

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/20022> holds various files of this Leiden University dissertation.

Author: Chahine, Nadine

Title: Reading Arabic : legibility studies for the Arabic script

Date: 2012-10-25

The Arabic Script

The Arabic script belongs to the group of Semitic scripts, which are descendents of the Aramaic offspring of the Phoenician alphabet. The Arabic script, like the Aramaic and Phoenician alphabets, is represented mainly by consonants in a system usually referred to as abjad.

The Arabic script was first used to write the Arabic language. With the spread of Muslim conquests, it grew to encompass many Semitic and non-Semitic languages such as New Persian, Turkish, several Berber languages in North Africa; Pashto, Kurdish, and Baluchi in Iran; Urdu, Sindhi and Kashmiri; Dravidian Moplah; Sulu, Malagasy and Malay as well as Swahili, Kanuri, Hausa, and Fulani in central Africa, and Harari in Ethiopia. As such, the Arabic script today comes right after Latin as the most “widespread segmental script in the world” (Kaye, 2006, p. 133). It is estimated that there are around 1 billion users of the Arabic script worldwide (Amara & Bouslama, 2003).

Characteristics

Like other Semitic scripts such as Hebrew and Syriac, Arabic is written from right to left and with the lines assembled from top to bottom. The numerals, usually referred to as Hindi, are written from left to right (Fig. 1.1).

Many Arabic words are written in a continuous pen stroke without lifting the pen, with the dots and marks added later. This means that the letters are connected to each other, and each letter can occur in four different possible scenarios depending on its position in the word: initial, middle, final, or isolated position. These positions usually affect the form that the each letter can take. For example, the letter Heh has four very distinct forms in its initial, middle, final, or isolated instances (Fig. 1.2a). Other letters usually retain a certain level of similarity across different positions where the initial form is similar to the middle, and the final is similar to the isolated one (Fig. 1.2b). The majority of Arabic calligraphic styles have variations for each one of the positions that a letter can fall in depending on the letters that come before and after it. These are sometimes expressed typographically via ligatures.

The classification of the contextual alternates into these four positions is a typographic interpretation of the Arabic script. Most typefaces will have only one shape per each of the positions. However, the forms of characters as evident in manuscripts are much more dynamic and variant than in the typographic norms that have been established till today. For example, the letter Jim can have a multitude of different

Fig. 1.1 Arabic is written from right to left, and the numerals are written from left to right.



Fig. 1.2 Some Arabic letters change their basic form significantly (a) while others do not (b).

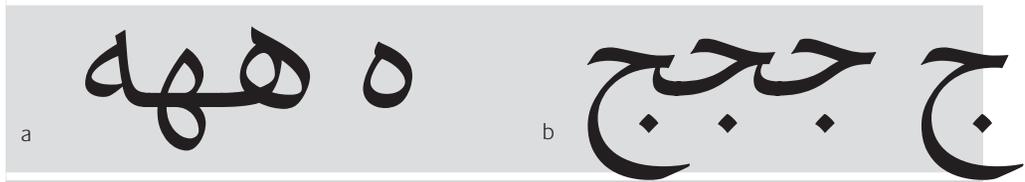


Fig. 1.3 Arabic short vowels, which are usually added above or below letters, are often not shown.



forms depending on its neighbouring characters (Fig. 1.4), but most typefaces will show only four different forms and perhaps a handful of ligatures. This is why the majority of typefaces available today look more stiff and repetitive in comparison to the manuscript calligraphic styles from which they have been derived.

Still, this division is a helpful way to analyze the script as it establishes a framework that can be later expanded as necessary. In some cases such as the styles referred to as Kufi, there are just those four forms to speak of in the first place. The transition of Arabic letterforms from calligraphic origins to typographic representations will be expanded on in the next chapter, when the anatomy of letterforms is discussed.

Of the basic twenty-eight letters, six do not join to the ones on their left. These letters only have two basic forms, a final and an isolated form (Fig. 1.5). As such, the Arabic script is partially connected and its words are characterized by an alternating rhythm of connected strings of letters where the break in connectivity can occur anywhere in the string. Though early examples of Arabic manuscripts showed an equal distance between strings within and outside of words, as well as the breaking of words over two lines, this practice fell out of common use very quickly and the typographic norm today is to cluster strings into words that are separated by a clear word space with no option for hyphenation.

Basic Elements

The claim that the Arabic script is made up of twenty-eight letters is indeed deceiving. In reality, there are several more letters that make up the repertoire needed to accurately represent the Arabic language as a start. Conspicuously missing from the count of twenty-eight is the Alif Maksura **آ**, which is the un-dotted version of Yeh **ي**. Another absentee is the Teh Marbuta **ة** which is written like the Heh **ه** but with two dots on top. Several other characters are also needed such as the isolated Hamza **ء** which denotes a glottal stop, as well as the different variations of the way it can be written in a word. When coupled with other base characters, the Hamza is treated as a combining mark. Other combining marks include the Madda **ا** and Hamzat al-



Fig. 1.4 A 16th century Thuluth and Naskh calligraphy practice sheet showing many variations per each letter position of the Beh and Jim depending on what comes before and after. Nasser D. Khalili Collection: CAL67

28

-  Dot(s) used to create new letters
-  Letters using same basic shape
-  Letters only connecting on the right
-  Letters that occur at word end only
-  Hamza Mark combinations

Fig 1.5 This is where the counting stops with 28 letters. The remaining letters are however needed for the proper representation of the Arabic language.

Final forms	Medial forms	Initial forms	Isolated forms		
	ا			ا	Alif
●	ب	ب	ب	ب	Beh
●	ت	ت	ت	ت	Teh
●	ث	ث	ث	ث	Theh
●	ج	ج	ج	ج	Jim
	ح	ح	ح	ح	Hah
●	خ	خ	خ	خ	Khah
	د			د	Dal
●	ذ			ذ	Thal
	ر			ر	Reh
●	ز			ز	Zeh
	س	س	س	س	Sin
●	ش	ش	ش	ش	Shin
	ص	ص	ص	ص	Sad
●	ض	ض	ض	ض	Dad
	ط	ط	ط	ط	Tah
●	ظ	ظ	ظ	ظ	Thah
	ع	ع	ع	ع	Ayn
●	غ	غ	غ	غ	Chayn
●	ف	ف	ف	ف	Feh
●	ق	ق	ق	ق	Qaf
	ك	ك	ك	ك	Kaf
	ل	ل	ل	ل	Lam
	م	م	م	م	Mim
●	ن	ن	ن	ن	Nun
	ه	ه	ه	ه	Heh
	و			و	Waw
●	ي	ي	ي	ي	Yeh
	ى			ى	Alef Maksura
●	ة			ة	Teh Marbuta
				ء	Hamza
■	آ			آ	Alif Madda
■	أ			أ	Alif Hamza Above
■	ؤ			ؤ	Waw Hamza
■	إ			إ	Alif Hamza Below
■	ئ	ئ	ئ	ئ	Alif Maksura Hamza

wasl $\dot{\text{ا}}$. All of these letters and marks are necessary and play an important linguistic role and so it is more precise to speak of thirty-one letters in total (Fig. 1.5).

Arabic also employs a number of combining marks that serve as short vowels. These are the Fatha $\acute{\text{ا}}$ (e as in ten), Damma $\mathring{\text{ا}}$ (u as in put), Kasra $\underset{\cdot}{\text{ا}}$ (i as in pit), and Sukoun ا° (denotes a lack of vowel). Another mark is the Shadda $\text{ا}^{\text{ˆ}}$, which is used for doubling a consonant and is normally combined with a short vowel. Other marks are the nunation marks (addition of the final nun) referred to in Arabic as Tanween and these are: Fathateyn $\text{ا}^{\text{ˆ}}$, Dammateyn $\text{ا}^{\text{ˆ}}$, and Kasrateyn $\text{ا}^{\text{ˆ}}$ (Fig. 1.3).

As mentioned at the beginning of this chapter, Arabic is a consonantal script. This means that words are represented by their consonants and the vowels are added in as marks. Three letters also double as long vowels and these are the Alif, Waw, and Yeh. The short vowels are usually dropped out in everyday literature but are normally included in children or beginner books and in poetry and religious texts. Proficient readers are able to read un-vocalised text without any difficulty although the vowels are in certain cases indicated so to avoid confusion.

The earliest Arab manuscripts often did not include any vocalization. Indeed, during the early days in Islamic history, it was considered impolite to vocalize letters as this was a sign of lack of respect to the addressee who was deemed to be lacking in linguistic prowess and evidently in need of help with reading. It is also said that the vowel marks were first added to Arabic after the Arabs had started mixing with other nations and their knowledge of how to properly vocalize text was affected (Yacoub, 1986). The early vowels were written as coloured dots (often in red) above or below the characters. This is not to be confused with the dots as known today since early Arabic manuscripts at that period were also un-dotted.

The presence and abundance of dots in the Arabic script is one of its most striking features. Of the basic thirty-one letters, sixteen are dotted (Fig. 1.5). The dots vary in number and are positioned either above or below the main body of the character. Their role is to distinguish between characters that share the same basic form. This system of dotting came into further use when the Arabic script was expanded to support more languages. The extra letters needed were in most cases based on existing forms but with a different number and/or position of dots. The abundant use of dots is a testament to a major shortcoming of the Arabic script and that is that there are very few basic shapes to use in the first place. For example, the tooth (the initial or middle form seen in Beh) is used for six different letters and this is for the Arabic language only. This characteristic of the script is challenging for beginning readers who would have to memorize the different sounds that the position of the dot corresponds to. It is especially so when these letterforms start forming complex ligatures, and one has to guess which dot belongs to which character.

The Arabic script today includes hundreds of characters and extra combining marks that are used to represent the non-Arabic languages that it supports. The Unicode consortium also defines extra notations that are strictly for Quranic usage. Punctuation marks similar to those used in Latin have also been incorporated into the system with some degree of adaptation. The French guillemots are used for quotes; the question mark, comma, and semi-colon are the mirrored version of the Latin ones; the percent sign uses the Hindi zero; the period, ellipses, and exclamation mark remain the same.

Fig. 1.6 The use of Arabic Kufi lettering on ceramics: one of the many artistic applications of the Arabic script in everyday objects. Metropolitan Museum of Art collections.



Cultural Ties

The Arabs in pre-Islamic times relied on an oral tradition. Literary works such as poetry were memorised and passed on from one generation to the next. The rise and spread of Islam changed all that. The text of the Quran and the teachings of the Prophet had to be written down and saved for future generations. This heralded a new age in Arab culture where the skill to read and write became highly valued and respected. And so, the Arabic script sprang to maturity and became the vessel through which the holy words were manifested. Consequently, the art of calligraphy as a whole took on a spiritual role by association. Calligraphers were praised and supported by the rulers, and were known to even take on high posts in government such as the famed calligrapher *Abu 'Ali Muhammad ibn 'Ali*, usually referred to as *Ibn Muqla*, who served as vizier three times during the Abbassid period between 928 and 936 AD (Blair, 2006, p. 157).

Another contributing factor to the rise of calligraphy in Arab culture was the ban on figural representation in Muslim art. The role of calligraphy then expanded to encompass the writings of holy texts as well as the decoration of buildings, tapestry, metal, ceramics and many more artefacts that were used in everyday life. The inclusion of a decorative band of calligraphy on a household item or at the entrance of a building provided a blessing or *barakah* to the owners.

As such, the Arabic script is more than the sum of its characters. It has strong religious and cultural ties that have on the one hand propelled it into the highest level of art forms, and at other times have proved to be difficult to negotiate. The

questions of script reform that had risen in the middle of the twentieth century, or any that could still arise today, have to be navigated through that focus. Given the fact that the Arabic script is the one in which the Quran is written, it would be almost impossible to try to change or reform anything in it.

Why would the script need to be reformed? As mentioned earlier, the Arabic script has just a handful of vowels. These are not enough to represent the vowel sounds present in the dialects spoken today. This means that if one were to attempt to write down the dialect of a certain region in the Arab world, one would not be able to fully represent it using the Arabic script. A simple example is the way the word *bab* (door) is pronounced in the Lebanese dialect. In Modern Standard Arabic, it rhymes with the English word “bad”, but in Lebanese the Alif sound here is similar to the French *é* sound but is a bit more protracted. In the Syrian dialect, the Alif sound is more like the French *a*. This distinction between the pronunciations of the vowels is lost in writing. Given that there are numerous dialects spoken across the region, the ability to properly represent the spoken forms of Arabic is an important requirement for linguists, and possibly for the entire population.

Resulting Complexity

The implications of all these characteristics are manifold. The fact that the Arabic script is partially connected, that the vowels are usually left out, and the abundant use of dots as the distinguishing factor among letters, has resulted in a high level of complexity from both linguistic and typographic points of view. On the other hand, the combination of these characteristics has provided raw material for centuries of artistic manipulation and experimentation. The Arabic script has lent itself to a large variety of different calligraphic styles, and supported many different languages and cultures. It has crowned works of art and architecture for centuries, and has ingrained itself into the everyday life of its readers in ways that very few other art forms have. It has played a role in shaping a rich heritage and has been the vessel for great literary pieces. On the other hand, the complexity of the writing system has been often criticized by students, teachers, and linguists alike. The balancing of these two forces is a quite delicate task. For these reasons and many more, the Arabic script provides a very wide and interesting field of study.

