Title: Proposal to Encode the Letters GYAN and TRA for Limbu in the UCS
Source: Anshuman Pandey (pandey@umich.edu)
Status: Individual Contribution
Action: For consideration by UTC and WG2
Date: 2011-01-14

1 Introduction

This is a proposal to encode two new characters in the Limbu block of the Universal Character Set:

GLYPH	CHARACTER NAME	GC	CCC	BIDI	MIRRORED
ᤖ	LIMBU LETTER GYAN	Lo	0	L	N
ᤗ	LIMBU LETTER TRA	Lo	0	L	N

Technically, these letters are ligatures that represent consonant conjuncts: ᤖ GYAN = <ᤕ JA + *virāma* + ᤣ YAN>; ᤗ TRA = <ᤓ TA + *virāma* + ᤠ RA>. However, as the encoding for Limbu is not based upon the *virāma* model, it is necessary to include these conjuncts as independent letters.

2 Background

There are three major historical varieties of Limbu, which may be labeled using the names of the individuals to whom their development is attributed: Śirijaṅgā, Cemajoṅ, and Subba. The Limbu script encoded in the UCS is based largely upon the script reformed by B. B. Subba in the 1970s. This version is adequate for writing contemporary Limbu, but is missing some letters found in the repertoires of Śirijaṅgā and Cemajoṅ.

Two such characters are the letters ᤖ GYAN and ᤗ TRA. These letters are part of a revised Limbu script defined by Īmāna Siṃha Cemajoṅ. They appear in charts published in his *Limbu-Nepali-English Dictionary* (1961) and *Kirat Grammar* (1970), excerpts of which are shown here in Figure 1 and Figure 2. In his *Grammar of Limbu*, George van Driem writes that as part of Cemajoṅ's expansion of Śirijaṅgā's original script, he added letters to accommodate 'Nepali characters', such as *jña* and *tra* (1987: xxv). When Subba began publishing textbooks in Limbu in the 1970s, he revised Cemajoṅ's script and removed the letters for *jña* and *tra* (van Driem 1987: 555). Thus, their usage was rendered obsolete in comtemporary Limbu.

These letters were discussed by Boyd Michailovsky and Michael Everson in their proposal to encode Limbu in the UCS, but were not proposed for inclusion because "it is not clear that they have ever actually been used in writing Limbu" (2002: 8). They do, however, propose the inclusion of obsolete characters, which are "necessary for publishing older texts." The letters GYAN and TRA should be considered for inclusion in this light. They may be obsolete and perhaps never used beyond script charts, but they are attested nevertheless.

Furthermore, these two letters appear in non-Limbu language books. They are shown in a chart of Limbu in a book about the Sunuwar, who belong to the Mahakiranti language community (see Figure 3). The Limbu script has been promoted as a pan-Kiranti script since the 1990s and these two letters may indeed be used in the local orthographies for other Kiranti languages.

3 Implementation Details

3.1 Character Names

The name GYAN is derived from Cemajoṅ's nomenclature. It aligns with Limbu ᤊ YAN, which is an element of the conjunct represented by GYAN. The name TRA adheres to UCS naming conventions for Indic scripts.

3.2 Allocation

The letters GYAN and TRA may be encoded at the code points U+191D and U+191E, respectively.

3.3 Collation

There is no information regarding collation for GYAN and TRA. The chart in Figure 1 shows TRA placed after VA and GYAN after HA. On the other hand, Figure 3 shows the letters placed in a sequence after HA. These order should not be considered as actual practice; Michailovsky and Everson write that the locations may be chosen simply "to fill out the last two rows of alphabet tables" (2002: 8).

Logical sort orders for the letters are as follows:

1. GYAN should be sorted after ᤈ JA.
2. TRA should be sorted after ᤋ TA. As it is an independent letter it may also be appropriate to sort it before the sequence <ᤋ TA + ◌ᤪ SUBJOINED LETTER RA>, which may be semantically and phonetically equivalent to TRA.

3.4 Annotation

In the UCS names list for Limbu, the entries for these characters should appear under a new heading to be named 'Consonant Ligatures'. Annotations for the entries may be added in order to describe the consonant sequences represented by each conjunct.

4 References

ईमान सिंह चेमजोङ् [Cemajoṅ, Īmāna Siṃha]. 1961 [2018]. ᤕᤠᤰᤌᤢᤱ-ᤐᤧᤏᤧ-ᤔᤡᤰᤑᤢ᤺ᤗᤠ-ᤐᤠᤴᤆᤧᤁᤠᤘᤠ *Yākthuṅ-pene-mikphu'lā pānchekavā* [लिम्बू-नेपाली-अँगरेजी शब्दकोश / Limbu-Nepali-English Dictionary]. काठमाण्डू: नेपाल एकेडेमी.

———. 1970 [2027]. ᤋᤢᤶ ᤕᤠᤰᤌᤢᤱ ᤜᤢᤵᤐᤠᤴ ᤛᤠᤵᤗᤠ *Tum yākthuṅ huppān sāplā*. [किरात ब्याकरण (लिम्बू) / The Kirat Grammar (Limbu)]. डार्जिलिंग: जसहांग मादेन.

van Driem, George. 1987. *A Grammar of Limbu*. Mouton Grammar Library, 4. Berlin: Walter de Gruyter & Co.

Michailovsky, Boyd and Michael Everson. 2002. "Revised proposal to encode the Limbu script in the UCS" (N2410 L2/02-055). February 5, 2002. `http://std.dkuug.dk/JTC1/SC2/WG2/docs/n2410.pdf`

प्रधान 'बमप', बुद्धिमान [Pradhāna 'Bamapa', Buddhimāna]. ed. 1999. किरातवंशी सुनुवार (मुखिया) कोइँच् (विविध खोज एवं संग्रह) [Sunuwar (Mukhiya) Koinch of Kirat Clan]. गान्तोक: सिक्किम सुनुवार (मुखिया) कोइँच् बु.

Figure 1: Excerpt from a chart of Limbu showing the letters ᤕ GYAN and ᤖ TRA (from Cemajoṅ 1961: 22).

Figure 2: Excerpt from a chart of Limbu showing the letters ᤖ GYAN and ᤘ TRA (from Cemajoṅ 1970: 14).

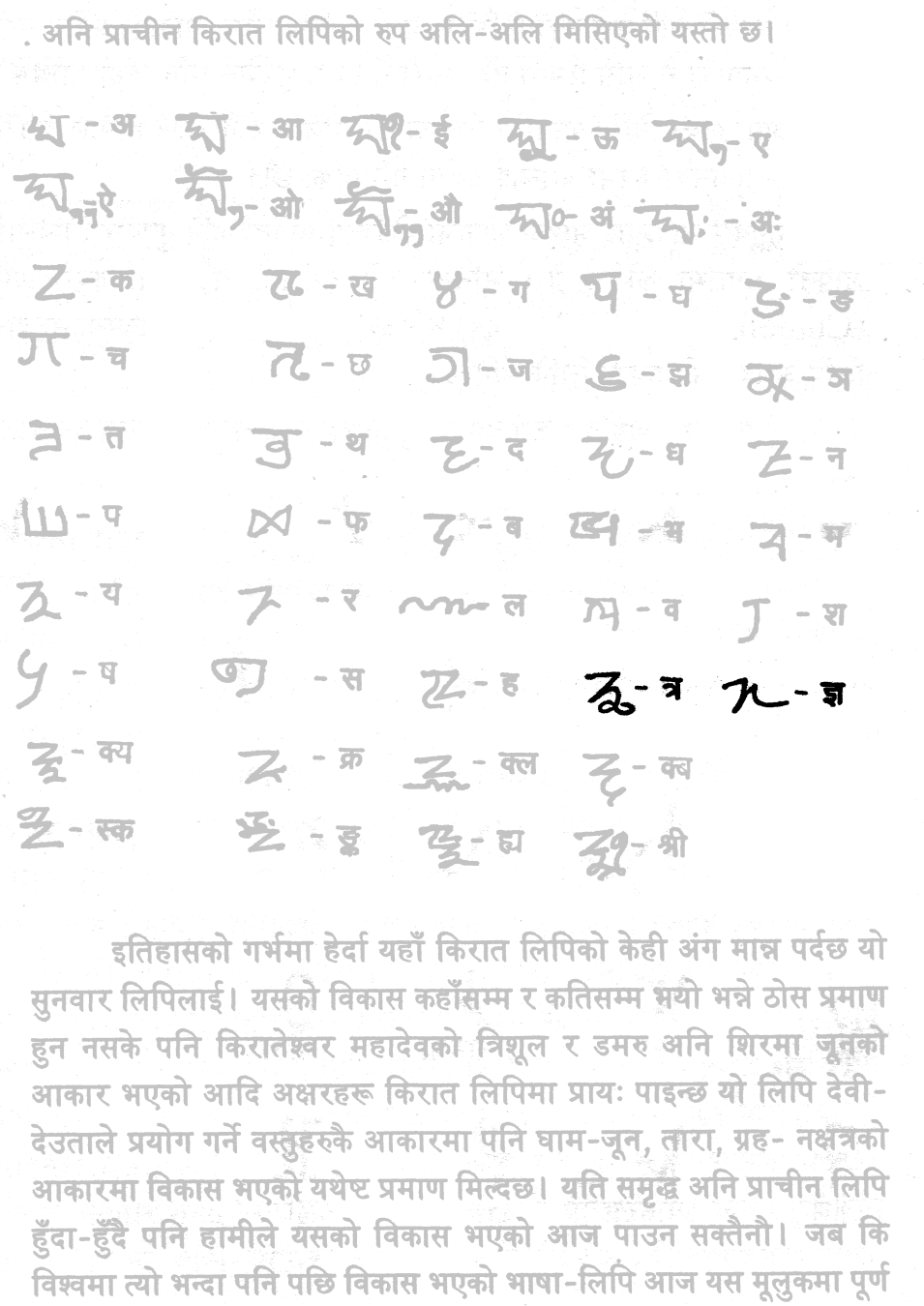

Figure 3: Excerpt from a chart of Limbu showing the letters ᤕ GYAN and ᤖ TRA (from Pradhāna 1999: 369).