Arabic alphabet

From Wikipedia, the free encyclopedia

The **Arabic alphabet** (Arabic: أَبْجَدِيَّة عَرَبِيَّة
'abjadiyyah 'arabiyyah) or **Arabic abjad** is
the Arabic script as it is codified for writing the
Arabic language. It is written from right to left,
in a cursive style, and includes 28 letters.
Because letters usually[1] stand for consonants,
it is classified as an abjad.

Arabic abjad

العربية

Type	Abjad
Languages	Arabic
Time period	400 to the present
Parent systems	Proto-Sinaitic - Phoenician - Aramaic - Syriac - Nabataean - Arabic abjad
Child systems	N'Ko alphabet
ISO 15924	`Arab, 160`
Direction	Right-to-left
Unicode alias	Arabic
Unicode range	U+0600 to U+06FF (http://www.unicode.org/charts/PDF/U0600.pdf) U+0750 to U+077F (http://www.unicode.org/charts/PDF/U0750.pdf) U+08A0 to U+08FF (http://www.unicode.org/charts/PDF/U08A0.pdf) U+FB50 to U+FDFF (http://www.unicode.org/charts/PDF/UFB50.pdf) U+FE70 to U+FEFF (http://www.unicode.org/charts/PDF/UFE70.pdf) U+1EE00 to U+1EEFF (http://www.unicode.org/charts/PDF/U1EE00.pdf)

Note: This page may contain IPA phonetic symbols.

Arabic alphabet

ا ب ت ث ج ح
خ د ذ ر ز س
ش ص ض ط ظ ع

<div dir="rtl">

غ ف ق ك ل

م ن ه و ي

</div>

History · Transliteration

Diacritics · Hamza ء

Numerals · Numeration v · т · е (//en.wikipedia.org/w/index.php?title=Template:Arabic_alphabet&action=edit)

Contents

Consonants

The Arabic alphabet has 28 letters. Adaptations of the Arabic script for other languages, such as Persian, Ottoman, Sindhi, Urdu, Malay or Pashto, Arabi Malayalam, have additional letters, shown below. There are no distinct upper and lower case letter forms.

Many letters look similar but are distinguished from one another by dots (*'i'jām*) above or below their central

part, called *rasm*. These dots are an integral part of a letter, since they distinguish between letters that represent different sounds. For example, the Arabic letters transliterated as *b* and *t* have the same basic shape, but *b* has one dot below, ب, and *t* has two dots above, ت.

Both printed and written Arabic are cursive, with most of the letters within a word directly connected to the adjacent letters.

Alphabetical order

There are two main collating sequences for the Arabic alphabet:

- The original *'abjadī* order (أَبْجَدِي), used for lettering, derives from the order of the Phoenician alphabet, and is therefore similar to the order of other Phoenician-derived alphabets, such as the Hebrew alphabet. In this order letters are also used as numbers.

- The *hijā'ī* (هِجَائِي) or *'alifbā'ī* (أَلِفْبَائِي) order shown in the table below, used where lists of names and words are sorted, as in phonebooks, classroom lists, and dictionaries, groups letters by similarity of shape.

The *'abjadī* order is not a simple historical continuation of the earlier north Semitic alphabetic order, since it has a position corresponding to the Aramaic letter *samek*/*semkat* ס, yet no letter of the Arabic alphabet historically derives from that letter. Loss of *samek* was compensated for by the split of *shin* שׁ into two independent Arabic letters, ش (*shīn*) and س (*sīn*) which moved up to take the place of *samek*.

The most common *'abjadī* sequence is:

غ	ظ	ض	ذ	خ	ث	ت	ش	ر	ق	ص	ف	ع	س	ن	م	ل	ك	ي	ط	ح	ز	و	ه	د	ج	ب	أ
gh	ẓ	ḍ	dh	kh	th	t	sh	r	q	ṣ	f	'	s	n	m	l	k	y	ṭ	ḥ	z	w	h	d	j	b	'

Note: In this sequence, and all those that follow, the letters are presented in Arabic writing order, i.e., right to left. The Latin script transliterations are also in this order, with each placed under its corresponding letter. Thus, the first letter of the sequence is "أ"(') at the right, and the last letter in the sequence is "غ"(gh), at the left.

This is commonly vocalized as follows:

> *'abjad hawwaz ḥuṭṭī kalaman saʿfaṣ qarashat thakhadh ḍaẓagh.*

Another vocalization is:

> *'abujadin hawazin ḥuṭiya kalman saʿfaṣ qurishat thakhudh ḍaẓugh*

Another *'abjadī* sequence (probably older, now mainly confined to the Maghreb), is:[2]

ش	غ	ظ	ذ	خ	ث	ت	س	ر	ق	ض	ف	ع	ص	ن	م	ل	ك	ي	ط	ح	ز	و	ه	د	ج	ب	أ
sh	gh	ẓ	dh	kh	th	t	s	r	q	ḍ	f	'	ṣ	n	m	l	k	y	ṭ	ḥ	z	w	h	d	j	b	'

which can be vocalized as:

> *'abujadin hawazin ḥuṭiya kalman ṣaʿfaḍ qurisat thakhudh ẓaghush*

Modern dictionaries and other reference books do not use the *'abjadī* order to sort alphabetically; instead, the

newer *hijā'ī* order (with letters partially grouped together by similarity of shape) is used:

ي	و	ه	ن	م	ل	ك	ق	ف	غ	ع	ظ	ط	ض	ص	ش	س	ز	ر	ذ	د	خ	ح	ج	ث	ت	ب	أ
y	w	h	n	m	l	k	q	f	gh	'	ẓ	ṭ	ḍ	ṣ	sh	s	z	r	dh	d	kh	ḥ	zh	th	t	b	'

Another kind of *hijā'ī* order used to be widely used in the Maghreb until recently when it was replaced by the Mashriqi order:[2]

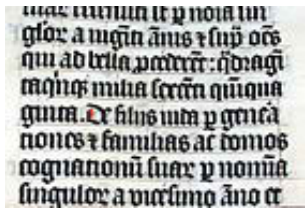| ي	و	ه	ش	س	ق	ف	غ	ع	ض	ص	ن	م	ل	ك	ظ	ط	ز	ر	ذ	د	خ	ح	ج	ث	ت	ب	أ
y	w	h	sh	s	q	f	gh	'	ḍ	ṣ	n	m	l	k	ẓ	ṭ	z	r	dh	d	kh	ḥ	zh	th	t	b	'

Letter forms

Unlike cursive writing based on the Latin alphabet, the standard Arabic style is to have a substantially different shape depending on whether it will be connecting with a preceding and/or a succeeding letter, thus all primary letters have conditional forms (allographs), depending on whether they are at the beginning, middle or end of a word, so they may exhibit four distinct forms (*initial*, *medial*, *final* or *isolated*). However, six letters (و ز ر ذ د ا) have only an isolated or final form, and so force the following letter (if any) to take an initial or isolated form, as if there were a word break. For example, أرارات (Ararat) has only isolated forms, because each letter cannot be connected to its adjacent one.

Some letters look almost the same in all four forms, while others show considerable variation. Generally, the initial and middle forms look similar except that in some letters the middle form starts with a short horizontal line on the right to ensure that it will connect with its preceding letter. The final and isolated forms, are also similar in appearance but the final form will also have a horizontal stroke on the right and, for some letters, a loop or longer line on the left with which to finish the word with a subtle ornamental flourish. In addition, some letter combinations are written as ligatures (special shapes), including *lām-'alif*.[3]

Table of basic letters

Notes

- See the article *Romanization of Arabic* for details on various transliteration schemes; however, Arabic language speakers don't follow a standardized scheme when transcribing names. Also names are regularly transcribed as pronounced locally, not as pronounced in Literary Arabic (if they were of Arabic origin).

- Regarding pronunciation, the phonemic values given are those of the pronunciation of Literary Arabic, the standard which is taught in universities. In practice, pronunciation may vary considerably from region to region, because Literary Arabic isn't anyone's native language. For more details concerning the pronunciation of Arabic, consult the articles *Arabic phonology* and *varieties of Arabic*.

- The names of the Arabic letters can be thought of as abstractions of an older version where they were meaningful words in the Proto-Semitic language. Names of Arabic letters may have quite different names popularly, but they are not provided in the article.

Calligraphy

- Arabic
- Chinese
- Georgian
- Indian
- Japanese
- Korean
- Kufic
- Nepalese
- Persian
- Sini
- Tibetan
- Western
- Mongolian

For example: ح *ḥā'* is most commonly known in Egypt as: IPA: [ħɑ]; in Lebanon: IPA: [ħe]. ز has two Literary Arabic names: *zayn/zāy* and called by Egyptians: IPA: [zeːn].

- Six letters (ا د ذ ر ز و) don't have a distinct medial form and have to be written with their final form without being connected to the next letter. Their initial form matches the isolated form.

Arabic letters usage in Literary Arabic

Name	Translit.	Value (IPA)	Contextual forms			Isolated
			Final	**Medial**	**Initial**	
'alif	' / *ā*	various, including /aː/ [a]	ـا	ـا	ا	ا
bā'	*b*	/b/ (sometimes /p/ in loanwords)[b]	ـب	ـبـ	بـ	ب
tā'	*t*	/t/	ـت	ـتـ	تـ	ت
thā'	*th* (also *ṯ*)	/θ/	ـث	ـثـ	ثـ	ث
jīm	*j* (also *ǧ, g*)	[d͡ʒ] ~ [ʒ] ~ [g] [c]	ـج	ـجـ	جـ	ج
ḥā'	*ḥ*	/ħ/	ـح	ـحـ	حـ	ح
khā'	*kh* (also *ḫ, k̲*)	/x/	ـخ	ـخـ	خـ	خ
dāl	*d*	/d/	ـد	ـد	د	د
dhāl	*dh* (also *ḏ*)	/ð/	ـذ	ـذ	ذ	ذ
rā'	*r*	/r/	ـر	ـر	ر	ر
zayn / zāy	*z*	/z/	ـز	ـز	ز	ز
sīn	*s*	/s/	ـس	ـسـ	سـ	س
shīn	*sh* (also *š*)	/ʃ/	ـش	ـشـ	شـ	ش
ṣād	*ṣ*	/sˤ/	ـص	ـصـ	صـ	ص

ḍād	ḍ	/dˤ/	ض	ـض	ـضـ	ضـ
ṭāʾ	ṭ	/tˤ/	ط	ـط	ـطـ	طـ
ẓāʾ	ẓ	[ðˤ] ~ [zˤ]	ظ	ـظ	ـظـ	ظـ
ʿayn	ʿ	/ʕ/	ع	ـع	ـعـ	عـ
ghayn	gh (also ġ, ḡ)	/ɣ/ (sometimes /g/ in loanwords)[c]	غ	ـغ	ـغـ	غـ
fāʾ	f	/f/ (sometimes /v/ in loanwords)[b]	ف [d]	ـف	ـفـ	فـ
qāf	q	/q/ (sometimes /g/ in loanwords)[c]	ق [d]	ـق	ـقـ	قـ
kāf	k	/k/ (sometimes /g/ in loanwords)[c]	ك	ـك	ـكـ	كـ
lām	l	/l/	ل	ـل	ـلـ	لـ
mīm	m	/m/	م	ـم	ـمـ	مـ
nūn	n	/n/	ن	ـن	ـنـ	نـ
hāʾ	h	/h/	ه	ـه	ـهـ	هـ
wāw	w / ū / aw	/w/, /uː/, /aw/, sometimes /u/, /o/, and /oː/ in loanwords	و	ـو	ـو	وـ
yāʾ	y / ī / ay	/j/, /iː/, /aj/, sometimes /i/, /e/, and /eː/ in loanwords	ي [e]	ـي	ـيـ	يـ

- **^a** *Alif* can represent many phonemes in Literary Arabic:

  1. Without diacritics: ا
     - initially: *a, i*   /a, i/ or sometimes silent in the definite article ال *(a)l-*
     - medially or finally: *ā*   /aː/.

  2. *Alif* with *hamzah* above: أ
     - initially: *'a, 'u*   /ʔa, ʔu/
     - medially or finally: *'a*   /ʔa/.

ا

3. *’Alif* with *hamzah* under: ٳ
   - initially: *’i* /ʔi/; doesn't appear medially or finally (see hamza).

آ

4. *’Alif* with *maddah* : آ
   - initially, medially or finally: *’ā* /ʔaː/.

- ^**b** /p/ and /v/ can be represented by پ and ڤ/ڦ or if unavailable, ب and ڥ/ف are used, respectively.

- ^**c** For Arabic language speakers, the phoneme /g/ can be represented using different letters, depending on local dialects. ج is normally used in Egypt, also sometimes Yemen and Oman. ق is used where it represents the [g] in local dialects. ك or غ are used where /g/ doesn't exist in local dialects. Other letters such as گ, ڭ or ڤ may also be used, but are not regarded as standard Arabic letters. Likewise, where ج represents [g], it can be also used for /ʒ/~/d͡ʒ/, or the letter چ can be used in Egypt.

- ^**d** *Fā’* and *qāf* are traditionally written in northwestern Africa as ڥ and ڧ ڧ ق, respectively, while the latter's dot is only added initially or medially.

- ^**e** *Yā’* in the isolated and the final forms in handwriting and print in Egypt, Sudan and sometimes other places, is always undotted ى, making it only contextually distinguishable from *’alif maqṣūrah*.

See also Additional letters below.

Further notes

- The letter *’alif* originated in the Phoenician alphabet as a consonant-sign indicating a glottal stop. Today it has lost its function as a consonant, and, together with *ya’* and *wāw*, is a *mater lectionis*, a consonant sign standing in for a long vowel (see below), or as support for certain diacritics (*maddah* and *hamzah*).

- Arabic currently uses a diacritic sign, ء, called *hamzah*, to denote the glottal stop [ʔ], written alone or with a carrier:
  - alone: ء ;
  - with a carrier: أ إ (above or under a *’alif*), ؤ (above a *wāw*), ئ (above a dotless *yā’* or *yā’ hamzah*).

In academic work, the *hamzah* (ء) is transliterated with the modifier letter right half ring (ʾ), while the modifier letter left half ring (ʿ) transliterates the letter *‘ayn* (ع), which represents a different sound, not found in English.

- Letters lacking an initial or medial version are never linked to the letter that follows, even within a word. The *hamzah* has a single form, since it is never linked to a preceding or following letter. However, it is sometimes combined with a *wāw*, *yā’*, or *’alif*, and in that case the carrier behaves like an ordinary *wāw*, *yā’*, or *’alif*.

- The shape of the final *yā’* is always undotted ى in both print and handwriting in Egypt and Sudan, mainly.

Modified letters

The following are not individual letters, but rather different contextual variants of some of the Arabic letters.

Conditional forms				Name	Translit.	Phonemic Value (IPA)
Isolated	Final	Medial	Initial			
آ	ـآ	ـآ	آ	'alif maddah	*'ā*	/ʔaː/
ة	ـة			tā' marbūṭah	*h* or *t* / *h* / *ẗ*	/a/, /at/
ى	ـى			'alif maqṣūrah[4]	*ā* / *ỳ*	/aː/

Ligatures

The use of ligature in Arabic is common. There is one compulsory ligature, that for *lām + 'alif*, which exists in two forms. All other ligatures (*yā' + mīm*, etc.) are optional.

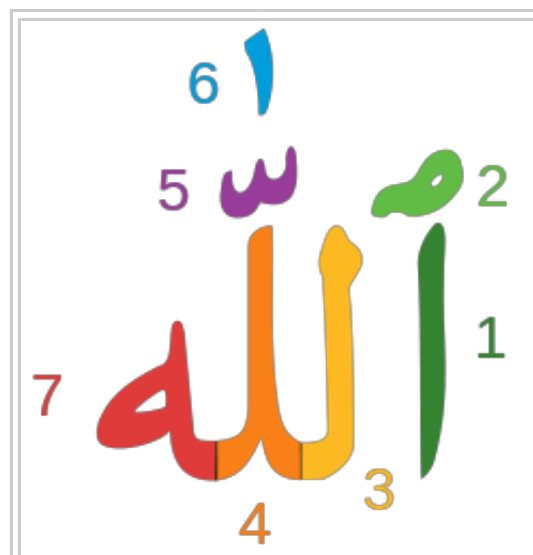| Contextual forms				Name
Final	Medial	Initial	Isolated	
لا		لا		lām + 'alif

Components of a ligature for "Allah":
1. alif
2. hamzat waṣl (همزة وصل)
3. lām
4. lām
5. shadda (شدة)
6. dagger alif (ألف خنجرية)
7. hāʾ

A more complex ligature that combines as many as seven distinct components is commonly used to represent the word *Allāh*.

Unicode primary range for basic Arabic language alphabet is the U+06xx range. Other ranges are for compatibility to older standards and do contain some ligatures. The only compulsory ligature for fonts and text processing in the basic Arabic language alphabet range U+06xx are ones for *lām + alif*. All other ligatures (*yā' + mīm*, etc.) are optional. Example to illustrate it is below. The exact outcome may depend on your browser and font configuration.

- *lām + 'alif*

لا

Note: Unicode also has in its Presentation Form B FExx range a code for this ligature. If your browser and font are configured correctly for Arabic, the ligature displayed above should be identical to this one, U+FEFB ARABIC LIGATURE LAM WITH ALEF ISOLATED FORM:

لا

- U+0640 ARABIC TATWEEL + *lām* + *'alif*

لـا

Note: Unicode also has in its Presentation Form B U+FExx range a code for this ligature. If your browser and font are configured correctly for Arabic, the ligature displayed above should be identical to this one:

- U+FEFC ARABIC LIGATURE LAM WITH ALEF FINAL FORM

لا

Another ligature in the Unicode Presentation Form A range U+FB50 to U+FDxx is the special code for glyph for the ligature *Allāh* ("God"), U+FDF2 ARABIC LIGATURE ALLAH ISOLATED FORM:

الله

This latter ligature code again is a work-around for the shortcomings of most text processors, which are incapable of displaying the correct vowel marks for the word *Allāh* in Koran. Because Arabic script is used to write other texts rather than Koran only, rendering *lām + lām + hā'* as the previous ligature is considered faulty:[5] If one of those fonts are installed on a computer (mry_KacstQurn, KacstOne, DejaVu Sans, Scheherazade, Lateef) the right will appear without automatically adding gemination mark and superscript Alef.

- *lām + lām + hā'*

  لله or لله

- *'alif + lām + lām + hā'*

  الله or الله

- *'alif + lām +* U+0651 ARABIC SHADDA + U+0670 ARABIC LETTER SUPERSCRIPT ALEF + *hā'*

  الله (*DejaVu Sans* and *KacstOne* don't show the added superscript Alef)

An attempt to show them on the faulty fonts without automatically adding the gemination mark and the

superscript Alef, although may not display as desired on all browsers, is by adding the `U+200d` (Zero width joiner) after the first or second *lām*

- (*'alif* +) *lām* + *lām* + `U+200d` ZERO WIDTH JOINER + *hā'*

لله الله

Gemination

Further information: Shadda

Gemination is the doubling of a consonant. Instead of writing the letter twice, Arabic places a *W*-shaped sign called *shaddah*, above it. Note that if a vowel occurs between the two consonants the letter will simply be written twice. The diacritic only appears where the consonant at the end of one syllable is identical to the initial consonant of the following syllable. (The generic term for such diacritical signs is *ḥarakāt*).

General Unicode	Name	Transliteration
0651 ّ ّ	shaddah	(consonant doubled)

Nunation

Main article: Nunation

Nunation (Arabic: تنوين *tanwīn*) is the addition of a final -*n* to a noun or adjective. The vowel before it indicates grammatical case. In written Arabic nunation is indicated by doubling the vowel diacritic at the end of the word.

Vowels

Users of Arabic usually write long vowels but omit short ones, so readers must utilize their knowledge of the language in order to supply the missing vowels. However, in the education system and particularly in classes on Arabic grammar these vowels are used since they are crucial to the grammar. An Arabic sentence can have a completely different meaning by a subtle change of the vowels. This is why in an important text such as the *Qur'ān* the three basic vowel signs (see below) are mandated, like the *ḥarakāt* and all the other diacritics or other types of marks, for example the cantillation signs.

Short vowels

Further information: Arabic diacritics

In the Arabic handwriting of everyday use, in general publications, and on street signs, short vowels are typically not written. On the other hand, copies of the *Qur'ān* cannot be endorsed by the religious institutes that review them unless the diacritics are included. Children's books, elementary-school texts, and Arabic-language grammars in general will include diacritics to some degree. These are known as "vocalized" texts.

Short vowels may be written with diacritics placed above or below the consonant that precedes them in the syllable, called *ḥarakāt*. All Arabic vowels, long and short, follow a consonant; in Arabic, words like "Ali" or "alif", for example, start with a consonant: *'Aliyy, 'alif*.

Short vowels (fully vocalized text)	Name	Trans.	Value
064E	faṭḥah	*a*	/a/
064F	ḍammah	*u*	/u/
0650	kasrah	*i*	/i/

Long vowels

In the fully vocalized Arabic text found in texts such as Koran, a long *ā* following a consonant other than a *hamzah* is written with a short *a* sign (*fatḥah*) on the consonant plus an *'alif* after it; long *ī* is written as a sign for short *i* (*kasrah*) plus a *yā'* ; and long *ū* as a sign for short *u* (*ḍammah*) plus a *wāw*. Briefly, $^a a = ā$, $^i y = ī$ and $^u w = ū$. Long *ā* following a *hamzah* may be represented by an *'alif maddah* or by a free *hamzah* followed by an *'alif*.

The table below shows vowels placed above or below a dotted circle replacing a primary consonant letter or a *shaddah* sign. For clarity in the table, the primary letters on the left used to mark these long vowels are shown only in their isolated form. Please note that most consonants do connect to the left with *'alif*, *wāw* and *yā'* written then with their medial or final form. Additionally, the letter *yā'* in the last row may connect to the letter on its left, and then will use a medial or initial form. Use the table of primary letters to look at their actual glyph and joining types.

Long vowels (fully vocalised text)	Name	Trans.	Value
064E 0627 اَ	faṭḥah ʼalif	ā	/aː/
064E 0649 ىَ	faṭḥah ʼalif maqṣūrah	ā / á	/aː/
064F 0648 وُ	ḍammah wāw	ū	/uː/
0650 064A يِ	kasrah yāʼ	ī	/iː/

In unvocalized text (one in which the short vowels are not marked), the long vowels are represented by the vowel in question: *ʼalif*, *ʼalif maqṣūrah* (or *yaʼ*), *wāw*, or *yāʼ*. Long vowels written in the middle of a word of unvocalized text are treated like consonants with a *sukūn* (see below) in a text that has full diacritics. Here also, the table shows long vowel letters only in isolated form for clarity.

Combinations وا and يا are always pronounced *wā* and *yā* respectively, the exception is when وا is the verb ending, where ʼalif is silent, resulting in *ū*.

Long vowels (unvocalized text)	Name	Trans.	Value
0627 ا	(implied faṭḥah) ʼalif	ā	/aː/
0649 ى	(implied faṭḥah) ʼalif maqṣūrah	ā / aỳ	/aː/
0648 و	(implied ḍammah) wāw	ū / uw	/uː/
064A ي	(implied kasrah) yāʼ	ī / iy	/iː/

In addition, when transliterating names and loanwords, Arabic language speakers write out most or all the

vowels as long (*ā* with ١ *'alif*, *ē* and *ī* with ي *ya'*, and *ō* and *ū* with و *wāw*), meaning it approaches a true alphabet.

Diphthongs

The diphthongs /aj/ and /aw/ are represented in vocalized text as follows:

Diphthongs (fully vocalized text)	Name	Trans.	Value
064E 064A يَ	fatḥah yā'	*ay*	/aj/
064E 0648 وَ	fatḥah wāw	*aw*	/aw/

Vowel omission

An Arabic syllable can be open (ending with a vowel) or closed (ending with a consonant):

- open: CV [consonant-vowel] (long or short vowel)
- closed: CVC (short vowel only)

In closed syllables, we can indicate that the closing consonant does not carry a vowel by marking it with a diacritic called *sukūn* (ْ   ) to remove any ambiguity, especially when the text is not vocalized. A normal text is composed only of series of consonants; thus, the word *qalb*, "heart", is written *qlb*. The *sukūn* indicates where not to place a vowel: *qlb* could, in effect, be read *qalab* (meaning "he turned around"), but written with a *sukūn* over the *l* and the *b* (قَلْبْ), it can only have the form *qVlb*. This is one step down from full vocalization, where the vowel *a* would also be indicated by a *fatḥah*: قَلْبْ.

The *Qur'ān* is traditionally written in full vocalization. Outside of the *Qur'ān*, putting a *sukūn* above a *yā'* (representing /i:/), or above a *wāw* (representing /u:/) is extremely rare, to the point that *yā'* with *sukūn* will be unambiguously read as the diphthong /aj/, and *wāw* with *sukūn* will be read /aw/. For example, the letters *m-w-s-y-q-ā* (موسيقى with an *'alif maqṣūrah* at the end of the word) will be read most naturally as the word *mūsīqā* ("music"). If one were to write a *sukūn* above the *wāw*, the *yā'* and the *'alif*, one would get مْوْسيقْ, which would be read as **mawsayqāy* (note however that the final *'alif maqṣūrah*, because it is an *'alif*, never takes a *sukūn*). The word, entirely vocalized, would be written as مُوسِيقَى. The Koranic spelling would have no *sukūn* sign above the final *'alif maqṣūrah*, but instead a miniature *'alif* above the preceding *qāf* consonant, which is a valid Unicode character but most Arabic computer fonts cannot in fact display this miniature *'alif* as of 2006.

No *sukūn* is placed on word-final consonants, even if no vowel is pronounced, because fully vocalized texts are always written as if the *'I'rāb* vowels were in fact pronounced. For example, *'Aḥmad zawj sharrīr*, meaning "Ahmed is a wicked husband", for the purposes of Arabic grammar and orthography, is treated as if still pronounced with full *'I'rāb*, i.e. *'Aḥmadu zawjun sharrīrun* with the complete desinences.

General Unicode	Name	Translit.	Phonemic Value (IPA)
0652 ْ	sukūn	(no vowel with this consonant letter or diphthong with this long vowel letter)	∅
0670 ٰ	’alif above	*ā*	/aː/

The *sukūn* is also used for transliterating words into the Arabic script. The Persian word ماسک (*mâsk*, from the English word "mask"), for example, might be written with a *sukūn* above the س to signify that there is no vowel sound between that letter and the ک.

Additional letters

Regional variations

- ڢ – a Maghrebi variation of the letter ف (*fā’*) .

- ڧ and ڡ – a Maghrebi variation of standard letter ق (as a rule, dotless in isolated and final positions and dotted in the initial and medial forms ڧ‍ ‍ڧ‍ ‍ڡ).

Additional modified letters, used in non-Arabic languages, or in Arabic for transliterating names, loanwords, spoken dialects only, include:

Sometimes used for writing names, loanwords and dialects

- ڤ – (not to be confused with ڨ) used in Kurdish language when written in Arabic script and sometimes used in Arabic language to represent the sound /v/ when transliterating names and loanwords in Arabic. Also used in writing dialects with that sound.[6] (Usually the letter ف (*fā’*) transliterates /v/.) Also used as *pa* in the Jawi script. The phoneme /v/ in Tunisia and some other regions of Maghreb is rendered using ڥ

- پ – used to represent the phoneme /p/ in Persian, Urdu, and Kurdish; sometimes used in Arabic language when transliterating names and loanwords, although Arabic mostly substitutes /b/ for /p/ in the transliteration of names and loanwords. So, "7up" can be transcribed as سفن أب or سڤن أپ.

- چ – used to represent /t͡ʃ/ ("ch"). It is used in Persian, Urdu, and Kurdish and sometimes used when transliterating names and loanwords in Arabic. In the Iraqi spoken dialect it may be used, especially when referring in the feminine, although it is rarely written, as well as rarely used in the Maghrebi spelling. Nevertheless, Arabic usually substitutes other letters in the transliteration of names and loanwords: normally the combination تش (*tā’-shīn*) is used to transliterate the /t͡ʃ/, as in "Chad". In Egypt چ is used for /ʒ/ (or /d͡ʒ/, which is approximated to [ʒ]). In Israel, it's used to render /g/ in Arabic language, for example on roadsigns.

- *Ca* in the Jawi script.

- گ – used to represent /g/. Normally used in Persian, Kurdish, and Urdu.[6] Often names and loanwords with /g/ are transliterated in Arabic with ك (*kāf*), ق (*qāf*), غ (*ghayn*) or ج (*jīm*), which may or may not change the original sound. In Egypt ج is normally pronounced [g].

- ڨ – a Maghrebi letter, sometimes used for [g] (not to be confused with ڤ). In Tunisia it is sometimes used to represent the phoneme /g/. In final and isolate form it has the form which resembles the letter ق *qāf* whence it derives.

- پش – a Maghrebi letter for [t͡ʃ].

Used in languages other than Arabic

Main article: Arabic script

Numerals

Main articles: Western Arabic numerals and Eastern Arabic numerals

There are two main kinds of numerals used along with Arabic text; Western Arabic numerals and Eastern Arabic numerals. In most of present-day North Africa, the usual Western Arabic numerals are used. Like Western Arabic numerals, in Eastern Arabic numerals, the units are always right-most, and the highest value left-most.

Letters as numerals

Main article: Abjad numerals

In addition, the Arabic alphabet can be used to represent numbers (Abjad numerals). This usage is based on the *'abjadī* order of the alphabet. ا *'alif* is 1, ب *bā'* is 2, ج *jīm* is 3, and so on until ي *yā'* = 10, ك *kāf* = 20, ل *lām* = 30, …, ر *rā'* = 200, …, غ *ghayn* = 1000. This is sometimes used to produce chronograms.

Western (Maghreb, Europe)	Central (Mideast)	Eastern/Indian (Persian, Urdu)
0	٠	٠
1	١	١
2	٢	٢
3	٣	٣
4	٤	۴
5	٥	۵
6	٦	۶
7	٧	٧
8	٨	٨
9	٩	٩

History

Main article: History of the Arabic alphabet

The Arabic alphabet can be traced back to the Nabataean alphabet used to write the Nabataean dialect of Aramaic. The first known text in the Arabic alphabet is a late 4th-century inscription from *Jabal Ramm* (50 km east of *'Aqabah*) in Jordan, but the first dated one is a trilingual inscription at Zebed in Syria from 512. However, the epigraphic record is extremely sparse, with only five certainly pre-Islamic Arabic inscriptions surviving, though some others may be pre-Islamic. Later, dots were added above and below the letters to differentiate them. (The Aramaic language had fewer phonemes than the Arabic, and some originally distinct

Aramaic letters had become indistinguishable in shape, so that in the early writings 15 distinct letter-shapes had to do duty for 28 sounds; cf. the similarly ambiguous Pahlavi alphabet.) The first surviving document that definitely uses these dots is also the first surviving Arabic papyrus (PERF 558), dated April 643, although they did not become obligatory until much later. Important texts like the *Qur'ān* were and still are frequently memorized, especially in Qur'an memorization, a practice which probably arose partially from a desire to avoid the great ambiguity of the script.

Later still, vowel marks and the *hamzah* were introduced, beginning some time in the latter half of the 7th century, preceding the first invention of Syriac and Hebrew vocalization. Initially, this was done by a system of red dots, said to have been commissioned by an Umayyad governor of Iraq, *Ḥajjaj ibn Yūsuf*: a dot above = *a*, a dot below = *i*, a dot on the line = *u*, and doubled dots indicated nunation. However, this was cumbersome and easily confusable with the letter-distinguishing dots, so about 100 years later, the modern system was adopted. The system was finalized around 786 by *al-Farāhīdī*.

Arabic printing presses

Although Napoleon Bonaparte generally is given the credit with introducing the printing press to Egypt, upon invading it in 1798, and he did indeed bring printing presses and Arabic script presses, to print the French occupation's official newspaper *Al-Tanbiyyah* (*The Courier*), the process was started several centuries earlier.
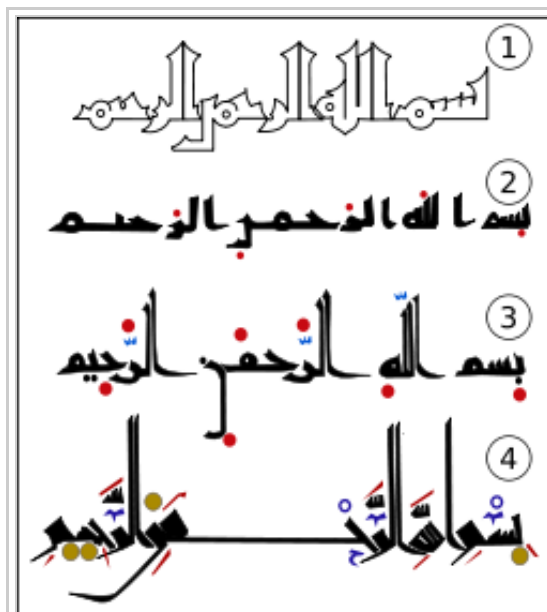
Gutenberg's invention of the printing press in 1450 was followed up by Gregorio de Gregorii, a Venetian, who in 1514 published an entire prayer book in Arabic script entitled *Kitab Salat al-Sawa'i* intended for the eastern Christian communities. The script was said to be crude and almost unreadable.[*citation needed*]

Famed type designer Robert Granjon working for Cardinal Ferdinando de Medici succeeded in designing elegant Arabic typefaces and the Medici press published many Christian prayer and scholarly Arabic texts in the late 16th century.

The first Arabic books published using movable type in the Middle East were by the Maronite monks at the *Maar Quzhayy*

Evolution of early Arabic calligraphy (9th–11th century). The *Basmala* is taken as an example, from Kufic *Qur'ān* manuscripts. (1) Early 9th century script used no dots or diacritic marks;[7] (2) and (3) in the 9th–10th century during the Abbasid dynasty, Abu al-Aswad's system used red dots with each arrangement or position indicating a different short vowel. Later, a second system of black dots was used to differentiate between letters like *fā'* and *qāf*;[8][8] (4) in the 11th century, al-Farāhīdī's system) dots were changed into shapes resembling the letters to transcribe the corresponding long vowels. This system is the one used today.[9]

Monastery in Mount Lebanon. They transliterated the Arabic language using Syriac script. It took a fellow goldsmith like Gutenberg to design and implement the first true Arabic script movable type printing press in the Middle East. The Greek Orthodox monk *Abd Allah Zakhir* set up an Arabic language printing press using movable type at the monastery of Saint John at the town of *Dhour El Shuwayr* in Mount Lebanon, the first homemade press in Lebanon using true Arabic script. He personally cut the type molds and did the founding of the elegant typeface. He created the first true Arabic script type in the Middle East. The first book off the press was in 1734; this press continued to be used until 1899.[10]

Computers and the Arabic alphabet

The Arabic alphabet can be encoded using several character sets, including ISO-8859-6, Windows-1256 and

Unicode (see links in Infobox, above), in the latter thanks to the "Arabic segment", entries U+0600 to U+06FF. However, neither of these sets indicate the form each character should take in context. It is left to the rendering engine to select the proper glyph to display for each character.

For compatibility with previous standards, initial, medial, final and isolated forms can be encoded separately in Unicode; however, they can also be inferred from their joining context, using the same encoding. The following table shows this common encoding, in addition to the compatibility encodings for their normally contextual forms (Arabic texts should be encoded today using only the common encoding, but the rendering must then infer the joining types to determine the correct glyph forms, with or without ligation).

Unicode

Main article: Arabic characters in Unicode

As of Unicode 6.1, the following ranges encode Arabic characters:

- Arabic (0600-06FF)
- Arabic Supplement (0750-077F)
- Arabic Extended-A (08A0-08FF)
- Arabic Presentation Forms-A (FB50-FDFF)
- Arabic Presentation Forms-B (FE70-FEFF)
- Arabic Mathematical Alphabetic Symbols 1EE00-1EEFF)

The basic Arabic range encodes the standard letters and diacritics, but does not encode contextual forms (U+0621-U+0652 being directly based on ISO 8859-6); and also includes the most common diacritics and Arabic-Indic digits. U+06D6 to U+06ED encode Qur'anic annotation signs such as "end of *ayah*" ⟨◯⟩ and "start of *rub el hizb*" ⟨❂⟩. The Arabic Supplement range encodes letter variants mostly used for writing African (non-Arabic) languages. The Arabic Extended-A range encodes additional Qur'anic annotations and letter variants used for various non-Arabic languages. The Arabic Presentation Forms-A range encodes contextual forms and ligatures of letter variants needed for Persian, Urdu, Sindhi and Central Asian languages. The Arabic Presentation Forms-B range encodes spacing forms of Arabic diacritics, and more contextual letter forms. The Arabic Mathematical Alphabetical Symbols block encodes characters used in Arabic mathematical expressions.

See also the notes of the section on modified letters.

Keyboards

Keyboards designed for different nations have different layouts, so that proficiency in one style of keyboard such as Iraq's does not transfer to proficiency in another keyboard such as Saudi Arabia's. Differences can include the location of non-alphabetic characters.
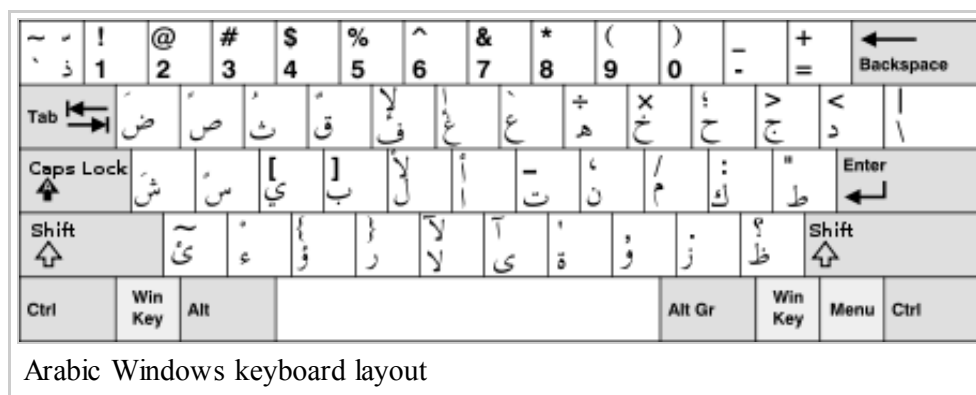

Arabic Mac keyboard layout

All Arabic keyboards allow typing Roman characters, e.g., for the URL in a web browser. Thus, each Arabic keyboard has both Arabic and Roman characters marked on the keys. Usually the Roman characters of an Arabic keyboard conform to the QWERTY layout, but in North Africa, where French is the most common

language typed using the Roman characters, the Arabic keyboards are AZERTY.

To encode a particular written form of a character, there are extra code points provided in Unicode which can be used to express the exact written form desired. The range *Arabic presentation forms A* (U+FB50 to U+FDFF) contain ligatures while the range *Arabic presentation forms B* (U+FE70 to U+FEFF) contains the positional variants. These effects are better achieved in Unicode by using the *zero-width joiner* and *non-joiner*, as these presentation forms are deprecated in Unicode, and should generally only be used within the internals of text-rendering software, when using Unicode as an intermediate form for conversion between character encodings, or for backwards compatibility with implementations that rely on the hard-coding of glyph forms.

Arabic Windows keyboard layout

Intellark imposed on a QWERTY keyboard layout.

Finally, the Unicode encoding of Arabic is in *logical order*, that is, the characters are entered, and stored in computer memory, in the order that they are written and pronounced without worrying about the direction in which they will be displayed on paper or on the screen. Again, it is left to the rendering engine to present the characters in the correct direction, using Unicode's bi-directional text features. In this regard, if the Arabic words on this page are written left to right, it is an indication that the Unicode rendering engine used to display them is out-of-date.[11][12]

There are competing on-line tools, e.g. Yamli editor (http://www.yamli.com/editor/) , allowing to enter Arabic letters without having Arabic support installed on a PC and without the knowledge of the layout of the Arabic keyboard.[13]

Handwriting recognition

The first software program of its kind in the world that identifies Arabic handwriting in real time has been developed by researchers at Ben-Gurion University (BGU).

The prototype enables the user to write Arabic words by hand on an electronic screen, which then analyzes the text and translates it into printed Arabic letters in a thousandth of a second. The error rate is less than three percent, according to Dr. Jihad El-Sana, from BGU's department of computer sciences, who developed the system along with master's degree student Fadi Biadsy.[14]

See also

- Arabic calligraphy

- Arabic diacritics (pointed vowels and consonants)
- Arabic numerals
- Arabic script about other languages written in Arabic letters.
- Arabic Unicode
- Arabic Chat Alphabet
- ArabTeX – provides Arabic support for TeX and LaTeX
- Rasm (unpointed consonants)
- Romanization of Arabic
- South Arabian alphabet
- Perso-Arabic script
- Additional Arabic Letters
- Arabic braille
- Algerian braille

References

1. ^ While there are ways to mark vowels, these are not always employed. Because of this, it is more exactly called an "impure abjad". See Abjad#Impure abjads for a discussion of this nomenclature.
2. ^ *a b* **(Arabic)** Alyaseer.net ترتيب المداخل والبطاقات في القوائم والفهارس الموضوعية Ordering entries and cards in subject indexes (http://alyaseer.net/vb/showthread.php?t=8807) Discussion thread *(Accessed 2009-October–06)*
3. ^ Rogers, Henry (2005). *Writing Systems: A Linguistic Approach*. Blackwell Publishing. p. 135.
4. ^ See the sub-paragraph about the *ʾalif maqṣūrah* at *aleph* in Arabic
5. ^ SIL International: This simplified style is often preferred for clarity, especially in non-Arabic languages (http://scripts.sil.org/cms/scripts/page.php?site_id=nrsi&id=ArabicFonts&_sc=1)
6. ^ *a b* Arabic Dialect Tutorial (http://www1.ccls.columbia.edu/~cadim/ArabicDialectTutorialAMTA2006.pdf)
7. ^ File:Basmala kufi.svg - Wikimedia Commons
8. ^ *a b* File:Kufi.jpg - Wikimedia Commons
9. ^ File:Qur'an folio 11th century kufic.jpg - Wikimedia Commons
10. ^ Arabic and the Art of Printing – A Special Section (http://www.saudiaramcoworld.com/issue/198102/arabic.and.the.art.of.printing-a.special.section.htm) , by Paul Lunde
11. ^ For more information about encoding Arabic, consult the Unicode manual available at The Unicode website (http://www.unicode.org/)
12. ^ See also Multilingual Computing with Arabic and Arabic Transliteration: Arabicizing Windows Applications to Read and Write Arabic & Solutions for the Transliteration Quagmire Faced by Arabic-Script Languages (http://www.lib.uchicago.edu/e/su/mideast/Multilingual_Computing_with_Arabic_and_Arabic_Transliteration.pdf) and A PowerPoint Tutorial (with screen shots and an English voice-over) on how to add Arabic to the Windows Operating System (http://www.lib.uchicago.edu/e/su/mideast/multilingual_computing_arabic.ppt) .
13. ^ Yamli in the News (http://www.yamli.com/press/)
14. ^ Israel 21c (http://www.israel21c.org/bin/en.jsp?enDispWho=Articles%5El1650&enPage=BlankPage&enDisplay=view&enDispWhat=object&enVersion=0&enZone=Technology&)

Arabic alphabet coloring pages for kids, find the letters and color. (http://umm4.com/?p=3897)

External links

- Learn Arabic alphabet with audio (http://www.naturalarabic.com/free_article.php?artid=10150)
- Interactive audio lesson for learning the Arabic alphabet (http://www.salaamarabic.com/lesson/plan/1490/isolated_consonants)

- Named Entity Recognition (http://onlinelibrary.wiley.com/doi/10.1002/asi.21090/abstract) – for a discussion of inconsistencies and variations of Arabic written text.
- Arabetics (http://arabetics.com) – for a discussion of consistency and uniformization of Arabic written text.
- The Arabic alphabet (writing letters) (http://www.arabic-keyboard.org/arabic/arabic-alphabet.php)
- Learn Arabic Letters and Quran Reading (http://www.essentialilm.com/read-arabic-quran.html)
- Open Source fonts for Arabic script (http://alefba.us/typography/free-arabic-persian-farsi-urdu-kurdish-fonts/)
- Online virtual keyboard to type and edit in arabic (http://www.dubaiblog.it/index.php/tastiera-araba-arabic-keyboard/) (in Italian).
- Arabic (http://www.dmoz.org/Science/Social_Sciences/Linguistics/Languages/Natural/Afro-Asiatic/Arabic/Online_Courses//) at the Open Directory Project

---

This article contains major sections of text from the very detailed article Arabic alphabet from the French Wikipedia, which has been partially translated into English. Further translation of that page, and its incorporation into the text here, are welcomed.

Retrieved from "http://en.wikipedia.org/w/index.php?title=Arabic_alphabet&oldid=537163606"
Categories: Arabic alphabets