

July 10, 2015

[illegible]

Others?

That is



My discussion on the possibility of replacement of NNBSpace with other character:

1. The NNBSpace is a space, not control character as it is defined in Unicode. we need here is one Mongolian Suffix Joiner (special control character) actually.
2. Some of the system software handles NNBSpace as a space (narrow non-break space) and ignore or filters out this character before come into MS-USE or Harfbuzz. In that case, we lost the important control characters in the text and could not get correct Mongolian forms actually. For example when we copy text from adobe PDF reader, the text will lost the NNBSpace.
3. NNBSpace had been filtered by the previous version browser like chrome and safari till last year and Mongolian text all get mistaken forms related with NNBSpace.
4. If it is possible to replace it with other character to define the function, I would like to propose to use U+180F, we can name it as Mongolian Suffix Joiner

Greg Eck's Response – Jul 8 2015

Another possible name is Mongolian Suffix Separator to pair it with the Mongolian Vowel Separator. Both actually join two parts of a word. Both are not spaces. Both are non-joining in the terms of ArabicSpacing.txt. One value of excluding the word "join" from the name is that it might reduce future confusion - a person might ask the question why is called a "Joiner" if it is actually "non-joining". We don't want to include the term "space" as it is not a space. If we use the term "separator" although it is still deficient, it does match the MVS terminology.

Mongolian Vowel Separator, Mongolian Suffix Separator.

Just a thought, but then we are getting ahead of ourselves.

I agree this, the “Mongolian Suffix Separator” is better than “Mongolian Suffix Joiner”.

My discussion on STANDARD NNBSM MONGOLIAN MODEL:

1. What is the NNBSM definition? According to Unicode definition, it is Narrow No-Break Space, narrow form of a no-break space, typically the width of a thin space or a mid space.
2. If we use it as Mongolian Suffix Joiner (no-break space), we should add the definition into the Unicode standard. In the case, all of system or developer will pay their attention to this is used in Mongolian and have some special definition not just a narrow no-break space.
3. When we use NNBSM in Mongolian, we should include the usage of the Character.
 - ① Is it only used for Mongolian Suffix? If so, can we define full collection of the Mongolian Suffix? Or We just consider the followed word after NNBSM is suffix ?

Greg Eck's Response – Jul 8 2015

I suggest that such a character be used only as the suggested name implies to "separate suffixes" or "join suffixes" to the stem or to each other. If there is another function that the character carries, then it should be reflected in a more generic name.

I think it should be up to the font designer to decide how tight he/she wishes to restrict the characters following this potential new character.

I agree this in this point. I will explain in the following content and get back to here.

- ② Is it can be used as other purpose in Mongolian like other Language. For example, we just use NNBSM as a "Narrow no-break Space", it can be used as connection of two words we do not want to separate them to two line or just want handle the two word as one word.

In this case, it is better to consider our NNBSM model will not impact our normal word forms. It is just impact the specific Mongolian Suffix.

Almas Font s are considered this possibility and our Font handle the Suffix only have been changed their first letter in special rule (Suffix Rule) and other followed normal words will not be changed the first letter.

For example:

Our Font shows like bellow

ᠠᠨᠠᠨᠠᠨ -- (1828+1823+182E+NNBSM+1824+1828)

ᠠᠨᠠᠨᠠᠨᠠᠨ -- (1828+1823+182E+NNBSM+1824+1837+1822+182C+1824)

But Mongolian Baiti shows this like bellow

לש

סוף סדר



I would like to ask all members to consider remaining this possibility for the further usages. The other languages are using this character as this kind of usage as its definition.

Greg Eck's Response – Jul 8 2015

I do not consider this to be a mistake, but rather a difference in implementation design. One font designer has decided to tightly restrict the characters following the NNBS. Another font designer has decided to form the grammatically correct suffixes correctly but not spell-check the text after the NNBS.

I consider these decisions to be independent of the model and left to the choices of the font designer.

I don't agree this in this point as well. Let me include my opinion in the following discussion.

Maybe the example listed above is not sufficient to explain what I want to discuss with all.

Additional example.

Our Font shows like bellow

לש לשון

מִיָּסָד וְעַד הַיּוֹם

But Mongolian Baiti shows this like bellow

לש לש

ॐ नमो भगवते वासुदेवाय

The (1828+1823+182E+NNBSP+1824+1837+1824+182D+1824) is Mongolian suffix listed in DS05. But it is not displaying same in our font and Microsoft Baiti Font (It is hard to be accepted)

The solution will be listed in the following contents.

My discussion on MONGOLIAN NNBS-CONNECTED SUFFIXES:

1. I have checked the DS05 listed Mongolian Suffixes list. It is covered most of the Mongolian Suffix well, compared with the our suffixes list find a little bit difference. For example, we included like bellow

ᠠᠨᠢᠨᠠᠨᠠᠨᠠ

But I think it is Ok for us if we list out the irregular first letter writing suffixes, it is enough for this Model document. Others will follow the normal word glyph rendering rules applied.

Greg Eck's Response – Jul 8 2015

Our grammar considers these forms to be pronouns and therefore not connected by the NNBS. I believe this is the same view as espoused by Professor Quejingzhabu also. Are there other words that are questionable?

Oh sorry, It is not pronoun actually, if it is spelled as 'ᠠᠨᠢᠨᠠᠨᠠᠨᠠ' it is pronoun. This three is Mongolian Suffix, and it is what we had been asked by the users to add to our Suffix on our Mongolian IME. The name of the Suffix I will confirm with Linguistic Expert and come back to you.



2. For the listed suffixes, would you please list out the Unicode in the Notes. Because, we find there significant difference in the encoding some of the suffixes. For example, we find following encoding difference

ᠠᠨᠢᠨᠠᠨᠠᠨᠠ

NNBS + 1822 + 1828

NNBS + 1822 + 1822 + 1828

NNBS + 1835 + 1822 + 1828

NNBS + 1836 + 1822 + 1828

We should define which is the correct encoding.

Greg Eck's Response – Jul 8 2015

Yes, good point.

I have modified DS05 with the code-points explicitly stated as suggested. There may be variation and differences of opinion on spellings. Please state differences so that we can discuss further. Are there other suffixes that others are processing that are not included on DS05?

Thank you very much for your understanding. I will my discussion for updated DS05 in following lines.

3. For my personal consideration is that we just define the listed suffixes encoding rule and will not impact other words. What I mean that we will not define the rule of single character display form rule followed by NNBSF.

4. >Our grammar considers these forms to be pronouns and therefore not
> connected by the NNBS. I believe this is the same view as espoused
> by Professor Quejingzhabu also. Are there other words that are questionable?

I have not listed all of the Suffixes what we are handling here, it is because we have just considered that the Mongolian linguistic Experts have their own consideration of which is suffix and which is not suffix. We do not need to roll into the endless argument of the linguistic opinion here. For example,

which is used in $\mathcal{M}_{\text{DS05}}$ is not included in DS05,



and some part says it is suffix and some part says it is not suffix. But it is really need a suffix character to handle it. But I can list out our additional suffix list here, but it is still not enough for all of the Mongolian Suffixes required by linguistic peoples.

What I want say here is that we need to declare **the basic rule of Mongolian Encoding**.

1. *The encoding rule follows the modern normal and publicly accepted display form comes first.*
2. *Abnormal display forms which are using in modern Mongolian will come to second. What I mention here modern Mongolian include the modern Todo and Shibe and Manchu.*
3. *Abnormal display forms which are not using in modern Mongolian will come to third.*
4. *Should consider the entire special requirement from linguistic experts and leave the implementation possibility for all of these requirements.*

PS. Please add your consideration here for getting quick solution for following part of the discussion includes the linguistic argument of which is which. Maybe Unicode have exact definition of this. If anyone know of that please let us know.

5. Let me list out our suffix list addition to DS05. The suffix name I can provide after refer to linguistic expert if necessary or appreciate anyone can contribute to this.

བྱོ་ ---- ལྟོ་ བྱོ་ བྱོ་ (202F+1824+1824, 202F+1826+1826) 
 ལྟོ་ ---- ལྟོ་ ལྟོ་ (202F+1828+1822) 
 ལྟོ་ ---- ལྟོ་ ལྟོ་ (202F+182E+1822+1828+1822) 
 ལྟོ་ ---- ལྟོ་ ལྟོ་ (202F+1834+1822+1828+1822) 
 ལྟོ་ ལྟོ་ - ལྟོ་ (202F+1833+1824+1828+1822, 202F+1833+1826+1828+1822) 
 ལྟོ་ ལྟོ་ - ལྟོ་ (202F+1832+1824+1828+1822, 202F+1832+1826+1828+1822) 
 ལྟོ་ ---- ལྟོ་ ལྟོ་ (202F+1833+1820, 202F+1833+1821) 
 ལྟོ་ ---- ལྟོ་ ལྟོ་ ལྟོ་ (202F+182F+1820, 202F+182F+1821) 
 ལྟོ་ ---- ལྟོ་ ལྟོ་ ལྟོ་ (202F+1830+1820+1828, 202F+1830+1821+1828) 
 ལྟོ་ ---- ལྟོ་ ལྟོ་ (202F+1834+1824, 202F+1834+1826) 
 ལྟོ་ ---- ལྟོ་ ལྟོ་ (202F+1830+1822+182D) 
 ལྟོ་ ---- ལྟོ་ ལྟོ་ (202F+1830+1822+182D+FVS2) 
 ལྟོ་ ---- ལྟོ་ (202F+1833+1820+182D) 
 ལྟོ་ ---- ལྟོ་ (202F+1833+1820+182D+FVS2) 
 ལྟོ་ ---- 2 ལྟོ་ (202F+1833+1824+182D+1820+1837) 
 ལྟོ་ ---- 1 ལྟོ་ (202F+1833+1826+182D+1821+1837) 

5. Let me bring following discussion here for more detailed discussion.

Our Font shows like bellow

ལྟོ་ ལྟོ་ -- (1828+1823+182E+NNBSP+1824+1828)

ལྟོ་ ལྟོ་ ལྟོ་ -- (1828+1823+182E+NNBSP+1824+1837+1824+182D+1824)

This is the correct display according to DS05 document.

But Mongolian Baiti shows this like bellow

ལྟོ་ ལྟོ་ -- (1828+1823+182E+NNBSP+1824+1828)

ལྟོ་ ལྟོ་ ལྟོ་ -- (1828+1823+182E+NNBSP+1824+1837+1824+182D+1824)

What I want to say here is that we should not just check the following character after NNBSP (or MSS) to select the form of the character. We should check if it is followed by the specific Suffixes or not and change the form of the character display. We will not impact first letter of other words except Suffixes.



What I want to say here is that the NNBS (or MSS) will only need to distinguish and handle the following display forms for 1824 or 1826 and display forms for 1836 and 1822 as well as display form of 1820. All of the other words followed after NNBS (or MSS) will be displayed same with original normal words form.

𐌲 (202F+ 1820)



𐌲 (202F+ 1821)

𐌲 (202F+ 1824)

𐌲 (202F+ 1826)

𐌲𐌶 (202F+ 1824+1828)

𐌲𐌶 (202F+ 1826+1828)

𐌲𐌶𐌵 (202F+ 1824+1833)

𐌲𐌶𐌵 (202F+ 1826+1833)

𐌲𐌶𐌵 (202F+ 1824+1824) 

𐌲𐌶𐌵 (202F+ 1826+1826) 

𐌲𐌶𐌵 (202F+ 1836+1822+1828)

𐌲𐌶𐌵𐌶 (202F+ 1822+1836+1820+1837)

𐌲𐌶𐌵𐌶 (202F+ 1822+1836+1821+1837)

𐌲𐌶𐌵𐌶 (202F+ 1822+1836+1820+1828)

𐌲𐌶𐌵𐌶 (202F+ 1822+1836+1821+1828)

ᠠᠨᠢ (202F+ 1836+1822)

ᠠᠨ (202F+ 1822)

All of the other Mongolian suffixes after NNBSP (or MSS) will be displayed as their original/normal forms and it is constructed with their prefix word together to be just like one word.

5. Let me list some suffixes what I want to discuss from the DS05 here.



ᠠᠨᠢᠨᠠᠨ (202F+ 1836+1822+1828)

ᠠᠨᠠᠨᠠᠨ (202F+ 1822+1836+1820+1837)

ᠠᠨᠠᠨᠠᠨ (202F+ 1822+1836+1821+1837)

ᠠᠨᠢᠨᠠᠨ (202F+ 1822+1836+1820+1828)

ᠠᠨᠢᠨᠠᠨ (202F+ 1822+1836+1821+1828)

ᠠᠨᠢ (202F+ 1836+1822)

Our fonts and our IME are same with this. Anyone have any different spells for these suffixes?

