

**ISO/IEC JTC1/SC2/WG2 N4218**

<b>Doc Type</b>	
<b>Title:</b>	<b>Proposal to add a Named UCS Sequence Identifier UYGHUR LETTERS</b>
<b>Source:</b>	<b>China</b>
<b>Action:</b>	<b>Review at WG2 Meeting</b>
<b>Distribution: SC2/WG2</b>	

The Unicode Technical Committee received input from Uyghur experts indicating the cultural importance of a text element referred to as vowel letters:

	Nominal form	Alphabet letters name	sequence characters	sequence characters code	Isolated form and represented code	Initial form and represented code	Medial form and represented code	Final form and represented code
1	ئا	Uyghur letter A	ئا+ئى	0626+0627	ئا FBEA ا FE8D			ئا FBEB ا FE8E
2	ئە	Uyghur letter AE	ئە+ئى	0626+06D5	ئە FBEC ە FEE9			ئە FBED ە FEEA
3	ئى	Uyghur letter E	ئى+ئى	0626+06D0	ئى FBF6 ې FBE4	ئى FBF8 ې FBE6	ئى FBD1 ې FBE7	ئى FBF7 ې FBE5
4	ئى	Uyghur letter I	ئى+ئى	0626+06D9	ئى FBF9 ى FEEF	ئى FBFB د FBE8	ئى FBD2 د FBE9	ئى FBFA ى FEF0
5	ئو	Uyghur letter O	ئو+ئى	0626+0648	ئو FBEE و FEED			ئو FBEE و FEEF
6	ئو	Uyghur letter U	ئو+ئى	0626+06C7	ئو FBF0 ۇ FBF0			ئو FBF0 ۇ FBF0
7	ئو	Uyghur letter OE	ئو+ئى	0626+06C6	ئو FBF2 ۆ FBD9			ئو FBF3 ۆ FBDA
8	ئو	Uyghur letter OU	ئو+ئى	0626+06C8	ئو FBF4 ۈ FBDB			ئو FBF5 ۈ FBDC
9	نک	Chaghatay letter NG	نک+ن	0646+06A9	نک FBD1	نک FBD2	نک FBDO	نک FBCF

This text element is perceived culturally to be a single character and is listed and named in "alphabet" lists. This text element has a representation in the UCS as a character sequence: <0626,0627>. And every vowel letters linker mode.

On that basis, the Unicode Consortium proposes addition of the following named UCS sequence identifier in Amendment 8 of ISO/IEC 10646:

Uyghur letter A : Nominal form ئا , sequence character <ئا>, sequence code <0626,0627> .

Isolated form ئا , sequence character <ئا>, sequence code <0626,0627>, represented code < FBEA >

Isolated form ا , sequence character <ا>, sequence code <0627>, represented code < FE8D >

final form ئا , sequence character <ئا>, sequence code <0626,0627>, represented code < FBEB >

final form ا , sequence character <ا>, sequence code <0627>, represented code < FE8E >

Uyghur letter AE : Nominal form ئە , sequence character <ئە>, sequence code <0626,06D5> .

Isolated form ئە , sequence character <ئە>, sequence code <0626,06D5>, represented code < FBEC >

Isolated form ە , sequence character <ە>, sequence code <06D5>, represented code < FEE9 >

final form ئە , sequence character <ئە>, sequence code <0626,06D5>, represented code < FBED >

final form ە , sequence character <ە>, sequence code <06D5>, represented code < FEEA >

Uyghur letter E : Nominal form ئې , sequence character <ئې>, sequence code <0626,06D0> .

Isolated form ئې , sequence character <ئې>, sequence code <0626,06D0>, represented code < FBF6 >

Isolated form ې , sequence character <ې>, sequence code <06D0>, represented code < FBE4 >

Initial form ئې , sequence character <ئې>, sequence code <0626,06D0>, represented code < FBF8 >

Initial form ې , sequence character <ې>, sequence code <06D0>, represented code < FBE6 >

Medial form ئې , sequence character <ئې>, sequence code <0626,06D0>, represented code < FBD1 >

Medial form ې , sequence character <ې>, sequence code <06D0>, represented code < FBE7 >

final form ئې , sequence character <ئې>, sequence code <0626,06D0>, represented code < FBF7 >

final form ې , sequence character <ې>, sequence code <06D0>, represented code < FBE5 >

Uyghur letter I : Nominal form ئى , sequence character <ئى>, sequence code <0626, 06D9> .

Isolated form ئى , sequence character <ئى>, sequence code <0626, 06D9>, represented code < FBF9 >

Isolated form ى , sequence character <ى>, sequence code <06D9>, represented code < FEEF >

Initial form ئى , sequence character <ئى>, sequence code <0626, 06D9>, represented code < FBFB >

Initial form ى , sequence character <ى>, sequence code < 06D9>, represented code < FBE8 >

Medial form ئى , sequence character <ئى>, sequence code <0626, 06D9>, represented code <FBD2 >

Medial form ى , sequence character <ى>, sequence code <06D9>, represented code < FBE9 >

final form ئى , sequence character <ئى>, sequence code <0626, 06D9>, represented code < FBFA >

final form ى , sequence character <ى>, sequence code <06D9>, represented code < FEF0 >

Uyghur letter O : Nominal form ئو , sequence character <ئو>, sequence code <0626,0648> .

Isolated form ئو , sequence character <ئو>, sequence code <0626,0648>, represented code < FBEE >

Isolated form و , sequence character <و> , sequence code <0648>, represented code < FEED >  
final form و , sequence character <و+ئ>, sequence code <0626,0648>, represented code < FBEF >  
final form و , sequence character <و> , sequence code <0648>, represented code < FEEE >

Uyghur letter U : Nominal form ئۇ , sequence character <ئ+ۇ>, sequence code <0626,06C7> .

Isolated form ئۇ , sequence character <ئ+ۇ>, sequence code <0626,06C7>, represented code < FBF0 >

Isolated form ئۇ , sequence character <ئ> , sequence code <06C7>, represented code < FBD7 >

final form ئۇ , sequence character <ئ+ۇ>, sequence code <0626,06C7>, represented code < FBF1 >

final form ئۇ , sequence character <ئ> , sequence code <06C7>, represented code < FBD8 >

Uyghur letter OE : Nominal form ئۆ , sequence character <ئ+ۆ>, sequence code <0626,06C6> .

Isolated form ئۆ , sequence character <ئ+ۆ>, sequence code <0626,06C6>, represented code < FBF2 >

Isolated form ئۆ , sequence character <ئ> , sequence code <06C6>, represented code < FBD9 >

final form ئۆ , sequence character <ئ+ۆ>, sequence code <0626,06C6>, represented code < FBF3 >

final form ئۆ , sequence character <ئ> , sequence code <06C6>, represented code < FBDA >

Uyghur letter OU : Nominal form ئۇ , sequence character <ئ+ۇ>, sequence code <0626,06C8> .

Isolated form ئۇ , sequence character <ئ+ۇ>, sequence code <0626,06C8>, represented code < FBF4 >

Isolated form ئۇ , sequence character <ئ> , sequence code <06C8>, represented code < FBDB >

final form ئۇ , sequence character <ئ+ۇ>, sequence code <0626,06C8>, represented code < FBF5 >

final form ئۇ , sequence character <ئ> , sequence code <06C8>, represented code < FBDC >

Chaghatay letter NG : Nominal form نك , sequence character <ن+ك>, sequence code <0646,06A9> .

Isolated form نك , sequence character <ن+ك>, sequence code <0646,06A9>, represented code < FBD1 >

Initial form نك , sequence character <ن+ك>, sequence code <0646,06A9>, represented code < FBD2 >

Medial form نك , sequence character <ن+ك>, sequence code <0646,06A9>, represented code < FBDO >

final form نك , sequence character <ن+ك>, sequence code <0646,06A9>, represented code < FBCF >

## Reference 1:

Table 1-2: Special vowel Scripts to Be Supplemented in ISO/IEC 10646 /Unicode for Uyghur languages

Language	Final form	Medial form	Initial form	Isolated form	Nominal form	No
Uyghur	ا FE8E    ئا FBEB			ا    ئا FE8D    FBEA	ئا	1
Uyghur	ە FEEA    ئە FBED			ئە    ە FBEC    FEE9	ئە	2
Uyghur	ې FBE5    ئې FBF7	بې FBE7    ئېب FBD1	ب    ئې FBE6    FBF8	بې    ئې FBE4    FBF6	ئېب	3
Uyghur	ى FEF0    ئى FBFA	ى FBE9    ئىب FBD2	د    ئى FBE8    FBF8	ى    ئى FEEF    FBF9	ئى	4
Uyghur	و FEEE    ئو FBEF			و    ئو FEED    FBEE	ئو	5
Uyghur	ۇ FBD8    ئۇ FBF1			ۇ    ئۇ FBD7    FBF0	ئۇ	6
Uyghur	ۈ FBDA    ئۈ FBF3			ۈ    ئۈ FBD9    FBF2	ئۈ	7

Uyghur	ۋ FBDC	ۋ FBF5			ۋ FBDB	ۋ FBF4	ۋ	8
--------	-----------	-----------	--	--	-----------	-----------	---	---

**FB50**

**Arabic Presentation Forms-A**

	FB5	FB6	FB7	FB8	FB9	FBA	FBB	FBC	FBD	FBE	FBF
0	آ FB60	ا FB60	آ FB70	چ FB80	ک FB90	ن FBA0	ے FBB0			و FBE0	ؤ FBF0
1	آ FB61	ا FB61	آ FB71	چ FB81	ک FB91	ن FBA1	ے FBB1			و FBE1	ؤ FBF1
2	ب FB62	ب FB62	ب FB72	ب FB82	ب FB92	ب FBA2				و FBE2	ؤ FBF2
3	ب FB63	ب FB63	ب FB73	ب FB83	ب FB93	ب FBA3			ك FBD3	و FBE3	ؤ FBF3
4	ب FB64	ب FB64	ب FB74	ب FB84	ب FB94	ب FBA4			ك FBD4	ي FBE4	ؤ FBF4
5	ب FB65	ب FB65	ب FB75	ب FB85	ب FB95	ب FBA5			ك FBD5	ي FBE5	ؤ FBF5
6	ب FB66	ب FB66	ب FB76	ب FB86	ب FB96	ب FBA6			ك FBD6	ي FBE6	ؤ FBF6
7	ب FB67	ب FB67	ب FB77	ب FB87	ب FB97	ب FBA7			و FBD7	ي FBE7	ؤ FBF7
8	ب FB68	ب FB68	ب FB78	ب FB88	ب FB98	ب FBA8			و FBD8	ي FBE8	ؤ FBF8
9	ب FB69	ب FB69	ب FB79	ب FB89	ب FB99	ب FBA9			و FBD9	ي FBE9	ؤ FBF9
A	ب FB6A	ب FB6A	ب FB7A	ب FB8A	ب FB9A	ب FBA0			و FBD0	ي FBE0	ؤ FBF0
B	ب FB6B	ب FB6B	ب FB7B	ب FB8B	ب FB9B	ب FBA1			و FBD1	ي FBE1	ؤ FBF1
C	ب FB6C	ب FB6C	ب FB7C	ب FB8C	ب FB9C	ب FBA2			و FBD2	ي FBE2	ؤ FBF2
D	ب FB6D	ب FB6D	ب FB7D	ب FB8D	ب FB9D	ب FBA3			و FBD3	ي FBE3	ؤ FBF3
E	ب FB6E	ب FB6E	ب FB7E	ب FB8E	ب FB9E	ب FBA4			و FBD4	ي FBE4	ؤ FBF4
F	ب FB6F	ب FB6F	ب FB7F	ب FB8F	ب FB9F	ب FBA5			و FBD5	ي FBE5	ؤ FBF5

## Reference 2: ISO/IEC JTC1/SC2/WG2 58 meeting 8 June 2011

### ISO/IEC JTC1/SC2/WG2 N4113

#### Ad hoc report on Uyghur

SOURCE: Tero Aalto, CSC – IT Center for Science

DATE: 8 June 2011

LOCATION: Finnish Standards Association SFS, Helsinki, Finland

#### PARTICIPANTS:

Tero Aalto (Finland)

Chen Zhuang (China)

Mike Ksar (independent)

Lu Qin (HKSAR)

Mikami Yoshiki (Japan)

Wushouer Silamu (China)

V. S. Umamaheswaran (Canada)

Yasen Yiming (China)

The ad hoc group met and discussed issues relating to three documents:

- the original proposal from China (“Preliminary Proposal for Encoding Special Scripts and Characters in UCS for Uyghur, Kazakh and Kirgiz” N3819)

- comments on this proposal (N3889)

- a new proposal from China (“Proposal to Encode Special Scripts and Characters in UCS for Uyghur language” N3919).

The discussion in the ad hoc group was based on the following document:

100078 N4067 – Proposal to Encode Special Scripts and Characters in UCS for Uyghur language

The eight characters proposed in N4067 can presently be produced as combinations of two existing characters. According to the proposal, these characters should, in addition, be encoded separately.

An argument supporting the proposal was the fact that the characters can have different numbers of shapes in different languages. However, this is not within the scope of UCS. Instead, different glyph variants are handled by fonts. According to the principle, anything that can be represented with a sequence will not be encoded as a separate character.

**The ad hoc group arrived at the following conclusions:**

100078 There is to be a named UCS Sequence Identifier (USI) for each of the proposed sequences.

- A USI is also to be determined for the character proposed in N4066 – PROPOSING TO SUPPLEMENT WITH THE SCRIPT AND CHARACTER OF Chaghatay Language.

100078 The contributor is requested to produce and submit another contribution defining the nine sequences for consideration in a future meeting.