

**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 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.

**A. Administrative**

|  |   |
|--|---|
| <b>1. Title:</b>   | <b>Proposal to encode combining double breve below</b>                              |
| 2. Requester's name:   | Rick McGowan, Peter Constable, Ken Whistler, and <i>Unicode Technical Committee</i> |
| 3. Requester type (Member body/Liaison/Individual contribution): | Individual contribution   |
| 4. Submission date:  | July 1, 2003  |
| 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 Diacritical Marks (0300-03FF) |
| 2. Number of characters in proposal:  | 1                                       |
| 3. Proposed category (see section II, Character Categories):  | A                                       |
| 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?   | YES                                     |
| If Yes, reference:  | <i>contains combining marks</i>         |
| 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?  | Unicode, Inc.                           |
| 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)?   | NO                                      |
| 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 <a href="http://www.unicode.org">http://www.unicode.org</a> for such information on other scripts. Also see <a href="http://www.unicode.org/Public/UNIDATA/UnicodeCharacterDatabase.html">http://www.unicode.org/Public/UNIDATA/UnicodeCharacterDatabase.html</a> 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: N2352-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09)

### C. Technical - Justification

|   |  |
|---|--|
| 1. Has this proposal for addition of character(s) been submitted before?<br>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.)?<br>If YES, with whom? _____<br>If YES, available relevant documents: _____  | NO   |
| 3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?<br>Reference: _____  | Users of IPA   |
| 4. The context of use for the proposed characters (type of use; common or rare)<br>Reference: _____   | _____  |
| 5. Are the proposed characters in current use by the user community?<br>If YES, where? Reference: _____   | YES<br>Users of IPA                                    |
| 6. After giving due considerations to the principles in <i>Principles and Procedures document</i> (a WG 2 standing document) must the proposed characters be entirely in the BMP?<br>If YES, is a rationale provided?<br>If YES, reference: Stay with other IPA characters and combining marks  | YES<br>YES   |
| 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?<br>If YES, is a rationale for its inclusion provided?<br>If YES, reference: _____   | NO<br>_____  |
| 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: _____  | NO<br>_____  |
| 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: _____  | YES (U+032E)<br>YES<br>See the body of proposal, below |
| 11. Does the proposal include use of combining characters and/or use of composite sequences (see clauses 4.12 and 4.14 in ISO/IEC 10646-1: 2000)?<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: _____ | YES<br>YES<br>Is a combining mark.<br>NO               |
| 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) _____   | NO<br>_____  |
| 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: _____   | NO<br>_____  |

## Proposal to encode combining double breve below

The combining double breve below is used in the handbook defining the International Phonetic Alphabet (IPA). It is also cited in Pullum and Ladusaw's *Phonetic Symbol Guide* published by the University of Chicago, 1996. The character appears on page 266, where they call it "Bottom Ligature".

In the Handbook of the IPA (1999), the character is actually misidentified.

Page 173 shows "Top tie bar", IPA #433, which is a diacritic to indicate various kinds of coarticulation (affricate or double articulation). This is a nonspacing double diacritic, = U+0361. Page 174 shows "Bottom tie bar", IPA #509, which is an indication of linking. That is a *spacing* extender, equivalent to U+203F UNDERTIE. (But the example shown on p. 203 is *not* this; instead p. 203 shows the nonspacing double diacritic in question, displaying *\*below\** two characters.) The IPA documentation does not properly distinguish between "tie bars" that behave this way and U+203F UNDERTIE. Unicode has to make this distinction, because nonspacing and spacing things are quite different in implementation and properties.

In concept the nonspacing tie bar *below* is simply an allograph of the nonspacing tie bar *above*. It is like the cedilla below the g (Figure 7-1, p. 162, The Unicode Standard 3.0) that jumps to an inverted comma *above* the g, because of the presence of a descender. However, because of properties and because of the nature of our encoding of nonspacing combining marks in particular, it makes sense to encode a separate *character* for this allograph.

**Similar Characters:** The proposed character is similar to U+032E COMBINING BREVE BELOW, but that mark is a single diacritic. Because of this difference in combining behavior, applying to two base characters, it must be encoded separately, as other double diacritics. Also similar is U+2323 SMILE. That character, however, is *not* a combining character at all, but rather an ordinary spacing character.

The proposed character name and representative glyph are as follows:



Example of usage:

kp

It is proposed to encode this character at U+035C.

Character properties are similar to other such double diacritical marks:

General Category is Mn  
Combining Type is 233.  
Bidi category is NSM.

### Contextual Examples from Published Sources

1. Catford, J.C. 1988. *A practical introduction to phonetics*. Oxford: Clarendon Press. (p. 104)

Now put a vowel before and after [pk] and say [apka], making sure that the two articulations are as nearly as possible coterminous, i.e. begin and end at the same time. Experiment with other co-articulated stops, [bg], [tk], [dg], [pt], [bd].

2. Lapidus, B.A., and S.V. Shevtsova. 1962. *A Russian-English dictionary for the foreign student of Russian*. Moscow: State Publishing House of Foreign and National Dictionaries. (p. 228)

епоупеи.  
наслѣдник [плас<sup>b</sup>l<sup>b</sup>éd<sup>b</sup>п<sup>b</sup>ik,  
-g] *n* -а, -ов heir (also fig).  
наслѣдств|о [плас<sup>b</sup>l<sup>b</sup>étstvə]  
*n* -а inheritance, heritage,  
legacy (also fig).

3. Odden, David. 1995. *Tone: African languages*. In *The handbook of phonological theory* (Blackwell handbooks in linguistics, 1), ed. by John A. Goldsmith, 444 - 75. Oxford: Blackwell Publishers. (p. 446)

As a second illustration of the difference between simple 2 tone and the 2-tone component of a 21 contour, a 2 tone downsteps to 3 after a 4 tone, so  $\text{ɔ}^2 \text{yi}^{21} \text{kla}^4 \text{fo}^2$  becomes  $\text{ɔ}^2 \text{yi}^{21} \text{kla}^4 \text{fo}^3$  "She didn't wait for Kla", but the 21 contour of  $\text{ɔ}^2 \text{yi}^{21} \text{kla}^4 \text{nye}^{21}$  "she doesn't hate Kla" does not downstep. Examples like these,