

**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: Proposal to Encode Additional Latin Orthographic Characters
2. Requester's name: Lorna A. Priest
3. Requester type (Member body/Liaison/Individual contribution): Individual contribution
4. Submission date: 31 October 2004
5. Requester's reference (if applicable): \_\_\_\_\_
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: Modifier letters, Latin Extended
2. Number of characters in proposal: 7
3. Proposed category (select one from below - see section 2.2 of P&P document):
 

|                                             |                                               |                                          |
|---------------------------------------------|-----------------------------------------------|------------------------------------------|
| A-Contemporary <u>x</u>                     | 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?<br>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.)?<br>If YES, with whom? <u>linguists</u><br>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?<br>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)<br>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?<br>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?<br>If YES, is a rationale provided?<br>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?<br>If YES, is a rationale for its inclusion provided?<br>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?<br>If YES, is a rationale for its inclusion provided?<br>If YES, reference: <u>(Cf. §F.1 of L2/04-047.)</u>                                               | <u>Yes</u><br><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?<br>If YES, is a rationale for its inclusion provided?<br>If YES, reference: <u>See comments in Section E</u>                                                                    | <u>Yes</u><br><u>Yes</u>                                                                                                              |
| 11. Does the proposal include use of combining characters and/or use of composite sequences?<br>If YES, is a rationale for such use provided?<br>If YES, reference: _____<br>Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?<br>If YES, reference: _____ | <u>No</u>                                                                                                                             |
| 12. Does the proposal contain characters with any special properties such as control function or similar semantics?<br>If YES, describe in detail (include attachment if necessary) _____                                                                                                                  | <u>No</u>                                                                                                                             |
| 13. Does the proposal contain any Ideographic compatibility character(s)?<br>If YES, is the equivalent corresponding unified ideographic character(s) identified?<br>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

|   | xx00 |
|---|------|
| 0 | !    |
| 1 | Ł    |
| 2 | ł    |
| 3 | Ł    |
| 4 | Ł    |
| 5 | Ɔ    |
| 6 | Ŕ    |
| 7 |      |
| 8 |      |
| 9 |      |
| A |      |
| B |      |
| C |      |
| D |      |
| E |      |
| F |      |

### D.1. Character Names

|      |                                                                                                                                                    |
|------|----------------------------------------------------------------------------------------------------------------------------------------------------|
| xx00 | MODIFIER LETTER STRAIGHT APOSTROPHE                                                                                                                |
| xx01 | LATIN CAPITAL LETTER DOUBLE-BARRED L                                                                                                               |
| xx02 | LATIN SMALL LETTER DOUBLE-BARRED L                                                                                                                 |
| xx03 | LATIN CAPITAL LETTER L WITH BAR <ul style="list-style-type: none"> <li>lowercase is U+019A ł</li> </ul>                                            |
| x004 | LATIN CAPITAL LETTER L WITH TILDE <ul style="list-style-type: none"> <li>lowercase is U+026B ł</li> </ul>                                          |
| xx05 | LATIN CAPITAL LETTER P WITH STROKE <ul style="list-style-type: none"> <li>lowercase is in the pipeline (proposed codepoint is U+1D7D) p</li> </ul> |
| xx06 | LATIN CAPITAL LETTER R WITH TAIL <ul style="list-style-type: none"> <li>lowercase is U+027D ɽ</li> </ul>                                           |

### D.3. Unicode Character Properties

xx00 should have a general category of Lm. Other properties should match those of similar characters, such as U+02BC MODIFIER LETTER APOSTROPHE.

xx02 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 Modifier Letters

MODIFIER LETTER STRAIGHT APOSTROPHE is used to represent the phoneme glottal stop. The intent is to have a character for glottal stop (other than the glottal stop phonetic symbol) that is unambiguous and will not be confused with apostrophe or quotation marks. There are already several different characters used to represent glottal stop: U+0027 APOSTROPHE, U+0294 LATIN LETTER GLOTTAL STOP, U+02bc MODIFIER LETTER APOSTROPHE, U+02be MODIFIER LETTER RIGHT HALF RING, U+2019 RIGHT SINGLE QUOTATION MARK. However, all of these existing representations involve a character that can be represented with a glyph either that is curved or that is straight but slanted. The distinction with MODIFIER LETTER STRAIGHT APOSTROPHE is that the user communities that developed this character (independently!) have explicitly intended this to be tapered, straight, vertical, and in many cases long precisely to distinguish it from apostrophes, quotation marks, accents and stress bars. Thus, it should not be unified with U+0027, U+02b9, U+02bc, U+02c8, U+02ca or U+2019. The curved shaped of U+0294 and U+02be would be considered by users to be a completely distinct character identity; unification with either of these characters is not an option.

MODIFIER LETTER STRAIGHT APOSTROPHE is used in many languages around the world, among them el chontal de la sierra de Oaxaca (figure 6) and Tlapaneco de Malinaltepec (figure 1) of Mexico, Ese Ejja of Bolivia (figure 2), Maya Mopán of Guatemala (figure 3), Dano (Upper Asano) of Papua New Guinea (figure 4) and FraFra of Ghana (figure 5).

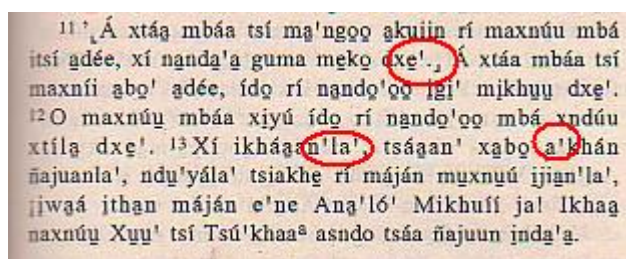


Figure 1. MODIFIER LETTER STRAIGHT APOSTROPHE (La Liga del Sembrador, 1994, p. 63) [Tlapaneco de Malinaltepec language of Mexico]

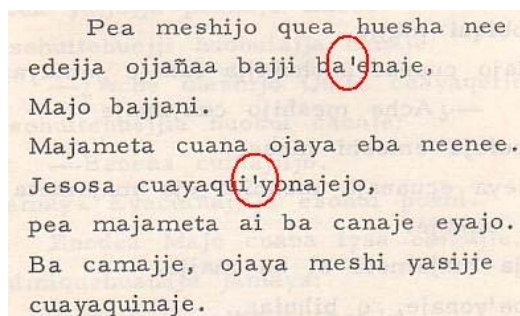


Figure 2. MODIFIER LETTER STRAIGHT APOSTROPHE (Imprenta Indigenista Boliviana, 1968, p. 13) [Ese Ejja language of Bolivia]

Los consonantes glotalizados c', ch', p', q'u, t', y tz' representan sonidos que se reproducen igual a c, ch, p, qu, t, y tz a diferencia de un cerrar y abrir de las cuerdas vocales durante la emisión del sonido, produciendo un sonido explosivo.

La w representa un sonido aproximando el sonido de la gu española en las palabras agua y guau (onomatopeyico de ladrar).

La x representa un sonido alveopaladal como la x que se encuentra en las palabras xeca, xara, xolco.

El saltillo (ˈ) representa una oclusiva glotal, un sonido cortado que se emite al sacar el aire de la garganta cuando se han cerrado y abierto las cuerdas vocales previamente. No se ha escrito el saltillo cuando ocurre en posición inicial de la palabra, antes de la vocal.

Figure 3. MODIFIER LETTER STRAIGHT APOSTROPHE (ILV, 1962, p.vi) [Maya Mopán language of Guatemala]

a'mine salani molo-lidine' inite' nene lelikisi makato'  
dalonido' minikeve. 7Ido' O'moso aza nasahili' o-nidine'  
ne'mo ei' amuzo've nene neniku' gono nodavo aza gor  
nemavosa a'mido' e'meto a'mine gamazi lamine' nene  
soto' ogo eze goboni ve ne'mine lolo' ogo nouve. 8Ee,  
nene gomu' minaa' uvodasi giluvoni' nene O'moso'mi  
mukikuti' lemo minuve minuvoni'za O'moso aza neni'  
do vevese-nedo nasahili' o-nidine' ne'mo Kilisitodati' r  
monovo danava lamine tamine' vai' goloso lo neivo laz

Figure 4. MODIFIER LETTER STRAIGHT APOSTROPHE (WHBL, 1989, p.762) [Dano (Upper Asano) language of Papua New Guinea]

9Yezu daa le yele yeti, "Ya tare sera la ya bahe  
Na'ayene pa'alego gee dola ya menhe pa'alego. 10Bala mea  
te Mosese daa yeti, 'Bo'ora ho so' la ho ma' gelema,' gee  
'Se'em n kan' kan'be'o bo' a so' bii a ma', la sontii ba ko'  
en.' 11Amaa, yamam yeti, te nera han yele a so' bii a ma'  
yeti, 'Sonsere woo te n wan songe ne ho la, la de la sela te  
n wan deke bo' Na'ayene.' 12Bala, sonsere te ya kan songe  
ya sodoma bii ya madoma n bala. 13Bala, ya menhe  
pa'alego wa zuo te ya bahe te Na'ayene Yela la ka tara  
vonre, gee te yela a se'a zo'e bona ana bala te ya eta."

Figure 5. MODIFIER LETTER STRAIGHT APOSTROPHE (IBS, 1986, p.87) [FraFra language of Ghana]

## E.2 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 (Ꞥ).



14“Daun kub dum laŋ midöm, pi  
göl rö lagnab; waiö yabil mideinab.  
Anib u rö nibi bi yad gau. Kale nibi bi  
gau mailö kalip u midpim. 15 Lam u  
lauöm rin migan gau lagpal; nibi bi  
niŋ aij gilaŋ a göm, ram raul adö aulik  
rik yubal. 16 Anib u rö, kale mid aij  
gipe, nibi bi niŋöm, Nap kumi kabö  
adö laŋ midöp hib u nöp dap ran löl.

Figure 10. LATIN CAPITAL LETTER L WITH BAR (BSPNG, in press, p. 8) [Kobon language of Papua New Guinea]

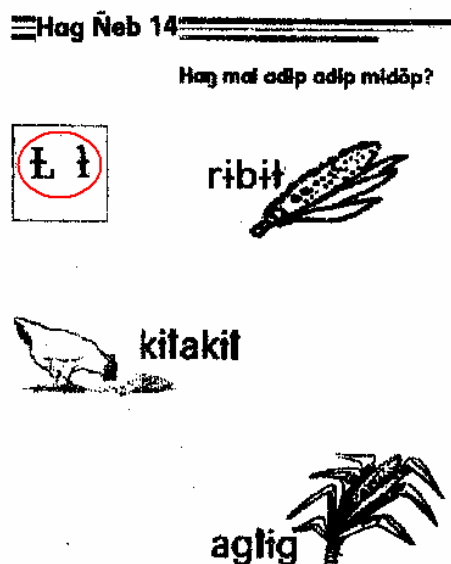


Figure 11. LATIN CAPITAL LETTER L WITH BAR (Manö Kamiŋ Mission Schools, 2003, p. 27) [Kobon language of Papua New Guinea]

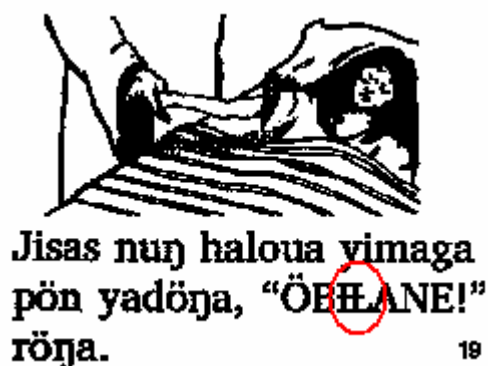


Figure 12. LATIN CAPITAL LETTER L WITH BAR (Yese, 2004, p. 19) [Haruai language of Papua New Guinea]

≡Hag Neb 21

Hag mai adip adip midöp?



bit



kanim at

Figure 13. LATIN CAPITAL LETTER L WITH TILDE (Manö Kamin 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.

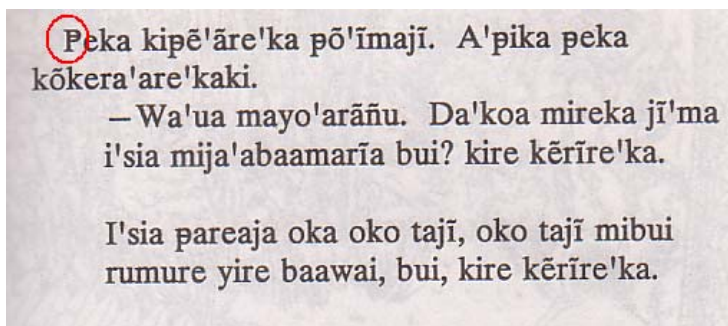


Figure 14. LATIN CAPITAL LETTER P WITH STROKE (Letuama, 1993, p. 15)

Īparīmarā naboebaarika. Tūparā makimarīka kīme  
 ārīwa'ri nayī'riberika. Pāperapuñua (la Biblia) narikae'ka  
 Tūpārā makire itarāñu sārīroyika pāperapuñuarā.  
 Nayī'riberikajika ruṑu. Nipamaki nare jā'me'kaki Jesure  
 ne'ewa'rirā Pīato pōirā. Romano ĩpamaki kīmae'ka.

Figure 15. LATIN CAPITAL LETTER P WITH STROKE (Editorial Townsend, 1987, p. 5)

**Adivinanzas**

1. Kūpajīka yīñika. Īpato ko'rojēka ñupua ĩmawai. Wa'iro'sia ĩmarī potojījī pō'ima jāre yīpō'imabiyawai. Dakoa?
2. Ñoaka ñime. Wa'iro'sia najāāriyapaika 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 ĩmaroyireka 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'ajeebarō'ōrā ruku pakiaki  
 wāsare sata'areka weje ñamajasirika

**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 ĩmaroyireka.
2. Pakiaki wāsare a'rīrijarireka wa'iro'sia \_\_\_\_\_  
 mo'arī ke'rīreka.

Figure 16. 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. Ꝛaṛ ṇopia ṇṅgi iduniya,  
 Igidi yorəbədia taltal,  
 Eda gero iduniya geləṇeto loman,  
 Walla ēdəñinia walla asana,  
 Illi Bapa gonto.
5. Yesu dəbia da-Daud didi aniruwa,  
 Rrəmwa elo, dibərīntia rīde,  
 Ltol lebara almalaiyēka yabəṭu.  
 Lorīdañ aḍuniya dagəra.

Figure 17. LATIN CAPITAL LETTER R WITH TAIL (SLC, 2000, p. 100) [Moro language]

#### 14. Seṭan yēbōdia ʔau dāñidi?

Seṭan yēbōdia leda ɲadāna, na yēbōdia isi yaleda irāmitu ɲen ɲaṇṭa alerte lwondata deṭam d-Almasiya na ʔaiyerāndeicu leda **R**-Rāmwa ʔa lēnāgulu alidi ɲen ɲeicia lāgeiyo Rāmwa, na yesākilo Rāmwa nano ēdāñēdāñin na ulāngələṅ.

*Margus 4:15*

Ŋəlwa iɲi ɲiṛṭu edad ɲarno leda ildi əlna ɲen, orn Seṭan nāyela taltal niyemate leda ɲəlwa iɲi ɲēnəṅənu enare enen.

Figure 18. LATIN CAPITAL LETTER R WITH TAIL (SLC, 1999, p. 27) [Moro language]

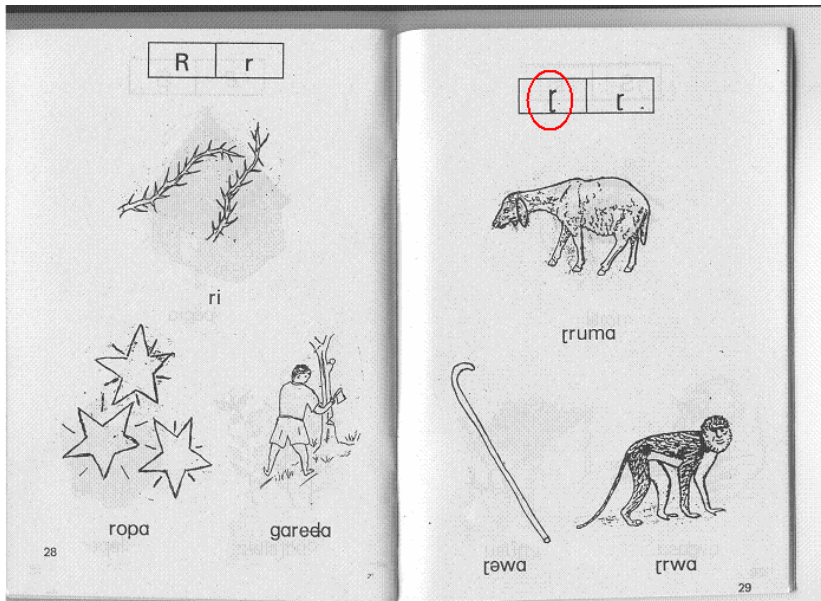


Figure 19. LATIN CAPITAL LETTER R WITH TAIL (Annggaico, 2000, p. 29) [Moro language]

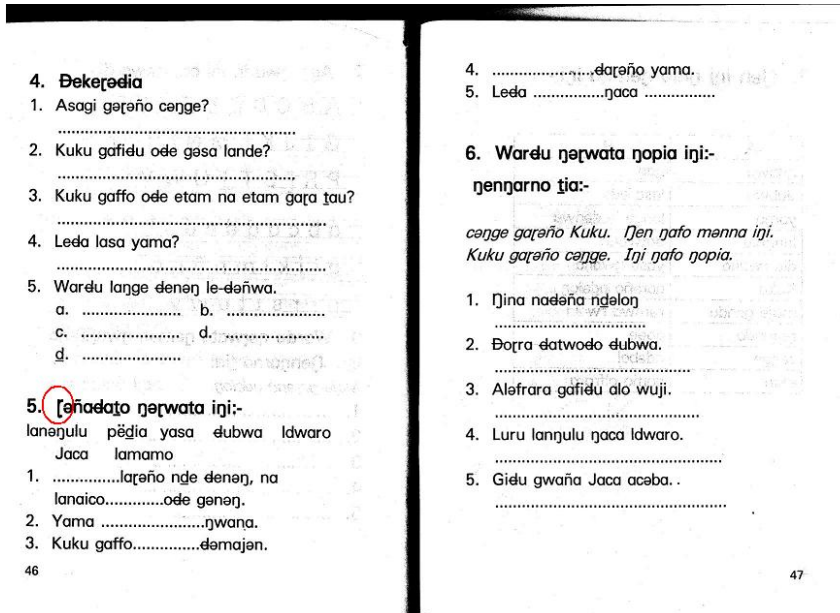


Figure 20. LATIN CAPITAL LETTER R WITH TAIL (MLC, 2002, p. 46) [Moro language]

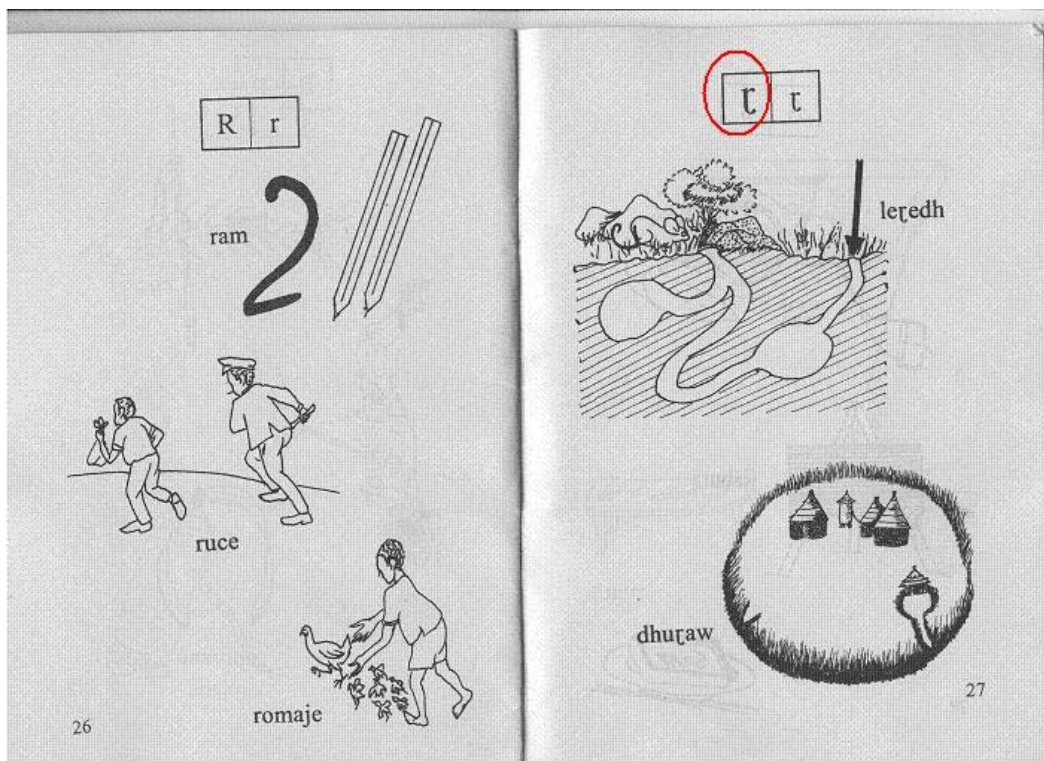


Figure 21. LATIN CAPITAL LETTER R WITH TAIL (Kafi, 1996, p. 27) [Heiban language]

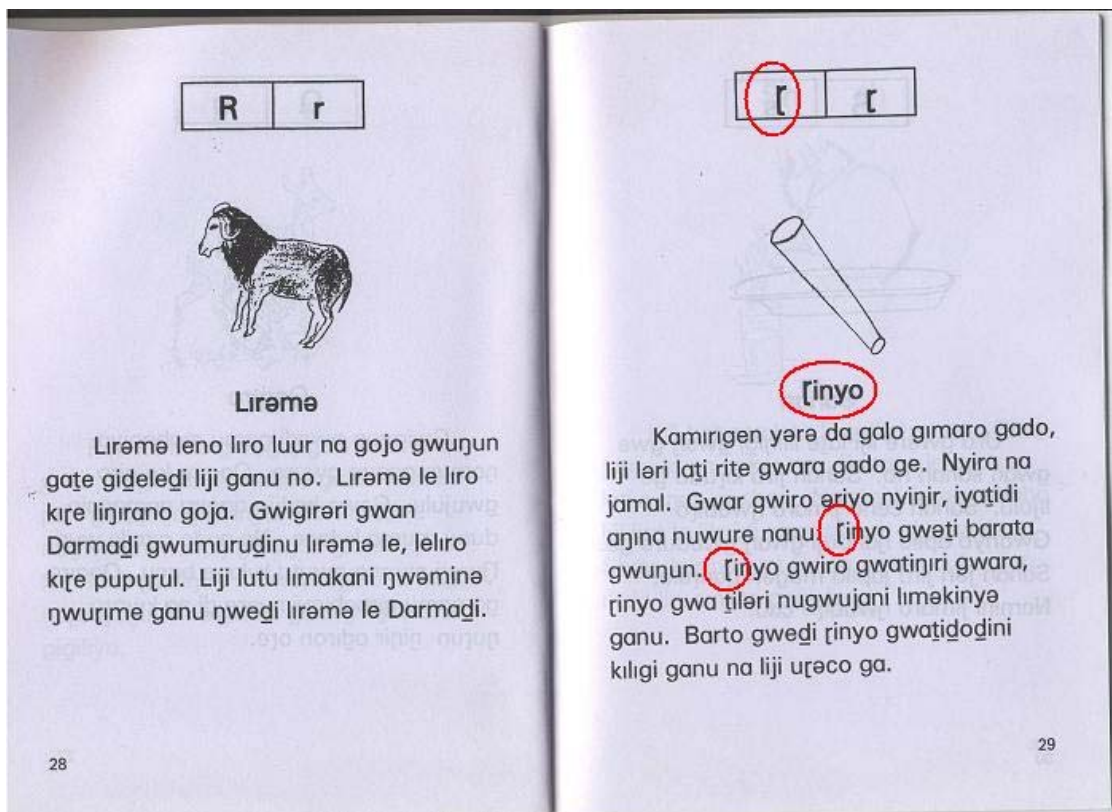


Figure 22. LATIN CAPITAL LETTER R WITH TAIL (Kodi, 2002, p. 29) [Otoro language]

## 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 Nombuta Rung Ik Kai*. Port Moresby-Lae, Papua New Guinea: The Bible Society of Papua New Guinea.
- \_\_\_\_\_. (in press). *Manö Kamiŋ*. Madang Province, Papua New Guinea: The Bible Society of Papua New Guinea.
- Camargo, Alberto Lleras (ed.). 1998. *Mawârürjarirã II: Cartilla para Segundo grado wejeñememajã oka / ũpairã oka*. Santafé de Bogotá, Colombia: Editorial Buena Semilla.
- Davies, H, John (trans. and ed). 2004. *Mani kabö magö ap ur ga* (Kobon Level 2 Reader). Dusi-Madang Province, Papua New Guinea.
- Editorial Townsend. 1987. *Jesucristore najããe'ka pōtojīkaka majaroka*. Meta, República de Colombia: Editorial Townsend.
- Imprenta Indigenista Boliviana. 1968. *Jesosa Cuayaquinaje, Navidad*. Cochabamba, Bolivia: Imprenta Indigenista Boliviana.
- Instituto Linguistico de Verano. 1962. *El Alfabeto Maya Mopán*. Guatemala, C.A.: Instituto Linguistico de Verano.
- International Bible Society. 1986. *Na'ayene Gŋpaalom*.
- 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.
- La Liga Bíblica. 1991. *Lataiqui' Loya'apa Ite al C'a Lixpic'epa Landios* (El Nuevo Testamento en el chontal de la sierra de Oaxaca). Mexico, D.F.: La Liga del Sembrador, A.C. y La Liga Bíblica.
- La Liga del Sembrador. 1994. *Ajngáa noxe' ri ni'tháán San Lúkas*. México: La Liga del Sembrador.
- Letuama, Miguel Antonio. 1993. *Ñamamajaroka*. Santafé de Bogotá, Colombia: Editorial Alberto Lleras Camargo.
- Manö Kamiŋ Mission Schools. 2003, revised. *Kobon Consonants: A Consonant Primer for use in Kobon Language Manö Kamiŋ Mission Schools in the Kobon-Huruai School District*. Madang Province, Papua New Guinea: Manö Kamiŋ Mission Schools.
- Meiwu, Kondii (transl and ed.). 1981. *Ek Poł 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.
- World Home Bible League. 1989. *Mono' Godolo Gosoho'*. South Holland, Ill.: World Home Bible League.
- Yese, Lukson (trans.). (2004, third printing). *Jairas halou nuŋ yi*. Dusi, Madang Province, Papua New Guinea: SIL.