

ISO/IEC JTC 1/SC 2/WG 2  
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS  
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646<sup>1</sup>

Please fill all the sections A, B and C below.  
Please read Principles and Procedures Document (P & P) from <http://www.dkuug.dk/JTC1/SC2/WG2/docs/principles.html> for guidelines and details before filling this form.

Please ensure you are using the latest Form from <http://www.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html>.  
See also <http://www.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html> for latest Roadmaps.

A. Administrative

1. Title: Revised Proposal to Encode Additional Latin Orthographic Characters  
2. Requester's name: UTC, Lorna A. Priest  
3. Requester type (Member body/Liaison/Individual contribution): Individual contribution  
4. Submission date: 31 October 2004 (revised 9 December 2004)  
5. Requester's reference (if applicable): L2/04-372R  
6. Choose one of the following:  
This is a complete proposal: Yes  
or, More information will be provided later: No

B. Technical – General

1. Choose one of the following:  
a. This proposal is for a new script (set of characters): No  
Proposed name of script:  
b. The proposal is for addition of character(s) to an existing block: Yes  
Name of the existing block: Latin Extended  
2. Number of characters in proposal: 5  
3. Proposed category (select one from below - see section 2.2 of P&P document):  
A-Contemporary x B.1-Specialized (small collection) B.2-Specialized (large collection)  
C-Major extinct D-Attested extinct E-Minor extinct  
F-Archaic Hieroglyphic or Ideographic G-Obscure or questionable usage symbols  
4. Proposed Level of Implementation (1, 2 or 3) (see Annex K in P&P document): 1  
Is a rationale provided for the choice? No  
If Yes, reference:  
5. Is a repertoire including character names provided? Yes  
a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document? Yes  
b. Are the character shapes attached in a legible form suitable for review? Yes  
6. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard? SIL International  
If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:  
7. References:  
a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided? Yes  
b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached? Yes  
8. Special encoding issues:  
Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)?  
Yes, suggested character properties are included, as are lower case representations (see section D)  
9. Additional Information:  
Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at <http://www.unicode.org> for such information on other scripts. Also see <http://www.unicode.org/Public/UNIDATA/UCD.html> and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

<sup>1</sup> Form number: N2652-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11)

### C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before? If YES explain _____	<u>No</u>
2. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)? If YES, with whom? <u>linguists</u> If YES, available relevant documents: <u>Email correspondence. See also References.</u>	<u>Yes</u>
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? Reference: <u>See comments in Section E</u>	<u>Yes</u>
4. The context of use for the proposed characters (type of use; common or rare) Reference: <u>Orthographic characters are used in literacy materials, liturgical books and general literature.</u>	<u>Common</u>
5. Are the proposed characters in current use by the user community? If YES, where? Reference: <u>See comments in Section E</u>	<u>Yes</u>
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP? If YES, is a rationale provided? If YES, reference: <u>If possible, should be kept with other related blocks in the BMP.</u>	<u>Preferably</u>
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)? <u>Preferably together with other related blocks</u>	
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence? If YES, is a rationale for its inclusion provided? If YES, reference: _____	<u>No</u>
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? If YES, is a rationale for its inclusion provided? If YES, reference: <u>(Cf. §F.1 of L2/04-047.)</u>	<u>Yes</u> <u>Characters with stroke or bar might be construed as precomposed forms of sequences with combining overlay marks.</u>
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character? If YES, is a rationale for its inclusion provided? If YES, reference: <u>See comments in Section E</u>	<u>Yes</u> <u>Yes</u>
11. Does the proposal include use of combining characters and/or use of composite sequences? If YES, is a rationale for such use provided? If YES, reference: _____ Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided? If YES, reference: _____	<u>No</u>
12. Does the proposal contain characters with any special properties such as control function or similar semantics? If YES, describe in detail (include attachment if necessary) _____	<u>No</u>
13. Does the proposal contain any Ideographic compatibility character(s)? If YES, is the equivalent corresponding unified ideographic character(s) identified? If YES, reference: _____	<u>No</u>

### D. Proposed Characters

A code chart and list of character names are shown on a new page.

**D.1. Proposed Characters**

	2C6
0	Ł
1	ł
2	Ł
3	Ṗ
4	Ṛ
5	
6	
7	
8	
9	
A	
B	
C	
D	
E	
F	

**D.1. Character Names**

U+2C60	LATIN CAPITAL LETTER DOUBLE-BARRED L
U+2C61	LATIN SMALL LETTER DOUBLE-BARRED L
U+2C62	LATIN CAPITAL LETTER L WITH MIDDLE TILDE <ul style="list-style-type: none"> <li>• lowercase is U+026B ł</li> </ul>
U+2C63	LATIN CAPITAL LETTER P WITH STROKE <ul style="list-style-type: none"> <li>• lowercase is in the pipeline (proposed codepoint is U+1D7D) Ṗ</li> </ul>
U+2C64	LATIN CAPITAL LETTER R WITH TAIL <ul style="list-style-type: none"> <li>• lowercase is U+027D Ṛ</li> </ul>

### D.3. Unicode Character Properties

U+2C61 should have a general category of Ll. Other properties for this character should match those of similar characters, such as U+0061 LATIN SMALL LETTER A.

Other characters should have a general category of Lu. Other properties for these remaining characters should match those of similar characters, such as U+0041 LATIN CAPITAL LETTER A.

### E. Other Information

#### E.1 Latin Extended

LATIN SMALL LETTER DOUBLE-BARRED L and LATIN CAPITAL LETTER DOUBLE-BARRED L are used orthographically in the Melpa and Nii languages of Papua New Guinea. They represent a velar fricative lateral. The Nii originally used one bar (Ł), but have more recently (since the early 90s) used the double-barred l (ll̥).

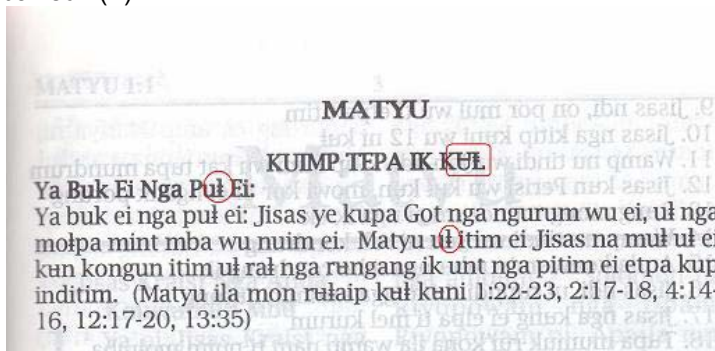


Figure 1. LATIN SMALL LETTER DOUBLE-BARRED L and LATIN CAPITAL LETTER DOUBLE-BARRED L (BSPNG, 1995, p. 1) [Melpa language of Papua New Guinea]

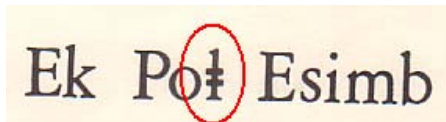


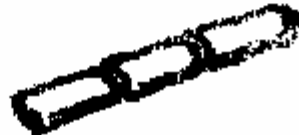
Figure 2. LATIN SMALL LETTER DOUBLE-BARRED L (Meiwu, 1981, title page) [Nii language of Papua New Guinea]

The Kobon language of Papua New Guinea uses LATIN CAPITAL LETTER L (alveolar lateral), LATIN CAPITAL LETTER L WITH BAR (retroflexed flapped lateral) **and** LATIN CAPITAL LETTER L WITH MIDDLE TILDE (alveopalatal lateral) in its orthography. This orthography has been in use for over 30 years. There are approximately 5,000 speakers of the Kobon language.

The lowercase counterpart to LATIN CAPITAL LETTER L WITH MIDDLE TILDE is U+026B.

≡ Hag Ńeb 21

Hag mai adip adip midöp?



bit



kanim ati

Figure 3. LATIN CAPITAL LETTER L WITH TILDE (Manö Kamiñ Mission Schools, 2003, p. 41) [Kobon language of Papua New Guinea]

LATIN CAPITAL LETTER P WITH STROKE is used orthographically by both the Letuama and Tanimuka (two distinct ethnic groups that speak the same language) of Colombia, South America. It represents a bilabial fricative. LATIN CAPITAL LETTER P WITH STROKE has been used in publications since 1987. The lowercase to this is already in the pipeline (proposed codepoint is U+1D7D). If this character is accepted, the upper and lower case should be placed contiguously.

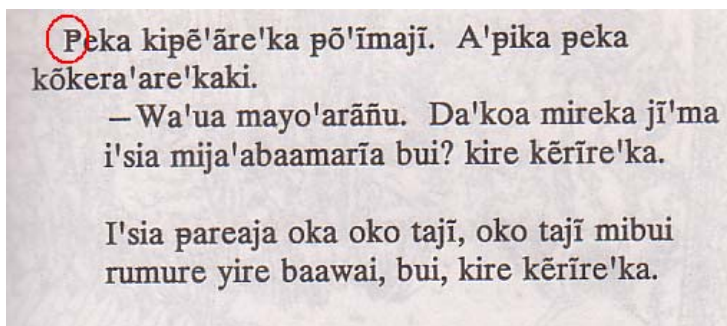


Figure 4. LATIN CAPITAL LETTER P WITH STROKE (Letuama, 1993, p. 15)

Īparīmarā naboebaarika. Tūparā makimarīka kīme  
 ārīwa'ri nayīrīberika. Papperapuñua (la Biblia) narikaek'a  
 Tūpārā makire itarāñu sārīroyika papperapuñuarā.  
 Nayīrīberikajika ruṣu. Nīpamaki nare jā'me'kaki Jesure  
 ne'ewa'rirā Pīato pōirā. Romano īpamaki kīmaek'a.

Figure 5. LATIN CAPITAL LETTER P WITH STROKE (Editorial Townsend, 1987, p. 5)

**Adivinanzas**

1. Kūpajīka yīñika. Īpato ko'rojēka nūpua īmawai. Wa'iro'sia īmarī potojījī pō'īmajāre yīpō'īmabiyawai. Dakoa?
2. Nōaka nīme. Wa'iro'sia najāriyapaika potojī yire ne'ewa'yawai. Dakoa?

**Ejercicio de puntuación — el punto**  
 (El profesor lo escribe en el tablero así y los alumnos lo escriben en sus cuadernos, colocando el punto donde falta.)

Wi'iareka kirīrāka pakiaki wāsare īmaroyireka Maē kirīrāre kērīreka, “Wa'iro'sia maro'si mo'arī ya'yu ōtomajā,” nare kērīreka Pakiaki wāsare a'rīrijarireka wa'iro'sia ō'ōmatorā mo'arī ke'rīreka Torā ruku maē ma'ajeabarō'ōrā ruku pakiaki wāsare sata'areka weje nāmajasirika

**Práctica**  
 (Llene el espacio con la palabra correcta según el cuento. El profesor escribe las frases en el tablero así:)

Nota: Los alumnos deben leer el cuento otra vez antes de hacer éste ejercicio.

1. Wi'iareka \_\_\_\_\_ pakiaki wāsare īmaroyireka.
2. Pakiaki wāsare a'rīrijarireka wa'iro'sia \_\_\_\_\_ mo'arī ke'rīreka.

Figure 6. LATIN CAPITAL LETTER P WITH STROKE (Camargo, 1998, p. 15)

LATIN CAPITAL LETTER R WITH TAIL is used orthographically in a number of Sudanese languages. There is confirmed orthographic usage in Heiban, Koalib, Moro and Otoro. Two different forms have been in use but Ꝛ is the preferred form (over ꝛ). The lowercase counterpart is U+027D.

4. Ꝛaṅr ṅopia ṅaṅgi iduniya,  
 Igidi yoṛəbədia taltal,  
 Eḍa gero iduniya geləṅeṭo loman,  
 Walla ēḍəñinia walla asana,  
 Illi Bapa gonto.
5. Yesu ḍəbia ḍa-Ḍaud ḍidi aṅiruwa,  
 Rrəmwa elo, ḍībərīntīa rīde,  
 Ltol lebara almalaiyəka yabəṭu.  
 Lorlḍaiñ aḍuniya ḍagəra.

Figure 7. LATIN CAPITAL LETTER R WITH TAIL (SLC, 2000, p. 100) [Moro language]



#### 14. Seṭan yēbōdia ṭau dēñidi?

Seṭan yēbōdia leda ṣadōna, na yēbōdia isi yaleda irāmitu ṣen ṣanta alerte lwōndata dētōm d-Almasiya na ṭaiyērāndeicu leda **R**-Rēmwa ṭa lēnāṣulu alidi ṣen ṣeicia lōgeiyo Rēmwa, na yēsākilo Rēmwa nano ēdēñēdēñin na ulōṣṣōlōṣ.

Margus 4:15

Ṣelwa inji ṣiṣṣṣu edad ṣarno leda ildi əlna ṣen, orn Seṭan nōyela taltal niyemate leda ṣelwa inji ṣēnāṣōnu enare enen.

Figure 8. LATIN CAPITAL LETTER R WITH TAIL (SLC, 1999, p. 27) [Moro language]

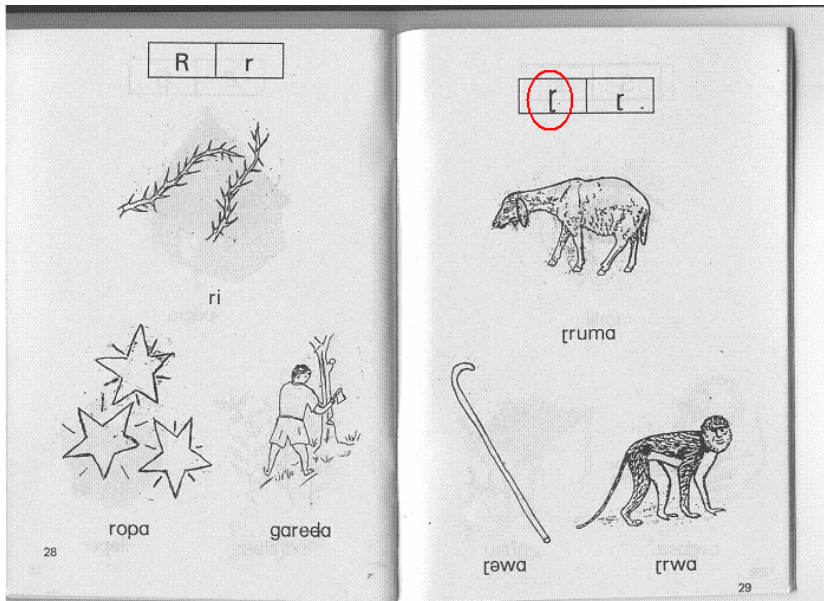


Figure 9. LATIN CAPITAL LETTER R WITH TAIL (Annggaico, 2000, p. 29) [Moro language]

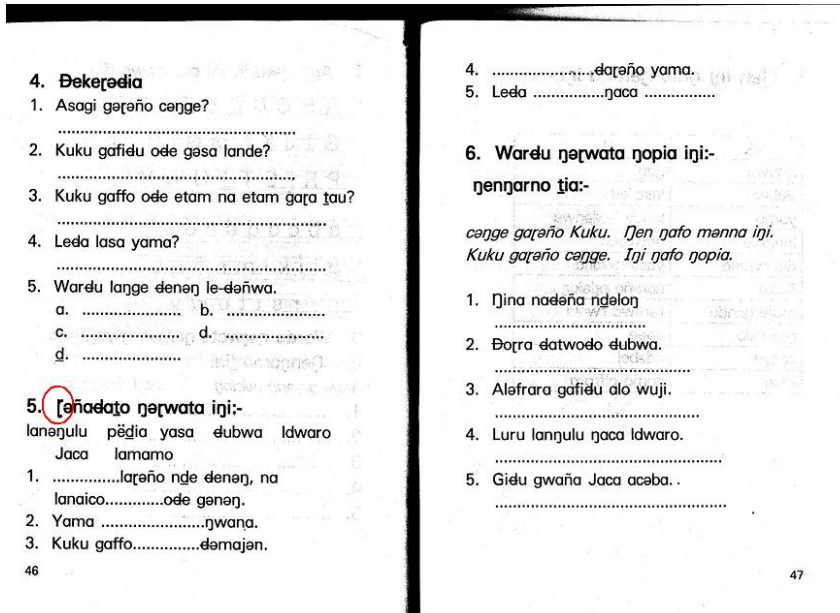


Figure 10. LATIN CAPITAL LETTER R WITH TAIL (MLC, 2002, p. 46) [Moro language]

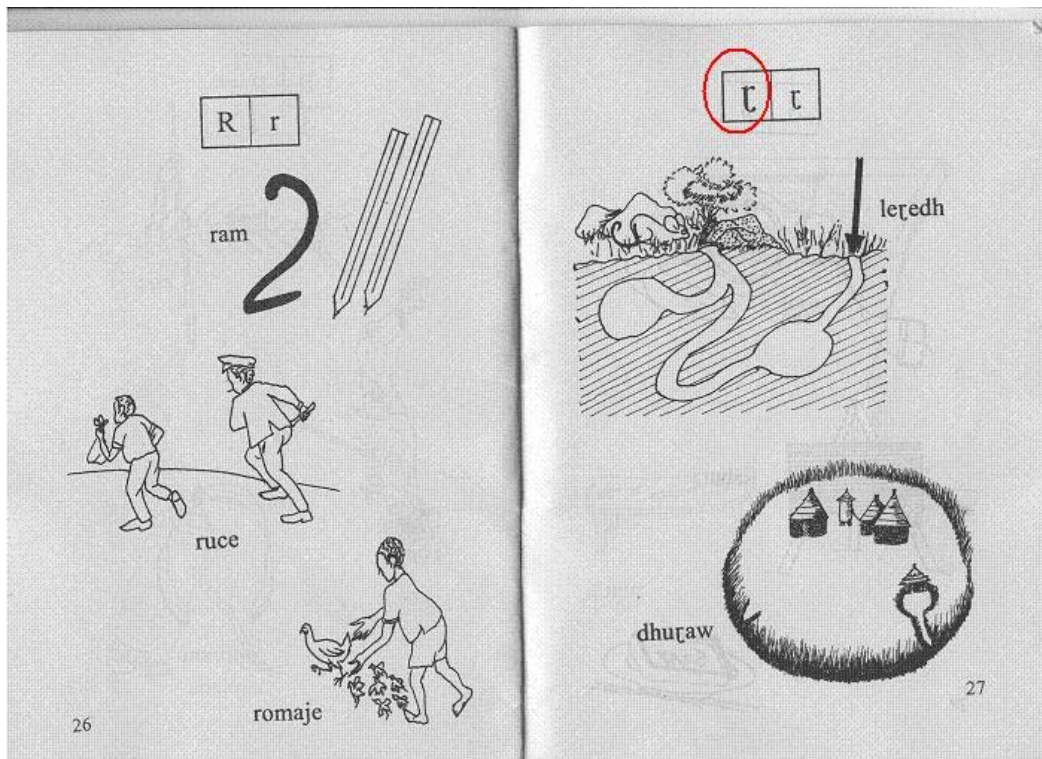


Figure 11. LATIN CAPITAL LETTER R WITH TAIL (Kafi, 1996, p. 27) [Heiban language]



Figure 12. LATIN CAPITAL LETTER R WITH TAIL (Kodi, 2002, p. 29) [Otoro language]



## F. References

- Annggaico, Isak Kuku, et. al. 2000. *Moro Alphabet Book, Second Edition*. Khartoum, Sudan: Khartoum Workshop Programme.
- The Bible Society of Papua New Guinea. 1995. *Got Nga Nombuŋa Rung Ik Kai*. Port Moresby-Lae, Papua New Guinea: The Bible Society of Papua New Guinea.
- Camargo, Alberto Lleras (ed.). 1998. *Mawãrĩjarĩ II: Cartilla para Segundo grado wejeñememajã oka / ùpairã oka*. Santafé de Bogotá, Colombia: Editorial Buena Semilla.
- Editorial Townsend. 1987. *Jesucristore najããe'ka pōtojĩkaka majaroka*. Meta, República de Colombia: Editorial Townsend.
- Kafi, Adam Kuku, et. al. 1996. *Heiban Alphabet Book, 1st Edition*. Khartoum, Sudan: Khartoum Workshop Programme.
- Kodi, Musa, et. al. 2002. *Otoro Alphabet Story Book*. Sudan: Sudan Workshop Programme.
- Letuama, Miguel Antonio. 1993. *Ñamamajaroka*. Santafé de Bogotá, Colombia: Editorial Alberto Lleras Camargo.
- Manö Kamiñ Mission Schools. 2003, revised. *Kobon Consonants: A Consonant Primer for use in Kobon Language Manö Kamiñ Mission Schools in the Kobon-Huruai School District*. Madang Province, Papua New Guinea: Manö Kamiñ Mission Schools.
- Meiwu, Kondii (transl and ed.). 1981. *Ek Poŋ Esimb*. Ukarumpa, Papua New Guinea: Summer Institute of Linguistics.
- Moro Language Committee. 2002. *Moro Primer, Third Edition*. Khartoum, Sudan: Khartoum Workshop Programme.
- Sudan Literature Centre. 1999 (reprinted from 1990). *Moro Catechism*. Nairobi, Kenya: Sudan Literature Centre.
- \_\_\_\_\_. 2000 (reprinted from 1991, 1996, 1997, 1998). *Moro Hymnal*. Nairobi, Kenya: Sudan Literature Centre.