

Summary of alternative encoding solutions to resyllabified compound words






To: MWG #3 (Unicode Mongolian Working Group Meeting 3), Ulaanbaatar
 From: Liang Hai / 梁海 <lianghai@gmail.com>
 Date: 5 April 2019

This document summarizes solutions mentioned on the first two days of MWG #3, about how the encoding of resyllabified compound words in Hudum can be improved. See Table 1 for comparison of how various structures affected by the stem-boundary are encoded in the three solutions.

- **Current:** The de facto standard today for comparison. May refer to MWG/3-N8R for introduction.
- **Shen Yilei:** Introduced in MWG/3-N9R4.
- **Bolorsoft:** Introduced in MWG/3-N7R2, implemented in released products. (The author referred to MongolianScript, version 4.00 for details not clarified in N7R2.)



There was a suggestion about using nirugu (U+180A MONGOLIAN NIRUGU), but no clarification about the exact behavior was provided. B. Enkhdalai also suggested that Inner Mongolian users could consider a change of orthography, using a hyphen to join the two parts in a name instead.

Table 1. Comparison

| | Pre-boundary Coda form of <i>n/g/d</i> | Boundary Invisible ↓ | Post-boundary Extra aleph in <i>a/i/o/u/ö/ü</i> | Post-boundary Extra yodh in <i>ö/ü</i> | Post-boundary Extra yodh in post-consonant <i>ö/ü</i> |
|-------------------|---|---|---|---|---|
| |  |  |  |  |  |
| Current | FVS | — | FVS | | FVS |
| Shen Yilei | contextual | — | SSBM | FVS | |
| Bolorsoft | contextual | ZWJ | contextual with out-of-set variants | | FVS |

See Figure 1 for how sequences of *-nö-* and *-nmö-* are usually written in words when a stem-boundary is not involved:

Figure 1. Ordinary written forms

 
-nö- *-nmö-*

* EOF *