

ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646¹

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 for Additional Latin Phonetic and Orthographic Characters
2. Requester's name: Lorna A. Priest
3. Requester type (Member body/Liaison/Individual contribution): Individual contribution
4. Submission date: 11 June 2004 (revised 26 July 2004)
5. Requester's reference (if applicable): L2/04-246R
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: Combining Diacritical Marks, Latin Extended, Phonetic Extensions
2. Number of characters in proposal: 30
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): 3 (some combining marks)
 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.

¹ 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)








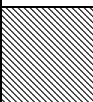
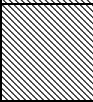







C. Technical - Justification

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>Phonetic symbols are in current use in linguistic journals and books; 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>No</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

	Xx0	xx1
0		B
1		U
2		Λ
3		@
4		Ⓐ
5		œ
6		Ǝ
7		ĵ
8		J
9		q
A		Q
B		f
C		R
D		ŷ
E		Ÿ
F		v

D.1. Character Names

xx00	COMBINING MACRON-ACUTE
xx01	COMBINING GRAVE-MACRON
xx02	COMBINING MACRON-GRAVE
xx03	COMBINING ACUTE-MACRON
xx04	COMBINING GRAVE-ACUTE-GRAVE
xx05	COMBINING ACUTE-GRAVE-ACUTE
xx06	COMBINING LATIN SMALL LETTER R BELOW
xx09	MODIFIER LETTER CHINANTEC TONE MARK BACKSLASH
xx0A	MODIFIER LETTER CHINANTEC TONE MARK VERTICAL BAR
xx0B	MODIFIER LETTER CHINANTEC TONE MARK SLASH
xx0C	MODIFIER LETTER CHINANTEC TONE MARK DOT VERTICAL BAR
xx0D	MODIFIER LETTER CHINANTEC TONE MARK DOT SLASH
xx0E	MODIFIER LETTER CHINANTEC TONE MARK DOT BAR
xx0F	MODIFIER LETTER CHINANTEC TONE MARK LOWER RIGHT CORNER ANGLE
xx10	LATIN CAPITAL LETTER B WITH STROKE <ul style="list-style-type: none"> • lowercase is 0180 b
xx11	LATIN CAPITAL LETTER U BAR <ul style="list-style-type: none"> • lowercase is 0289 u
xx12	LATIN CAPITAL LETTER TURNED V <ul style="list-style-type: none"> • lowercase is 028C v
xx13	LATIN SMALL LETTER AT
xx14	LATIN CAPITAL LETTER AT
xx15	LATIN SMALL LETTER E WITH STROKE
xx16	LATIN CAPITAL LETTER E WITH STROKE
xx17	LATIN SMALL LETTER J WITH STROKE
xx18	LATIN CAPITAL LETTER J WITH STROKE
xx19	LATIN SMALL LETTER Q WITH HOOK TAIL
xx1A	LATIN CAPITAL LETTER SMALL Q WITH HOOK TAIL
xx1B	LATIN SMALL LETTER R WITH STROKE
xx1C	LATIN CAPITAL LETTER R WITH STROKE
xx1D	LATIN SMALL LETTER Y WITH STROKE
xx1E	LATIN CAPITAL LETTER Y WITH STROKE
xx1F	LATIN SMALL LETTER V WITH CURL

D.3. Unicode Character Properties

xx00..xx05 should have a general category of Mn, and a canonical combining class of 230. Other properties should match those of similar characters, such as U+0300 COMBINING GRAVE ACCENT.

xx06 should have a general category of Mn, and a canonical combining class of 220. Other properties should match those of similar characters, such as U+0300 COMBINING GRAVE ACCENT.

xx09..xx0F should have a general category of Lm. Other properties for these characters should match those of similar characters, such as U+02BE MODIFIER LETTER RIGHT HALF RING.

xx10..xx12, xx14, xx16, xx18, xx1A, xx1C and xx1E should have a general category of Lu. Other properties for these characters should match those of similar characters, such as U+0041 LATIN CAPITAL LETTER A.

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

E. Other Information

E.1 Combining Diacritical Marks

The characters COMBINING MACRON-ACUTE, COMBINING GRAVE-MACRON, COMBINING MACRON-GRAVE, COMBINING ACUTE-MACRON are phonetic symbols which represent contour tone.

COMBINING MACRON-ACUTE and COMBINING GRAVE-MACRON are attested in the IPA Handbook. The other four contour-tone diacritics are not cited in the IPA handbook, but are widely used by linguists.

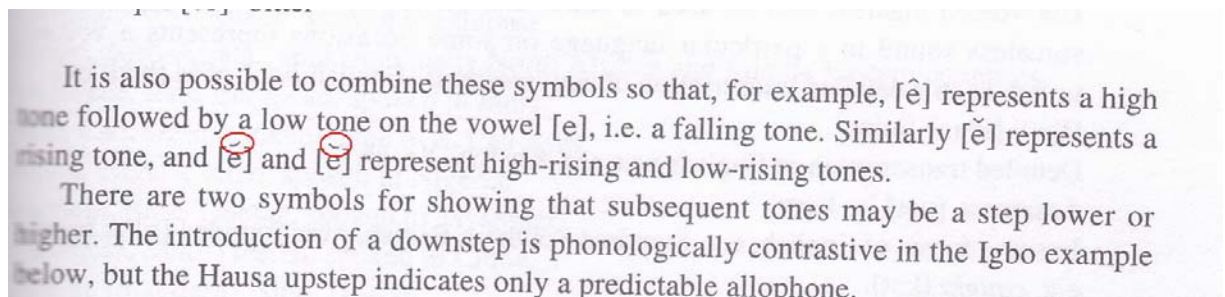


Figure 1. COMBINING MACRON-ACUTE and COMBINING GRAVE-MACRON (IPA, 1999, p. 23).

(96)	H	<i>bák</i>	'garden'
	M	<i>bān</i>	'cow with drooping horns'
	L	<i>bàk</i>	'guess!'
	H̄L	<i>bā̄n</i>	'servant'
	M̄L	<i>bw̄̃c</i>	'barren person'
	M̄H	<i>bā̄̃</i>	'arm'
	L̄H	<i>bȳ̃c</i>	'cow with horns straight out'

Figure 2. COMBINING MACRON-ACUTE and COMBINING MACRON-GRAVE (Gilley, 1992, p. 51).

abbreviations			
V	vowel	σ	syllable
V	verb	o	node
v.i	intransitive verb	`	low tone
v.r	reflexive verb	-	mid tone
v.tr	transitive verb	˘	high tone
VOC	vocative	◌̣	low-mid tone
VRT	verb root	◌̤	low-high tone
VSX	verb suffix	◌̥	mid-high tone
1	first person	◌̦	high-mid tone
2	second person	i	[i]
3	third person	u	[u]

Figure 3. COMBINING GRAVE-MACRON, COMBINING MACRON-ACUTE and COMBINING ACUTE-MACRON (Kutsch Lojenga, 1993, p. xii).

Inzá àmbò nàpfo	tùna	mà.
inzá àmbò ní -apfo	tù -na	mà
NEG dove	RSM-open:NAR	voice-3SG.LOG.POS even
<i>Dove didn't even make a sound.</i>		
Tdítò idre adyifò-sisì	kúbhingánga	tó 'ò, ka nángò:
tdítò idre adyifò-sisì	kúbhingá-nga	tó 'ò ka ní -angò
again ten sun -shadow at.night -NOM	POS in 3SG	RSM-crow:NAR
<i>Again, at four o'clock in the morning, Cock crowed:</i>		

Figure 4. COMBINING GRAVE-MACRON and COMBINING ACUTE-MACRON (Kutsch Lojenga, 1993, p. 408).

(103)	HĤLH	átùd̂	'duck'
	HĤLL	átê:ĝ	'goat'
	HĤMM	álwê:d̂	'crab'
	HLL	ácùñ̂	'black mound termite'
	HĤML	ágāk̂:i	'ravens'
	ML	ādù:l̂	'circular things'
	LMH	àbā:r̂	'fencing grass'
	LĤLL	àbûr̂	'reedbuck'
	LLL	àdùd̂	'large sieve'
	LH	jùb̂ó:w	'white calves' (R)
	LL̂H	dìd̂ÿ:k	'dark grey bulls'

Figure 5. COMBINING ACUTE-MACRON (Gilley, 1992, p. 53).

In addition to being used phonetically, COMBINING GRAVE-MACRON and COMBINING MACRON-GRAVE are also used orthographically in the Bette language of Nigeria.

¹⁶Ujeso a kung ityang anyin bun de le: "Usu 'ingbebb ken a ma yi lè ùtiá angwu biwom bi mbu-e há kangg. ¹⁷U fel gim ugim le: 'Mì yi lè 'ipi ashi m̀ ndor azi agi ye. Mi hà ba ndor āngi kung a? ¹⁸Ashi m̀ be shi 'ì yi le mi tya abel agi anyin wa he itang, m̀ tye abel asuso angi a dyam kangg. Há nà m̀ be ndor azi agi puu. ¹⁹Mi hà kung de 'uhwo-gi le, eng-heng, mi angwu liti li ngbebb, mi fel lè biwom gbamm abi m̀ be zī he aze gbamm angi ā be lè he 'ushu. Ki fel ashi m̀ sho 'unyong, zī bizizi, ngwo bingwungwo, gor igor.'

Figure 6. COMBINING GRAVE-MACRON and COMBINING MACRON-GRAVE (IBS, 1982, p. 198).

COMBINING GRAVE-ACUTE-GRAVE and COMBINING ACUTE-GRAVE-ACUTE are phonetic symbols which represent contour tone.

(327) <i>bàlóngó bǎkáé</i>	→	<i>bàlóngǎkáé</i>	'his book'
<i>bánà bǎmǒ</i>	→	<i>bánǎmǒ</i>	'other children'
<i>bǒmǎ bòtámǎ</i>	→	<i>bǒmǎtámǎ</i>	'another tree'
<i>bǎtswá là èmf</i>	→	<i>bǎtswémf</i>	'you who lead me away'

Figure 7. COMBINING GRAVE-ACUTE-GRAVE and COMBINING ACUTE-GRAVE-ACUTE (Burquest, 1998, p. 234).

(5) (a)	<i>bàlóngó</i>	<i>bǎkáé</i>	→	<i>bàlóngǎkáé</i>	"his book"
(b)	<i>bǒmǎ</i>	<i>bòmtámǎ</i>	→	<i>bǒmǎtámǎ</i>	"another tree"

Figure 8. COMBINING GRAVE-ACUTE-GRAVE and COMBINING ACUTE-GRAVE-ACUTE (Goldsmith, 1995, p. 446)

COMBINING LATIN SMALL LETTER R BELOW is used orthographically in four languages in Indonesia, namely Mongondow, Sangir, Siau and Talaud.

Buké i *Matius* muluhabaré si kité, Habaré'u Raluasé, humonéu i Yesus Kristus ko Datung Salamaté ko nipudiandingu Mawu Ruata. Bou i Yesus éé Mawu Ruata nungganapé apang ko nipudiandi-Né su ralungu Pudariandi Tebé su manga umaté-É. Maningu i Yesus nipuhana bou taumangu Yahudi, kaiso Habaré'u Raluasé éé mang balinébé ketang gunangu tau Yahudi, kaiso mang lai gunangu kase-lahé'u dunia.

Figure 9. COMBINING LATIN SMALL LETTER R BELOW (Lembaga Alkitab Indonesia, 2002, p. 1).

Hale sēmbau kawēhakeŋe nēnaungaŋ i Matius u Mawu Yesus kai Mananēntiro kasēllahenge, kŋ manēntiro u Mawu Ruata e kai mēpēparenta kere sēngkatau Ratu. Mawu Yesus maiaing pia kawasan'E waug'u mēlahē u mangalen Torat'u Mawu Ruata. Kaia-wokange wou manga tēntiron Mawu Yesus ene nipahia tuhu ti-monane. Ku pia e limang kamonaēng: (1) Tēntiro su wulude ku maanung kakanoa, tatuhu, manga pēnanaghuang, dingangu ti-monan pēbawiah'u manga raohong jamaat'u Mawu (pasalē 5-7); (2) manga tatēntiro si sire mapulo dua murit'u Mawu Yesus baugu pēkakoā munara (pasalē 10); (3) manga papinintu maanung kaka-koāu tempon pēpaparentan Duata kere sēngkatau Ratu (pasalē 13); (4) tēntiro maanun mangalen makoā murit'u Mawu Yesus (pasalē 18); ringangu (5) tēntiro maanun pēngēsueēngu dunia ringangu maanun darēntang Ahus'u Taumata (pasalē 24-25).

Figure 10. COMBINING LATIN SMALL LETTER R BELOW (Lembaga Alkitab Indonesia, 2003, p. 1).

WISARA TALAANE

Wuken Pandariaddi indi mabbatti'u pandariaddi ulilin pati'uppa, panataraṇṇu walahannu taumata, punnu paddarosa wuṇṇu asasusa su runia, see ereapa Ruata naaukka ringannu taumata. Wuken Pandariaddi indi wotongge tahianna sulalammu ruang kawageanganna, eteudde:

Figure 11. COMBINING LATIN SMALL LETTER R BELOW (Lembaga Alkitab Indonesia, 2003b, p. 1).

E.2 Latin Extended

MODIFIER LETTER CHINANTEC TONE MARK BACKSLASH, MODIFIER LETTER CHINANTEC TONE MARK VERTICAL BAR, MODIFIER LETTER CHINANTEC TONE MARK SLASH, MODIFIER LETTER CHINANTEC TONE MARK DOT VERTICAL BAR, MODIFIER LETTER CHINANTEC TONE MARK DOT SLASH, MODIFIER LETTER CHINANTEC TONE MARK DOT BAR, and MODIFIER LETTER CHINANTEC TONE MARK LOWER RIGHT CORNER ANGLE are orthographic characters which are used for the Ozumacín Chinantec language of Mexico.

32 Cha-ga'miih- jmeen' nuuhn' mada' kyah,
jēeh' ga' jmeeh' jēē- dsēn'.

Figure 12 Chinantec tone marks: BACKSLASH, SLASH, DOT VERTICAL BAR, DOT SLASH, DOT BAR, LOWER RIGHT CORNER ANGLE (Liga Bíblica, 2003, p. 864).

17 'Ta^ˈjmee- hnäh^ˈ huu^ˈdsē- he- ma-jañ mah^ˈ mi'llaan^ˈ ley kih^ˈ Moisé. Ni^ˈ sa^ˈ mi'llaan^ˈlä-jē^ˈ he- ga^ˈhlēēh^ˈ lä-jēē^ˈ hi- ga^ˈngēē^ˈ juu^ˈ kih^ˈ Dio. Sa^ˈ ma-jañ mah^ˈ mi'llaan^ˈ he- ja^ˈ, he- mi'ten^ˈ kih^ˈ juu^ˈ ja- ba' ja- ma-jañ. 18 Ma^ˈra^ˈ ju'dsoo^ˈ jwēēhn^ˈ hnäh^ˈ, ta^ˈko-jiih^ˈ cha- gyuu^ˈ hwa^ˈ la^ˈ, sa^ˈ hee^ˈ ni^ˈ koo^ˈ hwēh^ˈ mäh^ˈ, ni^ˈ koo^ˈ mi' punto he- na-sii^ˈ kih^ˈ ley ja^ˈ, jää^ˈ sa^ˈ ga^ˈ ma-lä-te^ˈ ga-lä-jē^ˈ. 19 He- lä-ja^ˈ, cher'mah^ˈ cha^ˈ hi- ga^ˈhlääh^ˈ koo^ˈ mada^ˈ pih^ˈ kih^ˈ ley la^ˈ, hi^ˈ cher'mah^ˈ lä-ja^ˈ mi-tēy^ˈ dsa^ˈ, leey^ˈ jaa^ˈ hi^ˈ mäh^ˈ ba' gyuu^ˈ jmi-gyuu^ˈ jee^ˈ jmee^ˈ Dio hih^ˈ. Pero cher'mah^ˈ hi^ˈ mi-te^ˈ kih^ˈ ley hi^ˈ mi-tēy^ˈ dsa^ˈ he- mi-tey^ˈ ka-läh^ˈ, hi^ˈ he^ˈ ba' lee^ˈ jøøh^ˈ gyuu^ˈ jmi-gyuu^ˈ jee^ˈ jmee^ˈ Dio hih^ˈ. 20 He-ja^ˈ hlēēhn^ˈ hnäh^ˈ, cher'mah^ˈ sa^ˈ ga^ˈjmeeh^ˈ hnäh^ˈ cha-ga^ˈmiih^ˈ ta' llu^ˈ lä-køh^ˈ te^ˈ he- jmee^ˈ dsa^ˈ hi^ˈ kyuu^ˈ ley hi^ˈ lä-køh^ˈ te^ˈ he- jmee^ˈ fariseo, sa^ˈ le^ˈ go'taah^ˈ hnäh^ˈ gyuu^ˈ jmi-gyuu^ˈ jee^ˈ jmee^ˈ Dio hih^ˈ.

Figure 13. Chinantec tone marks: VERTICAL BAR, SLASH, DOT VERTICAL BAR, DOT SLASH, DOT BAR, LOWER RIGHT CORNER ANGLE (Liga Bíblica, 2003, p. 11).

kih^ˈ dsa^ˈ. Dsa^ˈ he^ˈ laa^ˈ lē^ˈ laa^ˈ borrego hij^ˈ sa^ˈcha^ˈ juu^ˈ He-ja^ˈ ga^ˈløøy^ˈ heeyh^ˈ dsa^ˈ cha^ˈmiih^ˈ juu^ˈ. 35 Ma^ˈga^ˈhloo^ˈ ya^ˈnēē^ˈ discípulos koo^ˈ kih^ˈ Jesús. Ja^ˈ jēyh^ˈ siyh^ˈ Jesús: —Ga^ˈhloo^ˈ ba^ˈ la^ˈ jää^ˈ, sa^ˈcha^ˈ he^ˈ kuh^ˈ dsa^ˈ cha^ˈ jee^ˈ la^ˈ ni^ˈ miih^ˈ. 36 Chiuh^ˈ dsa^ˈ wa'lla^ˈnēē^ˈ dsa^ˈ lē^ˈkaa^ˈ jēh^ˈ na^ˈ, wa'lla^ˈlēy^ˈ he^ˈ kuyh^ˈ —jēh^ˈ discípulos siyh^ˈ Jesús. 37 —Hnēh^ˈ ba^ˈ kwa^ˈ he^ˈ kuyh^ˈ —jēh^ˈ Jesús. —¿Ha^ˈ lē^ˈ le^ˈ jnēēh^ˈ kih^ˈ ga^ˈlē^ˈjää^ˈ dsa^ˈ na^ˈ Lloo^ˈ koo^ˈ to^ˈ hña^ˈloo^ˈ denario lii^ˈ jnēēh^ˈ —jēh^ˈ discípulos. 38 —¿Joh^ˈ ba^ˈ he^ˈñiih^ˈ kyeh^ˈ hnēh^ˈ? Go^ˈjää^ˈ da^ˈ —jēh^ˈ Jesús. Ma^ˈga^ˈjää^ˈ discípulos, ja^ˈ jēyh^ˈ:

Figure 14. Chinantec tone marks: BACKSLASH, VERTICAL BAR, SLASH, DOT VERTICAL BAR, DOT SLASH, DOT BAR, LOWER RIGHT CORNER ANGLE (Liga Bíblica, 1990, p. 20).

MODIFIER LETTER CHINANTEC TONE MARK VERTICAL BAR is similar to U+02C8 MODIFIER LETTER VERTICAL LINE, but these can co-occur (U+02C8 would be used in dictionary entries to mark stressed syllables) and need to be distinguishable in encoding, and also be visually distinct. (There is also a horizontal-bar tone mark, but this can be represented as U+02C9 MODIFIER LETTER MACRON.)

LATIN CAPITAL LETTER B WITH STROKE is an orthographic character used for at least two languages

of Vietnam (Jorai and Katu). Lower case is represented in Unicode as U+0180 LATIN SMALL LETTER B WITH STROKE.

22 Mưi tangai dyoq, **B**ok manưih đư
dyoq tọt hót tóor atông gamak đah tồh,
đư kah ơi chaang atông, pi đó kiêng chik
chêk lăi Yêsu. Pi đó lăi mabeq buông
dyoq tadang. Pi đó nal Yêsu kah vôôiq
điq buông lợq manưih đư ting huiq lóom
đó; pi đó nal manưih đư ting huiq lóom
Yêsu, vôôiq mưi pi đó, hadyoq pi đó kah
nal leq Yêsu. 23 Lợq manua, nôôq buông
bral tâaq danuôr Têbêri, đăn điq sâng
ngaan manưih ơi cha beng mi lợq
kadóong, bêêl Yêsu ma-ơ mangaai Dyaang
Pleng. 24 **B**ok nôôq manưih đư ơi chêk
Yêsu, lăi kah ve dyoq đó, lợq nôôq đư
ting huiq lóom đó, hadyoq pi đó dyók

Figure 15. LATIN CAPITAL LETTER B WITH STROKE (Katu, 1978, p. 119).

<p><u>Tơlơi Hrăm Mrô Sa</u></p> <ol style="list-style-type: none"> 1. Ih hũa bống lai h ka? 2. Káo hũa bống lai h yoh. Káo phrăo hũa bống giông. 3. Ih Bống ănăm hơget? 4. Káo bống ănăm akan. Káo hũa asoi soh đư đư hơmăo ănăm ôh. 	<p><u>Bài thứ nhất</u></p> <p>Anh ăn cơm chưa? Tôi đã ăn rồi. Tôi mới ăn xong. Anh ăn với rau gì? Tôi ăn với cá. Tôi ăn cơm mắt mỡ; không có rau gì.</p>
<p><u>Tơlơi Hrăm Mrô Dua</u></p> <ol style="list-style-type: none"> 1. Ih hũa bống lai h ka? 2. Káo aka hũa ôh. 3. Rokáo kơ ih bống hũa hăng sang ăn káo mưi mưi anai hơ. 4. Ó, đui mưi. 	<p><u>Bài thứ hai</u></p> <p>Anh ăn cơm chưa? Tôi chưa ăn. Mời anh ăn với gia-dinh tôi tối nay. Được.</p>
<p><u>Tơlơi Hrăm Mrô Klăo</u></p> <p>Bống đổng wô.</p>	<p><u>Bài thứ ba</u></p> <p>Ăn thêm đi.</p>

Figure 16. LATIN CAPITAL LETTER B WITH STROKE (SIL, 1976, p. 1).

LATIN CAPITAL LETTER U BAR is used orthographically for a number of languages around the world. These include the Mesem and Melpa (see Figures 18 and 19) languages of Papua New Guinea, Sayula Popoluca of Mexico, the Badwe'e language of Cameroon (see Figure 20), the Budu language of Democratic Republic of Congo (see Figure 17), Comanche (see Figure 21) and Arhuaco of Colombia (see Figure 30). Lower case is represented in Unicode as U+0289 LATIN SMALL LETTER U BAR.

19 Kija malaika wa:mutisio bɔ: «Emɛ ma Gabhilieli. Emɛ kɔ
maakyia bumai apɛɛ ɛ Asɔbi. Iyo unatuma nɔwe kubia isau yɛɛ
ido ngia. 20 ɛna, moni ngika kɛsɛmɛ ka=kyiania akɔ wati wɔngɔ.
Nɔu wubhaya kɔmɔ moni ngika kɛsɛmɛ, wa-bio tɔ kɔ ijangi;
wambukubio tii bata kɔyaka kubɔya bisi ongo tɔ kɔ isiya moni
ngikunani ka=kyiania-ɔɔ.»

Figure 17. LATIN CAPITAL LETTER U BAR (WBT, 2000, p. 2).

kupa indimp nump mana ur," nitim.
nga upam. 20 salmoa ni koma nila mek purungi
Bawas nga upam, woti IK PENG KUL
Rump
1. Jisas Kraist nga anda kouwamin mba
2. Al ant ola ui orunga wu tama kat
3. Wamp nu tindi wu Jon Got nga ik pul nimba ngurum
4. Jisas Seitan ndi ola tapa kundram
5. Komnga tila ola mba wamp ik mbo inditim
6. Wamp karapa mba tap rapindumin
7. Kasa ndi wamp mat nga lump tak numpi nimba ik mbo inditim

Figure 18. LATIN CAPITAL LETTER U BAR (BSPNG, 1995, p. 1).

WU JONA KUN
AMP RUT
RAE NGA TIMAN

Figure 19. LATIN CAPITAL LETTER U BAR (WHBL, 1988, Cover).



Figure 20. LATIN CAPITAL LETTER U BAR (WBT, 2003, Back of front cover).

<p>U</p> <p>ubia (interj) oh! oh my! (exclamation of surprise used by women only). Ubia, kimaru maru. Oh my! They're coming. See ha?ii, yaa.</p> <p>uhpuitu (v), <i>pl</i> uhkooitu sleep. Ohnaa tsa? uhpuikatü. The baby is asleep. Kahnikuhatu tsa? sumu uhkoihkatü. Everyone in the house is asleep.</p>	<p>Uku tomopy (n) New Year's Day.</p> <p>ukubitsi (adj) young. Ukubitsi ma? oru nai?bi. That girl is young.</p> <p>Ukui yuba mua (n) August (<i>lit</i> new fall month).</p> <p>ukunaa (adv order) first. Ukunaa ma iiku. Let him in first.</p> <p>ukunanakatü (adj) young, youthful. Tuinuhpu? tsa? uru ukunanakatü. The boy is just young.</p> <p>uku numunuu, uku numu roopy-</p>
--	---

Figure 21. LATIN CAPITAL LETTER U BAR (Robinson, 1990, p.142).

2 Allah wir kwimɽu
 Kerreny nyiingə lurtunəgəno
 Kirem nyiingə limaamigna
 @alshaan shunu
 Nyiingə liɽimica rom-rom.

Figure 25. LATIN CAPITAL LETTER AT (Sudanese Church of Christ, 1998, p. 141).

Ŋiɽaŋal Ŋəthi Kiɽaab Kirllinəɓ Ter.

Isha@ya 45:5

Nyii kwir Kweeleny. Nə Allah ere naani wir ter
 kwokwony mac, illi ninyoro nyii ɽɽɽk.

Figure 26. LATIN SMALL LETTER AT (Frontline Fellowship, 1993, p. 3).

Ŋiɽaŋal ŋisaaw ŋəthi faathi

Ethi ne@mə kwəthi Kweeleny kwəri kwəni Yəcu
 Kwərɽɽɽɽ,

Nə ŋimɽi ŋəthi Allah tok,
 Nə ŋɽmat ŋəthi ɽigɽim ɽirllinəɓ ter tok,
 Ethi nanni dəŋgəri nə dok-dok.

Imba.

Figure 27. LATIN SMALL LETTER AT (Sudanese Church of Christ, 1998, p. 143).

The characters LATIN SMALL LETTER E WITH STROKE and LATIN CAPITAL LETTER E WITH STROKE are used orthographically in the Southeastern Tepehuan language of Mexico.

—Gatuuk dir bha jim gui' nax io'm jir ge'kam,
 cham tu' pui' na jax aañ, na ni ji'x kuñ cham jir
 jupaabkamu'n nañ dumaartii tu' suuska'n jup ba
 kookoxdha' git. Ba' aañ xib suudai' ki'n pix ka jam
 bopkon, gu' ba' gui' na pai' dhuuk ya aaya', pui' ji
 buusnia' ip na jam bopkona', na gu' ba jam

Figure 28. LATIN SMALL LETTER E WITH STROKE (SIL, 2000, p. 5)

Although there are currently no publications using LATIN CAPITAL LETTER E WITH STROKE, it should be encoded as it is an orthographic character. A dictionary is currently in the process of being published which uses LATIN CAPITAL LETTER E WITH STROKE.

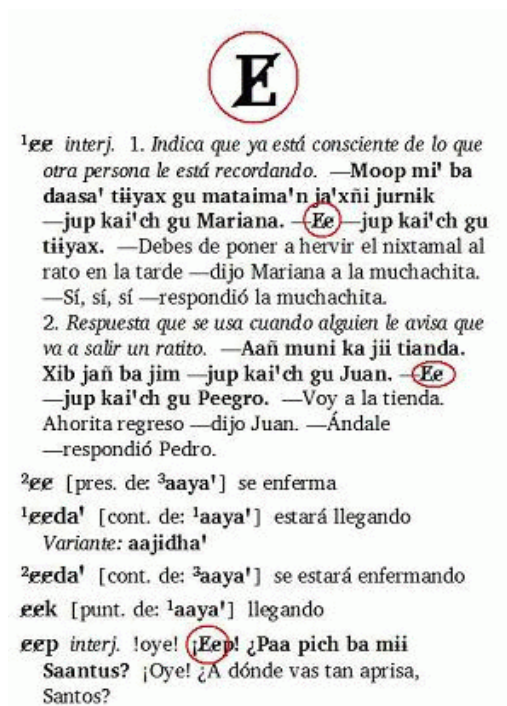


Figure 29. LATIN CAPITAL LETTER E WITH STROKE (Ramírez Solís, (in press), p. 53)

The characters LATIN SMALL LETTER J WITH STROKE and LATIN CAPITAL LETTER J WITH STROKE are used orthographically in the Arhuaco (Ethnologue code: ARH) language of Colombia. It is being used for a voiced alveopalatal affricate like the j of “just” in English. It has been used for perhaps the last 10 years, but little literature has been produced to date. The letter “j” also occurs in Arhuaco for borrowed words from Spanish.

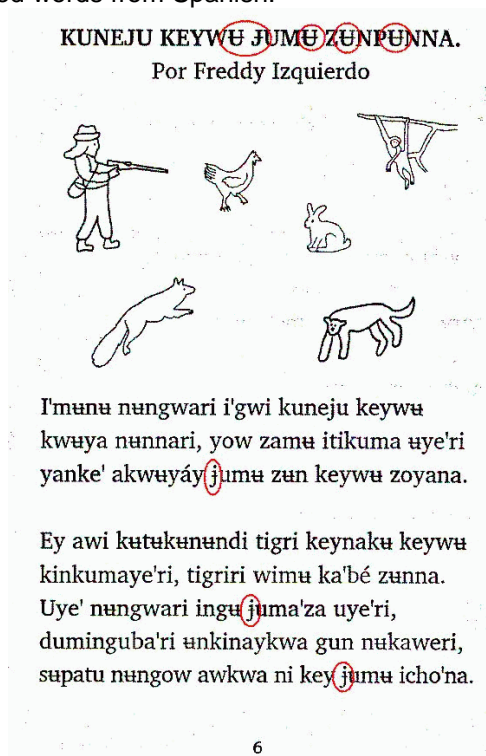


Figure 30. LATIN SMALL LETTER J WITH STROKE and LATIN CAPITAL LETTER J WITH STROKE (Fundación para el Desarrollo..., 2004, p. 6).

LATIN SMALL LETTER Q WITH HOOK TAIL and LATIN CAPITAL LETTER SMALL Q WITH HOOK were introduced by early Lutheran missionaries for the Numanggang language of Papua New Guinea (1930s or 1940s). After a number of decades of use, in 2002 the community decided to discontinue using LATIN SMALL LETTER Q WITH HOOK TAIL and LATIN CAPITAL LETTER SMALL Q WITH HOOK, but they are found in existing publications such song books and liturgy materials. They are needed for electronic publishing of archived documents. They are also found in Kâte language (distantly related to Numanggang) literature. As shown in Figure 31 and 33 “q” also appears in the literature and could not be considered an alternate form.

Eme ηonden-jeηicko ηokac moc juwec eki soc rohuc 25
 juhame jara 12 ewec. ı Eme dokta bocjahazi erao gie bāi- 26
 piehapie zāic wemo afecne manahuc juwec. Ā irec furi
 moneη mafa-ticne jarekume hesackewec. Eme sifu huc-
 ticne mi afeckehuc mâc qârâηkecnewec. ı Eme ηokac eki 27
 hâcne Jesure biηe manarâ ηic ηokac gâcne jaηere ηonden-
 ko juhuc râpefuηnao rarâ ηaqi-ticne bâuawec. ı I zîηuc 28
 muhuc ewec: «Ĵaqi-ticne sawa bâuarâ bianuocmu.» ı
 Bâuame izia socticne zickeme sahac-ticne bâgunuckeme 29

Figure 31. LATIN SMALL LETTER Q WITH HOOK TAIL (BFBS, 1965, p. 105)

Wofuη Jesu Kristo

Ere Miti Qârî Zi

Figure 32. LATIN CAPITAL LETTER SMALL Q WITH HOOK TAIL (BFBS, 1965, Title page).

OL SAMTING ı STAP LONG DISPELA BUK			
Sampela Leta Yu No Save			
Lesen 5	ŋ ŋ	37
Lesen 6	Q q	42
Lesen 7	Q, q	46
Lesen 8	Wok Bilong Traim	Qq/Qq	50

Figure 33. LATIN SMALL LETTER Q WITH HOOK TAIL and LATIN CAPITAL LETTER SMALL Q WITH HOOK TAIL (Hynum, 1989, p. 3).

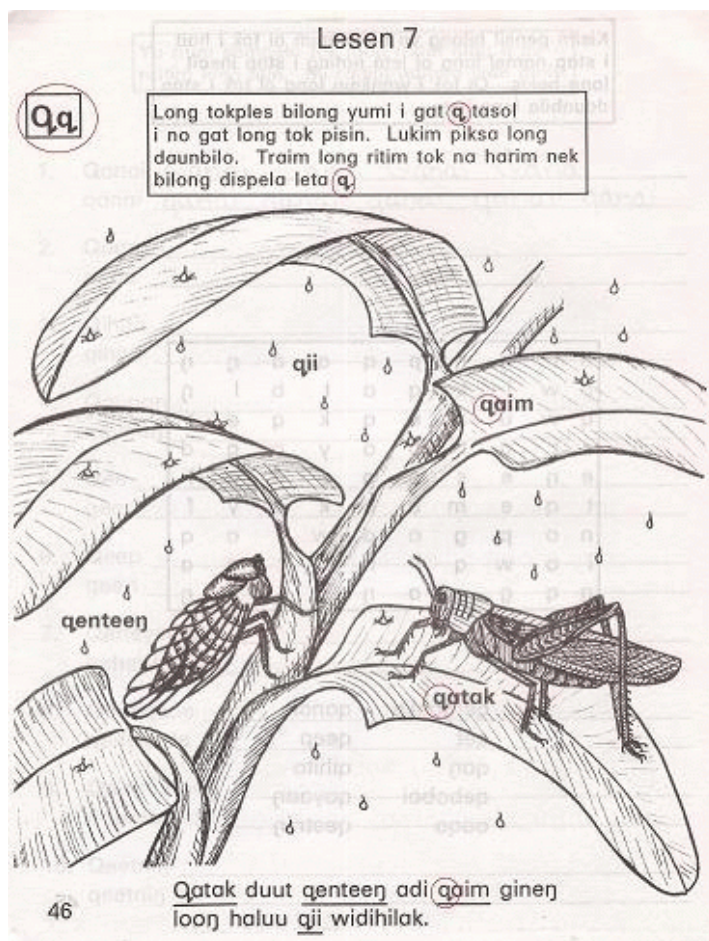


Figure 34. LATIN SMALL LETTER Q WITH HOOK TAIL and LATIN CAPITAL LETTER SMALL Q WITH HOOK TAIL (Hynum, 1989, p. 46).

LATIN SMALL LETTER R WITH STROKE and LATIN CAPITAL LETTER R WITH STROKE are used in the Kanuri orthography of Niger.

deri ngaye napkeno. Amde ci beremben daada. Misalnin
awo kada tayiro wulyeyeno. Curo mananjuyen: «Semondo
kagangowo, mana fanowo. Kam laa algama jejerduro kiluwo.
Jejerji jejerji adagai duwon, kasuni laa jawalnin fiyada. Daji
ngudoso lsane cadore sel gewo. Kasuni laa ye na kau kaua
katti ngewu bawoalan fiyada. Katti ngewu bawodero, nagatro
algamade fiyeno. Riyenniya, kausu jauje, ferdunju duno bawo
nangaro fimje kiau ngamgeno. Kasuni laade na ngiwian
fiyada. Ngiwide wuraje algamadea kimicciye kela cakkenni.
Amma kasuni laade na katti ngelaan fiyada. Algamade wuraje
nema tai daje kekko. Laa tadanju fiakku, laa firakku, laa mia
mia,» yeno. Daji isa temojiye: «Kam semonjuamade semonju
kagakco fanjo!» yeno.

Figure 35. LATIN SMALL LETTER R WITH STROKE and LATIN CAPITAL LETTER R WITH STROKE (SIL, 1998, p. 7).

The characters LATIN SMALL LETTER Y WITH STROKE and LATIN CAPITAL LETTER Y WITH STROKE are used orthographically in the Lubuagan Kalinga language of the Philippines. The Lubuagan people are strong in their determination to keep this symbol. No examples of the LATIN CAPITAL LETTER Y WITH STROKE are available: the y-stroke is never word-initial in Kalinga text, and so the capital would not occur in most body text. Like any Latin characters used orthographically, however, it can occur in document elements set in full caps, and so must also be encoded.

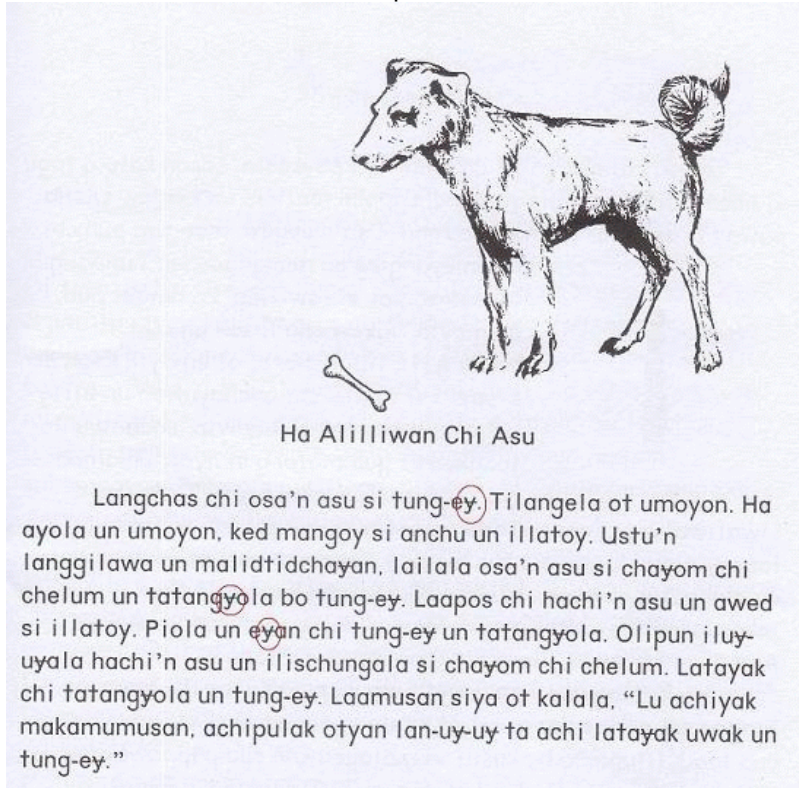


Figure 36. LATIN SMALL LETTER Y WITH STROKE (SIL, 1998, p. 1).

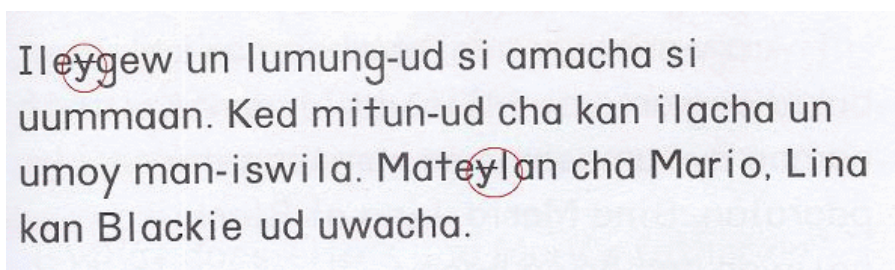


Figure 37. LATIN SMALL LETTER Y WITH STROKE (Pateuweg, 1998, p. 17).

E.3 Phonetic Extensions

LATIN SMALL LETTER V WITH CURL is a phonetic symbol used to indicate a labiodental flap. It is not approved IPA usage, yet it has been widely used, and is in current use, particular among Africanist linguists.

vb Voiced labiodental flap, made by moving the lower lip inwards behind the upper teeth and then flapping it against the upper teeth outwards (phonetic symbol **v̥**).¹ This sound seems to occur in ideophones only, e.g. *bávbú* (describing sud-

Figure 38. LATIN SMALL LETTER V WITH CURL (Hoffman, 1963, p. 25).

In Niger-Kordofanian almost all known examples are from Adamawa-Eastern. Once more Tucker and Bryan describe it as most common in the Ndogo-Sere group, examples being given from Sere, Mundu, Ndogo and Bai and once more in common words. Particularly interesting is its occurrence in Ndogo, Sere and Bai in the word of 'child' **v̥** an obvious cognate of the widespread Niger-Congo root *bɛ*. Richardson gives an example from an Adamawa language Kapere, a southern dialect of Mbum in which it occurs in the second person plural pronoun **v̥**. He also notes its occurrence in Ngbaka Mabo, a form of Mbaka Limba, and in the Yangere dialect of Banda, both of these once more languages of the Eastern division of Adamawa-Eastern.¹⁸

Figure 39. LATIN SMALL LETTER V WITH CURL (Greenberg, 1983, p. 11).

Sounds (DGBAKA MABO)

There are five phonemes **i**, **e**, **a**, **ɔ**, **ɯ** with subsidiary members in unstressed syllables, also **ɔ̃**, **ẽ**, **ã**. Amongst the consonants the following are noteworthy: **kɸ**, **gb̥**, **ɾ** (flapped **r**) **v̥** (flapped **v**), and **v** with syllabic value, e.g. *kɸázé* 'sleep', *nzɔg-bɛlé* 'knee', *ɯɾɯsé* 'woman', *v̥jíná* 'nine', *v̥vé* 'five'.

Tone. Both nominals and verbals may be categorized according to tone, which often distinguishes between lexical entities, e.g. *kɸáná* 'branch', *kɸáná* 'leaf'.

Word shape. No closed syllables occur. Simple words are generally of the shape CV or CVCV. In the latter case, when the second consonant is **l** or **ɾ** the two vowels are the same. VCV is very rare, e.g. *ɯwá* 'fire'. Compound words are very common, e.g. *kɸá-ná* 'branch' ('hand' + 'tree').

Figure 40. LATIN SMALL LETTER V WITH CURL (Richardson, 1957, p. 91).

321. 5. The characteristic consonant sounds **kp** and **gb**³ are found everywhere; 'b' and 'd' are found everywhere except in Zande. In addition a flapped l-sound is heard throughout (usually as a variant of **l** or **r**),⁴ while Ndogo and Kreish pronounce a flapped **v**-sound.

	MORU	BONGO	NDOGO	ZANDE
kp	ekpɪ (to cough)	ekpɪ (to help)	kpɪ (sour)	tikpo (salt)
gb	tómgbó (canoe)	'dugba (to catch)	gbɪ (to strike)	gbia (chief)
'b	'ba (home)	'be (home)	'bá (home)	
'd	la'di (to cook)	f'dí (to cook)	'dé (to speak)	
ɾ	kɾiɾf (charcoal)	ɾingədò (rhinoceros)	ɾɪɾ ⁵ (sweet)	ringara ⁶ (country)
v̥			v̥ (boy)	
			gevé (arrow) (Kreish)	

¹ See vocabulary in Fr. Ribero's *Elementi di Lingua Ndogo*.

² Unless, of course, drum signalling has been borrowed from a di-tonal people and adapted to Zande. Note that the Banda drum signalling analysed by the author in the Western District is tri-tonal.

³ Relatively rare in Moru-Madi, however.

⁴ **ɾ** is the normal sound in Zande, alternating with **r** and **l**. It is spelt **r**.

⁵ In Sere, where **ɾ** is mostly to be found.

⁶ Pronounced *ɾingara*.

Figure 41. LATIN SMALL LETTER V WITH CURL (Tucker, 1940, p. 65).

F. References

- Bible Society of Papua New Guinea. 1995. *Got Nga Nombuta Rung Ik Kai*. The New Testament. Port Moresby-Lae: The Bible Society of Papua New Guinea.
- Bible League. 1990. *Moreni Wam Makge KAyekgātn*. The Gospel as Written by Mark. South Holland, IL: The Bible League.
- British and Foreign Bible Society. 1965. *Wofuŋ Jesu Kristo Ere Miti Qāri ʒi*. The New Testament in Kâte. London: British and Foreign Bible Society.
- Burquest, Donald A. Second Edition; 1998. *Phonological Analysis, a Functional Approach*. Summer Institute of Linguistics.
- Frontline Fellowship. 1993. *Katikisim*. (Koulib Catechism). Nairobi, Kenya: Sudan Literature Centre.
- Fundación para el Desarrollo de los Pueblos Marginados. Edición provisional, 2004. *Ga'kɨnamɨ Rinre'takɨn Migwa'sukweyna*. Palabras que nos llevan a reflexionar. Bogotá, Colombia: Fundación para el Desarrollo de los Pueblos Marginados.
- Gilley, Leoma G. 1992. *An Autosegmental Approach to Shilluk Phonology*. Dallas: Summer Institute of Linguistics and The University of Texas at Arlington.
- Goldsmith, John A. (ed.). 1995. *The Handbook of Phonological Theory*. Oxford: Blackwell Publishers.
- Greenberg, J. 1983. *Some areal characteristics of African languages*. In Dihoff, I.R. (editor), *Current Approaches to African Linguistics*, 3-22. Dordrecht: Foris Publications.
- Hoffman, C. 1963. *A grammar of the Margi language*. London: Oxford University Press (for the International African Institute).
- Hynum, Barbara (et. al.). 1989. *Olanan! Tokples Bilong Mi I Swit Moa!* Transfer Primer. Ukarumpa, EHP, Papua New Guinea: Summer Institute of Linguistics.
- International Bible Society. 1982. *Kushia Unim ahwu 'ufufe*. The New Testament. East Brunswick, NJ: New York International Bible Society.
- International Phonetics Association. 1999. *Handbook of the International Phonetic Association: a guide to the use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press.
- Katu New Testament. 1978. *Taruuih Ađô Truuih Yêsu Krit*. Katu New Testament. Việt Nam: no publisher listed.
- Kutsch Lojenga, Constance. 1993. *Ngiti: A Central-Sudanic Language of Zaire*. Köln: Rüdiger Köppe Verlag.
- Lembaga Alkitab Indonesia. 2003. *Binohe Susi Baŋo Mapia*. Jakarta, Indonesia: Lembaga Alkitab Indonesia.

- _____. 2002. *Habaré'u Raḷuasé Pudariandi Wuhu*. Jakarta, Indonesia: Lembaga Alkitab Indonesia.
- _____. 2003b. *Wuke Susi Wattu'u Aḷaḷuassa*. Jakarta, Indonesia: Lembaga Alkitab Indonesia.
- Liga Bíblica. 2003. *Hmooh' hmëë' he' ga'jmee' Jesucristo Salmos*. El Nuevo Testamento y Salmos. México: La Liga Bíblica, A.C.
- _____. 1990. *Libro he' ga'si' Marcos la' lle' vida kih' Jesús*. El Evangelio según San Marcos en el chinanteco de Ozumacín. México: La Liga del Sembrador, A.C. y La Liga Bíblica Mundial del Hogar.
- _____. 1979. *Ruucasi Ooja Istumaasi Ivuaadana Suusicristo Aimaraitai Oidi Daama*. El Evangelio de Jesucristo Segun San Lucas. México: Liga Biblica Mundial del Hogar.
- Pateuweg, Josephine. 1998. *Si Titu*. Trilingual story book for Grade 1 students. Manila, Philippines: Summer Institute of Linguistics. Published in cooperation with the Commission on Philippine Languages and the Department of Education, Culture and Sports.
- Ramírez Solís, Cornelio, et al. (in press). *Diccionario tepehuano de Santa María Ocotán, Durango*. (Versión Preliminar.) D.R. por el Instituto Lingüístico de Verano, A.C., México, D.F.
- Richardson, I. 1957. *Linguistic survey of the northern Bantu borderland*. Volume 2. London: Oxford University Press (for the International African Institute).
- Robinson, Lila Wistrand, and James Armagost. 1990. *Comanche dictionary and grammar*. (Summer Institute of Linguistics publications in linguistics, 92.) Dallas: Summer Institute of Linguistics and University of Texas at Arlington.
- Sudanese Church of Christ. 1998. *Rilḡa Rëthi Allah*. (Hymnal in the Koalib language of Sudan). Nairobi, Kenya: Sudan Literature Centre.
- Summer Institute of Linguistics. 1998. *Cha Istorya si Sin-eygew*. Traditional Stories of Lubuagan. Written by school teachers of Lubuagan District. Manila, Philippines: Summer Institute of Linguistics. Published in cooperation with the Commission on Philippine Languages and the Department of Education, Culture and Sports.
- _____. 1976. *Jarai Language Lessons*. Summer Institute of Linguistics.
- _____. 1998. *Lawar Këji Isa Almasihuye*. Niamey, République du Niger. SIL.
- _____. 2000. *Na ja bopkon gu ja'tkam gu Juan na ba' mi' ba aaya' guch Xoi'kam*. (2nd edition.) Las Sagradas Escrituras para todos. Publicado por La Liga Bíblica, A.C., México, D.F.
- Tucker, A.N. 1940. *The Eastern Sudanic languages*. Volume 1. London: Oxford University Press (for the International African Institute).
- World Home Bible League. 1988. *Wu Jona Kən Amp Rut Ra-I Nga Timan*. The stories of Jonah and Ruth. South Holland, IL: World Home Bible League.

Wycliffe Bible Translators. 2003. *Kálara Zyěm: Edwán é mpún*. La Nouveau Testament en bádwe'e. Wycliffe Bible Translators.

_____. 2000. *Isa yee Ido ngia yi LUKA, Evangile selon Luc*. Orlando, FL: Wycliffe Bible Translators. Distribution: Projet de Traduction de la Bible et d'Alphabétisation en Langue Budu, CECCA/16, B.P. Isiro. République Démocratique du Congo.