

N2457

L2/02-204

ISO/IEC JTC 1/SC 2/WG 2

**PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646**

Please fill all the sections A, B and C below.

(Please read Principles and Procedures Document for guidelines and details before filling this form.)

See <http://www.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html> for latest Form.

See <http://www.dkuug.dk/JTC1/SC2/WG2/docs/principles.html> for latest Principles and Procedures document.

See <http://www.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html> for latest roadmaps.

(Form number: N2352-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09))

A. Administrative

1. Title: *Additional Double Diacritics*

2. Requester's name: *US National Body*

3. Requester type (Member body/Liaison/Individual contribution): *Member Body*

4. Submission date: *2002-05-10*

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: _____

B. Technical - General

1. (Choose one of the following:)

a. This proposal is for a new script (set of characters): _____

Proposed name of script: _____

b. The proposal is for addition of character(s) to an existing block: *YES*

Name of the existing block: *Combining Half Marks*

2. Number of characters in proposal: *3*

3. Proposed category (see section II, Character Categories): *Combining*

4. Proposed Level of Implementation (1, 2 or 3)
(see clause 14, ISO/IEC 10646-1: 2000): *3*

Is a rationale provided for the choice? _____

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 ISO/IEC 10646-1: 2000? *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?

The Unicode Consortium

If available now, identify source(s) for the font (include address,
e-mail, ftp-site, etc.) and indicate the tools used:

Asmus Freytag

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:

Proposal for 3 Additional Double Diacritics

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)?
The character behave just like the existing double diacriticals.

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 Unicode Character Database <http://www.unicode.org/Public/UNIDATA/UnicodeCharacterDatabase.html> 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. Has this proposal for addition of character(s) been submitted before? NO
If YES explain
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.)? YES
If YES, with whom? NATIONAL BODY
If YES, available relevant documents: SEE ATTACHMENTS
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? YES
Reference:
4. The context of use for the proposed characters (type of use; common or rare) INFREQUENT
Reference:
5. Are the proposed characters in current use by the user community? YES
If YES, where? Reference: ENGLISH DICTIONARIES, EDUCATIONAL USE
6. After giving due considerations to the principles in Principles and Procedures document (a WG 2 standing document) must the proposed characters be entirely in the BMP? YES
If YES, is a rationale provided? YES
If YES, reference: THERE ARE A SMALL NUMBER OF THESE CHARACTERS AND THEY
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)? YES
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence? NO
If YES, is a rationale for its inclusion provided?
If YES, reference:
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? NO
If YES, is a rationale for its inclusion provided?
If YES, reference:
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character? YES
If YES, is a rationale for its inclusion provided? YES
If YES, reference: APPEARANCE IS SIMILAR TO BREVE, ETC, BUT THE BEHAVIOR
11. Does the proposal include use of combining characters and/or use of

Proposal for 3 Additional Double Diacritics

composite sequences (see clauses 4.12 and 4.14 in ISO/IEC 10646-1: 2000)? YES

If YES, is a rationale for such use provided? YES

If YES, reference: _____

Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided? YES

If YES, reference: *SEE ATTACHMENT BELOW*

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) _____

13. Does the proposal contain any Ideographic compatibility character(s)? *NO*




If YES, is the equivalent corresponding unified ideographic character(s) identified? _____

If YES, reference: _____

L2 and the Unicode Technical Committee considered the issue of double diacritics at some length. There are three encoded double diacritics:

- U+0360 (◌̂) COMBINING DOUBLE TILDE
- U+0361 (◌̃) COMBINING DOUBLE INVERTED BREVE
- U+0362 (◌̄) {COMBINING DOUBLE RIGHTWARDS ARROW BELOW

But this does not include several double diacritics in common use in dictionaries and educational material. One proposal was to overload the COMBINING GRAPHEME JOINER to be able to represent these characters in a generative fashion using the current combining characters. Upon reflection, however, it was felt that since the set of double diacritics is rather small, and not likely to require many further additions, it would be better to complete the set with the known attested forms, which consist of the following.

	U+FE24 COMBINING DOUBLE BREVE
	U+FE25 COMBINING DOUBLE MACRON
	U+FE26 COMBINING DOUBLE LOW LINE

See the attachments for examples of usage, in Webster's Collegiate Dictionary and in the American Heritage Dictionary.

Pronunciation Key

The system of indicating pronunciations in the Dictionary is explained in the section headed "Pronunciation" in the "Guide to the Dictionary." The column below headed HD represents the pronunciation key used in the Dictionary. The symbols marked with an asterisk are discussed in the guide. Important insights into the theoretical basis of representing pronunciations as well as information on dialect variations may be obtained from the article "English Dialects" by Henry Lee Smith, Jr. The symbols in the right-

hand column, labeled T-S, are from the system of phonemes described by Professor Smith and George L. Trager and are widely used by linguists. (The Trager-Smith symbols, however, are for phonemes of English only.) The symbols are explained in Professor Smith's article. The center column, labeled IPA, contains symbols from the International Phonetic Alphabet, also widely used by scholars. The three systems do not precisely correspond, because they were conceived for somewhat different purposes.

spellings	HD	IPA	T-S
pat	ä	æ	æ
pay	ā	e	ey
care	*är	er, er	ehr, cyr, er
father	ä	ɑ, u	ah
bib	b	b	b
church	ch	tʃ	č
deed, milled	d	d	d
pet	ē	e	e
bee	e	i	iy
life, phasc, rough	f	f	f
gag	g	g	g
hat	h	h	h
which	hw	hw (also w)	hw
pit	*ī	i	i
pie, by	i	u	ay
pier	*īr	ir, ir	ihr, iyr, ir
Judge	j	dʒ	j
kick, cat, pique	k	k	k
lid, needle	*l (nēd'li)	l, l̥ [ˈnid]	l (not syllabic)
mum	m	m	m
no, sudden	*n (sūd'n)	n, ŋ [ˈsɑdŋ]	n (not syllabic)
thing	ng	ŋ	ŋ
pot, *harrid	ō	ɑ	a, o
toe, *hoarse	ō	o	ow
caught, paw, *for	ō	ɔ	oh, oh, o
noise	oi	ɔi	oy
took	oo	u	u
boot	oo	u	uw
out	ou	au	aw, əw
pop	p	p	p
roar	*r	r	r
sauce	s	s	s

spellings	HD	IPA	T-S
ship, dish	sh	ʃ	š
tight, stopped	t	t	t
thin	th	θ	θ
this	th	ð	ð
cut	ú	ʌ	ə
urge, term, firm, word, heard	*ür	ʊ, ʊr	ər, ohr
valve	v	v	v
with	w	w	w
yes	y	j	y
zebra, xylem	z	z	z
vision, pleasure,	zh	ʒ	ž
garage			
about, item, edible, *o		ɔ	ə, ɪ
gallop, circus			
butler	ər	ə	ər

FOREIGN

	HD	IPA
French ami	ä	a
French feu,	œ	œ
German schön		
French tu,	ü	y
German über		
German ich,	KH	x
Scottish loch		
French bon	N	ñ, ñ, ã, ã
French compiégne	y' (kôn-pyén'y')	ʃ

STRESS

Primary stress	ˈ	bi-ol'u-gy (bi-öl'a-jē)
Secondary stress	ˈ	bi'a-log'i-cal (bi'a-løj'i-ksl)

Note on Illustrations: Each illustration has been positioned as close as possible to the entry it illustrates. In those instances where close juxtaposition proved impossible, the illustration nevertheless always appears on the same page as the entry or on the page opposite.

MERRIAM-WEBSTER PRONUNCIATION SYMBOLS

See "Guide to Pronunciation", "Explanatory Notes", 2

- | | |
|---|---|
| <p>ə banana, collect, about</p> <p>ʊ, ɒ humdrum, about</p> <p>^a immediately preceding \l/, \n/, \m/, \ŋ/, as in battle, mitten, eaten, and sometimes cap and bells \-m/, lock and key \-ŋ-; immediately following \l/, \m/, \r/, as often in French table, prisme, titre</p> <p>ɔr operation, further, urger</p> <p>ʔr- } as in two different pronunciations
ʔ-r } of hurry \ˈhʊr-ə, ˈhʊ-rə\</p> <p>ɹ mat, map, mad, gag, snap, patch</p> <p>ā day, fade, date, aorta, drape, cape</p> <p>ä bother, cot, and, with most American speakers, father, cart</p> <p>â father as pronounced by speakers who do not rhyme it with bother</p> <p>au now, loud, out</p> <p>b baby, rib</p> <p>ch chin, nature \ˈnā-ʃar\ (actually, this sound is \t\ + \ʃh\)</p> <p>d did, adder</p> <p>e bet, bed, peck</p> <p>ē, ē beat, nosebleed, evenly, easy</p> <p>ē easy, meaty</p> <p>f fifty, cuff</p> <p>g go, big, gift</p> <p>h hat, ahead</p> <p>hw whale as pronounced by those who do not have the same pronunciation for both <i>whale</i> and <i>wail</i></p> <p>i tip, banish, active</p> <p>ī site, side, buy, tripe (actually, this sound is \ā\ + \i/, or \ā\ + \i\)</p> <p>j job, gem, edge, join, judge (actually, this sound is \d\ + \zh\)</p> <p>k kin, cook, ache</p> <p>k̄ German kch, Buch</p> <p>l lily, pool</p> <p>m murmur, dim, nymph</p> <p>n no, own</p> | <p>ⁿ indicates that a preceding vowel or diphthong is pronounced with the nasal passages open, as in French <i>un bon vin blanc</i> \œⁿ-bõⁿ-vaⁿ-blãⁿ\</p> <p>ŋ sing \ˈsɪŋ/, singer \ˈsɪŋ-ər/, finger \ˈfɪŋ-ɡər/, ink \ˈɪŋk/, thing \ˈθɪŋ\</p> <p>ō bone, know, beau</p> <p>ô saw, all, gnaw</p> <p>œ French <i>bœuf</i>, German <i>Hülle</i></p> <p>œ̄ French <i>feu</i>, German <i>Höhle</i></p> <p>oi coin, destroy, sawing</p> <p>p pepper, lip</p> <p>r rarity</p> <p>s source, less</p> <p>sh with nothing between, as in <i>shy</i>, <i>mission</i>, <i>machine</i>, <i>special</i> (actually, this is a single sound, not two); with a hyphen between, two sounds as in <i>death's-head</i> \ˈdeθs-hed\</p> <p>t tie, attack</p> <p>th with nothing between, as in <i>thin</i>, <i>ether</i> (actually, this is a single sound, not two); with a hyphen between, two sounds as in <i>knighthood</i> \ˈnɪt-hud\</p> <p><u>th</u> then, either, this (actually, this is a single sound, not two)</p> <p>ü rule, youth, union \ˈyūn-yən/, few \ˈfyū\</p> <p>û pull, wood, book, curable \ˈkyūr-ə-bəl\</p> <p>uc German <i>füllen</i>, <i>hübsch</i></p> <p>ǖ French <i>rue</i>, German <i>fühlen</i></p> <p>v vivid, give</p> <p>w we, away; in some words having final \(),̄\ a variant \ə-w\ occurs before vowels, as in \ˈfɪl-ə-wɪŋ/, covered by the variant \ə(-w)\ at the entry word</p> <p>y yard, young, cue \ˈkyū/, union \ˈyūn-yən/</p> <p>y indicates that during the articulation of the sound represented by the preceding character the front of the tongue has substantially the position it has for the articulation of the first sound of <i>yard</i>, as in French <i>digne</i> \dēnʲ\</p> <p>yü youth, union, cue, few, mute</p> <p>yû curable, fury</p> <p>z zone, raise</p> <p>zh with nothing between, as in <i>vision</i>, <i>azure</i> \ˈazh-ər/ (actually, this is a single sound, not two); with a hyphen between, two sounds as in <i>gazehound</i> \ˈgāz-haund\</p> |
|---|---|