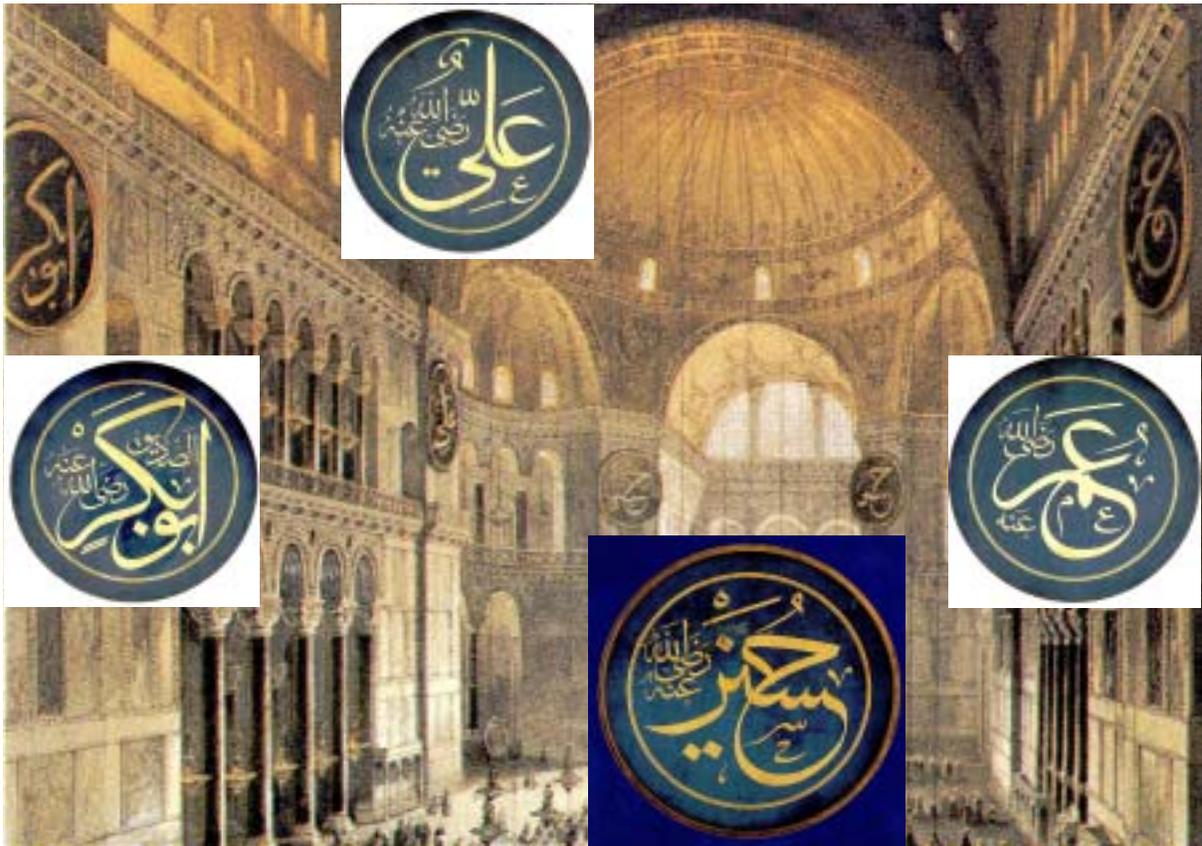


Authentic Arabic: a case study

2. TECHNICAL AND AESTHETIC CHALLENGES

Thomas Milo, DecoType, Amsterdam, the Netherlands

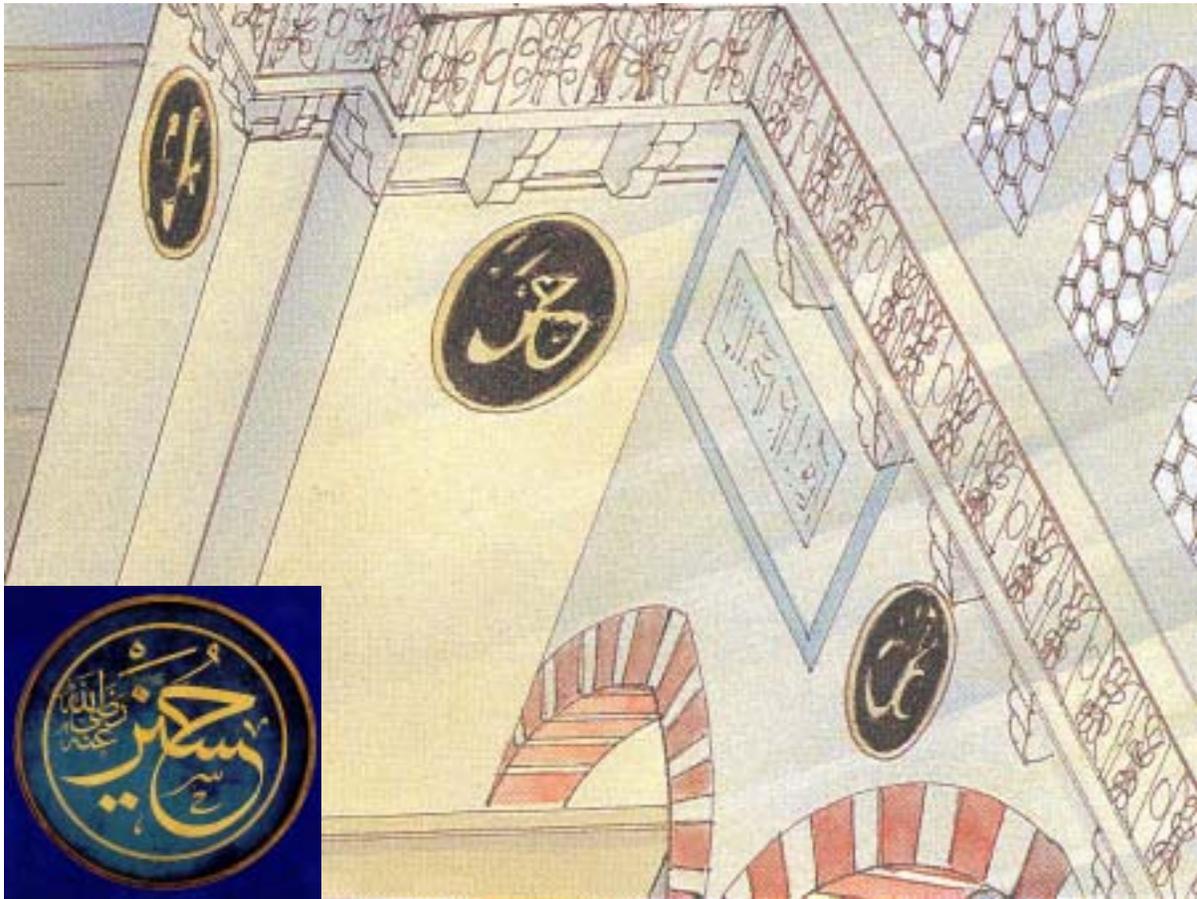
Before the invention of photography, 19th century travelers often were accompanied by artists. Their meticulous drawings reveal an interesting blind spot in these observers' minds. The famous David Roberts R.A. does not depict a single letter of Arabic¹. Others seriously try to reproduce Arabic script with varying success: this drawing of the interior of the Hagia Sophia Church, alias Aya Sofya Mosque², includes some of the large calligraphic tableaux with the names of the caliphs (visible are the names of: Ali, Umar, Husain, Hasan and Abu Bakr). The delicate beauty of the building is captured with an eye for subtle detail. None of that subtlety remains of the Arabic calligraphies. What does remain is the visual equivalent of Beethoven's *Für Elise* as played by a cell-phone. This alarming lack of perception still pervades all attempts to deal with Arabic script.



¹ See for instance David Roberts R.A., *The Holy Land, 123 coloured facsimile lithographs & The Journal from his Visit to the Holy Land*, Terra Sancta Arts Ltd, Israel, 1982 (first edition 1842)

² Caspare Fossati, *Die Hagia Sophia, nach dem tafelfwerk von 1852*, Harenberg Kommunikation, Dortmund 1980. For comparison similar, authentic tableaux are superimposed. They were taken from Nabil F. Safwat, *The Art of the Pen, Calligraphy of the 14th to 20th centuries*, Volume V of the Nasser D. Khalili Collection of Islamic Art, The Nour Foundation/Oxford University Press 1996.

A more recent illustration is a drawing of the interior of the Süleymaniye Mosque in Istanbul³. It shows little improvement in this respect. In fact it makes the earlier picture look accurate!

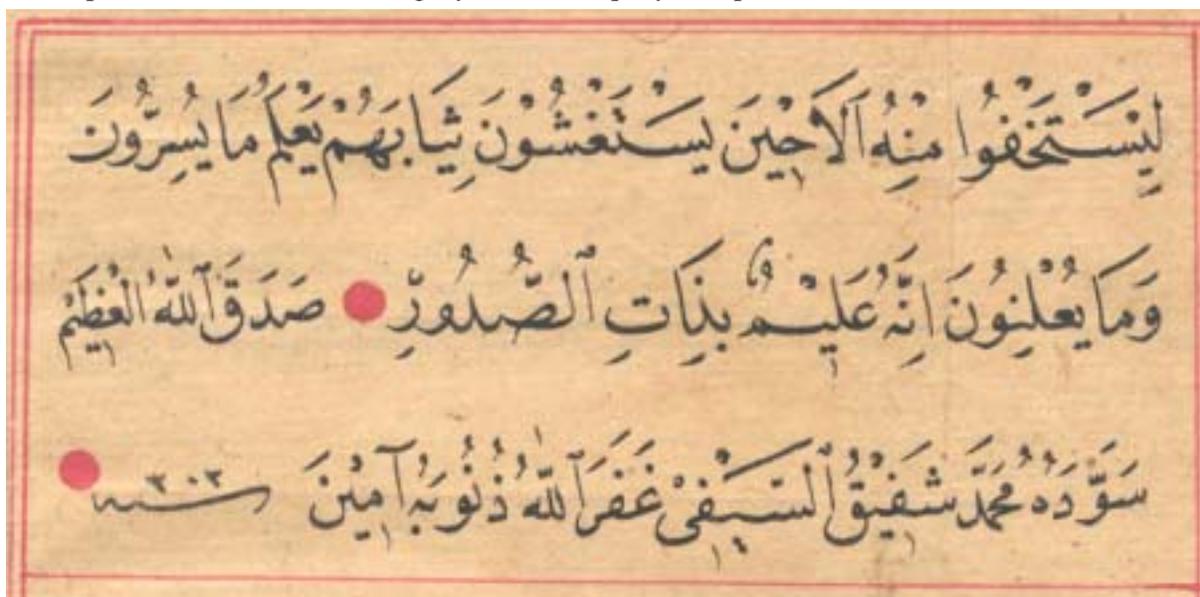


A cultural aspect

In the simple approach to Arabic script all attention goes to the *assimilation* of the letters. The four positions (beginning, middle, final and isolated) are presented as actual forms. However, the outcome is also determined by *dissimilation*. E.g., the sequence *beh-seen*  can easily be misread as *seen-beh*  as the letter *Seen* uses three strokes similar to that of the letter *beh* (or rather, of the *beh*-class). In such cases the letter *beh* gets either a dissimilar, raised stroke instead before, or a dissimilar stretched horizontal connection after *seen*. This essential reading aid and design feature is missing in all present font systems. This is just one example. The shaping of Arabic is governed by a set of rules that are both practical, in that they improve the legibility - and elegant, as they were laid down by people who rank among the world's greatest graphic artists. What sets the Arabic alphabet apart from all others is its development into an elaborate morphological system. It is the outcome of a conscious effort by Arab and Persian scholars to turn the

³ Fiona Macdonald & Mark Bergin, *A 16th Century Mosque, Inside Story*, The Salariya Book Company 1994

late Syriac-Aramaic script inherited by the Arabs into a finely balanced connected script as an expression of Islamic culture. From the sixteenth century onwards, Ottoman calligraphers developed a number of the existing styles into uniquely disciplined art forms.



Qur'an fragment with colophon identifying it as the handwriting of Mehmed Şefik (Muhammad Shafiq) Efendi, an Ottoman calligrapher in the school of Mustafa İzzet Efendi, whose handwriting was to become the model of 19th and 20th century Arabic typography.

Simplification attempts

Typesetters have been wrestling with for five centuries to handle Arabic script. By giving it a structure identical to that of Latin script all problems would be eliminated, of course. The known attempts were completely illegible, which may account for the lack of success of the designs of that nature.

Arabic Script



Understanding the structure

Writing Arabic involves more than just lining up letters. The connected letters assimilate with each other. They are highly adaptable, which makes it impractical to describe each variant individually. In Arabic script the graphic unit of writing is the letter compound: a string of connected letters. The concept of discrete or analytical letter permutations as with typewriters and in modern fonts did not exist amongst calligraphers. Traditional *mashq*, or writing exercises never show individual context variants, they are always made up of complete letter compounds. Another observation is that all letters are subject to the intricate shaping rules that balance between assimilation of distinct letters and dissimilation of featureless letters.

ن ن ن ن ن - ن ن ن ن - ن ن ن ن

Typesetting authentic Arabic

The apparent differences between authentic Naskh writing and mechanical Naskh reproductions can be attributed to a technical problem: it is very difficult to handle letter compounds in typography. Each compound forms a unit that relates to the baseline as a whole. E.g., each unit stands on a secondary baseline (in Naskh at an angle of approximately 5 degrees, in Ruq'ah slightly steeper). Between the words there are no orthographic spaces: the so-called final forms mark the words at the cutting point of the two baselines. In conventional typesetting of Arabic no such difference between main and secondary baseline is possible. The angle of the word groups is neutralized and the final forms of the word groups need to be enhanced by typographical spaces, as is the case with Latin script. The 19th century Ottoman type cutters who worked to emulate the elegance of Arabic script understood the synthetic nature of script. Far from using individual, analytic letterforms, they

designed an elaborate system of letter compound-segments to mould any occurring sequence of letters into a letter compound. These letter compound-segments resemble ligatures, but actually calling them ligatures denies the problem and its solution.

Ligatures are optional letter combinations in an otherwise analytical writing system. However, in the authentic Arabic script there is no such option.

Proper designs

At the end of the 19th there were very good Ottoman type designs with provisions to cope with the dual baseline of most Islamic scripts. There was a multilevel approach: for a given type, there were ten variants, each of them positioned on a higher baseline. These typefaces consisted of thousands of movable types and were highly complex. Because of this, they required specialized technical virtuosity based on thorough knowledge of the underlying calligraphic script. In the last quarter of the 19th previous century and the first half of the previous century this typography produced impressive results.



European developments

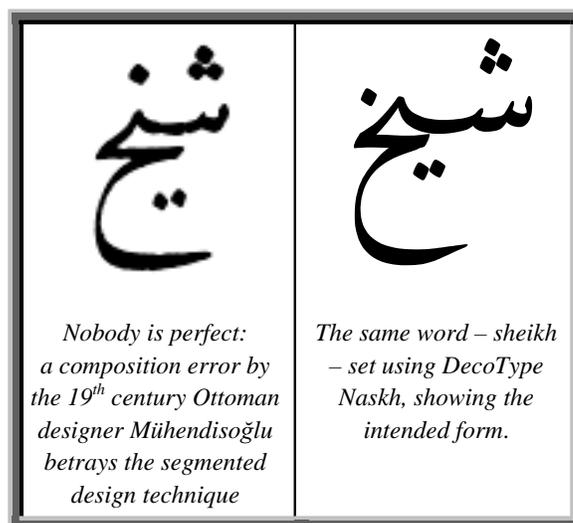
These developments took place in the second half of the nineteenth century. In Europe, typesetting with Arabic characters has been undertaken since the early 16th century. The Arabic types had a Western North-African appearance.

A possible explanation is that European type cutters might have had access to the Andalusian spoils of the Spanish Reconquista, and were consequently misled in their calligraphic styling. Anyhow, their designs were totally out of touch with Islamic taste at large and Middle Eastern taste in particular.

If these early attempts had any effect on the development of Islamic printing at all, it must have been negative. In the period that Middle Eastern calligraphy was reaching its zenith, European typography produced absolute monstrosities. For more than two centuries the Ottoman authorities opposed the large scale introduction of typesetting and printing of Islamic script. There are strong indications, that the low quality of the designs was a factor in delaying the acceptance of typography in the Islamic world.

Breakthrough in Istanbul

The first, short-lived, effort to print books with Arabic letters was in Istanbul by İbrahim Müteferrika, a Hungarian renegade in 1727. In the last decade of the 18th century, typesetting in the Ottoman Empire was taken up again and on a larger scale, culminating in the designs by Ohanis Mühendisoglu in the second half of the 19th century. These laid the basis for all modern Naskh typefaces. It took a scholar to handle the set of over 1500 movable types to construct each letter compound correctly. With the increased use of typography one can observe that the discrepancy between hand-written and typeset Naskh increases due to mistakes or ignorance of the design.



Evolution of present-day typefaces

The typographical development of the Naskh type is mainly the result of a combination of ignorance, and of compromises to bring the script in line with Latin script.

In many Naskh-derived fonts, letter compounds are replaced by simplified sequences of permutating letters as described in the beginning of this essay. Kerning (overlap) complications were avoided by adding 'false' connectors.

ك ر ا م ة
false connector false connector

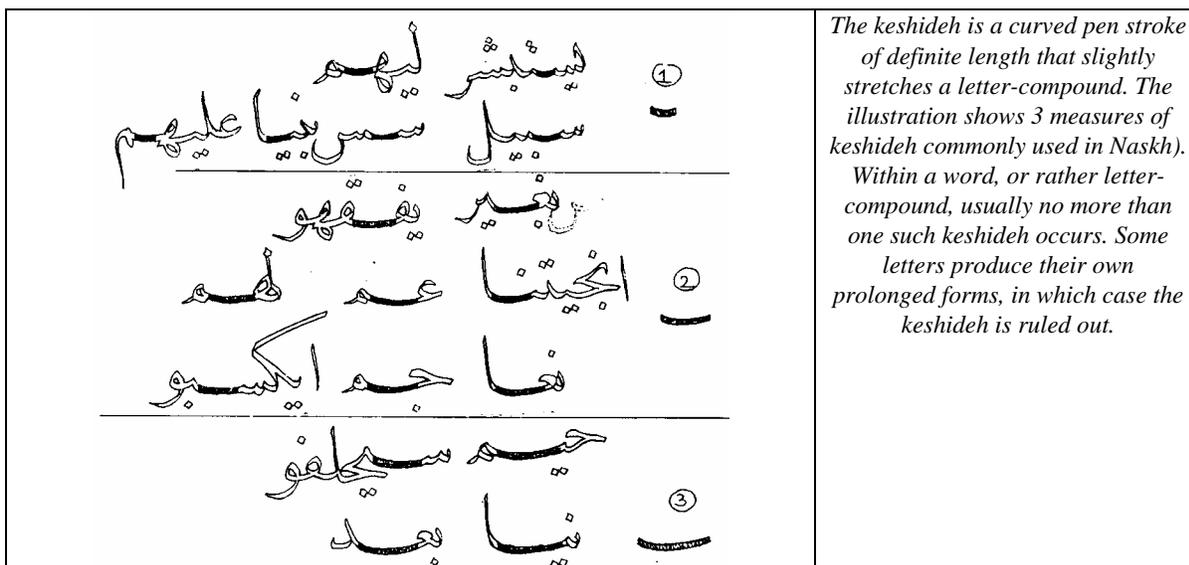
Typical example of machine-biased design strategy

ك ر ا م ة

Same word in authentic Naskh

The use of filler lines vs. calligraphic lengthening

Text block justification in Arabic involves more than in Latin script. The latter has two distinct mechanisms for justification: a *global* one, which indiscriminately inserts micro-spaces and a *specific* one to hyphenate words according to elaborate rules that vary from language to language. Islamic calligraphy has a device called the *keshideh*, a Persian and Ottoman Turkish term meaning 'stretching'. *Keshideh* is typeface-dependent, as the hyphen is language-dependent. That is, to get aesthetically acceptable results, a *keshideh* is placed according to a complex set of rules giving priority to certain letter combinations over others. These rules vary between calligraphic styles. The result is characteristically different for each kind of Arabic script. In other words, the *keshideh* is the equivalent of hyphenation and not of micro-justification.



Authentic set of long and short forms

Static versus dynamic Arabic typography

Given the fact that static, table-driven mechanisms to render Arabic authentically get swamped sooner or later, an alternative must be found. Logically, the way forward must be towards a *dynamic* solution for the problem. By careful analysis one can reconstruct the traditional rule systems that



used to create elegant compositions for any sequence of Arabic letters, with a minimum of glyphs. In this computer-aided approach, complex scripts are reduced to fonts made up of recurring shapes and a set of rules that govern the creating of compounds.

DecoType has worked along these lines for more than a decade. The first project concentrated on the Ruq'ah-script, with its characteristic sloping letter compounds. Another one covered the classic Naskh script, with its overwhelming complexity of letter compounds. Here follow some of the observations that we made.

Horizontally layered structure

Arabic script is constructed best following the pattern that evolved historically: in several horizontal layers surrounding the skeleton layer, each one adding an additional aspect of the script in a strict order. The best-known layers are that of the dot patterns and that of the vowels. In addition there is a separate layer for *shadda* or consonant enhancer.

For the more demanding Unicode requirements, dynamic shaping along layers is the only economic approach. Once the requirement of complete vowel treatment is added, a static system will become swamped.



In Urdu, the Arabic scripted language of Pakistan, clusters with larger numbers of dots occur. This can cause interesting design problems:

پریچ و خم کا پیچ

پریچ و خم کا پیچ و خم

Dynamic adjustment of dots of the word پیچ
in DecoType authentic Naskh adapted Urdu.

The problem: when rendered with DecoType Naskh, a static font
that comes with MS Office, the dots cannot be evenly spread,
because their positions are fixed.

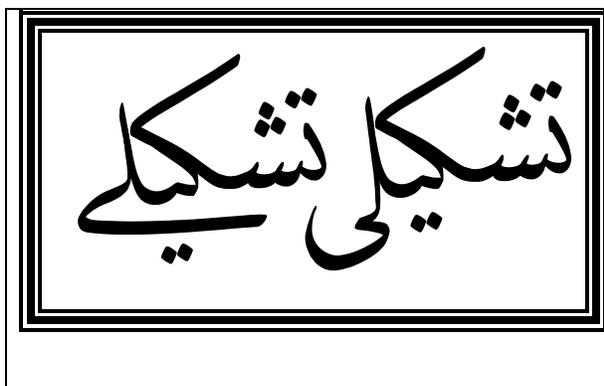
Calligraphic alternation

During the research it was also underscored, that the Islamic script is more than just a connective alphabet with contextual variation. In addition, many letters can be permuted with calligraphic alternatives, each of them havens an equivalent range of contextual variation. This phenomenon is completely different from anything known in the Latin alphabet. It enables synthesizing variant compounds out of the same abstract or analytical letter group. Calligraphic permutation opens the perspective of fine-tuning justification in typesetting with a mechanism that also occurs in manuscripts.

Two random words ضربه and سهولة, repeated seven times illustrating the use of alternation for justification text.

ضربه سهولة
ضربه سهولة
ضربه سهولة
ضربه سهولة
ضربه سهولة
ضربه سهولة
ضربه سهولة

A number of the extra letters required for extended Arabic is actually derived from calligraphic alternatives for the Arabic core alphabet. The representation forms for these letters can also be used to enrich the rendering of core Arabic. By doing so, complications that appear to be exotic are introduced to Arabic proper.



In this illustration, the availability of ے as an alternative for ی is made use of to create a calligraphic variation. The dots need to be placed dynamically to avoid clashes like this:

تشکیلی تشکیلی

To sum it all up, the future of complex scripts can be captured in fonts that are made up of recurring shapes and a set of rules that govern the creating of compounds.

Finally

In terms of encoding, Arabic is no different than any other alphabetic script, but care has to be taken to leave its graphical structure intact. The Unicode Standard is conceived for encoding raw text, not as a glyph list. Particularly attempts to fix the repertoire of Arabic letter alternations is a gross simplification and poses a long term threat to authentic reproduction of Arabic in the IT industry. Graphic representation of text remains outside the competence of Unicode proper.

The purpose of Unicode is to enable cultural diversity without imposing irrelevant constraints.