

## Proposal to encode a combining diacritical mark for Low German dialect writing

2008-10-25

Karl Pentzlin, [karl-pentzlin@europatastatur.de](mailto:karl-pentzlin@europatastatur.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:

- [1] CORNELISSEN, Georg, HONNEN, Peter, LANGENSIEPEN, Fritz (ed.):  
Das rheinische Platt · Eine Bestandsaufnahme  
(Handbuch der rheinischen Mundarten, Teil 1: Texte)  
Köln 1989, ISBN 3-7927-0689-X
- [2] HONNEN, Peter:  
Rheinische Dokumenta. Lautschrift für rheinische Mundarten  
Pulheim 1987, ISBN 3-7927-0947-3
- [3] [http://de.wikipedia.org/wiki/Rheinische\\_Dokumenta](http://de.wikipedia.org/wiki/Rheinische_Dokumenta) as of 2008-10-25 (*German*)
- [4] SCHMITZ, Friedhelm, HINTZEN, Josef:  
De Lüschtbank. Plattdeutsche Anekdoten aus Jüchen und Umgebung  
Jüchen 1985, private edition, retrieved 2008-10-25 at:  
<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.

Die **Rheinische Dokumenta** ist die Anfang der 1980er Jahre entwickelte und 1986 vom **Landschaftsverband Rheinland** vorgestellte Lautschrift für die **rheinischen Mundarten** des nördlichen und südlichen **Niederrheins**, des westlichen und zentralen Rheinlands, des **Bergischen Landes**, des **Westerwalds**, der westlichen und östlichen **Eifel**, der Unter- und der **Obermosel** und der Region um **Hunsrück** und **Nahe**.

Neben den im Hochdeutschen üblichen Buchstaben nebst Umlauten ä, ö und ü finden neun Sonderzeichen Verwendung:

ḅ ḁ ch ɛ ɔ ö ɾ s sch

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

#### Mundart: Mutterschied (Hunsrück)

*Im Schpääjtjoa, wän de Ääre in waa un di Sou füt genuuch schiin, is bai de Doorfmätska en Tärmiin foor de Schlachtdaach ousgemach woa. Am Qoment foorhea is di Muul med-em Schupkarē gehuult woa, de Wäschkäşel met Waşa gefilt un Haitmadriaal baraat geläät, das am friie Mōrjē angehaitst weere kōnt. Wän de Mätska kōm is, hot mişē deş Waşa kōchē. Wän-s dan şo weit waa, is di Muda oore de Fada in de Schdal gang un hot de Sou en Schtrik aan en Hinabään gebun un şē foa di Deer gefoort, wo şē dan koats an-ē Hooge oore Ring draan gebun is woa. Intwişchē hot şisch de Mätska met de Aks bewavent un hot doomet di Sou uf de Schdeere geschlaan. Un wän şē nit gleisch umgefal is, hōr-a noch-emōō noogeschlaan. Dan hōr-a das groos Mäşa genom un hot di Sou in de Hals eningeschdōch un di Schlaachōōra doorischgeschniit.\**

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

#### 495 Mandel (Vbg. Rüdeshheim)

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 de alt Qōba noch, de °Ängge, also maim Fader şai Fader, deer hat doo ove şai Schdup gehat, un de hot doo gewoont und also das Şäkrädärschē schdään [. . .], un isch waa doo şälischēmōō<sup>1</sup> finēf, şäks Joer alt. Ich kōnd-ē bisjē friier lääşē wii dii Nōrmaale, also ich kōnt mit finf Joer lääşē oōder, dāsch-ē<sup>2</sup>: un läişē, das waa drum, isch hōn halt gelääşē, şoo mach ich-s hōit noch, also isch lääşē tsäntnerwais Bischer

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

Ët woqr tseen op tseen, Johan woqr alt jans nārvōōs op de Bangk erōm-am-rōdschē. Hā kreesch wier ens de Uer eruus. Di woqr en am Ärtsälē, Ët woqr alt tswansisch op tseen. Op eemōōl şäät 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<sup>1</sup>**

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: 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:

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: 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) <input checked="" type="checkbox"/>	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? - tbd -

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? \_\_\_\_\_

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.

<sup>1</sup> Form number: N3102-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)

**C. Technical - Justification**

1. Has this proposal for addition of character(s) been submitted before? If YES explain		<i>no</i>
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>Peter Honnen, Landschaftsverband Rheinland</i> <i>n/a</i>	<i>yes</i>
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 proposal text</i>	<i>yes</i>
4. The context of use for the proposed characters (type of use; common or rare) Reference:	<i>see proposal text</i>	
5. Are the proposed characters in current use by the user community? If YES, where? Reference:	<i>see proposal text</i>	<i>yes</i>
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>keeping them together with all other diacritics for the Latin script</i>	<i>yes</i>
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?		<i>n/a</i>
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:		<i>no</i>
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:		<i>no</i>
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:		<i>no</i>
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:	<i>see proposal text</i> <i>see proposal text</i>	<i>yes</i> <i>yes</i> <i>yes</i>
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)		<i>no</i>
13. Does the proposal contain any Ideographic compatibility character(s)? If YES, is the equivalent corresponding unified ideographic character(s) identified If YES, reference:		<i>no</i>