

**Title:** Proposal to Change Category of U+02D7 from Sk to Lm  
**Author:** Martin Hosken  
**Action:** For consideration by UTC  
**Date:** 2011-04-18

**Proposal:** This document proposes to change the general category of U+02D7 MODIFIER LETTER MINUS SIGN from Sk to Lm

**Rationale:** There are a number of writing systems (ium-Thai, ptn-Thai, pdu-Latn) that use a hyphen character between syllables in a polysyllabic word. There is need therefore for a word extending hyphen type character. There are a number of options to addressing this need:

1. Encode new characters within the script blocks of the scripts using them.
2. Encode a single new character as a pan-script word extending hyphen.
3. Use an existing character (U+02D7 is proposed) but changes its properties to be word extending
4. Change the general category of U+02D7 as the most consistent approach to changing its word break property to word extending.

There are a number of hyphen characters already in the standard and adding more is liable to add to the confusion rather than to lessen it, especially due to any such characters being added being visually identical to 002D HYPHEN.

The choice of U+02D7 is appropriate in that its name indicates it as being a modifier letter, which while stretching interpretation, does at least imply it is a letter. Just as U+02BC MODIFIER LETTER APOSTROPHE is used in a number of languages as a full letter, U+02D7 can be thought of as a full letter between syllables. While U+02D7 can take an alternative glyph shape, this is a stylistic issue that is only of interest to technical use of the character. In its use between syllables, a user might say that a small serified dash would be stylistically inappropriate, but still legible.

The difficulty with using U+02D7 is that it is not word extending. In fact, there are no word extending hyphens or full minus signs in the standard. U+02D7 is the most appropriate character to be changed to be word extending. Even its current usage statement implies that it is associated with a previous word extending character and therefore should itself be word extending. The easiest way to achieve this is to change its general category from Sk to Lm. This would bring it into line with other similar modifier letters (e.g. U+02BC MODIFIER LETTER APOSTROPHE, 02C6 MODIFIER LETTER CIRCUMFLEX ACCENT) which are based on lower ASCII punctuation marks being used as letters.

U+02D7 is not known to be used other than in IPA, where changing it from Sk to Lm would have only a positive impact. Therefore there is currently no known negative impact on this change on existing implementations. U+02D7 has not undergone the Sk/Lm flipflop that other characters have gone through, and previous discussions have rarely, if ever, considered this character, apart from as part of a block of technical characters.

#### **Examples:**

Use of syllable breaking hyphen found in Pekon Kayan (pdu-Latn):

kukei dō tajā. 1  
dō mjè khu, dō 1  
ă ta-è ta-ŏ alè

Use of syllable break in Patani Malay (ptn-Thai):

ตา-โอะ เบ็ด; taut.

ตีมอ วิด (ป่อ); menimba.

ตือบอ เบื่อ (ปลา); menuba.

แตแฮ ตา-โอะ ธงเบ็ด, ปักเบ็ด  
(เพื่อจับปลา); menahan taut.

### **Bibliography:**

Manson, Ken. 2006. Pekon Kayan orthography. Chiang Mai: Payap University.

สำนักงานกองทุนสนับสนุนการวิจัยพจนานุกรม: ภาษามลายูปาตานี-ไทย-มาเลเซีย

ฉบับเพื่อการศึกษาและการสื่อสารเบื้องต้น

Lamnakngankongtun snapsnun ganwitjay Pojnanugrom: Pasa malayupatani-Thai-Malesia chabob  
pua gansuksa le gansuasanbuangdan