

翻字および数詞の問題について

I. アラビア語DBにおける「読み」、および「翻字表」の問題

a) 目録情報流通の観点から

アラビア語DBを構築するに当たっても、できるだけ多角的検索の余地を残したDBが望ましいことは、前回も述べたとおりである。(検索項目としては、書名のAlphabet読み、著者のAlphabet形表示部分、SH[Subject Heading]、アラビア語への翻訳書の場合は原書名、原著者名等の諸条件からの検索が挙げられる)

この観点からするとアラビア語のみで構成されるDBは汎用性に欠ける。(アラビア語の読める利用者にしか利用できない)また、多言語対応システムを持たない図書館からの利用に対しても、利用の可能性を閉ざすべきではない。翻字タイトルさえあれば、アラビア文字の部分はともかく、検索や同定することは可能である。

b) 図書館業務及び資料流通の観点から

書名および著者名部分のAlphabet翻字形は、アラビア語に堪能でない図書館員にとっては取り扱いの手掛かりとして必須である(所蔵検索、レファレンス、ILL、翻字形データを持つ他大学<海外を含む>図書館との照会・同定)。また多言語対応システムを持たない図書館に対しても、NACSIS-CATのデータを利用できる可能性を残す必要があるので、その意味でも翻字タイトルの存在は不可欠である。

c) ALA-LC方式翻字表採用の提案

ア) ALA-LC方式翻字表の普遍妥当性

翻字形のフィールドが必要だとすればどのような翻字が妥当であろうか? ALA-LC方式のアラビア語読み転写翻字方式は、我が国では最も人口に膾炙した方式であるから、実用性の側面から見れば他の方式に優先して採用すべきと考える。欧米でもこの翻字表の権威はゆるぎない。試みに、欧米の図書館にインターネット経由でアクセスして見れば、この翻字表が広く使用されていることがわかる筈である。

「目録情報の基準 第2版」の「11.2 日本語以外の外字」の項では「アラビア文字、デヴァナガリ文字等は、原則として、LCの採用する方式に準じて、当該言語に応じた翻字形を入力する(Cataloging Service Bulletinの各号を参照)」と明確に規定している。「目録情報の基準 第4版」ではその後の情勢の変化に応じて「UCSにない文字は、原則として、LCの採用する方式に準じて、当該言語に応じた翻字形を「」で囲んで入力する」と修正されているものの、こと翻字形に関する限りALA-LC翻字表の重要性には変化がみられない。

目録情報において採用すべき翻字表があるとすれば、それはALA-LCの翻字表であって、その他の翻字表ではありえない。国際的標準の翻字表を使用して、翻字形での情報流通を保証するか、それともすべてをOriginal scriptに委ねて、翻字形での情報流通を断念するののかということの問題にしていることをご理解いただきたい。

イ) 既存書誌との整合性

国情研のNACSIS-CAT内には、既にALA-LC翻字表によるアラビア語書誌が多数作られている。この資源を有効に利用しアラビア語目録へ変換していく作業の過程においても、読みフィールドにも同翻字表を用いることが、最も効率が良いと思われる。すなわち項目コピーで、翻字タイトルの項目へ翻字形をコピーし、しかる後にアラビア語化すべき部分はOriginal scriptで機械変換もしくは入力すれば良いのであって、他の翻字表を採用した場合は、わざわざALA-LC

翻字形を一々修正するという無駄な手間が発生する。それはいかにも非効率的な作業と言わざるをえない。また工に於いて触れる簡略翻字形にする場合には、ALC-LC 翻字形から母音を除去する作業が必要になる。

ウ) LA-LC 方式翻字表と書誌作成の実際

国情研参加図書館が、NAC SIS - CAT 内に書誌を作成する場合、参照 MARC 資源からの流用入力の有効利用可能性は常に念頭に置かれるべき問題である。アラビア語の参照 MARC 資源として、最初に挙げるべきは、US - MARC (= LC) であろう。作業効率から考えると、これら US - MARC は、NAC SIS - CAT に流用入力した上で、必要な部分をアラビア語に変換し、そのまま無修正で使える部分はそのまま残す方法が最も合理的である。(理由はイの項目で述べた) この意味で、ALA-LC 翻字表を標準的翻字表とし続けることが最もリスクや無駄が少ない。(他の翻字表を採用した場合には、読みの部分は他の翻字表によって再翻字や手直しする必要が生じる。また工に於いて触れる簡略翻字形にする場合には、ALC-LC 翻字形から母音を除去する作業が必要になる。) ちなみに US - MARC の作成者である Library of congress は、今のところ翻字形書誌を Original script に変更する予定はなく、将来とも翻字形式で参照 MARC を供給し続けるとのことである。

エ) 簡略翻字形の問題点

簡略翻字形の採用に関しては、さらに検討すべき問題が存在する。既存の ALA-LC 翻字形データとの整合性、および、今後作成される書誌に対して起こる、AL や SH 等においてアラビア語にあたる言葉を標目とする場合の表記の問題である。1 つの書誌に対して 2 度 (Original script をふくめれば 3 度) の検索を行わなければ正確な登録状況がわからない目録 DB でいいのか? また、これを過渡期的なものとして解決するとしても、AL や SH における ALA-LC 翻字形との整合性をどうとるのか? TR の検索補助的記述としての簡略翻字であればともかく、AL や SH において(しかもアラビア語にあたる言葉に関してのみ)これを採用することには根拠が無い。しかし、現行どおり AL や SH において ALA-LC 翻字形を用いるのであれば、TR に簡略翻字形を採用した場合、2 つの翻字形が混在することになる。これは、前述の既存書誌との整合性以上の書誌記述としての根本的な矛盾である。

以上の理由により、必須項目となる翻字表としては、簡略翻字(子音のみを記述)方式では実用面からも、書誌記述の整合性という点からも問題がある。簡略翻字がいまだ市民権を得た翻字表でないことも、採用するのに躊躇されるところである。

オ) LA-LC 方式翻字形の検索機能面における問題提起

ALA-LC 翻字形に問題がないわけではない。アラビア文字系言語で必ず使われる Alif や Ayn の 2 文字は、LC 翻字表上では Alphabet とは独立した記号が当てられている。そしてこれら 2 文字は目録 DB や NAC SIS - CAT では、EXC 文字の一部になっており、対応する正規化文字がないため、通常のキ - ボ - ドタイピングからは検索できないのが現状である。しかし、この文字を入れずに検索すれば正確な検索は難しく、検索洩れを生じる可能性が大きい。Web-Cat ではこの 2 文字を省略した形でも検索を可としているが、NAC SIS - CAT では未解決である。

・数詞の扱いの問題

a) 前提条件の整理

国情研の NAC SIS - CAT 内に書誌を作成する以上、そして AACR 2 を適用する以上、同じ AACR 2 を適用する他の言語 DB 目録との整合性を図る必要がある。

非 Roman-alphabet 使用言語ではロシア語、ギリシア語、中国語、それに現在策定中の韓国(朝鮮)語を参照の対象とすることが可能である。しかしながら、中国語・韓国(朝鮮)語は、デ - タベ - ス構造上の参考にはできるが、適用目録規則が NCR (=「日本目録規則」) 1987 年版なので、若干事情は異なる。

b) 数詞に関する提案

ここで、東京外国語大学図書館が「第一回アラビア文字文献入力規則WG」で提示した案より数詞に関する部分を再録する。

[VOL] 巻号を表す言葉はアラビア語、数字は西欧式アラビア数字

[ED] 版次(刷次)・改訂・増補等の内容を表す言葉はアラビア語、数字は西欧式アラビア数字

[PUB] アラビア語を基本とする(一部 AACR2 を適用) 刊年は西欧式アラビア数字を用いる。
(図書の記述に従ってイスラム暦と西暦を列記する場合もありうる)

[PTBL] 基本的にはアラビア語、巻号数をあらわす数字は西欧式アラビア数字

[AL] 著者名の翻字形部分

上記のとおり西欧式アラビア数字の適用を主とするが、その理由は以下のとおりである。

ア) AACR2 との整合性

AACR2 では「C. 4 東洋諸言語の数字」の項では数詞につき、細かく規定している。詳しくは英米目録規則第2版 p. 620 を参照していただきたいが、今回のアラビア語DB に際して特に参考になる項目のみを引用したい。

C. 4 A アラビア文字、極東諸語、ギリシア語、ヘブライ語、インド数字などの史料を目録する際には、以下の規則で指示するように、言語の数字を炉-間数字または西欧式のアラビア数字に代える。

C. 4 D 次の書誌的記述エリアでは、西欧式のアラビア数字に代える

- 1) エリアの版または刷表示的要素
- 2) 特定の条項が別の指示をしていない場合(3.3.B2 参照)、資料(または出版物の種類特性エリア)
- 3) 出版、頒布などのエリアの出版年、頒布年などの要素
- 4) 形態的記述エリア
- 5) シリ-ズエリアのシリ-ズ番号

このように AACR2 には細かく数詞について規則が定められている。そして AACR2 を採用した以上、これに正確に準拠した方が良いのではないかというのが第一の理由である。

イ) 多言語環境であるがゆえの Global standard としての利便性 = 公共性

それぞれの国の文化を尊重して、それぞれの言語で使用されている数詞はそのまま書いてあると通りに使用すべきであるという意見には一定の根拠があり、理解することはできる。そこでアラビア文字書誌にはアラビア数詞という意見は一応もっともではあるが、実際にはアラビア文字だけの問題ではなく、他の AACR2 目録規則を使用する言語(例えばロシア語・ギリシア語) 書誌との整合性も考える必要がある。これらの書誌では数詞は全て西欧式アラビア数字を用いている。

しかし、ここに挙げたロシア語・ギリシア語は西欧文化圏に属するので、東洋諸語とは事情が異なるという反論もあるやもしれない。けれどもアラビア文字書誌で定められる書誌構造が、他の東洋諸言語書誌のモデルケース(- 各言語それぞれの事情を加味したにしても-) になることを考えると、ことはそう簡単ではではない。

東南アジア諸国のうちでも、Roman-alphabet を採用した国々(フィリピン・インドネシア・マレ-シア・ベトナム)を除く他の諸国、すなわちラオス、タイ(クメ-ルはほぼタイに同じ)・ミャンマ-等の言語ではそれぞれ独自の数詞を持ち、発音もそれぞれ異なっている。さらにデ-ヴァナ-ガリ-文字を使用するインド諸言語では、ほぼヒンディ-語と同じ数詞を用いるが、異体字も極めて多いとのことである。

言語毎に異なる数詞を用いるという原則を採用した場合、ある言語を理解している利用者は当該言語の数詞を読めても、他の言語の利用者には、その数詞が読めないということにはならないだろうか? いかな Polyglot といえども、全部の数詞が読めるとは限らないのである。まして上記 I の a・b でも触れたように、これら多言語図書を扱う図書館員にまでそのような言語能力を求

めるのは現状では無理がある。

仮にそれぞれの言語の数詞を用いるにしても、異体字に対して、その文字フォントを探して用いるべきか、あるいは文字フォントが存在しない場合は標準的な字体に直して入力すべきかという問題も生じる。その場合は決して、図書に書いてある通りに入力することにはならない。

1 から 0 までの数詞さえ覚えれば、当該言語の数詞を扱えるようになるかということ、そう単純ではない。基数の位取りはほとんどが西欧式アラビア数字とおなじく、アラビア数詞・インド数詞のように 1 桁ずつが上がっていくので、ペ - ジ数を西欧式アラビア数字に置き換える作業はさほど困難ではないが、読み方に関しては和数字同様、「十」に当たる文字を間に挟んで（ - 「二十六」のように - ）表現する言語（タイ語・ミャンマ - 語）も存在するので、一筋縄にはゆかない。序数に関してはいうまでもなく、各言語それぞれの読み方がある。

各言語には、その言語の数詞をという理論は理解できなくもない。しかし数詞を含めて多言語を Original script で入力できる環境が整いつつある現在、かえって図書館員はとまどいを感じている。多数言語の数詞を全て理解する能力を図書館員に求めるのが無理である以上、ここは西欧式アラビア数詞の Global standard としての利便性 = 公共性を採るべきではないか。その方が国情研の「学術情報資源の全国的な共同利用」という理念にもかなうはずである。

c) 数詞情報の重要性和書誌構造上の優先度

目録情報流通・資料流通の諸場面において数詞の重要性は非常に高い。確かに資料を欲するのは記述言語に精通した者であろうが、求められた資料を検索・照会・同定し、実際の利用に供するのは図書館員である。言語に精通した者が介在しない状態での確実な流通を確保するためには、共通認識手段としての西欧式アラビア数詞はぜひとも必要と考えている。

また、数詞に関わるフィールドである、 . b) の ~ のフィールド、および のシリーズ番号入力部分は、国情研の書誌構造では複数持つことが許されていない。この書誌構造下において数詞を Original script で記述した場合、多言語対応ができていない図書館においては、記述の内容が判読できないことになる。そして、先に述べた数詞の重要性を鑑みれば、このような状態が望ましくないことは明白である。書誌構造上 1 つしか入力フィールド持ち得ない以上、基本的にどのような環境であっても判読できる記述手段 = 西欧式アラビア数字を採るべきである。

. 終わりに

これまで、主として、NACCIS-CAT への接続館の環境は様々で多言語環境に全ての図書館が対応するには相当の時間とコストがかかるという側面から、ALA-LC 翻字表や西欧式アラビア数詞の必要性を説いてきた。

それでは、ALA-LC 翻字表や西欧式アラビア数詞が必要なのは過渡的措置であって、多言語対応が完了すればそれらの項目は不要になるのかということ、そういう訳でもないだろう。

これまで述べてきたように、非西欧言語に疎遠な図書館員にとっては、ALA-LC 翻字表や西欧式アラビア数詞は依然として、検索・同定・流通のための重要なキ - である。これら ALA-LC 翻字表や西欧式アラビア数詞は非西欧世界と西欧世界とをつなぐ重要な橋渡し (=ブリッジ) であるとともに、両方の世界を包括するための共通認識手段でもある。国際社会で通用する目録を目指すのであれば、必要不可欠な情報であるといえよう。

了