ISO/IEC JTC 1/SC 2/WG 2

PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS

FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646[1]

Please fill Sections A, B and C below. Section D will be filled by SC 2/WG 2.

A. Administrative

1. Title: Arabic Letter Final/Isolated Kaf Sign

2. Requester's names: Vladimir Ivanov, Irina Novgorodova

3. Requester type (Member body/Liaison/Individual contribution): Individual Contribution

4. Submission date: 2001-08-01

5. Requester's reference (if applicable):

6. (Choose one of the following :)

        This is a complete proposal: Yes;

        More information will be provided later: No

B. Technical - General

1. (Choose one of the following:)

a.   This proposal is for a new script (set of characters):   No

     Proposed name of script:

b.   The proposal is for addition of character(s) to an existing block: Yes

     Name of the existing block: Arabic

2. Number of characters in proposal: 1

3. Proposed category (see section II, Character Categories):   A

4. Proposed Level of Implementation (see clause 15, ISO/IEC 10646-1):

     Is a rationale provided for the choice?

        If Yes, reference:

5. Is a repertoire including character names provided?:   Yes

     a. If YES, are the names in accordance with the 'character naming guidelines'

       in Annex K of ISO/IEC 10646-1?

     b. Are the character shapes attached in a reviewable form?   Yes

6. Who will provide the appropriate computerized font (ordered preference: True Type,

     PostScript or 96x96 bit-mapped format) for publishing the standard?

                                    True Type, Vladimir Ivanov

     If available now, identify source(s) for the font (include address, e-mail,

     ftp-site, etc.) and indicate the tools used:

        11 Mokhovaya, ISAA pri MGU, Moscow, 103911, Russia, iranorus@online.ru

---

[1] (Form number: N-1116-F 1994-10-14; Revised 1995-01-27, 1995-04-05, 1996-04-25, 1996-08-19)

Tools: Fontographer, VOLT

7. References:

a. Are references (to other character sets, dictionaries, descriptive texts etc.)
    provided?   Yes (see References)

b. Are published examples (such as samples from newspapers, magazines, or
    other sources) of use of proposed characters attached?   Yes (see Description)

8. Special encoding issues:

Does the proposal address other aspects of character data processing (if applicable) such as input,
presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information):
                                              Yes (see Description)

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before?   No
    If YES explain_____

2. Has contact been made to members of the user community (for example: National
    Body, user groups of the script or characters, other experts, etc.)?   Yes
        If YES, with whom?_____ see Description
        If YES, available relevant documents? _____

3. Information on the user community for the proposed characters (for example: size,
    demographics, information technology use, or publishing use) is included?   Yes
    Reference: _____MS Encarta Interactive World Atlas 2000

4. The context of use for the proposed characters (type of use; common or rare)   Common
    Reference: _____ see Description

5. Are the proposed characters in current use by the user community?   Yes
    If YES, where? Reference: _____Kabul (Afghanistan) (see Description)

6. After giving due considerations to the principles in N 1352 must the proposed
    characters be entirely in the BMP?   Yes, U+06EE
        If YES, is a rationale provided?   Yes
            If YES, reference: _____ see Description

7. Should the proposed characters be kept together in a contiguous range (rather than
    being scattered)?   N/A

8. Can any of the proposed characters be considered a presentation form of an existing
    character or character sequence?   No
        If YES, is a rationale for its inclusion provided?   _____
            If YES, reference: _____

9. Can any of the proposed character(s) be considered to be similar (in appearance
    or function) to an existing character?   Yes
        If YES, is a rationale for its inclusion provided?   Yes
            If YES, reference: _____ see Description

10. Does the proposal include use of combining characters and/or use of composite

sequences (see clause 4.11 and 4.13 in ISO/IEC 10646-1)?	Yes
If YES, is a rationale for such use provided?	Yes
If YES, reference:	see Description
Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?	Yes
If YES, reference:	see Description
11. Does the proposal contain characters with any special properties such as control function or similar semantics?	No
If YES, describe in detail (include attachment if necessary)	

D. SC 2/WG 2 Administrative (To be completed by SC 2/WG 2)

1. Relevant SC 2/WG 2 document numbers:

2. Status (list of meeting number and corresponding action or disposition):

3. Additional contact to user communities, liaison organizations etc:

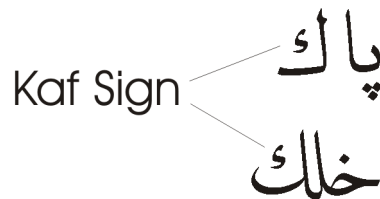4. Assigned category and assigned priority/time frame:

Description

(To read Pashto examples below demo font "Paxto Normal" must be installed. This demo font and electronic version of this Proposal can be found at iranorus@online.ru)

Pashto is the native language of Pashtuns or Afghans in Afghanistan and Pakistan. The lower limit of their population can be estimated as 23 millions, though much higher figures are suggested. All native speakers of Pashto can be divided into 3 main dialects; 2 of them are spread over the territory of Afghanistan: Western (centered in Kandaghar), Eastern (with center in Kabul, The Capital of the country). Those two dialect groups use the same orthographic rules. Another dialect group is outside of Afghanistan (in Pakistan). The latter uses different orthographic rules. The population of Afghanistan is estimated to be as large as 24,792,375; and the population of Kabul — 700,000 (MS Encarta Interactive World Atlas 2000).

The Alphabet of Pashto is based on Persian alphabet (32 letters) and consists of 42 main letters. Number of letters varies from source to source.

This Proposal is dedicated to Arabic Letter Final/Isolated Kaf Sign (thereafter shortened to Kaf Sign) — a graphic element put in the middle of Arabic Letter Kaf, when it appears in its final and/or isolated position (see Picture 1).

Picture 1 Final/Isolated Kaf Sign

This element can be seen in Unicode characters U+0643, U+06AC, U+06AD, U+06AE, U+FBD3, U+FBD4 etc. We could not find Kaf Sign as a standalone character neither in Unicode charts, nor in the list of proposals.

Kaf sign is used in Pashto to form letter Dze. If we use the IPA transcription, its name is pronounced [ʒē] in both dialects of Afghanistan. It is pronounced [zē] and sometimes [zīm] in the dialects of Pakistan, because affricate [ʒ] is simplified to [z] in this region. Four glyphs of this letter (from right to left: initial, medial, final and isolated) are shown on Picture 2.

ځ ځ ځ ځ

Picture 2 Glyphs of letter Dze in Pashto

Kaf Sign has no known special name neither in Arabic (the language it was borrowed from), nor in Pashto. We asked about its name 3 Arabs and 3 Afghans, all of them referred to it as "Hamza", being fully aware, that its form is quite different from usual Hamza, which can be seen in Unicode characters U+0654, U+0655, U+0674 etc.

Kaf Sign is similar in shape and name to Hamza, but differs in function. According to the orthographic rules, adopted in Afghanistan, its function is formation of letter Dze, while Hamza in Pashto script is used:

a) as a diacritical mark above some letters in Arabic loan words like *influence*, *responsible*, *trace* etc.

b) as a diacritical mark above letter Ya in Pashto verb endings like *he is*, *you do*, *you become*, where it in combination with Ya represents sound [əy].

On Picture 3 a fragment from a textbook is shown (2, p.35).

Kaf Sign in Letter Dze

Kaf Sign in Kaf

Ya with Hamza

Picture 3 A fragment from Pashto textbook

It can be seen, that Kaf Signs play two roles here: a) form letter Kaf and b) form letter Dze. In both cases Kaf Sign is obligatory and cannot be omitted. And it is opposed to Hamza, that is written over Ya, which represents pseudo-diphthong [y].

Similar situation can be observed on Picture 4. This fragment was taken from the popular magazine "Zhwandun" (1, p.22). Kaf Sign in letter Gaf is optional and can be omitted.

The same orthographic rule for forming letter Dze is characteristic for books (3) and (4) (see References). According to Prof. K.A.Lebedev (Moscow State University) such a rule was recommended by Pashto Tol na (Pashto Academy of Science).

Unfortunately, there is pro et contra. Not all the books published in Kabul follow this recommendation and use Kaf Sign to form letter Dze. A good counterexample is book (5), where Hamza is used for this purpose. We think that the reason for this was that this huge handwritten book (5) required a great amount of work. The recommendation of Academy was relatively new tendency at the time it was published. And the most of the work had been done before such a recommendation took place. Books with Hamza in letter Dze are outnumbered, but still can be found in Kabul.

To verify our observations we interrogated 5 Afghans, living in Moscow. Their description is given in the table below. The question to them was: "What element is preferable in letter Dze: Hamza or Kaf Sign?" Their answers are given in the table.

Native speaker	Place of birth and study	Choice	Comments
P1	Kabul	Kaf Sign	
P2	Kandaghar	Hamza	
P3	Hilmand	Hamza	Said: "Kaf Sign is no mistake, but Hamza is better".
P4	Paktia	Kaf Sign	
D1	Kabul	Kaf Sign	Has good proficiency in Pashto, but his mother tongue is Dari.

From our discussion with native speakers we came to conclusion, that there are more proponents of Kaf Sign in the capital of the country. The probability of Kaf Sign usage outside of the capital region is lower. Beside that, we put the same question to a Pashtun from Peshawar Region, who never lived in Afghanistan. He wouldn't accept Kaf Sign at all.

Picture 4 A fragment from magazine "Zhwandun"

In Russia and former Soviet Union many books were published on Pashto language and literature. The list of several Russian books, where Kaf Sign is systematically used in Pashto graphics to form Letter Dze, is shown below. Prof. K.A.Lebedev, who teaches Pashto at University for more than 40 years, has the following opinion: "Discernment of Kaf Sign is evidence of good scientific level, and, on the contrary, disregard of it shows ignorance of orthographic rules and academic recommendations".

Summary

Kaf Sign is used to form letter Dze:

1) In Pashto books, published in Kabul. The community of contemporary native users can be estimated to enumerate dozens of thousands at least.

2) In descriptions and textbooks of Pashto, published in Russia and NIS Countries. The community of users consists of thousands of specialists.

In the books and magazines belonging to these two categories Kaf Signs is strongly opposed to Hamza, which has other functions in Pashto script.

Acknowledgements

Demo Font

Demo font "Paxto Normal" contains zero width Kaf Sign encoded at U+06EE as a diacritical mark. The character can be inserted into the text through Alt+X combination plus code or through Insert/Symbol etc. in MS Word XP or Word 2000. Letter Dze can be formed either by inputting two characters: U+062D and   U+06EE (Result=) or by inputting one character   at U+0682, which was put there for demo purposes. In future implementations it would be more convenient for a Pashto user to input this letter at one key stroke.

References

Pashto

1) Popular Magazine "Zhwandun" (Life), 28th of Saratan, 1360 (1981), Kabul; ‚1360   28 ‚

2) Textbook „Pashto j ba" (Pashto language) for the 3rd form. By M.U.Nezhand, Gh.S.Sharifi, P.M.Wardak. Kabul, 1359 (1980);   ‚   :   ‚
1359 ‚   ‚   ‚

3) University Publication "The Dragon of Egoism", Department of Philology and Humanities, Kabul, 1356 (1977); 1356 ‚   ‚   ‚   ‚

4) Collected songs in Dari and Pashto   ‚
1337   20 ‚   ‚   . :   ‚

5) Pashto-Pashto Descriptive Dictionary (in 5 vols.), Kabul, 1981;   ‚   ―
1360 ‚   ‚

Russian

6) Aslanov M.G. Pashto-Russian Dictionary. 50000 words, 5000 copies, Moscow, 1966;
Асланов М.Г. Афганско-русский словарь (пушту). 50000 слов, тираж 5000 экз., Москва, 1966

7) Lebedev K.A. Pashto Grammar. Moscow University, 1988, 500 copies; ISBN 5-211-00500-7
Лебедев К.А. Грамматика языка пушту. МГУ, 1988, тираж 500 экз.

8) Lebedev K.A. Modern Pashto Syntax. Moscow, 1996, 500 copies; ISBN 5-02-017896-9;
Лебедев К.А. Синтаксис современного языка пушту, Москва, 1996, тираж 500 экз.

9) Lebedev K.A. Pashto-Russian Dictionary. Moscow, 2000, ISBN 5-8463-0029-4;
Лебедев К.А. Пушту-русский словарь, Муравей-Гайд, Москва, 2000.

10) Yatsevich L.S., Ostapenko V.B. Textbook of Pashto (Basic course, Part 1), Moscow, 2001, 1000 copies; ISBN 5-8463-0045-6; Яцевич Л.С., Остапенко В.Б. Учебник языка пушту (Основной курс, I часть)

11) Novgorodova I.Yu. Textbook of Pashto (Introductory course), Moscow, 2001 (in Press);
Новгородова И.Ю. Учебник языка пушту (Вводный курс), Москва, 2001 (в печати).