

Dot orientation and similarity in Arabic script

Meikal Mumin, M.A. (meikal.de/profile)

Institut für Afrikanistik, Universität zu Köln

&

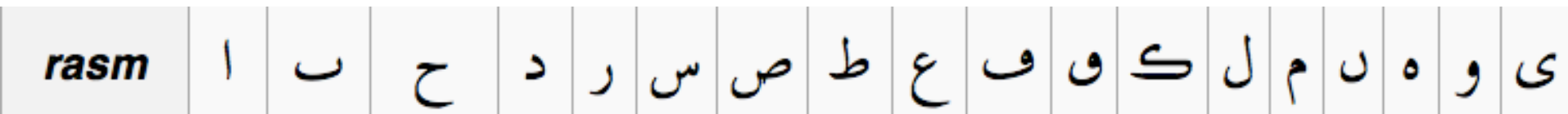
Dipartimento di Studi e Ricerche su Africa e Paesi Arabi,
Università degli Studi di Napoli L'Orientale

Types of dots encountered in Arabic script

- ▶ Two types of dots must be separated
 - ▶ ب (U + 06A2) vs. ب (U + 06A1 & U + 065C)
- ▶ Most dots are *nuqṭa* '[diacritical] dot' or *ʾiḡām* 'providing letters with diacritical dots'
 - ▶ Few dots are *taškīl* 'Phonetic aids (Lit. forming)' or *ḥarakāt* 'Vowel marks (lit. motions)'
 - ▶ In hand-writing, these *taškīl* / *ḥarakāt* are traditionally separated from *nuqṭa* / *ʾiḡām* by red color
- ▶ In Arabic script, dots are frequently used as modifiers of basic letter shapes
 - ▶ Only *nuqṭa* / *ʾiḡām* are used to form new letters, while *taškīl* / *ḥarakāt* are used to modify existing letters
 - ▶ Among *taškīl* / *ḥarakāt*, only one dot occurs regularly – the so called *Warṣ* dot <◌◌>
 - ▶ Other exceptions include *sukūn* <◌◌>, which in hand-writing could be reduced to just a dot

Arabic script as an extensible writing system: *nuqta* / ʾiʿǧām

- ▶ Originally *rasm* ‘Basic letter shapes (lit. drawing, outline, or pattern)’ were written without any (diacritical) dots



- ▶ *nuqta* were introduced from the 7th century onward (Bauer 1996:559) raising the specification of *rasm*, i.e. making letter shapes unambiguous
 - ▶ *nuqta* were introduced to facilitate standardization of the spelling of *Qurʿān*
- ▶ Effectively, this increased the number of letters from 12/15 to 28

Dots in Arabic script based orthographies

- ▶ *nuqṭa* / ʾiḡām were recognized by users of Arabic script wanting to write other languages as a prominent strategy
 - ▶ By far the most widely encountered strategy for formation of new letters
 - ▶ At least 57 out of 90 new letters encountered in Africa only are modifications of *nuqṭa* / ʾiḡām
 - ▶ In analogy with the original dotting patterns dotting was added modifying phonological parameters (Daniels 1997), e.g. 2 dots = Alveolar, 3 dots = Inter-dental
- ▶ Number of dot varies from 1-4(5)

The *warš* dot

- ▶ The *warš* dot became employed in other Arabic script writing, often much indicating /e/
- ▶ Originally technographic sign indicating the correct pronunciation (*tağwīd*) of *Qurʿānic* texts, which specifically indicates the *ʾimāla*, i.e., raising of the vowel [a] in certain contexts
 - ▶ So called after *ʿUtmān ibn Saʿīd Warš al-Qurašī al-Miṣrī* (d. 812), whose *qirāʾāt* 'reading', i.e. exegetical tradition of the *Qurʿān* became the most widespread tradition in North and West Africa
 - ▶ Used in Arabic language only in certain *Qurʿānic* manuscripts following the *Warš* tradition

Further letter formation strategies

- ▶ Apart from dotting, other strategies for formation of new letters are encountered in Arabic based script, including
 - ▶ Empty ring/*sukūn* attached to a letter ت
 - ▶ Hacek/caron ت
 - ▶ Parallel گ and crossing ل strokes
 - ▶ Superscripting of existing letters including
 - ▶ *tā'* ت
 - ▶ *ṭā'* ط
 - ▶ *mīm*
 - ▶ *nūn* ن
 - ▶ *hā'*

Orientation of dotting patterns

- ▶ Dotting in single groups ب or in two groups ف on multiple sides (of the letter shape)
 - ▶ Dotting groups left and right ب or above and below ش
 - ▶ Dotting groups can be opposite ق or not ص
- ▶ Within groups, dots can be rotated
 - ▶ Vertically ذ
 - ▶ Horizontally ز
 - ▶ Diagonally گ
- ▶ Within groups, dots can be mirrored/inverted ث or simply point upwards ژ as default rotation

Number of dots

- ▶ Dots can number from 0 (absence of or removal of dots) to 4 within groups

0	1	2	3	4
و	م	ن	ه	و

- ▶ And can number up to five across two groups

1 + 1	1 + 2	2 + 1	3 + 1	3 + 2	2 + 3
ف	ب	ق	ف	ث	ضي

Summary of dot orientation and similarity in Arabic script

- ▶ Wide variety of dotting patterns are encountered in Arabic based script
- ▶ While graphic similarity might be an issue for the non-user, there is currently no indication that dotting patterns are confused by local users
- ▶ In an IT context, confusability would be highly dependent on
 - ▶ the linguistic background/aptitude with Arabic script of the user
 - ▶ the fonts employed by the clients, e.g. the browser
- ▶ The Arabic Case Study Team of the Internationalized Domain Names Variant Issues Project, had prepared an Issues Report containing an appendix of codepoints, for which they foresaw problems of confusability due to similarity in dotting patterns

Appendix B.2. Confusable Similar Shape with Difference in Dot Orientation

Appendix B.2. Confusable Similar Shape with Difference in **Dot Orientation**

Unicode	Characters
i) U+062A	i) ﺃ
ii) U+067A	ii) ﺃ
i) U+062B	i) ﺃ
ii) U+067D	ii) ﺃ
i) U+063C	i) ﻙ
ii) U+0764	ii) ﻙ
i) U+064A	i) ﻱ
ii) U+06D0	ii) ﻱ
iii) U+067B	iii) ﺏ
i) U+067E	i) ﺏ
ii) U+0752	ii) ﺏ
i) U+0683	i) ﺎ
ii) U+0684	ii) ﺎ

i) U+0686	i) ﺎ
ii) U+0758	ii) ﺎ
i) U+068E	i) ﺎ
ii) U+068F	ii) ﺎ
i) U+06A0	i) ﺎ
ii) U+075F	ii) ﺎ
i) U+06B2	i) ﺎ
ii) U+06B3	ii) ﺎ
i) U+075D	i) ﺎ
ii) U+075F	ii) ﺎ
i) U+0697	i) ﺯ
ii) U+076B	ii) ﺯ

Thank you for your
kind attention

For more information on
TASIA, kindly consult
tasia.org

References

- ▶ Daniels, Peter T. (1990): Fundamentals of Grammatology. In *Journal of the American Oriental Society* 110 (4), pp. 727–731.
- ▶ Bauer, Thomas (1996): “Arabic Writing”. In Peter T. Daniels, William Bright (Eds.): *The World's Writing Systems*. New York, Oxford: Oxford University Press, pp. 559–564.