

アラビア系文字の基礎知識

2005/6/6 ver. 1.00

(株)ワイズ 道広勇司

- アラビア文字はアラビア語を表記するための文字システムですが、ペルシア語、ウルドゥー語、インドネシア語、ウイグル語など多様な言語に広がっていきました。これらの文字システムは少しずつ違っていますが、出自が同じで共通点も多いので、アラビア系文字と総称します。
- 本資料はアラビア系文字の書字システムについての基礎知識をまとめたもので、もともと、多言語組版研究会 (<http://www.antenna.co.jp/ml/>) のセミナー「アラビア系文字の基礎知識」(2005年1月24, 31日) で配布した資料に加筆修正を施したものです。
- 本文中で “[23]” のようになっているのは、文献番号です。
- 筆者は文字学の専門家ではありませんし、本資料で取り扱うさまざまな言語を習得していません。誤りのご指摘や疑問点については筆者宛 (mich@moji.gr.jp) をお願いします。多言語組版研究会のメーリングリスト参加者の方はそちらでもお受けします。
- 用語については、日本で流布している用語が分からないために勝手に作ったものもあります。また、従来の用語が不適切なものは問題を指摘して代案を示しました。これらについてもご批判をお願いいたします。
- 本稿の執筆のために多数の文献を参照しましたが、互いに矛盾する記述や明らかな誤りが散見され、何が正しいのか推定するのが困難でした。そこで、なるべく文献番号で出典を明らかにするよう努力しましたが、とりわけ巻末の用語集については時間的制約のためほとんど出典が書けませんでした。
- 本資料は XSL Formatter V3.2 を用いて作成しました。

目 次

I アラビア系文字の世界	6
1 使用言語と地域	6
1.1 アフロアジア語族セム語派	6
1.2 印欧語族イラン語派	6
1.3 印欧語族インド語派	7
1.4 アルタイ語族テュルク諸語	7
1.5 ドラヴィダ語族	7
1.6 古語, アラビア系文字廃止などの例	7
2 起源と発展・伝播	8
2.1 文字の起こり	8
2.2 クルアーンの集成と正書法の確立	8
2.3 イスラームの拡大とアラビア文字の普及	10
3 特徴	10
3.1 形態的特徴	10
3.2 構造的特徴	10
II 文字	11
1 基本の文字	11
1.1 基字と識別点	12
1.2 'alif	12
1.3 fā' と qāf	13
1.4 fā' と ghayn	13
1.5 ʾayn	13
1.6 kāf と lām	13
1.7 mīm	13
1.8 hā'	13
1.9 文字の名称	14
1.10 文字の配列	14
2 単子音文字の仕組み	15
2.1 単子音文字とは	15
2.2 重子音	15
2.3 長母音の表記	15
2.4 セム語の特徴	16
3 続け書き	16
3.1 具体例	16
3.2 続け書きルールの詳細	16
4 合字	17
4.1 ラーム・アリフ合字	17
4.2 他の合字	17
4.3 合字の抑制方法	18
4.4 合字の書体依存性	18
5 カシーダ	18
III 付加記号と句読点類	19
1 記号の種類	19
2 母音記号類	19
2.1 /a, i, u/を表す記号	20
2.2 母音が無いことを表す記号	20
2.3 重子音記号	20
2.4 /ā/の綴りに使う記号	20
3 読誦指示のための記号	20

4	句読点類	20
IV	数字	22
1	字体	22
2	用途	22
3	書字方向	23
V	書体	24
1	筆記具	24
2	書道における書体	24
2.1	ナスフ体	24
2.2	スルス体	25
2.3	クーファ体	25
2.4	ナスタアリーク体	25
2.5	ルクア体	25
3	印刷書体	25
3.1	漢字書体	25
3.2	ラテン文字書体	26
3.3	アラビア系文字書体	26
4	ボールドと斜体	27
4.1	ボールド	27
4.2	斜体	27
5	和文との混植	28
5.1	文字サイズの釣り合い	28
5.2	行間	29
VI	組版・入力・校正上の注意	30
1	識別点の有無	30
1.1	アラビア語	30
1.2	ペルシア語	30
2	字形の混同	30
2.1	fā' と ghayn の両接形	30
2.2	lām-mīm 合字	30
2.3	tā'	31
VII	アラビア語	32
1	言語の概要	32
1.1	使用地域と人口	32
1.2	言語の系統	32
1.3	音韻	32
2	文字	32
2.1	アリフ・マッダ	33
2.2	アリフ・マクスーラ	33
2.3	ター・マルブータ	33
2.4	ハムザ	33
2.5	太陽文字と月文字	33
3	記号と句読点類	34
3.1	母音記号	34
3.2	シャッダ	34
3.3	ワスラ	34
3.4	クルアーン読誦のための記号	34
3.5	句読点類	34
3.6	慣用句の省略記法	35
4	書体	35
5	正書法	35

5.1	分かち書き	35
VIII	ペルシア語	36
1	言語の概要	36
1.1	使用地域と人口	36
1.2	言語の系統	36
1.3	音韻	36
2	文字	37
2.1	文字表の配列	38
2.2	字体についての注意	38
2.3	無音の he	39
3	記号と句読点類	39
3.1	母音記号	39
3.2	ハムゼ	40
3.3	他の記号	40
3.4	句読点類	40
4	書体	40
5	正書法	41
5.1	アラビア語との綴りの異同	41
5.2	語中の /i/	41
5.3	続け書きの分離	41
5.4	その他	42
6	ダリー語について	42
IX	ウルドゥー語	44
1	言語の概要	44
1.1	使用地域と人口	44
1.2	言語の系統	44
1.3	音韻	44
2	文字	44
2.1	反り舌音の文字	44
2.2	有気音を表す文字	45
2.3	二つの ye	45
2.4	鼻母音を表す nūn	45
2.5	異なる文字なのか異体字か	45
3	記号	45
3.1	ハムザ	45
3.2	アリフ・マクスーラの表記	46
3.3	母音記号	46
3.4	句読点類	46
4	書体	46
5	正書法	47
5.1	ハムザの有無	47
5.2	複合語の分かち書き	47
X	現代ウイグル語	48
1	言語の概要	48
1.1	使用地域と人口	48
1.2	言語の系統	48
1.3	文字の変遷	48
1.4	音韻	49
2	文字	49
2.1	母音字	50
2.2	ä と h	50
2.3	識別点のない文字 i	51

3	記号	51
3.1	ハムザ	51
3.2	ハイフン	51
4	書体	51
5	略語	51
XI	Unicode	52
1	言語	52
2	似た文字の使い分け	53
2.1	yā' の仲間	53
2.2	hā' の仲間	54
2.3	kāf の仲間	54
2.4	nūn の仲間	54
2.5	fā' の仲間	55
2.6	qāf の仲間	55
2.7	wāw の仲間	55
3	文字の採用方針	55
XII	参考文献	57

I アラビア系文字の世界

1 使用言語と地域

「アラビア文字」と「アラビア語」を混同してはならない。アラビア語以外にもアラビア系文字で書く言語はあるし、アラビア語がヘブライ文字やラテン文字で書かれることもある。英語の文献に“Arabic”とあったとき、アラビア文字なのかアラビア語なのか、よく注意する必要がある。英語で書き分ける必要がある場合は Arabic script, Arabic language とすればよい。

なお、ペルシア語、ウルドゥー語、パシュトー語…を表記するためのアラビア文字を特にペルシア文字、ウルドゥー文字、パシュトー文字…などと呼ぶ⁽¹⁾。狭義のアラビア文字はアラビア語を表記するためのものであり、広義のアラビア文字はこれらの総称である。

本稿では広義のアラビア文字を指す言葉として、「アラビア系文字」を使うことにする。

アラビア系文字は以下に述べるように極めて多様な言語で用いられており、地域的にも中東を中心に広範囲に分布している。ラテン文字に次いで広く普及した文字であるといえる。

1.1 アフロ=アジア語族セム語派

この語派に属する言語を「セム語」と総称する（セム語という一つの言語があるわけではないので注意されたい）。

セム語は、3 個前後の子音の組み合わせが単語の語彙的意味を担い、母音と接辞が単語の文法的意味を担う、という特徴を持つ。アラビア語以外のセム語として、ヘブライ語、アムハラ語などがあるが、これらは現在、普通にはアラビア系文字を用いない。アラビア系文字を用いるセム語に次のものがある。

- アラビア語：アラブ 22 ヶ国の言葉。またクルアーン⁽²⁾(القرآن al-Qur'ān) の言語。

1.2 印欧語族イラン語派

この語派に属する言語を「イラン語」と総称する⁽³⁾。次の項で挙げる印欧語族インド語派の言語とは近い関係にある。アラビア系文字を用いるイラン語として、次のものがある。

- ペルシア語：イラン、アフガニスタン（アフガニスタンではとくにダリー語と呼ぶ）
- パシュトー語：アフガニスタン、パキスタンほか
- クルド語：トルコ、イラク、イラン、シリアほか（ただしトルコ領内ではラテン文字、旧ソ連領内ではキリル文字を使用）
- バルーチ語：パキスタン、イラン

(1)ただし、アラビア系文字を用いるすべての言語が〇〇文字という固有の文字名称を持っているわけではない。また、現代ウイグル語の文字はアラビア系文字であるが、名称は「ウイグル文字」（これはソグド系文字である）ではなく「新ウイグル文字」（またはウイグル新字）であることに注意。

(2)コーランのこと。より原語に近い「クルアーン」の表記が専門書などで使われつつある。

(3)印欧語族（インド・ヨーロッパ語族とも）には、イラン語派・インド語派（両者を合せてインド・イラン語派と呼ぶ）のほかにイタリック語派（ラテン語、フランス語、イタリア語など）、スラヴ語派（ロシア語、ポーランド語など）、ゲルマン語派（ドイツ語、スウェーデン語など）といったさまざまなグループがある。これらはすべてただ一つの言語「印欧祖語」から分かれたと考えられている。

1.3 印欧語族インド語派

インド語派にはサンスクリット語やヒンディー語など、インドの多数の言語が含まれるが、これらの多くはインド系文字⁽⁴⁾を用いる。アラビア系文字を用いるのは以下に挙げる言語である。おおざっぱにはイスラーム教徒の多い言語と思えばよい。

- ウルドゥー語：パキスタン（公用語）、インド
- パンジャーブ語：パキスタン
- スインド語：インド、パキスタン
- カシュミール語：インド、パキスタン

1.4 アルタイ語族テュルク諸語

アルタイ語族という概念が成り立つかどうかはまだはっきりせず、アルタイ諸語といったほうが無難らしい。ともかくその中のトルコ（＝テュルク、Türk）系の諸言語をテュルク諸語という。シベリアからアナトリア半島（現在のトルコ共和国）まで広範囲に分布している。日本語と文法のよく似た膠着語（こうちゃくご）である⁽⁵⁾。

- ウイグル語：中国新疆ウイグル自治区ほか
- カザフ語・キルギス語：中国（旧ソ連ではキリル文字）
- アーゼルバーイジャン語：イラン（旧ソ連ではラテン文字）

1.5 ドラヴィダ語族

インダス文明を築いたと考えられているドラヴィダ人の言語。アーリア人がインドに進入すると、ドラヴィダ人の多くが南下した。そのため、現在南インドにタミール語、テルグ語、マラヤーラム語などのドラヴィダ語を話す人々がいるが、これらの言語はインド系文字を用いる。唯一バルーチスターン（イランとパキスタンの国境あたり）に逃げたブラーフイー人の言語は現在アラビア系文字を使っている。

- ドラヴィダ諸語は日本語と同じく膠着語であり、大野晋氏が日本語との系統関係を論じている。
- ブラーフイー語：パキスタン

1.6 古語、アラビア系文字廃止などの例

ここでは、既に廃れた言語や、政治的理由などによりアラビア系文字を使用しなくなった言語を取り上げる。旧ソ連邦では、レーニンの頃にラテン文字化させられ、スターリンの頃にキリル文字化された言語がいくつかある。

- チャガタイ語：テュルク系言語で、ペルシア語・アラビア語の語彙を多量に受容して成立した文語。19世紀まで中央アジアで共通語として使われた。
- スペイン語：スペイン語のアラビア系文字表記はアルハミーヤと呼ばれる。キリスト教支配下の隠れイスラーム教徒により用いられた。
- タージーク語：ペルシア語の一方言で、タージーク共和国の公用語。現在はキリル文字を使用。ただし、中国のタージーク人はアラビア系文字を使っているらしい。
- ウズベク語：ウズベキスタンほか。現在はラテン文字を使用。
- トルクメン語：トルクメニスタンほか。現在はラテン文字を使用。

(4) ブラーフミー文字およびそこから派生した諸文字。極めて多様な種がある。梵字もその一つ。

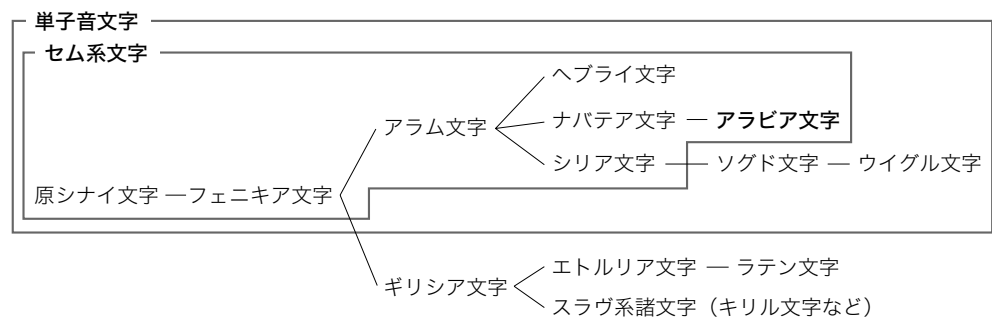
(5) 日本語は、たとえば「明日」という名詞に、「に」「は」「まで」といった接辞を組み合わせて付加し、「明日は」「明日には」「明日まで」「明日までに」「明日までは」「明日までには」といった句を作る。この種の文法をもつ言語を膠着語と呼ぶ。

- タタール語：タタールスタンほか。現在、キリル文字からラテン文字へ移行中。
 - カザフ語：カザフスタンほか。現在はキリル文字を使用。
 - インドネシア語：現在はラテン文字を使用。
 - スワヒリ語：アフリカ東岸。バントゥー諸語の一つ。現在はラテン文字を使用。
 - ハウサ語：現在はラテン文字を使用。
 - オスマン・トルコ語：現在のトルコ共和国のトルコ語の元になった、オスマン朝の言語。1928年にラテン文字化。
 - 漢語：漢語のアラビア系文字表記は小児錦（シャオアルチン）などと呼ばれる。一部のイスラーム教徒の間で使われてきたが、廃れつつある。地域によって正書法も異なり、文字名称も小児錦のほかいろいろある。
- 網羅的なリストは文献 [1] の「アラビア文字」を参照。

2 起源と発展・伝播

2.1 文字の起こり

アラビア文字は、フェニキア文字、アラム文字の流れを汲み、直接的にはナバテア文字を元にしてアラビア語を书写するために生み出された文字とされる。6世紀の碑文が残っている。



アラム文字とギリシア文字の系統（説明に必要な部分のみ）

この系統樹で、セム語の表記に使われた諸文字をセム系文字という。これらは単子音文字であり、各文字がもっぱら子音しか表さない。

一方、ギリシア人は、フェニキア文字の中でギリシア語に不要の文字を母音字として用い、世界で初めて子音・母音のすべての音素を表記する音素文字体系（狭義のアルファベット）を作り上げた。これがギリシア文字である。

ナバテア文字は22文字からなるが、アラビア語は28の子音音素を持つため、6字足りない。そこで、あぶれた音素は似た音の字で代用した。しかも、ナバテア文字の中には、字形が似すぎて実際上区別がつかなくなってしまった文字がいくつかあった。以上により、一つの字形がいくつかの音に対応することになり、多数の同綴異語（どうてついで）⁽⁶⁾を抱えることになった。この欠点は、後に識別点を導入することで解決した（後述）。

セム系文字は当初ばらばらに書かれていたが、ナバテア文字は続け書きすることがあり、これがアラビア文字にも取り入れられ、標準的な書き方となった。

2.2 クルアーンの集成と正書法の確立

こうして出来たアラビア文字は、7世紀にイスラームが誕生すると大きく発展していくことになる。

⁽⁶⁾ 同じ綴りをもつ異なる語をこう呼ぶよう提案する。英語なら lead ([li:d] 先頭, [led] 鉛) がそうである。

啓示

610年、ヒラー山の洞窟で瞑想していたメッカ（アラビア語ではマッカ）の商人ムハンマド⁽⁷⁾（محمد Muḥammad）の元に天使ジブリール（＝ガブリエル、マリアに受胎告知した天使に同じ）を通じてアラビア語で啓示が降った。啓示は継続して降り、ムハンマドはやがて預言者としての自覚を持つ。これがイスラームの始まりであった。啓示はムハンマドが亡くなる632年まで継続し、そのつど側近が記憶したり、秘書らが棗椰子（なつめやし）の葉、石片、獣骨、羊皮紙などに記録したりしたという。

この頃のアラビア文字はかなり不完全なものであったらしい [8]。

クルアーン

ムハンマドが亡くなると、カリフ（アラビア語ではハリーフ）と呼ばれる代理人が立てられた。最初の4人のカリフ（順に、アブー・バクル、ウマル、ウスマーン、アリー）をとくに正統カリフと呼ぶ（シーア派はアリー以外を認めないが）。

初代正統カリフのアブー・バクルは啓示の散逸を恐れ、ウマル（後の第2代正統カリフ）の進言を受けてザイド・イブン・サービト（ムハンマドの秘書の一人）にその集成を命じた。こうして書物としてまとめられたのがクルアーンである。第3代正統カリフ、ウスマーンは、これを元に、再びザイド・イブン・サービトに命じて権威ある原本としてのクルアーン写本を完成させた（651年）。この版をウスマーン本という。ウスマーンはこれをいくつも書写させ、各地に送った [8]。現物は残っていないが、現代まで営々と受け継がれてきたクルアーンテキストは、このウスマーン本である。

この時期まで、識別点や母音記号の使い方は十分に確立しておらず、ウスマーン本には、識別点も母音記号も使われていない。

母音記号と識別点

短母音を表記せず、一つの文字がいくつもの子音音素に対応するのでは、テキストの正確な読解・解釈に著しい支障を来すことになる。

そこで、イラク州知事の要請を受けたアブー・アスワド・ドゥアリー（688年没）が、クルアーンに母音点を付ける作業を行った。加点によって母音を表記することはそれまでにも行われていたが、彼によりシステムティックな方法が確立した。/a, i, u/音を表すのに、それぞれ文字の上、下、左に点を付加するのである [8]。これが現代の母音記号の元になった。

次に、ナスル・ビン・アースィム（707年没）が、同じ字形の子音を区別するための点（識別点）をクルアーンに付加する作業を行った。識別点のアイデアも、イスラーム以前からシリア文字などで使われており、ムハンマドの存命中に既にアラビア語でも行われていた証拠がある [8]。

この頃は、識別点は文字と同じ黒インクで、母音記号はそれと区別して赤インクで書くといったことが行われた。

識別点や母音記号を正確に扱うためには、アラビア語の音声や文法についての正確な知識が欠かせない。このためにアラブ文法学が発展することになった（8世紀頃）。

クルアーンのアラビア語は、綴りと発音がずれているが、クルアーンは聖書と違って神の言葉そのものとされているので、既に確定した綴りを改変することは許されなかった。そこで、特定の文字に特別な識別点を導入したり、新たな記号を案出したりすることで綴りと発音を対応づけた（具体的な内容は「アラビア語」の章で扱う）。このようにしてアラビア語の正書法が徐々に確立し、現在に到っている。

⁽⁷⁾ 日本では長らくマホメットと呼ばれてきたが、原語に近い「ムハンマド」の表記が近年普及しつつある。

2.3 イスラームの拡大とアラビア文字の普及

キリスト教が布教のために聖書の翻訳を積極的に進めたのと対照的に、イスラームはクルアーン
の翻訳を認めていない。神の言葉そのものであるから、「内容が伝わればそれでよい」というわけ
にいかないのだ。現在、クルアーンは日本語を含むさまざまな言語に翻訳されているが、それは
「〇〇語版クルアーン」などではなく、クルアーンの内容を〇〇語で解説したもの、とされている。

それゆえ、イスラーム教徒は母語がなんであれ、アラビア語でクルアーンを学び、朗読するのが
建前である。これにより、アラビア語はイスラーム世界全体に広く普及することになった⁽⁸⁾。

イスラームは、正統カリフ、ウマルとウスマーンの時代に、征服によって急速にその版図を広げ
たが、その後も商業活動などさまざまな要因により、アジア・アフリカ・ヨーロッパに広まった。
現在、世界最大のイスラーム教徒人口を抱える国はインドネシア（2億人弱）、以下インド、パキ
スタン、バングラデシュ（いずれも1億人以上）と続く。

このようにイスラームとアラビア語が普及する過程で、①無文字言語にアラビア文字が取り入れ
られたり、②旧来の文字を廃してアラビア文字表記が採用されたりした。後者の代表例はペルシア
語（旧来はパフラヴィー文字）である。これら、新規にアラビア文字を採用した言語は、文字を追
加するなどして自らの言語をうまく表記できるように工夫した。

3 特徴

3.1 形態的特徴

- 右から左へ書く（RL）
ただし、数字は左から右
書籍は右開きとなる。
- 続け書きする
- 分かち書きする
- 縦に積む合字⁽⁹⁾(ligature) がある
- 識別点だけで区別される文字がある
- 多様な書体があり、書体によってはコンピューター処理が極めて難しい
- 行末にかかった単語を分割しない（ただし現代ウイグル語は例外）

3.2 構造的特徴

- 単子音文字：基本的には全文字がもっぱら子音を表す。
ただし、ウイグル語、ウズベク語などは母音字を導入している。
- 長母音は子音字を流用して表記する。
- 短母音を表記するために母音記号を使うことがある。

⁽⁸⁾しかし、もちろんイスラーム教徒みながアラビア語を解するはずはない。ウラマー（聖職者・神学者）でさえ、アラビア語の発
音がろくにできないことも珍しくないそうである。とはいえ、ほとんどのイスラーム教徒は何らかの形でアラビア語になじんで
いるはずである。

⁽⁹⁾アラビア系文字でいう合字はラテン文字のそれとはやや異なるし、「縦に積む」という表現が適切かどうか疑問がある。

II 文字

ここでは、アラビア語を念頭に置き、基本の文字 28 字について解説する。ここで述べることは、他のアラビア系文字にもある程度あてはまる。アラビア語特有の事柄についてはアラビア語の章で取り上げる。

1 基本の文字

アラビア語の基本の文字は、次の 28 字である。

文字	名称		転写	右接形	両接形	左接形
ا	الف	'alif		ا		
ب	باء	bā'	b	ب	بب	بب
ت	تاء	tā'	t	ت	تت	تت
ث	ثاء	thā'	th	ث	ثث	ثث
ج	جيم	jīm	j	ج	جج	جج
ح	حاء	ḥā'	ḥ	ح	حح	حح
خ	خاء	khā'	kh	خ	خخ	خخ
د	دال	dāl	d	د		
ذ	ذال	dhāl	dh	ذ		
ر	راء	rā'	r	ر		
ز	زاي	zāy	z	ز		
س	سين	sīn	s	س	سس	سس
ش	شين	shīn	sh	ش	شش	شش
ص	صاد	ṣād	ṣ	ص	صص	صص
ض	ضاد	ḍād	ḍ	ض	ضض	ضض
ط	طاء	ṭā'	ṭ	ط	طط	طط
ظ	ظاء	ẓā'	ẓ	ظ	ظظ	ظظ
ع	عين	'ayn	'	ع	عع	عع
غ	غين	ghayn	gh	غ	غغ	غغ
ف	فاء	fā'	f	ف	فف	فف
ق	قاف	qāf	q	ق	قق	قق
ك	كاف	kāf	k	ك	كك	كك
ل	لام	lām	l	ل	لل	لل
م	ميم	mīm	m	م	مم	مم
ن	نون	nūn	n	ن	نن	نن

ه	هاء	hā'	h	هـ	هـ	هـ
و	واو	wāw	w	و		
ي	ياء	yā'	y	ي	يـ	يـ

ここに挙げた文字名称は正則アラビア語（フスハー）⁽¹⁰⁾のものである。字体や接続の仕方は書体によって少し異なるが、本稿では（多くの入門書と同じく）ナスフ体を前提として接続形を示している。ナスフ体はアラビア語でもペルシア語でも印刷物の本文に使う基本書体である。本稿に出てくるアラビア系文字は、Microsoft の Arabic Typesetting というナスフ体フォントで組んである。文字配列は歴史的にはいろいろあったが、上に挙げたのは正則アラビア語の標準的なものである。

なお、アラビア語では、このほかにハムザという記号⁽¹¹⁾や、特別な働きをもつ異体字などが使われるが、これらはアラビア語の章で取り上げる。

この表で、左接形・両接形・右接形は、分かりやすさのために接続部の線を長めにしている。

1.1 基字と識別点

表をひと目見て分かるとおおり、識別点の数しか変わらない文字が多数ある。既に述べたように、これらはもともと同じ字形で別の音を表わしていた文字に識別点を導入して別字としたものである。

たとえば، ㇀㇀㇀はいずれも㇀を基本の形としている。識別点を取り去った形を**基字**（英語では basic shape）という。なお、左接形・両接形（後述）では، ㇀も同じ基字となる。

1.2 'alif

'alif（アリフ）という文字名称は「牡牛」という意味のセム語であり、字形のルーツは牛の頭部を象った象形文字にある。形は変わってもその名称と音価は多くのセム語で保存された。この文字がギリシア文字のアルファ（Α, α）になったことはよく知られているが、よく誤解されているように、'alif の音価が /a/ だというわけではない。セム系文字における 'alif は声門閉鎖音という子音を表す文字であり、ギリシア人がそれを /a/ を表すのに借用しただけである。

※声門とは

2枚の筋肉膜である「声帯」に挟まれた呼吸の通り道を声門という。声門は開閉自在で、息を詰めると完全に閉じるし、食べ物を飲むときは完全に開く。声門をわずかに開いて呼吸を強く通すと声帯が振動し、音が出る。これが母音や有声子音の発声の原理である。

※声門閉鎖音とは

声門を調音点とする閉鎖音を声門閉鎖音という。日本語も含め、多くの言語で声門閉鎖音が観察されるが、これが音素として認められる言語は限られている。たとえば日本語の場合、「アッ!」というときの母音の前後などに声門閉鎖音が開かれるが、声門閉鎖音の有無で区別される単語は存在しない。声門閉鎖音はラテン文字では /ʔ/ で表す習慣である。この場合のアポストロフィーは記号ではなく文字の扱いとなる。

'alif は声門閉鎖音の文字であると書いたが、実はアラビア語についてはやや事情が複雑になる。クルアーンにおいて声門閉鎖音が文字に表記されない例外的な綴りが多々あり⁽¹²⁾、綴りを変更せずに発音との対応を取るため、ハムザという記号を導入し、これが声門閉鎖音を表すとした [1]。このため、アラビア語では 'alif は声門閉鎖音の文字とは言えなくなっている（ペルシア語では声門閉鎖音に対応する）。詳しくは「アラビア語」の章で述べる。

⁽¹⁰⁾ アラビア語の共通語。

⁽¹¹⁾ ハムザを記号ではなく文字とする立場もあるが、本稿では記号の扱いとする。

⁽¹²⁾ これはメッカを含むアラビア半島西部の方言を反映したものである。

'alif のもう一つの働きは、長母音 /ā/ を表すことだが、これは後述する。

1.3 fā' と qāf

fā' (ファー) ف と qāf (カーフ) ق は基字が違うように見えるが、もともとは同じである。

なお、これらの識別点は歴史的に揺れてきたが、現在は fā' が「上に一点」、qāf が「上に二点」となっている。ただし、マグリブ（西方の意）と呼ばれる西北アフリカでは、古い方式の一つが残り、それぞれを ب, ق のように書く。

1.4 fā' と ghayn

fā' (ف) と ghayn (غ) の両接形はそれぞれ ڤ と ڤ であり、基本的な違いは屈曲が無いかあるかということだけである。ghayn のほうは次に述べるように、ループをつぶすデザインもある。

1.5 ayn

'ayn (アイン) は、有声咽頭摩擦音を表す子音字である。咽頭摩擦音とは、舌の根元を奥（咽頭壁）のほうへ引いてせばめを作る摩擦音である。ラテン文字表記では、/ʕ/ または /ħ/ で表す。前者の記号はアポストロフィーを左右反転させた形であり、Unicode では U+201B: SINGLE HIGH-REVERSED-9 QUOTATION MARK として定義されている。後者はもちろん左シングルクォーテーションである。

右接形・両接形では、書体により頭部をループにするデザインとつぶすデザインがある (ghayn も同じ)。

1.6 kāf と lām

kāf (カーフ) の字形 ڤ は、中に s 字状の線が入っている。これは、ラーム (ل) との区別を明確にするために付けられたものである (かつてラームは現在ほど下に大きく膨らんでいなかった)。この s 字状の線の正体は、kāf の左接形 ڤ である。

1.7 mīm

mīm (ミーム) は、独立形と左接形ではループ部分を描く向きが逆になっている。また、右接形ではループをつぶすことが多い。

1.8 hā'

hā' (ハー) の独立形は ھ だが、辞書などでアラビア文字について記述する場合に、この文字が単独で書かれる場合は ھ という字形が好んで使われる。しかし、もちろん単語中に独立形が現れるときは、たとえばペルシア語の ھ (māh; 月) のように必ず ھ の字形でなければならない。

そこで、独立形には以下の 2 種類があると考えてもよいかもしれない⁽¹³⁾。

- 無接形 (ا): 単語中に書かれる場合の字形。

具体的には、語末に位置し、前文字が左接しない文字 (3 節参照) の場合。

- 孤立形 (ا): 単語の構成要素ではなく、文字が単独で書かれる場合の字形。

具体的には、(1)辞書のラベル文字、(2)略号 (たとえば年号がヒジュラ暦 [هجريه hijriyah] であることを示す هـ), (3)簡条書きの頭にアラビア文字を使う場合、などである。

(13) この「無接形」「孤立形」という用語は筆者が仮に名づけるものであり、業界で使われているわけではない。

一般の文字は無接形＝孤立形であり、hā'に限っては、無接形と異なる孤立形が使われる場合がある、というわけである。

孤立形に ھ が使われる理由として、「数字の 5 (٥) とまぎらわしいから」と言われているが、それだけが理由ではないように思う。

なお、孤立形として、代用的に左接形 (ه) を用いることが活版時代から行われてきた。歴史的には、印刷物ではむしろこちらが普通ではないかと思うが、格好の良さから考えれば ھ を使うべきものと思う。

hā'の両接形には二つの字体、ه と ه がある⁽¹⁴⁾。使い分けのルールがあるのかどうか、筆者には分からなかった。

1.9 文字の名称

基本的には個々の文字の名称の先頭の子音が文字の音価である。たとえば bā' (ب) の音価は /b/ である。このような対応をアクロフォニー (acrophony⁽¹⁵⁾) という。日本語では、文脈に応じて「頭音字法」とか「頭音原理」と呼ぶ。

文字名称を仮名で表記するのに一つ問題がある。たとえば ح, خ, ه の名称はそれぞれ hā', khā', hā' であるが、これを仮名で書けばいずれも「ハー」となり、区別がつかない。本稿において文字名称をすべてラテン文字表記したのはそのためである。

文字の名称は言語によって異なる。たとえばアラビア語の bā' はペルシア語ではふつう be である。

Unicode の文字名称はおおむねペルシア語風の文字名を英語風に表記したもののように見受けられるが、どのような命名方針なのか筆者は知らない。たとえば ب (Ar: bā', Fa: be) の Unicode 名は“ARABIC LETTER BEH”である (この最後の H は無音の H [「ペルシア語」章「文字」節参照] と思われる)⁽¹⁶⁾。ن (Ar, Fa: nūn) の Unicode 名は“ARABIC LETTER NOON”となっており、長母音 /ū/ にラテン文字“oo”を対応させるあたり、英語風表記といえる。

1.10 文字の配列

文字表に現れる文字の順序にはいくつかの種類がある。

アブジャド

伝統的に重要な配列として、アブジャド (أبجد 'abjad) というものがある。これはその並びの最初の 4 文字が ا, ب, ج, د であることによる命名である。アラム文字の配列に由来し、ギリシア文字の配列とも符合する。

「アブジャド」という言葉は、アラビア文字に基づく数の表記システムの名前でもある。つまり、'alif ا に 1, bā' ب に 2, jim ج に 3, dāl د に 4, というように数価をあてはめる。もちろん、1 から 9 までのほかに、10, 20, ..., 90, 100, 200, ..., 900 といった大きな数も文字に割り当てられており、これらの組み合わせで数を表記する。

現在の標準的配列

本節の表に挙げたのは現在の正則アラビア語で最も標準的な配列であり、辞書などもこれに基づいている。この配列は、同じ基字をもつ字がかたまりになるようアブジャドを並び換えたものである。

いずれの方式も、kāf ك, lām ل, mīm م, nūn ن の並びをもち、ラテン文字の K, L, M, N の並びとびつたり一致しているが、これは偶然ではない。

(14) このあたりの話はすべてナスフ体を前提としている。他の書体ではまた話が違ってくる。

(15) 接辞 *acro-* には「先端」という意味がある。

(16) 簡便のため、アラビア語を Ar, ペルシア語を Fa (Farsi より) と略す。

この順序はペルシア語やウルドゥー語においてもおおむね保たれている。

2 単子音文字の仕組み

2.1 単子音文字とは

アラビア文字を含むセム系文字はどれも単子音文字とよばれ、各文字がいずれももっぱら子音を表すという特徴をもつ。これはどういうことを意味するのか？

単子音文字の感じをつかむために、アラビア文字をラテン文字に置き換えて考えてみよう（このような置き換えを**転写**という）。アラビア語でペンのことを *qalam* という。この単語を表すために、単子音文字では QLM と表記する（ラテン文字転写は大文字で書く習慣がある）。

ちなみにアラビア文字表記は قلم である。

2.2 重子音

もう一つ例を挙げる。人名のムハンマドは *muḥammad* と発音する。ここで、“ḥ” は “h” とは異なる音を表す⁽¹⁷⁾。ムハンマドの表記は MḤMMD となりそうなものだが、アラビア文字を含む多くのセム系文字では重子音（ここでは mm）をまとめて 1 文字で表記するので、実際には MḤMD となる。アラビア文字表記は محمد である。

セム系文字の開発者たちは、同じ子音が二つ並んでいるとは認識せず、一つの子音が伸びた、あるいは強くなったものと捉えていたのだろう（そう捉えるほうがむしろ自然かもしれない）。

このように、単子音文字では当然ながら同綴異語ができやすい。単語や文法を知らなければ発音すら分からない。

2.3 長母音の表記

アラビア文字は、母音を一切表さないのではなく、長母音に限っては、特定の子音文字を流用して表記する。例を挙げよう。

アラビア文字の非常に古い書体にクーフィー (كوفي *kūfī*)⁽¹⁸⁾ というものがある。この語は KWIFY と表記する。W, Y は本来は子音字だが、ここでは長母音 /ū, ī/ を表すために流用されているのである。このような流用は他のセム系文字にも見られる。

アラビア文字では、長母音の表記に使う文字は 'alif (ا), wāw (و), yā' (ي) の三つあり、それぞれ、/ā, ū, ī/ の表記に用いる。なお、'alif のラテン文字転写はアポストロフィー (') であることに注意されたい。「本」という意味の *kitāb* は KT'B となる。

もう一つの解釈

しかし、これらは本当に子音字を長母音の表記に流用したものなのだろうか。セム系文字の開発者たちには、/ū/ や /ī/ ではなく、/uw/ や /iy/ と聞こえていたのではないだろうか。そうだとすれば、流用などではなく、子音字を原則通りに書いたことになる。/ū, ī/ と /uw, iy/ は、音声学的には異なった音だが、おそらくアラブ人にとっても日本人にとっても聞き分けたり発音し分けたりするのは困難だろう。

(17) ラテン文字に置き換えるとき、文字数が足りないで、このような下に点のある文字などが使われる。

(18) イラクの古都クーファ (Kūfa) で発達したのでこの名がある。本資料の他の箇所ではクーファ体という呼び方をしている。

2.4 セム語の特徴

単子音文字はこのように発音を不完全にしか表記しないが、それでも十分に実用的であるのは、セム語に共通して見られる次のような特徴のためだと考えられている。

たとえば、アラビア語で *kataba* という動詞がある。これは「書く」という動詞の原形だが、「彼は書いた」という意味の活用形でもある。他の活用形は、たとえば *katabat*（彼女は書いた）、*yaktubūna*（彼らは書く）というように、[k, t, b] の部分が共通している。

また、派生名詞として既に見た *kitāb*（本）や、*kātib*（作家）、*maktab*（事務所）、*maktabat*（図書館）などが挙げられるが、やはり [k, t, b] という共通の子音の組を持っている。

この子音の組を語根と呼ぶ。語根から派生形や活用形を作るやり方はおおむね規則的である。言語学的に言えば、語根が単語の語彙的意味を担い、そのまわりの母音や接辞が文法的意味を担う、ということになる。したがって、アラビア語の単語の綴りは、語根文字のまわりに限られた種類の文字（長母音の文字と接辞に現れる文字）がまとわりついた分かりやすい構造をしている。

以上に加えて、アラビア語の短母音が/a, i, u/の3種類しかないことも、幸いしている。

では、ペルシア語やウルドゥー語、ウイグル語のような、セム語でない言語や、多くの母音を持つ言語ではどうか、ということは個別に検討する必要がある。

3 続け書き

アラビア文字は（おおざっぱに言えば）単語ごとに続け書きする。ラテン文字における筆記体のごとくに。しかし、ラテン文字の筆記体との大きな違いは、続け書きの先頭・中間・末尾で字形が変わるという点である。このそれぞれの字形を左接形・両接形・右接形と呼ぶ。頭字形・中字形・末字形という呼び方もある⁽¹⁹⁾。一方、文字が単独で書かれるときの字形を独立形と呼ぶ。「左接」とか「両接」などというのは、「左に接続する」「両方に接続する」というような意味だと思えばよい。

アラビア文字は、古くは単語間にスペースを空けずに書くことがあったが⁽²⁰⁾、それでもこの続け書きのルールがあるために、単語の境界が推測できる。我が国の仮名書道においても、文節（？）の区切りで続け書き（連綿）を切るといった書き方があり、これによく似ている。

3.1 具体例

では具体例を見てみよう。「書く」という意味の *kataba*（ラテン文字転写は KTB）を取り上げる。要は *kāf* (ك), *tā'* (ت), *bā'* (ب) を右から左へ並べればよい。それぞれが左接形、両接形、右接形となる。

ك ت ب ← ← ك ت ب ← ← ك ت ب

それぞれの字の左接形、両接形、右接形がどのような形かは、文字表から拾えばよい。なお、手書きの場合にどのような書き順となるかは、文献[]などを参照されたい。

3.2 続け書きルールの詳細

さきほど「単語ごとに続け書きする」と書いたのは雑な表現であり、実際にはもう少し複雑である。

(19) 専門書を含む多くの本に、語頭形・語中形・語末形という用語で書かれているが、不適切な呼び方である。あとで分かるように、単語の先頭・中間・末尾に対応しているとはいえないから。

(20) ラテン文字もかつては語間を空けずに書かれた。

左接しない文字

まず、「左接しない文字」という一群の文字がある。さきに掲げた文字表の中では، و ز بر ب د با の6文字がそれである。単語の中であっても、この文字の後に続け書きはいったん途切れる⁽²¹⁾。

たとえば「本」という意味の kitāb は次のようになる（ラテン文字転写は KT'B）。

ك ت ا ب ← ك ت ا ب ← ك ت ا ب

'alif (ا) が左接しない文字のため，bā' (ب) から新たに続け書きを始めることになるが，その後ろにはもはや文字がないので，結局 bā' が独立形で書かれることになったわけである。

草書体の場合

この続け書きのルールは，厳密に言えば書体によって異なる。ディーワーン体やシェキャステ・ナスタアリーク体のような草書⁽²²⁾では，上に「左接しない文字」として挙げた文字の後ろも続け書きする。しかし，この手の書体はもともと機械的に描画するのは困難で，印刷書体としては使われていない（将来はコンピューターで組めるようになるだろう）。

分かち書きしない語

また，前後の単語と続け書きするもの場合もある。たとえばアラビア語では，定冠詞 ال ('al) や，一文字からなる前置詞は必ず後ろの単語に続けて書く。具体的にはアラビア語の章で扱う。

単語の途中で切る場合

ペルシア語では複合語などの場合に単語の途中で続け書きを切ることがある。詳細はペルシア語の章を参照。

4 合字

特定の二つ以上の文字が隣り合ったときに，それぞれの文字の単純なつなぎ合わせではない字形を用いることがある。これを合字 (ligature) という。ラテン文字における合字の概念とは少し違うが，印刷用語としては合字と呼ばれているようだ。

4.1 ラーム・アリフ合字

lām (ل) のあとに 'alif (ا) が続くとき，綴りはいになりそうに思えるが，実際にはこの形は決して使わず，合字にしなければならない。これは義務的である。

ラーム・アリフ合字の独立形は لا، 右接形は لاである。

このあたりの事情も，厳密に言えば書体依存性がある。

4.2 他の合字

ラーム・アリフ以外の合字は美的観点からなされるもので，義務的ではない。合字を使った場合にどのような字形になるか，いくつか例を挙げる。

تاس ← تاس ← تخ ← حمام ← حمام ← كا ← كا ← ني ← ني ← م ← م

こういった合字の使用・不使用の好みは言語や文化圏によって違うかも知れないが，調べがつかなかった。

(21) それゆえ，ペルシア語ではこれらの文字を「ちぎれ文字」(حروف مقطعه horūfe moqatta'e) と呼ぶ（元はアラビア語と思うが調べつかず）。

(22) アラビア系文字にも，速書きのために文字を崩して出来た書体があり，これを一般にシェキャステ (شكسته šekasteh；「壊された」という意味のペルシア語。古い発音ではシカスタ) という。これを草書体と訳すのは問題ないように思う。

4.3 合字の抑制方法

合字が使われていると、初学者は読みに困難を感じる。そこで、アラビア語などの学習書・辞書においてはラーム・アリフ合字以外の合字を一切使わないものも多い⁽²³⁾。

Unicode 対応の組版アプリで、合字を使わないように組ませるには、各文字の間に“ゼロ幅結合子” (ZERO WIDTH JOINER; 以下 ZWJ) と呼ばれる文字を挿入すればよい。このキャラクターは幅を持たずグリフも持たないが、前後のアラビア文字と接続する性質を持つ。

ただし、左接しない文字の後ろには入れてはならない。さもないとその後ろの文字が右接形ないし両接形になってしまうからである。

たとえば、「風呂」というような意味の ḥammām を، حم ではなく حمام のように組ませるには、以下のように (この図は左から右へ並べてある) 入力すればよい。

ح ZWJ م ZWJ ا م

ZWJ は U+200D と定義されているので、HTML 文書や XML 文書では ‍ のように表現することができる。

また、ペルシア語キーボードにはこの文字が割り当てられており (IME やアプリによる)、キーボードから直接入力できる。

4.4 合字の書体依存性

どの文字が隣あったときにどういう字形になるかは、書体に強く依存する。Unicode ではアラビア系文字の Presentation Forms として、各種の合字が挙げられているが、中途半端であり、多書体にわたる普遍性があるわけでもない。

5 カシーダ

ラテン文字では、ジャスティファイを実現するために、ワードスペースを伸縮させる。その際、調整量を小さくするため、行末にかかった単語を特定位置で分割 (word division) し、前綴りにハイフンを付ける (hyphenation)。

それに対し、アラビア系文字では語分割は行わず (現代ウイグル語は例外)、基本的に単語の長さを伸ばすことでジャスティファイを実現する⁽²⁴⁾。そのために単語の中の水平部分に、カシーダ (كشيدة kashīdah) と呼ばれる短い水平線を挿入する。

一般に、地図に書かれる山脈、河川、街路などの名称は、必要に応じて長く書かれなければならない。日本語の地図では文字間を広げることでこれを実現するが、アラビア語ではカシーダを使って単語を引き伸ばすことで実現する。



イランの地図：中央に「ホルモズ海峡」(霍尔摩兹海峡 tangeye hormoz) が丸く引き伸ばされて書かれている。

⁽²³⁾ 逆に言えば、合字を使わないとかっこ悪く見えることがある。

⁽²⁴⁾ 同時にワードスペースを広げることも行うようだが、調査不足のため詳細不明。

III 付加記号と句読点類

1 記号の種類

一口に記号といっても、以下のようにいろいろある。

- 母音記号類
- 読誦指示のための記号
- 句読点類 (punctuation marks)

2 母音記号類

アラビア語・ペルシア語では短母音は文字に表さないが、場合により、文字の上下に記号を付けて表すことがある⁽²⁵⁾。

たとえば、「飲む」という意味のアラビア語、shariba に母音記号を付けて記せば شَرِبَ となる⁽²⁶⁾。母音記号を付けるのは、だいたい以下の場合である。

(1) クルアーン

アラビア語を母語としないイスラーム教徒であっても正確に読める必要があるので、クルアーンには完全に母音記号を付ける。

(2) 小学校の教科書

子供の語彙が十分でないことを考慮して小学校の教科書には母音記号を付ける、とよく言われている。実際には国によって事情は違い、アラブでは半数ほどの国で実施されているらしい。イランでは1年生の教科書は付けるそうである。

(3) アラビア語等の学習書

語学書は当然、非ネイティブが使うものなので、母音記号を付ける場合が多い。しかし、辞書の用例などは母音記号を付けないことが多いので、発音は個々の単語をいちいち引かないと分からない。

(4) 書道作品

書道作品では装飾的要素として母音記号を付けるのが普通である。母音記号のほか、綴りの中の文字を取り出し元の文字の近くに独立形で（あるいは変形して）小書きしたり、意味のない記号を置いたり、すき間を点で埋めたりする。装飾の付け方は書体にも依る。たとえば、ナスターアーク体の書道では母音記号を付けないほうが普通である。

(5) 同綴異語の区別

文脈からは判断が付きにくい同綴異語の場合に、最小限の箇所に母音記号を付けて区別することがある。

たとえば、アラビア語の kataba（〔彼は〕書いた）と kutiba（書かれた）はどちらも كَتَبَ で、たいていは文脈で判断がつくが、分かりにくい場合に kutiba を كُتِبَ と書くことがある（らしい）。

(6) 難読語

難しい動植物名や外国語の固有名詞など、ネイティブの大人でも知っているとは限らないような語に母音記号を付けることがあるようだが、筆者にはよく分からない。

以下では、母音記号とその仲間を具体的に概観する。詳細はそれぞれの言語の章で扱う。

⁽²⁵⁾ 新ウイグル文字は短母音も文字として完全に表記するので、母音記号はない。

⁽²⁶⁾ 余談だが、英語の sherbet と syrup は、この動詞の派生名詞 شربة sharbah と شراب sharāb（いずれも飲み物を指す）に由来する。

2.1 /a, i, u/を表す記号

アラビア語の短母音 /a, i, u/ を表す記号は、以下のように三つある。

記号	音価	名 称
َ	[a]	ファトハ (فتحة fathāh)
ِ	[i]	カスラ (كسرة kasrah)
ُ	[u]	ダンマ (ضمة ḍammah)

ファトハとダンマは文字の上に置き、カスラは文字の下に置く。例として「書かれた」という意味の kutiba を母音記号を付けて記せば كُتِبَ になる。

これらの記号の字形の起源は、それぞれアラビア文字 'alif (ا), yā' (ي), wāw (و) である。

2.2 母音が無いことを表す記号

母音が無いことを示す記号は ' であり、スクーン (sukūn) と呼ぶ。文字の上に置く。

2.3 重子音記号

先に見たように、同じ子音が連続する Muḥammad や awwal (「第一の」) のような単語では、子音を二つ書くことはせず، محمد MHMD とか اول 'WL のように綴る。

このような重子音をシャッダ (شدة shaddah, 原義は「激しさ」) というが、文字表記上は区別しない。これを明確にするため、文字の上に ˆ という記号を付ける。シャッダ記号を付すことをタシユデード (تشديد tashdīd) という。

シャッダは子音に関する記号だが、母音記号の仲間に入れて解説している本が多い。母音記号が通常は書かれないのと同じく、シャッダも通常は書かない。

2.4 /ā/ の綴りに使う記号

ā という音はどのように綴られるだろうか。声門閉鎖音で始まっているからまず 'alif, そして、長母音 /ā/ だからさらに 'alif が来て ʾʾ となりそうなものだが、そうはならない。

この音の綴りは複数あり、クルアーンでは典型的には 'alif 一つで書かれてきた。たとえば, al-Qurʾān が القرآن といった具合である。しかし, ā であることを明確にするため, 'alif の上にマッダ (مدة maddah) と呼ばれる記号を付け, ʾā と書くことになった。先の例では القرآن となる。

今日ではマッダを省略して書くことはない。

3 読誦指示のための記号

クルアーンには、通常のアラビア語文書には記されないような記号が多数使われているが、その多くはどのように読むかを細かく指示するためのものである。調査不足のため、ここでは取り上げない。

4 句読点類

多くの文字体系がそうであるが、アラビア系文字も近代に入るまでいわゆる句読点類 (punctuation marks) は使われなかった⁽²⁷⁾。

アラビア系文字の句読点類も、一足先にそれを整備していたヨーロッパ列強が植民地主義とともに持ち込んだものと思われる。

言語によって違うが、アラビア系文字でもパーレンやピリオド、カンマ、コロンのような記号が使われており、用法もよく似ている。ただし、カンマの向きは「,」を180度回転させたような形「٫」となっている。疑問符が「؟」、セミコロンが「؛」となっているのも面白い。これらの違いはもちろん書字方向に由来する。

アラビア語・ペルシア語において、引用符として、ダブルクォーテーションマークよりもギョメ《》が使われているあたり、フランスの影響がうかがえる。ただし、見てのとおり、通常のギョメ《》と違って丸みを帯びたデザインが普通のようなものである。

(27) 日本語も例外ではない。なお、最近まで目上の人への手紙に句点・読点を使うと失礼に当たるとされていたのは、句読点の歴史が浅いことにも関係がある。

IV 数字

我々が「アラビア数字」とか「算用数字」と呼んで日常的に用いている数字（0, 1, 2, 3, 4, 5, 6, 7, 8, 9）は、インドで発明されて8世後半にイスラーム世界に入り、そこから12世紀頃にヨーロッパに伝わったものである。単に数字が伝わったというよりも10進法に基づく数の表記システム（記数法）が伝わったと捉えるべきである。漢数字やローマ数字とは根本的に異なる記数法であることに注意されたい⁽²⁸⁾。

1 字体

数字の字体が伝播とともに変わっていったので、現在、インド、中東、ヨーロッパではかなり違った字体の数字が用いられている。アラビア語とペルシア語でも一部の字体（4, 5, 6）が違っている。ここでは、ラテン文字、アラビア文字、ペルシア文字、デーヴァナーガリー文字（ヒンディー語）の数字の字体の違いを見てみよう。

なお、ここに挙げた字形は、それぞれの文字の数字が長い年月をかけて変化し、現代において書体としてデザインされたものであるから、この表が字形の変遷を表わしているわけではない。

ラテン文字	0	1	2	3	4	5	6	7	8	9
アラビア文字	٠	١	٢	٣	٤	٥	٦	٧	٨	٩
ペルシア文字	۰	۱	۲	۳	۴	۵	۶	۷	۸	۹
デーヴァナーガリー文字	०	१	२	३	४	५	६	७	८	९

ウルドゥー語では、ペルシア文字とほぼ同じだが、4はやや異なる字体を用いる（技術的問題により掲出しない）。

これらの違いは、単に字体の違いであり、数の表記システムとしては変わりが無い。そこで、本稿では記数法として「インド・アラビア数字」という用語を用い、字体の違いをいうときは、それぞれ「デーヴァナーガリー文字の数字」「アラビア文字の数字」「ラテン文字の数字」という呼び方をすることにする⁽²⁹⁾。

欧米と日本ではラテン文字の数字を「アラビア数字」と呼んでいる。そして、アラビア文字の数字をヨーロッパでは「インド数字」と呼ぶ傾向がある。アラブでも実はこれと同じ用語が使われることがあるようだが、詳細は知らない。

2 用途

アラビア文字・ペルシア文字の数字は、本文中で数を表すのに用いられるほか、当然ながら書籍のノンブルや簡条書きのラベルにも用いられる。

⁽²⁸⁾ただし、漢数字で1968を一九六八などと書くのは、記数法としてはアラビア数字とまったく同じである。

⁽²⁹⁾「〇〇文字の数字」という呼び方もいま一つ妥当性に欠けるが、さりとて適切な用語が無いのが現状である。

3 書字方向

アラビア系文字の中にあっても、数字は左から右へ書く。左が大きい桁、右が小さい桁である。つまり、書字方向は我々の書く算用数字と何ら変わらない。したがって、文章中に2桁以上の数が出てくると、ペンは行きつ戻りつすることになる。

一見不合理のようにも見えるが、我々が縦組みの書籍の中で英単語を90度回転させてまで横組みすると大差無い。左に大きい桁を置くというのはまわりの文字に依らない約束ごとなのだ。

V 書体

イスラーム世界ではカリグラフィー（以下、書道⁽³⁰⁾）が高度に発達した。その理由の一つは、イスラームでは偶像崇拝を禁じているため宗教画というものが基本的にはありえず⁽³¹⁾、神の偉大さ・公正さ・慈悲などを表現しようとする力が文字芸術に向かったことだと言われている。その通りなのだろう。

とりわけクルアーン（の章句または全体）をいかに美しく書くかということは、書道の中心テーマであり続けた。

1 筆記具

アラビア書道では、写真のようなカラム（قلم qalam）と呼ばれる、葦の茎の一端を平らに削って先を斜めにカットしたペンを用いる（写真は竹製）。



伝統的な筆記具「カラム」

先端の形状により太い線と細い線のコントラストが生まれる仕組みは西洋のカリグラフィー（以下、ラテン文字書道と呼ぶことにする）と同じである。ラテン文字書道のペンは、先端の紙に接する部分がペンの軸に対して垂直だが、アラビア書道では斜めになっているところが大きく違っている。これは、ラテン書道では用紙とペン軸の関係を一定に保って書くのが基本であるのに対し、アラビア書道では回転を多用することと関係があるのだろう。先端のカットの角度および厚みは、書く書体や文字の大きさによって変える。

2 書道における書体

筆者にはアラビア系文字の書体を解説する力はないが、代表的な書体の一部を簡単に説明する。

2.1 ナスフ体

ナスフ体（نسخي Naskhi）はアラビア語圏やイランで印刷物の本文に圧倒的に多く使われている書体である。端正で簡素なので習得しやすく読みやすい。アラビア語・ペルシア語の学習書は、ナスフ体で組まれている。比較的活字が作りやすい（活字という制約の枠内で表現しやすい）書体でもある。

伝統的にはクルアーンの写本に多用された。

⁽³⁰⁾ calligraphy という語は語源的には美しく文字を書くことを意味し、「書道」の意味範囲と近い。しかし、この語を聞いて連想するのはもっと狭い意味ではないだろうか。狭義の calligraphy は、用紙に線状接触する先端を持つ筆記具によるものに限られ、毛筆・硬筆書道を含まない。アラビア系文字の書道は狭義のカリグラフィーにあてはまる。

⁽³¹⁾ とはいえ、現在イランの街角では預言者ムハンマドや第4代正統カリフ、アリーの肖像画（もちろん想像図）が売られているし、宗教的場面を描いたミニアチュール（細密画）も作られてきた。何事も例外はつきものだ。

2.2 スルス体

スルス体 (ثلوث Thuluth) は、クルアーンの章句を書くために開発された書体で、鋭く緊張感のある美しさを持つ。印刷物では、タイトル・見出しによく使われる。スルスとは 1/3 という意味で、これはかつて、基準のカラム (ペン) に対し 1/3 の幅を持つカラムで書いた書体であったことに由来している [8]。スルス体からはさまざまな書体が生まれた。

2.3 クーファ体

クーファ体 (كوفي Kūfi) はイラクの古都クーファで発展した書体であり、アラビア文字の歴史のごく初期からある。画線が太く、力強い書体である。極めて多様なバリエーションを生んだ。その中には筆書書体もあれば、装飾的なレタリング書体もあり、現代でもたとえば社名のロゴなどいろいろな場面で使われている。看板などにおいてはヨーロッパのサンセリフ体や日本のゴシック体に近い使われ方をしているようである。

2.4 ナスタアリーク体

ナスタアリーク体 (نستعلیق Nasta'liq) はペルシア語圏で開発されたのでペルシア書体 (فارسی Fārsī) ともいう。たおやかな曲線をもつ流麗な書体である。現在もイランはもとより、かつてのペルシア語文化圏 (イラン, アフガニスタン, パキスタンなど) で多用されている。

2.5 ルクア体

ルクア体 (رقة Ruq'a) は極めて簡素な書体で、流麗さや豪華さは無いが力強く直線的な美しさがある。看板に使われたり、値札などに使われたりする。また、アラブにおける日常の手書き文字はルクア体をベースにしている。

二つあるいは三つの識別点を簡略化して線で表現するなど、速書きのための工夫がされている。

3 印刷書体

ここでいう印刷書体とは、活版や写植、コンピューター組版などにおける書体のことである。英語で言えば typeface であり、書道でいう書体 (writing style) とは分けて考えなければならない。漢字やラテン文字と比較すると分かりやすいので、回りくどくなるが、これらと対照しながら考えてみる。

3.1 漢字書体

まず漢字書体について考えてみよう。印刷書体である明朝体はどんな筆記具を使っても決して手では書けない⁽³²⁾。ウロコ (横画の右端にある三角形の部分) や払いの形状は、もちろん毛筆の書法に由来するものではあるが、著しく様式化されており、もはや毛筆で再現することはできなくなっている。宋朝体などについても同じことがいえる。

それに対し、楷書体などの毛筆書体は、書道の書体と印刷書体の間に大きな開きはなく、おおむね「書くとおりの」形を印刷文字として作る⁽³³⁾。

また、毛筆とは無関係の書体として、ゴシック体などさまざまな書体がある⁽³⁴⁾。

(32) 「描く」ことならできる。これはレタリングである。

(33) ただし、印刷書体では正方形の枠にうまく収まるように文字の形を決めるという点で書道の書体とはやはり違う。

(34) 正確に言えば、ゴシック体といえども筆の痕跡をもつものがあり、まったく無関係ともいえない。

以上をまとめると、印刷書体には、(1) 書道の書体に由来しながらも、独自の様式化を経たもの、(2) 書道の書体にほぼ忠実なもの、(3) 書道とは無関係に創案されたもの、があるといえる。

3.2 ラテン文字書体

ラテン文字はどうか。ローマン体は印刷書体である。その大文字はローマ時代に石碑に彫刻された書体に基づいており、(手書きと無関係ではないが) 手で書けるものではない。小文字はカロリング朝風ミヌスケール (Carolingian minuscule) という手書き書体に基づいているが、様式化されている。ニコラ・ジャンソン (Nikolas Jenson) が作った世界最初のローマン体活字は、まだ手書き風の特徴を色濃く残しているが、ローマン体は時代とともに様式化が進んでいく。

それに対し、カーシヴ (cursive) と分類される書体は美しさと実用性 (速筆性) を兼ね備えた手書き書体であり (とはいえ日常書体ではなく訓練された写字生が書写を行うためのもの)、これがほぼそのまま印刷書体となっている。つまり、漢字の毛筆書体と同じく、書道の書体と印刷書体の間に大きな開きが無い。

ラテン文字でも、書道とはまったく無関係の書体として、サンセリフなどさまざまな書体がある。

3.3 アラビア系文字書体

漢字とラテン文字を調べたと同じ観点でアラビア系文字を考える。

アラビア系文字にも、書道とは無関係の印刷書体が作られているのは確かである。Tahoma のようなサンセリフも作られている。

しかし、多くの印刷書体は書道の書体をそのまま再現しようとしているように思われる。それだけ書道が求心力を持っているといえるのではないか。

ただし、歴史的にはアラビア系文字の活版書体のデザインは稚拙なものが多かった。例外的に、19 世紀前半にオスマン朝の書家エフェンディーが書いた文字を元にした活版書体は高く評価されている [4]。

Times New Roman と Tahoma

Windows でアラビア語を扱ったことのある人にとって、一番なじみがあるアラビア系文字書体は、OS に初めから入っている Times New Roman と Tahoma ではないだろうか。いずれもラテン文字・ギリシア文字・キリル文字の書体にアラビア系文字を追加したものである。それぞれの書体で صباح الخير (おはよう) を組んだサンプルを次の図に示す。

Times New Roman صباح الخير

Tahoma صباح الخير

Times New Roman は、元は Monotype 社の活版書体であり、ラテン文字の設計はスタンリー・モリスン (Stanley Morison) が主導した。新聞の本文に使う書体として開発されたものの、広い分野で使われ、現在にいたるまで人気がある。

しかし、アラビア系文字の部分はとても美しいとはいえず⁽³⁵⁾、間に合わせに作った感を免れない。黎明期を過ぎたいま、もはやアラビア系文字の印刷物に Times New Roman を用いるべきではない。

Tahoma のラテン文字・ギリシア文字・キリル文字は現代の書体デザイナー、マシュー・カーター (Matthew Carter) がデザインしたもので、端正な美しさと柔らかさがある。アラビア系文字は Monotype 社のスタッフがそれに合わせてデザインしたとされるが [4]、両者のデザインに一貫性が

(35) 少ないグリフ数でどうにか表示できるよう、文字の形を著しく歪めている。

あるようには見えない。共通点はサンセリフであることと、カウンター（画線に囲まれた領域）が広めであることぐらいではないか。アラビア系文字は漫画チックな（もっと言えばふざけた）印象を与える。あまりよく出来た書体とは思えず、印刷物では使い道が無い。

ただし、画面表示用としては別の価値がある。ビットマップフォントがよく出来ているので、小さい文字サイズでも非常に見やすく、ウェブのブラウザには適していると思う。

まともなデジタルフォントはあるか

デジタルフォントで、書体デザインの優れたものはどれか、という問いに筆者は答えられない（自信をもって言えるほど書体が分かってはいない）。

しかし、DecoType 社の一連の書体、DecoType Naskh, 同 Naskh Swashes, 同 Thuluth は一見してかなり品質が高いと思う（残念ながら市販はされていない）。Lotus Linotype や Naskh Ahmad Arabic も、使って恥ずかしくないナスフ体であると思う。グリフ数の多さで言えば Microsoft の Arabic Typesetting となる。本稿はさまざまな言語を扱うので、不可避免的に Arabic Typesetting を採用した。

4 ボールドと斜体

アラビア系文字でも、サンセリフ体などでは、一つの書体で複数のウエイト（太さ）の書体をデザインすることが考えられる。

4.1 ボールド

しかし、カリグラフィックな書体⁽³⁶⁾ではどうだろう。書道理論的には、カラムの先端の幅では文字の大きさが決まってしまうため、同じサイズで太さだけが異なる文字を書くことはしない。

したがって、カリグラフィックな書体にウエイトの概念を持ち込むのは無理がある。しかしながら、活版時代からラテン文字に倣ってボールド体が作られてきた。一つの書体に、(1)通常の太さのもの(2)太めにデザインしたものの二種類を用意することはかまわないのではないだろうか。

機械的ボールド

太くデザインしたフォントがあればよいが、無ければアプリの機能で機械的に太らせることになる。こうして作った機械的ボールド体を揶揄して一般に poor man's bold と呼ぶ。

どんな文字の書体においても、poor man's bold はあまり好ましくないが、アラビア系文字の場合とくに気をつけるべき点がある。それは、fā' (ف) や wāw (و) のループ（閉じた円状部分）がつぶれやすいこと、thā' (ث) などの三つ組の識別点がくっつきやすいことである。

実際、アラビア語やペルシア語の教科書・辞書においてこれが使われ、文字がつぶれて汚らしく読みづらい例を非常によく見かける。ほんのちょっとしたタイポグラフィーの不具合でも初学者は著しく困難を感じるものなので、よく注意していただきたい。

どうしても poor man's bold を使うのであれば、文字のサイズや印刷の再現性を十分に考慮して、その書体でどこまで太らせられるのか、よく調べるべきである。

4.2 斜体

ラテン文字の場合

ラテン文字では斜体を二種類に分類することができる。一つは直立した書体を機械的に傾けたもの。これはスラント (slant) と呼ばれる。細かく分ければ、機械的に傾けたようにデザインし、そ

⁽³⁶⁾ここではカラムで書かれたような（カリグラフィーに忠実な）デザインの書体を指す。

それをフォント化したものと、直立書体を写植の変形レンズや DTP の変形機能を使ってその場で機械的に傾けたものがある。後者を機械スラントなどと呼ぶ。もう一つはイタリック体であり、これは斜めに傾いた独立の書体であり、ローマン体とは別の起源を持つ⁽³⁷⁾。

イタリック体は、ローマン体の文章の中で、強調部分・引用部分や書名、固有名詞、学名などに用いる。

アラビア文字の場合

アラビア系文字についてはどうか。

ナスフ体やスルス体は縦線が一定の傾きを持っているが、これをラテン文字におけるイタリック体と同列に扱うことはできない。また、ラテン文字と同様の目的で二書体を混植することも、筆者の知るかぎりでは無い。

しかし現在、文献リストにおける書籍名などで機械スラントが用いられることはある。用法はラテン文字におけるイタリック体と同じである。これはおそらく DTP 以降の現象と思う。カリグラフィックな書体を機械スラントにすると非常に醜く、筆者は賛成しかねる。ラテン文字の諸概念を、普遍的と勘違いして安直に他の文字文化へ持ち込むべきでない⁽³⁸⁾。

しかし、イランの印刷物でそういうものを見たことがあるので、現実に使われているようだ。

5 和文との混植

和文とアラビア系文字を混植する際、ラテン文字の場合と同様に和垂間を少し空けるのが望ましい。このほか、文字サイズとベースラインの調整、そして行間設定にも注意を要する。

5.1 文字サイズの釣り合い

文字サイズとは何だろうか。

和文の場合、(プロポーションナル書体⁽³⁹⁾を除いて)文字は正方形の枠を基準としてその中に収まるようデザインする。組版上はこの正方形がぴったり接するようにグリフを並べるのが基本の組み方である。つまり、この枠はデザインの基準でもあり、組版の基準でもある。活版の場合はこの正方形は活字の軸(ボディー)そのものである。写植以降は現実のボディーを持たないが、基準となる正方形を仮想ボディーと呼んできた。

ともかく、この(仮想)ボディーの一辺の長さが文字サイズなのである。10ptの文字とは、この長さが10ptである文字に他ならない。

欧文の場合、文字幅が文字によって異なるが、活版では活字の(字面に対する)天地方向の大きさは一定であり、これをもって文字サイズとする。写植以降はボディーがないが、書体デザイナーが10ptと定めた大きさが10ptなのだ。これはアラビア系文字についても同じである。

和文の中にアラビア系文字を混植する場合、それぞれの文字サイズが同じだと、アラビア系文字のほうが小さく見えるのが普通である。これはもちろん書体に強く依存する。

本稿では、和文書体に「本明朝-M小がな」、アラビア系文字書体に「Arabic Typesetting」を用いており、和文に対してアラビア系文字を1.3倍にして混植している。それでもなお、アラビア系文字のほうが小さく見えるだろう。Arabic Typesettingはアラビア系文字書体の中でもとりわけ小振りに見える書体だからだ⁽⁴⁰⁾。ではもっとアラビア系文字を大きくしたらどうか。そうすると母音記

⁽³⁷⁾ なお、イタリック体は本来は小文字書体である。大文字のイタリックは後でできたもので、デザイン的にはスラントに過ぎない。

⁽³⁸⁾ しかし、この流れは止められそうにない。デーヴァナーガリー文字においても、文献名に斜体を用いたものを見かける。日本語の場合、印刷物においてはさすがにこのような馬鹿げた斜体は使われていないが、ウェブの世界では珍しくなくなっている。

⁽³⁹⁾ 主として仮名に固有の文字幅を持たせた書体

号やシャッダなどが重なったときに、アラビア系文字の部分が天地に大きくなって前後の行に重なる恐れがあるのである。

5.2 行間

一般にアラビア系文字は行間を広めにとらざるを得ない。背の高い文字に گ などがあり、底の深い文字には ع などがある。こういった文字に母音記号などが付くと、さらに上や下に張り出すので、前後の行でぶつからないようにしなければならない。厄介なのは和文中に混植する場合で、アラビア文字に合わせると和文がスカスカになってしまう。とって、アラビア文字が現れる行の前後だけ行間を開くのも格好が悪い。

⁽⁴⁰⁾ なお、Arabic Typesetting のラテン文字は、アラビア系文字に合うように小振りに作られているので、同じ文字サイズの普通のラテン文字フォントと比較して非常に小さく見える。

VI 組版・入力・校正上の注意

1 識別点の有無

識別点の有無はよく気をつけなければならないが、これが間違っている日本の印刷物をよく見かける。

1.1 アラビア語

アラビア語の手書き原稿では、ある種の識別点を省略することがよくあるので、注意を要する。具体的には、(1) yā' (ي) の独立形・右接形における識別点、(2) ター・マルブータ (ة) の識別点、である。それぞれ、識別点を省略すると別の文字になってしまうが、ネイティブなら正しく読める。つまり、単語を知っていないと、原稿を見て正しく入力することすらできないわけである。

1.2 ペルシア語

ペルシア語では ye (ی) の独立形・右接形には識別点を付けないが、付けているものをよく見かける。これはアラビア語キーボードでペルシア語を入力したときによく起こる問題である。また、OS やブラウザーなどの環境によってはこの文字が正しく出ないため、ウェブでは確信犯的にアラビア語の yā' (ي) を使っていることが多い。

2 字形の混同

2.1 fā' と ghayn の両接形

既に述べたが、fā' と ghayn の両接形はよく似ている。文字が小さすぎたり製版が不良だったりして印刷が不鮮明だと見分けが付かなくなるので注意。実際、辞書や参考書でこれが見分けにくいものが多くあり、非常に困る。ネイティブは問題なく読めても初学者にはたまったものではない。

2.2 lām-mīm 合字

次の図版は、ペルシア語の学習書から取ったものである。

(200%に拡大)

ペルシア語で「はい」という意味の **بلى** (bale) が組んであるが、すき間があいているように見える。これはどうしたことだろうか。実は手動写植機で打つ字を間違えたのである。lām (ل) の両接形 (ل) を打ったつもりで、形の似た lām-mīm 合字の左接形 (ل) を打ってしまったものだ。この本では徹頭徹尾このように打っていた。

2字目が mīm (م) の合字は mīm の部分が簡素な形をしているので、慣れないと分かりにくいかもしれない。

この種の誤植は、XSL Formatter や Microsoft WORD のような、続け書きを自動的に処理するアプリでは生じ得ないが、何らかの事情で手作業で字形を選択して組む場合には注意を要する。

2.3 ṭā'

ṭā'の独立形 (ط), 左接形 (ط), 両接形 (ط), 右接形 (ط) は, 互いによく似ている。lām-mīm 合字について述べたと同じような注意が必要である。実際, アラビア系文字を解説した書籍で, これらを混同して組まれたものをよく見かける。

VII アラビア語

1 言語の概要

1.1 使用地域と人口

アラビア語は六つある国連公用語の一つであり、アラブ 22 ヶ国などで使われている。また、聖典クルアーンの言語として、世界各地のイスラーム教徒に親しまれている。母語話者は 1 億 5000 万人程度と推定されている。共通文語ともいえるフスハー（正則アラビア語；ニュースや演説などかきこまった場面では口頭でも使われる）と、それぞれの地域で話されている口語（アーンミーヤ）に分けられる。書かれるのは基本的にはフスハーであるが、エジプトなどのアーンミーヤはある程度標準的な書き方があり、文章に記される。しかし、本稿はフスハーのみを対象とする。

1.2 言語の系統

アフロ=アジア語族セム語派に属す。セム語には他に、ヘブライ語、アラム語（イエス・キリストが話した言語）などがある。

イスラームの拡大にともなって、他の言語の話者にも普及し、ペルシア語やトルコ語などに大きな影響を与えた。また逆に、それらの言語からの語彙の流入も起こった。

1.3 音韻

子音はアラビア文字 28 字に対応する 28 個（ただし、'alif が声門閉鎖音に対応すると考えた場合）。母音は短母音が/a, i, u/, 長母音が/ā, ī, ū/の計 6 個。

アラビア語に特徴的な子音として咽頭化音が挙げられる。/t, d, s, z, th, dh/ のそれぞれに対応する咽頭化音がある。咽頭化音とは調音時に舌根を咽頭に近づけて発音する子音である⁽⁴¹⁾ [16]。

2 文字

アラビア語特有の文字の問題について述べる。

アラビア文字の文字数は、後述するハムザを除けば 28 字とされており、「文字」の章に掲げた表の通りである。しかし、ハムザのほか、この 28 字の中に数えられていない文字ならびに記号付き文字が存在する。これをまず表に示す。

文字	名称		右接形	両接形	左接形
آ	الف مده	'alif maddah	آ		
ى	الف مقصورة	'alif maqṣūrah	ى		
ة	تاء مربوطة	tā' marbūṭah	ة		
ء	همزة	hamzah			

初期のアラビア語の表記は変則的な点が多いあり、クルアーンの表記も綴りと発音が必ずしもきれいに対応していない。これにはいくつかの原因があり、(1)元になったナバテア語の綴りをそ

(41)咽頭化音はアラビア語を母語としない者には発音が困難なので、アラブ人は自らを「ダード (ض dād) の民」(اهل الضاد 'ahl 'al-ḡād) と呼び、アラビア語を「ダードの言葉」(لغة الضاد lughat 'al-ḡād) と呼んでいる。ḡād が「難しい音 (咽頭化音)」の代表なのだろう。

のまま用いた, (2) 文脈形と停止形とで発音が変わるアラビア語特有の音韻現象を反映した, (3) 声門閉鎖音が脱落したり前後の音が長母音化するアラビア半島西部の発音を反映したことなどにある [1]。こういった問題への対処として上記のような文字・記号ができたわけである。

こうして確立した正書法は, 現代まで基本的に継承されている。

2.1 アリフ・マッダ

綴りと発音の食い違いが大きいのは, 長母音 /ā/ の表記である。このうち, /ā/ が 'alif 一字で表されているものについては, その 'alif の上にマッダ (مَدَّة maddah) と呼ばれる記号 (◌) を付加した。例: قرآن qur'ān (クルアーン)

マッダは, クルアーンでは ي の上に付くこともあるが, 例外的。この場合, ي に識別点は付かない。

2.2 アリフ・マクスーラ

語末が長母音 /ā/ であるにも関わらず, これを yā' (ي) で表記する単語がある。これは語源であるナバテア語の綴りを反映したものとされる。例: موسى (mūsā, モーセ; 人名)

発音どおりならアリフ (ا) を書くべきところなので, アリフの変形と見立てて, これをアリフ・マクスーラ (الف مقصورة 'alif maqṣūrah) と呼ぶ。「縮められたアリフ」の意である。

yā' は, 識別点の導入時より ي と表記されるようになったが, アリフ・マクスーラは現在でも識別点を付けないので, 容易に区別できる。

ただし, 手書きでは ي のヌクタを省略することがあり, その場合は単語を知らないと区別できない。アリフ・マクスーラは U+0649 である。

2.3 ター・マルブータ

女性名詞および形容詞の女性形の語尾において, 綴り上は ha' (ه) で書かれるのに, /t/ の音を表している文字がある。これはのちに識別点を二つ載せて: と書かれるようになった。音価がこと同じであり, 形があたかも t を丸めたように見えるのでター・マルブータ (تاء مربوطة tā' marbūṭah) と呼ばれる。「結ばれたター」の意である。

このような変則的な綴りはアラビア語の音韻現象に基づいている。用語集の「ワクフ」の項にその事情を書いたので参照されたい。

2.4 ハムザ

ハムザ (همزة hamzah) という記号 (ء) も, アラビア語の正書法が確立してゆくなかで創案された記号であり, 声門閉鎖音を表す。

セム系文字の伝統からいえば声門閉鎖音は 'alif が表すはずであるが, クルアーンではこれが表記されていないことが多い。そこで, 適宜ハムザを単独で置いたり, ا, و, ي の上にハムザを載せることで声門閉鎖音を表すことにした。ただし, 語頭の /i/ については, 'alif の下に置き, اِ のようにする。また, ハムザの載った ي は識別点を付けないことに注意されたい。

ハムザの導入により, 'alif 自体は音価を持たないことになった。「ハムザを載せる台」に成り下がったのである。

母音記号を付けない一般の文書であってもハムザは省略しない。

2.5 太陽文字と月文字

アラビア文字 28 字のうち, 次の 14 字を太陽文字という。

ن، ل، م، ط، ض، ص، ش، س، ز، ر، ذ، د، ث، ت

残りの14字は月文字（または太陰文字）である。

太陽文字か月文字かは、実は文字の問題というよりアラビア語の音韻の問題であり、タイポグラフィ上は知らなくてよいが、簡単に説明する。太陽文字で始まる単語に定冠詞'al (ال) が付くと、lの音はその太陽文字の音に同化する。たとえば、「アッサラーム」はsalām سلامに定冠詞がついて、al-salām→assalāmとなったものである。本稿では、この手の単語のラテン文字表記をal-salāmの形で示すので、発音とのずれに注意されたい。

学習者は真っ先に太陽文字を覚えなければならないが、何のことはない、lām (ل) に調音点の近い文字が太陽文字なのである。

3 記号と句読点類

3.1 母音記号

タンウィーン

母音 /a, i, u/ を表す母音記号、ファトハ、カスラ、ダンマを「文字」の章で扱ったが、アラビア語ではこのそれぞれに対応するタンウィーン記号というものがある。詳細は用語集を参照されたい。字形は、基本的に母音記号を二つ重ねた形であるが、ダンマのタンウィーン記号はさまざまな変種がある。ナスフ体では◌◌の形が一般的。

短剣アリフ

アラビア語における長母音 /ā/ の表記は不完全で、あるべき'alif がしばしば省略されている。そこで、/ā/ の存在を示唆するために、該当子音字の上に'alif を小書きする。例：الرحمن al-rahmān 「慈悲深きもの（＝アッラー）」⁽⁴²⁾

この記号の名称がいまいちはっきりしないが、本稿では仮に「短剣アリフ」（出典失念）とした。短剣アリフを省略するか否かは他の母音記号に準じる。

3.2 シャツダ

「文字」の章で見た如く、同じ子音が重なるとき、綴り上は子音が一つであるのと変わらないので、重子音であることを明示するために文字の上にシャツダという記号 (◌◌) を載せる。シャツダを省略するか否かは母音記号に準じる。

3.3 ワスラ

語頭の声門閉鎖音で始まる音節のうち、前の単語と続けて読む場合に、脱落するものがある。このような場合に、その'alif の上にワスラ (◌◌) という記号を載せ、◌◌のようにする。たとえば、定冠詞ال'al の /a/ や人名によく出てくる ابن'ibn の /i/ などは、先行語と続けて読むときに脱落する。用語集の「ワスラ」参照。

3.4 クルアーン読誦のための記号

クルアーンを正確に読誦するための記号がいくつかあるが、調査不足のため省略する。

3.5 句読点類

句読点類には次のものがある。

⁽⁴²⁾ この例では、他の母音記号は省略した。また al-rahmān の実際の発音は、r が太陽文字なので arrahmān となる。

記号	意味	アラビア語名称	説明
.	終止符	نقطة nuqṭah	使い方は英語などと同様。Unicode ではラテン文字と共用。
،	カンマ	فاصلة fāṣilah	使い方は英語などと同様。
!	感嘆符	علامة التعجب 'alāmat al-ta'ajjub	使い方は英語などと同様。Unicode ではラテン文字と共用。
؟	疑問符	علامة الاستفهام 'alāmat al-istifhām	使い方は英語などと同様。
:	コロン	النقطتان al-nuqṭatāni	使い方は英語などと同様。Unicode ではラテン文字と共用。
؛	セミコロン		使い方は英語などと同様。
« »	ギョメ	مزدوجان muzdawjāni	引用符。デザイン的には、ラテン文字のようにとがったものではなく、二重パーレンのように丸みを帯びたものが普通。Unicode ではラテン文字と共用。
—	ダッシュ		調査不足のため詳細不明。

3.6 慣用句の省略記法

「～など」を表す英語の「etc」は、ラテン語の「et cetera (およびそのような [他のもの])」からの借用語である。これと似たものがアラビア語にもある。الع と (上に線を引いて) 書けば、これは إلى آخره 'ilā 'ākhirihī のことで、「…等々」という意味である [1]。

このように、慣用句の省略記法は上に線を引いて示す。

4 書体

印刷物の本文の基本書体はナスフ体であり、見出しにはスルス体が使われる。看板にはディーワーン体やクーフア体、ルクア体なども使われる。書体の使い分けについては調査不足。

5 正書法

正書法の詳細は調べきれなかった。正書法のうち、個別の単語をどう綴るかという問題（正綴法）については、「文字」の節で部分的に述べた。ここでは分かち書きについて述べる。

5.1 分かち書き

アラビア語は、かつてはワードスペース無しに書かれたが、現代では基本的に単語間に空白を入れる。しかし、定冠詞 ال'al の後や、一文字からなる前置詞، wa (～と)、ب bi (～において) などは、後ろの単語に続けて書く。例：اهلاً وسهلاً 'ahlan wa sahlan (ようこそいらっしゃい)、النيل al-nīl⁽⁴³⁾(ナイル川)

⁽⁴³⁾ 子音の同化により、実際の発音は annīl となる。

VIII ペルシア語

1 言語の概要

1.1 使用地域と人口

ペルシア語 (فارسی Fārsī) はイラン・イスラーム共和国の唯一の公用語で、人口約 7000 万人のうちの約半数がこれを母語とする。識字率は過去半世紀に急速に向上した。1956 年には都市部で 33.3%、農村部で 6% だったが [1]、現在は全体で 80% 程度である。

アフガニスタンで話されるペルシア語はダリー語 (دري Dari) と呼ばれ、イランのペルシア語とは発音や語彙に違いがあるが、方言程度の差である。ダリー語は、パシュトー語 (パシュトゥーン人の母語) とともにアフガニスタンの公用語であるだけでなく、民族間の共通語的役割をもち、バイリンガルも多い。ダリー語はペルシア文字で書かれる。アフガニスタンの (とりわけ女子の) 識字率はかなり低いと推定されている。

タジキスタンで話されるペルシア語はタージーク語と呼ばれ、イランのものとは発音や語彙に違いがあるが、これも方言程度の差である。タジキスタンは旧ソ連邦に属していたため、タージーク語は現在キリル文字で書かれている。

上記以外にもウズベキスタンなど中央アジアなどにペルシア語の話者がいる。

1.2 言語の系統

印欧語族イラン語派に属し、パシュトー語、クルド語などと近い。

7 世紀にペルシアの王朝、サーサーン朝がアラブ・イスラーム軍に敗れ (642 年、ネハーヴァンドの戦い)、その支配を受けた。これにより中世ペルシア語が多量のアラビア語彙を受容して近世ペルシア語が成立したとされる⁽⁴⁴⁾。これに伴い、それまでのパフラヴィー文字 (アラム文字を改変したもの) を廃し、アラビア系文字を採用した。

アラビア語とペルシア語の関係は、漢語と日本語の関係に似ている。漢語系語彙抜きに現代日本語が成り立たないように、近世ペルシア語からアラビア語彙を取り去ることはもはやできない。

近世ペルシア語には千年以上の歴史があるが、現在のものをとくに現代ペルシア語と呼ぶ。本稿で扱うのは現代ペルシア語である。日本語がこの千年の間に語彙・文法・発音の面で大きく変化したのに比べ、ペルシア語はさほど変化していない。

1.3 音韻

現代ペルシア語の音素は以下のとおりである [11]。

子音は /p, b, t, d, č, j, k, g, f, v, s, z, š, ž, x, q, m, n, l, r, y, h, ʔ/ の 23 個。

母音は、短母音 /a, e, o/, 長母音 /ā, ī, ū/, 二重母音 /ey, ow, ay, ây, uy, oy/ である。

それぞれがどのような音であるかには立ち入らないが、/ā/ についてだけ注意する。この音を /ā/ のように書かないのは、これが短母音 /a/ を伸ばした音ではないからである。

ペルシア語の音の表記法はいろいろある。本稿で採用した表記法は、都合によりアラビア語とペルシア語でまったく違っている。たとえば、アラビア語の表記に用いた kh (無声軟口蓋摩擦音) はペルシア語にもあるが、表記は x としている。

(44) サーサーン朝崩壊後、最初のペルシア語文献が現れるまでに約 200 年を要しており、この期間を「沈黙の二世紀」と呼ぶ。

2 文字

ペルシア語を表記するためのアラビア系文字をとくにペルシア文字という。基本の文字表を次に示す。

文字	名称	転写	右接形	両接形	左接形
ا	'alef	'	ا		
ب	be	b	ب	ب	ب
پ	pe	p	پ	پ	پ
ت	te	t	ت	ت	ت
ث	se	s	ث	ث	ث
ج	jīm	j	ج	ج	ج
چ	če	č	چ	چ	چ
ح	he hottī	h	ح	ح	ح
خ	xe	x	خ	خ	خ
د	dāl	d	د		
ذ	zāl	z	ذ		
ر	re	r	ر		
ز	ze	z	ز		
ژ	že	ž	ژ		
س	sīn	s	س	س	س
ش	šīn	š	ش	ش	ش
ص	sād	s	ص	ص	ص
ض	dād	d	ض	ض	ض
ط	tā	t	ط	ط	ط
ظ	zā	z	ظ	ظ	ظ
ع	'eyn	'	ع	ع	ع
غ	qeyn	q	غ	غ	غ
ف	fā	f	ف	ف	ف
ق	qāf	q	ق	ق	ق
ک	kāf	k	ک	ک	ک
ل	lām	l	ل	ل	ل
م	mīm	m	م	م	م
ن	nūn	n	ن	ن	ن
و	vāv	v	و		

ه	he havvez	h	ه	ه	ه
ی	ye	y	ی	ی	ی

アルファベットは、元祖アラビア文字の 28 字に、ペルシア語特有の子音/p, ç, ž, g/を表すための 4 字 پ, چ, ز, گを追加した 32 字⁽⁴⁵⁾。追加文字の前三者は既存文字の識別点を増やしたもの。گは既存文字 (ک) に短い線を追加したものである。この短い線を母音記号のファトハと混同してはならない。

子音音素がアラビア語よりずっと少ないにも関わらず文字数が増えているのは、アラビア語からの借用語の綴りを温存するためである。そのため以下のように多数の同音の文字がある。

/f/ : ع, ف

/t/ : ط, ت

/s/ : ص, ث, س

/h/ : ح, ه

/z/ : ظ, ز, ذ

以上のうち、アラビア語からの借用語でない元来のペルシア語の綴りには、基本的には上のリストの各行の先頭の文字 (ت, س, ز) が使われる。つまりリストの各行の 2 番目以降の文字 (ط, ث など) を含む単語はたいていアラビア語からの借用語である。ただし、わずかに例外もある。صد (sad: 百) は元来のペルシア語だが ص を含む。

ه と ح は発音が同じため、文字名称がどちらも he となるが、これをそれぞれへ・ハヴヴェズ (he havvez), へ・ホッティー (he hottī) と呼んで区別する。

なお、ع と ق は本来別の子音だが、現代ペルシア語では区別が無くなっている。この音素を /q/ のように書く。ただし、地方方言やダリー語などにはこの区別が残っている。

2.1 文字表の配列

文字表の配列において、ペルシア語で追加された 4 文字はそれぞれの基字のグループの末尾に挿入されている。

それ以外でのアラビア語との順序の違いは hā'/he (ه) と yā'/ye (ی) の順序が入れ代わっていることだけである。

2.2 字体についての注意

ye の字体

ペルシア語では ye (ی) の独立形・右接形に識別点を付けない。

言語	独立形	右接形	両接形	左接形
アラビア語	ي	ي	ي	ي
ペルシア語	ی	ی	ی	ی

アラビア語の yā' (ي) との違いは機能面ではとくに無く、単なる異体字と考えてよい。しかし、Unicode ではペルシア語の ye のためにコードポイントが用意されており、U+06CC ARABIC LETTER FARSI YEH と定義されている。

識別点を付けないので、語末ではアリフ・マクスーラと区別がつかない。

⁽⁴⁵⁾ これらは一度に追加されたのではなく、さまざまな試みを経て徐々に整備されたものである。

kâf の字体

アラビア語の kâf の独立形と右接形は ك, ڪ であるが、ペルシア語ではどちらかと言えば ک, ڪ のほうが普通である（これも単なる異体字に過ぎない）。

言語	独立形	右接形	両接形	左接形
アラビア語	ك	ڪ	ڪ	ڪ
ペルシア語	ک	ڪ	ڪ	ڪ

日本人むけ学習書では、どういうわけか、独立形を ک としておきながら、右接形を ڪ とする、いわば混用した組版をよく見かける。

2.3 無音の he

語末の he havvez (ه) は、/h/ の音を表わす場合と、それ自体は音価を持たない場合とがある。前者を有音の he、後者を無音の he と呼ぶ。無音の he の場合、わずかな例外を除いて、その前の音節の母音は /e/ である。したがって、無音の he は、母音 /e/ を示唆する符丁と見することもできる。

有音の例として、ماه/mâh/「月」、無音の例として、سه/se/「三」、بچه/bačče/「子供」を挙げておく。

3 記号と句読点類

3.1 母音記号

母音記号はアラビア語とほぼ同じであるが、表す母音が異なる。基本的には、ファトハ、カスラ、ダンマが /a, e, o/ に対応している。スクーンの記号・用法はアラビア語と同じである。

アラビア語とは記号名称の発音がやや違うので、その対照表を掲げておく。

記号	アラビア語	ペルシア語
ˆ	فتحة fathāh ファトハ	فتحه fathe ファトヘ
˙	كسرة kasrah カスラ	كسره kasre キャスレ
˘	ضمة ḍammah ダンマ	ضمه zamme ザンメ
◌	سكون sukūn スクーン	سکون sokūn ソクーン

タンヴィーン記号

アラビア語からの借用語で、名詞の対格・非限定の形に由来する副詞が多数ある。アラビア語文法では対格・非限定は -an で終わり、ファトハのタンヴィーン記号が付くが、これを借用したペルシア語においてもやはり表記する。タンヴィーンはペルシア語ではタンヴィーン (tanvīn) となる。

母音記号を書かない一般文書においても、普通は省略しない。たとえば、تقریب taqrīb 「近似、接近、概算」の対格 تقریبان taqrībān は「およそ、だいたい」という意味の副詞になる。

ファトハ以外のタンヴィーン記号はペルシア語では使わない。

さて、アラビア語の正書法の原則からいえば、このタンヴィーン記号は 'alef ではなくその前の文字に付くはずである。しかし、ペルシア語では（アラビア語にも見られるが）上の例のように 'alef の上にタンヴィーン記号を載せたもののほうがよく目にする。どちらが正しいかについての合意はどうか無きようである。しかし、少なくとも一つの文書を通して統一する必要がある。

タンウィーンはもともとアラビア語文法上の現象だが、そのうち名詞の対格・非限定の形をした副詞がペルシア語に入ったもので、もはやファトハのタンウィーン記号は単なるお約束でしかない。母音記号の原則から外れた書き方をするのはそのような事情もあってのことではないだろうか。

3.2 ハムゼ

アラビア語のハムザはペルシア語ではハムゼ (همزة hamze) という。

ペルシア語では、語頭の'alifの上下に付けたりすることはない。アラビア語と違って'alif自体が声門閉鎖音を表す子音字だからである。語中では、アラビア語と同じく'alif (ا), vāv (و), ye (ی)の上に付けることがある。しかし、アラビア語と違い、'alefの下に付けることはない。語末に単独のハムゼが置かれることもある（正書法の項参照）。

無音の he につくハムゼ

ハムゼのペルシア語特有の用法として、語末の無音の he の上に置く場合がある。これについて述べる前にエザーフェについて説明しなければならない。エザーフェは、名詞同士や名詞と形容詞を結ぶ働きをする母音 /e/ のことである。たとえば私 (من man) の名前 (اسم 'esm) であれば'esme man のように、'esm の末尾に /e/ が入る。この場合はエザーフェは表記されず、من اسم のように 2 語を並べただけの綴りとなる。

先行語が母音で終わっている場合には /e/ ではなく /ye/ が付くが、この場合は綴り上、何らかの表示がある。語末が /e/ の語は、綴りは無音の he で終わっているが、これにエザーフェ /ye/ が付くと、he の上にハムゼが置かれる。たとえば、「私の家 (خانه xâne)」であれば، من خانه xâneye man となる。しばしば省略されるので، من خانه と綴っても間違いではないが、少なくとも文書全体で統一する必要はある。

このハムゼは声門閉鎖音とはとくに関係が無い。

3.3 他の記号

シャッダ、マッダ（それぞれシャッデ、マッデと呼ぶ）はアラビア語と同じ用法である。ワスラなどアラビア語特有の音韻現象に関係する記号は当然使わない（ただし、ファトハのタンウィーン記号は使う。これは先に述べた）。

3.4 句読点類

句読点類にはコンマ、コロン、セミコロン、ピリオド、括弧、引用符、疑問符、感嘆符などがあり、字形・用法ともにアラビア語と変わらないようである。「アラビア語」章の対応する節参照。

4 書体

ペルシア語の基本書体はナスタアリーク体 (نستعلیق nasta'liq) であり、写本から看板に到るまで広く使われてきた。書籍のタイトルなどでは、これの草書体であるシェキヤステ・ナスタアリーク体 (شکسته نستعلیق šekasteye nasta'liq) もよく使われる（下図）。



シェキヤステ・ナスタアリーク体による書名 نقش و نگاره های ایران naqš wa negârâyê 'îrân (イランの絵と文様)

しかし、ナスタアリーク体の活字は（おそらく）作られなかったので、印刷物ではナスフ体が使われた。デジタル組版が普及した現在のイランの印刷物でも本文書体はナスフ体である。

現在、ナスタアリーク体のデジタル書体がいくつも作られてきているので、今後普及すると見られるが、書籍や新聞の本文書体として使われていくかどうかは分からない。

5 正書法

5.1 アラビア語との綴りの異同

先に述べたように、アラビア語からの借用語は基本的には綴りを温存しているが、いくつかの違いがある。

省略された'alif

長母音 /ā/ があるにも関わらず、'alif が書かれていない単語、たとえば لکین (lākin; しかし) は、'alef を付加し، لاکین (lāken) とする⁽⁴⁶⁾。

語末のハムザ

アラビア語では الفاء ('alifbā; アルファベット) のように語末にハムザを持つ単語があり、ペルシア語でも同じく الفاء ('alefbā) と書かれてきたが、近年、このハムザを省略する傾向がある。現在では省略するほうが普通であるらしい。

アラビア語ではこのハムザは声門閉鎖音として発音されるが、現代ペルシア語では発音されない。

ター・マルブータ

アラビア語からの借用語の語末のター・マルブータはペルシア語では二つの可能性がある。

まず、/t/ の音になるもの。これは te (ت) で表す。例：حركه ḥarakah (動き, 母音) → حرکت harakat
もう一つは、子音が無く前の音節が /e/ で終わるもの。これは he havvez (ه) で表す。つまり識別点が取れるのである。例：فتحة (fatḥah; ファトハ) → فاتحه (fathe; ファトヘ)

この he havvez はいわゆる無音の he である。

綴りの温存

アラビア語の綴りをそのまま用いるものも、もちろん多い。アリフ・マクスーラは 'alef にせず、そのままである。したがって、語末の ye (ی) と区別が付かない。

5.2 語中の /i/

語頭の /i/ は ی と綴られるが、語中では二種類の表記方法がある。

一つは یی (ハムザつき ye の後に ye), もう一つは ی (ye を重ねる) である。たとえば آینه/آئینه ('ā'īne; 鏡)。近年では後者の綴りが普通になってきているらしい。

5.3 続け書きの分離

ペルシア語では複合語などの場合に、一つの単語でありながら途中で続け書きを切るものがある。これはアラビア語にはないルールである。

複合語

たとえば آب ('āb; 水) と شدن (šodan; なる) を組み合わせた آبشدن ('ābšodan; 溶ける) という動詞では、構成要素の間で続け書きが切れている（しかし，آبشدن という綴りもあるようだ）。Unicode の枠組みでは、このような組版を行うには、ゼロ幅非結合子 (ZWNJ: Zero Width Non-Joiner, U+200C) という不可視文字を間に挟むことになっている。

⁽⁴⁶⁾ 必ずそうするのかどうかは調査不足。

このような綴りを確かめようとして辞書を繙くと，*آب شدن* のように間にワードスペースを入れたものも見受けられる。ワードスペースの有無は辞書によってバラバラなのが現状である。辞書編纂者が2語とみなせばワードスペースを入れ，複合語と見なせばスペースは入れずに続け書きを切っけて掲げる。複合語か否かは簡単に決まることではなく，時代とともに変わりもする⁽⁴⁷⁾。イランにそれを決める国家機関があるわけでもない⁽⁴⁸⁾。同様の問題はわかち書きする他の言語にもあるが（例：*postcard* か *post card* か），英語などに比べて揺れの幅がずっと大きいようだ。

なお，組版・印刷が悪いためにワードスペースが空いているのかどうかすら判断つかない辞書・参考書も少なくない。

無音の *he* の後

無音の *he* で終わる単語に何らかの別の単語や接辞が付いて出来た単語の場合，その *he* のあとで続け書きを切断することがある。

たとえば，*بچه* /*bačče*/ 「子供」に複数形語尾 *ها* (*hâ*) を付けた「子供たち」は *بچه‌ها* /*baččehâ*/ と綴る。

また，曜日の名称のうち，日曜日から木曜日までは，数詞のあとに *شنبه* /*šambe*/ を付けた形をしていて，「日曜日」なら「一」を表す *یک* /*yek*/ と合成して *یکشنبه* /*yekšambe*/ という綴りである。しかし，火曜日「三」の意の *سه* /*se*/ が無音の *he* で終わっているので，ここで続け書きを切断し，*سه‌شنبه* /*sešambe*/ と綴る。このあたりも揺れがありそうだが，調べきれなかった。

動詞の接頭辞 *mī*

動詞の現在形や過去進行形に使われる *می* /*mī*/ という接頭辞がある。この接頭辞は続け書きを切断する書き方と続ける書き方がある。

たとえば，「私はする」という意味の /*mīkonam*/ であれば，*می‌کم* でも *میکنم* でもどちらでもよい。しかし，文書全体を通して統一すべきである。

切断した方が初学者に分かりやすいのは言うまでもない。たとえば，初学者は *میبینید* *mībinid*（あなたは見る）という綴りを見て戸迷うが，*می‌بینید* なら比較的に見やすい。

5.4 その他

前置詞 *be*

ペルシア語では，前置詞は一つの単語としてワードスペースを空けて書くのが基本である。しかし，「～へ」という意味の前置詞 *به* *be* については，後ろの単語と続けて書く場合がある。そのときは *he* (ه) が消える。

たとえば，「家へ」(*be xāne*) は *به خانه* とも *بخانه* とも綴られる。無論，初学者には分離してくれたほうがありがたい。

続けるなら文書全体で統一すべきなのか，また，他にもこのような前置詞があるのかどうか，については調べがつかなかった。

6 ダリー語について

ダリー語はペルシア語の一方言であり，アフガニスタンの共通語の役割を果たしている言語である。イランのペルシア語とは語彙や発音等に違いがあるが⁽⁴⁹⁾，ダリー語の文法書などを見るかぎり，正書法など文字に関する面では概ね同じと想像される（調査不足）。

(47) そもそも昔はワードスペースなどなかったものであり，このような正書法は近代になってできたものではないか。

(48) イラン言語・文学アカデミー (فرهنگستان زبان و آدب فارسی) という機関はあるが，正書法を定めているわけではない。

(49) 面白いことに，発音の違いは主として母音なので，綴りはほとんど同じである。

イランの現代ペルシア語に比べると、母音が二つ多い。これは長母音の /ē/ と /ō/ であり、それぞれ \bar{e} , \bar{o} で表す。したがって、長母音表記の \bar{e} には /ē, i/ の二つの音が、 \bar{o} には /ō, ū/ の二つの音が対応することになる。

また、現代ペルシア語が qāf (ق) と qeyn (ق) の発音を区別しなくなってしまったのに対し、ダリー語では現在も別の子音である。

IX ウルドゥー語

1 言語の概要

1.1 使用地域と人口

パキスタン（人口1億4872万人〔2004年6月〕）の国語はウルドゥー語だが、母語話者は約1000万人に過ぎない。一方、インドには約5000万人のウルドゥー語話者がいる。

パキスタンの人口の半分以上を占めるのはパンジャーブ人で、パンジャーブ語を母語とする⁽⁵⁰⁾。しかし、パンジャーブ州でも教育はウルドゥー語と英語で行われるので、ウルドゥー語を第2言語とする者の数は非常に多い。ウルドゥー語は多民族国家の共通語的役割を持っている。日本ではわりあい古くから研究されており、文献もいろいろある。

1.2 言語の系統

印欧語族インド語派に属し、ヒンディー語とは非常に近く、文法はほぼ共通している。しかし、10世紀以降イスラーム化が進み、多量のペルシア語彙（この中には多量のアラビア語彙が含まれる）を受容して現在のウルドゥー語が成立したようである。

1.3 音韻

サンスクリット語、ヒンディー語と同じく、子音に有気音／無気音の区別をもち、反り舌音、鼻母音がある⁽⁵¹⁾。

母音は、鼻母音化を捨象すれば /a, ā, i, ī, u, ū, e, ē, o, ō, ai, au/ の12個あり⁽⁵²⁾、サンスクリット語やヒンディー語の母音体系にほぼ対応する。

2 文字

ウルドゥー語を表すためのアラビア系文字をとくにウルドゥー文字と呼ぶ。

正書法は、ウルドゥー語促進協会が1920年代に標準化しようとしたものの[6]、現在もなお確立はしてはならず、揺れがある。

文字は、ペルシア語の32文字に反り舌音を表す3文字を追加した35字としている。したがって、ペルシア語と同じく同音異字を多数もつ。

2.1 反り舌音の文字

反り舌音を表す文字は ٽ /t̪/, ڍ /d̪/, ڙ /ʒ/ の三つであり、それぞれ، ٽ (の基字)، ٽ, ڙ に小書きの ٽ を載せた形である。Unicode の U+0615 ARABIC SMALL HIGH TAH はこの記号によく似ているが、まったく別の役割の記号なので注意されたい。

19世紀までは小書きの ٽ の代わりに四つ点を使った文字 (ٽ, ڍ, ڙ) もあり、並行して使われた。

⁽⁵⁰⁾ パンジャーブ語は、パキスタンではシャームキー文字というアラビア系文字で書かれ、インドではグルムキー文字というインド系文字で書かれる。シャームキー文字について取り上げる余裕は無かった。文献[6]参照。

⁽⁵¹⁾ このあたりの音声学用語は文献[16]を参照。

⁽⁵²⁾ 二重母音 /ai, au/ は、それぞれ /ae, ao/ とも。

2.2 有気音を表す文字

ヒンディー語を表記するデーヴァナーガリー文字では有気音と無気音には別の文字を使う。これに対し、ウルドゥー文字では、無気音の子音字の後に *do čašmī he* (二眼の *he*) と呼ばれる文字 (ح) を並べることで有気音を表す⁽⁵³⁾。*do čašmī he* は、本来はナスタアリーク体において、*h* の両接形・右接形に現れるうる異体字だが、ウルドゥー文字ではこれらに別の役割を持たせている。Unicode では、*do čašmī he* は U+06BE として符号化されている。

2.3 二つの *ye*

長母音 /i, ē/ および二重母音 /ai/ はいずれも *ye* (ی) で表記されるが、語末においては /ē, ai/ に限り、*ye* の異体字 ے を用いる。この文字を「大きいイエ (بڑی ے *baṛī ye*)」と呼ぶ。通常の *ye* は「小さいイエ (چھوٹی ے *chōṭī ye*)」である。

この二つの字体は、アラビア語・ペルシア語においては単なる書き方の違いであり、機能面での違いはない。たとえば、ペルシア語で人名の「アリー」は、通常は علی だが、手書きではしばしば美的観点から ے とも書かれる。それに対し、ウルドゥー語の正書法ではこの二つの字体は機能の違いなのである。

2.4 鼻母音を表す *nūn*

デーヴァナーガリー文字では、鼻母音化を表すためにチャンドラ・ビンドゥ (月の点) と呼ばれる記号を文字の上に付加するが、ウルドゥー語では *nūn ḡunna* ヌーン・グンナという文字 (ن) を挿入して示す。これは *nūn* (ن) から識別点を取り去ったものであり、Unicode では U+06BA として符号化されている。この文字も左接形・両接形においては識別点を付けるので、通常の *nūn* と区別がつかなくなる⁽⁵⁴⁾。したがって、正確な発音は単語ごとに個別に覚えるしかない。

2.5 異なる文字なのか異体字か

以上に挙げた (1) 通常の *he* と *do čašmī he*, (2) 小さい *ye* と大きい *ye*, (3) 通常の *nūn* と *nūn ḡunna*, の各ペアは、「ウルドゥー文字は全部で 35 文字ある」といったときに、別個には数えられていないということに注意すべきである。ウルドゥー語の文字表ではこれらはあくまで異体字ということなのだろう。しかし、機能的には別の文字とみなすほうが自然ではないだろうか

3 記号

ウルドゥー文字にもいろいろな記号があるが、調査不足。

3.1 ハムザ

ウルドゥー文字におけるハムザは、もはや声門閉鎖音という子音を表す記号ではなく、単に母音が連続することを示す。たとえば、「取る」という意味の動詞 *lenā* の完了分詞・男性複数形 *lie* は、/l/ に対応する ل と語末の /e/ に対応する ے の間に、ハムザつきの *ye* (ی) を立て、 ے と綴る [6]。なお、この単語には異綴が存在するが、それは正書法の節で取り上げる。

⁽⁵³⁾ *do, čašm* は、ペルシア語でそれぞれ「2」「目」の意味である。

⁽⁵⁴⁾ 文字コードの観点から言えばそうなのだが、伝統的な文字観(?)からは「*nūn ḡunna* は語末にしか現れない」とも表現できる。

3.2 アリフ・マクスーラの表記

ペルシア文字と同じく ye (ی) の右接形・独立形に識別点を付けないので、語末ではアリフ・マクスーラとイエの区別がつかなくなってしまう。

そのため、ウルドゥー文字では、アリフ・マクスーラの上にアラビア語と同じく「短剣のアリフ」を付ける。例：تعالی (ta'ālā; 至高なる〔神の形容〕)

3.3 母音記号

母音記号は概ねアラビア語・ペルシア語と同様のようなのだが、調べきれなかった。

3.4 句読点類

引用符はギユメではなく、いわゆるダブルクォーテーションマーク “ ” を用いる。英領インド時代の英語の影響だろうか？

終止符はピリオドでなく、ベースライン上に置く短い線-である。

文献[13]には、「記号」として、次の12の句読点類が挙げられている。

،	コンマ	سکتہ
؛	セミコロン	وقفہ
:	コロン	رابطہ
-	終止符	ختمہ
؟	疑問符	سوالیہ
!	感嘆符	فجائیہ ، ندائیہ
[] ()	括弧	قوسین
“ ”	引用符	واوین
...	リーダー	قطار
—	ダッシュ	خط
-	ハイフン	خط ربط
—	横線 (書名等の上)	لکیر

4 書体

13世紀以降、ナスフ体が使われてきたが、ペルシア語を公用語としたムガル朝時代にナスターアリーク体が一般化した[] [6]。書籍・新聞からトラックの装飾まで⁽⁵⁵⁾ナスターアリーク体が使われており、この書体へのこだわりはイランよりも強いように思われる。ペルシア語で、手書きはナスターアリーク体が基本書体でありながら、印刷物の本文はほとんどナスフ体であることと対照的である。ナスターアリーク体の活字を作るのは困難なため、近年まで新聞は手書き(石版印刷もしくは平版オフセット印刷)であったという⁽⁵⁶⁾。ナスターアリーク体のデジタル書体は、1981年のNūrī

(55) パキスタンのトラック野郎たちは、ナスターアリーク体を使ったド派手な装飾を愛車に施している。

(56) ナスフ体の活字が使われたこともある。また、文献[1]には、19世紀にタイプライター用活字として「ナスフ体とナスターアリーク体の中間的な独特の形ができた」とある。

Nasta'liq を嚆矢として、パキスタンで盛んに開発されてきている。それとともに新聞や書籍でコンピュータ組版が急速に普及した。

筆者の見たところ、パキスタンのナスターリーク体とイランのそれは美の基準が違うようである。同じ漢字の楷書体でも日中韓で書法が微妙に違うようなものであろうか。

5 正書法

筆者はウルドゥー語を全く解さないため、正書法についてもよく分からないが、文献 [6] には綴りの揺れの問題として次の二点が挙げられている。これらは、ペルシア語の章の「語中の /i/」「続け書きの分離」の章で取り上げた問題と似ている。

いずれも、原文の図版ではナスターリーク体で例が示されているが、本稿では組版の便宜上、ナスフ体で示す。

このような綴りの揺れがいつまでも統一されないのは、インド、パキスタンいずれにおいても綴りを決定し、強制する機関が存在しないことが一因であるらしい。

5.1 ハムザの有無

「記号」の節で、lie という単語の綴りを ل としたが、これの異綴として、ل があるという。これはどういうことなのか。実際の発音において、/i/ と /e/ の間に“わたり音”として /y/ が生じることがあり⁽⁵⁷⁾、これが綴りに反映したものであるという。

この二つの綴りはどちらもよく使われている。

5.2 複合語の分かち書き

「店」という意味の دكان dukān に、「持つ」という意味の دار dār を付けてできる「店主」dukān-dār は、دكاندار のように一単語でも、دار دكان のように分かち書きされた二語でも綴られる。

(57) このような現象は多くの言語に見られる。日本語でも「試合」がごく自然に「シヤイ」のような発音になることがある。

X 現代ウイグル語

1 言語の概要

現代ウイグル語で「現代ウイグル語」を *ھازىرقى زامان ئۇيغۇر تىلى* *hazirqi zaman uyghur tili* という⁽⁵⁸⁾。言語学的に隔たりのある古ウイグル語と区別してこう呼ぶ。

本稿の内容は大部分が文献 [14] に基づいているが、この本はウルムチ方言について書かれているので、以下の説明もそれが前提である。

1.1 使用地域と人口

話者は推定 750 万人以上で、その大部分が中華人民共和国 新疆（しんきょう）ウイグル自治区（いわゆる東トルキスタン）に住む。

1.2 言語の系統

アルタイ語族テュルク諸語の一つ。ウズベク語と非常に近く、トルコ共和国で話されているトルコ語とも近い。ただし、アルタイ語族という概念が成立しうるかどうか、まだはっきりしていないらしい。

テュルク諸語は膠着語であり、文法構造は日本語や朝鮮語とよく似ている。

語彙はアラビア語・ペルシア語やロシア語、中国語からの借用語を多く含む。

1.3 文字の変遷

現代ウイグル語を表記するために現在採用されているアラビア系文字を「新ウイグル文字」とか「ウイグル新字」と呼ぶ。

古代ウイグル語ではウイグル文字⁽⁵⁹⁾が使われていたが、テュルク（ここではテュルク諸語を話す人々の意）のイスラーム受容が進むと、ウイグル語を含むテュルク諸語はアラビア系文字で書かれるようになった。このときの正書法は、アラビア語・ペルシア語の語彙を綴りを温存して表記するやり方であり、テュルク語の発音を表す上では多数の同音異字を抱えている。この点、ペルシア文字やウルドゥー文字と同様である。

しかし、現代ウイグル語を表記する文字はこの 100 年の間に目まぐるしく変化する。

まず、20 世紀前半に、綴りを改革することによって現代ウイグル語に適したアラビア系文字表記を作ろうという試みがなされた。具体的には、同音異字の廃止や母音字の導入である。これが現在の新ウイグル文字につながっている。その一方で、ソ連の影響によりラテン文字による表記やキリル文字による表記も行われるようになった。

1950 年に上述のアラビア系文字の新字母表が確定し、以後新疆全域で現在のような新ウイグル文字が公式に用いられるようになった [6]。

ところが、1956 年にはソ連の影響でキリル文字が公式採用され、1958 年にはピンイン⁽⁶⁰⁾に基づいたラテン文字表記が導入された。このラテン文字表記は 1974 年から本格的に使われ始めた [22]。

(58) この単語の順に「今の」「時代」「ウイグル」「ことば」といったところである。

(59) ソグド文字がウイグル語に適用されて次第に変化したものがウイグル文字で、縦書きし、行は左から右へ配す。これをモンゴル語に適用したものがモンゴル文字であり、それをさらに改変して満州語に適用したものが満州文字である。

(60) 現代中国の共通語である「普通話」を表記するためのラテン文字。

しかし、1982年に記述のアラビア系文字表記を元にした現在の新ウイグル文字が採用され、現在に到っている。正書法は何度も改訂され、文字コードも含めて現在も議論が続いている。

なお、現在テュルク諸語の中で、カザフスタン共和国のカザフ語、キルギス共和国のキルギス語、ウズベキスタン共和国のウズベク語などはアラビア系文字を用いず、キリル文字ないしラテン文字を用いている（これらの言語は他の地域ではアラビア系文字を用いる場合がある）。

1.4 音韻

母音は/a, ä, e, i, o, ö, u, ü/の八つ。このうち、/e/は/a, ä/の変異音であり、母音音素は全部で七つである。ウムラウトの有無は前舌／後舌の違いに対応している。

2 文字

新ウイグル文字の最大の特徴は、母音字の導入にある。母音記号を使ったり子音字を流用するのではなく、八つの母音のすべてに独自の文字を割り当てている。

新ウイグル文字は、現代に作られただけあって、過去のアラビア系文字の伝統を破り、あくまで現代ウイグル語の表記に必要な文字だけを採用している。

元祖アラビア文字→ペルシア文字→ウルドゥー文字という発展の系譜においては、必要な文字を付加しながら、不要な文字（その言語の音を表すのに不可欠ではない文字）も削除はしなかったが、新ウイグル文字はこれと好対照をなす⁽⁶¹⁾。

文字は全部で32あり、うち八つが母音字である。

次に文字表を掲げるが、字形に疑問のある文字もある。これについては後述する。

文字	名称	転写	右接形	両接形	左接形
ا	ئا'a	a	ا		
ە	ئەä	ä	ە		
ب	بېbe	b	ب	ب	ب
پ	پېpe	p	پ	پ	پ
ت	تېte	t	ت	ت	ت
ج	جېje	j	ج	ج	ج
چ	چېche	ch	چ	چ	چ
خ	خېx	x	خ	خ	خ
د	دەde	d	د		
ر	رەre	r	ر		
ز	زەze	z	ز		
ژ	ژەzhe	zh	ژ		
س	سېse	s	س	س	س
ش	شېshe	sh	ش	ش	ش

(61)ただし、古いウイグル語のアラビア系文字表記では、伝統的な、文字を追加する方式が採られていた。新ウイグル文字はまったく新しいアラビア系文字表記である。

غ	غې ghe	gh	غ	غ	غ
ف	فې fe	f	ف	ف	ف
ق	قې qe	q	ق	ق	ق
ك	كې ke	k	ك	ك	ك
گ	گې ge	g	گ	گ	گ
ڭ	ڭې nge	ng	ڭ	ڭ	ڭ
ل	لې le	l	ل	ل	ل
م	مې me	m	م	م	م
ن	نې ne	n	ن	ن	ن
ه	هې he	h	ه	ه	ه
و	و' o	o	و		
ۇ	ۇ' u	u	ۇ		
ۈ	ۈ' ö	ö	ۈ		
ۊ	ۊ' ü	ü	ۊ		
ۋ	ۋ' we	w	ۋ		
ې	ې' e	e	ې	ې	ې
ى	ى' i	i	ى	ى	ى
ي	ي' ye	y	ي	ي	ي

2.1 母音字

母音 /a, ä, e, i, o, ö, u, ü/ を表す母音字は、それぞれ، ە، ې، ى، ۈ، ۇ، ۊ， ۈ である。これらはすべて単独の文字である。ۈ はワーウの上にダンマがついたものと同じ形をしているが、新ウイグル文字においてはそのような「子音字+母音記号」という解釈をしてはいけない。あくまでこれで一つの文字なのである。

2.2 ä と h

母音字の中で気をつけなければならないのが、/ä/ を表す ە である。独立形はアラビア語の hā' と同じ形をしているが、左接しない。右接形は ھ である。

一方、子音字である ھ は、左接形 ھ، 両接形 ھ، 右接形 ھ となる。この文字が Unicode でどこに割り当てられているのか、いまいちよく分からない。各字形はウルドゥー語の do čašmī he (二眼のへー) と同じ型のように見えるが、こちらは U+06BE: ARABIC LETTER HEH DOACHASHMEE と定義されていて、ウルドゥー語で使うこと、有気音の表記に使うことは書かれているが、ウイグル語の名は挙げられていない。

すると、アラビア文字の普通の hā' (U+0647) だろうか。しかし、こちらは両接形が ھ となったり、独立形が ھ となる可能性があり、とくにウイグル語専用のシステムでない場合に望ましい字形で組まれないおそれがある。

2.3 識別点のない文字 i

/i/を表す文字は ى であるが、これは左接形・両接形においても識別点が無いので、ペルシア文字・ウルドゥー文字における *ye* とは違う。このため ئىسسىق (/issiq/, 「暑い・熱い」) のような文字の切れ目が分かりにくい綴りが存在する。

Unicode においては、どうやらこれは U+0649: ARABIC LETTER ALEF MAKSURA を用いるようである。というのは、他にそれらしい文字が見当たらないうえ、この文字の説明として“represents YEH-shaped letter with no dots in any positional form”となっているから。もちろん、位置による字形の変化が共通しているから共用できるというだけのことであって、現代ウイグル語にとっては決してアリフ・マクスーラではない。

3 記号

3.1 ハムザ

ハムザは ا at (馬) や سائت sa'ät (時計), سائت san'ät (芸術) のように、 ا または ئ の形でしか現れない。これらは、 ى の左接形・両接形である。

これらが声門閉鎖音を表わしているのかどうか、文献で確認できなかった。音声学的には声門閉鎖音が聞かれる場合とそうでない場合があるのではないかと思う。音素としては声門閉鎖音は認められていないようである。

しかしながら、正書法上は母音字が語頭に立つことはなく、「馬」の例で示したように、母音字の前に必ず ى を置く。また、「時計」の例で示したように、母音字が直接隣り合うこともなく、必ず間に ى を挟む。さらに、「芸術」の例で示したように、子音に母音が続く場合でも、異なる音節に分かれているときは必ず間に ى を挟む⁽⁶²⁾。

ややこしく感じるかもしれないが、要は母音で始まる音節には頭に ى を付けよということである。ハングルにおけるゼロ子音文字の問題とよく似ている。

3.2 ハイフン

新ウイグル文字にはハイフネーションがある。つまりラテン文字の言語と同様、行末にかかった単語を分割し、前綴の末尾に短い水平線を置く。分割可能な箇所はやはり音節の境界である。

4 書体

基本書体はテズズ体 (tüz xät, 文字どおりには「直立体」の意)。ナスフ体*を元に行っていると思われるが、ナスフ体のような傾きは見られず、画線が水平・垂直に伸びているのが特徴。

5 略語

頭文字をとった略語 (英語の acronym に相当) がある。

「新疆ウイグル自治区」は شىنجاڭ ئۇيغۇر ئاپتونوم رايونى شىنجاڭ Shinjang Uyghur Aptonom Rayoni だが、これを略して شى ئۆنار と書く。頭文字を取るといっても、声門閉鎖音に当たる文字 ئ で始まっている語はその次の母音字も含めて採っているということに注意されたい。

(62) この例は日本語で「三愛」のローマ字表記を「san'ai」とするのに似ている。

XI Unicode

Unicode では、アラビア系文字は以下の三つのエリアに収録されている。

- Arabic (0600–06FF)
- Arabic Presentation Forms-A (FB50–FDFF)
- Arabic Presentation Forms-B (FE70–FEFF)

このうち、文字として用いるのは“Arabic”に含まれるものだけである。他の“Presentation Forms”（表示形）というの、左接形とか合字といった、組版上必要な字形を文字のように扱うためにある（ただし、あらゆる合字が収録されているわけではない。中途半端である）。

コードチャートが <http://www.unicode.org/charts/> で公開されているので、参照されたい。

Unicode における文字名称は、ヤケクソ (?) で付けられたものが多数あるので、これを文字学的に正当な名称であると考えてはならない⁽⁶³⁾。たとえば、U+06CE (ﻱ) は ARABIC LETTER YEH WITH SMALL V となっているが、ﻱ の上に載っているのはもちろんラテン文字の v ではなく、V 字状の記号という程度の意味である。また、U+06C8 (ﻯ) は ARABIC LETTER YU となっているが、母音 /u/ を YU と表記するのは違和感がある。しかし、それも無理からぬことで、もともとその文字を使っている人々（民族）が名前を与えていない文字や記号にも名称を作らねばならず、しかもラテン文字 26 字で表現できるように、となるとこうならざるを得ないのかもしれない。

いずれにせよ、Unicode 名は便宜的なものと考えるべきである。

1 言語

Unicode のアラビア系文字がカバーしている言語は、正則アラビア語（フスハー）のほかに、以下のものがある（コードチャートに出てくる文字の説明に書かれているもののみを挙げる）。

- Adighe（アディゲ語；コーカサス）
- Baluchi（バルーチ語；イラン、パキスタン）
- Berber（ベルベル語；北アフリカ）
- Dargwa（ダルグワ語？；コーカサス）
- Ingush（イングーシ語；コーカサス）
- Kashmiri（カシュミール語；インド・パキスタン）
- Kazakh（カザフ語）
- Kirghiz（キルギス語）
- Kurdish（クルド語）
- Lahnda（ラーンダー語；パキスタン）
- Maghrib Arabic（マグリブ⁽⁶⁴⁾のアラビア語）
- Moroccan Arabic（モロッコのアラビア語⁽⁶⁵⁾）
- old Hausa
- old Malay
- old Urdu

⁽⁶³⁾ 権威ある団体 (?) が制定したということで、Unicode 名が無批判に広まってしまふことを筆者は恐れる。

⁽⁶⁴⁾ モロッコ・アルジェリア・チュニジア、ときにリビアを含むアフリカの北西部地域。(al-)Maghrib المغربは「去る、(日が) 没する」という意味の غرب gharaba の派生語であり、「日の没する所」転じて「西」という意味の مغرب maghrib に定冠詞が付いたもの。

⁽⁶⁵⁾ こういう区分があるということは Maghrib Arabic が指すものはモロッコ以外のマグリブということだろう。

2.2 hā' の仲間

アラビア語の hā' (ه) に、形の上で似た文字と機能的に関連のある文字を挙げる。こちらも網羅はしていない。

U+0629 ARABIC LETTER TEH MARBUTA ه ّ

Ar の tā' marbūtah。語末でしかあり得ないので、左接形・両接形は無い。

U+0647 ARABIC LETTER HEH ه هـ

Ar, Fa の hā'。単独で現れるときには ه の字形を取ることが多い。

U+06BE ARABIC LETTER HEH DOACHASHMEE هـ شمش

Ur において有気音を示す do čašmī he。この字体は U+0647 の異体である。Uy ではこれを子音字 he に用いるのではないかと思う。

U+06C1 ARABIC LETTER HEH GOAL ه ٲ

Ur における he。U+0647 の異体字であり、本来はナスタアリーク体の字体である。Unicode のコードチャートでは両接形に ٲ 状の付加物は無いが、Arabic Typsetting フォントにはある。

U+06C2 ARABIC LETTER HEH GOAL WITH HAMZA ABOVE ه ٲٲ

Ur で、U+06C1 にハムザが載った場合。語末でしかあり得ないので、左接形・両接形は無い。

U+06C3 ARABIC LETTER TEH MARBUTA GOAL ه ّٲ

Ur における語末のター・マルブータ。左接形・両接形は無い。

U+06D5 ARABIC LETTER AE ه ٲٲ

Uy における母音字 ä。Ar, Fa の子音字 ه に似ているが、左接しない。したがって、語中でもそこで続け書きが切れる。

2.3 kāf の仲間

アラビア語の kāf (ك) の仲間を挙げる。こちらも網羅はしていない。

U+0643 ARABIC LETTER KAF ك كـ ك

Ar の kāf。

U+06A9 ARABIC LETTER KEHEH ك كـ كـ كـ

本質的には U+0643 と同じ文字で、独立形・右接形が異体である。おそらくはナスタアリーク体に由来する字体。Fa ではこちらを用いる。

U+06AA ARABIC LETTER SWASH KAF كـ كـ

Ar における ك ʾ の異体字。語頭にしか使われられないので、独立形・左接形のみを挙げた。純粋に視覚効果的に選択される字体だと思うが、何か使い分けの規則があるかもしれない。

U+06AD ARABIC LETTER NG كـ كـ كـ كـ

Uy の nge。この音は語頭には現れない。

U+06AF ARABIC LETTER GAF كـ كـ كـ كـ

Fa, Ur, Uy などの gāf。

2.4 nūn の仲間

nūn の仲間の文字を挙げる。

U+0646 ARABIC LETTER NOON ن نـ نـ نـ

Ar の nūn。

U+06BA ARABIC LETTER NOON GHUNNA نـ نـ نـ نـ

Ur のヌーン・グンナ。語末でのみ通常のヌーンと区別される。

2.5 fā' の仲間

アラビア語の fā' の異体字を挙げる。

U+0641 ARABIC LETTER FEH ف ف ف ف

Ar (フスハー) の fā'。

U+06A2 ARABIC LETTER FEH WITH DOT MOVED BELOW ف ف ف ف

マグリブにおける U+0641 の異体字。

2.6 qāf の仲間

アラビア語の qāf の異体字を挙げる。

U+0642 ARABIC LETTER QAF ق ق ق ق

Ar (フスハー) の qāf。

U+06A7 ARABIC LETTER QAF WITH DOT ABOVE ق ق ق ق

マグリブにおける U+0642 の異体字。

コードチャートには U+06A8 ARABIC LETTER QAF WITH THREE DOTS ABOVE (ق) という文字があり、チュニジアのアラビア語で用いるという。これが qāf の異体字なのかどうかは調査不足で不明。

2.7 wāw の仲間

U+0624 ARABIC LETTER WAW WITH HAMZA ABOVE ؤ ؤ

U+0648 にハムザが付いたもの。Ar, Fa, Ur など で用いる。声門閉鎖音を表す。

U+0648 ARABIC LETTER WAW و و

Ar, Fa, Ur などの wāw。Uy ではこれを母音字 o に用いるようだ。

U+06C6 ARABIC LETTER OE ؤ ؤ

Uy の母音字 ö。

U+06C7 ARABIC LETTER U ؤ ؤ

Uy の母音字 u。形の上では wāw に母音記号ダンマがついたものと同じだが、使い分けなければならない。コードチャートではなぜか注釈に「Kirghiz」とのみあり、ウイグル語の名が挙がっていない。

U+06C8 ARABIC LETTER YU ؤ ؤ

Uy の母音字 ü。

U+06CB ARABIC LETTER VE ؤ ؤ

Uy の子音字 ve。

3 文字の採用方針

前節を見ても分かるように、Unicode の文字の採用方針は一貫性に欠けている。これは既存コードとの互換性に配慮した結果だろうか。

たとえば本質的には同じ文字である U+064A (ي) と U+06CC (ي) に別のコードポイントを割り当てている。アラビア語とペルシア語で書き方がやや違っているだけの文字を区別しているわけで、Unicode が前提としている character/glyph モデル (囲み記事参照) に従っていないことになる。二種類の kāf (ك) や二系列の数字 (U+0660-9 と U+06F0-9) などと同様。

※character/glyph モデルとは

筆者はこれを正確に説明する能力に欠けるが、簡潔に言えば文字概念として character と glyph を分離する考え方。たとえばラテン文字の A の小文字には、大雑把に言って a と α の二つの字体があるが⁽⁶⁸⁾、これは glyph の違いである。こういった差異を捨象した文字概念が character である。Unicode をはじめ、多くの文字コードは、glyph ではなく character の集合にコード番号を振ったものである。もちろん、character と glyph の概念が常にきれいに分離できるとは限らない。

また、語末において（つまり独立形・右接形の場合に）この U+06CC と全く区別の付かない U+0649（アリフ・マクスーラ）にも別のコードポイントを割り当てている。character/glyph モデルから言えばこれらを区別することに意味はある。しかし、運用はやや難しい。ペルシア語ではどちらも用いるが、間違ってもタイプしてもディスプレイやゲラではまったく区別が付かず、検索やソーティングの障害になるだろう⁽⁶⁹⁾。プログラムで自動判定させようとするなら、ペルシア語の辞書が必要になる。

もっと深刻なのが現代ウイグル語の場合である。母音 i を表す文字が現代ウイグル語専用には用意されていないようで、四つの字形のパターンが唯一一致している文字は U+0649（本来はアリフ・マクスーラ）しかない。もしこれを用いるのだとすれば、U+0649 は本質的に異なる二つの文字を兼ねていることになり、character/glyph モデルを無視していることになる。

⁽⁶⁸⁾ それぞれ「二階建の a」「一階建の a」などと呼ぶ。ローマン体では前者、イタリック体では後者が普通。サンセリフではどちらもよくある。インクナブラ（1500年までに刊行された活版印刷物）には、両者を混ぜて使っているものもある。

⁽⁶⁹⁾ もちろん、両者を同一視するアルゴリズムを採用すれば事足りるが、ウェブ検索エンジンや DB ソフト、ワープロ、表計算ソフトなどはそうなっているだろうか？

XII 参考文献

- [1]『世界文字辞典』, 言語学大辞典別巻, 河野 六郎, 千野 栄一, 西田 龍雄 編著, 三省堂 (2001), 本体価格 48,000 円
古今東西の諸文字について言語学の立場からまとめた辞典。
「アラビア文字」, 「ウルドゥー文字」, 「ジャウィ文字」, 「シンディー文字」, 「スラヴ語の文字」, 「東南アジア島嶼部の文字」, 「トルコ語の文字」, 「ナバテア文字」, 「西アジアの文字」, 「パシュトー文字」, 「東アジアの諸文字」, 「フィリピンの文字」, 「ペルシア文字」の項参照。
- [2]『文字の世界史』, ルイ=ジャン・カルヴェ 著, 矢島 文夫 監訳, 河出書房新社 (1998), 本体価格 2,800 円
世界の主要な文字の構造と起源・発展を解説。わりとよくまとまっている。
- [3]『世界の文字』講座言語 5, 西田 龍雄 編, 大修館書店 (1981), 本体価格 3,300 円
- [4] *Arabic Typography — a comprehensive sourcebook*, Huda Smitshuijzen AbiFarès 著, Saqi Books (2001)
題名にタイポグラフィーとあるが, 内容の大部分が文字そのものと書体についてであり, 組版に関する記述は少ない。アラビア系文字印刷の歴史に一章を割いているほか, 書体や書体デザイナーについて具体的に列挙して記述されている。
- [5]『中国におけるアラビア文字文化の諸相』, 町田 和彦, 黒岩 高, 菅原 純 共編, 東京外国語大学 アジア・アフリカ言語文化研究所 (2003)
GICAS 「『小児錦』文字資料コーパス構築へ向けた資料収集とデジタル化」プロジェクト主催のワークショップの報告。次の文献も同じ。
- [6]『周縁アラビア文字の世界—規範と拡張—』, 町田 和彦, 黒岩 高, 菅原 純 共編, 東京外国語大学 アジア・アフリカ言語文化研究所 (2004)
本稿では本書のほか, ワークショップで配布された資料も参照した。
- [7]『アラビア文字を書いてみよう読んでみよう』, 本田 孝一, 師岡 カリマ・エルサムニー 著, 白水社 (1999), 本体価格 1,800 円
アラビア文字の最も優れた入門書。ナスヒー体とルクア体の書き方, 読み方が分かる。
- [8]『イスラム書道芸術大鑑』, イスラム歴史・芸術・文化研究センター編・監修, 本田 孝一 訳・解説, 平凡社 (1996), 本体価格 56,311 円
トルコでまとめられたアラビア書道作品の豪華本。書道史の解説もある。訳者は日本のアラビア書道の第一人者。
- [9]『図説アラビア文字事典』, ガブリエル・マンデル・ハーン 著, 緑 慎也 訳, 矢島 文夫 監修, 創元社 (2004), 本体価格 2,800 円
原著は *Arabic Script: Styles, Variants, and Calligraphic Adaptations*, Gabriel Mandel Khan, Abbeville Press (2001)。
- [10]『アラビア語・ペルシア語・ウルドゥー語対照文法』, 黒柳 恒男 著, 大学書林 (2002), 本体価格 8,000 円
- [11]「現代ペルシア語の音とカナ表記」, 上岡 弘二, 吉枝 聡子, 『アジア・アフリカ言語文化研究』, vol. 60 (2000), pp. 169–235
現代ペルシア語を仮名でどう表記するかというテーマで書かれたものだが, 結果的に現代ペルシア語の発音についての極めて詳細な分析報告となっている。旧説を覆す記述も多く, 必読。
- [12]『基礎ウルドゥー語』, 鈴木 斌 著, 大学書林 (1986), 本体価格 3,800 円

- [13]『ウルドゥー語文法の要点』, 鈴木 斌 著, 大学書林 (1996), 本体価格 5,040 円
- [14]『現代ウイグル語四週間』, 竹内 和夫 著, 大学書林 (1991), 本体価格 7,500 円
- [15]『パシュトー語文法入門』, 縄田 鉄男 著, 大学書林 (1985)
- [16]『日本語音声学入門』, 斎藤 純男 著, 三省堂 (1997), 本体価格 2,000 円
題名は「日本語」と付いているが, 内容はほとんど「音声学入門」である。アラビア語の咽頭化音も含め, 音声が網羅的に解説されている。音韻論 (音素論) についての簡単な解説もある。
- [17] *A Dictionary of Modern Written Arabic* (Arabic-English), Hans Wehr, ed. by J. Milton Cowan, 4th ed., Spoken Language Services, Inc., Ithaca, New York (1994).
用語などのアラビア語の確認は主にこの亜英辞書によった。
- [18] Unicode のコードチャート
<http://www.unicode.org/charts/>
これは必ず参照すべきである。
- [19] 「Unicode と XSL によるアラビア語の組版」(2004.3.1)
第 10 回多言語組版研究会の資料。文字と記号の合成や書字方向制御といった点について参考になる。
- [20] 展覧会「点と線 アラビア文字の旅」サイト
<http://www.gicas.jp/a-moji/>
2004 年 11 月 24 日～12 月 22 日に東京外国語大学 アジア・アフリカ言語文化研究所で行われた展覧会のサイト。
- [21] The Urdu alphabet
<http://www.unics.uni-hannover.de/nhtcapri/urdu-alphabet.html>
- [22] 現代ウイグル語字母の変遷史
<http://www3.aa.tufs.ac.jp/~sugawara/uigitil/script/script.htm>
- [23] 新疆研究室
<http://www.uighur.jp/>
新ウイグル文字について「現代ウイグル語とコンピュータ」などが必読。

付録 用語集

●文字名称のみはラテン文字表記を基本とする（仮名では kāf と qāf などの区別がつかぬ為）

●ター・マルブータ* (i) で終わる単語のラテン文字表記には揺れがあるが、これを(h)のように括弧つきで示した。

●言語名略号 Ar: アラビア語, En: 英語, Fa: ペルシア語, Ur: ウルドゥー語, Fr: フランス語

欧 文

'abjad → アブジャド

acrophony → アクロフォニー

'alef → 'alif

'alif الف アラビア文字*「アリフ」(ا)。左接しない。本来は声門閉鎖音を表す子音字だが、アラビア語ではハムザ*の台として使われる（この場合'alif 自体は音価を持たないとされる）など、原則からやや外れている。また、長母音 /ā/ の表記にも使われる。Fa: アレフ ('alef)。

U+0627 ARABIC LETTER ALEF (ا)

'alifbā' الباء 「アルファベット」に相当する概念。アラビア語ではアリフバーという。

Fa: アレフバー ('alefbā)

→アルファベット

'alif maqṣūra(h) アリフ・マクスーラ (ى)。「縮められた'alif」の意。語末にのみ現れる。文字としては識別点*の無い yā* (ي) と同じだが、長母音 /ā/ を表し、発音上は'alif*であるかのように扱う。ナバテア語*の綴りを反映したものとされる。

アラビア文字の総数 28 字の中には数えられていない。これは,'alif の異体字として扱われているからだろう。

Fa: アレフ・マクスーレ ('alef maqṣūre)。

U+0649 ARABIC LETTER ALEF MAKSURA (ى)

aljamía → アルハミーヤ

the Arabic numerals → インド・アラビア数字

Arabic script → アラビア文字

Aramaic script → アラム文字

'ayn عين アラビア文字「アイン」(ع)。アラビア語では有声咽頭摩擦音を表す (ḥā* ح の有声音)。ラテン文字に転写する場合, /j/ を用いるが, 左シングルクオート /j/ で代用することも多い。

U+0639 ARABIC LETTER AIN (ع)

U+201B SINGLE HIGH-REVERSED-9 QUOTATION MARK (ﻻ)

bā' باء アラビア文字「バー」(ب)。文字名称は beh とも。

U+0628 ARABIC LETTER BEH

baṛī yē → 大きいイェー

basic shape → 基字

basmaḷa → バスマラ

caliph, calif → カリフ

ḥōṭī yē → 小さいイェー

composition 組版。

consonant 子音。

consonant enhance mark シャッダ*とスクーン*の総称。なぜこのように呼ぶのかはよく分からない。

qād قاد アラビア文字「ダード」(ض)。アラビア語では /d/ の咽頭化音 /d/ を表す。

U+0636 ARABIC LETTER DAD

dāl دال アラビア文字「ダール」(د)。音価は /d/。左接しない。

U+062F ARABIC LETTER DAL

ḍamma → ダンマ

decimal comma 小数点として用いるコンマ。アラビア語の数表記では、小数点にはピリオドではなくコンマを用い、3 桁区切りはスペースを用いる（いわゆるヨーロッパ式）。このコンマはファースィラ*と違ってラテン文字のコンマと同じ向きを持つ。

U+066B ARABIC DECIMAL SEPARATOR (٫)

dhāl ذال アラビア文字「ザール」(ذ)。アラビア語では有声歯舌摩擦音 /ð/ を表す。左接しない。文字名称は thal とも。

U+0630 ARABIC LETTER THAL

diacritic dots → 識別点

Dīwānī → ディーワーン体

do čašmī he → 二眼の he

dotless nūn → 点無しヌーン

fā' فاء アラビア文字「ファー」(ف)。

U+0641 ARABIC LETTER FEH

→ マグリブ体

Fārsī ペルシア語*, ペルシア文字*。

fatha → ファトハ

figure 数字。数の表記体系の意味ではなく、個々の数字を指す。→ numerals

final form → 右接形

free standing form → 独立形

fuṣḥā' → フスハー

ghayn غين アラビア文字「ガイン」(غ)。アラビア語では有声軟口蓋摩擦音 /ɣ/ を表す (khā* خ の有声音)。文字名称は ghain とも。

U+063A ARABIC LETTER GHAIN

Ghubār, Ghubārī → グバル体

hā' هاء アラビア文字「ハー」(ه)。仮名表記では ḥā' や khā' と区別が付かないことに注意。文字名称は heh とも。→ tā' marbūṭa

U+0647 ARABIC LETTER HEH

ḥā' حاء アラビア文字「ハー」(ح)。アラビア語では無声咽頭摩擦音を表す。仮名表記では hā' や khā' と区別が付かないことに注意。文字名称は hah とも。→ 'ayn

U+062D ARABIC LETTER HAH

hamza(h) → ハムザ

ḥaraka(h) → ハラカ

ḥarakāt حركات harakah の複数形。

ḥarf حرف (個々の) 文字。英語の letter に相当。複数形は ḥurūf (حروف)。「活字」という意味もある。「アラビア文字」などというときの集合的な「文字」は khatt*。

Fa : ハルフ (harf) [複] ホルーフ (horūf)

Indic numerals →インド・アラビア数字

initial form →左接形

jazm جزم →ジャズム

Jawi →ジャウイ

jīm جيم アラビア文字「ジーム」(ج). 文字名称は jeem と
も。

U+062C ARABIC LETTER JEEM

kāf كاف アラビア文字「カーフ」(ك). 仮名表記では qāf と
区別が付かないことに注意。

U+0643 ARABIC LETTER KAF

kaliph, kalif →カリフ

kashīda(h) →カシーダ

kasra(h) →カスラ

khā' خاء アラビア文字「ハー」(خ). アラビア語では無声
軟口蓋摩擦音を表す。仮名表記では hā' や ḥā' と区別が付
かないことに注意。文字名称は khah とも。

U+062E ARABIC LETTER KHAH

khatt خط (1) 文字 (英語の script* に相当), (2) (この後に
書体名を続けて)「~体」, (3) 書道。

原義は「線」。一つ一つの文字を指す言葉は ḥarf*。

knot ع の両接形 ㄷ などに見られる結び目状の部分。

Kūfi →クーファ体

Kufic →クーファ体

lām لام アラビア文字「ラーム」(ل).

U+0644 ARABIC LETTER LAM

ligature →合字

long vowel 長母音。

madda(h) →マッダ

Maghrib, Maghreb →マグリブ

Maghribī →マグリブ体

medial form →両接形

mīm ميم アラビア文字「ミーム」(م). 文字名称は meem と
も。

U+0645 ARABIC LETTER MEEM

Muhaqqaq →ムハッカク体

Nabataean script →ナバテア文字

Naskh, Naskhī →ナスフ体

Nasta'liq →ナスタアリーク体

numerals 数字。個々の数字ではなく数の表記体系を指
す。例: Arabic numerals (アラビア数字), Roman
numerals (ローマ数字)。

→figure, インド・アラビア数字

nūn نون アラビア文字「ヌーン」(ن). 文字名称は noon と
も。

U+0646 ARABIC LETTER NOON

nuqṭa(h) →ヌクタ

the Ottoman Empire オスマン帝国 (1300 頃–1922)。首都
イスタンブルはこの時代のアラビア書道の中心地であっ
た。

→オスマン語

Phoenician script →フェニキア文字

phoneme →音素

punctuation marks 句読点類。ラテン文字なら、ピリオ
ド、カンマ、コロン、セミコロン、引用符、ダッシュな
どの総称。

qāf قاف アラビア文字「カーフ」(ق). アラビア語では無
声軟口蓋破裂音を表す。仮名表記では kāf と区別が付か
ないことに注意。

U+0642 ARABIC LETTER QAF (ق)

→マグリブ体

qalam →カラム

rā' راء アラビア文字「ラー」(ر). 左接しない。文字名称
は reh とも。

U+0631 ARABIC LETTER REH

Rayhān →ライハーン体

Riqā' →リカーウ体

root →語根

Ruq'a(h) →ルクア体

ṣād صاد アラビア文字「サード」(ص). アラビア語では /s/
の咽頭化音 /s/ を表す。文字名称は単に sad とも。

U+0635 ARABIC LETTER SAD

script 「アラビア文字」「ラテン文字」というときの「文
字」。個々の文字は letter。区別を明確にするため、日本
語でも「スクリプト」の訳語を当てることがある。

また、書体名のあとに付けて「~体」という意味にも使
われる。例: Naskh script (ナスフ体)

Ar : خط khatt

Semitic script →セム系文字

shadda(h) →シャッダ

shakl →母音記号

Shikaste(h), Shekaste(h) →シェキャステ体

shīn شين アラビア文字「シーン」(ش). 文字名称は sheen
とも。

U+0634 ARABIC LETTER SHEEN

short 'alif →短剣アリフ

short vowel 短母音。

sīn سين アラビア文字「スイーン」(س). 文字名称は seen
とも。

U+0633 ARABIC LETTER SEEN

Sindhi スインド語*, スインド文字*

standing form →独立形

sukūn →スクーン

sūra(t) →スーラ

tā' تاء アラビア文字「ター」(ت). 文字名称は teh とも。
仮名表記では tā'* (ط) と区別が付かないことに注意。

U+062A ARABIC LETTER TEH

tā' تاء アラビア文字「ター」(ٹ)。アラビア語では /t/ の咽頭化音を表す。文字名称は tah とも。仮名表記では tā* (ㄊ) と区別が付かないことに注意。

U+0637 ARABIC LETTER TAH

tā' marbūṭa ター・マルブータ。「結ばれた tā* (ㄊ)」の意。文字としては元来 hā* (ه) であるが、音価が /t/ であるため、のちに二つの識別点*を載せたものである。

女性名詞や形容詞女性形の語尾に現れる。

アラビア文字の総数 28 字の中には数えられていない。これは tā' の異体字として扱われているからだろう。

→ワクフ

Ta'liq →タアリーク体

tajwid →タジュウィード

tanwīn →タンウィーン

tashdīd →シャッド

tashkīl →タシュキール

taṭwīl →カシーダ

thā' ثاء アラビア文字「サー」(ث)。アラビア語では無声歯舌摩擦音 /θ/ を表す。文字名称は theh とも。

U+062B ARABIC LETTER THEH

Thuluth →スルス体

type 活字。

typesetting 植字 (文字組版)。

vocalisation marks →母音記号

vowel 母音

vowelisation 母音記号*を付けること。

vowelisation marks →母音記号

waqf →ワクフ

waṣla(h) →ワスラ

wāw واو アラビア文字「ワーウ」(و)。アラビア語では子音 /w/ を表すほか、長母音 /ū/ の表記に使われる。左接しない。Fa : vāv (ヴァーヴ)。

U+0648 ARABIC LETTER WAW

writing mode →書字方向

xiǎo'érjīng →小児錦 (シャオアルチン)

yā' ياء アラビア文字「ヤー」(ي)。アラビア語では子音 /y/ を表すほか、長母音 /ī/ の表記に使われる。文字名称は yeh とも。→アリフ・マクスーラ

U+064A ARABIC LETTER YEH

zā' ظاء アラビア文字「ザー」(ظ)。アラビア語では /z/ の咽頭化音 /z/ を表す。文字名称は zah とも。

U+0638 ARABIC LETTER ZAH

zāy زاي アラビア文字「ザーイ」(ز)。左接しない。文字名称は zain とも。

U+0632 ARABIC LETTER ZAIN

zabar زبر →ザバル

zēr →ゼール

あ

アイン →'ayn

アクロフォニー acrophony 表意文字から表音文字を生み出す際にその文字が表す単語の頭の音をその音価とする方法。例えば、アラム語では家 (bet) をかたどった文字が /b/ の音を表すのに使われた (これがギリシア文字の β [ベータ] の起源)。日本語では頭音字法ともいう。

また、文字の音価が文字名称の頭の音に一致していることをこう呼ぶ (頭音原理)。

アブジャド 'abjad アラビア文字を用いた数の表記体系。

アブド・アル=マリク 'Abd 'al-Malik ウマイヤ朝 (661–750, 首都ダマスカス) 第 5 代カリフ*。在位 685–705。イラク総督ハッジャージュ・ブン・ユースフ*を重用。

アラビア系文字 広義のアラビア文字*。つまり、アラビア語を表記するためのアラビア文字だけでなく、それを他の言語に適用・改変した諸文字を合わせた総称。ペルシア文字*, パシュト=文字*, ウルドウ=文字*, シャムキー=文字*, スインド文字*, ジャウイ*, チャガタイ文字, 新ウイグル文字*, カザフ文字*, アルハミーヤ*, 小児錦*などがある。

アラビア書道 Arabic Calligraphy 広義にはアラビア系文字*によるカリグラフィー全般を指す。狭義にはアラビア語*のものをいう。ペルシア語*のものは特にペルシア書道と呼ぶこともある。

イスラーム*以前にも文字を美しく書く努力はなされてきたであろうが、書道として発展していくのはイスラーム以後、とりわけウマイヤ朝 (661–750) に入ってからのものである。

芸術としてのアラビア書道のテーマ (題材) は、当初はもっぱらクルアーン*の章句であり、これは今日に到るまで主要なテーマであり続けている。しかし、ペルシア語文化圏では 10 世紀頃 (?) から詩が重要なテーマであるなど、地域差もある。このことは書体についても言え、地域や言語によって好まれる書体が異なる (同じアラブ圏でも東と西の差は歴然としている)。また、同じ地域・言語でも、題材や場面による書体の使い分けがある。

書家: イブン・ムクラ*, イブン・アル=バウワブ*, ヤークート・アル=ムスタアスィミー*

書体: グバル体*, クーフア体*, シェキヤステ体*, ジャリー・ディーワーン体*, スルス体*, タアリーク体*, タウキーウ体*, 中国書体*, ディーワーン体*, テュズ体*, ナスタアリーク体*, ナスフ体*, マグリブ体*, ムハッカク体*, ライハーン体*, リカーウ体*, ルクア体*

→カラム, トウグラ, ヌクタ, バスマラ, 六体

アラビア数字 →インド・アラビア数字

アラビア文字 Arabic script 狭義には、アラビア語を表記するために用いられる、ナバテア文字*に由来する単子音文字*。セム系文字*の一つである。5–6 世紀に成立し、7 世紀以降、正書法が整備された。イスラームの拡大に伴って、他の言語の表記にも用いられるようになったが、広義にはこれらも含めていう。本稿では広義のアラビア文字をいうときには「アラビア系文字*」の語を用いている。

シリア系文字に由来するとする説もあるが、今日ではあまり支持されていないらしい。

アラム文字 Aramaic script 古代オリエントで広く通用したアラム語を表記するための、フェニキア文字*から派生した文字。アラム文字からは極めて多数の文字が派生した。

アリフ →'alif

アリフ・マクスーラ →'alif maqṣūrah

アルハミーヤ aljamía スペイン語を表記するためのアラビア系文字*。キリスト教支配下の隠れイスラーム教徒（表向きはキリスト教に改宗）によって用いられた。

アラビア語に無い音は、たとえば /p/ 音ならシャツダ*を用いてこのように表記する。

マグリブ体*で書かれたので、fā*は **ف**、qāf*は **ق** と書かれる。

アル=ハリール・ブン・アフマド 8世紀にアラブ文法学・アラビア語正書法の確立に貢献。

アルファベット alphabet 音素文字体系をいう。したがってセム系文字*などの単子音文字*も含む。ただし、狭義には母音をも表記するギリシア文字の系統のものに限定的に使われることもある。語源的にはギリシア文字の最初の二文字の名称アルファ (αλφα) とベータ (βητα) をつなげたものである。

アラビア語*にはこれに対応する言葉としてアリフバー (الفاء'alifbā') があるが、これももちろん、アルファとベータに対応するアラビア文字の最初の二文字の名称をつなげたものである。ペルシア語ではアレフバー ('alefbā) という。

アーンミーヤ フスハー*に対し、それぞれの地域で話されている個別の口語アラビア語をいう。たとえばエジプトのアーンミーヤ (つまりエジプト方言) では、**ح** を /j/ ではなく /g/ の音で発音する。文章は普通はフスハーで書かれるため、アーンミーヤは正書法が確立していない。しかし、たとえばエジプトのアーンミーヤでは標準的な書き方ができつつあるらしい。

イスラーム 'islām اسلام 預言者ムハンマド*が創始した一神教。アラビア文字の確立・発展の原動力となった宗教である。マホメット教と呼ばれたこともあるが、イスラームでは預言者に神性は無く、イスラーム教徒はこのような呼び方を嫌うので用いるべきでない。また、回教とも呼ばれるが、これは中国国内のイスラーム化した民族である回族の宗教という意味であり、適当な呼称ではない。

En: Islam

イブン・ムクラ 'ibn Muqla(h) ابن مقلة アッバース朝 (751-1258) 時代の大書家兄弟の名。作品は現存していない。兄のアブー・アリー・ムハンマド・ビン・アリー (940年没) は同朝の宰相であり、ヌクタ*の大きさを基準に各書体のプロポーションを理論化した。アラビア書道*への二人の寄与の区分は判然としていない。

イブン・アル=パウワーブ 'ibn 'al-Bawwāb ابن البواب ブワイフ朝 (932-1062) の書家。1032年没。イブン・ムクラ*の書道理論を発展させた。また、クルアーン*を暗記して 64

回書したといわれる。ナスフ体*、ムハッカク体*で著名。

インド・アラビア数字 10進法に基づき、10個の記号(つまり数字)を用いる数の表記システムがインドで発明され、8世紀後半にアラブに伝わった。これが11世紀頃(?)ヨーロッパに伝わり、日本でも「算用数字」と呼ばれて広く使われるに到っている。しかし、現在の(1)ヒンディー語の数字、(2)アラビア語・ペルシア語の数字、(3)ヨーロッパの諸言語の数字は、字体*がかなり違っている(表参照)。そこで、それぞれをどのように呼ぶかが問題となる。

欧米・日本・アラブともに、(3)の字体のものを「アラビア数字」と呼ぶことが多い。その場合、アラブでは(2)の字体のものを「インド数字」と呼ぶ。このような用語では、(1)の字体のものを何と呼ぶべきか。

数学者は、字体の相違を無視して、数の表記システムとして「インド・アラビア数字」という用語を使うことがある。筆者としては、これに倣ったうえで、(1)-(3)の字体の区別を「デーヴァナーガリー文字の数字」「アラビア文字の数字」「ラテン文字の数字」と呼びたいと思う。なお、ペルシア語圏では4,5,6の字体がアラビア語圏とは違っており、これはペルシア文字の数字と呼びたい。ウルドゥー語においてはさらに4,7の字体が若干違う。

ラテン文字	0	1	2	3	4	5	6	7	8	9
アラビア文字	٠	١	٢	٣	٤	٥	٦	٧	٨	٩
ペルシア文字	۰	۱	۲	۳	۴	۵	۶	۷	۸	۹
デーヴァナーガリー文字	०	१	२	३	४	५	६	७	८	९

Unicode においては上記の4スク립トについてはそれぞれ別にコードポイントを持つが、ウルドゥー語については特別なコードポイントを持たない。

数字は、アラビア系文字*と一緒に使う場合でも左から右へ書く。数字の後ろに%が来る場合は ١٩% のように%を数字列の左側に置く。

→decimal comma

ウイグル新字 =新ウイグル文字

右接形 アラビア系文字*の続け書きにおいて、前文字から続いてきてその文字で書き終える(続け書きを終える)場合の字形*をいう。

たとえば **ع ج ب** の右接形は **ع ج ب** である。右接形は、独立形の形をよく保持しており、前から続く線を付加したに過ぎないものが多い。

末字形とも呼ばれる。「語末形」なる用語が用いられていることがあるがこれは適切でない。なぜなら、語末においても前文字が〈左接しない文字*〉の場合は独立形*で書かれるし、自身が左接しない文字の場合は語中でも右接形が現れうるからである。

En: final form

→左接形, 両接形

ウルドゥー文字 ウルドゥー語を表記するためのアラビア系文字*。有気音を表すのに、対応する無気音の文字のあとに do čašmī hē(二眼の hē*) と呼ばれる文字 (ه) を付け、2字で1子音を表す。ここに、アラビア文字が持っていた、1字で1子音を表すという原則は破られている。

大きいイェー baṛī yē بڑی یے ウルドゥー文字*で、語末の /e, ai/ を表すための yē (ی) の異体字 (ع)。

→小さいイェー

U+06D2 ARABIC LETTER YEH BARREE (ع)

オスマン語 トルコ系の言語で、アラビア語・ペルシア語の語彙を多量に受容して成立したオスマン朝の宮廷の言語。アラビア系文字で表記された。その正書法は煩雑で、いくつもの音価をもつ文字があったり、アラビア語・ペルシア語からの外来語は、オスマン語における発音を無視して原綴りどおりに記すなど、習得には難しい面があった。20世紀に入ってラテン文字化し、現在のトルコ語の正書法ができた。

音素 phoneme 音素を正確に説明するのは難しいが、簡単に言えば、一つの言語において現れる音声をリストアップし、意味の区別に寄与しない差異を無視してグループ分けしたそれぞれが音素である。音韻ともいう。

たとえば、「三代」「三階」「三枚」の「ん」は音声学的にはまったく違った音であるが、日本語においてはこの違いは後続子音に引きずられて自然に起こるもので、語の識別には関わっていない。そこで、日本語学ではこれら三つの音を同じ音素とみなす。

音素記号は、/k/ のように // に挟んで表記することになっている。これに対し、音声記号は [k] のように [] に挟んで表記する。

か

カウサーニ qawsāni قوسان ラテン文字の括弧 () とその形状・用法ともにほぼ同じ記号。قوس (qawisa : 曲がった) の双数形*である。

カザフ文字 カザフ語を表記するためのアラビア系文字*。
飾り括弧 al-muzharrāni المزهران クルアーン*を引用するための括弧。

カシーダ kashīda(h) ジャスティファイの都合などにより単語全体の長さを長くするために文字の間に挿入される短い横線 (.)。ペルシア語で「引き伸ばされた」という意味のキャッシュデー (کشیده kashīde) がアラビア語化したもの。アラビア語ではタトウィール (تطويل taṭwīl) ともいう。例：مصر→مصر (Miṣr, エジプトの意)

地図で地名を引き伸ばして組む場合にも使う。

書道においても、装飾性などの理由から特定の箇所を横に引き伸ばして書くことがよく行われるが、この考えを活版やタイプライターの世界に持ち込んだのがカシーダではないかと思う。

U+0640 ARABIC TATWEEL (.)

カスラ kasra(h) كسرة 母音記号の一つ ()。基本的な役割は、アラビア語では短母音 /i/, ペルシア語では同 /e/ を表すこと。形はファトハ*と同じ。語源は「割ること」、つまり口を割る(横に薄く開ける)母音が /i/ だから。原則として文字の下に置くが、シャッダ*と合成する場合に限り、シャッダの直下に置く。

ペルシア語ではキャスレ (كسره kasre) のほか、「下部」という意味のズィール (زیر zīr) ともいう。ウルドゥー語ではゼール (زیر zēr) である。

カスラのタンウィーン* () は /in/ の音を表す。原則として文字の下に置くが、シャッダ*と合成する場合に限り、シャッダの直下に置く。

U+0650 ARABIC KASRA ()

U+064D ARABIC KASRATAN ()

括弧 →カウサーニ

カラム qalam قلم ペン一般を指すアラビア語だが、とくに葦(あし)の茎の先端を平らに削った筆記具をいう。



アラビア書道に用いるカラム (これは竹製)

カリフ khalīfa(h) خليفة 元来はムハンマド*の後継者・代理人を指す。アラビア語ではハリーフア。カリフはその訛りである。第4代までのカリフをとくに正統カリフと呼び、順にアブー・バクル、ウマル、ウスマーン、アリーである。ウスマーンの時代に啓示の本格的な集成が行われ、クルアーン*の権威ある原本となった。これをウスマーン本と呼ぶ。

En : caliph, calif, kaliph, kalif

感嘆符 ʾalāmat at-taʾajjub علامة التعجب ラテン文字の感嘆符と形状・用法ともにほぼ同じ。アラビア語ではアラーマ・タッタアッジュブ (「感嘆の記号」の意)。Unicode ではラテン文字の感嘆符と共用。

カンマ →ファースイラ, decimal comma

基字 basic shape 識別点を取り去ったアラビア文字*の字形。たとえばب と ث は同じ基字 (ب) を持つ。

疑問符 ʾalāmat al-istifhām علامة الاستفهام ラテン文字の疑問符を左右反転したような字形。用法はほぼ同じ。

U+061F ARABIC QUESTION MARK (؟)

キャスレ kasre كسره =カスラ*

ペルシア語読み。

ギユメ muzdawjāni مزدوجان フランス語のギユメ (guillemet) が導入されたもので、引用符として用いる。アラビア語ではムズダウジャーニという。مزدوج (muzdawij : 二重の) の双数形*である。Unicode ではラテン文字のギユメと共用だが、「»」ではなく二重パーレンのような字形«»」が一般的である。

グバール体 Ghubār غبار ナスフ体*から生まれた、極めて小さく書くための書体。原義は「塵」。Ghubārī غباريとも。

クーファ体 kūfī クーフィー体とも。マッカ体, マディーナ体, バスラ体と並んで、アラビア文字*の最初期の書体。他の3書体は廃れたが、クーファ体は発展を続け、

さまざまなバリエーションを生み、今日でも多用されている。イラクの古都クーファで発展したのでこの名がある。

دعوة إلى الحروف العربية

クーファ体の例

『アラビア文字を書いてみよう読んでみよう』 [7], p. 92, 本田孝一筆

クラーン al-Qur'ān القرآن イスラームの聖典。「読誦されるもの」の意より。610年から632年にかけて、預言者ムハンマドにアラビア語で下された啓示をのちに集成したもの。114の章（スーラ*）よりなる。当初は識別点*も母音記号*もなしにまとめられたが、正確な読みの必要性から、のちに識別点・母音記号などを導入した。これにより、フスハー*の正書法が確立した。

合字 ラテン文字においては、(1) フランス語において文法的な理由で用いる œ (oeを連結したもの) や、(2) 印刷上の都合と美的観点から用いられる fiなどをいう。

続け書き*するアラビア系文字*においては、(1) lām* (ل) に'alif* (ا) が続く場合に義務的に用いられる字形群（独立形*：لا, 右接形*：لا), および(2) 美的観点から二つ以上の文字を組み合わせた特別な字形をいう。

En: ligature

→ラーム・アリフ合字

小書きのミーム クラーン*の正確な発音指示のために用いられる記号。nūn (ن) の上に小さく mīm (م) を書く。bā' (ب) の直前の nūn を /m/ で発音することを示す。U+06E2 ARABIC SMALL HIGH MEEM ISOLATED FORM (ﻥ̣)

語根 アラビア語の多くの単語は2~4の子音が意味の中核を担っており、それらにつく母音の変化および接辞の付加などにより変形・派生語を作る。この子音列を語根と呼ぶ。

たとえば kataba (書く) という動詞の語根は k-t-b である。ここから kitābun (本), kitāban (本を), kitābin (本の), kitābatun (書くこと), kātibun (作家), maktabun (事務所), maktabatun (図書館, 書店), maktūbun (書かれた) といった単語が派生する。最後の二例のように、母音を入れ換えるだけでなく語頭に他の音節を付ける派生形もある。語根 k-t-b を表すのに√KTB という表記が用いられることがある。

En: root.

コーラン →クラーン

コロ al-nuqṭatāni النقطتان 句読点の一つ。ラテン文字のものと同形状・用法ともにおおむね同じ。アラビア語ではヌクタターニという（ヌクタ*の双数形*）。Unicode ではラテン文字のコロンと共用。

コンマ →ファースィラ

さ

ザイド・イブン・サービト Zayd 'ibn Thābit زيد بن ثابت ムハンマド*の秘書の一人。ペルシア語、ギリシア語など数ヶ国語に通じ、ムハンマド*に送られてくる外国語文書の翻訳を行っていたとされる。初代カリフ*であるアブー・バクルの命によりクラーン*を集成した。

左接形 アラビア系文字*の続け書き*において、その文字から書き始めて後ろの文字に続く場合の字形*をいう。頭字形とか左接形とも呼ばれる。「語頭形」なる用語が用いられることがあるがこれは適切ではない。なぜなら、語頭においても（左接しない文字*）は独立形で綴られるし、左接形は語中でも現れうるからである（前文字が左接しない文字の場合）。

En: initial form

→独立形, 右接形, 両接形, 左接しない文字

左接しない文字 アラビア系文字*の続け書き*において、特定の文字は後ろに続けない。その特定の文字とは、アラビア語においてはا (合字لの場合を含む), و, ز, ر, د, ذ である。他のアラビア系文字でも同様で、たとえばペルシア文字ではこれにځが付け加わる、といった具合である。これらを本稿では「左接しない文字」と呼ぶことにする。

Fa: خط مقطعه (khatte moqatta'e)

ザバル zabar →ファトハ

ザンメ zamme =ダンマ*

ペルシア語読み。

小児錦 シェオアルチン（ピンイン表記では xiǎo'érbīng）と読む。中国のイスラーム教徒の間で使われる、漢語のアラビア系文字*表記。小児錦の他にも「小經 (xiǎojīng)」などいくつかの呼び方がある。単一の正書法があるわけではない。

シェキャステ・タアリーク体 =ディーワーン体*

→シェキャステ体

シェキャステ体 Shekasteh 「崩された」という意味のペルシア語であり、書体については本来草書体一般を指す用語だが、単にシェキャステ体といえ、ナスタアリーク体の草書（シェキャステ・ナスタアリーク体）を指すのが普通。シェキャステ・ナスタアリーク体は17世紀にヘラート（現アフガニスタン内）で発達した書体で、ペルシア語圏ではナスタアリーク体とともに広く使われている。

この手の草書体は、（左接しない文字*）も後ろに続けて書く。

シカスタ体 =シェキャステ体*

アラビア語もしくはダリー語読みである。

字形 本稿では二つの意味で用いている：(1) アラビア系文字*における独立形*・左接形*・両接形*・右接形*の別, (2) 文字の具体的な形状。

→字体

識別点 判別点ともいう。初期のアラビア文字*では一つの文字に複数の子音が対応しており、文字の音価は文脈で判断しなければならなかったが、のちに文字の上または下に1個ないし3個の点を付けて区別するようにした。これが識別点である。アラビア語では単にヌクタ*（「点」の意）という。

識別点のアイデアはシリア文字を参考にしたものといわれる。

アラビア文字がペルシア語*、ウルドゥー語*、オスマン語、ウイグル語などに適用されるとき、識別点の付加その他により文字種を増やした。

En: diacritic dots

字体 文字の形についての抽象的な概念であって、ある文字をその文字として認識するために最低限必要な文字の視覚的構造をいう。たとえば明朝体の「島」もゴシック体の「島」も、字形*は異なるが字体は同じである。一方、「島」と「嶋」は、意味・発音・起源ともに同じであって、本質的に同じ文字といえるが、字体は異なる。このようなとき「島」と「嶋」は互いに異体字であるという（「嶋」を本字、「島」を異体字とする用語法もある）。こういった諸概念はアラビア系文字*にあてはめることができる。

二つの文字の字形が与えられたとき、それらが同じ字体であるかどうかは社会的な合意によって決まることであり、機械的に決めることはできない。

書体と字体の概念は完全に直行してはならず、書体の違いが字体の違いを引き起こすこともある。たとえば、ナスフ体とナスタアリーク体の kāf 』を比較してみれば分かる。

文字の具体的な形状について「字体」と表現する人が多いが、これは誤りである。

ジャウィ Jawi マレー語（ムラユ語、マレーシア語、インドネシア語ともいう）を表記するためのアラビア系文字*。オリジナルのアラビア文字に加え、ج (/c/)、ڠ (/ny/)、گ (/g/)、ڠ (/ng/)、ف (/p/)、و (/v/) を用いる。注意すべきは、ڠ の左接形、両接形がڠ、ڠではなくڠ、ڠとなることである。

シャクル shakl = 母音記号*

ジャズム jazm = スクーン*

語源は「切断する」。

シャッド shadda(h) شدة 重子音記号 (˘)。文字の上に置き、子音を二つ重ねることを意味する。形は頭文字 ش 由来（識別点*の無い形の頭部）。語源は「激しさ」。タシュディード (تشديد tashdid) ともいう（語根*が同じであることに注意）。

ファトハ*と合成するときはファトハをこの上に置く。カスラ*と合成するときはカスラをこの下に置く（したがってシャッド無しの場合のカスラの位置とは異なる）。カスラのタンウィーン*の場合も同様。

なお、/iy/, /ūw/ を表すのにもシャッド付きの ي، و を用いる。例：ياباني yābāniyun（日本人）

アルハミーヤ*においてはシャッドは全く違う役割をもつ。Fa: シャッド (شدة šadde)。

シャームキー文字 パンジャブ語を表記するためのアラビア系文字*。

ジャリー・ディーワーン体 ディーワーン体*の変種で、装飾性が高い書体。画線の間の空間をヌクタ*や飾り線で埋めつくす点に特徴がある。ディーワーン体との違いはそれだけではなく、文字の骨格や画線の先端の形状も異なる。

重子音記号 = シャッド

終止符 'alāmat 'rsq;al-waqf علامة الوقف アラビア語・ペルシア語では終止符としてピリオド* (.) を用いるが、ウルドゥー語では「.」を用いる。

U+06D4 ARABIC FULL STOP (.)

書字方向 文字および行を配列する方向。左から右に文字を並べるのを「左横書き」、その反対を「右横書き」と呼ぶ（左右どちらから起こすか、による命名）。アラビア系文字*は右横書きであるが、数字の部分は左横書きになるので、全体としては双方向的である。

ウイグル文字、モンゴル文字は縦書きで行を左から右へ配列するが、これは右横書きであったソグド文字を反時計回りに90°回転させて書いたことに由来する。

En: writing mode

新ウイグル文字 新疆ウイグル自治区で使われている、現代ウイグル語を表記するためのアラビア系文字*。母音に対し文字を割り当てている。したがって、もはや単子音文字*ではない。

なお、単に「ウイグル文字」といった場合、通常はソグド文字を改良してウイグル語を表記するようにしたものを指し、まったく別物である。「新ウイグル文字」はこれと区別するための呼称である。ウイグル新字ともいう。

シンド文字 → スィンド文字

スィーバワイヒ Sibawayhi アル=ハリール・ブン・アフマド*の弟子。793年没。アラブ文法学、アラビア語正書法の確立に貢献。名は「リンゴの香り」という意味の雅号である。

スィンド文字 西インド、スィンド地方のスィンド語 (Sindhi; スィンディー語、シンディー語とも) を表記するためのアラビア系文字*。スィンド語はもともとランダー文字、グルムキー文字、デーヴァナーガリー文字などで書かれていたが、イスラーム化によってペルシア文字*が浸透し、これを改良してスィンド文字とした。

スィンド語は反り舌音をもち、短母音/長母音が対をなし、有気音/無気音の対立があるなど、ヒンディー語、ウルドゥー語と音韻的に似た点が多い。

En: Sindhi script

数字 → インド・アラビア数字

スクリプト → script

スクーン sukūn 母音なし記号 (◌̣)。文字の上に置き、その子音が後ろに母音を伴わないことを表す。語源は「静か

な」。ジャズム*ともいう。ウルドゥー語ではもっぱらジャズムと呼ばれている(たぶん)。

当初, ḥā' حの頭部をとった形(𐤀)をしていたが, のちに丸みを帯び, ついには小円状になった。ウルドゥー文字*では逆v字型のスクーンが使われるらしい。

Fa: ソクーン (sokūn)。

U+0652 ARABIC SUKUN (◌ْ)

スーラ sūra(t) سورة クルアーン*における「章」。クルアーンは114章からなる。英語でもchapterとはいわず, suraという。

スルス体 Thuluth ثلث 六体*の一つ。「書体の母」と呼ばれ, ここから多数の書体が派生した。名前は「1/3」を意味するثلثthulthから。これは, ある基準値の1/3の幅のカラム*を用いることを元来は意味した。

Fa: ソルス体。

正則アラビア語 =フスハー*

正則語 =フスハー*

セム系文字 単にセム文字ともいう。ごくおおざっぱに言えば, セム諸語で用いられた諸文字で, 原カナン文字に起源を持つヘブライ文字, サマリア文字, シリア文字, マンダ文字, アラビア系文字*, エチオピア文字などを指す。字形だけでなく, 単子音文字*であるといった構造的共通点を持つ。

En: Semitic script

ゼール zēr =カスラ*

双数形 アラビア語の名詞・形容詞には, 文法上の数に関して三つの区分(単数, 双数, 複数)がある。双数は二つのもの, 複数はいくつ以上のものに使われる。数によって単語はその形を変える。このうち双数の場合の語形を双数形とよぶ。

ソクーン sokūn سكون →スクーン

ペルシア語読み。

ソルス体 solsi ثلثي =スルス体*

ペルシア語読み。

た

タアリーク体 Ta'liq تعلق (1) 12世紀頃にイランで誕生し, 13世紀に確立したとされる書体。公文書に最もよく用いられた。(2) オスマン朝ではナスターアリーク体*の意味。

太陰文字 =月文字*

太陽文字 アラビア語の28文字のうちの14字: ت, ث, د, ذ, ر, ز, س, ش, ص, ض, ط, ظ, ع, غ, ف, ق, ك, ل, م, ن, ه, و, ي, かなわち, 調音点がlām لに近いものの文字である。これ以外の14字を月文字*または太陰文字という。

定冠詞 الのつく単語の語頭が太陽文字である場合, لはその太陽文字に同化する。例: السلام (al+salām) で السلام (as-salām)。

タウキーウ体 tawqī' توقيع 六体*の一つ。公文書・勅令・書簡に用いられた書体。タウキーウという語には署名という意味がある。

ター・マルブータ →tā' marbūṭah

タシュキール tashkil تشكيل 母音記号*を付すこと。

タシュディード tashdīd تشديد =シャッダ*

タトウィール taṭwīl تطويل =カシーダ*

タンウィーン tanwīn تنوين 名詞・形容詞の語尾の母音の後に/n/音が付くこと, およびそれを表す記号。正書法上, この/n/音は文字には表わさないの, 補助記号であるタンウィーンが使われる。タンウィーンとは「ヌーン化」というような意味。アラビア語の名詞・形容詞は限定(特定のもの)／非限定(一般的なものを指す)の区別があるが, 非限定の場合にヌーン化する。例: kitābun (ある一冊の本) / al-kitābu (その一冊の本)

三つの母音記号ファトハ*, カスラ*, ダンマ*それぞれに対応したタンウィーンがある。

タンウィーンはアラビア語特有の現象であるが, ファトハのタンウィーン(対格を表す)を持つ語が副詞としてペルシア語に入っており, これはペルシア語の正書法でもタンウィーン記号を付ける。

→ワクフ

ダンマ ḍamma(h) アラビア語では短母音 /u/ を表す母音記号。文字の上に置く。形は文字 wāw* و, に由来。語源は「すぼめること」, つまり口をすぼめる母音が /u/ だから。ペルシア語では短母音 /o/ または二重母音 /ow/ を表す。ペルシア語では ḍammah が訛ってザンメ (ضمه zamme) というほか, 「前部」という意味のピーシュ (پیش pīsh) ともいう。ウルドゥー語ではペーシュ (پیش pēsh) である。ダンマのタンウィーン (◌ُ) は /u/ の音を表す。もともとは他のタンウィーンと同様に, ダンマを二つ横に並べた形状 (◌ُ◌ُ) だが, のちに二つ巴紋のような形状 (◌ُ◌ُ) が現れ, 現在では ◌ُ が一般的である。字形のバリエーションはこれ以外にもある。

U+064C ARABIC DAMMATAN (◌ُ)

短剣アリフ 母音記号*の一種。本来, 長母音 /ā/ を表す'alif が書かれてしかるべき位置に書かれない綴りの場合に, それと分からせるために子音字の上に書かれる短い'alif 状の記号。クルアーンの不完全表記をカバーする。「短いアリフ」ともいう。例: لكن lākin (しかし), موسى mūsā (人名: モーセ)

U+0670 ARABIC LETTER SUPERSCRIPT ALEF (◌ْ)

単子音文字 各文字がもっぱら子音のみを表す文字体系。セム系文字は基本的に単子音文字である。ただし, たとえばアラビア文字*では, 長母音 /ā, ī, ū/ を表すために子音字 و, ي, و を援用するなど, 単子音文字の原則から離れた点もある。こうした援用は他のセム系文字*にも見られる。ヒエログリフは単子音文字であり, これが強い影響を与えたとされる原シナイ文字, およびそこから派生したフェニキア文字*, アラム文字*などの多くが単子音文字である。

ただし, フェニキア文字から派生したギリシア文字では母音字を導入し, ここから派生したエトルリア文字, ラテン文字, キリル文字などはすべて母音字をもつ。

新ウイグル文字*はアラビア系文字*であるが、母音字を導入しているため、単子音文字ではない。

小さいイエー *čhöpī yē* چۈپۈي يە ウルドゥー文字*において、文字 *yē* は子音 /y/ および母音 /i, ē, ai/ を表すが、語末の /ē, ai/ に限って「大きいイエー*」と呼ばれる異体字 (چۈپۈي) を用いる。それ以外の *yē* の字体 (چۈپۈي) を「小さいイエー」と呼ぶ。

中国書体 中国のイスラーム教徒の間で用いられた書体。

中字形 = 両接形*

月文字 アラビア語の文字 28 字のうち、太陽文字*以外の 14 字をいう。太陰文字とも。

続け書き アラビア系文字*はラテン文字における筆記体や仮名の連綿のように、続け書きを行う。手で書く場合は一続きの線を書いてから縦画・識別点*などを書き足す。各文字は、続け書きの始め・中・終わりのいずれかによって字形*を変える。その文字から書き始めて後ろへ続く場合の字形を「左接形*」、前文字から続いてきて後ろへ続く場合を「両接形*」、前文字から続いてきてその文字で終わる場合を「右接形*」と呼ぶ。

続け書きが途切れるのは以下の箇所である：(1) (左接しない文字*) の後ろ、(2) 単語の間。単語の間には基本的に空白を入れる。しかしアラビア語の定冠詞 *al* や、一文字からなる前置詞は必ず後ろの単語に続けて書く。また、ペルシア語で単語では、空白は入れないが続け書きを切断する綴りがある。このように、空白を入れるか入れないか、続け書きするかしないかのルールが言語ごとにいろいろあるので、注意しなければならない。

停止形 → ワクフ

ディーワーン体 *Diwānī* ديواني タアリーク体*の草書として作られた書体。それゆえ、シェキヤステ・タアリーク体とも呼ばれる。15 世紀にオスマン朝において公用文書・勅令などに用いられた。ディーワーンは行政機関の意味である (オスマン朝においては国会会議)。速書きできる事務書体という面もあるが、装飾性が高く、書道でも好んで使われる。



ディーワーン体の例

(『アラビア文字を書いてみよう読んでみよう』 [7], p. 92, 本田孝一筆)

テュズ体 *tüz xät* تۈز خەت 新疆ウイグル自治区で現代ウイグル語を表記するための書体。文字どおりには「直立体」の意。ナスフ体*を元にしていていると思われるが、*'alif** (ا), *lām** (ل) などの縦画がほとんど垂直で、*bā** (ب) などの底部がまっすぐ水平に伸びている特徴がある。

ياپون تلسدن ئاساس

テュズ体の例 : yapon tilidin 'asas (日本語基礎)

点無しヌーン ウルドゥー語で用いられる、*nūn** (ن) の異体字 (نۈ). 鼻母音化を表す。デーヴァナーガリー文字

のチャンドラ・ビンドゥと同じ働きである。意味的には通常の *nūn* とはまったく異なる文字だが、文字表では異体字の扱いである。左接形* (نۈ), 両接形* (نۈ) において、通常の *nūn* と区別が付かない (というより、書き手の意識としては語末でしか書き分けないということなのだろう)。ウルドゥー語ではヌーン・グンナという。

En : dotless nūn

U+06BA ARABIC LETTER NOON GHUNNA (نۈ)

頭音原理 *acrophony* → アクロフォニー

頭音字法 *acrophony* → アクロフォニー

トゥグラ *tughrā'* طغراء 元来はオスマン朝のスルタン (君主) が自らの名をある様式で装飾的にまとめた一種の署名であって、その位置づけは我が国の花押 (かおう) に似ている。のちにバスマラ*などの文言をトゥグラのように書く芸術が行われ、アラビア書道*の一様式となった。トルコ語ではトゥグラという。

頭字形 = 左接形

独立形 アラビア系文字*は前後の文字と続け書きし、それによって字形を変えるが、とくに単独の文字がとる字形を独立形とよぶ。

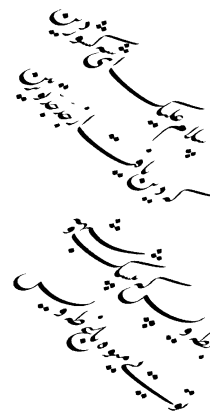
hā' の独立形は ه であり、語中に独立形で現れる場合はこの字体であるが、辞書の見出しや何らかの略号として単独で書かれるときは ه という字体が使われることがある。

En : free standing form

→ 左接形, 両接形, 右接形, 続け書き

な

ナスタアリーク体 *Nasta'liq* نستعلیق タアリーク体*が発展した書体。15 世紀にミール・アリー・タブリーズィーが考案し、ミール・エマード・カズヴィーニー (1615 没) が完成したとされる。ペルシア語圏、ウルドゥー語圏での基本書体。オスマン朝ではタアリーク体と呼ばれた。



ナスタアリーク体の例

(ミール・エマード筆、四行詩を *čalipā* という形式で書いたもの)

ナスフ体 *Naskhī* نسخي 六体*の一つで、クルアーン*の書写のために開発された書体。判読・筆記とも比較的習得しやすい。ナスヒー体とも。11 世紀頃から、クルアーン*の書体は次第にクーファ体*からナスフ体に置き換えられていった。この頃のナスフ体をとくに古ナスフ体と呼ぶ。

アラビア語圏（一部地域を除く）の基本書体であり、印刷物にも広く使われる。ペルシア語圏でも印刷物はナスフ体が基本書体である。

ナバテア語 セム語の一つ。ナバタイ語とも。ペトラを中心に栄えたナバテア王国（200BC頃～AD100頃）の言語。

ナバテア文字 ナバテア語*を表記するためのセム系文字*。ナバタイ文字とも。アラム文字の流れを汲む。アラビア文字*はナバテア文字をアラビア語に適用したものとされるが、両者の境目は必ずしも判然としない。

二眼の hē ウルドゥー文字*の一つ。直前の子音字が有気音であることを示す。独立形*：ه، 左接形*：ه، 両接形*：ه، 右接形*：ه。

それに対し、ウルドゥー文字の通常の hē は、独立形*：ه، 左接形*：ه، 両接形*：ه， 右接形*：هである。

二眼の hē は、「ウルドゥー文字 35 字」の中で通常の hē と別個には数えられていない。

U+06BE ARABIC LETTER HEH DOACHASHMEE (ه)

U+06C1 ARABIC LETTER HEH GOAL (ه)

ヌクタ nuqta(h) نكتة アラビア語で「点」という意味の一般名詞だが、アラビア系文字*に関していうときには、とくに識別点*およびピリオド*を指す。アラビア書道*では文字のプロポーシオンを決定するのに、ヌクタの大きさを基準にする。

Fa：ノクテ (noqte)。

ヌーン・グンナ →点無しヌーン

ノクテ noqte نكتة →ヌクタ

ペルシア語読み。

は

バスマラ basmala(h) بسملة 次の句：بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ (慈悲深く慈悲あまねきアッラーの御名において)

クルアーン*の各章（スーラ*）は、第9章を除きバスマラで始まる。食事・講演の前など、何かを始めるときに唱えられるほか、手紙・書類の始め、映画のオープニングなどに掲げられる。また、アラビア書道*に好んで取り上げられる題材でもある。

ハッジャージュ・ブン・ユースフ Ḥajjāj bun Yūsuf حجاج بن يوسف
ウマイヤ朝のカリフ*であるアブド・アル＝マリク*に仕えたイラク総督（661-714）。アラビア文字*の識別点*を確立したと言われる。

ハムザ hamza(h) همزة 声門閉鎖音を表す記号 (ء)。アラビア語では、(1)'alif* (ا)の上または下、(2)識別点*無しの上 yā* (ي)の上、(3)wāw* (و)の上に付加するか、(4)単独で使われる。形は'ayn* (ع)の頭部に由来('ayn*と声門閉鎖音は調音点が近い)。Fa：ハムゼ。

U+0621 ARABIC LETTER HAMZA (ء)

U+0623 ARABIC LETTER ALEF WITH HAMZA ABOVE (أ)

U+0624 ARABIC LETTER WAW WITH HAMZA ABOVE (ؤ)

U+0625 ARABIC LETTER ALEF WITH HAMZA BELOW (إ)

U+0626 ARABIC LETTER YEH WITH HAMZA ABOVE (ئ)

U+0654 ARABIC HAMZA ABOVE (ْ)

U+0655 ARABIC HAMZA BELOW (ِ)

ハラカ ḥaraka(h) حركة 母音記号*。本来「動き」という意味の名詞だが、「母音即ち口の動き」というところから母音記号をも指すようになった。複数形はハラカート (حركات ḥarakāt)

Fa：ハラキヤト (حركة ḥarakat), [複]ハラカート (حركات ḥarakāt)

ハリーフ →カリフ

パーレン parenthesis →カウサーニ

判別点 →識別点

ピリオド nuqta(h) نقطة 句読点の一つ。形状・用法ともにラテン文字におけるピリオドとおおむね同じ。アラビア語ではヌクタ*という。

→終止符

ファースイラ faṣīla(h) فاصلة アラビア語におけるコンマ（カンマ）。ラテン文字のそれを180°回転したような形状（.）をもつ。用法はラテン文字のそれとおおむね同じ。فضل (faṣala：分離する、区切る)の派生語。なお、これに対応するペルシア語 فاصله fāsele は、コンマではなく「(字間の)スペース」の意である。→decimal comma

Fa：ヴィールグール (ویرگول vīrgūl) <Fr

ファトハ faṭḥa(h) فتحة 短母音 /a/ を表す母音記号 (ا)。形は'alif*が起源。語源は「開くこと」、つまり口を広く開ける母音が /a/ だから。

ペルシア語ではファトヘ (fathe) のほか、「上部」という意味のザバル (زabar) ともいう。ウルドゥー語ではザバルである。

U+064E ARABIC FATHA (ا)

ファトハのタンウィーン* (اِ) は /an/ の音を表す。他のタンウィーンとは異なり、タンウィーンのつく文字の後ろに正書法上、さらに'alif を付け加える（いくつかの条件の場合には付けない）。近年、タンウィーンをこの付加した'alif の上に載せたものが見られる。おそらく誤りだが、ペルシア語では非常によく見かける。

حقیقة ḥaqīqatan (実際のところ) のようにター・マルブータ*の上に付ける場合もある。

U+064B ARABIC FATHATAN (اِ)

フェニキア文字 フェニキア人の用いたセム系文字*で、アラビア文字*の先祖。単子音文字*であることや字形・文字名称・文字順序などの特徴がフェニキア文字からアラビア文字に引き継がれた。

フスハー fuṣḥā فصحاء 正則語または正則アラビア語と訳されている。古典北アラビア語に基づいた、アラビア語の共通文語だが、公的な場やニュースでは口頭でも使われ

る。クルアーンのアラビア語はフスハーであり、その規範でもある。反対語はアーンミーヤ*。

二つ目の hē → 二眼の hē

文脈形 → ワクフ

ペルシア語 Fārsī فارسی 本稿でいうペルシア語は、言語学的には近世ペルシア語を指す。サーサーン朝が7世紀、アラブ・イスラームに征服されたのを契機に、それまでの中世ペルシア語が多量のアラビア語集を受容して近世ペルシア語が成立し、現在に到っている。中世ペルシア語はアラム文字系のパフラヴィー文字で書かれたが、近世ペルシア語は一貫してアラビア系文字で書かれている。サーサーン朝の崩壊から、ペルシア語文献が現れるまでに約200年が経過しており、この期間を沈黙の二世紀と呼ぶ。

ペルシア文字 近世ペルシア語を表記するためのアラビア系文字*。オリジナルのアラビア文字に پ/p/, چ/č/, ژ/z/, گ/g/の4文字を追加している。これらの文字は一度に追加されたのではなく13世紀から15世紀ごろにかけて徐々に使われだしたらしい。

アラビア語(フスハー*)とペルシア語の音素体系が異なるため、同じ文字でも発音は異なるものがある。正書法上もさまざまな違いがある。

母音記号 shakl شکل ファトハ*, カスラ*, ダンマ*とそれらのタンウィーン*, およびスクーン*, そして短剣アリフ*の総称。これ以外の記号も含めていうことがある。

→ハラカ, タシュキール

母音点 アラビア文字*の母音記号*が今日の形を取る以前は、点を付加して短母音を表記する方式があった。文字の上, 中, 下に一つの点を打つと、それぞれ/a, u, i/を表す。文字の上に点二つを横に並べて打つと、タンウィーンを表す。識別点*とまぎらわしいため、母音点は赤などの色インクを使って書かれた。現在は使われない。

ま

マグリブ 'al-Maghrib المغرب 「日の没する地」転じて「西方」を意味するアラビア語だが、具体的にはチュニジア, アルジェリア, モロッコなどの西北アフリカを指す。もとベルベル人の土地だが、7世紀以降にアラブ人が進入し、徐々にイスラーム化・アラブ化した。イスラーム圏のなかでも政治的・文化的に独自の発展を遂げた。書体や正書法も独自の様式をもつ。

マグリブ体 Maghribī مغربي クーフア体*の一つを元に、マグリブ*において作られた一群の書体。ナスフ体*などとの違いは単に書体としての違いにとどまらず、fā* (ف) は ب, qāf* (ق) は ق と書かれる。これは古い方式の識別点*が、他の書体では廃れ、マグリブにおいては残ったからである。

また、マグリブでは文字表の順序も他の地域とは異なっている。

末字形 = 右接形*

マッダ madda(h) مدّة alif*の上に置かれる記号。マッダ付き'alif (ا) はそれ単独で /ā/ の音になる。理論的には子音文字としての'alif (ا) に、長母音を示す'alif が続いたものと考えてよい。形は مد に由来するとする説がある。語源は「伸ばすこと」(/a/を伸ばすから)。

なお、クルアーンの章名 يس におけるように, ي の上に付くこともある(例外的)。Fa: マッデ (madde), Ur: マッド (madd)。

マホメット →ムハンマド

短いアリフ = 短剣アリフ*

ムハッカク体 Muḥaqqaq محقق 六体*の一つで、「書体の父」と呼ばれ、さまざまな書体を派生した。ムハッカクの本義は「確立された」「成し遂げられた」。

ムハンマド Muḥammad イスラーム*を創始した預言者(570-632)。「マホメット」の表記が長らく定着していたが、近年、原語に近い「ムハンマド」の表記が普及しつつある。イスラームでは、ムハンマドをアブラハム(アラビア語ではイブラーヒーム), モーセ(同ムーサー), イエス(同イーサー)らの預言者の系列の最後の預言者としている。

ムハンマド自身は文盲であったが、文字の重要性を見抜き、ザイド・イブン・サービト*ら優秀な秘書官を抱えた。また、子弟に読み書きを奨励したとされる。

や

ヤーカート・アル=ムスタアシミ Yāqūt 'al-Musta'īmī ياقوت المستعصي 1298年没。アッバース朝最後のカリフ*に仕えた大書家。六体*を確立し、とくにスルス体*で著名。モンゴル軍の襲撃時にも超然として書を行っていたという伝説がある。

ら

ライハーン体 Rayhānī ريحاني 六体*の一つで、ムハッカク体*から派生したクルアーン*筆写用の書体。書家の名(ريحان Rayhān)から。

ラーム・アリフ合字 lām (ل) の次に'alif (ا) が来る場合、必ず合字にしなければならない。ナスフ体*の場合、独立形は لا, 右接形は لا である。左接はしない。独立形の場合、2本の画線が根元で交差しているため、理論的には右の角(つ)が'alif, 左の角が lām のはずであるが、現在は右の角が lām, 左の角が'alif という合意がある。母音記号などを付加する位置は、この解釈に基づいて行う。

リカーウ体 Riqā' رفاق 六体*の一つで、タウキーウ体*から派生した速書き書体。単語としては Ruq'ah の複数形だが、ルクア体*とはまったく別の書体。

両接形 アラビア系文字*の続け書き*において、前文字と後文字の両方につながる字形。中字形ともいう。例: ب, ع の両接形はそれぞれ مع ب, مع ع

「語中形」という用語が使われることもあるが、不適切である。なぜなら語中であっても左接形*／右接形*が現れうるから。

〈左接しない文字*〉は当然両接形を持たない。両接形が現れるのは、(1)前文字が左接する文字であり、かつ(2)自身も左接する文字であり、かつ(3)後ろに文字が来る場合に限る。

En : medial form

ルクア体 Ruq'a(h) رُقعة 簡素で力強い骨格を持った書体。速書きのための工夫がなされている。垂れ幕や商品の値札など広範に使われている。たとえば、二つの識別点*は横線で、三つの識別点は山形で表す。また、ヌーンの独立形・右接形では画線の末尾を屈曲させて識別点を表す。地域にもよるが、現在アラブで行われている手書き文字はルクア体の骨格に基づいている。

リカーウ体*とは別物だがよく混同される。

六体 'aqlāmi sitta(h) اقلام ستة アラビア書道*の基本6書体で、スルス体*、ナスフ体*、ムハッカク体*、ライハーン体*、タウキーウ体*、リカーウ体*を指す。六書ともいう。

→ヤーカート・アル・ムスタアスイミー

わ

ワクフ waqf وقف 元来はアラビア語で「休止」の意。単語を単独で発音したり、文中の意味の切れ目で息を継ぐ場合のアラビア語の読み方のルール。ワクフの場合の発音を「停止形」、そうでない場合を「文脈形」と呼ぶ。

タンウィーン*のない単語がワクフで読まれると、語末の母音が消える。

ファトハ*のタンウィーン*がワクフで読まれると、/an/ が /ā/ になる（これが、ファトハのタンウィーンにおいて'alif*が書かれている理由である）。

ダンマ*とカスラ*のタンウィーンがワクフで読まれると、/un, in/ の音が消える。例：طَالِبٌ (男子学生) tālibun→tālib

なお、tā' marbūṭa* (ة) で終わる単語がワクフで読まれる場合、ةではなく。として読む。これがそもそも tā' marbūṭa が（識別点*無しの時代に）。として書かれた理由である。さらに、口語では /h/ 音も消える傾向にある。

例：طَالِبَةٌ (女子学生) tālibatun→tālibah→tāliba

アラビア語の単語をラテン文字や片仮名で表記する場合、ワクフの読み方を元に行っていることが多いが、この h を表記しないこともある。例：shaddah/shadda

ワスラ waṣla(h) وصلة 語頭の'alif*に載せる記号 (i)。これがついた場合、その音節は脱落する。形はوصلةのصの頭部に由来。

定冠詞الのlは、先行語があると必ず脱落する。

例：وَعَلَيْكُمْ السَّلَامُ で wa'alaykumussalām.

U+0671 ARABIC LETTER ALEF WASLA (i)

ワーウ →wāw