

DIN EN 13710

ICS 01.140.20; 35.040

Ersatz für
DIN V ENV 13710:2007-09**Europäische Sortierregeln –
Sortierung von lateinischen, griechischen, kyrillischen, georgischen und
armenischen Schriftzeichen;
Deutsche Fassung EN 13710:2011**

European Ordering Rules –
Ordering of characters from Latin, Greek, Cyrillic, Georgian and Armenian scripts;
German version EN 13710:2011

Règles de classement européen –
Classement des caractères latins, grecs, cyrilliques, géorgiens et arméniens;
Version allemande EN 13710:2011

Gesamtumfang 56 Seiten

Normenausschuss Informationstechnik und Anwendungen (NIA) im DIN

Nationales Vorwort

Dieses Dokument (EN 13710:2011) wurde im Technischen Komitee CEN/TC 304 „Informations- und Kommunikationstechnik — Europäische Lokalisierungsanforderungen“ erarbeitet, dessen Sekretariat vom DIN (Deutschland) gehalten wird.

Das zuständige deutsche Gremium ist der Arbeitskreis NA 043-01-29-01 AK „Codierte Zeichensätze“ im Normenausschuss „Informationstechnik und Anwendungen (NIA)“.

Änderungen

Gegenüber DIN V ENV 13710:2007-09 wurden folgende Änderungen vorgenommen:

- a) ENV 13170:2000-12 *Europäische Sortierregeln — Sortierung von lateinischen, griechischen und kyrillischen Schriftzeichen*, und CR 14400:2001-12 *Europäische Sortierregeln — Sortierung von lateinischen, griechischen, kyrillischen, georgischen und armenischen Schriftzeichen* zusammengefasst;
- b) Teile des Dokuments inhaltlich überarbeitet und an den Stand der Technik angepasst.

Frühere Ausgaben

DIN V ENV 13710: 2007-09

Deutsche Fassung

Europäische Sortierregeln - Sortierung von lateinischen, griechischen, kyrillischen, georgischen und armenischen Schriftzeichen

European Ordering Rules - Ordering of characters from Latin, Greek, Cyrillic, Georgian and Armenian scripts

Règles de classement européen - Classement des caractères latins, grecs, cyrilliques, géorgiens et arméniens

Diese Europäische Norm wurde vom CEN am 5. Februar 2011 angenommen.

Die CEN-Mitglieder sind gehalten, die CEN/CENELEC-Geschäftsordnung zu erfüllen, in der die Bedingungen festgelegt sind, unter denen dieser Europäischen Norm ohne jede Änderung der Status einer nationalen Norm zu geben ist. Auf dem letzten Stand befindliche Listen dieser nationalen Normen mit ihren bibliographischen Angaben sind beim Management-Zentrum des CEN-CENELEC oder bei jedem CEN-Mitglied auf Anfrage erhältlich.

Diese Europäische Norm besteht in drei offiziellen Fassungen (Deutsch, Englisch, Französisch). Eine Fassung in einer anderen Sprache, die von einem CEN-Mitglied in eigener Verantwortung durch Übersetzung in seine Landessprache gemacht und dem Management-Zentrum mitgeteilt worden ist, hat den gleichen Status wie die offiziellen Fassungen.

CEN-Mitglieder sind die nationalen Normungsinstitute von Belgien, Bulgarien, Dänemark, Deutschland, Estland, Finnland, Frankreich, Griechenland, Irland, Island, Italien, Kroatien, Lettland, Litauen, Luxemburg, Malta, den Niederlanden, Norwegen, Österreich, Polen, Portugal, Rumänien, Schweden, der Schweiz, der Slowakei, Slowenien, Spanien, der Tschechischen Republik, Ungarn, dem Vereinigten Königreich und Zypern.



EUROPÄISCHES KOMITEE FÜR NORMUNG
EUROPEAN COMMITTEE FOR STANDARDIZATION
COMITÉ EUROPÉEN DE NORMALISATION

Management-Zentrum: Avenue Marnix 17, B-1000 Brüssel

Inhalt

	Seite
Vorwort	3
Einleitung.....	4
1 Anwendungsbereich	5
2 Normative Verweisungen	5
3 Begriffe	5
4 Normkonformität.....	6
5 Anpassbarkeit	6
6 EOR Delta-Tabelle.....	7
Anhang A (informativ) Prinzipien hinter den Europäischen Sortierregeln.....	14
A.0 Einleitung.....	14
A.1 Begriffe	14
A.2 Vorbereitende Schritte	19
A.3 Das mehrstufige Sortierverfahren.....	20
A.4 Erste Sortierebene	21
A.5 Zweite Sortierebene.....	22
A.6 Dritte Sortierebene.....	23
A.7 Vierte Sortierebene.....	23
A.8 Spezielle Reihenfolgen für die alphabetische Anordnung.....	24
Anhang B (informativ) Anordnung unter Berücksichtigung der Wortgrenzen	37
B.1 Veränderte Begrifflichkeit	37
B.2 Kurzbeschreibung	37
B.3 Beispiel für die Anordnung unter Berücksichtigung von Wortgrenzen im Gegensatz zur Anordnung ohne deren Berücksichtigung	37
B.4 Vereinfachte Anordnung unter Berücksichtigung von Wortgrenzen.....	37
Anhang C (informativ) Sortierung nach Stellung und Schriftschnitt.....	38
C.1 Hintergrund	38
C.2 Empfohlene Regeln	38
Anhang D (informativ) Mehrschrift-Sortierung bei einer vorherrschenden Schrift.....	39
D.1 Hintergrund	39
D.2 Vorgeschlagene Schritte.....	39
D.3 Explizite Transliteration	39
Anhang E (informativ) Festlegen des auf EOR basierenden nationalen Delta	40
E.1 Hintergrund	40
E.2 Strukturierte Spezifikation	40
E.3 Maschinell lesbare Spezifikation.....	40
E.4 Beispiel 1: Nationales Delta für Deutsch.....	41
E.5 Beispiel 2: Nationales Delta für Norwegisch	43
Anhang F (informativ) Moderne Europäische Schriften/MES.....	45
Anhang G (informativ) EOR Delta in LDML Syntax.....	47
Literaturhinweise	54

Vorwort

Dieses Dokument (EN 13710:2011) wurde vom Technischen Komitee CEN/TC 304 „Informations- und Kommunikationstechnik — Europäische Lokalisierungsanforderungen“ erarbeitet, dessen Sekretariat vom DIN gehalten wird.

Diese Europäische Norm muss den Status einer nationalen Norm erhalten, entweder durch Veröffentlichung eines identischen Textes oder durch Anerkennung bis September 2011, und etwaige entgegenstehende nationale Normen müssen bis September 2011 zurückgezogen werden.

Es wird auf die Möglichkeit hingewiesen, dass einige Texte dieses Dokuments Patentrechte berühren können. CEN [und/oder CENELEC] sind nicht dafür verantwortlich, einige oder alle diesbezüglichen Patentrechte zu identifizieren.

Dieses Dokument ersetzt ENV 13710:2000 und CR 14400:2001.

Gegenüber ENV 13710:2000 wurden in der vorliegenden Ausgabe folgende Änderungen vorgenommen:

- a) ENV 13170:2000 und CR 14400:2001 zusammengefasst;
- b) Teile des Dokuments inhaltlich überarbeitet und an den Stand der Technik angepasst.

Entsprechend der CEN/CENELEC-Geschäftsordnung sind die nationalen Normungsinstitute der folgenden Länder gehalten, diese Europäische Norm zu übernehmen: Belgien, Bulgarien, Dänemark, Deutschland, Estland, Finnland, Frankreich, Griechenland, Irland, Island, Italien, Kroatien, Lettland, Litauen, Luxemburg, Malta, Niederlande, Norwegen, Österreich, Polen, Portugal, Rumänien, Schweden, Schweiz, Slowakei, Slowenien, Spanien, Tschechische Republik, Ungarn, Vereinigtes Königreich und Zypern.

Einleitung

Diese Europäische Norm enthält Regeln für die Sortierung multilingualer Auflistungen in eine eindeutig festgelegte, intuitive Abfolge. Diese Regeln sind für Daten aus verschiedenen europäischen Sprachen vorgesehen, die in eine eindeutige Sortierreihenfolge gebracht werden müssen, die es Anwendern mit unterschiedlichem kulturellem Hintergrund ermöglicht, Informationen leicht zu finden. Zugleich bildet die Norm die Grundlage für die Festlegung sprachenspezifischer Profile, wobei die Regeln einer gegebenen Sprachgemeinschaft ebenso wie der gesamte paneuropäische Zeichensatz auf eine europaweit einheitliche Art und Weise berücksichtigt werden.

Diese Regeln wurden geprüft und in zwei vorherigen Spezifikationen weitgehend angenommen, in ENV 13170:2000 *Europäische Sortierregeln — Sortierung von lateinischen, griechischen und kyrillischen Zeichen*, und ihrer begleitenden und ergänzenden Spezifikation CR 14400:2001 *Europäische Sortierregeln — Sortierung von lateinischen, griechischen, kyrillischen, georgischen und armenischen Zeichen*. Diese Europäische Norm führt diese beiden Dokumente in eine technisch in hohem Maße aufwärtskompatible Norm zusammen.

Diese Europäische Norm gilt für zwei verschiedene Zielgruppen, für Softwareentwickler auf der einen Seite, und für Anwender von Sortieranwendungen auf der anderen Seite.

Softwareberater benötigen eindeutige maschinell verarbeitbare Vorlagen, die in vorhandenen und zukünftigen Anwendungen verwendet werden können. Das kann am besten geschehen, indem eine europäische Standard-Sortiertabelle in der Syntax von zwei international wichtigen Spezifikationen festgelegt wird:

- in der internationalen Sortiernorm ISO/IEC 14651:2007 *International string ordering and comparison — Method for comparing character strings and description of the common template tailorable ordering*, wovon der gegenwärtige Standard ein „Delta“ ist (siehe Abschnitt 6);
- in der *Locale Data Markup Language (LDML)*, die (unter anderem) verwendet wird, um die anpassbaren Daten für *Unicode Common Locale Data Repository (CLDR)* zu erfassen. CLDR beruht auf *Unicode Collation Algorithm (UCA)*, die technisch gesehen ein Profil von ISO/IEC 14651:2007 ist (siehe informativer Anhang G).

Anwender ohne speziellen Hintergrund in der Informationstechnik benötigen hingegen eine Erklärung der Prinzipien, die sich mehr an den vorhandenen nationalen Sortiernormen oder der einschlägigen Praxis orientiert. Da anpassbare Sortiertabellen für Nichtfachleute schwer zu lesen sind, wird in den informativen Anhängen eine Erklärung derjenigen Regeln gegeben, die bei der Gestaltung der anpassbaren Sortiertabelle verwendet werden. Anwendern, die die Syntax der anpassbaren Sortiertabelle nicht beherrschen, wird geraten, zunächst diese Anhänge zu lesen.

Der Hauptteil dieser Norm legt für Zeichenfolgen eine alphabetische Anordnung fest. Der informative Anhang A präsentiert ergebnisgleiche Regeln in einer für Nichtfachleute leichter zugänglichen Form. Der informative Anhang B behandelt alphabetische Anordnungen Wort nach Wort als eine spezielle Form der Sortierung mit mehreren Sortierschlüsseln. Der informative Anhang C erklärt weitere Kriterien alphabetischer Anordnung. Der informative Anhang D erläutert eine populäre Alternative zum Hauptteil, nämlich die Vermischung mehrerer Schriften durch implizierte Transliteration. Der informative Anhang E gibt eine Anleitung zur Anwendung dieser Europäischen Norm als Grundlage für die Darstellung nationaler Deltas. Der informative Anhang F listet den zugrunde liegenden Zeichenvorrat auf, um die Konsultation zu erleichtern. Der informative Anhang G stellt das formale Delta in der LDML-Syntax dar.

In Übereinstimmung mit der Praxis von ISO/IEC 14651 werden die Zeichen als UXXXX benannt, wobei X für eine beliebige Hexadezimalzahl steht und der Ausdruck sich auf die Codierung dieses Zeichens in der ISO/IEC 10646 bezieht. Diese Übereinkunft wird überall in dieser Europäischen Norm verwendet.

1 Anwendungsbereich

Diese europäische Norm legt die Reihenfolge zwischen zwei Zeichenfolgen fest, die aus Zeichen des *Modern European Scripts (MES)*-Zeichenvorrates aus ISO/IEC 10646:2003 oder Untermengen davon entstammen.

ANMERKUNG Der Zeichenvorrat 283 *Modern European Scripts (MES)* aus ISO/IEC 10646:2003 wurde ursprünglich im CEN Workshop Agreement 13873:2000 als *Multilingual European Subsets* aus ISO/IEC 10646 als *Multilingual European Subset Nummer 3* festgelegt und wurde danach als Zeichenvorrat in den Anhang A von ISO/IEC 10646:2003 eingeführt, neben den verwandten Zeichenvorräten MES-1 und MES-2.

Die in dieser Europäischen Norm angegebenen Regeln zur alphabetischen Anordnung sind nur für Datenaufstellungen in mehr als einer europäischen Sprache vorgesehen und nur dann, wenn diese Daten für ein multikulturelles Publikum vorgesehen sind. Sie ergänzen bestehende nationale Normen oder Verfahren in der Praxis.

2 Normative Verweisungen

Die folgenden zitierten Dokumente sind für die Anwendung dieses Dokuments erforderlich. Bei datierten Verweisungen gilt nur die in Bezug genommene Ausgabe. Bei undatierten Verweisungen gilt die letzte Ausgabe des in Bezug genommenen Dokuments (einschließlich aller Änderungen).

ISO/IEC 10646:2003, *Information technology — Universal Multi-Octet Coded Character Set (UCS)*

ISO 12199:2000, *Alphabetical ordering of multilingual terminological and lexicographical data represented in the Latin alphabet*

ISO/IEC 14651:2007, *Information technology — International string ordering and comparison — Method for comparing character strings and description of the common template tailorable ordering*

3 Begriffe

Für die Anwendung dieses Dokuments gelten die Begriffe nach ISO/IEC 10646:2003 und ISO/IEC 14651:2007.

3.1

Zeichen

Element aus einer festgelegten Menge von Elementen, das für die Organisation, die Steuerung oder die Darstellung von Daten verwendet wird

[ISO/IEC 10646:2003]

3.2

Zeichenfolge

eine Folge von Zeichen, die als einzelne Einheit betrachtet wird

[ISO/IEC 14651:2007]

3.3

Sortiersymbol

Symbol, das dazu verwendet wird, den Stellenwert eines Sortierelements festzulegen

[ISO/IEC 14651:2007]

3.4

Sortierelement

Folge von einem oder mehreren Zeichen, die als einzelne Einheit zum Sortieren betrachtet wird

[ISO/IEC 14651:2007]

3.5

Sortiertabelle

Zuordnung von Sortierelementen zu wertenden Elementen

[ISO/IEC 14651:2007]

ANMERKUNG Eine besondere Sortiertabelle ist die in Anhang A von ISO/IEC 14651 verwendete anpassbare Sortiertabelle (Common Template Table CTT), um die Zuordnung von Sortierelementen zu wertenden Elementen darzustellen.

3.6

Delta

Aufstellung der Abweichungen zwischen einer gegebenen Sortiertabelle und einer anderen

[ISO/IEC 14651:2007]

ANMERKUNG Die gegebene Sortiertabelle bildet, zusammen mit einem gegebenen Delta, eine neue Sortiertabelle. Wenn in dieser Europäischen Norm nicht anders festgelegt, bezieht sich der Begriff „Delta“ immer auf die Abweichungen von der allgemeinen Sortiertabelle (Common Template Table), wie sie in ISO/IEC 14651 festgelegt ist.

3.7

alphabetische Anordnung

Vorgang, bei dem für zwei Zeichenfolgen festgelegt ist, ob die erste kleiner als, gleich oder größer als die zweite Zeichenfolge ist

[ISO/IEC 14651:2007]

3.8

Sortieren

Darstellung von Informationen auf strukturierte Weise

ANMERKUNG Sortieren kann die Unterteilung von Informationen nach Themengebieten beinhalten, z. B. mehrere Register in einem Buch, die Unterteilung des Telefonbuchs nach Städten oder verschiedene Kataloge in einer Bibliothek. Die alphabetische Anordnung ist im Regelfall ein integraler Bestandteil dieses Verfahrens.

4 Normkonformität

Um konform mit dieser Europäischen Norm zu sein, muss eine Anwendung die Anforderungen nach ISO/IEC 14651:2007, Abschnitt 6, erfüllen, und die anpassbare Sortiertabelle ISO14651_2006_TABLE1 verwenden, die nach der Anwendung der EOR Delta Tabelle in Abschnitt 6 dieser Europäischen Norm festgelegt ist. Eine gleichwertige Beschreibung der resultierenden angepassten Sortiertabelle muss ebenso konform mit dieser Europäischen Norm sein.

5 Anpassbarkeit

Diese Europäische Norm legt Europäische Sortierregeln (EOR) fest, die nach ISO/IEC 14651 (vgl. ebenso den informativen Anhang E) an die Bedürfnisse jedes europäischen Landes angepasst werden kann.

Diese Europäische Norm sieht nicht vor, nationale Normen oder Traditionen im Anwendungsfeld der alphabetischen Anordnung zu beeinflussen, da ihr Anwendungsbereich die alphabetische Anordnung multilingualer Daten ist. Dennoch werden nationale Normen dazu angeregt, ihre nationalen Sortierregeln in dieser Europäischen Norm darzustellen, indem sie mittels einer formalisierten Syntax (Delta) an deren Bedürfnisse angepasst werden, wie im informativen Anhang E dargestellt ist. Auf diese Weise werden die entsprechenden Sortierregeln automatisch maschinell verarbeitbar und können in den internationalen Speicher der Sprach- und Gebietseinstellungsdaten eingegliedert werden, was eine umfassendere Unterstützung nationaler Sortier-normen über Softwareprodukte ermöglicht.

6 EOR Delta-Tabelle

ANMERKUNG Die Syntax der Sortiertabelle ist in ISO/IEC 14651 festgelegt.

```
%% EOR's EORDeltaTable
%
%% European Ordering Rules.
%
% EOR delta for MES-3 from ISO/IEC 14651:2007's CTT (ISO14651_2006_TABLE1_en.txt).
%
% EOR uses four levels for comparison

reorder-after <BASE> % Introduce the LIG weight.
collating-symbol <LIG>
<LIG>
reorder-end

reorder-after <VRNT3> %Introduce more variants
collating-symbol <VRNT4>
collating-symbol <VRNT5>
collating-symbol <VRNT6>
collating-symbol <VRNT7>
collating-symbol <VRNT8>
collating-symbol <VRNT9>
<VRNT4>
<VRNT5>
<VRNT6>
<VRNT7>
<VRNT8>
<VRNT9>
reorder-end

reorder-after <S0584> %Introduce a weight for U0587 ARMENIAN SMALL LIGATURE ECH YIWN
collating-symbol <ECH-YIWN>
<ECH-YIWN>
reorder-end

reorder-after <SFFFF>

order_start forward:forward:forward:forward

% Non-alphanumeric characters (including some modifier letters):
% The DRACHMA SIGN is already in ISO14651_2006 ignorable on levels 1-3
<U0024> IGNORE:IGNORE:IGNORE:<U0024> % DOLLAR SIGN
<U00A2> IGNORE:IGNORE:IGNORE:<U00A2> % CENT SIGN
<U00A3> IGNORE:IGNORE:IGNORE:<U00A3> % POUND SIGN
<U00A4> IGNORE:IGNORE:IGNORE:<U00A4> % CURRENCY SIGN
<U00A5> IGNORE:IGNORE:IGNORE:<U00A5> % YEN SIGN
<U20A0> IGNORE:IGNORE:IGNORE:<U20A0> % EURO-CURRENCY SIGN
<U20A1> IGNORE:IGNORE:IGNORE:<U20A1> % COLON SIGN
<U20A2> IGNORE:IGNORE:IGNORE:<U20A2> % CRUZEIRO SIGN
<U20A3> IGNORE:IGNORE:IGNORE:<U20A3> % FRENCH FRANC SIGN
```

<U20A4> IGNORE:IGNORE:IGNORE:<U20A4> % LIRA SIGN
<U20A5> IGNORE:IGNORE:IGNORE:<U20A5> % MILL SIGN
<U20A6> IGNORE:IGNORE:IGNORE:<U20A6> % NAIRA SIGN
<U20A7> IGNORE:IGNORE:IGNORE:<U20A7> % PESETA SIGN
<U20A8> IGNORE:IGNORE:IGNORE:<U20A8> % RUPEE SIGN
<U20A9> IGNORE:IGNORE:IGNORE:<U20A9> % WON SIGN
<U20AA> IGNORE:IGNORE:IGNORE:<U20AA> % NEW SHEQEL SIGN
<U20AB> IGNORE:IGNORE:IGNORE:<U20AB> % DONG SIGN
<U20AC> IGNORE:IGNORE:IGNORE:<U20AC> % EURO SIGN
<U20AD> IGNORE:IGNORE:IGNORE:<U20AD> % KIP SIGN
<U20AE> IGNORE:IGNORE:IGNORE:<U20AE> % TUGRIK SIGN
<U20AF> IGNORE:IGNORE:IGNORE:<U20AF> % DRACHMA SIGN
<U20B0> IGNORE:IGNORE:IGNORE:<U20B0> % GERMAN PENNY SIGN
<U20B1> IGNORE:IGNORE:IGNORE:<U20B1> % PESO SIGN
<U20B2> IGNORE:IGNORE:IGNORE:<U20B2> % GUARANI SIGN
<U20B3> IGNORE:IGNORE:IGNORE:<U20B3> % AUSTRAL SIGN
<U20B4> IGNORE:IGNORE:IGNORE:<U20B4> % HRYVNIA SIGN
<U20B5> IGNORE:IGNORE:IGNORE:<U20B5> % CEDI SIGN

% Modifier letters that are not ignorable in ISO14651_2006_TABLE1_en.txt
<U02B0> IGNORE:IGNORE:IGNORE:<U02B0> % MODIFIER LETTER SMALL H
<U02B1> IGNORE:IGNORE:IGNORE:<U02B1> % MODIFIER LETTER SMALL H WITH HOOK
<U02B2> IGNORE:IGNORE:IGNORE:<U02B2> % MODIFIER LETTER SMALL J
<U02B3> IGNORE:IGNORE:IGNORE:<U02B3> % MODIFIER LETTER SMALL R
<U02B4> IGNORE:IGNORE:IGNORE:<U02B4> % MODIFIER LETTER SMALL TURNED R
<U02B5> IGNORE:IGNORE:IGNORE:<U02B5> % MODIFIER LETTER SMALL TURNED R WITH HOOK
<U02B6> IGNORE:IGNORE:IGNORE:<U02B6> % MODIFIER LETTER SMALL CAPITAL INVERTED R
<U02B7> IGNORE:IGNORE:IGNORE:<U02B7> % MODIFIER LETTER SMALL W
<U02B8> IGNORE:IGNORE:IGNORE:<U02B8> % MODIFIER LETTER SMALL Y
<U02BB> IGNORE:IGNORE:IGNORE:<U02BB> % MODIFIER LETTER TURNED COMMA
<U02BC> IGNORE:IGNORE:IGNORE:<U02BC> % MODIFIER LETTER APOSTROPHE
<U02BD> IGNORE:IGNORE:IGNORE:<U02BD> % MODIFIER LETTER REVERSED COMMA
<U02BE> IGNORE:IGNORE:IGNORE:<U02BE> % MODIFIER LETTER RIGHT HALF RING
<U02BF> IGNORE:IGNORE:IGNORE:<U02BF> % MODIFIER LETTER LEFT HALF RING
<U02C0> IGNORE:IGNORE:IGNORE:<U02C0> % MODIFIER LETTER GLOTTAL STOP
<U02C1> IGNORE:IGNORE:IGNORE:<U02C1> % MODIFIER LETTER REVERSED GLOTTAL STOP
<U02D0> IGNORE:IGNORE:IGNORE:<U02D0> % MODIFIER LETTER TRIANGULAR COLON
<U02D1> IGNORE:IGNORE:IGNORE:<U02D1> % MODIFIER LETTER HALF TRIANGULAR COLON
<U02E0> IGNORE:IGNORE:IGNORE:<U02E0> % MODIFIER LETTER SMALL GAMMA
<U02E1> IGNORE:IGNORE:IGNORE:<U02E1> % MODIFIER LETTER SMALL L
<U02E2> IGNORE:IGNORE:IGNORE:<U02E2> % MODIFIER LETTER SMALL S
<U02E4> IGNORE:IGNORE:IGNORE:<U02E4> % MODIFIER LETTER SMALL REVERSED GLOTTAL STOP
<U02EE> IGNORE:IGNORE:IGNORE:<U02EE> % MODIFIER LETTER DOUBLE APOSTROPHE

<U0294> IGNORE:IGNORE:IGNORE:<U0294> % LATIN LETTER GLOTTAL STOP
<U0295> IGNORE:IGNORE:IGNORE:<U0295> % LATIN LETTER PHARYNGEAL VOICED FRICATIVE
<U0296> IGNORE:IGNORE:IGNORE:<U0296> % LATIN LETTER INVERTED GLOTTAL STOP
<U0298> IGNORE:IGNORE:IGNORE:<U0298> % LATIN LETTER BILABIAL CLICK
<U02A1> IGNORE:IGNORE:IGNORE:<U02A1> % LATIN LETTER GLOTTAL STOP WITH STROKE
<U02A2> IGNORE:IGNORE:IGNORE:<U02A2> % LATIN LETTER REVERSED GLOTTAL STOP WITH STROKE

%%

% Latin

% Almost all changes here result from CEN/TC304's resolution

% for the Latin script part of the Modern European Scripts / MES-3 to

% treat only the letters a to z and thorn as distinct on the first

% level and treat other combinations as variants or ligatures

<U0250> <S0061>; "<BASE><VRNT1>"; "<MIN><MIN>"; <U0250> % LATIN SMALL LETTER TURNED A

<U0251> <S0061>; "<BASE><VRNT2>"; "<MIN><MIN>"; <U0251> % LATIN SMALL LETTER ALPHA

<U0252> <S0061>; "<BASE><VRNT3>"; "<MIN><MIN>"; <U0252> % LATIN SMALL LETTER TURNED ALPHA

<U0299> <S0062>; "<BASE><VRNT1>"; "<MIN><MIN>"; <U0299> % LATIN LETTER SMALL CAPITAL B

<U0180> <S0062>; "<BASE><VRNT2>"; "<MIN><MIN>"; <U0180> % LATIN SMALL LETTER B WITH STROKE

<U0243> <S0062>; "<BASE><VRNT2>"; "<CAP><MIN>"; <U0243> % LATIN CAPITAL LETTER B WITH STROKE

<U0253> <S0062>; "<BASE><VRNT3>"; "<MIN><MIN>"; <U0253> % LATIN SMALL LETTER B WITH HOOK

<U0181> <S0062>; "<BASE><VRNT3>"; "<CAP><MIN>"; <U0181> % LATIN CAPITAL LETTER B WITH HOOK

<U0183> <S0062>; "<BASE><VRNT4>"; "<MIN><MIN>"; <U0183> % LATIN SMALL LETTER B WITH TOPBAR

<U0182> <S0062>; "<BASE><VRNT4>"; "<CAP><MIN>"; <U0182> % LATIN CAPITAL LETTER B WITH TOPBAR

<U0188> <S0063>; "<BASE><VRNT1>"; "<MIN><MIN>"; <U0188> % LATIN SMALL LETTER C WITH HOOK

<U0187> <S0063>; "<BASE><VRNT1>"; "<CAP><MIN>"; <U0187> % LATIN CAPITAL LETTER C WITH HOOK

<U0255> <S0063>; "<BASE><VRNT2>"; "<MIN><MIN>"; <U0255> % LATIN SMALL LETTER C WITH CURL

<U0297> <S0063>; "<BASE><VRNT3>"; "<MIN><MIN>"; <U0297> % LATIN LETTER STRETCHED C

% <VRNT1> is used for U00F0 LATIN SMALL LETTER ETH (already in CTT)

<U0256> <S0064>; "<BASE><VRNT2>"; "<MIN><MIN>"; <U0256> % LATIN SMALL LETTER D WITH TAIL

<U0189> <S0064>; "<BASE><VRNT2>"; "<CAP><MIN>"; <U0189> % LATIN CAPITAL LETTER AFRICAN D

<U0257> <S0064>; "<BASE><VRNT3>"; "<MIN><MIN>"; <U0257> % LATIN SMALL LETTER D WITH HOOK

<U018A> <S0064>; "<BASE><VRNT3>"; "<CAP><MIN>"; <U018A> % LATIN CAPITAL LETTER D WITH HOOK

<U018C> <S0064>; "<BASE><VRNT4>"; "<MIN><MIN>"; <U018C> % LATIN SMALL LETTER D WITH TOPBAR

<U018B> <S0064>; "<BASE><VRNT4>"; "<CAP><MIN>"; <U018B> % LATIN CAPITAL LETTER D WITH TOPBAR

<U0221> <S0064>; "<BASE><VRNT5>"; "<MIN><MIN>"; <U0221> % LATIN SMALL LETTER D WITH CURL

<U018D> <S0064>; "<BASE><VRNT6>"; "<MIN><MIN>"; <U018D> % LATIN SMALL LETTER TURNED DELTA

<U02A5> "<S0064><S007A>"; "<BASE><BASE><VRNT4>"; "<COMPAT><COMPAT><COMPAT>"; <U02A5> % LATIN SMALL LETTER DZ DIGRAPH WITH CURL

<U02A4> "<S0064><S007A>"; "<BASE><BASE><VRNT5>"; "<COMPAT><COMPAT><COMPAT>"; <U02A4> % LATIN SMALL LETTER DEZH DIGRAPH

<U01DD> <S0065>; "<BASE><VRNT1>"; "<MIN><MIN>"; <U01DD> % LATIN SMALL LETTER TURNED E

<U018E> <S0065>; "<BASE><VRNT1>"; "<CAP><MIN>"; <U018E> % LATIN CAPITAL LETTER REVERSED E

<U0259> <S0065>; "<BASE><VRNT2>"; "<MIN><MIN>"; <U0259> % LATIN SMALL LETTER SCHWA

<U018F> <S0065>; "<BASE><VRNT2>"; "<CAP><MIN>"; <U018F> % LATIN CAPITAL LETTER SCHWA

<U025B> <S0065>; "<BASE><VRNT3>"; "<MIN><MIN>"; <U025B> % LATIN SMALL LETTER OPEN E

<U0190> <S0065>; "<BASE><VRNT3>"; "<CAP><MIN>"; <U0190> % LATIN CAPITAL LETTER OPEN E

<U0258> <S0065>; "<BASE><VRNT4>"; "<MIN><MIN>"; <U0258> % LATIN SMALL LETTER REVERSED E

<U025A> <S0065>; "<BASE><VRNT5>"; "<MIN><MIN>"; <U025A> % LATIN SMALL LETTER SCHWA WITH HOOK

<U025C> <S0065>; "<BASE><VRNT6>"; "<MIN><MIN>"; <U025C> % LATIN SMALL LETTER REVERSED OPEN E

<U025D> <S0065>; "<BASE><VRNT7>"; "<MIN><MIN>"; <U025D> % LATIN SMALL LETTER REVERSED OPEN E WITH HOOK

<U025E> <S0065>; "<BASE><VRNT8>"; "<MIN><MIN>"; <U025E> % LATIN SMALL LETTER CLOSED REVERSED OPEN E

<U029A> <S0065>; "<BASE><VRNT9>"; "<MIN><MIN>"; <U029A> % LATIN SMALL LETTER CLOSED OPEN E

<U0192> <S0066>; "<BASE><VRNT1>"; "<MIN><MIN>"; <U0192> % LATIN SMALL LETTER F WITH HOOK

<U0191> <S0066>;"<BASE><VRNT1>";"<CAP><MIN>";<U0191> % LATIN CAPITAL LETTER F WITH HOOK

<U0261> <S0067>;"<BASE><VRNT1>";"<MIN><MIN>";<U0261> % LATIN SMALL LETTER SCRIPT G

<U0262> <S0067>;"<BASE><VRNT2>";"<MIN><MIN>";<U0262> % LATIN LETTER SMALL CAPITAL G

<U01E5> <S0067>;"<BASE><VRNT3>";"<MIN><MIN>";<U01E5> % LATIN SMALL LETTER G WITH STROKE

<U01E4> <S0067>;"<BASE><VRNT3>";"<CAP><MIN>";<U01E4> % LATIN CAPITAL LETTER G WITH STROKE

<U0260> <S0067>;"<BASE><VRNT4>";"<MIN><MIN>";<U0260> % LATIN SMALL LETTER G WITH HOOK

<U0193> <S0067>;"<BASE><VRNT4>";"<CAP><MIN>";<U0193> % LATIN CAPITAL LETTER G WITH HOOK

<U029B> <S0067>;"<BASE><VRNT5>";"<MIN><MIN>";<U029B> % LATIN LETTER SMALL CAPITAL G WITH HOOK

<U0263> <S0067>;"<BASE><VRNT6>";"<MIN><MIN>";<U0263> % LATIN SMALL LETTER GAMMA

<U0194> <S0067>;"<BASE><VRNT6>";"<CAP><MIN>";<U0194> % LATIN CAPITAL LETTER GAMMA

<U0264> <S0067>;"<BASE><VRNT7>";"<MIN><MIN>";<U0264> % LATIN SMALL LETTER RAMS HORN

<U01A3> <S0067>;"<BASE><VRNT8>";"<MIN><MIN>";<U01A3> % LATIN SMALL LETTER OI

<U01A2> <S0067>;"<BASE><VRNT8>";"<CAP><MIN>";<U01A2> % LATIN CAPITAL LETTER OI

<U029C> <S0068>;"<BASE><VRNT1>";"<MIN><MIN>";<U029C> % LATIN LETTER SMALL CAPITAL H

<U0266> <S0068>;"<BASE><VRNT2>";"<MIN><MIN>";<U0266> % LATIN SMALL LETTER H WITH HOOK

<U0267> <S0068>;"<BASE><VRNT3>";"<MIN><MIN>";<U0267> % LATIN SMALL LETTER HENG WITH HOOK

<U0265> <S0068>;"<BASE><VRNT4>";"<MIN><MIN>";<U0265> % LATIN SMALL LETTER TURNED H

<U02AE> <S0068>;"<BASE><VRNT5>";"<MIN><MIN>";<U02AE> % LATIN SMALL LETTER TURNED H WITH FISHHOOK

<U02AF> <S0068>;"<BASE><VRNT6>";"<MIN><MIN>";<U02AF> % LATIN SMALL LETTER TURNED H WITH FISHHOOK AND TAIL

<U0195> "<S0068><S0076>";"<BASE><BASE>";"<MIN><MIN>";<U0195> % LATIN SMALL LETTER HV

<U01F6> "<S0068><S0076>";"<BASE><BASE>";"<CAP><MIN>";<U01F6> % LATIN CAPITAL LETTER HWAIR

<U0131> <S0069>;"<BASE><VRNT1>";"<MIN><MIN>";<U0131> % LATIN SMALL LETTER DOTLESS I

<U026A> <S0069>;"<BASE><VRNT2>";"<MIN><MIN>";<U026A> % LATIN LETTER SMALL CAPITAL I

<U0268> <S0069>;"<BASE><VRNT3>";"<MIN><MIN>";<U0268> % LATIN SMALL LETTER I WITH STROKE

<U0197> <S0069>;"<BASE><VRNT3>";"<CAP><MIN>";<U0197> % LATIN CAPITAL LETTER I WITH STROKE

<U0269> <S0069>;"<BASE><VRNT4>";"<MIN><MIN>";<U0269> % LATIN SMALL LETTER IOTA

<U0196> <S0069>;"<BASE><VRNT4>";"<CAP><MIN>";<U0196> % LATIN CAPITAL LETTER IOTA

<U029D> <S006A>;"<BASE><VRNT1>";"<MIN><MIN>";<U029D> % LATIN SMALL LETTER J WITH CROSSED-TAIL

<U025F> <S006A>;"<BASE><VRNT2>";"<MIN><MIN>";<U025F> % LATIN SMALL LETTER DOTLESS J WITH STROKE

<U0284> <S006A>;"<BASE><VRNT3>";"<MIN><MIN>";<U0284> % LATIN SMALL LETTER DOTLESS J WITH STROKE AND HOOK

<U0199> <S006B>;"<BASE><VRNT1>";"<MIN><MIN>";<U0199> % LATIN SMALL LETTER K WITH HOOK

<U0198> <S006B>;"<BASE><VRNT1>";"<CAP><MIN>";<U0198> % LATIN CAPITAL LETTER K WITH HOOK

<U0138> <S006B>;"<BASE><VRNT2>";"<MIN><MIN>";<U0138> % LATIN SMALL LETTER KRA

<U029E> <S006B>;"<BASE><VRNT3>";"<MIN><MIN>";<U029E> % LATIN SMALL LETTER TURNED K

% <VRNT1> is used for U0140 LATIN SMALL LETTER L WITH MIDDLE DOT (already in CTT)

<U029F> <S006C>;"<BASE><VRNT2>";"<MIN><MIN>";<U029F> % LATIN LETTER SMALL CAPITAL L

<U019A> <S006C>;"<BASE><VRNT3>";"<MIN><MIN>";<U019A> % LATIN SMALL LETTER L WITH BAR

<U023D> <S006C>;"<BASE><VRNT3>";"<CAP><MIN>";<U023D> % LATIN CAPITAL LETTER L WITH BAR

<U026B> <S006C>;"<BASE><VRNT4>";"<MIN><MIN>";<U026B> % LATIN SMALL LETTER L WITH MIDDLE TILDE

<U026C> <S006C>;"<BASE><VRNT5>";"<MIN><MIN>";<U026C> % LATIN SMALL LETTER L WITH BELT

<U026D> <S006C>;"<BASE><VRNT6>";"<MIN><MIN>";<U026D> % LATIN SMALL LETTER L WITH RETROFLEX HOOK

<U0234> <S006C>;"<BASE><VRNT7>";"<MIN><MIN>";<U0234> % LATIN SMALL LETTER L WITH CURL

<U019B> <S006C>;"<BASE><VRNT8>";"<MIN><MIN>";<U019B> % LATIN SMALL LETTER LAMBDA WITH STROKE

<U026E> "<S006C><S007A>";"<BASE><BASE><VRNT5>";"<MIN><MIN><MIN>";<U026E> % LATIN SMALL LETTER LEZH

<U0271> <S006D>";"<BASE><VRNT1>";"<MIN><MIN>";<U0271> % LATIN SMALL LETTER M WITH HOOK
<U026F> <S006D>";"<BASE><VRNT2>";"<MIN><MIN>";<U026F> % LATIN SMALL LETTER TURNED M
<U019C> <S006D>";"<BASE><VRNT2>";"<CAP><MIN>";<U019C> % LATIN CAPITAL LETTER TURNED M
<U0270> <S006D>";"<BASE><VRNT3>";"<MIN><MIN>";<U0270> % LATIN SMALL LETTER TURNED M WITH LONG LEG

<U0149> <S006E>";"<BASE><VRNT1>";"<MIN><MIN>";<U0149> % LATIN SMALL LETTER N PRECEDED BY APOSTROPHE
<U0274> <S006E>";"<BASE><VRNT2>";"<MIN><MIN>";<U0274> % LATIN LETTER SMALL CAPITAL N
<U0272> <S006E>";"<BASE><VRNT3>";"<MIN><MIN>";<U0272> % LATIN SMALL LETTER N WITH LEFT HOOK
<U019D> <S006E>";"<BASE><VRNT3>";"<CAP><MIN>";<U019D> % LATIN CAPITAL LETTER N WITH LEFT HOOK
<U019E> <S006E>";"<BASE><VRNT4>";"<MIN><MIN>";<U019E> % LATIN SMALL LETTER N WITH LONG RIGHT LEG
<U0220> <S006E>";"<BASE><VRNT4>";"<CAP><MIN>";<U0220> % LATIN CAPITAL LETTER N WITH LONG RIGHT LEG
<U0273> <S006E>";"<BASE><VRNT5>";"<MIN><MIN>";<U0273> % LATIN SMALL LETTER N WITH RETROFLEX HOOK
<U0235> <S006E>";"<BASE><VRNT6>";"<MIN><MIN>";<U0235> % LATIN SMALL LETTER N WITH CURL
<U014B> <S006E>";"<BASE><VRNT7>";"<MIN><MIN>";<U014B> % LATIN SMALL LETTER ENG
<U014A> <S006E>";"<BASE><VRNT7>";"<CAP><MIN>";<U014A> % LATIN CAPITAL LETTER ENG

% <VRNT1> is used for U0153 LATIN SMALL LIGATURE OE (already in CTT)
<U0254> <S006F>";"<BASE><VRNT2>";"<MIN><MIN>";<U0254> % LATIN SMALL LETTER OPEN O
<U0186> <S006F>";"<BASE><VRNT2>";"<CAP><MIN>";<U0186> % LATIN CAPITAL LETTER OPEN O
<U0275> <S006F>";"<BASE><VRNT3>";"<MIN><MIN>";<U0275> % LATIN SMALL LETTER BARRED O
<U019F> <S006F>";"<BASE><VRNT3>";"<CAP><MIN>";<U019F> % LATIN CAPITAL LETTER O WITH MIDDLE TILDE
<U0277> <S006F>";"<BASE><VRNT4>";"<MIN><MIN>";<U0277> % LATIN SMALL LETTER CLOSED OMEGA
<U0223> <S006F>";"<BASE><VRNT5>";"<MIN><MIN>";<U0223> % LATIN SMALL LETTER OU
<U0222> <S006F>";"<BASE><VRNT5>";"<CAP><MIN>";<U0222> % LATIN CAPITAL LETTER OU

<U0276> "<S006F><S0065>";"<BASE><VRNT1><BASE>";"<COMPAT><COMPAT><COMPAT>";<U0276> % LATIN LETTER SMALL CAPITAL OE

<U01A5> <S0070>";"<BASE><VRNT1>";"<MIN><MIN>";<U01A5> % LATIN SMALL LETTER P WITH HOOK
<U01A4> <S0070>";"<BASE><VRNT1>";"<CAP><MIN>";<U01A4> % LATIN CAPITAL LETTER P WITH HOOK
<U0278> <S0070>";"<BASE><VRNT2>";"<MIN><MIN>";<U0278> % LATIN SMALL LETTER PHI

<U02A0> <S0071>";"<BASE><VRNT1>";"<MIN><MIN>";<U02A0> % LATIN SMALL LETTER Q WITH HOOK

<U0280> <S0072>";"<BASE><VRNT1>";"<MIN><MIN>";<U01A6> % LATIN LETTER SMALL CAPITAL R
<U01A6> <S0072>";"<BASE><VRNT1>";"<CAP><MIN>";<U01A6> % LATIN LETTER YR
<U0279> <S0072>";"<BASE><VRNT2>";"<MIN><MIN>";<U0279> % LATIN SMALL LETTER TURNED R
<U027A> <S0072>";"<BASE><VRNT3>";"<MIN><MIN>";<U027A> % LATIN SMALL LETTER TURNED R WITH LONG LEG
<U027B> <S0072>";"<BASE><VRNT4>";"<MIN><MIN>";<U027B> % LATIN SMALL LETTER TURNED R WITH HOOK
<U027C> <S0072>";"<BASE><VRNT5>";"<MIN><MIN>";<U027C> % LATIN SMALL LETTER R WITH LONG LEG
<U027D> <S0072>";"<BASE><VRNT6>";"<MIN><MIN>";<U027D> % LATIN SMALL LETTER R WITH TAIL
<U027E> <S0072>";"<BASE><VRNT7>";"<MIN><MIN>";<U027E> % LATIN SMALL LETTER R WITH FISHHOOK
<U027F> <S0072>";"<BASE><VRNT8>";"<MIN><MIN>";<U027F> % LATIN SMALL LETTER REVERSED R WITH FISHHOOK
<U0281> <S0072>";"<BASE><VRNT9>";"<MIN><MIN>";<U0281> % LATIN LETTER SMALL CAPITAL INVERTED R

% <VRNT1> is used for U00DF LATIN SMALL LETTER SHARP S (already in CTT)
% <VRNT2> is used for U017F LATIN SMALL LETTER LONG S (already in CTT)
<U0282> <S0073>";"<BASE><VRNT3>";"<MIN><MIN>";<U0282> % LATIN SMALL LETTER S WITH HOOK
<U0283> <S0073>";"<BASE><VRNT4>";"<MIN><MIN>";<U0283> % LATIN SMALL LETTER ESH
<U01A9> <S0073>";"<BASE><VRNT4>";"<CAP><MIN>";<U01A9> % LATIN CAPITAL LETTER ESH

<U01AA> <S0073>;"<BASE><VRNT5>;"<MIN><MIN>;<U01AA> % LATIN LETTER REVERSED ESH LOOP
 <U0285> <S0073>;"<BASE><VRNT6>;"<MIN><MIN>;<U0285> % LATIN SMALL LETTER SQUAT REVERSED ESH
 <U0286> <S0073>;"<BASE><VRNT7>;"<MIN><MIN>;<U0286> % LATIN SMALL LETTER ESH WITH CURL

<U0167> <S0074>;"<BASE><VRNT1>;"<MIN><MIN>;<U0167> % LATIN SMALL LETTER T WITH STROKE
 <U0166> <S0074>;"<BASE><VRNT1>;"<CAP><MIN>;<U0166> % LATIN CAPITAL LETTER T WITH STROKE
 <U01AB> <S0074>;"<BASE><VRNT2>;"<MIN><MIN>;<U01AB> % LATIN SMALL LETTER T WITH PALATAL HOOK
 <U01AD> <S0074>;"<BASE><VRNT3>;"<MIN><MIN>;<U01AD> % LATIN SMALL LETTER T WITH HOOK
 <U01AC> <S0074>;"<BASE><VRNT3>;"<CAP><MIN>;<U01AC> % LATIN CAPITAL LETTER T WITH HOOK
 <U0288> <S0074>;"<BASE><VRNT4>;"<MIN><MIN>;<U0288> % LATIN SMALL LETTER T WITH RETROFLEX HOOK
 <U01AE> <S0074>;"<BASE><VRNT4>;"<CAP><MIN>;<U01AE> % LATIN CAPITAL LETTER T WITH RETROFLEX HOOK
 <U0236> <S0074>;"<BASE><VRNT5>;"<MIN><MIN>;<U0236> % LATIN SMALL LETTER T WITH CURL
 <U0287> <S0074>;"<BASE><VRNT6>;"<MIN><MIN>;<U0287> % LATIN SMALL LETTER TURNED T

<U02A8> " <S0074><S0063>;"<BASE><BASE><VRNT2>;"<COMPAT><COMPAT><COMPAT>;<U02A8> % LATIN SMALL LETTER TC DIGRAPH WITH CURL

<U0289> <S0075>;"<BASE><VRNT1>;"<MIN><MIN>;<U0289> % LATIN SMALL LETTER U BAR
 <U0244> <S0075>;"<BASE><VRNT1>;"<CAP><MIN>;<U0244> % LATIN CAPITAL LETTER U BAR
 <U028A> <S0075>;"<BASE><VRNT2>;"<MIN><MIN>;<U028A> % LATIN SMALL LETTER UPSILON
 <U01B1> <S0075>;"<BASE><VRNT2>;"<CAP><MIN>;<U01B1> % LATIN CAPITAL LETTER UPSILON

<U028B> <S0076>;"<BASE><VRNT1>;"<MIN><MIN>;<U028B> % LATIN SMALL LETTER V WITH HOOK
 <U01B2> <S0076>;"<BASE><VRNT1>;"<CAP><MIN>;<U01B2> % LATIN CAPITAL LETTER V WITH HOOK
 <U028C> <S0076>;"<BASE><VRNT2>;"<MIN><MIN>;<U028C> % LATIN SMALL LETTER TURNED V
 <U0245> <S0076>;"<BASE><VRNT2>;"<CAP><MIN>;<U0245> % LATIN CAPITAL LETTER TURNED V

<U028D> <S0077>;"<BASE><VRNT1>;"<MIN><MIN>;<U028D> % LATIN SMALL LETTER TURNED W
 <U01BF> <S0077>;"<BASE><VRNT2>;"<MIN><MIN>;<U01BF> % LATIN LETTER WYNN
 <U01F7> <S0077>;"<BASE><VRNT2>;"<CAP><MIN>;<U01F7> % LATIN CAPITAL LETTER WYNN

<U028F> <S0079>;"<BASE><VRNT1>;"<MIN><MIN>;<U028F> % LATIN LETTER SMALL CAPITAL Y
 <U01B4> <S0079>;"<BASE><VRNT2>;"<MIN><MIN>;<U01B4> % LATIN SMALL LETTER Y WITH HOOK
 <U01B3> <S0079>;"<BASE><VRNT2>;"<CAP><MIN>;<U01B3> % LATIN CAPITAL LETTER Y WITH HOOK
 <U028E> <S0079>;"<BASE><VRNT3>;"<MIN><MIN>;<U028E> % LATIN SMALL LETTER TURNED Y
 <U021D> <S0079>;"<BASE><VRNT4>;"<MIN><MIN>;<U021D> % LATIN SMALL LETTER YOGH
 <U021C> <S0079>;"<BASE><VRNT9>;"<CAP><MIN>;<U021C> % LATIN CAPITAL LETTER YOGH

<U01B6> <S007A>;"<BASE><VRNT1>;"<MIN><MIN>;<U01B6> % LATIN SMALL LETTER Z WITH STROKE
 <U01B5> <S007A>;"<BASE><VRNT1>;"<CAP><MIN>;<U01B5> % LATIN CAPITAL LETTER Z WITH STROKE
 <U0225> <S007A>;"<BASE><VRNT2>;"<MIN><MIN>;<U0225> % LATIN SMALL LETTER Z WITH HOOK
 <U0224> <S007A>;"<BASE><VRNT2>;"<CAP><MIN>;<U0224> % LATIN CAPITAL LETTER Z WITH HOOK
 <U0290> <S007A>;"<BASE><VRNT3>;"<MIN><MIN>;<U0290> % LATIN SMALL LETTER Z WITH RETROFLEX HOOK
 <U0291> <S007A>;"<BASE><VRNT4>;"<MIN><MIN>;<U0291> % LATIN SMALL LETTER Z WITH CURL
 <U0292> <S007A>;"<BASE><VRNT5>;"<MIN><MIN>;<U0292> % LATIN SMALL LETTER EZH
 <U01B7> <S007A>;"<BASE><VRNT5>;"<CAP><MIN>;<U01B7> % LATIN CAPITAL LETTER EZH
 <U01EF> <S007A>;"<BASE><VRNT5><CARON>;"<MIN><MIN><MIN>;<U01EF> % LATIN SMALL LETTER EZH WITH CARON
 <U01EE> <S007A>;"<BASE><VRNT5><CARON>;"<CAP><MIN><MIN>;<U01EE> % LATIN CAPITAL LETTER EZH WITH CARON
 <U01B9> <S007A>;"<BASE><VRNT6>;"<MIN><MIN>;<U01B9> % LATIN SMALL LETTER EZH REVERSED
 <U01B8> <S007A>;"<BASE><VRNT6>;"<CAP><MIN>;<U01B8> % LATIN CAPITAL LETTER EZH REVERSED
 <U01BA> <S007A>;"<BASE><VRNT7>;"<MIN><MIN>;<U01BA> % LATIN SMALL LETTER EZH WITH TAIL
 <U0293> <S007A>;"<BASE><VRNT8>;"<MIN><MIN>;<U0293> % LATIN SMALL LETTER EZH WITH CURL

% Greek

% ISO14651_2006_TABLE1_en.txt now contains the tailorings of CR 14400 in its CTT

% Full conformance with GOST requirements for Cyrillic letters

<U0453> <S0452>;"<BASE><VRNT1>";"<MIN><MIN>";<U0453> % CYRILLIC SMALL LETTER GJE

<U0403> <S0452>;"<BASE><VRNT1>";"<CAP><MIN>";<U0403> % CYRILLIC CAPITAL LETTER GJE

<U045C> <S045B>;"<BASE><VRNT1>";"<MIN><MIN>";<U045C> % CYRILLIC SMALL LETTER KJE

<U040C> <S045B>;"<BASE><VRNT1>";"<CAP><MIN>";<U040C> % CYRILLIC CAPITAL LETTER KJE

% Georgian: Identical to ISO14651_2006_TABLE1_en.txt

% Armenian:

<U0587> <ECH-YIWN>;<BASE>;<CAP>;<U0587> % ARMENIAN SMALL LIGATURE ECH YIWN

reorder-end %% for EOR's EORDeltaTable

Anhang A (informativ)

Prinzipien hinter den Europäischen Sortierregeln

A.0 Einleitung

Dieser Anhang beabsichtigt, die in Abschnitt 6 enthaltene Information und die ihr zugrunde liegenden Prinzipien in einer leichter verständlichen Form zu erläutern. Diejenigen Leser, die sich nicht für die Einzelheiten der EDV-Umsetzung interessieren, können diese eher traditionelle Darstellung als eine gültige Interpretation des normativen Teils dieser Europäischen Norm ansehen.

A.1 Begriffe

In diesem informativen Anhang gelten die folgenden Begriffe zusätzlich zu denen im normativen Teil dieser Europäischen Norm (siehe Abschnitt 3).

A.1.1

Ziffer

eines der Zeichen

0 (U0030), 1 (U0031), 2 (U0032), 3 (U0033), 4 (U0034), 5 (U0035), 6 (U0036), 7 (U0037),
8 (U0038), 9 (U0039)

A.1.2

Buchstabe

Zeichen, das in der Schrift (alleine oder in Kombination mit anderen) Laute oder Lautfolgen einer natürlichen Sprache darstellt

ANMERKUNG In dieser Norm sind alle Zeichen des *Multilingual European Subset Nr. 3* gemeint, deren Name die Wörter `LETTER` oder `LIGATURE` enthält.

A.1.3

Buchstabe der ersten Stufe

Zeichen, das in der folgenden Auflistung von Buchstaben enthalten ist:

Lateinische Schrift:

a (U0061),	A (U0041),	b (U0062),	B (U0042),	c (U0063),	C (U0043),	d (U0064),	D (U0044),
e (U0065),	E (U0045),	f (U0066),	F (U0046),	g (U0067),	G (U0047),	h (U0068),	H (U0048),
i (U0069),	I (U0049),	j (U006A),	J (U004A),	k (U006B),	K (U004B),	l (U006C),	L (U004C),
m (U006D),	M (U004D),	n (U006E),	N (U004E),	o (U006F),	O (U004F),	p (U0070),	P (U0050),
q (U0071),	Q (U0051),	r (U0072),	R (U0052),	s (U0073),	S (U0053),	t (U0074),	T (U0054),
u (U0075),	U (U0055),	v (U0076),	V (U0056),	w (U0077),	W (U0057),	x (U0078),	X (U0058),

y (U0079), Y (U0059), z (U007A), Z (U005A), þ (U00FE), Þ (U00DE)

Griechische Schrift:

α (U03B1), Α (U0391), β (U03B2), Β (U0392), γ (U03B3), Γ (U0393), δ (U03B4), Δ (U0394),
ε (U03B5), Ε (U0395), Ϝ (U03DC), Ϛ (U03DA), Ϝ (U03B6), Ζ (U0396), η (U03B7), Η (U0397),
θ (U03B8), Θ (U0398), ι (U03B9), Ι (U0399), κ (U03BA), Κ (U039A), λ (U03BB), Λ (U039B),
μ (U03BC), Μ (U039C), ν (U03BD), Ν (U039D), ξ (U03BE), Ξ (U039E), ο (U03BF), Ο (U039F),
π (U03C0), Π (U03A0), Ϛ (U03DE), ϑ (U03C1), Ρ (U03A1), σ (U03C3), Σ (U03A3), τ (U03C4),
Τ (U03A4), υ (U03C5), Υ (U03A5), ϕ (U03C6), Φ (U03A6), χ (U03C7), Χ (U03A7), ψ (U03C8),
Ψ (U03A8), ω (U03C9), Ω (U03A9), Ϟ (U03E0)

ANMERKUNG 1 Stigma Ϛ (U03DA / U03DB), Koppa / Qoppa Ϛ (U03DE / U03DF) und Sampi Ϟ (U03E0 / U03E1) sind archaische Buchstaben, die jetzt als Zahlzeichen verwendet werden. Digamma Ϝ (U03DC) wird in keiner modernen Sprache verwendet.

ANMERKUNG 2 Im Zeichenvorrat 9 „Griechische und koptische Zeichen“ aus ISO/IEC 10646:2003 enthält MES auch eine Anzahl koptischer Buchstaben. Ihre Sortierung ist in ISO/IEC 14651 festgelegt.

Kyrillische Schrift:

а (U0430), А (U0410), ә (U04D1), Ă (U04D0), ä (U04D3), Ä (U04D2), ə (U04D9), Ə (U04D8),
ä (U04DB), Ă (U04DA), æ (U04D5), Æ (U04D4), б (U0431), Б (U0411), в (U0432), В (U0412),
г (U0433), Г (U0413), ғ (U0493), Ғ (U0492), Һ (U0495), Һ (U0494), д (U0434), Д (U0414),
ђ (U0452), Ѓ (U0402), ҙ (U0499), Ҙ (U0498), е (U0435), Е (U0415), ё (U04D7), Ё (U04D6),
є (U0454), Є (U0404), ж (U0436), Ж (U0416), ӗ (U04DD), Ӗ (U04DC), ӗ (U0497), Ӗ (U0496),
з (U0437), З (U0417), ӓ (U04DF), Ӛ (U04DE), s (U0455), S (U0405), ӓ (U04E1), Ӛ (U04E0),
и (U0438), И (U0418), ӗ (U04E5), Ӛ (U04E4), i (U0456), I (U0406), ӗ (U0457), Ӛ (U0407),
ӗ (U0439), Ӛ (U0419), ӗ (U048B), Ӛ (U048A), j (U0458), J (U0408), к (U043A), К (U041A),
қ (U049B), Қ (U049A), ӗ (U04C4), Ӛ (U04C3), ӗ (U04A1), Ӛ (U04A0), ӗ (U049F), Ӛ (U049E),
к (U049D), К (U049C), л (U043B), Л (U041B), ӗ (U04C6), Ӛ (U04C5), ӗ (U0459), Ӛ (U0409),
м (U043C), М (U041C), ӗ (U04CE), Ӛ (U04CD), н (U043D), Н (U041D), ӗ (U04CA), Ӛ (U04C9),
ң (U04A3), Ӛ (U04A2), ӗ (U04C8), Ӛ (U04C7), ӗ (U04A5), Ӛ (U04A4), ӗ (U045A), Ӛ (U040A),
о (U043E), О (U041E), ö (U04E7), Ö (U04E6), ӗ (U04E9), Ӛ (U04E8), ӗ (U04EB), Ӛ (U04EA),