Proposal to encode a combining diacritical mark for Low German dialect writing
2008-10-25
Karl Pentzlin, karl-pentzlin@europastatur.de

The following combining diacritical mark is proposed:

U+1DFD   COMBINING TRIPLE BREVE BELOW
   → 035C combining double breve below
   • to be applied to the first letter of any three-letter combination
   • used with "sch", "Sch", and "SCH" in Low German dialect writing

The combining triple breve below is used in the "Rheinische Dokumenta" [2], an orthography promoted by the "Landschaftsverband Rheinland" (a cultural authority of the southern part of the German federal state Nordrhein-Westfalen) during the 1980s for the local Low German dialects.

It is used there to mark the trigraph "sch" (which is used to denote the voiceless ʃ sound, like in standard High German orthography) denoting the voiced ʒ sound. Thus, it may be applied to "sch" in lower case, "Sch" in title case, and "SCH" in upper case.

The proposed properties are:

1DFD;COMBINING TRIPLE BREVE BELOW;Mn;233;NSM;;;;;N;;;;;

The triple breve below is to be input after the first character of the combination it applies to, analogous to the double diacritics like U+035C COMBINING DOUBLE BREVE BELOW.

Thus, the correct way to enter a "sch" trigraph with triple breve below is:
U+0073 U+1Dxx U+0063 U+0068.

Bibliography:


http://www.beepworld.de/memberdateien/members85/friedhelmschmitz/dielnschbank1.pdf

---

Der stimmhafte Reibelaut sch (Garage) wird durch einen untergesetzten Bogen gekennzeichnet. Dieser Laut ist in den Mundarten des südlichen Rheinlands häufig anzutreffen: intswische (inzwischen), gleisch (gleich).

Fig. 1: Sample from [2], p. 16, introducing the proposed diacritical mark.

Neben den im Hochdeutschen üblichen Buchstaben nebst Umlauten a, o und u finden neun Sonderzeichen Verwendung:

\[
\begin{align*}
\text{a} & \quad \text{ä} & \quad \text{ch} & \quad \text{e} & \quad \text{ö} & \quad \text{o} & \quad \text{ö} & \quad \text{r} & \quad \text{s} & \quad \text{sch}
\end{align*}
\]

**Fig. 2:** Sample from [3], explaining the Rheinische Dokumenta and showing the proposed diacritical mark.

Mundart: Mutterschied (Hunsrück)

*Im Schpääüoa, wän dë Ääre in waa un di Šou fät genuuch schiën, is bai dë Doorfmäska en Tärmiin foor dë Schlachtadaach ausgemach woa. Am Šöment foorehe is di Muul med-ën Schupkarë gehuult woa, dë Wäschkäsèl met Wasa gefilt en Haitsmadriaal baraat gelää, das am frië Mëörje angehaisst weerë könt. Wän dë Mätska küm is, hot misë des Wasa koche. Wän-s dan so weit waa, is di Muda qörtë dë Fada in de Schdal gang un hot de Šou en Schtrik aan en Hinabddän gebun un se foa di Deer gefoor, wo se dan koats an-ë Hooge qörtë Ring draan gebun is woa. Intwische hot sisch dë Mätska met dë Aks bëwaven et hot ñömet di Šou uf dë Schdeere geschlaan. Un wän se nit gleisx umgefal is, hör-a noch-emëp röppgeschlaan. Dan hör-a das groos Mäsa genoom un hot di Šou in de Hals eningeschloch un di Schlaachoopra doorischgeschniit.*

**Fig. 3:** Sample from [2], p. 14, showing the proposed diacritical mark.

495 Mandel (Vbg. Rüdesheim)

S Enck, Helmut, Jg. 1928, geb. in Mandel, lebt auch dort, Winzer.
A 1982, 3.54 min.
T Weihnachten, Mandel um 1933, spontan gesprochen.

Doo waa dë alt Qoba noch, dë °Ängge, alëo maim Faderšai Fader, deër hat dë oqë sa Schdup gëhat, un dë hoët doo gëwqont und alëo das Šäkrädarsche schdäään [. . .], un isch waa dëo sälischëmëpp° finef, sëks Joër alt. Ich kënd-ë bisë frïer laaëse wii dë Normaalë, alëo ich kont mit finf Joër läaëse oqëder, dässch-ë°: un läišë, das waa drum, isch hon halt gelääse, soo mach ich-s hoët noch, alëo isch läaëse tsäntnerwais Bischër

**Fig. 4:** Sample from [1], p. 495, showing the proposed diacritical mark.

Et woqr tseen op tseen, Johan woqr alt jans näröös op de Bangk èröm-am-rödsche. Hää kreisch wiër éns dë Ugr èruus. Di woqren am Ärtsälë, et woqr alt tswanaisch op tseen. Op eemqol sät Hain:

**Fig. 5:** Sample from [4], p. 14, showing the proposed diacritical mark.
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 ensure you are using the latest Form from http://www.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html for latest Roadmaps.

A. Administrative

| 1. Title: | Proposal to encode three combining diacritical marks for German dialect writing |
| 2. Requester's name: | Karl Pentzlin |
| 3. Requester type (Member body/Liaison/Individual contribution): | Individual Contribution |
| 4. Submission date: | 2008-10-25 |
| 5. Requester's reference (if applicable): | |
| 6. Choose one of the following: | |

| 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: | 1DC0-1DFF Combining Diacritical Marks Supplement |

| 2. Number of characters in proposal: | 1 |

| 3. Proposed category (select one from below - see section 2.2 of P&P document): | |

| A-Contemporary | 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. 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 |

| 5. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard? | |

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

| 6. 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 |

| 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)? | yes |

| 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 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. | |

---

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? Peter Honnen, Landschaftsverband Rheinland
   If YES, available relevant documents: n/a

3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? yes
   Reference: see proposal text

4. The context of use for the proposed characters (type of use; common or rare) is included? yes
   Reference: see proposal text

5. Are the proposed characters in current use by the user community? yes
   Reference: see proposal text

6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP? yes
   If YES, a rationale provided: keeping them together with all other diacritics for the Latin script

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? no
   If YES, a rationale for its inclusion provided?
   If YES, reference:

9. Can any of the proposed characters be encoded using a composed character sequence of other existing characters or other proposed characters? no
   If YES, 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? no
    If YES, a rationale for its inclusion provided?
    If YES, reference:

11. Does the proposal include use of combining characters and/or use of composite sequences? yes
    If YES, a rationale for such use provided?
    If YES, reference: see proposal text
    Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided? yes
    If YES, reference: see proposal text

12. Does the proposal contain characters with any special properties such as control function or similar semantics? no
    If YES, describe in detail (include attachment if necessary)

13. Does the proposal contain any Ideographic compatibility character(s)? yes
    If YES, is the equivalent corresponding unified ideographic character(s) identified?
    If YES, reference: