

Improvement of Mongolian Unicode Block

ᠠᠨᠢᠯᠠᠯᠤᠯᠤᠰ
ᠠᠨᠠᠵᠢᠨ
ᠠᠨᠠᠵᠢᠨ
ᠠᠨᠠᠵᠢᠨ
ᠠᠨᠠᠵᠢᠨ
ᠠᠨᠠᠵᠢᠨ

CITA-WG, Ulaanbaatar 2019



Contents

- Format Controls - Improvement
- Reason that why is decided to use special format control MVS
- New code points are required for Todo and Manchu-Sibe

د استقبیل

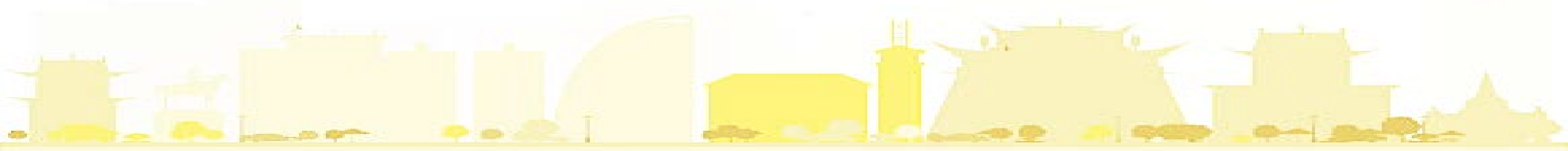
- مہینوں کی تعداد میں اضافہ
- مہینوں کی تعداد میں اضافہ - مہینوں کی تعداد میں اضافہ
- مہینوں کی تعداد میں اضافہ - مہینوں کی تعداد میں اضافہ
- مہینوں کی تعداد میں اضافہ - مہینوں کی تعداد میں اضافہ



Format Controls - Improvement

- MVS, **NNBSP**, FVS1, FVS2 and FVS3
- MVS, FVS1, FVS2 and FVS3
(reduced the quantity of format controls)
- MVS – used not as a Vowel Separator, but as a Suffix Separator
(solved problems caused by NNBSP)

• - /



- MVS **ᠮᠤᠨᠰᠤ** – **ᠮᠤᠨᠰᠤ** **ᠲᠤᠮᠤᠭ᠎ᠠ** **ᠠᠯᠢᠮᠠ** **ᠪᠠᠢᠨᠠ** **ᠬᠦᠷᠬᠢᠷᠢᠷᠢ** **ᠶ᠋ᠠ**

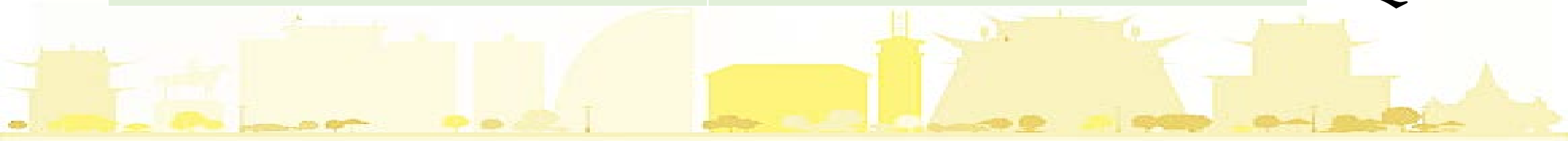
MVS – never used as a Vowel Separator!

Standard Font	Custom New Font
<p>байин᠎ᠠ, абаг᠎ᠠ, агулжай᠎ᠠ, алим᠎ᠠ, хүрхир᠎ᠢᠷᠢᠷᠢᠶ᠋ᠠ</p>	<p>байина, абага, агулжайа, алима, хүрхирэ, аха,</p>
<p>MVS is used as vowel separator</p>	<p>There is no any format control used as vowel separator.</p>

MVS –used as a Suffix Separator only!

Standard Font	Custom New Font
<p>НУТУГ~УН, ХҮМҮС~И, ЖИЛАГ~а~ЙИ, НИГҮР~ТЭЙ, • • •</p>	<p>НУТУГУН, ХҮМҮСИ, ЖИЛАГАЙИ, НИГҮРТЭЙ, • • •</p>
<p>NNBSP is used as Suffix Separator</p>	<p>No NNBSP, but MVS used instead of NNBSP as Suffix Separator</p>

- MVS - NNBSP



Reason: Why it is decided to use special format control MVS

- Mongolian, Todo and Manchu-Sibe scripts has one common character or code point for a letter “a - ᠠ” i.e. U+1820.

ᠠᠨᠠᠨᠠᠨ ᠠᠨᠠᠨᠠᠨ ᠠᠨᠠᠨᠠᠨ ᠠᠨᠠᠨᠠᠨ ᠠᠨᠠᠨᠠᠨ U+1820
ᠠᠨᠠᠨᠠᠨ «ᠠ» ᠠᠨᠠᠨᠠᠨ ᠠᠨᠠᠨᠠᠨ ᠠᠨᠠᠨᠠᠨ
ᠠᠨᠠᠨᠠᠨ ᠠᠨᠠᠨᠠᠨ ᠠᠨᠠᠨᠠᠨ ᠠᠨᠠᠨᠠᠨ ᠠᠨᠠᠨᠠᠨ
ᠠᠨᠠᠨᠠᠨ :

Reason: Why it is decided to use special format control MVS

Mongolian		Todo	Manchu-Sibe
U+1820	U+1821	U+1820	U+1820
ᠰ	ᠰ	ᠰ	ᠰ
ᠰᠢᠰᠦ	ᠰᠢᠰᠦ	ᠰᠢᠰᠦ	ᠰᠢᠰᠦ

ᠨᠠᠮᠤᠴᠢ ᠰᠢᠪᠦ ᠮᠠᠨᠤ ᠮᠣᠩᠭᠣᠯ ᠤᠯᠤᠰ

• ᠮᠠᠨᠤᠰᠢᠪᠦ ᠰᠢᠪᠦ ᠮᠠᠨᠤᠮᠣᠩᠭᠣᠯ ᠤᠯᠤᠰ ᠨᠠᠮᠤᠴᠢ ᠰᠢᠪᠦ ᠮᠠᠨᠤ ᠮᠣᠩᠭᠣᠯ ᠤᠯᠤᠰ ᠨᠠᠮᠤᠴᠢ ᠰᠢᠪᠦ ᠮᠠᠨᠤ ᠮᠣᠩᠭᠣᠯ ᠤᠯᠤᠰ



Pros and Cons

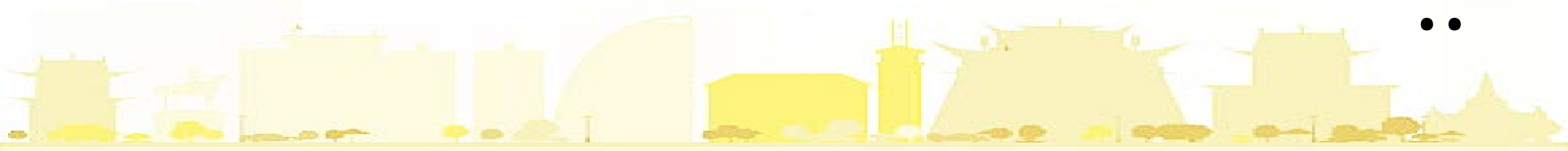
- **Pros.**

- Quantity of Format Controls is reduced
- Problems caused by NNBSF is solved
- Searching of words is becoming easy

- **Cons**

- If decided the editing of existing data according to the new rule of font then needs some volume of works should be done.

ଅନୁରୋଧ କରାଯାଉଛି ଯେ ଏହି ନିୟମାବଳୀକୁ ଗ୍ରହଣ କରାଯାଏ :



How to migrate existing data to new rule

- **It is very easy.**

1. Replace all occurrences of MVS within whole text with empty string, i.e. delete all MVS.
2. Replace all NNBSF with MVS.
3. May be some small editing is required

هل هو ممكن؟
■ في حقهم يا سيدي!



શ્રીગણેશાય નમઃ
શ્રીગણેશાય નમઃ
શ્રીગણેશાય નમઃ

Thanks for your attention!

