SLOVENSKI STANDARD
SIST ENV 13710:2003

01-oktober-2003

Evropska pravila za razpored – Razpored za latinsko, grško in cirilsko pisavo

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

Ta slovenski standard je istoveten z: ENV 13710:2000

ICS:

01.140.20	Informacijske vede	Information sciences
35.040	Nabori znakov in kodiranje informacij	Character sets and information coding

SIST ENV 13710:2003 **en**

EUROPEAN PRESTANDARD

PRÉNORME EUROPÉENNE

EUROPÄISCHE VORNORM

ENV 13710

December 2000

ICS 01.140.20; 35.040

English version

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

This European Prestandard (ENV) was approved by CEN on 4 February 2000 as a prospective standard for provisional application.

The period of validity of this ENV is limited initially to three years. After two years the members of CEN will be requested to submit their comments, particularly on the question whether the ENV can be converted into a European Standard.

CEN members are required to announce the existence of this ENV in the same way as for an EN and to make the ENV available promptly at national level in an appropriate form. It is permissible to keep conflicting national standards in force (in parallel to the ENV) until the final decision about the possible conversion of the ENV into an EN is reached.

CEN members are the national standards bodies of Austria, Belgium, Czech Republic, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Luxembourg, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland and United Kingdom.

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

Management Centre: rue de Stassart, 36 B-1050 Brussels

Ref. No. ENV 13710:2000 E

Foreword

This European Prestandard has been prepared by Technical Committee CEN/TC 304 "Information and communications technologies - European localization requirements", the secretariat of which is held by STRI.

According to the CEN/CENELEC Internal Regulations, the national standards organizations of the following countries are bound to announce this European Prestandard: Austria, Belgium, Czech Republic, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Luxembourg, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland and the United Kingdom.

This European Prestandard is intended to facilitate cross border communications and data exchange and to ensure that European cultural requirements are safeguarded in the increasingly interconnected world of today. It provides rules for ordering multilingual European texts and data into a single sequence. These rules come into effect if data from different languages must be brought into a predictable order that makes it easy for users to find information, which is often the case in pan-European applications.

This European Prestandard does not intend to influence, let alone substitute itself for, national standards or customs in this field. Nevertheless, national standards have the opportunity to adapt this European Prestandard by declaring a formalized set of deviation rules ("delta") if they so wish.

This is an ENV and is thus not binding on the National Member Bodies of CEN. It exists for the purpose of validating its usefulness before it might be proposed for an EN in due time. At that stage, the sample repertoire which the current Prestandard is using, could be replaced by a more formally defined character repertoire, if by then in existence.

Sorting assists users by presenting information in a structured way. This may include the subdivision of information by subject matters, e. g. by having several registers in a book, by splitting a phone book into several sections, one for each town that falls into its purview or by having multiple indices in a library. *Ordering* – the arrangement of information in alphabetical sequence – is in most circumstances an integral part of this procedure.

This European Prestandard must cater for two mutually exclusive demands: Implementers need clear guidelines and data which can readily be used in existing and future ordering applications. This can best be done by defining a European default ordering table in the syntax of the forthcoming ordering standard *ISO/IEC FDIS 14651*, of which the present standard is a "profile". Users with no specific IT-background, however, need an explanation of the principles in a form more in line with existing national ordering standards or relevant practice. As tailoring tables in the syntax of *ISO/IEC FDIS 14651* can be difficult to read for human readers, an explanation of the principles behind that table is given in the informative annexes. They are written in a more general style and users not familiar with the formal syntax of the tailoring table are advised to consult those annexes first. A web site on this subject is hosted by the Icelandic Standards Organization STRÍ for further reference.[1]

---

[1]At present STRÌ can be accessed under http://www.stri.is

1 Scope

This European Prestandard specifies the sequence to be established by alphabetical ordering of multilingual data composed of characters comprised in the *Multilingual European Subset Number 2* or subsets thereof. *MES-2* is a collection has been proposed for addition to Annex A of *ISO/IEC 10646-1:2000*.

NOTE The *Multilingual European Subset Number 2* is usually termed MES-2 and was formerly known as the *Minimum European Subset* (MES). Cf. *ENV 1973:1995*. MES-2 basically covers the Latin, Greek and Cyrillic letters needed in European data interchange and commerce and a variety of symbols. A CEN workshop agreement on the *Multilingual European Subsets of ISO/IEC 10646* has been published as *CEN ISSS CWA 13873*.

The ordering rules given here are only intended for data in more than one European language. They are not meant to influence, let alone replace existing national standards or practices.

The main part of this European Prestandard specifies letter-by-letter ordering of character strings. Informative Annex A presents equivalent information in a more readily accessible way. Informative Annex B deals with word-by-word ordering as a special form of ordering with multiple keys. Informative Annex C explains the use of further ordering criteria. Informative Annex D presents a widely used alternative to the main part, namely the amalgamation of several scripts in one index via implicit transliteration. Informative Annex F, finally, presents the information inherent in section 6 of the body of this European Prestandard in a formally equivalent, though condensed, form.

NOTE This European Prestandard is designed for upwards compatibility towards larger repertoires which may include combining diacritics. A sample of such a repertoire is the *Multilingual Subset No. 3 (MES-3)* of *CEN ISSS CWA 13783*.

Following the practice of *ISO/IEC FDIS 14651* characters are referenced as UXXXX where *X* stands for any hexadecimal digit and refers to the value of that character in *ISO/IEC 10646-1:2000*.This convention is used throughout this European Prestandard.

2 Normative references

This European Prestandard incorporates by dated or undated reference provisions from other publications. These normative references are quoted at the appropriate places in the text, and the publications are listed hereafter.

All standards are subject to revision. Dated references do not always refer to subsequent amendments of the publication in question. Undated references always refer to the latest edition.

ISO/IEC 10646-1:2000, Information Technology – Universal Multi-Octet Coded Character set (UCS). Second edition.

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

ISO/IEC FDIS 14651, International string ordering and comparison — Method for comparing character strings and description of the common template tailorable ordering

NOTE To be published

3 Definitions

For the purpose of this European Prestandard the following definitions of *ISO/IEC 10646-1* and of *ISO/IEC FDIS 14651* apply:

3.1
character
A member of a set of elements used for the organisation, control, or representation of data. [ISO/IEC 10646-1]

NOTE For the purpose of this European Prestandard always a member of the *Multilingual European Subset No 2.*

3.2
character string
a sequence of characters [ISO/IEC FDIS 14651]

3.3
delta
differences from a given collation table. The given collation table, together with a given delta, forms a new collation table. Unless otherwise specified in the European Prestandard, the term " delta" always refers to differences from the Common Template Table as defined in ISO/IEC FDIS 14651 [ISO/IEC FDIS 14651]

3.4
ordering
a process by which two strings are determined to be in exactly one of the relationships of less than, greater than or equal to another [ISO/IEC FDIS 14651]

4 Conformance

In order to be conformant to this European Prestandard an application shall meet the requirements prescribed in section 6 of ISO/IEC FDIS 14651 and use the default table of section 6 or an equivalent description of the information contained therein.

5 Tailorability

The European Ordering Rules defined in this standard can be taken as a template which can be tailored to the needs of any European country in the manner specified by *ISO/IEC FDIS 14651* .

6 Default Table

NOTE For the syntax of the table please consult *ISO/IEC FDIS 14651.*

```
%% EOR's EORTable
%
%% European Ordering Rules
%
% EOR delta for MES-2 from ISO/IEC 14651:2000's CTT (ISO14651_2000_TABLE1).
%
% This delta gives only the actual changes from the first edition of the CTT
% for punctuation and modifier letters, but includes explicit tailoring lines
% for all digits and letters in MES-2 even where no substantial change (or
% addition) is intended.
%
% Note that here new collating symbols are used instead of the CTT ones.
%
% This delta is equivalent to EORDeltaTable (Annex F) if only MES-2 characters
% occur. It is usually NOT equivalent to EORTable if any character(s) outside
% of MES-2 occur, due to the use of new collation weight names in EORTable.


reorder-after <BLK> %

%% Declarations of the third level ordering weight names
collating-symbol <ACCENT>
```

```
collating-symbol <SMALL>
collating-symbol <COMPAT>
collating-symbol <SUPERSCRIPT>
collating-symbol <SUBSCRIPT>
collating-symbol <CAPITAL>
collating-symbol <FRACTION>

%% Weight assignments of the third level ordering weights
<ACCENT> % or modification weight at level 2
<SMALL>
<COMPAT>
<SUPERSCRIPT>
<SUBSCRIPT>
<CAPITAL>
<FRACTION>

reorder-after <BASE> %

%% Declarations of the second level ordering weight names
collating-symbol <BASE>
collating-symbol <LIGATED>
collating-symbol <PSILI>
collating-symbol <DASIA>
%% Compliance with the principles of ISO/IEC FCD 14651 requires that
%% OXIA and ACUTE be treated identically
collating-symbol <ACUTE>
%% Compliance with the principles of ISO/IEC FCD 14651 requires that
%% VARIA and GRAVE be treated identically
collating-symbol <GRAVE>
collating-symbol <PERISPOMENI>
collating-symbol <TONOS>
%% Compliance with the principles of ISO/IEC FCD 14651 requires that
%% PROSGEGRAMMENI and YPOGEGRAMMENI be treated identically
collating-symbol <YPOGEGRAMMENI>
%% BREVE is unified with the Greek VRACHY
%% This does not influence the ordering results
collating-symbol <BREVE>
collating-symbol <CIRCUMFLEX>
collating-symbol <CARON>
collating-symbol <RING_ABOVE>
%% DIAERESIS is unified with the Greek DIALYTICA
%% This does not influence the ordering results
collating-symbol <DIAERESIS>
collating-symbol <DOUBLE_ACUTE>
collating-symbol <TILDE>
collating-symbol <DOT_ABOVE>
collating-symbol <CEDILLA>
collating-symbol <COMMA_BELOW>
collating-symbol <OGONEK>
collating-symbol <MACRON>
collating-symbol <STROKE>
collating-symbol <MODIFIED>
collating-symbol <MODIFIED2>

%% Weight assignments for the second level ordering weights

<BASE>
<LIGATED>

<PSILI>
<DASIA>
<ACUTE>
<GRAVE>
<PERISPOMENI>
<TONOS>
<YPOGEGRAMMENI>
<BREVE>
<CIRCUMFLEX>
<CARON>
<RING_ABOVE>
```

```
<DIAERESIS>
<DOUBLE_ACUTE>
<TILDE>
<DOT_ABOVE>
<CEDILLA>
<COMMA_BELOW>
<OGONEK>
<MACRON>
<STROKE>
<MODIFIED>
<MODIFIED2>


reorder-after <S0000> %

%% Declarations of the first level ordering weight names

%%% Digits
collating-symbol <0>
collating-symbol <1>
collating-symbol <2>
collating-symbol <3>
collating-symbol <4>
collating-symbol <5>
collating-symbol <6>
collating-symbol <7>
collating-symbol <8>
collating-symbol <9>

%%% Latin
collating-symbol <a>
collating-symbol <b>
collating-symbol <c>
collating-symbol <d>
collating-symbol <e>
collating-symbol <f>
collating-symbol <g>
collating-symbol <h>
collating-symbol <i>
collating-symbol <j>
collating-symbol <k>
collating-symbol <l>
collating-symbol <m>
collating-symbol <n>
collating-symbol <o>
collating-symbol <p>
collating-symbol <q>
collating-symbol <r>
collating-symbol <s>
collating-symbol <t>
collating-symbol <u>
collating-symbol <v>
collating-symbol <w>
collating-symbol <x>
collating-symbol <y>
collating-symbol <z>
collating-symbol <thorn>

%%% Greek
collating-symbol <alpha>
collating-symbol <beta>
collating-symbol <gamma>
collating-symbol <delta>
collating-symbol <epsilon>
collating-symbol <digamma>
collating-symbol <stigma>
collating-symbol <zeta>
collating-symbol <eta>
collating-symbol <theta>
collating-symbol <iota>
```

```
collating-symbol <kappa>
collating-symbol <lamda>
collating-symbol <mu>
collating-symbol <nu>
collating-symbol <xi>
collating-symbol <omicron>
collating-symbol <pi>
collating-symbol <koppa>
collating-symbol <rho>
collating-symbol <sigma>
collating-symbol <tau>
collating-symbol <upsilon>
collating-symbol <phi>
collating-symbol <chi>
collating-symbol <psi>
collating-symbol <omega>
collating-symbol <sampi>

%%% Cyrillic
collating-symbol <cyr_a>
collating-symbol <cyr_a_with_breve>
collating-symbol <cyr_a_with_diaeresis>
collating-symbol <cyr_schwa>
collating-symbol <cyr_schwa_with_diaeresis>
collating-symbol <cyr_a_ie>
collating-symbol <cyr_b>
collating-symbol <cyr_v>
collating-symbol <cyr_ghe>
collating-symbol <cyr_ghe_with_stroke>
collating-symbol <cyr_ghe_with_middle_hook>
collating-symbol <cyr_d>
collating-symbol <cyr_dje>
collating-symbol <cyr_ze_with_descender>
collating-symbol <cyr_ie>
collating-symbol <cyr_ie_with_breve>
collating-symbol <cyr_ukrainian_ie>
collating-symbol <cyr_zhe>
collating-symbol <cyr_zhe_with_diaeresis>
collating-symbol <cyr_zhe_with_descender>
collating-symbol <cyr_ze>
collating-symbol <cyr_ze_with_diaeresis>
collating-symbol <cyr_dze>
collating-symbol <cyr_abkhasian_dze>
collating-symbol <cyr_i>
collating-symbol <cyr_i_with_diaeresis>
collating-symbol <cyr_byelorussian-ukrainian_i>
collating-symbol <cyr_yi>
collating-symbol <cyr_short_i>
collating-symbol <cyr_je>
collating-symbol <cyr_k>
collating-symbol <cyr_k_with_descender>
collating-symbol <cyr_k_with_hook>
collating-symbol <cyr_bashkir_k>
collating-symbol <cyr_k_with_stroke>
collating-symbol <cyr_k_with_vertical_stroke>
collating-symbol <cyr_l>
collating-symbol <cyr_lje>
collating-symbol <cyr_m>
collating-symbol <cyr_n>
collating-symbol <cyr_n_with_descender>
collating-symbol <cyr_n_with_hook>
collating-symbol <cyr_n_ghe>
collating-symbol <cyr_nje>
collating-symbol <cyr_o>
collating-symbol <cyr_o_with_diaeresis>
collating-symbol <cyr_barred_o>
collating-symbol <cyr_barred_o_with_diaeresis>
collating-symbol <cyr_p>
collating-symbol <cyr_p_with_middle_hook>
collating-symbol <cyr_r>
```

```
collating-symbol <cyr_s>
collating-symbol <cyr_s_with_descender>
collating-symbol <cyr_t>
collating-symbol <cyr_t_with_descender>
collating-symbol <cyr_tshe>
collating-symbol <cyr_u>
collating-symbol <cyr_short_u>
collating-symbol <cyr_u_with_diaeresis>
collating-symbol <cyr_u_with_double_acute>
collating-symbol <cyr_straight_u>
collating-symbol <cyr_straight_u_with_stroke>
collating-symbol <cyr_f>
collating-symbol <cyr_h>
collating-symbol <cyr_h_with_descender>
collating-symbol <cyr_shha>
collating-symbol <cyr_tse>
collating-symbol <cyr_t_tse>
collating-symbol <cyr_che>
collating-symbol <cyr_che_with_diaeresis>
collating-symbol <cyr_che_with_descender>
collating-symbol <cyr_khakassian_che>
collating-symbol <cyr_che_with_vertical_stroke>
collating-symbol <cyr_abkhasian_che>
collating-symbol <cyr_abkhasian_che_with_descender>
collating-symbol <cyr_dzhe>
collating-symbol <cyr_sha>
collating-symbol <cyr_shcha>
collating-symbol <cyr_hard_sign>
collating-symbol <cyr_yeru>
collating-symbol <cyr_yeru_with_diaeresis>
collating-symbol <cyr_soft_sign>
collating-symbol <cyr_e>
collating-symbol <cyr_yu>
collating-symbol <cyr_ya>
collating-symbol <cyr_abkhasian_ha>
collating-symbol <cyr_palochka>
```



```
%%%Weight assignments for the first level ordering weights

%%% Digits
<0>
<1>
<2>
<3>
<4>
<5>
<6>
<7>
<8>
<9>

%%% Latin
<a>
<b>
<c>
<d>
<e>
<f>
<g>
<h>
<i>
<j>
<k>
<l>
<m>
<n>
```

```
<o>
<p>
<q>
<r>
<s>
<t>
<u>
<v>
<w>
<x>
<y>
<z>
<thorn>

%%% Greek
<alpha>
<beta>
<gamma>
<delta>
<epsilon>
<digamma>
<stigma>
<zeta>
<eta>
<theta>
<iota>
<kappa>
<lamda>
<mu>
<nu>
<xi>
<omicron>
<pi>
<koppa>
<rho>
<sigma>
<tau>
<upsilon>
<phi>
<chi>
<psi>
<omega>
<sampi>

%%% Cyrillic
<cyr_a>
<cyr_a_with_breve>
<cyr_a_with_diaeresis>
<cyr_schwa>
<cyr_schwa_with_diaeresis>
<cyr_a_ie>
<cyr_b>
<cyr_v>
<cyr_ghe>
<cyr_ghe_with_stroke>
<cyr_ghe_with_middle_hook>
<cyr_d>
<cyr_dje>
<cyr_ze_with_descender>
<cyr_ie>
<cyr_ie_with_breve>
<cyr_ukrainian_ie>
<cyr_zhe>
<cyr_zhe_with_diaeresis>
<cyr_zhe_with_descender>
<cyr_ze>
<cyr_ze_with_diaeresis>
<cyr_dze>
<cyr_abkhasian_dze>
<cyr_i>
```

```
<cyr_i_with_diaeresis>
<cyr_byelorussian-ukrainian_i>
<cyr_yi>
<cyr_short_i>
<cyr_je>
<cyr_k>
<cyr_k_with_descender>
<cyr_k_with_hook>
<cyr_bashkir_k>
<cyr_k_with_stroke>
<cyr_k_with_vertical_stroke>
<cyr_l>
<cyr_lje>
<cyr_m>
<cyr_n>
<cyr_n_with_descender>
<cyr_n_with_hook>
<cyr_n_ghe>
<cyr_nje>
<cyr_o>
<cyr_o_with_diaeresis>
<cyr_barred_o>
<cyr_barred_o_with_diaeresis>
<cyr_p>
<cyr_p_with_middle_hook>
<cyr_r>
<cyr_s>
<cyr_s_with_descender>
<cyr_t>
<cyr_t_with_descender>
<cyr_tshe>
<cyr_u>
<cyr_short_u>
<cyr_u_with_diaeresis>
<cyr_u_with_double_acute>
<cyr_straight_u>
<cyr_straight_u_with_stroke>
<cyr_f>
<cyr_h>
<cyr_h_with_descender>
<cyr_shha>
<cyr_tse>
<cyr_t_tse>
<cyr_che>
<cyr_che_with_diaeresis>
<cyr_che_with_descender>
<cyr_khakassian_che>
<cyr_che_with_vertical_stroke>
<cyr_abkhasian_che>
<cyr_abkhasian_che_with_descender>
<cyr_dzhe>
<cyr_sha>
<cyr_shcha>
<cyr_hard_sign>
<cyr_yeru>
<cyr_yeru_with_diaeresis>
<cyr_soft_sign>
<cyr_e>
<cyr_yu>
<cyr_ya>
<cyr_abkhasian_ha>
<cyr_palochka>

collating-element <U000D_U000A> from "<U000D><U000A>"

reorder-after <SFFFF> % The only place where we can put the order_start line.

order_start forward;forward;forward;forward
```

```
% Reweighted (relative to ISO14651_2000_TABLE1) non-alphanumeric
% characters (including some modifier letters):


% Currency signs (DRACHMA SIGN is not in ISO14651_2000_TABLE1):
<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
<U20A3> IGNORE;IGNORE;IGNORE;<U20A3> % FRENCH FRANC SIGN
<U20A4> IGNORE;IGNORE;IGNORE;<U20A4> % LIRA SIGN
<U20A7> IGNORE;IGNORE;IGNORE;<U20A7> % PESETA SIGN
<U20AC> IGNORE;IGNORE;IGNORE;<U20AC> % EURO SIGN
<U20AF> IGNORE;IGNORE;IGNORE;<U20AF> % DRACHMA SIGN

% General catagory Lm (M.L. DOUBLE APOSTROPHE is not in ISO14651_2000_TABLE1):
<U02BB> IGNORE;IGNORE;IGNORE;<U02BB> % MODIFIER LETTER TURNED COMMA
<U02BD> IGNORE;IGNORE;IGNORE;<U02BD> % MODIFIER LETTER REVERSED COMMA
<U02BC> IGNORE;IGNORE;IGNORE;<U02BC> % MODIFIER LETTER APOSTROPHE
<U02EE> IGNORE;IGNORE;IGNORE;<U02EE> % MODIFIER LETTER DOUBLE APOSTROPHE



reorder-after <U0030>   % After DIGIT ZERO, just to make the 4th level
            % weights heavier than for punctuation

%% Digits (no substantial change from ISO14651_2000_TABLE1)

<U0030> <0>;<BASE>;<SMALL>;<U0030> % DIGIT ZERO
<U0031> <1>;<BASE>;<SMALL>;<U0031> % DIGIT ONE
<U00B9> <1>;<BASE>;<SUPERSCRIPT>;<U00B9> % SUPERSCRIPT ONE
<U00BD> "<1><2>";"<BASE><BASE>";"<FRACTION><FRACTION>";<U00BD> % VULGAR FRACTION ONE
HALF
<U00BC> "<1><4>";"<BASE><BASE>";"<FRACTION><FRACTION>";<U00BC> % VULGAR FRACTION ONE
QUARTER
<U215B> "<1><8>";"<BASE><BASE>";"<FRACTION><FRACTION>";<U215B> % VULGAR FRACTION ONE
EIGHTH
<U0032> <2>;<BASE>;<SMALL>;<U0032> % DIGIT TWO
<U00B2> <2>;<BASE>;<SUPERSCRIPT>;<U00B2> % SUPERSCRIPT TWO
<U2082> <2>;<BASE>;<SUBSCRIPT>;<U2082> % SUBSCRIPT TWO
<U0033> <3>;<BASE>;<SMALL>;<U0033> % DIGIT THREE
<U00B3> <3>;<BASE>;<SUPERSCRIPT>;<U00B3> % SUPERSCRIPT THREE
<U00BE> "<3><4>";"<BASE><BASE>";"<FRACTION><FRACTION>";<U00BE> % VULGAR FRACTION
THREE QUARTERS
<U215C> "<3><8>";"<BASE><BASE>";"<FRACTION><FRACTION>";<U215C> % VULGAR FRACTION
THREE EIGHTHS
<U0034> <4>;<BASE>;<SMALL>;<U0034> % DIGIT FOUR
<U0035> <5>;<BASE>;<SMALL>;<U0035> % DIGIT FIVE
<U215D> "<5><8>";"<BASE><BASE>";"<FRACTION><FRACTION>";<U215D> % VULGAR FRACTION FIVE
EIGHTHS
<U0036> <6>;<BASE>;<SMALL>;<U0036> % DIGIT SIX
<U0037> <7>;<BASE>;<SMALL>;<U0037> % DIGIT SEVEN
<U215E> "<7><8>";"<BASE><BASE>";"<FRACTION><FRACTION>";<U215E> % VULGAR FRACTION
SEVEN EIGHTHS
<U0038> <8>;<BASE>;<SMALL>;<U0038> % DIGIT EIGHT
<U0039> <9>;<BASE>;<SMALL>;<U0039> % DIGIT NINE


%% Latin
% Almost all substantial changes here result from CEN/TC304's resolution
% for the Latin script part of MES-2 to treat only the letters a to z and
% thorn as distinct on the first level and have LETTER AE treated
% as a ligature, similar to how LIGATURE OE is treated in the CTT.
% Note that H WITH CARON is not in ISO14651_2000_TABLE1.

<U0061> <a>;<BASE>;<SMALL>;<U0061> % LATIN SMALL LETTER A
<U0041> <a>;<BASE>;<CAPITAL>;<U0041> % LATIN CAPITAL LETTER A
<U00AA> <a>;<BASE>;<SUPERSCRIPT>;<U00AA> % FEMININE ORDINAL INDICATOR
```

<U00E1> <a>;"<BASE><ACUTE>";"<SMALL><ACCENT>";<U00E1> % LATIN SMALL LETTER A WITH ACUTE
<U00C1> <a>;"<BASE><ACUTE>";"<CAPITAL><ACCENT>";<U00C1> % LATIN CAPITAL LETTER A WITH ACUTE
<U00E0> <a>;"<BASE><GRAVE>";"<SMALL><ACCENT>";<U00E0> % LATIN SMALL LETTER A WITH GRAVE
<U00C0> <a>;"<BASE><GRAVE>";"<CAPITAL><ACCENT>";<U00C0> % LATIN CAPITAL LETTER A WITH GRAVE
<U0103> <a>;"<BASE><BREVE>";"<SMALL><ACCENT>";<U0103> % LATIN SMALL LETTER A WITH BREVE
<U0102> <a>;"<BASE><BREVE>";"<CAPITAL><ACCENT>";<U0102> % LATIN CAPITAL LETTER A WITH BREVE
<U00E2> <a>;"<BASE><CIRCUMFLEX>";"<SMALL><ACCENT>";<U00E2> % LATIN SMALL LETTER A WITH CIRCUMFLEX
<U00C2> <a>;"<BASE><CIRCUMFLEX>";"<CAPITAL><ACCENT>";<U00C2> % LATIN CAPITAL LETTER A WITH CIRCUMFLEX
<U00E5> <a>;"<BASE><RING_ABOVE>";"<SMALL><ACCENT>";<U00E5> % LATIN SMALL LETTER A WITH RING ABOVE
<U00C5> <a>;"<BASE><RING_ABOVE>";"<CAPITAL><ACCENT>";<U00C5> % LATIN CAPITAL LETTER A WITH RING ABOVE
<U01FB> <a>;"<BASE><RING_ABOVE><ACUTE>";"<SMALL><ACCENT><ACCENT>";<U01FB> % LATIN SMALL LETTER A WITH RING ABOVE AND ACUTE
<U01FA> <a>;"<BASE><RING_ABOVE><ACUTE>";"<CAPITAL><ACCENT><ACCENT>";<U01FA> % LATIN CAPITAL LETTER A WITH RING ABOVE AND ACUTE
<U00E4> <a>;"<BASE><DIAERESIS>";"<SMALL><ACCENT>";<U00E4> % LATIN SMALL LETTER A WITH DIAERESIS
<U00C4> <a>;"<BASE><DIAERESIS>";"<CAPITAL><ACCENT>";<U00C4> % LATIN CAPITAL LETTER A WITH DIAERESIS
<U01DF> <a>;"<BASE><DIAERESIS><MACRON>";"<SMALL><ACCENT><ACCENT>";<U01DF> % LATIN SMALL LETTER A WITH DIAERESIS AND MACRON
<U01DE> <a>;"<BASE><DIAERESIS><MACRON>";"<CAPITAL><ACCENT><ACCENT>";<U01DE> % LATIN CAPITAL LETTER A WITH DIAERESIS AND MACRON
<U00E3> <a>;"<BASE><TILDE>";"<SMALL><ACCENT>";<U00E3> % LATIN SMALL LETTER A WITH TILDE
<U00C3> <a>;"<BASE><TILDE>";"<CAPITAL><ACCENT>";<U00C3> % LATIN CAPITAL LETTER A WITH TILDE
<U01E1> <a>;"<BASE><DOT_ABOVE><MACRON>";"<SMALL><ACCENT><ACCENT>";<U01E1> % LATIN SMALL LETTER A WITH DOT ABOVE AND MACRON
<U01E0> <a>;"<BASE><DOT_ABOVE><MACRON>";"<CAPITAL><ACCENT><ACCENT>";<U01E0> % LATIN CAPITAL LETTER A WITH DOT ABOVE AND MACRON
<U0105> <a>;"<BASE><OGONEK>";"<SMALL><ACCENT>";<U0105> % LATIN SMALL LETTER A WITH OGONEK
<U0104> <a>;"<BASE><OGONEK>";"<CAPITAL><ACCENT>";<U0104> % LATIN CAPITAL LETTER A WITH OGONEK
<U0101> <a>;"<BASE><MACRON>";"<SMALL><ACCENT>";<U0101> % LATIN SMALL LETTER A WITH MACRON
<U0100> <a>;"<BASE><MACRON>";"<CAPITAL><ACCENT>";<U0100> % LATIN CAPITAL LETTER A WITH MACRON
<U00E6> "<a><e>";"<LIGATED><LIGATED>";"<SMALL><SMALL>";<U00E6> % LATIN SMALL LETTER AE
<U00C6> "<a><e>";"<LIGATED><LIGATED>";"<CAPITAL><CAPITAL>";<U00C6> % LATIN CAPITAL LETTER AE
<U01FD> "<a><e>";"<LIGATED><LIGATED><ACUTE>";"<SMALL><SMALL><ACCENT>";<U01FD> % LATIN SMALL LETTER AE WITH ACUTE
<U01FC> "<a><e>";"<LIGATED><LIGATED><ACUTE>";"<CAPITAL><CAPITAL><ACCENT>";<U01FC> % LATIN CAPITAL LETTER AE WITH ACUTE
<U01E3> "<a><e>";"<LIGATED><LIGATED><MACRON>";"<SMALL><SMALL><ACCENT>";<U01E3> % LATIN SMALL LETTER AE WITH MACRON
<U01E2> "<a><e>";"<LIGATED><LIGATED><MACRON>";"<CAPITAL><CAPITAL><ACCENT>";<U01E2> % LATIN CAPITAL LETTER AE WITH MACRON

<U0062> <b>;<BASE>;<SMALL>;<U0062> % LATIN SMALL LETTER B
<U0042> <b>;<BASE>;<CAPITAL>;<U0042> % LATIN CAPITAL LETTER B
<U1E03> <b>;"<BASE><DOT_ABOVE>";"<SMALL><ACCENT>";<U1E03> % LATIN SMALL LETTER B WITH DOT ABOVE
<U1E02> <b>;"<BASE><DOT_ABOVE>";"<CAPITAL><ACCENT>";<U1E02> % LATIN CAPITAL LETTER B WITH DOT ABOVE

<U0063> <c>;<BASE>;<SMALL>;<U0063> % LATIN SMALL LETTER C
<U0043> <c>;<BASE>;<CAPITAL>;<U0043> % LATIN CAPITAL LETTER C

<U0107> <c>;"<BASE><ACUTE>";"<SMALL><ACCENT>";<U0107> % LATIN SMALL LETTER C WITH ACUTE
<U0106> <c>;"<BASE><ACUTE>";"<CAPITAL><ACCENT>";<U0106> % LATIN CAPITAL LETTER C WITH ACUTE
<U0109> <c>;"<BASE><CIRCUMFLEX>";"<SMALL><ACCENT>";<U0109> % LATIN SMALL LETTER C WITH CIRCUMFLEX
<U0108> <c>;"<BASE><CIRCUMFLEX>";"<CAPITAL><ACCENT>";<U0108> % LATIN CAPITAL LETTER C WITH CIRCUMFLEX
<U010D> <c>;"<BASE><CARON>";"<SMALL><ACCENT>";<U010D> % LATIN SMALL LETTER C WITH CARON
<U010C> <c>;"<BASE><CARON>";"<CAPITAL><ACCENT>";<U010C> % LATIN CAPITAL LETTER C WITH CARON
<U010B> <c>;"<BASE><DOT_ABOVE>";"<SMALL><ACCENT>";<U010B> % LATIN SMALL LETTER C WITH DOT ABOVE
<U010A> <c>;"<BASE><DOT_ABOVE>";"<CAPITAL><ACCENT>";<U010A> % LATIN CAPITAL LETTER C WITH DOT ABOVE
<U00E7> <c>;"<BASE><CEDILLA>";"<SMALL><ACCENT>";<U00E7> % LATIN SMALL LETTER C WITH CEDILLA
<U00C7> <c>;"<BASE><CEDILLA>";"<CAPITAL><ACCENT>";<U00C7> % LATIN CAPITAL LETTER C WITH CEDILLA
<U2105> "<c><o>";"<BASE><BASE>";"<COMPAT><COMPAT>";<U2105> % CARE OF

<U0064> <d>;<BASE>;<SMALL>;<U0064> % LATIN SMALL LETTER D
<U0044> <d>;<BASE>;<CAPITAL>;<U0044> % LATIN CAPITAL LETTER D
<U010F> <d>;"<BASE><CARON>";"<SMALL><ACCENT>";<U010F> % LATIN SMALL LETTER D WITH CARON
<U010E> <d>;"<BASE><CARON>";"<CAPITAL><ACCENT>";<U010E> % LATIN CAPITAL LETTER D WITH CARON
<U1E0B> <d>;"<BASE><DOT_ABOVE>";"<SMALL><ACCENT>";<U1E0B> % LATIN SMALL LETTER D WITH DOT ABOVE
<U1E0A> <d>;"<BASE><DOT_ABOVE>";"<CAPITAL><ACCENT>";<U1E0A> % LATIN CAPITAL LETTER D WITH DOT ABOVE
<U0111> <d>;"<BASE><STROKE>";"<SMALL><ACCENT>";<U0111> % LATIN SMALL LETTER D WITH STROKE
<U0110> <d>;"<BASE><STROKE>";"<CAPITAL><ACCENT>";<U0110> % LATIN CAPITAL LETTER D WITH STROKE
<U00F0> <d>;"<BASE><MODIFIED>";"<SMALL><ACCENT>";<U00F0> % LATIN SMALL LETTER ETH
<U00D0> <d>;"<BASE><MODIFIED>";"<CAPITAL><ACCENT>";<U00D0> % LATIN CAPITAL LETTER ETH

<U0065> <e>;<BASE>;<SMALL>;<U0065> % LATIN SMALL LETTER E
<U0045> <e>;<BASE>;<CAPITAL>;<U0045> % LATIN CAPITAL LETTER E
<U00E9> <e>;"<BASE><ACUTE>";"<SMALL><ACCENT>";<U00E9> % LATIN SMALL LETTER E WITH ACUTE
<U00C9> <e>;"<BASE><ACUTE>";"<CAPITAL><ACCENT>";<U00C9> % LATIN CAPITAL LETTER E WITH ACUTE
<U00E8> <e>;"<BASE><GRAVE>";"<SMALL><ACCENT>";<U00E8> % LATIN SMALL LETTER E WITH GRAVE
<U00C8> <e>;"<BASE><GRAVE>";"<CAPITAL><ACCENT>";<U00C8> % LATIN CAPITAL LETTER E WITH GRAVE
<U0115> <e>;"<BASE><BREVE>";"<SMALL><ACCENT>";<U0115> % LATIN SMALL LETTER E WITH BREVE
<U0114> <e>;"<BASE><BREVE>";"<CAPITAL><ACCENT>";<U0114> % LATIN CAPITAL LETTER E WITH BREVE
<U00EA> <e>;"<BASE><CIRCUMFLEX>";"<SMALL><ACCENT>";<U00EA> % LATIN SMALL LETTER E WITH CIRCUMFLEX
<U00CA> <e>;"<BASE><CIRCUMFLEX>";"<CAPITAL><ACCENT>";<U00CA> % LATIN CAPITAL LETTER E WITH CIRCUMFLEX
<U011B> <e>;"<BASE><CARON>";"<SMALL><ACCENT>";<U011B> % LATIN SMALL LETTER E WITH CARON
<U011A> <e>;"<BASE><CARON>";"<CAPITAL><ACCENT>";<U011A> % LATIN CAPITAL LETTER E WITH CARON
<U00EB> <e>;"<BASE><DIAERESIS>";"<SMALL><ACCENT>";<U00EB> % LATIN SMALL LETTER E WITH DIAERESIS
<U00CB> <e>;"<BASE><DIAERESIS>";"<CAPITAL><ACCENT>";<U00CB> % LATIN CAPITAL LETTER E WITH DIAERESIS
<U0119> <e>;"<BASE><OGONEK>";"<SMALL><ACCENT>";<U0119> % LATIN SMALL LETTER E WITH OGONEK
<U0118> <e>;"<BASE><OGONEK>";"<CAPITAL><ACCENT>";<U0118> % LATIN CAPITAL LETTER E WITH OGONEK