# Proposal to Update Properties for Two Khmer Characters

To: UTC via Script Ad Hoc Status: Individual Contribution Date: 2023-04-28

Authors: Steven R. LOOMIS <<u>srl295@codehivetx.us</u>> (Individual Unicode Member)

## 1. Summary

This is a request for modifications to the sentence break properties of two Khmer Characters.

- U+17D4 1 KHMER SIGN KHAN
- U+17D5 1 KHMER SIGN BARIYOOSAN

## 2. Proposal

Add to PropList.txt

```
# PropList.txt
17D4..17D5 ; Sentence_Terminal # Po [2] KHMER SIGN KHAN..KHMER SIGN
BARIYOOSAN
```

Which will cause an update to SentenceBreakProperty.txt

# SentenceBreakProperty.txt
17D4 ; STerm # Po KHMER SIGN KHAN
17D5 ; STerm # Po KHMER SIGN BARIYOOSAN

## 3. Background

Note: This issue was originally filed as <<u>https://unicode-org.atlassian.net/browse/CLDR-16220</u>> by Mr. Makara SOK of SIL.

#### 3.1 NamesList

The NamesList already indicates the functions of these characters.

17D4 KHMER SIGN KHAN
 \* functions as a full stop, period
17D5 KHMER SIGN BARIYOOSAN
 \* indicates the end of a section or a text

### 3.2 Example Text

This example text contains multiple sentences, with a U+17D4 (1) at the end of each sentence, and a U+17D5 (1) at the end of the chapter. There two three variations for 1:

- 1. ปี
- 2. Tem
- 3. <sup>1</sup>10w

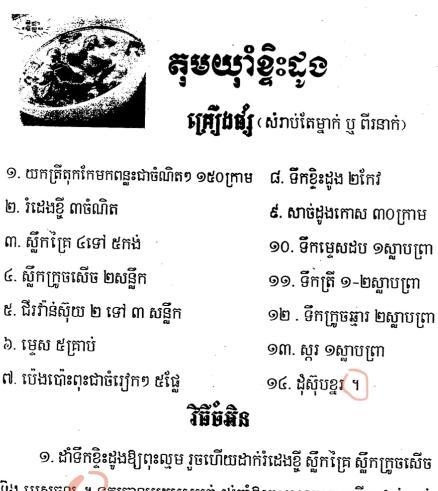
YouVersion	Search Bible.com	Get the app	$\bigoplus$	ΞQ			
ម៉ាថាយ 1 	- GKH	В		•			
<i>គេនឹងហៅ៖</i> (អេម៉ាញូអែ <sup>24</sup> យ៉ូរែ គាត់ គឺគាត់	<i>ស្ត្រីព្រហ្មចារិម្នាក់នឹងមានផ្ទៃពោះ រ</i> ទាមរបស់កូននោះថា អេម៉ាញូអែល" ល មានអត្តនំយប្រែថា ព្រះគង់នៅជ លបក្រោកពីដំណេក ក៏ធ្វើដូចដែលទ ទទួលនាងជាប្រពន្ធ <sup>25</sup> ប៉ុន្តែគាត់មិន ាងសម្រាលបានកូនប្រុសម្នាក់ <sup>==</sup> ។	។ 📟 ាមួយយើង)។ ទុតសួគិរបស់ព្រះអម្ចាស់បាន ទបានរួមដំណេកជាមួយនាន	នបង្គាប់ ងទេ				
	HMER BIBLE ©, Copyright © 20 <u>e About ព្រះគម្ពីរខ្</u> វែរសាកល	19 by Global Bible Initi	ative				
https://www.bible.com/bible/2287/MAT.1.GKHB							
	Explore ម៉ាថាយ 1 l	y Verse					

Example 1 https://www.bible.com/bible/2287/MAT.1.GKHB

Example 2 shows U+17D5 at the end of a chapter, plus 0<sup>th</sup> showing the end of a book.

#### Proposal to Update Properties for Two Khmer Characters

YouVersion	Search Bible.com	C	Get the app	⊕	≡2
យ៉ូហ	ាន ទី៣ 1	► GKHB	-	-	
	ពាក្យជម្រាបលា				
<sup>13</sup> ខ្ញុំមានសេចក្តីជាច្រើនទៀតដែលត្រូវសរសេរមកអ្នក ប៉ុន្តែខ្ញុំមិនចង់សរសេរ មកអ្នកដោយទឹកខ្មៅ និងប៉ាកកាទេ។ <sup>14</sup> ផ្ទុយទៅវិញ ខ្ញុំសង្ឃឹមថានឹងមកជួបអ្នកភ្លាម ១ ហើយនិយាយគ្នាផ្ទាល់មាត់។ <sup>15</sup> សូមឲ្យមានសេចក្តីសុខសាន្តដល់អ្នក! ពួកមិត្តសម្លាញ់របស់យើង ក៏ផ្កាំ សួរសុខទុក្ខអ្នកដែរ។ សូមអ្នកជួយសួរសុខទុក្ខមិត្តសម្លាញ់របស់យើងម្នាក់១ផង៕៚ GLOBAL KHMER BIBLE ©, Copyright © 2019 by Global Bible Initiative Learn More About ព្រះគម្ពីរខ្មែរសាកល					
<b>Explore</b> ឃ៉ូហាន ទី៣ <b>1 by Verse</b>					
< http://www.htt	ttps://www.bible.co	om/bible/2287/3JN.1	.GKHB		>
	Example 2 <u>https://w</u>	ww.bible.com/bible/2287	//3JN.1.GKHB		



១. ដាទកខ្ទះដូងឱ្យពុះល្មម រួចហើយដាក់រំដេងខ្ចី ស្លឹកគ្រៃ ស្លឹកក្រូចសើច រឹង ម្ទេសចូល ។ ទុកចោលមួយស្របក់ រង់ចាំឱ្យពុះស្រួលបួល ហើយដាក់សាច់ ព្រី ប៉េងប៉ោះ និងសាច់ដូងចូលបន្ថែម ។

២. បន្ថែមរសជាតិដោយស្ករ ទឹកម្ទេស ទឹកក្រូចឆ្នារ ទឹកត្រី និង រៀស៊ីបន្តរ ។ ភ្លូក្សរសជាតិសាកល្បងមើល ដួសដាក់ចាន រោយស្លឹកជីរវ៉ាន់ស៊ុយពី ហើ ហើយទទួលទានក្តៅ១មានរសជាតិឈ្ងួយឆ្ងាញ់បំផុត ។

**ចំណាំ :** រយ:ពេលចំអិន ១៥នាទី🕐





Example 3, a scanned Khmer book showing end of paragraph marks

## 4. Other Consideration

4.1: ATerm was considered, but ATerm is a fixed list in <u>UAX#29</u> and seems to only interact differently than STerm for numeric and upper/lower letters.

4.2: The codepoint U+17D8 1001 discouraged in the names list. But, the sequence 1001 (1+00+1) means "etcetera" and so break rules may need to be evaluated. N is already handled as complex text.

Sometimes written -໙- but with ຳ໙ຳ at the end of a sentence. The document below shows that there are two variations of ຳ໙ຳ. It can be either ຳ໙ຳ or ຳໜຳ.

(LOL) E (LDL) (LOL) E (LDL) (LOL) E (LDL) (៤) ប្រយុរ្ភល (១០១) ឬ(១០១) ១ណូសញ្ញាមួយបែប សម្រាប់បើបំព្រញាសេចក្តី หลังเจี้ ឬ เ/ อีรอเลีย, เป็รสสกุร จ ลุภมเนีย ชา ๆ เข ๆ (พรชา เข้) ค์ตาร, เป็รสหกุญ ตุรุจาม สลักรณ์เอรเซียร์เซีย์ (ๆ ณ ๆ) เอะ, ຈອນ:ເທົ່: ថេដាប់បេដាដែលទំយកទៅពៀតមានលើតមន គំសៀវភៅសរសេរ, សៀវភៅអាន, ដង់ញាតិពរ បនាត, ទ័កទៅ 9-05 តន៍ភាសាបាលី (ចិនរ ຄອາເກາະຄຳ:

4.3: The question was asked, are there other contexts where 1 is used as an abbreviation?

Makara SOK writes: "Not that I know of. Period (.) is used in abbreviation if a symbol is needed in between letters."

The conclusion is that there are no known contexts where I is used as an abbreviation.