TO: Unicode Technical Committee

FROM: Greg Eck (with input from members of the W3C Mongolian discussion list)

DATE: 26 January 2016

SUBJECT: Mongolian discussion documents

This document includes the following:

1. "Baluda Refinements"

Request to change General Category for Mongolian MONGOLIAN LETTER ALI GALI BALUDA U+1885 and MONGOLIAN LETTER ALI GALI THREE BALUDA U+1886 (page 1)

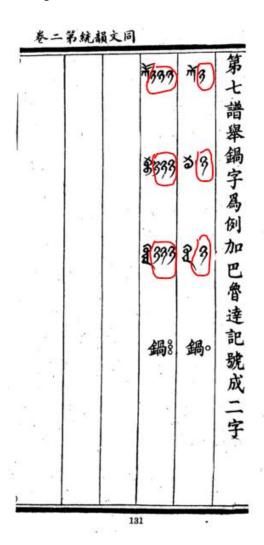
- 2. "FVS Mis-match" (pp. 2-5)
- 3. Contents page (page 6)
 - a. Exhaustive Analysis of Original Standardized Variants Listing (pp. 7-9)
 Mongolian Velar Family of Glyphs (pp. 9-13)
 Proposal for Unification of Mongolian Velar Block (pp. 13-14)
 Action items for Unification of the Mongolian Velar Block (page 14)
 Unicode Isolate Glyphs without FVS Assignment (pp. 14-15)
 - b. Additional New FVS Assignments (pp. 15-17)
 - c. Additional New FVS Control Character (page 17)
- 4. Standardized Variants Table (pp. 18-30)

This table has two views:

- (a) The first view is sorted on the CodePoint sequence.
- (b) The second view is sorted on the particular grouping under review (eg. first is the Velar Glyph Family, second is the Unicode Isolates, third are the new proposals for assignment, etc.)

Baluda Refinements

The Baluda pair (U+1885 / U+1886) has never been implemented correctly. This was brought to light in 2012 in meetings with Professor Quejingzhabu when he informed us that the baluda actually is written to the right of the top-most letter of a given word. Since that time, others have confirmed the same thought. Andrew West has put forth the following interesting finding¹:



PLUSS.

Proposal: Change the general category from Lo (Letter, Other) to Mn (non-spacing mark) so that the character can be processed as a diacritical mark rather than a letter drawn on the stem. The Dagalga (U+18A8) is currently being processed as a mark successfully.

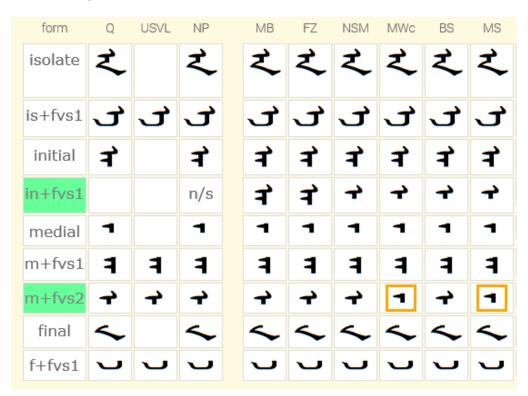
¹ http://www.babelstone.co.uk/Mongolian/TWYT 130.jpg

FVS MIS-MATCH

(see DS00 FVSMisMatch.pdf)

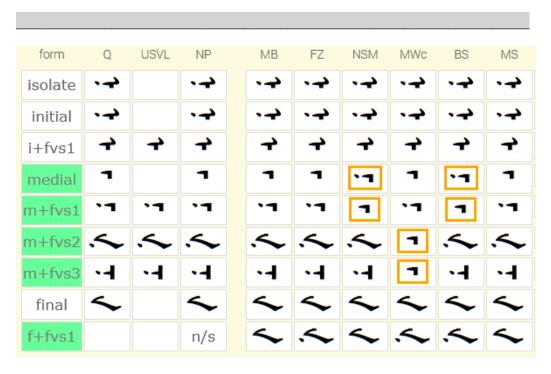
There are two opinions on how to determine the position of a glyph before/after the NNBSP/MVS. The first method as written up in the earliest documentation (MGWBM/TR170) seems to hold to the idea that anything between the first letter and the last letter of a Mongolian word is in the medial position. If there is a NNBSP or a MVS control character in the middle of the word, the letters adjacent to the NNBSP/MVS are still considered to be in the medial position. The second method as implemented by font developers (at least those in our forum) is to consider the letters pre/post NNBSP/MVS as final/initial. In other words, the letter before an NNBSP/MVS (assuming that the string of letters preceding the NNBSP/MVS has more than one letter) is considered to be in the final position. A letter following an NNBSP/MVS is considered to be in the initial position (assuming that the string of letters following the NNBSP/MVS has more than one letter). The Microsoft Universal Shaping Engine (aka Uniscribe) and the Google Harfbuzz rendering engine both follow the second method. This has created a "disconnect" between those holding to the first method and those holding to the second method. There are several cases where we can see that a font has placed the FVS assignment in two locations, thus following both methods rather than holding to one opinion.

Proposal: Leave the fonts as they are (in running text implementation). Glyph shaping in running text displays fine. It is the FVS specification where we find variation and duplication. Change the FVS specification to match the actual positions as determined by the font rendering engines. Given such a specification, the individual font developers will know how to change their FVS implementations in their individual fonts.



- U+1820-A (post-NNBSP) by the MGWBM, TR170, and current Chinese Standard: **Medial** ZWJ+1820+FVS2+ZWJ

 Third Medial Form
- U+1820-A (post-NNBSP) by the Universal Shaping Engine (Uniscribe) / Harfbuzz: Initial
- Proposed



- U+1828-N (pre-MVS) by the MGWBM, TR170, and current Chinese Standard: **Medial**ZWJ+1828+FVS2+ZWJ

 Third Medial Form
- U+1828-N (pre-MVS) by the Universal Shaping Engine (Uniscribe) / Harfbuzz: Final

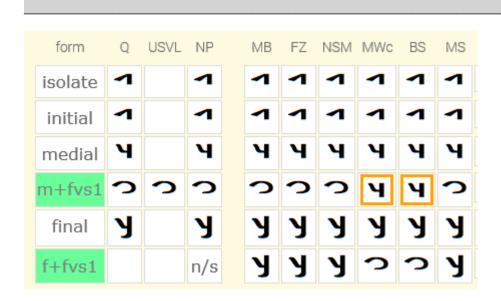


- U+182C-Q (pre-MVS) #1 by the Universal Shaping Engine (Uniscribe) / Harfbuzz: Final
- Proposed ZWJ+182C+FVS1 Second Final Form

- U+182C-Q (pre-MVS) #2 by the MGWBM, TR170, and current Chinese Standard: **Medial** ZWJ+182C+FVS3+ZWJ Fourth Medial Form
- U+182C-Q (pre-MVS) #2 by the Universal Shaping Engine (Uniscribe) / Harfbuzz: Final
- Proposed (x_ First Final Form



- U+182D-G (pre-MVS) by the Universal Shaping Engine (Uniscribe) / Harfbuzz: Final



• U+1835-J (pre-MVS) by the MGWBM, TR170, and current Chinese Standard: Medial

ZWJ+1835+FVS1+ZWJ

Λ (x__)

Second Medial Form

- U+1835-J (pre-MVS) by the Universal Shaping Engine (Uniscribe) / Harfbuzz: Final
- ZWJ+1835+FVS1

Proposed

Second Final Form



- U+1836-Y (pre-MVS) by the MGWBM, TR170, and current Chinese Standard: Medial
 ZWJ+1836+FVS2+ZWJ
 Third Medial Form
- U+1836-Y (pre-MVS) by the Universal Shaping Engine (Uniscribe) / Harfbuzz: Final

• Proposed

ZWJ+1836

° (x__) Final Form

This paper has three sections:

- I. Exhaustive Analysis of Original Standardized Variants Listing
 - A. Column 1 Graphic image of the Unicode code-point in question
 - B. Column 2 Unicode sequence assigned to a given positional variant (eg. 1820 180B, 1825 180C, 182D 180D, etc.)
 - C. Column 3 Variant position (ie. Isolate, initial, medial or final)
 - D. Column 4 Positional variant glyph image/text
 - E. Column 5 Unicode formal name of the code-point followed by phrase describing the particular variant
 - F. Superfluous assignments / Troublesome assignments
 - 1. Mongolian Velar Family of Glyphs
 - 2. Unicode Isolate Glyphs without FVS Assignment
- II. Additional New FVS Assignments
- III. Additional New FVS Control Character

Exhaustive Analysis of Original Standardized Variants Listing 1

Column 1 - No issues

Column 2 – One issue (182D-GA) Masculine final form - ; Feminine final form -

• There is confusion at the feminine final 182D-GA. This is easily the most complex variant in the Mongolian script context-wise and possibly specification-wise as well. The MGWBM places the variant at 182D+180C. The TR170 specification places the variant at the182D+180B. The reason that the MGWBM skips one FVS slot (FVS1) and places the feminine form at 182D+180C is that the specification has taken into account the need for the default context to be over-ridden in special cases. There are several cases where the default context determining the final feminine form is present, but the correct form is the masculine. This case would be handled by the 182D+FVS1-Final over-ride. The MGWBM notation uses the parenthesis, as below, to say that the masculine is both the final default as well as the FVS1 assigned glyph. This form of notation is confusing in its terse-ness, however concise in its descriptive power. This is what I teach my students:

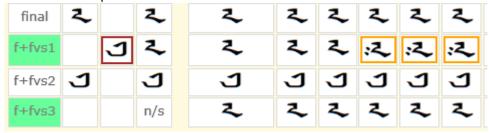
ZWJ+182D

ZWJ+182D+FVS1

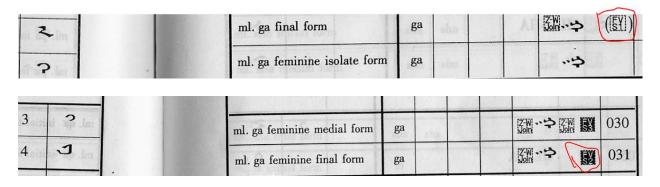
ZWJ+182D+FVS2

ZWJ+182D+FVS2

The font implementations under inspection are all in complete agreement as shown at the Ishida Font Comparator site²



MGWBM follows ...



TR170 follows ...

¹ http://www.unicode.org/Public/UNIDATA/StandardizedVariants.html as of January 20, 2016.

² http://r12a.github.io/scripts/mongolian/variants.html

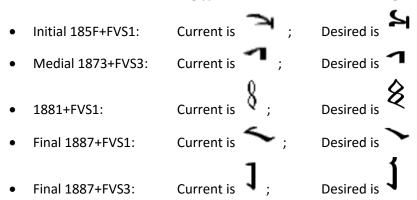
45	·.\$	MONGOLIAN LETTER GA	-		**	FIRST INITIAL FORM		**	Z-W Join		GA
			44		4	SECOND INITIAL FORM		**	Z-W Join	FV S1	GA
				6	7	FIRST MEDIAL FORM	Z-W Join	**	Z-W Join		GA
				26	77	SECOND MEDIAL FORM	Z-W Join	**	Z-W Join	FV S1	GA
				27	•	THIRD MEDIAL FORM	Z-W Join	**	Z-W Join	FV S2	GA
				28	2	FINAL FORM	Z-W Join	**			GA
			137		ņ	FEMININE ISOATE FORM		*			GA
				30	r	FEMININE MEDIAL FORM	Z-W Join	.⇒	Z-W Join	FV S3	GA
				31	77	FEMININE FINAL FORM	Z-W Join	**	(FV S1	GA

PROPOSAL: Change the FVS assignment for the feminine final GA from 182D+FVS1 to 182D+FVS2.

Column 3 - One issue of incorrect positional assignment, as follows ...

• The 1887 isolate is actually a final. This was brought to light in discussions with Professor Quejingzhabu late last year (2015). The question now, is how to assign it as a final since all of the FVS slots are taken up. Suggested solution is the specification of a new FVS4.

Column 4 – Five issues where the glyph is either incorrect or needing refinement, as follows ...



Column 5 - Eleven issues where the description could be sharpened, as follows ...

- Medial 1828+FVS4 the term "separate" is used only once in the entire block; it is felt that
 in keeping to similar descriptions such as "first", "second", "third", "fourth", there is greater
 clarity. Proposed description is "fourth medial form" rather than "separate form".
- Final 182A+FVS1 the term "alternate" again is used only in this one occasion; for the same reason as above the proposed description is "second final form" rather than "alternate form".
- Isolate 182C+FVS1 the main feature used in the description is easily seen to be position (first, second, third, fourth, fifth); the additional feature of gender if used in the description should not be as prominent as the main feature of position; if used, it should be used consistently throughout the specification a study should be done to determine whether this is indeed the case. Proposal is to leave the description as tight as possible and use only position and leave out gender in this case "second isolate form".
- Final 182D+FVS2 for the same reason as above the proposal is to leave the term "feminine" out and use the description "third final form" instead of "feminine form". There

- will be cases where this "feminine" form is used in words that are actually masculine also, so again the addition of gender directly in the description is a bit cumbersome.
- Medial 182D+FVS3 for the same reason above, the proposal is to use the description of "fourth medial form" rather than "feminine form".
- Initial 184D+FVS1 for the same reason above, the proposal is to use the description of "second initial form" rather than "feminine form".
- Medial 184D+FVS1 for the same reason above, the proposal is to use the description of "second medial form" rather than "feminine form".
- Final 1874+FVS1 for the same reason above, the proposal is to use the description of "second final form" rather than "feminine first final form".
- Medial 1874+FVS2 for the same reason above, the proposal is to use the description of "third medial form" rather than "feminine first medial form".
- Final 1874+FVS2 for the same reason above, the proposal is to use the description of "third final form" rather than "feminine second final form".
- Medial 1874+FVS3 for the same reason above, the proposal is to use the description of "fourth final form" rather than "feminine second final form".

Mongolian Velar Family of Glyphs

The Mongolian Script standards (MGWBM³ and the TR170⁴) have gone through a lot of change from definition to the stage of actual implementations. You can see the initial design thoughts in the early standards. You can trace the history of what worked and what did not work. One thing that seems to have not worked was the use of the feminine isolates as an initial default. At least in the Chinese Standard, you can see the move away from this idea to the more workable masculine as the display default. I have never had a full round of discussion about this process, but it seems that this is what happened. Given this premise then, it is reasonable that we see gaps and inconsistencies around the area of the feminine isolates for the Mongolian QA/GA, Todo QA/GA, Sibe KA/GA/HA, and the Manchu KA.

This is a very special class of variants. There is no question that they all need assignment for purposes of teaching heuristic, for dictionary headings, and such. There is also the possibility of over-specification however as currently seen in the Todo pair. I think it is wise to look at these all together and determine an appropriately economic as well as effectual assignment. Given the need to display these forms in dictionaries, we have a clear justification for the stand-alone isolate. There is another area that covers probably more scope however. That is the area of ligaturing. The feminine forms of these glyphs all ligature readily. And we might say more precisely, they only ligature except when used in an isolate setting. This means that they never appear in initial or medial positions except in ligatured context. And of course, they would never appear in a final context. So, we might consider these feminine glyphs to have virtual forms in the initial and medial contexts. As we have already agreed that the isolate form is justified, we might consider, for purposes of economics, not assigning FVS initial/medial sequences (which should be almost identical), but allowing the isolate form to represent the virtual forms at the initial / medial contexts. A precedent for this approach is found currently in the Mongolian 182C/182D pair where there is no initial/medial FVS assignment found for these feminine glyphs. This representation could be annotated allowing for the overall reduction of at least two FVS assignments per letter.

³ http://www.babelstone.co.uk/Mongolian/MGWBM.html

⁴ ftp://ftp.iist.unu.edu/pub/techreports/report170a.tgz

There is another case where this method has not been followed resulting in inconsistent specification. This is the Todo pair at U+184D/U+184E. According to the MGWBM/TR170/USVL we find one initial specification and one medial specification at U+184D. However at U+184E, the sister letter, we find no initial/medial feminine form assignment. Instead we find one isolate feminine FVS assignment. This method of assignment is exactly the opposite of the sister letter, 184D.

Looking at the attached chart, we can start with the Mongolian pair QA/GA to set the nomenclature paradigm. Though current documentation is shy about specifying an isolate for a consonant, all font designers know that each position must have a display form. Therefore, specify the QA initial default (which is also the Unicode presentation glyph) as the masculine isolate glyph. This is followed by the feminine isolate closest to the default – the undotted form taking the first FVS1 slot. The remaining dotted feminine isolate takes the FVS2 slot. Following the same pattern for the Mongolian 182D-GA, we take the initial default which is masculine as the default isolate. The feminine glyph closest in feature to the default isolate is the dotted isolate – so this glyph takes the FVS1 slot. The remaining isolate, the undotted feminine form takes the FVS2 slot. This pattern, which almost matches the earliest documentation then gives us a template to follow in the FVS assignments of the Todo pair, the Sibe three and the Manchu single.

MONGOLIAN 182C-QA/182D-GA (TR170)

-											
44	\$	MONGOLIAN LETTER QA	-		\$	FIRST INITIAL FORM		\$	Z-W Join		QA
			45			SECOND INITIAL FORM		\$	Z-W Join	FV S1	QA
				6	4	FIRST MEDIAL FORM	Z-W Join	\$	Z-W Join		QA
				26	:4	SECOND MEDIAL FORM	Z-W Join	\$	Z-W Jon	FV S1	QA
				27	:2.	THIRD MEDIAL FORM	Z-W Join	\$	Z-W Join	FV S2	QA
				28	2	FOURTH MEDIAL FORM	Z-W Join	\$	Z-W Join	FV S3	QA
			137		12	FEMININE FIRST ISOLATE FORM		4			QA
				29	(:3)	FEMININE SECOND ISOLATE FORM		\$		FV S1	QA
45	**	MONGOLIAN LETTER GA	-		٠.٠	FIRST INITIAL FORM		ふ	Z-W Join		GA
			44		\$	SECOND INITIAL FORM		**	Z-W Join	FV S1	GA
				6	1	FIRST MEDIAL FORM	Z-W Join	**	Z-W Join		GA
				26	;	SECOND MEDIAL FORM	Z-W Join	*	Z-W Jon	FV S1	GA
				27	:2,	THIRD MEDIAL FORM	Z-W Join	ふ	Z-W Join	FV S2	GA
				28	2	FINAL FORM	Z-W Join	*			GA
			137		(r	FEMININE ISOATE FORM		**			GA
				30) ر	FEMININE MEDIAL FORM	Z-W Join	ふ	Z-W Join	FV S3	GA
				31	J	FEMININE FINAL FORM	Z-W Join	ふ		FV S1	GA

MONGOLIAN 182C-QA/182D-GA (MGWBM)

	(編) 8	一题	an I	mue form	1	2
Oall	3	A PER YEAR	3005	137	ed tobs	(3)
023		& Well-year	(00)		2	:5
045	ψ٠.	ML. GA	ad		mod fair	٠.ఫ
	10	o Si ad ad	ui l	044	and feib	\$
024		o to al ad	ńd	032 - 3	orbot for	3
7	IXII	中間	ed	075 mol	1	:=
	10/87	area .	101		2	:2
	政	0 (G)	iq l	044 – 1	nol lam	2
100		ar light	29	137	al forms	(2)

Todo 184D-QA / 184E-GA (TR170)

	BA	SIC CHARACTERS				VARIANT FORM	4S			
No	CHARACTER	NAME	No	No GLYPH NAME				RU	JLE	
77	ःव	MONGOLIAN LETTER TODO QA			ःब	INITIAL FORM		ःब	Z-W Jon	
				26	:4	MEDIAL FORM	Z-W Join	ःब	Z-W Jon	
			59	1	<u>س</u>	FEMININE INITIAL FORM		ःब	Jon.	FV S1
			59	7	D	FEMININE MEDIAL FORM	2-W Join	:4	Z-W Jon	FV S1
				54) ၅	FEMININE FINAL FORM	29W Join	ःब		
78	∘∓	MONGOLIAN LETTER TODO GA			°₹	INITIAL FORM		۰4	Z-W Join	
				83	∘я	FIRST MEDIAL FORM	2-W Join	٠Ŧ	Z-W Jon	
				84	4	SECOND MEDIAL FORM	Jon.	৽ⅎ	29W Jon	FV S1
				85	乜	FINAL FORM	Z-W Join	۰ 4		
			137		(V)	FEMININE ISOLATE FORM		°व		

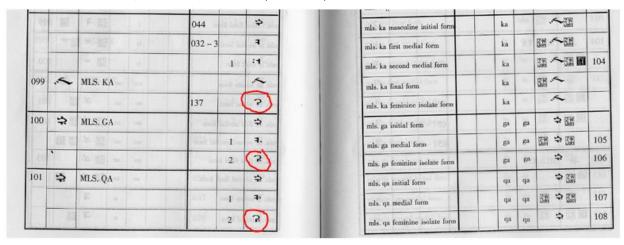
TODO 184D-QA / 184E-GA (MGWBM)

7	ःन'	MLT. QA	99		-	ःव ।	mlt. qa initial form	qa	: **
	1 4	BOUTH TE	-	045 – 1	-	:4	mlt. qa medial form with dots	qu	SH : → SH
	163	3		059	(2)	mlt. qa feminine initial form	qu	:7 38 11
				059	1	37	mlt. qa feminine medial form	qa	器 计图图
_	1 52	等 题		059 - 1	lban k	3	mlt. qa final form	qa	國 : 4 *
078	॰ न	MLT. GA	100	413,522	ip time	॰वं	mlt. ga initial form	gn	• - ₹ <u>₹</u> 8
		9 Al		035%	1	۰٦	mlt. ga first medial form	ga	SN 04 SN
		\$- All	100	To de la constitución de la cons	2	14	mlt. ga second medial form	ga	82 ° ₹ 82 8

• SIBE 1863-QA / 1864-GA / 1865-HA (TR170)

99			44		\$	INITIAL FORM		:~	Z-W Join			КА	
				6	1	FIRST MEDIAL FORM	2500 Sloini	:~	(29)06 (Join)			КА	
				104	:-	SECOND MEDIAL FORM	25W Join	:~	Jon:	FV S1		КА	
	:~	MONGOLIAN LETTER SIBE KA			:~	FINAL FORM	Join	:~				КА	
			137		(2)	FEMININE ISOLATE FORM		:~				КА	
100	₽	MONGOLIAN LETTER SIBE GA	-		#	INITIAL FORM		₽	29W Jon			GA	GA
				105	4.	MEDIAL FORM	Jon.	₽	Jon.			GA	GA
				106	(3)	FEMININE ISOLATE FORM		≉				GA	GA
101	≉	MONGOLIAN LETTER SIBE HA	-		32	INITIAL FORM		₽	ZSWI Sloin			на	на
				107	4.	MEDIAL FORM	25W Sloini	₽	Spine Spine			на	на
				108	(3)	FEMININE ISOLATE FORM		₽				на	НА

MONGOLIAN 1863-QA / 1864-GA / 1865-HA (MGWBM)



MANCHU 1874-KA (TR170)

116			44		\$	INITIAL FORM		:4	ZW Jon			КА
				6	1	FIRST MEDIAL FORM	Jone	:4	Jon:			КА
				26	:1	SECOND MEDIAL FORM	25W Join	:2	Z-W Join	S1		ка
	:2	MONGOLIAN LETTER MANCHU KA			:2	FINAL FORM	25WF Join	:4				КА
			137		\bigcirc	FEMININE ISOLATE FORM		:4				КА
				30	2	FEMININE FIRST MEDIAL FORM	Jone	:4	Join:	FV S2		КА
				120	:2	FEMININE SECOND MEDIAL FORM	2500 Join	:4	Join:	FV S3		КА
				121	2	FEMININE FIRST FINAL FORM	Jon.	:4		S1		ка
				122	c;	FEMININE SECOND FINAL FORM	Soni Sloni	:4		FV S2		КА

• MANCHU 1874-KA (MGWBM)

	RA	F6 4		094 - 3	and in	10.00	mlm. i third final form	i i/	额日	131
111		F 112 at		044	n Nove	\$	mlm. ka masculine initial form	ka	233	
	7	MUS. BAI	100	032 - 3		1	mlm. ka masculine first medial form	ka	雄之雄	
	138	Britis John	100	045 - 1	hili	:4	mlm. ka masculine second medial form	ka	湖 ス湖	
16	.2	MLM. KA	Arts .		1.2	2	mlm. ka masculine final form	ka	强之	9
(D MI)	-	OLIAN LETTER	MANCHU	+				10.2-01.61	<u> </u>	
D MI	-	OLIAN LETTER	MANCHU	137	rick factor	(2)	mtm. ka feminine isolate form	ka	2	-
ML	E .		Ala I	137 045 - 3	rick factor	2	mlm. ka feminine isolate form	1 2 2 200	2 3 3 2 5	
MI	E .		Ala I	1	end have	\sim	THE RESIDENCE OF THE PERSON OF	kı	000.00	-
D MI	E .		Ala I	1	of bits	2	mlm. ka feminine medial form	kı	羅之區	XIXI

• AliGali 1889-KA (added for sake of completion although there is nothing needed here in the FVS area)

Proposal for Unification of the Mongolian Velar Block

•	182C (x)	21	MONGOLIAN	LETTER QA
	182C+FVS1		ر _(x)	Second Isolate Form
	182C+FVS2		ζ (x)	Third Isolate Form
	182C+ZWJ		↑ (x)	First Initial Form
•	182D (x)	ù	MONGOLIAN	LETTER GA
	182D+FVS1		ر. (x)	Second Isolate Form
	182D+FVS2		ζ (x)	Third Isolate Form
	182D+ZWJ		, (x)	First Initial Form
•	184D (_x)	77	MONGOLIAN	LETTER TODO QA
	184D+FVS1		6 (x_)	Second Isolate Form
	184D+ZWJ) (_x)	Initial Form
•	184E (_x)	ᡏ	MONGOLIAN	LETTER TODO GA
	184E+FVS1		ر ₋ (x_)	Second Isolate Form
	184E+ZWJ		Tr (_x)	Initial Form
•	1863 (_x_)	\',	MONGOLIAN	LETTER SIBE KA
	1863		介 (_x_)	First Isolate Form
	1863+FVS1		ر (_x_)	Second Isolate Form
	1863+ZWJ		? (_x_)	Initial Form
•	1864 (_xx)	₹	MONGOLIAN	LETTER SIBE GA
	1864+FVS1		رد (_xx)	Second Isolate Form
	1864+ZWJ		? (_xx)	Initial Form
•	1865 (_xx)	ᡥ	MONGOLIAN	LETTER SIBE HA
	1865+FVS1		(xx)	Second Isolate Form

	1865+ZWJ		? (_xx)	Initial Form
•	1874 (x)	ĽŸ	MONGOLIAN L	ETTER MANCHU KA
	1874		↑ (x)	First Isolate Form
	1874+FVS1		(x)	Second Isolate Form
	1874+ZWJ		A (x)	Initial Form
•	1889 (xx_)	Ç	MONGOLIAN L	ETTER ALI GALI KA
	1889+ZWJ		ر (xx)	Initial Form

Action Items for Unification of the Mongolian Velar Block

- U+182C-Mongolian QA
 - Change 182C+FVS1 from the dotted feminine isolate to the undotted isolate as the second isolate form
 - Add 182C+FVS2 dotted feminine isolate as the third isolate form
- U+182D-Mongolian GA
 - Add 182D+FVS1 dotted feminine isolate as the second isolate form
 - Add 182D+FVS2 undotted feminine isolate as the third isolate form
- U+184D-Todo QA
 - Delete 184D+FVS1 feminine initial
 - Delete 184D+FVS1 feminine medial
 - Add 184D+FVS1 feminine isolate as the second isolate form
- U+184E-Todo GA
 - Add 184E+FVS1 feminine isolate as the second isolate form
- U+1863-Sibe KA
 - Add 1863+FVS1 first feminine isolate as the first isolate form
 - Add 1863+FVS2 second feminine isolate as the second isolate form
- U+1864-Sibe GA
 - Add 1864+FVS1 second feminine isolate as the second isolate form
- U+1865-Sibe HA
 - Add 1864+FVS1 second feminine isolate as the second isolate form
- U+1874-Manchu KA
 - Add 1874+FVS1 first feminine isolate as the first isolate form
 - Add 1874+FVS2 second feminine isolate as the second isolate form
- U+1889-AliGali KA
 - No action necessary

NOTE: You may be asking why there is a difference of treatment at 1863 and 1874? Why is there a first isolate in these two cases and not in the others? These are two special cases where the Unicode ascribed code-point glyph is not the initial but instead another positional variant – in this case the final variant. Whereas in other cases the first isolate is also the initial/Unicode glyph and there is no need to state that it is the first isolate – in this case the first isolate must be explicitly assigned or there will be no way to display the glyph except through simulation with the ZWJ. This begs the question, then, as to whether we should also assign FVS positions to the 15 code-points whose Unicode glyphs are also undisplayable except by simulating the position with ZWJ. This is the area as discussed below.

Unicode Isolate Forms without FVS assignments

This is an FYI at present. The issue is that there is no easy way for the user to type these code-point's Unicode form without having intrinsic understanding of the block. The user will have to use a ZWJ combination. The user will have to know the position of the particular code-point that is the Unicode form. It is felt that this should be a simple task to write any given Unicode form whereas in reality it is not.

• 1824 (Initial) – already included as a proposed new variant

•	1826 (Initial) – already included as a propos	ed new varia	nt
•	184A (Final)	ca new varia	
•	184B (Final)		
•	184F (Final)		
•	1859 (Final)		
•	185D (Final)		
•	185E (Second Medial)		
•	1861 (Medial)		
•	1862 (Final)		
•	1863 (Final)		
•	1873 (Final)		
•	1874 (Medial)		
•	1875 (Final)		
•	1876 (Medial)		
A dditic	onal New FVS Assignments		
Additio	onal New F V5 Assignments		
•	New FVS2 assignment at <u+1822-medial><</u+1822-medial>	<u+180c> Th</u+180c>	ree out of six font developers
	implement this double-tooth U+1822 form.		·
	ZWJ+1822+FVS2+ZWJ		Third Medial Form
	ZWJ+1822+FVS2+ZWJ	'' (x)	Tima Mediai Form
•	New FVS3 assignment at <u+1822-medial><</u+1822-medial>	<u+180d> Th</u+180d>	is variant is needed to over-ride
	default context of the medial I in words such		
			-
	ZWJ+1822+FVS3+ZWJ	' (x)	Fourth Medial Form
•	New FVS1 assignment at <u+1822-final><u< td=""><td>±180B> Thic</td><td>1822 final variant has been attested</td></u<></u+1822-final>	±180B> Thic	1822 final variant has been attested
•	by various independent sources.	11000> 11113	1822 Illiai variant has been attested
	ZWJ+1822+FVS1	-	Coond Final Form
	ZWJ+1822+FVS1	(x)	Second Final Form
•	New FVS1 assignment at <u+1824-isolate><</u+1824-isolate>	:U+180B> Thi	is variant is needed to display the
	Unicode isolate form of U+1824. This is also		
	1824+FVS1	-	Second Isolate Form
	1024 ⁺ F V 51	(x)	Second Isolate Form
•	New FVS1 assignment at <u+1824-final><u< th=""><th>±190B> Thic</th><th>variant is needed for foreign</th></u<></u+1824-final>	±190B> Thic	variant is needed for foreign
•	spellings such as "Nassau".	+1000> 11113	variant is needed for foreign
	, •	_	a 15' 15
	ZWJ+1824+FVS1	0 (x)	Second Final Form
	New 5VC2 resignment at 41,4024 Finals 41	. 1000x This	
•	New FVS2 assignment at <u+1824-final><u< th=""><th></th><th>•</th></u<></u+1824-final>		•
	independent sources as a variant of the dot	tea genitive o	case U/UE after the sequence
	<u+1828-n><nnbsp>.</nnbsp></u+1828-n>	^	
	ZWJ+1824+FVS2	è (x)	Third Final Form
•	New FVS2 assignment at <u+1826-isolate><</u+1826-isolate>		
	Unicode isolate form of U+1826. This is also	•	
	1826+FVS2	1 0 √ (x)	Third Isolate Form

•	New FVS3 assignment at <u+1826-medial><u+180d> This variant is needed to over-ride the</u+180d></u+1826-medial>
	context so that we can use the DUEGER with or without the NNBSP.

New FVS2 assignment at <U+1826-Final><U+180C> This variant has been attested by several independent sources as a variant of the dotted genitive case U/UE after the sequence <U+1828-N><NNBSP>.

New FVS4 assignment at <U+1828-Medial><U+180F> This default form is needed with FVS assignment to over-ride native context for such rare cases as ANAR.

New FVS1 assignment at <U+1829-Isolate><U+180B> This isolate is evidenced in the dictionary "mongyol nanggiyad üsüg-ün toli bičig" (蒙漢字典) published in 1928. This is shown in context at http://www.cneas.tohoku.ac.jp/staff/hkuri/articles/a62houkoku14.pdf



New FVS2 assignment at <U+182C-Initial><U+180C> This variant is needed for the special case of QQIR.

^(*) Third Initial Form

New FVS3 assignment at <U+182D><U+180C> This variant is needed for foreign loan words such as GRAM.

^(*) Third Initial Form

New FVS4 assignment at <U+182D><U+180F> This default form is needed with FVS assignment to over-ride native context for such rare cases as COGAGULA.

Fifth Medial Form

New FVS1 assignment at <U+1843><U+180B> This variant has been attested to.



' (x_) Second Final Form

New FVS3 assignment at <U+1887><U+180F> This is currently listed as an isolate in the Chinese Standard. Professor Quejingzhabu says that this is a mistake and is actually a final.

「「x」 Fifth Final Form

New FVS Control Character

• There are three instances now where the FVS4 may be needed, one of which is definitely confirmed. The current <U+1887-Isolate><FVS1> is actually a final as confirmed by Professor Quejingzhabu. There is discussion ongoing now as to whether this control character should actually be something of a "Default Variant Selector" as there are at least four cases where this feature could be helpful. One of our forum members would argue that there are 40+ cases where such a feature could be helpful. The problem is that we will still be needing a more generic new control character to handle instances like the U+1887. Additional information forthcoming regarding the New 2016 Chinese Standard will be helpful in directing this proposal also.

180F





MONGOLIAN FREE VARIATION SELECTOR FOUR

Further Research Necessary

- 180A-Finals
- 1860-ThirdFinalForm
- 1868/1869-Medials is the order correct?
- 1873-FourthFinalForm
- 1888-SecondMedialForm
- 18A0-SecondInitialForm
- 18A0-SecondMedialForm
- 18A1-SecondInitialForm
- 18A1-SecondMedialForm

Code-Point Sequence	Glyph	Code	Unicode Name/ Variant Name	Code Legend
180F	FVS4	II	Monglian Free Variation Selector Four	F1 – Velar Family
1820+FVS1	Ð	Х	Second Isolate Form	F2 – Unicode Isolates
1820+FVS1	Ŀ	Х	Second Medial Form	I – Proposed New Variants
1820+FVS2	1	MM	Third Medial Form	II – Proposed New FVS
1820+FVS1)	Х	Second Final Form	MM – FVS MisMatch
1821+FVS1	117	Х	Second Initial Form	X - Previous Assignment
1821+FVS1)	Х	Second Final Form	
1822+FVS1	٣	Х	Second Medial Form	
1822+FVS2	77	I	Third Medial Form	
1822+FVS3	7	I	Fourth Medial Form	
1822+FVS1	7	I	Second Final Form	
1823+FVS1	٥	Х	Second Medial Form	
1823+FVS1	D	Х	Second Final Form	
1824+FVS1	1 0	F2	Second Isolate Form	
1824+FVS1	ط	х	Second Medial Form	
1824+FVS1	D	I	Second Final Form	
1824+FVS2	è	I	Third Final Form	
1825+FVS1	57	х	Second Medial Form	
1825+FVS2	701	х	Third Medial Form	
1825+FVS1	6	х	Second Final Form	
1826+FVS1	16	х	Second Isolate Form	
1826+FVS2	107	F2	Third Isolate Form	
1826+FVS1	σ,	Х	Second Medial Form	

1826+FVS2	न्द	х	Third Medial Form
1826+FVS3	ь	I	Fourth Medial Form
1826+FVS1	6	х	Second Final Form
1826+FVS2	è	I	Third Final Form
1828+FVS1	1	х	Second Initial Form
1828+FVS1	5	х	Second Medial Form
1828+FVS2	Į,	MM	Third Medial Form
1828+FVS3	7	х	Fourth Medial Form
1828+FVS4	F	I	Fifth Medial Form
1829+FVS1	TBD	I	Second Isolate Form
182A+FVS1	9	х	Second Final Form
182C+FVS1	c	F1	Second Isolate Form
182C+FVS2	č	F1	Third Isolate Form
182C+FVS1	ŷ	х	Second Initial Form
182C+FVS2	Ç	I	Third Initial Form
182C+FVS1	n.	х	Second Medial Form
182C+FVS2	7	MM	Third Medial Form
182C+FVS3	Ķ	MM	Fourth Medial Form
182D+FVS1	c	F1	Second Isolate Form
182D+FVS2	č.	F1	Third Isolate Form
182D+FVS1	Ω	х	Second Initial Form
182D+FVS2	٠	I	Third Initial Form
182D+FVS1	'n	Х	Second Medial Form
182D+FVS2	Ķ	MM	Third Medial Form
182D+FVS3	Ç	х	Fourth Medial Form
182D+FVS4	۳	х	Fifth Medial Form

182D+FVS1	₩,	х	Second Final Form
182D+FVS2	7	х	Third Final Form
1830+FVS1	✓	х	Second Final Form
1830+FVS2	H	х	Third Final Form
1832+FVS1	4	х	Second Medial Form
1833+FVS1	চ	х	Second Initial Form
1833+FVS1	চ	х	Second Medial Form
1833+FVS1	વ	х	Second Final Form
1835+FVS1	C	MM	Second Medial Form
1836+FVS1	7	х	Second Initial Form
1836+FVS1	7	х	Second Medial Form
1836+FVS2	C	MM	Third Medial Form
1838+FVS1	Ь	х	Second Medial Form
1838+FVS1	Ф	х	Second Final Form
1843+FVS1	L	I	Second Final Form
1844+FVS1	۲	х	Second Medial Form
1845+FVS1	닐	х	Second Medial Form
1846+FVS1	þ	х	Second Medial Form
1847+FVS1	P	х	Second Isolate Form
1847+FVS1	۴	х	Second Medial Form
1847+FVS2	Ь	Х	Third Medial Form
1847+FVS1	6	Х	Second Final Form
1848+FVS1	₽	Х	Second Medial Form
1849+FVS1	1 _P	Х	Second Isolate Form
1849+FVS1	Ф	Х	Second Medial Form

184A	را ا	F2	Final Form
184B	Б	F2	Final Form
184D+FVS1	6	F1	Second Isolate Form
184E+FVS1	c	F1	Second Isolate Form
184E+FVS1	۴ı	Х	Second Medial Form
184F	4	F2	Final Form
1859	3	F2	Final Form
185D+FVS1	٢	х	Second Medial Form
185D	<i>?</i>	F2	First Final Form
185D+FVS1	1	Х	Second Final Form
185E+FVS1	7	F2	Second Medial Form
185E+FVS2	7	Х	Third Medial Form
185E+FVS1	1	Х	Second Final Form
185E+FVS2	•	х	Third Final Form
1860+FVS1	Ь	Х	Second Medial Form
1860+FVS1	6	х	Second Final Form
1860+FVS2	Ó	Х	Third Final Form
1861	σ τ	F2	Medial Form
1862	Ē	F2	Final Form
1863+FVS1	c	F1	Second Isolate Form
1863+FVS1	Σ	х	Second Medial Form
1863	Ŋ,	F2	Final Form
1864+FVS1	건	F1	Second Isolate Form
1865+FVS1	૯	F1	Second Isolate Form
1868+FVS1	9	Х	Second Initial Form

			,
1868+FVS1	চ	Х	Second Medial Form
1868+FVS2	σ ι	Х	Third Medial Form
1869+FVS1	4	Х	Second Initial Form
1869+FVS1	र्	Х	Second Medial Form
186F+FVS1	л	Х	Second Initial Form
186F+FVS1	ъ	х	Second Medial Form
1873+FVS1	7	х	Second Medial Form
1873+FVS2	77	х	Third Medial Form
1873+FVS3	7	Х	Fourth Medial Form
1873	Ç	F2	First Final Form
1873+FVS1	J	х	Second Final Form
1873+FVS2	•	х	Third Final Form
1873+FVS3	4,	х	Fourth Final Form
1874+FVS1	r	F1	Second Isolate Form
1874+FVS1	Ţ,	х	Second Medial Form
1874+FVS2	٠	х	Third Medial Form
1874+FVS3	Ç	х	Fourth Medial Form
1874	ιζ	F2	First Final Form
1874+FVS1	9	Х	Second Final Form
1874+FVS2	Ù	Х	Third Final Form
1875	ک	F2	Final Form
1876+FVS1	r	Х	Second Initial Form
1876	t	F2	First Medial Form
1876+FVS1	r	Х	Second Medial Form
1880+FVS1	0	х	Second Form

1881+FVS1	₩	х	Second Form
1887+FVS1	,	х	Second Final Form
1887+FVS2)	Х	Third Final Form
1887+FVS3	7	х	Fourth Final Form
1887+FVS4	L,	I	Fifth Final Form
1888+FVS1	٢	Х	Second Medial Form
1888+FVS1	•	х	Second Final Form
188A+FVS1	ķ	х	Second Initial Form
188A+FVS1	۴,	Х	Second Medial Form
18A0+FVS1	Ý	Х	Second Initial Form
18A0+FVS1	Ý	х	Second Medial Form
18A1+FVS1	ಈ	х	Second Initial Form
18A1+FVS1	हिं	Х	Second Medial Form

Code-Point Sequence	Glyph	Code	Unicode Name/ Variant Name	Code Legend
180F	FVS4	II	Mongolian FVS4	F1 - Velar Family
182C+FVS1	Ç	F1	Second Isolate Form	F2 – Unicode Isolates
182C+FVS2	č	F1	Third Isolate Form	I – Proposed New Variants
182D+FVS1	c	F1	Second Isolate Form	II - Proposed New FVS
182D+FVS2	č	F1	Third Isolate Form	MM – FVS MisMatch
184D+FVS1	6	F1	Second Isolate Form	
184E+FVS1	c	F1	Second Isolate Form	
1863+FVS1	c	F1	Second Isolate Form	
1864+FVS1	۲	F1	Second Isolate Form	
1865+FVS1	C	F1	Second Isolate Form	
1874+FVS1	Ç	F1	Second Isolate Form	
1859	Ŋ	F2	Final Form	
1861	σ,	F2	Medial Form	
1862	Ē.	F2	Final Form	
1863	Ŋ,	F2	Final Form	
1873	Ç	F2	First Final Form	
1874	*	F2	First Final Form	
1875	ک	F2	Final Form	
1876	¥	F2	First Medial Form	
1824+FVS1	10	F2	Second Isolate Form	
1826+FVS2	105	F2	Third Isolate Form	
184A	F	F2	Final Form	
184B	٩	F2	Final Form	

184F	4	F2	Final Form	
185D	<i>?</i>	F2	First Final Form	
185E+FVS1	7	F2	Second Medial Form	
1822+FVS1	J	I	Second Final Form	
1822+FVS2	77	I	Third Medial Form	
1822+FVS3	7	I	Fourth Medial Form	
1824+FVS1	٥	I	Second Final Form	
1824+FVS2	è	I	Third Final Form	
1826+FVS2	è	I	Third Final Form	
1826+FVS3	ь	I	Fourth Medial Form	
1828+FVS1	TBD	I	Second Isolate Form	
1828+FVS4	٢	I	Fifth Medial Form	
182C+FVS2	c	I	Third Initial Form	
182D+FVS2	٠	I	Third Initial Form	
1843+FVS1	L	I	Second Final Form	
1887+FVS4	IJ	I	Fifth Final Form	
1820+FVS2	1	MM	Third Medial Form	
1828+FVS2	> ;	MM	Third Medial Form	
182C+FVS2	4	MM	Third Medial Form	
182C+FVS3	` ?:	MM	Fourth Medial Form	
182D+FVS2	Ķ	MM	Third Medial Form	
1835+FVS1	C	MM	Second Medial Form	
1836+FVS2	C	MM	Third Medial Form	
1820+FVS1	ŋ	Х	Second Isolate Form	
1820+FVS1	т	Х	Second Medial Form	
1820+FVS1	י	Х	Second Final Form	

1821+FVS1	1+	х	Second Initial Form
1821+FVS1)	х	Second Final Form
1822+FVS1	7	х	Second Medial Form
1823+FVS1	ط	Х	Second Medial Form
1823+FVS1	O	Х	Second Final Form
1824+FVS1	٦	х	Second Medial Form
1825+FVS1	₩.	х	Second Medial Form
1825+FVS1	6	х	Second Final Form
1825+FVS2	·στ	х	Third Medial Form
1826+FVS1	16	х	Second Isolate Form
1826+FVS1	σ-	х	Second Medial Form
1826+FVS1	6	х	Second Final Form
1826+FVS2	·στ	х	Third Medial Form
1828+FVS1	1	х	Second Initial Form
1828+FVS1	7	х	Second Medial Form
1828+FVS3	Ţ	х	Fourth Medial Form
182A+FVS1	Б	х	Second Final Form
182C+FVS1	ı;	х	Second Medial Form
182C+FVS2	Ω	х	Third Initial Form
182D+FVS1	Ω	х	Second Initial Form
182D+FVS1	ı.	х	Second Medial Form
182D+FVS1	7	Х	Second Final Form
182D+FVS2	5	х	Third Final Form
182D+FVS3	Ç	Х	Fourth Medial Form
182D+FVS4	г	х	Fifth Medial Form
182D+FVS2 182D+FVS3	ر ب	x x	Third Final Form Fourth Medial Form

		·
✓	Х	Second Final Form
予	х	Third Final Form
4	х	Second Medial Form
চ	х	Second Initial Form
চ	х	Second Medial Form
વ	х	Second Final Form
7	х	Second Initial Form
7	х	Second Medial Form
Ф	х	Second Medial Form
6	х	Second Final Form
٦	Х	Second Medial Form
Ä	х	Second Medial Form
đ	х	Second Medial Form
Ίų	х	Second Isolate Form
प्	х	Second Medial Form
6	х	Second Final Form
Ф	х	Third Medial Form
क्	х	Second Medial Form
10	х	Second Isolate Form
đ	Х	Second Medial Form
۲۰	Х	Second Medial Form
٦	Х	Second Medial Form
1	Х	Second Final Form
1	х	Second Final Form
7	Х	Third Medial Form
	1 < 1	ト X X X X X X X X X X X X X X X X X X X

185E+FVS2	,	х	Third Final Form	
	,	^		
1860+FVS1	ь	Х	Second Medial Form	
1860+FVS1	ଚ	Х	Second Final Form	
1860+FVS2	Ď	Х	Third Final Form	
1863+FVS1	Ŀ	Х	Second Medial Form	
1868+FVS1	đ	х	Second Initial Form	
1868+FVS1	ما	х	Second Medial Form	
1868+FVS2	Ь	х	Third Medial Form	
1869+FVS1	वं	х	Second Initial Form	
1869+FVS1	۵۱,	х	Second Medial Form	
186F+FVS1	Л	х	Second Initial Form	
186F+FVS1	ъ	х	Second Medial Form	
1873+FVS1	7	х	Second Medial Form	
1873+FVS1	1	х	Second Final Form	
1873+FVS2	77	х	Third Medial Form	
1873+FVS2	•	х	Third Final Form	
1873+FVS3	7	х	Fourth Medial Form	
1873+FVS3	4	Х	Fourth Final Form	
1874+FVS1	ı,	х	Second Medial Form	
1874+FVS1	9	х	Second Final Form	
1874+FVS2	r	Х	Third Medial Form	
1874+FVS2	ÿ	х	Third Final Form	
1874+FVS3	¿	Х	Fourth Medial Form	
1876+FVS1	b	Х	Second Initial Form	
1876+FVS1	b	х	Second Medial Form	

1880+FVS1	0	х	Second Form	
1881+FVS1	<₩	х	Second Form	
1887+FVS1	,	х	Second Final Form	
1887+FVS2)	х	Third Final Form	
1887+FVS3	ţ	х	Fourth Final Form	
1888+FVS1	٢	х	Second Medial Form	
1888+FVS1	•	х	Second Final Form	
188A+FVS1	ķ	х	Second Initial Form	
188A+FVS1	Ķ	х	Second Medial Form	
18A0+FVS1	Ý	х	Second Initial Form	
18A0+FVS1	Ý	х	Second Medial Form	
18A1+FVS1	d e	х	Second Initial Form	
18A1+FVS1	्राः	х	Second Medial Form	