraphed title page. The letter shīn (ش) shares the form of sīn with three dots added above. It can also be written as a long line, with the dots retained, but this is exceedingly rare, and again will only appear in calligraphy.

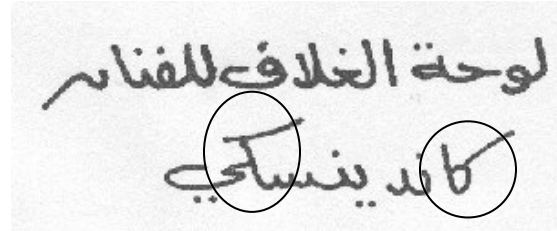


Fig. 7: Note the two kāfs (ك or ك) on the second line. The first (on the right) is followed by an alif. The second looks like a horizontal line over a vertical one.

A kāf (ك) followed by an alif (ا) (كا) sometimes looks like the lam-alif ligature (لا) with the alif sticking straight up. Depending on the elaborateness of the calligraphy, it can be reminiscent of the German *eszet* (ß), but slightly split at the top. This letter form is more common in Persian texts but can appear in Arabic (Figure 7).

The lam-alif ligature mentioned above (لا) can also appear with the hamzah (ء)

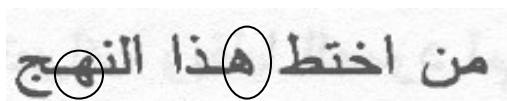


Fig. 8: An initial and a medial ʾ hāʾ, which appear as ھ and ھ respectively.

(glottal stop sound) above or below the alif (لاء, لاء) or with a ~ maddah above (لاء). These alifs would be Romanized as “a,” “i,” and “ā,” respectively.

The hā' (هـ), which appears as a double loop in the medial position (هـ), can also be written as a dagger point dipping below the line in the medial position (Figures 8 and 9). Again, this will most often appear in calligraphed title pages.

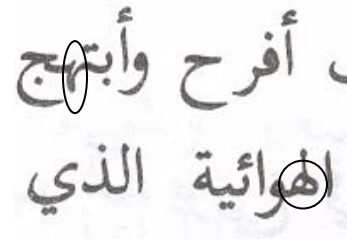


Fig. 9: Two kinds of medial هـ hā' – one with loops above and below (bottom right) and one that looks like a dagger below the line (top left).

The hamzah (ء) is not a letter, but it can be one of the three radicals of a root (see next section). It represents a glottal stop, like the cutting off of air in the middle of “uh-oh.” It may be written by itself on a line or be seated on an alif (ا), wāw (و), or ي, إ, أ : ا/ى/ي.

Once one has a sense of how to Romanize Arabic in order to search for cataloging copy, one next needs to look up words in the dictionary in order to find out which vowels to use in the Romanization. Without the help of a dictionary, choosing the correct unwritten vowels is simply a matter of trial and error.

The “Root” of the Problem: Using Dictionaries

Unfortunately, using an Arabic-to-English dictionary to assist in Romanization (or interpreting the subject matter or terminology for description) is not a straightforward operation, as it would be for a European language. The standard dictionary for cataloging Arabic materials in the U.S. and for Arabic study in general is Hans Wehr’s *A Dictionary of Modern Written Arabic*, third edition. The Wehr dictionary reads from left to right like an English-language book, i.e., when closed and face up, the spine is on the left. The alphabetization is in Arabic letter order (the ALA-LC Romanization table for Arabic is in this order). The cataloger who is new to the Arabic language may find it most helpful to

add labeled tabs to the pages at the beginning of each letter to aid in the navigation of the unfamiliar alphabet, and to provide a quick reference to the alphabetical order.

Unlike most dictionaries, however, Wehr is organized alphabetically not by the spelling of each word, but by the root of the word. This is discussed in more detail in the following section. Briefly, however, each Arabic word has a three, occasionally four, letter root with a basic meaning. To use this dictionary, you must determine the root of the word. This is a difficult task for anyone who doesn't speak the language, but with trial and error and the use of the strategies given below, many words can be found. Caution must be exercised, as Wehr doesn't use exactly the same Romanization scheme as the ALA/LC Romanization. However, it is similar. One important difference is that ALA/LC Romanizes the tā' marbūṭah (ة) as "ah" while Wehr Romanizes it simply as "a."

Determining the Root of a Word

Most Arabic words derive from a three-radical root which expresses a basic meaning (a small percentage have four-radical roots, but let's not address that). Radicals can be any of the letters as well as the ء hamzah. The three radicals are laid onto different patterns with varying vowels and additional consonants. The letters and other marks that can be added to build different forms or derivative words are: ء, ّ (see below), ا, ل, س, ت, ث, ج, م, ن, ه (ة) (see below), و, and ي.

The ّ shaddah is a diacritical mark which indicates a doubled letter. It is not itself a letter and cannot be a radical of a root. The letter ه with two dots above (ة) is called the tā' marbūṭah or "tied-up t" and only appears at the end of a word. It indicates a feminine form or non-human plural. This version is never a radical of a root.

Here are a few examples of words derived from common roots:

From the root ب ت ك :

1. كتب kataba to write (he/it wrote)
2. كاتب kātib writer
3. مكتوب maktūb written
4. مكتبة maktabah library

From the root د ر س :

5. درس darasa to study (he/it studied)
6. درّس darrasa to teach
7. دراسة dirāsah study (noun)
8. مدرسة madrasah school

The examples for numbers one and five are simple. They are the basic form of the root, which is always a verb in the third person past (perfect) tense. This represents the citation form, which in English is the infinitive. When there are only three letters, it's usually easy to guess the root. If one of the root letters appears to be a vowel, however (ا, و, or ي) try looking in the dictionary for both و and ي in that position in the root, as these vowels shift.

Number 2, كاتب, is a noun. The letters shown individually are ك ا ت ب . Examine again the list of letters on the previous page that can be added to a root to derive different words. ك and ب are not on that list, so one knows that they must be the first and last radicals of the root. Instead of looking for ك ا ب , one would look in the dictionary under ك و ب , where there are borrow words such as كوب kūb (cup) but nothing resembling ك ا ت ب and ك ي ب , which does not have an entry. Finally, one would look under ب ت ك , where one would find كاتب in the third column of words under the root كتب, p. 813 of Wehr.

Number 3, مكتوب, is a little different. Again, ك and ب are two of the radicals, but one can't be sure whether ك is the first or second. Therefore the root could be م ك ب, م ك ت ب, or م ب ك. A م often appears at the beginning of an active or passive participle – a “doer” or something acted upon, or a word that describes a place where something is done. So, there's a good chance that in a word beginning with م, the م is not one of the radicals. One would want to try searching for the word under the roots ك ت ب and ب ك ت.

مكتبة, number 4, appears easy to figure out now. The ك, ب, and ة can't be radicals and the م is likely an added letter, so one would look under the root ك ت ب before trying م ك ب or م ك ت.

Number 6, درس, is a verb, and it appears very similar to the basic verb form, درس, with the addition of the doubling diacritic shaddah (ّ). This shaddah will often not be written in, however, so one would have to try to discover from the context which it will be.

Number 7 is دراسة. Of these letters, the only one that one can rule out as a radical of the root is the ة. The possible roots are د ر ا (where the ا is standing in for a و or ي) and د ر س. One should try the non-vowel option, د ر س, first, followed by د ر و and د ر ي.

Finally, number 8, مدرسة. Compare it with number 4, which also begins with a م and ends with a ة. One would want to look for this word under د ر س before trying م د ر. د and ر are not on the list of letters and marks that can be added to derive different words, so they both must be radicals of the root.

A Dictionary Exercise

To learn to navigate a dictionary entry in Wehr, it would be useful to look at a sample, the set of entries for the root ك ت ب ka ta ba, beginning on page 812 of the dictionary. The basic form of the word is كَتَبَ kataba, meaning “he wrote.” The verbal noun that one would translate as the infinitive or a gerund is the maṣḍar, and it is listed in parentheses after the Romanization of the basic form of the verb and a single vowel. In this case the vowel is u, which indicates the vowel used in the second syllable when conjugating the verb in the present tense (for instance, يَكْتُبُ yaktub he writes). The Arabic spelling and Romanization for the maṣḍar and any varying forms are given next. In the case of kataba one will notice that there are three possible maṣḍars which mean “to write” or “writing.” Katb is spelled the same as kataba (كَتَب), so the Arabic is not given again. Kitba and kitāba (kitbah and kitābah in ALA/LC Romanization) are two other possibilities.

After the maṣḍar, the definition or equivalent verbs are given, using the English infinitive. Thus kataba may be translated as “to write” as well as various other related verbs, such as “to inscribe”, “to register”, or “to foreordain” (think of “it is written”). For many words, you may find that the simplest or most commonly used meaning is not the first one given, so skim all the possibilities and think about the context of the word when guessing at the best meaning (Hujelan, 2004). When “ ٥ sth.” or “ ٥ s.o.” is listed in parentheses following a verb or list of verbs, it means that the verb is transitive and may take an object. Prepositions which can, or must, be used for certain meanings are also listed in parentheses.

Within this first entry for a root one will notice Roman numerals, each followed by another set of definitions. These stand for the different patterns on which the root consonants can be overlaid to change the basic meaning. This will be returned to later, but for now notice that *kataba* can be put into patterns I (the basic meaning), II, III, IV, VI, VIII, and X, while *kabala* كبل, in the first column on the same page, can only be patterns I, II, and III. This will vary for each verb.

After the first entry which deals with the verbal forms of the root, there is a page or so of noun, adjective, and adverbial forms derived from *kataba* (the last part of the

كتب entry is a reference at the top of page 814). The first word is كتاب *kitāb*.

Although the first definition given is the most elemental, “a piece of writing,” today the word most often means simply “book.” Notice that the plural form, كتب *kutub* is given immediately after the Romanization of the singular. Arabic often employs “broken plurals,” which are formed by a change in the middle of the word instead of something suffixed at the end, as is most common in English. One will need to look at both singular and plural forms when looking up a word. There are some circumstances under which a plural is regularly changed at the end. If the word ends in a ة, the plural will often be formed by changing the ة into an alif and a tā’, that is from ة to ات. Find the entry for مكتبة *maktabah*. The plural is listed as “-āt” in this case instead of being spelled out. The correct spelling of the plural is مكاتبات *maktabāt*. Masculine human plurals are sometimes formed by adding the suffixes ون -ūn or ين -īn.

Try finding these words in the **كتب** entry in Wehr: **كتبي**, **مكاتب**, and **كاتبات** (answers in Appendix B). For more information on using the dictionary, see its Introduction and list of abbreviations, p. VII-XVII.

Verb Patterns

Now that one has a sense of the layout of the Wehr dictionary and individual entries, let's return to the knotty problem of the different verb patterns. There are 10 possible verb patterns for each root, although no root can be put into all 10. In Wehr these patterns are denoted by Roman numerals in the first section of an entry under a root. The possible verbs are not spelled out in the dictionary; one would have to go to a list of the patterns and then apply conjugation. The patterns in third person masculine singular (he did, he does) are given in Appendix C along with the prefixes and suffixes of conjugation, but it will be easier, if one is really trying to find a verb, to refer to a book which spells out all the possible forms. An excellent work is John Mace's *Arabic Verbs and Essential Grammar*, from the *Teach Yourself* series.

These patterns are usually shown with the root **فعل** fa'ala, i.e., to make or do. The **ف**, **ع**, and **ل** stand in for the three radicals of whichever root you are deriving the verb from. Form II is **فَعَّل** fa''ala. To make the form II of the root **كتب** kataba, we just need to add the **ّ** to double the middle radical: **كَتَّب** kattaba (to make someone write something).

Just one more example before one is tempted to avoid verbs altogether. The last form, X, is the longest. Its pattern is **إِسْتَفْعَل** istaf'ala. If one substitutes the root radicals **ك ت ب** for **ع ل ف** you get **إِسْتَكْتَب** istaktaba. Luckily, most book titles do not contain

verbs, so you shouldn't have to try figuring them out often when Romanizing a few words in order to search for cataloging copy. Many maṣḍars are listed separately in Wehr, but you may want to refer to the maṣḍar patterns in Appendix C of this guide.

Arabic and Muslim Names

Modern Arabic names generally consist of personal names and surnames, with the surname being the filing element. Before surnames began to be adopted around 1800, however, Arabic names were generally made up of a variety of elements, without a standard last name. AACR2R Rule 22.22B1 says that the filing element of a classical Arabic name, one without an obvious surname, should be the element by which the person is best known. Rules 22.22C1 and 22.22D1 detail what elements should be left out and in what order the remaining elements should be given. The authority file for a classical author may look very different from the name given on the title page:

Author's name on title page (Romanized):

Abī Ḥanīfah al-Nu‘mān ibn Muḥammad ibn Manṣūr ibn Aḥmad
ibn Ḥayyūm al-Tamīmī al-Maghribī

Name in authority file:

al-Nu‘mān ibn Muḥammad, Abū Ḥanīfah, d. 974.

Luckily, a copy cataloger does not have to establish authority files. One may, however, need to try an element of a name in a search of a bibliographic utility. Thus it would be useful to understand the possible complexity of classical names when comparing a catalog record with the book in hand. A book of Arabic names alphabetized in their Romanized forms can be a great help when you need to include a name in a search. Salahuddin Ahmed's *A Dictionary of Muslim Names* is a good choice, and includes Arabic spelling and alternate Romanizations for most names. Some popular

names that derive from the Torah or Bible are also listed, as they are popular among Muslims, Christians and Jews in the Middle East. داود Dāwūd (David) and عيسى ‘Iṣá (Jesus) are but two examples. Forms and spellings of names can vary somewhat, however, so one should prefer spelling given in the book in hand or a name authority file over what is found in a reference work.

Persons who have become known in the West under other versions or Romanizations of their names will often be established under those names. The medieval philosopher Ibn Sīnā, for instance, has long been known in the Europe and America as Avicenna. Books about the former Egyptian president Gamal Abdel Nasser will have a Marc Field 600 with “Nasser, Gamal Abdel, 1918-1970,” while the title may refer to him as “Jamal ‘Abd al-Nāsir” according to proper ALA-LC Romanization. Egyptians tend to pronounce the letter ج as a hard g instead of a j, but ALA-LC always Romanizes the letter as j in descriptive cataloging. Accordingly, the cataloger does not bear the burden of speculating on whether an author is Egyptian and how he or she personally pronounces his or her name.

Strategies for Cataloging Copy Searches

Obviously, if the book in hand has an ISBN, LCCN (Library of Congress Control Number), added title page in English, or another readily accessible search element, the cataloger should attempt to use it in a search before braving the perilous shoals of Romanization.

For the non-Arabic speaking cataloger, Romanization of Arabic words is largely a matter of trial and error. If one cannot determine the root of the word, or which

vocalization of the word is appropriate, with the help of the dictionary, one will need to start trying out possible Romanizations.

OCLC's Connexion gives one the option of derived searching or keyword searching in various categories (title, author, publisher, ISBN, etc.) The experienced cataloger who is used to derived searching may nonetheless find it easier to use keyword fields and search for, for example, a word from the title and the author's first or last name, and perhaps the city of publication as well. Any element that can be speedily Romanized is fair game, and will assist the cataloger in zeroing in on a record.

Connexion's keyword searching is always within a certain category, so a cataloger who is not sure whether what they are searching is a name or part of the title, for instance, may prefer to begin searching in WorldCat. WorldCat accesses the same database of records as Connexion or Passport, but offers uncontrolled keyword searching.

Like many OPACs, OCLC's WorldCat does not automatically display vernacular information or diacritics. However, it has the capability to display diacritics on all Arabic records and vernacular for the smaller percentage of records which have parallel vernacular fields. Having this information available can be a great help as one tries out Romanizations, so one should set WorldCat to display them when entering a session. In the top right corner of the screen of the main page there is a link labeled "Options." One should choose this link and then under "Character set" switch from Standard to Unicode. Finally, choose the "Set" button, or "Set and save" for an individual account login. If one does not have an individual account (as is the case in most libraries), one will need to reset this option at the beginning of each WorldCat session. Records will now automatically display any vernacular information and diacritics in Romanized fields, in

Arabic and in other languages. When searching in WorldCat or in Connexion, diacritics are not searched. Thus one searching for “maṣāḍir” or “masadir” will get the same results for each. However, it helps to see them in the results so that one can distinguish between letters.

As one begins to Romanize title information, one may find that the title is written in elaborate calligraphy. If it looks like trouble, one should look for the title written elsewhere in a simpler design, for instance on the cover, spine, half-title page, or title page verso. This can help in figuring out the title as presented on the prescribed source of information.

In Connexion 1.3 and future versions, Arabic vernacular can be searched as well as input in parallel fields (Appendix D). However, only about a fifth of records for Arabic items have vernacular information, so it would not be a complete search. Connexion has a transliterator tool that automatically generates parallel fields in Arabic from Romanized fields. It is supported on operating systems Windows 2000 and Windows XP. However, it may also transliterate subfield indicators, which then need to be restored to English.

If a cataloger feels confident about editing or adding data due to language study or the help of a native speaker, he or she may wish to add Arabic typing capability to his or her computer. A search engine will turn up directions for doing so; one good set is located at http://web.syr.edu/%7Ediekemar/How_to_type_Arabic_characters.pdf. In Connexion, diacritics are put in after the corresponding letter, unlike in Passport, where they are put in before. They appear above or below the letter rather than next to it.

Connexion offers a button marked ALA which displays ALA/LC diacritics that can be inserted with mouse clicks. For those who prefer typing commands, here are the keystrokes for diacritics used in Arabic Romanization:

<ctrl><alt><E>	macron (line above letter)
<ctrl><alt><R>	dot below letter
<ctrl><alt><0> (zero)	‘ayn (ع)
<ctrl><alt><.> (period)	hamzah (ء) (glottal stop)
<ctrl><alt>	acute (used over a for ا alif maqṣūrah)

Conclusion

Catalogers are frequently faced with material that is challenging because of its format, ambiguities, or language, and must collect tools and strategies for dealing with a variety of situations. Sometimes, however, these tools prove inadequate without special knowledge possessed by the cataloger, as in the case of a language like Arabic, which is difficult to Romanize. It is hoped that this guide will serve as a useful implement to facilitate access to the actual cataloging tools necessary for Arabic copy cataloging searches, and that by extension it will help widen access to Arabic materials in Western libraries.

Works Cited

- American Library Association (2004). *Anglo-American cataloging rules*, 2nd edition, 2002 revision. Chicago: American Library Association.
- Beall, J. (1991). Describing the foreign language skills of catalogers in academic libraries. *Cataloging & Classification Quarterly*, 14(1), 39-47.
- Ben-Gurion University combined catalog* (2004). Retrieved March 15, 2005 from Ben-Gurion University of the Negev Web site: <http://dahab.bgu.ac.il/>
- Chressanthis, J. (2001). Cataloging and classifying non-English translations: An overview. *Technical Services Quarterly*, 18(4), 27-33.
- Da Conturbia, S. (1992). Who catalogs foreign-language materials?: A survey of ARL libraries. *Technical Services Quarterly*, 10, 15-30.
- Houissa, A. (1991). Arabic personal names: Their components and rendering in catalog entries. *Cataloging & Classification Quarterly*, 13(2), 3-22.
- Hujelan, N. (2004). *Arabic-English dictionary: How to use "Wehr" dictionary*. Retrieved January 22, 2005, from Indiana University Bloomington Web site: http://www.iub.edu/~smac/exam/arabic_dictionary.pdf
- Jajko, E.A. (1993). Cataloging of Middle Eastern materials (Arabic, Persian, and Turkish). *Cataloging & Classification Quarterly*, 17(1/2), 133-147.
- Khurshid, Z. (2002). Arabic script materials: Cataloging issues and problems. *Cataloging & Classification Quarterly*, 34(4), 67-77.
- Khurshid, Z. (1998). ARABMARC: A long way to go. *World Libraries*, 8(2), 60-68.
- Khurshid, Z. (2002). From MARC to MARC 21 and beyond: Some reflections on MARC and the Arabic language. *Library Hi Tech*, 20(3), 370-377.
- LC is first to use new RLIN Arabic-script cataloging system (1991). *Library of Congress Information Bulletin*, 50, 445.
- Online Computer Library Center (2004). *Arabic cataloging: Cataloging software for Arabic-language materials*. Retrieved March 15, 2005 from OCLC United States Web site: <http://www.oclc.org/libproxy.lib.unc.edu/arabic/>

Riedlmayer, A.J. (1999). "Ghost cataloging" and other tales of degraded access. *Art Documentation*, 18(1), 29-34.

Schultze, L.M. (1995). *A Beginner's guide to copy cataloging in OCLC/PRISM*. Englewood, Co.: Libraries Unlimited, Inc.

Ugalde, M. and Etcheverria, K. (1998) Mendigainera: Scaling the peak of the Basque library backlog. *Library Resources & Technical Services*, 42(4), 304-312.

Vernon, E. (1996). Decision-making for automation: Hebrew and Arabic script materials in the automated library. *Occasional Papers*, Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign (205).

Vernon, E. (1991). Hebrew and Arabic script materials in the automated library: The United States scene. *Cataloging and Classification Quarterly*, 14(1), 49-67.

Weihs, J. (1998). Interfaces: Three tales of multilingual cataloging. *Technicalities*, 18(10), 6-8.

Reference Works About Arabic for the Cataloger

Ager, S. (2004). *Omniglot: A guide to written language*. Web site:
<http://www.omniglot.com>

Ahmed, S. (1999). *A dictionary of Muslim names*. New York: New York University Press.

Alrabiah, T. (1992). *Gregorian-Hijri dates converter*. Web site:
<http://www.rabiah.com/convert/convert.php3>

American Library Association and Library of Congress (2002). *Arabic*. Web site:
<http://www.loc.gov/catdir/cpsd/romanization/arabic.pdf>

Answers.com. *Arabic alphabet*. Answers.com Web site:
<http://www.answers.com/topic/arabic-alphabet-2>

Barry, R.K. (Ed.) (1997). *ALA-LC Romanization tables: Transliteration schemes for non-Roman scripts*. Washington: Cataloging Distribution Service, Library of Congress.

Diekema, A. *How to type Arabic characters*. Retrieved February 21, 2005 from Web site:
http://web.syr.edu/~diekemar/How_to_type_Arabic_characters.pdf

Ja'far, M. (2002). *Arabic calligraphy: Naksh script for beginners*. New York: McGraw Hill.

Mace, J. (1999). *Arabic verbs and essential grammar*. London: Teach Yourself Books.

Mandel Khan, G. (2000). *Arabic script: Styles, variants, and calligraphic adaptations*. New York: Abbeville Press.

Wehr, H., Cowan, J. M. (ed.) (1976). *A Dictionary of modern written Arabic*. Ithaca, NY: Spoken Language Services, Inc.

APPENDIX A: Christian and Islamic Calendar Months

Gregorian Calendar Months (ميلادي or Christian year, usually marked .م)
This information was adapted from Jajko (1993).

Month (English)	Eastern Mediterranean	Egypt/Libya
January	كانون الثاني Kānūn al-Thānī	يناير Yanāyir
February	شبات Shubāt	فبراير Fibrāyir
March	آذار A`dhār	مارس Māris
April	نيسان Nīsān	أبريل Abrīl
May	أيار/أيّار Iyyār/Ayyār	مايو Māyū
June	حزيران Hazīran	يونيو/يونية Yūniyū/Yūniyah
July	تموز Tammūz	يوليو/يولية Yūliyū/Yūliyah
August	آب A`b	أغسطس Aghusṭus
September	أيلول Aylūl	سبتمبر Sibtimbar/Sibtimbir
October	تشرين الأول Tishrīn al-Awwal	أكتوبر Uktūbar/Uktūbir
November	تشرين الثاني Tishrīn al-Thānī	نوفمبر Nūfimbar/Nūfimbir
December	كانون الأول Kānūn al-Awwal	ديسمبر Dīsīambar/Dīsīmbir

Month (English)	Algeria/Tunisia	Morocco
January	جانفي Jānfi	يناير Yanāyir
February	فيفري Fīfri	فبراير Fabrāyir
March	مارس Māris	مارس Māris
April	أفريل Afrīl	أبريل Abrīl
May	ماي Māy	مايو Māy
June	جوان Juwān	يونيو Yūnyūh
July	جويلية Juwīliyah	يوليو Yūlyūz
August	أوت Ūt	غشت Ghūsht
September	سبتمبر Sibtimbar	شتنبر Shitanbar
October	أكتوبر Uktūbar	أكتوبر Uktūbar
November	نوفمبر Nūfanbar	نونبر Nwinanbar
December	ديسمبر Dīsīambar	ديجنبر Dījanbar

Hijriyah (Islamic) Calendar Months (marked هجرية or هـ.)

محرم	Muḥarram
صفر	Ṣafar
ربيع الأول	Rabī‘ al-Awwal
ربيع الثاني	Rabī‘ al-Thānī
جمادى الأولى	Jumādā al-Ulā
جمادى الآخرة	Jumādā al-Akhirah
رجب	Rajab
شعبان	Sha‘bān
رمضان	Ramaḍān
شوال	Shawwāl
ذو القعدة	Dhū al-Qa‘dah
ذو الحجة	Dhū al-Ḥijjah

APPENDIX B: Answers to Dictionary Exercise (page 30)

1. كُتُبِي kutubī bookseller, book dealer (last catchword on p. 812 of Wehr).
2. مَكَاتِب makātib libraries, bookstores, desks (plural of مَكْتَبَة, first catchword in the second column of p. 813). Notice that مَكْتَبَات is another possible way of pluralizing the word.
3. كَاتِبَات kātibāt woman secretaries, authoresses, writers (plural of كَاتِبَة, the feminine form of كَاتِب, fourth catchword from the bottom of the second column on p. 813). The order of these definitions is a testament to the age of the dictionary, translated from German to English in 1960 and last revised in 1976.

APPENDIX C: Verb Patterns and Conjugation

The Verb Patterns (I is irregular but has three common forms, IX is very rare and is omitted here)

Maṣḍar (verbal noun) (usu. listed as a noun in Wehr)	Present Tense	Past Tense	Form
(varies)	يَذْرُسُ yafʿul	فَعَّلَ faʿala	I
(varies)	يَذْرَسُ yafʿal	فَعَّلَ faʿila	I
(varies)	يَذْرَسُ yafʿil	فَعَّلَ faʿala	I
نَفَعِيلٌ tafʿīl	يُفَعِّلُ yufaʿʿil	فَعَّلَ faʿʿala	II
مُفَاعِلَةٌ mufāʿilah	يُفَاعِلُ yufāʿil	فَاعَلَ fāʿala	III
إِفْعَالٌ ifʿāl	يُفْعِلُ yufʿil	أَفْعَلَ afʿala	IV
تَفَعَّلٌ tafaʿʿul	يَتَفَعَّلُ yatafaʿʿal	تَفَعَّلَ tafaʿʿala	V
تَفَاعُلٌ tafāʿul	يَتَفَاعُلُ yatafāʿal	تَفَاعَلَ tafāʿala	VI
نُفْعَالٌ infiʿāl	يَنْفَعِلُ yanfaʿil	إِنْفَعَلَ infaʿala	VII
إِفْتِعَالٌ iftiʿāl	يَفْتَعِلُ yaftaʿil	إِفْتَعَلَ iftaʿala	VII
إِسْتِفْعَالٌ istifʿāl	يَسْتَفْعِلُ yastafʿil	إِسْتَفْعَلَ istafʿala	X

Conjugation

Prefixes and suffixes are added to the measure to conjugate a verb for different persons.

There are dual and female plural forms, but these are rare and are omitted here.

Past tense (the perfect tense)

تُ	-tu	I	نا	-nā	we
تَ	-ta	you (masc.)	تُمْ	-tum	you (pl.)
تِ	-ti	you (fem.)			
اَ	-a	he	وا	-ū	they
تَ	-at	she			

Present tense (the imperfect tense)

أَ	a-	I	نَ	na-	we
تَ	ta-	you (masc.)	تَونَ	ta- - ūn	you (pl.)
تَينَ	ta- -īn	you (fem.)			
يَ	ya-	he	يَونَ	ya- - ūn	they

APPENDIX D: OCLC Connexion Record with Parallel Vernacular Fields

OCLC Connexion		Page 1 of 1			
OCLC 55071699 Held by NOC - 3 other holdings					
Rec stat n	Entered 20040328	Replaced 20040504			
Type a	ELvl 5	Srce	Audn	Ctrl	Lang ara
BLvl m	Form	Conf 0	Biog	MRec	Ctry le
Desc a	Cont b	GPub	LitF 0	Indx 0	
	Ills abh	Fest 0	DtSt s	Dates 2004 ,	
010	2003489893				
040	DLC-R #c DLC #d NOC				
066	#c (3 #c (4				
042	lcode				
090	DS247.9.N35 #b R553 2004				
049	NOCC #c 1 [00024830408] [#xkew]				
⌠100 1	ركابي، كريم طلال.				
⌠100 1	Rikābī, Karīm Ṭalāl.				
⌠245 1 2	تأليف كريم طلال #c / التطورات السياسية الداخلية في نجد، 1283-1319 هـ/1865-1902 م				
	الركابي؛ تقديم ومراجعة عبد الله بن محمد المنيف.				
⌠245 1 3	al-Taṭawwurāt al-siyāsiyah al-dākhiḻiyah fī Najd, 1283-1319 H/1865-1902 M / #c ta'īf Karīm Ṭalāl al-Rikābī ; taqdīm wa-murāja'at 'Abd Allāh ibn Muḥammad al-Munīf.				
⌠250	1 الطبعة.				
⌠250	al-Ṭab'ah 1.				
⌠260	بيروت : #b ،الدار العربية للموسوعات #c 2004.				
⌠260	Bayrūt : #b al-Dār al-'Arabīyah lil-Mawsū'āt, #c 2004.				
300	270 p. : #b ill., maps, facsim. ; #c 24 cm.				
520	Najd (Saudi Arabia); politics and government; history; 1865-1902.				
504	Includes bibliographical references (p. 247-264).				
500	Abstract in English (p. 264-265).				
500	Islamic date of publication: 1425.				
651	0 Najd #z Saudi Arabia #x Politics and government #y 19th century.				
651	0 Najd #z Saudi Arabia #x Politics and government #y 20th century.				
651	0 Najd #z Saudi Arabia #x History #y 19th century.				
651	0 Najd #z Saudi Arabia #x History #y 20th century.				
Delete Holdings-	Export-Label-Produce- C	Submit-Replace-Report	Error-	Update Holdings-	Validate- Workflow-In Process
about:blank					3/21/2005

APPENDIX E: تم بحمد الله Tamm bi-ḥamdu lillāh

These words, usually in calligraphy, are sometimes found on the final page of a book. They mean “Completed by the grace of God.” (تم الحمد لله tamm al-ḥamdu lillāh is a common variant with the same meaning.)

