

Qurʻānic typography comes of age: Aesthetics, layering, and paragraph optimization in ConTeXt

Idris Samawi Hamid

1 The background of Oriental TeX

Attempts to integrate scripts beyond the Latin into the TeX universe are nearly as old as TeX itself. In the case of Right-to-Left (RTL) scripts as Arabic script, the inadequacy of the original TeX to the task was pointed out bluntly by Knuth himself back in 1987.¹ Since then, the heroic efforts of the ArabTeX and Omega projects took large strides in the way of extending TeX to support Arabic-script typesetting. On the other hand, by the early 2000s the realization of a paradigm capable of capturing the fullness of the Arabic script and its sophistication still seemed a long ways away. Combined with other challenges, e.g., critical-edition typesetting, so much work remained to be done.

In the winter of 2005–6, this author, along with Hans Hagen and Taco Hoekwater, initiated a very ambitious attempt to address the challenges of Arabic-script and critical-edition typesetting in the context of a radical overhaul and extension of TeX that would affect and potentially benefit virtually every corner of the TeX paradigm. This overarching context constitutes the ongoing LuaTeX project. By virtue of a major grant from Colorado State University in the spring of 2006, since extended by the generosity of US TUG, DANTE, and private donors, Oriental TeX has both served as the midwife of LuaTeX as well as having reached major mileposts in its particular goals pertaining to Arabic-script and critical-edition typesetting and typography.

2 Mileposts in the Oriental TeX project

2.1 The Qurʻān test

The most relevant torture test of Arabic-script typesetting and typography involves capturing the nuances of the Arabic used in traditional Qurʻānic script, a task that involves much by way of multiple layers of diacritics, paragraph optimizations using stretching and shape alternates, as well as multilayered coloring.

The Oriental TeX project is proud to announce that it has reached the milestone of being able to represent these aspects of Qurʻānic typography, marking a major outward milestone in the forward movement of Oriental TeX. There is still a ways to go, but in

the current visual results we can confidently say that we are “over the hump” so to speak.

2.2 Infrastructure

The visuals achieved by Oriental TeX build on an extensive infrastructure, involving the following mileposts, both achieved and in progress:

- Aleph + pdf ϵ -TeX + native UTF-8 + Lua = LuaTeX (done);
- OpenType + language processing in Lua \Rightarrow ConTeXt MkIV (beta)
- Development of Husayni font family (flagship font nearly complete);
- Paragraph optimization and justification model (initial implementation);
- Bidirectional model (under development — no more naive mirroring primitives but a full typographical framework for bidi);
- Structural-element/critical-apparatus control \Rightarrow CriTeXt (under development)

2.3 Documentation

In addition to the numerous papers, presentations, and other documentation by Hans Hagen, I am working on the following documentation tasks for a wider audience:

- *An Ontology of Arabic-script Typography* (in development);
- *CriTeXt: The Critical Edition Module for ConTeXt* (white paper available, unpublished)
- *Typographical Aesthetics and Engineering: Structured and Automated Authoring in ConTeXt* (in development: this will be perhaps the first book on ConTeXt for a general audience)

3 Qurʻānic typography

Qurʻānic typography involves getting the following tasks done.

3.1 Control of aesthetics

Given an Arabic string, there is often more than one way to aesthetically represent it. Which way one chooses depends on the context. This goes beyond the mere availability of glyph alternates as in Latin. Put another way, just as one must choose the right font for a given typographical task in Latin script, so also must one choose the right *set of aesthetic features* for a given Arabic-script task. This principle can be applied to Latin, especially in the form of calligraphy fonts, but it’s nowhere near as important an issue.

Available Arabic fonts generally mimic Latin fonts in that they have little-to-no flexibility in this regard. That is, given a font, its aesthetic style is

¹ See p. 157 of Knuth’s *Digital Typography*, CSLI Publications, Stanford, 1987.

generally fixed. OpenType allows for far more flexibility of aesthetic sets within Arabic-script, providing more culturally authentic possibilities.

In ConTeXt MkIV, we have something called *fontfeatures*. Using fontfeatures, we can define, add, and subtract *sets* of OpenType lookups — called *features* to create myriads of output possibilities.

In the case of Qurʾānic typography, we will illustrate a default feature set, and see what happens when we

- subtract and add features,
- stretch glyphs using alternates.

3.2 Layering

Arabic script is mostly consonantal. The vowels are generally not letters.² And the consonants are characterized by ambiguity: given a letter shape, it may represent two or more actual letters. One disambiguates by means of *identity marks*, mostly in the form of dots.

In historical texts, including the Qurʾān, sometimes only the consonantal layer is represented in the text, sometimes only either the vowel or the identity-marks layer. Sometimes the two layers clash and we have to adjust the shape of the consonant or the mutual positions of characters belonging to the two diacritics layers.

3.3 Paragraph optimization

Given that Arabic script in general does not accept hyphenation, getting even color in a paragraph will use alternate glyph substitution, or stretched glyph substitution. The first involves an entirely different shape, e.g.,

ك versus ك

whereas the latter involves a stretch of the existing character, e.g.,

الحمد versus الحمد

One can also combine this with changes in aesthetics, e.g.,

الحمد versus الحمد

Getting all of this under control is still in the experimental stages of implementation. The plethora of possibilities for optimization, plus the good job that the current paragraph builder does, leads us to focus on a line-by-line optimization after the initial paragraph is optimized by TeX. We illustrate this by showing some real-life samples from the Qurʾān: from Sūrah’s Fāṭiḥah and Baqarah.

² There are three consonants that sometimes function as vowels as well, analogous to the letter ‘y’ in English.

4 Aesthetics and layering

4.1 Fontfeatures: default

Given a font, it should have a default aesthetic behavior. To that end, we define a default fontfeature set:

```
\definefontfeature[husayni-default]
[analyze=yes, mode=node, language=dflt,
script=arab, ccmp=yes, init=yes, medi=yes,
fina=yes, rlig=yes, calt=yes, salt=yes,
anum=yes, ss01=yes, ..., ss60=yes, ...,
js16=yes, kern=yes, curs=yes, mark=yes,
mkmk=yes, tlig=yes,
colorscheme=husayni:default]
```

For illustration we have shown excerpts from the list of dozens of possible features. Features to be disabled can be commented out, or ‘yes’ changed to ‘no’.

Also note the `colorscheme` key at the end. This refers to a mechanism for the coloring of the various layers of the text as we mentioned earlier.

4.2 Subtracting aesthetics

Let’s get a barebones, minimalist Arabic-script implementation by subtracting features. Thus the key command here is `\subff`:

```
\switchtobodyfont[husayni-default,40pt]
\definefontfeature [first_order]
[script=arab,dlig=yes,ss01=yes,
ss03=yes,ss07=yes,ss10=yes,ss12=yes,
ss15=yes,ss16=yes,ss19=yes,ss24=yes,
ss25=yes,ss26=yes,ss27=yes,ss31=yes,
ss34=yes,ss35=yes,ss36=yes,ss37=yes,
ss38=yes,ss41=yes,ss42=yes,ss43=yes]
\subff{first_order}
```

Our ConTeXt input text, typed in Arabic (omitted in other examples):

العالمين العالمين \crlf

مُتَهَجِّدٌ سَيَنْفَجِرُ لُؤَاتِيخ

And the result:

العالمين الْعَالَمِينَ
مُتَهَجِّدٌ سَيَنْفَجِرُ لُؤَاتِيخ

4.3 Default aesthetics

Let’s show the default aesthetics:

```
\switchtobodyfont[husayni-default,40pt]
```

The result:

Study the differences between the first two words once we introduce vowels: one of them has a mild stretch to accommodate a lower vowel.

4.4 Advanced aesthetics

For more fancy text, let's add some features. Note the stacking of characters, giving in some cases a slanted feel to the text. The key command here is `\adfff`:

```
\switchtobodyfont [husayni-default,40pt]
\definefontfeature [stack:haa:multi-level]
[script=arab,ss05=yes,ss06=yes,ss09=yes,
ss13=yes,ss17=yes,ss40=yes]
\adfff{stack:haa:multi-level}
```

4.5 Stretched and alternate aesthetics

Again, we are adding, this time adding stretched-alternate and glyph-alternate features:

```
\switchtobodyfont [husayni-default,40pt]
\definefontfeature [maximal_stretching]
[script=arab,ss05=yes,ss09=yes,ss06=yes,
ss13=yes,ss17=yes,ss40=yes,js13=yes,js14=yes,
```

```
js16=yes,js05=yes]
\adfff{maximal_stretching}
```

This example in part illustrates that `TEX` is informed about where it is legal to stretch and where it is illegal. This word, third from the right in the above example, shows three stretches; in real life we would use only one of those three for any given stretched instantiation of the word.

4.6 The base consonant layer

Here we get the base consonants only:

```
\switchtobodyfont [husayni-default,40pt]
\definefontfeature [consonant]
[script=arab,ss61=yes,ss49=yes,ss52=yes]
\adfff{consonant}
```

4.7 The identity marks layer

Let's disambiguate the characters, and use colors to illustrate the difference:

```
\switchtobodyfont [husayni-default,40pt]
\setfontcolorscheme [1]
\definefontfeature [identity]
[script=arab,ss49=yes,ss52=yes]
\adfff{identity}
```

4.8 The vowels layer

In the history of the Qur^ānic text, the vowels were developed before the identity marks, so vowels without identity marks is not an abstract example:

```
\switchtobodyfont [husayni-default,40pt]
\setfontcolorscheme[1]
\definefontfeature [vowel]
[script=arab,ss61=yes]
\addff{vowel}
```

الْعَلَمِينَ مَهْجِدٌ سَدَفِرُ
لَوَاسِحُ

4.9 Full layering

Let's put it all together, with full layering and color distinctions:

```
\switchtobodyfont [husayni-default,40pt]
\setfontcolorscheme[1]
```

الْعَلَمِينَ مَهْجِدٌ سَدَفِرُ
لَوَاسِحُ

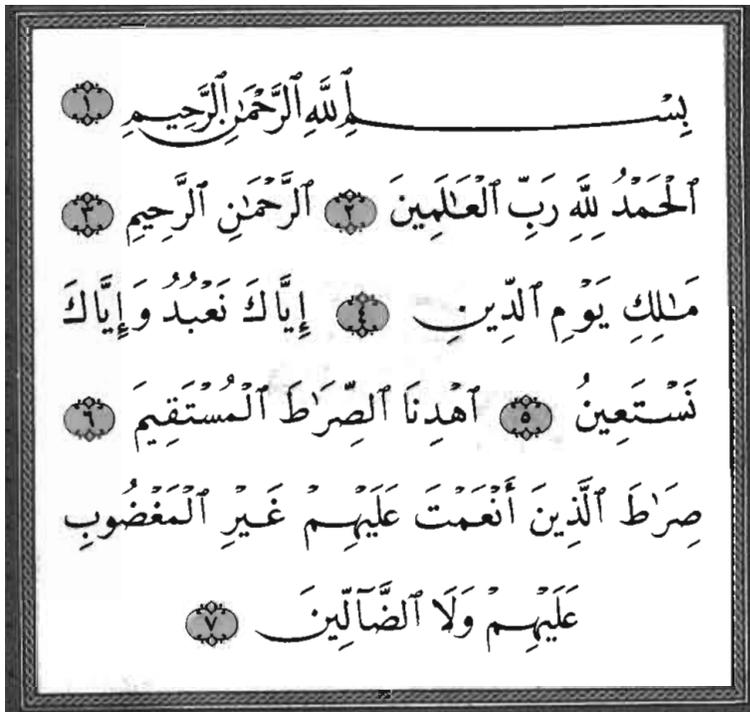


Figure 1: Sūrah Fātiḥah, from the 1924 Egyptian Edition of the Qurʾān.

5 Optimization and sample Qurʾānic typography

5.1 Sūrah Fātiḥah

Let's look at the first page of the handwritten frontispiece of the standard Egyptian edition of the Qurʾān. The entire edition is typeset, except for the two first pages that constitute the frontispiece (see Figure 1).

As simple as the handwritten version seems, it contains numerous subtleties that make imitation using traditional typesetting technology virtually impossible. Figure 2 presents it in Oriental T_EX, with full layering.

But we can be creative and imitate other styles as well. An Iranian style is shown in Figure 3, and an Indo-Pakistani style in Figure 4.

5.2 Sūrah Baqarah

From the frontispiece, we move to more straightforward paragraph typography. We illustrate this with two aesthetic sets from Sūrah Baqarah, with optimization applied to get square paragraphs that fill the page, as well as full layering (Figures 5 and 6).

6 Conclusion

As you can see, Oriental T_EX has come a long way. There is much work ahead, including further refinements to the aesthetics and the optimization routines, but it is certainly gratifying to see such sophisticated real-world results. Qurʾānic typography has finally come of age!

◊ Idris Samawi Hamid
Colorado State University
ishamid (at) colostate dot edu



Figure 2: Sūrah Fātiḥah, based on the 1924 Egyptian Edition of the Qurʾān, done in Oriental T_EX.

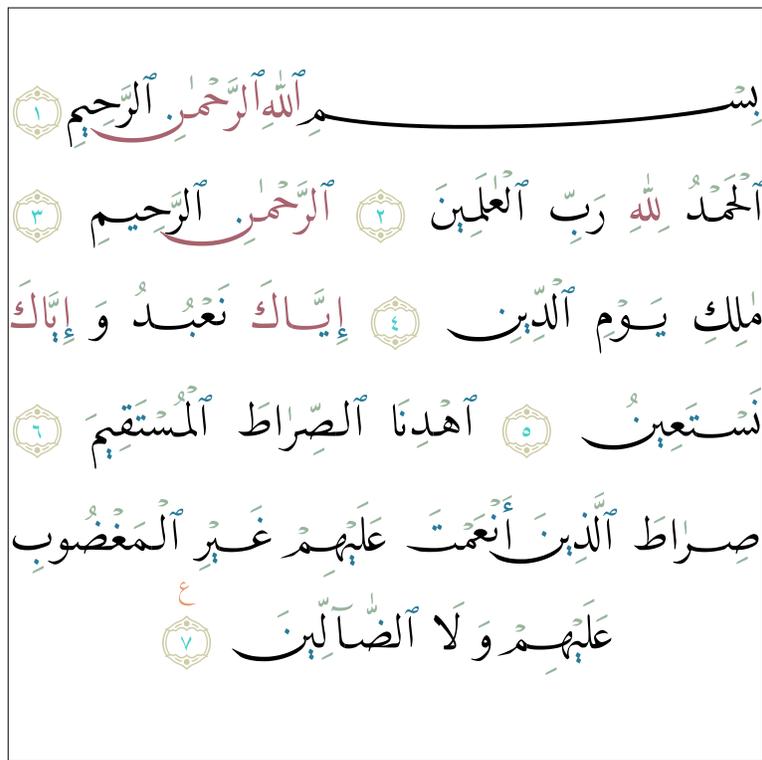


Figure 3: Sūrah Fātiḥah, in a more Iranian style.



Figure 4: Sūrah Fātiḥah, in a more Indo-Pakistani style.



Figure 5: Sūrah Baqarah, in a basic Egyptian style.



Figure 6: Sūrah Baqarah, in a slightly different vowelization style, with more character stacking.