Title: Changing the properties of U+06DE from a combining mark to a spacing symbol

Date: 2010-06-06

Introduction

To help the devout Muslims recite the whole Koran during the holy month of Ramadan, the Koran is divided into thirty parts. These parts, each called a *juz'* (جزء), are different from the traditional 114 *surah*s (سورة): they divide the Koran into roughly equal amounts of text, so the devout can spend an almost equal amount of time each day for their recitations. To further help timing, each *juz'* is further divided into two *hizb*s (حزب), and each *hizb* into four *rub'-al-hizb*s (ربع الحزب). In printed Korans, this sectioning is typically specified in the margins, with a marker inside running text to point the exact location of one *rub'-al-hizb* ending and the next starting (see Figure 1).

The Unicode character U+06DE ARABIC START OF RUB EL HIZB is used for that in-text marker. In printed Korans, the symbol appears in running text by itself, always next to an end-of-ayah marker. The rendering behavior is the same as U+06E9 ARABIC PLACE OF SAJDAH: the character just sits there and does not interact with other adjacent text.

But in Unicode, U+06DE is specified to be a combing character. We believe this to be a simple mistake, since it was encoded next to U+06DD ARABIC END OF AYAH (originally a combining character too). Looking at the reference glyph for U+06DE also confirms this: there is a dotted circle in the glyph, but there is no way something could fit in there.
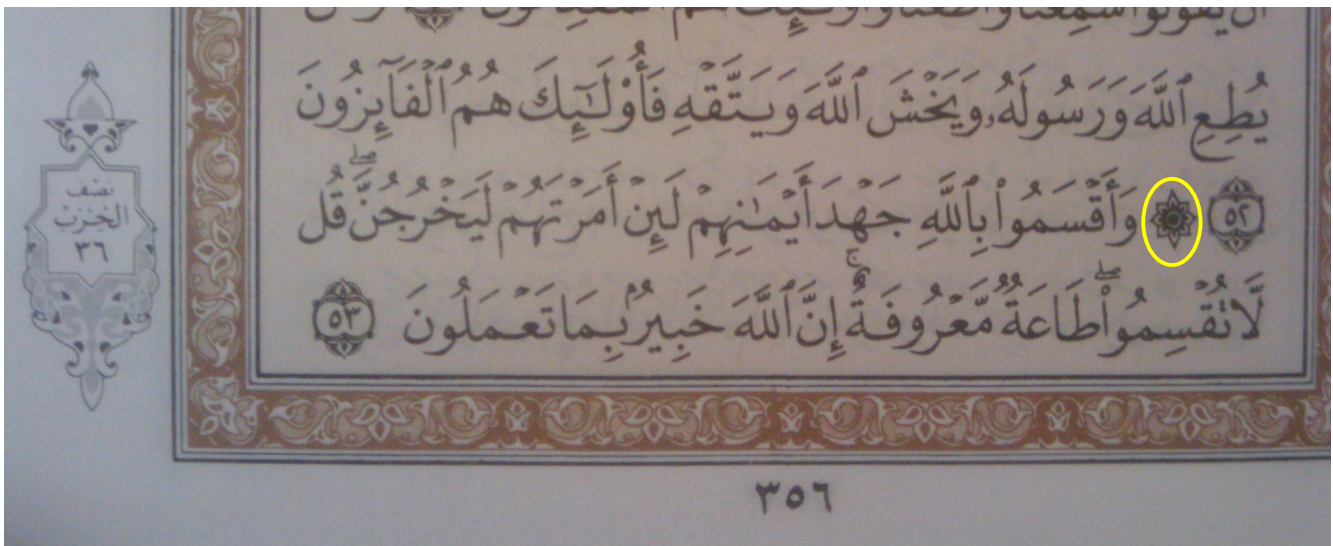

Figure 1. Sample from a Koran printed in Iran [1, page 356]. The character U+06DE ARABIC START OF RUB EL HIZB is visible at the beginning of line -2, after an End of Ayah marker. The large decoration at the left margin of the page helps readers find the inline marker easier. The text in the center of the decoration reads "half of *hizb* 36" (نصف الحزب ٣٦), indicating that the reader is at approximately 71/120 of the whole text of the Koran.

Recommendation

- Change the properties of U+06DE ARABIC START OF RUB EL HIZB from that of a combining mark to a normal spacing symbol. All character properties could be copied from U+06E9 ARABIC PLACE OF SAJDAH. (This does not create any problems for normalization, as the character already has a canonical combining class of 0.) This means a general category of "So", a bidi class of "ON", etc.

- Change the glyph for U+06DE ARABIC START OF RUB EL HIZB to remove the dotted circle.

References

1. *Qor'ān-e Karim*, calligrapher: Uthman Taha. Osve (a subsidiary of Orgranization for Awqaf and Charitable Affairs), Tehran, 1375 AP [≈1986 CE].