

Arabic diacritics

Arabic alphabet	
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 ا	
Arabic script	
•	History
•	Transliteration
•	Diacritics
•	Hamza
•	Numerals
•	Numeration
•	v
•	t
•	e ^[1]

The Arabic script has numerous diacritics, including **i'jam** (إِعْجَام) (*i'jām*, consonant pointing), and **tashkil** (تَشْكِيل) (*tashkīl*, supplementary diacritics). The latter include the **ḥarakāt** (حَرَكَات) (vowel marks; singular: *ḥarakah* (حَرَكَة)).

The Arabic script is an impure abjad, where short consonants and long vowels are represented by letters but short vowels and consonant length are not generally indicated in writing. *Tashkīl* is optional to represent missing vowels and consonant length. Modern Arabic is nearly always written with consonant pointing, but occasionally unpointed texts are still seen. Early texts such as the Qur'an were initially written without pointing, and pointing was added later to determine the expected readings and interpretations.

Tashkil (marks used as phonetic guides)

The literal meaning of *tashkīl* is 'forming'. As the normal Arabic text does not provide enough information about the correct pronunciation, the main purpose of *tashkīl* (and *ḥarakāt*) is to provide a phonetic guide or a phonetic aid; i.e. show the correct pronunciation. It serves the same purpose as furigana (also called "ruby") in Japanese or pinyin in Mandarin Chinese for children who are learning to read or foreign learners.

The bulk of Arabic script is written without *ḥarakāt* (or short vowels). However, they are commonly used in some religious texts that demand strict adherence to pronunciation rules such as Qur'an (الْقُرْآن) (*al-Qur'ān*). It is not uncommon to add *ḥarakāt* to hadiths (الْحَدِيث) (*al-ḥadīth*; plural: *aḥādīth*) as well. Another use is in children's literature. *Harakat* are also used in ordinary texts when an ambiguity of pronunciation might arise. Vowelled Arabic dictionaries provide information about the correct pronunciation to both native and foreign Arabic speakers.

Short vowels can be included in cases where readers could not easily resolve word ambiguity from context alone, or simply wherever their writing might be considered aesthetically pleasing.

An example of a fully *vocalised* (vowelised or *vowelled*) Arabic from the *Qur'ān* (*Al-Fatiha* 1:1):

﴿ بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ ﴾

bism Allāh al-Raḥmān al-Raḥīm

In the Name of Allāh, the Most Gracious, the Most Merciful...

Some Arabic textbooks for foreigners now use *ḥarakāt* as a phonetic guide to make learning reading Arabic easier. The other method used in textbooks is phonetic romanisation of unvocalised texts. Fully vocalised Arabic texts (i.e. Arabic texts with *ḥarakāt*/diacritics) are sought after by learners of Arabic. Some online bilingual dictionaries also provide *ḥarakāt* as a phonetic guide similarly to English dictionaries providing transcription.

Harakat (short vowel marks)

The *ḥarakāt*, which literally means 'motions', are the short vowel marks. There is some ambiguity as to which *tashkīl* are also *ḥarakāt*; the *tanwīn*, for example, are markers for both vowels and consonants.

Fathah

ˆ
—

The **fathah** (فَتْحَة) is a small diagonal line placed *above* a letter, and represents a short /a/. The word *fathah* itself (فَتْحَة) means *opening*, and refers to the opening of the mouth when producing an /a/. Example with *dāl* (henceforth, the base consonant in the following examples): ⟨د⟩ /da/.

When a **fathah** is placed before the letter ⟨ا⟩ (*alif*), it represents a long /a:/ (as in the English word "father"). Example: ⟨ا⟩ /da:/. The *fathah* is not usually written in such cases.

Kasrah

˘

A similar diagonal line *below* a letter is called a **kasrah** (كَسْرَة) and designates a short /i/. Example: ⟨د⟩ /di/.

When a *kasrah* is placed before the letter ⟨ي⟩ (*yā*), it represents a long /i:/ (as in the English word "steed"). Example: ⟨دي⟩ /di:/. The *kasrah* is usually not written in such cases but if *yā* is pronounced as a diphthong /aj/, *fathah* should be written on the preceding consonant to avoid mispronunciation. The word *kasrah* means 'breaking'.

Dammah

◌

The **dammah** (ضَمَّة) is a small curl-like diacritic placed above a letter to represent a short /u/. Example: ⟨دو⟩ /du/.

When a **dammah** is placed before the letter ⟨و⟩ (*wāw*), it represents a long /u:/ (as in the English word "blue"). Example: ⟨دوو⟩ /du:/. The *dammah* is usually not written in such cases but if *wāw* is pronounced as a diphthong /aw/, *fathah* should be written on the preceding consonant to avoid mispronunciation.

Maddah

ˆ

The **maddah** (مَدَّة) is a tilde-like diacritic which can appear only on top of an *alif* and indicates a glottal stop /ʔ/ followed by a long /a:/.

In theory, the same sequence /ʔa:/ could also be represented by two *alifs*, as in *⟨اا⟩, where a hamza above the first *alif* represents the /ʔ/ while the second *alif* represents the /a:/. However, consecutive *alifs* are never used in the Arabic orthography. Instead, this sequence must always be written as a single *alif* with a *maddah* above it—the combination known as an *alif maddah*. Example: ⟨قُرْآن⟩ /qur'ʔa:n/.

Dagger alif

ا

The **superscript (or dagger) alif** (أَلِفٌ خَنْجَرِيَّةٌ) (*alif khanjarīyah*), is written as short vertical stroke on top of a consonant. It indicates a long /a:/ sound where *alif* is normally not written, e.g. ⟨هَادِهًا⟩ (*hādihā*) or ⟨رَحْمَانًا⟩ (*rahmān*).

The dagger *alif* occurs in only a few words, but these include some common ones; it is seldom written, however, even in fully vocalised texts. Most keyboards do not have dagger *alif*. The word Allah (الله) (*Allāh*) is usually produced automatically by entering *alif lām lām hā*. The word consists of *alif* + ligature of doubled *lām* with a *shaddah* and a *dagger alif* above *lām*.

Alif waslah

آ

The **waṣlah** (وَصَلَّة), **alif waṣlah** (ألف وصالَة) or **hamzat waṣl** (هَمْزَة وصال) looks like a small letter *ṣād* on top of an *alif* (آ) (also indicated by an *alif* (ا) without a *hamzah*). It means that the *alif* is not pronounced, e.g. (بِاسْم).

It only occurs in the beginning of words (can occur after prepositions and the definite article). It is commonly found in imperative verbs, the perfective aspect of verb stems VII to X and their verbal nouns (*maṣdar*). The *alif* of the definite article is considered a *waṣlah*.

It occurs in phrases and sentences (connected speech, not isolated/dictionary forms) 1) to replace the elided hamza whose alif-seat has assimilated to the previous vowel, e.g., في اليمن or في أليمن (alif by itself is also indicative of the connecting hamza); 2) in hamza-initial imperative forms following a vowel, especially following the conjunction و, e.g., قُمْ و اشرب الماء.

Sukun

◌ْ

The **sukūn** (سُكُون) is a circle-shaped diacritic placed above a letter. It indicates that the consonant to which it is attached is not followed by a vowel; this is a necessary symbol for writing consonant-vowel-consonant syllables, which are very common in Arabic. Example: (دَد) *dad*.

The **sukūn** may also be used to help represent a diphthong. A *fathah* followed by the letter (ي) (*yā'*) with a sukūn over it indicates the diphthong *ay* (IPA /aj/). A *fathah* followed by the letter (و) (*wāw*) with a *sukūn* indicates /aw/.

Tanwin (final postnasalized or long vowels)

◌ً ◌ٍ ◌ٌ

The three vowel diacritics may be doubled at the end of a word to indicate that the vowel is followed by the consonant *n*. These may or may not be considered *ḥarakāt*, and are known as **tanwīn** (تَنْوِين), or *nunation*. The signs indicate, from right to left, *-un*, *-in*, *-an*.

These endings are used as non-pausal grammatical indefinite case endings in literary Arabic or classical Arabic (triptotes only). In a vocalised text, they may be written even if they are not pronounced (see *pausa*). See *i'rāb* for more details. In many spoken Arabic dialects, these endings are absent. Many Arabic textbooks introduce standard Arabic without these endings. The grammatical endings may not be written in some vocalized Arabic texts, as knowledge of *i'rāb* varies from country to country, and there is a trend in simplifying the Arabic grammar.

The sign (◌ْ) is most commonly written in combination with (ا) (*alif*), (ة) (*tā' marbūṭah*) or stand-alone (ة) (*hamzah*). *Alif* should always be written (except for words ending in *tā' marbūṭah*, *hamzah* or diptotes), even if *an* is not. Grammar cases and *tanwīn* endings in indefinite triptote forms:

- *-un*: nominative case;
- *-an*: accusative case, also serves as an adverbial marker;
- *-in*: genitive case.

Shaddah (consonant gemination mark)



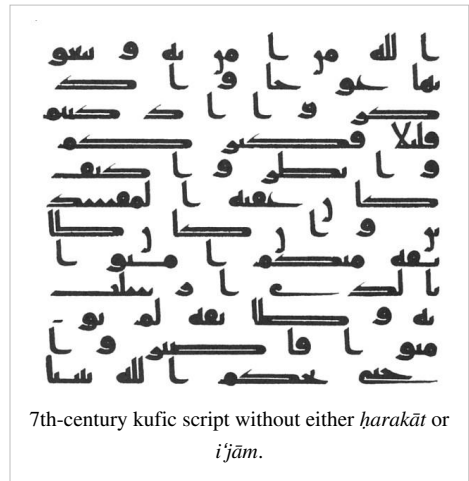
The **shadda** or **shaddah** (شَدَّة) (*shaddah*), or **tashdid** (تَشْدِيد) (*tashdīd*), is a diacritic shaped like a small written Latin "w".

It is used to indicate gemination (consonant doubling or extra length), which is phonemic in Arabic. It is written above the consonant which is to be doubled. It is the only *ḥarakah* that is sometimes used in ordinary spelling to avoid ambiguity. Example: <د> /dd/; *madrasah* (مَدْرَسَة) ('school') vs. *mudarrisah* (مُدْرِسَة) ('teacher', female).

I'jam (phonetic distinctions of consonants)

The *ijam* (إِعْجَام) (*i'jām*) are the pointing diacritics that distinguish various consonants that have the same form (*rasm*), such as <ب> /b/, <ت> /t/, <ث> /θ/, <ن> /n/, and <ي> /j/. Typically *ijam* are not considered diacritics but part of the letter.

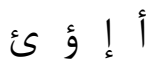
Early manuscripts of the *Qur'ān* did not use diacritics either for vowels or to distinguish the different values of the *rasm*. Vowel pointing was introduced first, as a red dot placed above, below, or beside the *rasm*, and later consonant pointing was introduced, as thin, short black single or multiple dashes placed above or below the *rasm* (image). These *i'jām* became black dots about the same time as the *ḥarakāt* became small black letters or strokes.



7th-century kufic script without either *ḥarakāt* or *i'jām*.

Typically, Egyptians do not use dots under final *yā'* (ي), both in handwriting and in print, as *alif maqṣūrah* (ألف مقصورة). This practice is also used in copies of the *muṣḥaf* (Qur'ān) scribed by 'Uthman Ṭāhā. The same unification of *yā* and *alif maqṣūrā* has happened in Persian, resulting in what the Unicode Standard calls "arabic letter farsi yeh", that looks exactly the same as *yā* in initial and medial forms, but exactly the same as *alif maqṣūrah* in final and isolated forms (ي ي ي).

Hamza (glottal stop semi-consonant)



Although often a diacritic is not considered a letter of the alphabet, the **hamza** (هَمْزَة) (*hamzah*, glottal stop), often does stand as a separate letter in writing, is written in unpointed texts, and is not considered a *tashkīl*. It may appear as a letter by itself or as a diacritic over or under an *alif*, *wāw*, or *yā*.

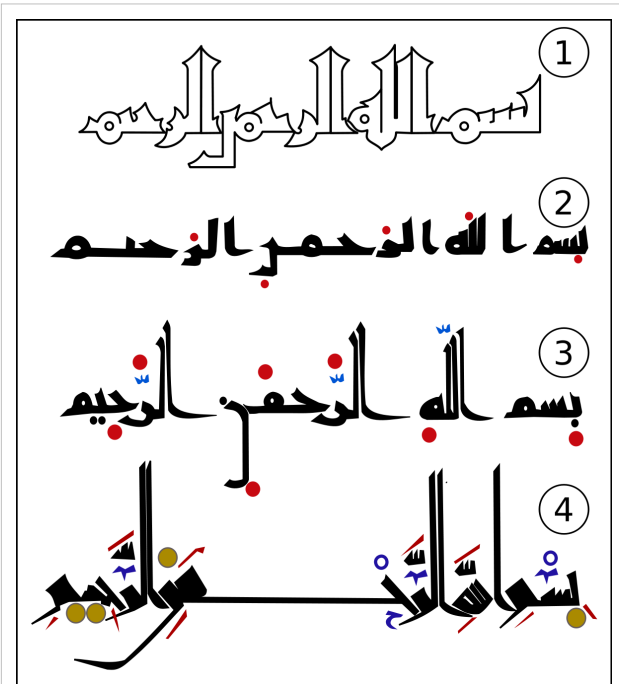
Which letter is to be used to support the *hamzah* depends on the quality of the adjacent vowels.

- If the syllable occurs at the beginning of the word, the glottal stop is always indicated by hamza on an *alif*.
- if the syllable occurs in the middle of the word, *alif* is used only if it is not preceded or followed by /i/ or /u/.
- If /i(:)/ is before or after the glottal stop, a *yā* with a *hamzah* is used (the two dots which are usually beneath the *yā'* disappear in this case): <ي>.
- If /u(:)/ is before or after the glottal stop, a *wāw* with a *hamzah* is used: <ؤ>.

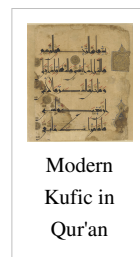
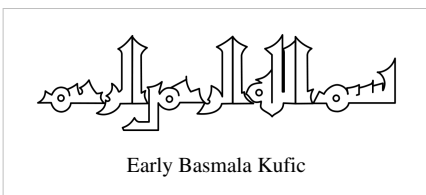
Consider the following words: <أخ> /ʔax/ ("brother"), <إِسْرَائِيل> /ʔisra:ʔi:l/ ("Israel"), <أم> /ʔumm/ ("mother"). All three of above words "begin" with a vowel opening the syllable, and in each case, *alif* is used to designate the initial glottal stop (the *actual* beginning). But if we consider *middle* syllables "beginning" with a vowel: <نشأة> /naʔa/ ("origin"), <إِسْرَائِيل> /ʔisra:ʔi:l/ ("Israel" — notice the /ʔi:l/ syllable), <رؤوس> /ruʔu:s/ ("heads", singular <رأس> /raʔs/), the situation is different, as noted above. See the comprehensive article on *hamzah* for more details.

History

According to tradition, the first to commission a system of *harakat* was Muawiyah I of the Umayyad dynasty, when he ordered Ziad Ibn Abih, his *wālī* in Basra (governed 664–673), to find someone who would devise a method to transcribe correct reading. Ziad Ibn Abih, in turn, appointed Abu al-Aswad al-Du'ali for the task. Abu al-Aswad devised a system of dots to signal the three short vowels (along with their respective allophones) of Arabic. This system of dots predates the *i'jām*, dots used to distinguish between different consonants.



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



Abu al-Aswad's system

Abu al-Aswad's system of Harakat was different from the system we know today. The system used red dots with each arrangement or position indicating a different short vowel.

A dot above a letter indicated the vowel *a*, a dot below indicated the vowel *i*, a dot on the side of a letter stood for the vowel *u*, and two dots stood for the *tanwīn*.

However, the early manuscripts of the Qur'an did not use the vowel signs for every letter requiring them, but only for letters where they were necessary for a correct reading.

Al Farahidi's system

This is the precursor to the system we know today. *al-Farāhīdī* found that the task of writing using two different colours was tedious and impractical. Another complication was that the *i'jām* had been introduced by then, which, while they were short strokes rather than the round dots seen today, meant that without a color distinction the two could become confused.

Accordingly he replaced the *ḥarakāt* with small superscript letters: small alif, *yā'*, and *wāw* for the short vowels corresponding to the long vowels written with those letters, a small *s(h)īn* for *shaddah* (geminate), a small *khā'* for *khafīf* (short consonant; no longer used). His system is essentially the one we know today.^[2]

References

[1] http://en.wikipedia.org/w/index.php?title=Template:Arabic-script_sidebar&action=edit

[2] Versteegh, 1997. *The Arabic language*. p 56ff.

External links

- Online Arabic Diacritic Tool (<http://www.arabic-keyboard.org/tashkeel/>)
- Interactive lesson for learning Arabic short vowels (<http://www.salaamarabic.com/lesson/plan/1495/vowels>)
- Free Comprehensive Reference of Arabic Grammar (<http://arabic.tripod.com/>)
- Basic Introduction To Arabic Short Vowels (<http://arabicgenie.com/blog/2009/09/arabic-short-vowels-what-you-need-to-know-about-tashkeel-and-harakaat>)
- Vocalised Arabic (and other) texts online (for children) (<http://www.childrenslibrary.org/icdl/SimpleSearchCategory?ids=&langid=309&pnum=1&cnum=1&text=&lang=English>)
- Fully vocalised, transliterated and translated online Qur'ān with audio (http://transliteration.org/quran/WebSite_CD/MixNoble/Fram2E.htm)
- Sakhr Multilingual Dictionary (uses Harakat) (<http://qamoos.sakhr.com/>)
- web-based Arabic Notepad that supports full vowelization (tashkiil) (<http://www.eiktub.com>)

Article Sources and Contributors

Arabic diacritics *Source:* <https://en.wikipedia.org/w/index.php?oldid=591799586> *Contributors:* Abjklam, Academic Challenger, Adjwilli, Aeusoes1, Agari, Al-Andalusi, Aleverde, AlexanderKaras, Ankimai, Aris riyanto, Athenodorus, Atitarev, BD2412, Badagnani, Chem1, Daniel Case, David.s.kats, Dawright12, Dbachmann, DePiep, Denelson83, Deor, Djihed, Downtown dan seattle, Dzozar, Enaya g, Erutuon, FilipeS, Fitoschido, Flewis, Gangasrotagati, Gregkaye, Gyopi, Hakeem.gadi, Herbivore, Hmains, Hvn0413, Ian Pitchford, Jchthys, Jerzy, Jheald, Johanna-Hypatia, KSA13, Kaatibah, Kbdank71, Kbh3rd, Kenrick95, Khalid hassani, Kwamikagami, König Alfons der Viertelvorzwölfte, Lockesdonkey, Lothar von Richthofen, M. Adiputra, MER-C, Mahmudmasri, Mairi, MosheA, Mpatel, Mr Stephen, Nickjamil, Obersachse, Omergold, OrenBochman, OwenX, Pejmany, Per Aspera, Phil Boswell, Ptcamn, Quuxplusone, Reaper Eternal, Rjwilmsi, Sannse, Shaalank, Snowolf, Stevertigo, Strabismus, Supertouch, Syncategoremata, Tesseran, Thecurran, TomeHale, TreasuryTag, Valodzka, Vedek Dukat, Verdy p, Westonmr, Who, Xevorim, Zakuragi, Zerokitsune, خالد حسني, 98 anonymous edits

Image Sources, Licenses and Contributors

Image:Kufic Quran 7th Cent.jpg *Source:* https://en.wikipedia.org/w/index.php?title=File:Kufic_Quran_7th_Cent.jpg *License:* Public Domain *Contributors:* Estr4ng3d, FellFairy

Image:Arabic script evolution.svg *Source:* https://en.wikipedia.org/w/index.php?title=File:Arabic_script_evolution.svg *License:* Creative Commons Attribution-Share Alike *Contributors:* Abusyliman, AnonMoos, Bennylin, Hakeem.gadi, Judithcomm, YassineMrabet

Image:Basmala kufi.svg *Source:* https://en.wikipedia.org/w/index.php?title=File:Basmala_kufi.svg *License:* Creative Commons Attribution-ShareAlike 3.0 Unported *Contributors:* Baba66, Wst, 1 anonymous edits

Image:Kufi.jpg *Source:* <https://en.wikipedia.org/w/index.php?title=File:Kufi.jpg> *License:* Public Domain *Contributors:* Baba66, Degeefe

Image:Qur'an folio 11th century kufic.jpg *Source:* https://en.wikipedia.org/w/index.php?title=File:Qur'an_folio_11th_century_kufic.jpg *License:* Public Domain *Contributors:* Baba66, Grenavitar, Obayd, Rlevse, خالد حسني, 1 anonymous edits

License

Creative Commons Attribution-Share Alike 3.0
[//creativecommons.org/licenses/by-sa/3.0/](https://creativecommons.org/licenses/by-sa/3.0/)