

Universal Multiple-Octet Coded Character Set
 International Organization for Standardization
 Organisation Internationale de Normalisation
 Международная организация по стандартизации

Doc Type: Working Group Document

Title: Proposal to encode a Latin Capital Letter L with Belt

Source: Dept. of Linguistics & TESOL, The University of Texas at Arlington, USA

Author: Joshua M Jensen, Karl Pentzlin

Status: Expert Contribution

Action: For consideration by JTC1/SC2/WG2 and UTC

Date: 2012-02-08

1. Introduction

1.1 Use in the Alabama (Alibamu) language (akz)

The Alabama language (ISO 639-3: akz, also known as Alibamu) is spoken by about 200 people in south eastern Texas, USA. Its Latin orthography uses the letter U+026C LATIN SMALL LETTER L WITH BELT.

The orthography uses capital letters, in a way similar to English (see e.g. fig. 3). As words can start with l with belt (see fig. 2), the capital letter is needed.

As the capital form of this letter is not yet contained in Unicode until now, it is proposed here.

Examples of use are given in fig. 1–4.

The proposed letter is needed for an ongoing project at the University of Texas at Arlington, USA, regarding the Alabama language. Therefore, the letter was separated from a not yet submitted proposal draft collecting letters for orthographies of North American indigenous languages not yet included in Unicode, to enable its availability without delay.

1.2 Use in other North American indigenous languages

Fig. 5 shows the use in a dictionary for the Koasati Language (ISO 639-3 cku).

Also, according to information compiled on <http://www.languagegeek.com> (retrieved 2010-04-26), the l with belt is used in following languages:

Coast Salish: Hən̓q̓əmin̓əm (Halkomelem; ISO 639-3: hur)

Coast Salish: Nəx̣ʷṣʷəỵəm̓úcən (Klallam; ISO 639-3: clm)

Coast Salish: ʔayʔəjuθəm/Saʔuʔtx̣ʷ (Sliammon, Comox, Homalco, Klahoose; ISO 639-3: coo)

While the use of capital letters is not directly attested in the sources given there, it is used in capitalized English texts using words of these languages in their original spelling (see fig. 6). It is also presumed that any proper names used by their original spelling in English text will have their first letter capitalized, independent whether this would occur in the original orthography of the source language.

2. Proposed Character

Block: Latin Extended-D

Addition for North American indigenous orthographies

Ł U+A7xx LATIN CAPITAL LETTER L WITH BELT

• lower case is U+026C LATIN SMALL LETTER L WITH BELT

Properties:

A7xx;LATIN CAPITAL LETTER L WITH BELT;Lu;0;L; ; ; ; ;N; ; ; ;026C;

Properties to be changed for existing characters:

026C;LATIN SMALL LETTER L WITH BELT;Ll;0;L; ; ; ; ;N;LATIN SMALL LETTER L BELT;;A7xx;;A7xx

3. Examples and Figures

Fig. 1: Alabama alphabet, retrieved 2012-02-08 from:
<http://www.ling.unt.edu/~montler/Alabama/Dictionary/AIAlphabet.htm>

The Alabama alphabet and key to pronunciation

Letter	Letter Name	Example	Translation
a	eyka	intakba	`stomach'
b	biika	kolbi	`basket'
ch	chiika	chaaha	`tall'
f	efka	ifa	`dog'
i	iika	nipo	`meat'
k	kiika	akaaka	`chicken'
l	elka	laana	`yellow'
ł	łihka	łato	`fish'

m	imka	mikko	`chief'
n	inka	naani	`man, male'
o	ooka	oki	`water'
p	piika	paani	`creek'
s	eska	sobayli	`to know'
t	tiika	taata	`father'
w	dabiyo	wohka	`to bark'
y	wayka	yimmi	`to believe'

Fig. 2: Retrieved 2012-02-08 from: <http://www.ling.unt.edu/~montler/Alabama/Dictionary/>
 Note the "?" in the letter selection line where a capital L with belt belongs.
 As this character is missing in Unicode, the representation of that letter at this place did not survive the production of the Internet page.
 The page is shown as it appeared after clicking on that character.



Fig. 3: From p.79 (Alabama chapter) of:
 Hardy, Heather K., and Janine Scancarelli (eds.) 2005. Native Languages of the Southeastern United States. Native Languages of the Southeastern United States. Lincoln, NE: U of Nebraska Press.
 Showing a capital L with belt as the first character.

(20) **𐌌óok** *chokfikok,* *nahokon*
𐌌óka-t *chokfi-ka-o-k* *naho-ka-o-n*
 again-T rabbit-DEM-DET-K AUX-DEM-DET-N

paani *chobakon* *chofotlit* *abanàtlok,*
pahni *choba-ka-o-n* *chofot-li-t* *abanàt-li-o-k*
 stream big-DEM-DET-N jump.once-LI-T cross-LI-FOC-K
 once again Rabbit, that other one, that big stream, he jumped across
 it and

Fig. 4: Explanation of use of L with belt, on p. viii of Dictionary of the Alabama Language (Sylestine, Hardy, and Montler), University of Texas Press, 1993.

Only one letter has had to be added to the usual English letters. This letter represents a special kind of /l/ that Alabama has in addition to a regular /l/ (this sound is described in more detail in section 1.4). The standard symbol for this sound is an /l/ with a sort of 'bow' or bar across the middle [ɫ] (capitalized as Ľ) and its name is 'barred l'. This letter can be

Fig. 5: Excerpt from p.128 of a dictionary for the Koasati Language (ISO 639-3 cku): Kimball, Geoffrey D. 1994. Koasati Dictionary. University of Nebraska Press.

<p>their arrows on it and completely wounded it. inʔaki-ikson V, no class to be stingless. inʔaki:sa N stinging insect. [Slow pronunciation is /inʔakiʔi:sa/.] ʔaliká N homesickness. ʔató N fish. ʔatobá á:tipa N shark. ʔatbómma N redfish. ʔati:pa nihtá N Good Friday. [Lit. fish eating day. Refers to the food eaten by the local French Catholics due to the rules of fast and abstinence.] ʔatlawísta N, pl minnows sg ʔatosí. ʔatlocá N buffalo fish. ʔatlocák cofkoninánnat acihbá:hoscok ilpá. <i>The buffalo fish is bony, and is extremely hard to eat.</i> ʔató bitka N fish dance. [Swanton] ʔató imahissí N wormseed <i>Chenopodium ambrosioides.</i> [Folsom-Dickerson] MEDICAL USE: Worms - the leaves</p>	<p>ʔató stímokwáyka incofá N fishhook. ʔató stintálka N weir. ʔatokkí N, AM- kidney. ʔatópi N ringworm. ʔatóʔpik lapá:lin SV2AII, CA- to have ringworm. ʔatóʔpik calapá:l. <i>I have ringworm.</i> ʔána N horsefly. ʔaská:cin V3B vn ʔaskacilk ʔaská:ci to splash someone passing. ʔaslí:cin V2B vn ʔaskacá an ʔaslí:ci to splash out. Okhayihmofihnask, mobí:lak ʔaslí:ci atá:y. <i>Because it is overly muddy, the car goes splashing.</i> ʔátka N snorer. ʔátkan V3A vn ʔathilká an ʔátka to snore. ʔatóʔfin V2AII vn ʔatóʔka an ʔatóʔfi 1 to melt something. Tokohlí:ci anó:ci:fó:kok, campo:l-okcí ikbá:ci ʔatóʔci:fó:kok o:cikkí:litiska:p paʔá:ta stonohkí:con. <i>When you have popped it all, when you have warmed and melted syrup, you can</i></p>
---	--

Fig. 6: from <http://www.languagegeek.com/salishan/sliammon.html> , retrieved 2010-04-26

ʔAYʔAJUΘƏM – SAƷUƷTX^W
 COMOX – HOMALCO – KLAHOOSE – SLIAMMON

 [ʔayʔajuθəm/SaƷuƷtx^w Keyboards are available](#)

**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:	<i>Proposal to encode a Latin Capital Letter L with Belt</i>
2. Requester's name:	<i>Joshua M Jensen, Karl Pentzlin</i>
3. Requester type (Member body/Liaison/Individual contribution):	<i>Expert contribution</i>
4. Submission date:	<i>2012-02-08</i>
5. Requester's reference (if applicable):	<i>Dept. of Linguistics & TESOL, The University of Texas at Arlington, USA (J. J.)</i>
6. Choose one of the following: This is a complete proposal:	<input type="checkbox"/> Yes
(or) More information will be provided later:	<input type="checkbox"/>

B. Technical – General

1. Choose one of the following:		
a. This proposal is for a new script (set of characters):	<input type="checkbox"/> No	
Proposed name of script:		
b. The proposal is for addition of character(s) to an existing block:	<input type="checkbox"/> Yes	
Name of the existing block:		
2. Number of characters in proposal:	<input type="checkbox"/> 1	
3. Proposed category (select one from below - see section 2.2 of P&P document):		
A-Contemporary <input checked="" type="checkbox"/>	B.1-Specialized (small collection) <input type="checkbox"/>	B.2-Specialized (large collection) <input type="checkbox"/>
C-Major extinct <input type="checkbox"/>	D-Attested extinct <input type="checkbox"/>	E-Minor extinct <input type="checkbox"/>
F-Archaic Hieroglyphic or Ideographic <input type="checkbox"/>	G-Obscure or questionable usage symbols <input type="checkbox"/>	
4. Is a repertoire including character names provided?	<input type="checkbox"/> Yes	
a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?	<input type="checkbox"/> Yes	
b. Are the character shapes attached in a legible form suitable for review?	<input type="checkbox"/> Yes	
5. Fonts related:		
a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?	<i>TBD</i>	
b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):	<i>TBD</i>	
6. References:		
a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?	<input type="checkbox"/> Yes	
b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?	<input type="checkbox"/> Yes	
7. 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)?	<input type="checkbox"/> No	

8. 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 Unicode Character Database (<http://www.unicode.org/reports/tr44/>) and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

C. Technical - Justification

¹ Form number: N3902-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03, 2008-05, 2009-11, 2011-03)

1. Has this proposal for addition of character(s) been submitted before? If YES explain		No
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? If YES, available relevant documents:	 <i>See Requester's Reference</i> <i>see text</i>	Yes
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? Reference:	<i>See text</i>	Yes
4. The context of use for the proposed characters (type of use; common or rare) Reference:	<i>see text</i>	common
5. Are the proposed characters in current use by the user community? If YES, where? Reference:	<i>see text</i>	Yes
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:	 <i>To keep it in line with similar characters</i>	Yes Yes
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?		n/a
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:		No
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:		No
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:		No
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:		No
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)		No
13. Does the proposal contain any Ideographic compatibility characters? If YES, are the equivalent corresponding unified ideographic characters identified? If YES, reference:		No