



CENTRO DE LINGUÍSTICA DA UNIVERSIDADE NOVA DE LISBOA
Linha de Investigação 4 - Linguística Histórica
<http://www.fcsh.unl.pt/clunl/linha4.html>

THE PORTUGUESE MEDIEVAL FONT PROJECT AND THE MEDIEVAL UNICODE FONT INITIATIVE

António Emiliano, Universidade Nova de Lisboa, ah.emiliano@fcsh.unl.pt
Susana Pedro, Universidade Lusófona (Lisboa), susana.t.pedro@netc.pt

The Portuguese Medieval Font (PMF) was created for making narrow paleographic transcriptions of Latin-Portuguese and Portuguese documents from the 9th century to the mid-14th century. During that time span the dominant scripts were the Visigothic script (from the 9th to the 12th century) and the Carolingian script (from the 2nd half of the 11th century onwards), with several varieties thereof. Much that can be said about Portugal regarding those scripts applies also to the several regions (realms) of Spain, with the exception of Catalonia (where the Carolingian script was introduced much earlier than in the rest of medieval Spain).

The PMF is a 8-bit font with 256 character slots. The outlook of the font is deliberately neutral (monospaced and sans serif), and no efforts were made to imitate medieval writing in any way. The characters were designed from scratch and no scans of actual manuscripts were used.

The font contains several series of characters that the authors consider to be representative, meaningful and contrastive (in the structuralist sense of the word) in terms of the typographic representation of Medieval Portuguese manuscripts. The glyphs that were included in the font should be regarded as "typical forms". Many graphic aspects were deliberately left out, such as ligatures and modified lettershapes arising from the cursive nature of some of the scripts. A Unicode version of the font will allow the inclusion of more characters (especially ligatures) and a broadened time interval, with the inclusion of special characters from varieties of the gothic script.

The design of the font was done by a digital typographer (from a Portuguese webdesign company that was hired for this task), and the scientific coordination of the project was done by António Emiliano, a linguist, and Susana Pedro, a paleographer. The Portuguese Medieval Font Project is one of the ongoing projects of Sub-unit 4 (Historical Linguistics) of the Centro de Linguística da Universidade Nova de Lisboa, a research centre funded by the Portuguese government.

Character series and names

We propose that the more basic letter-characters be classified under the following headings, which refer to the different alphabets (i.e. different versions of the Roman alphabet) that coexisted in Medieval writing:

- ❖ Capital letters
- ❖ Small letters
- ❖ Small capital letters
- ❖ Uncial letters
- ❖ Small uncial letters
- ❖ Enlarged small letters
- ❖ Combining (supralinear) small letters
- ❖ Ligatures

Following Unicode practice the name for each letter-character will be formed by adding the word "LATIN" at the beginning, and adding the special name of each character at the end; in the last but one series on the list the word "combining" must precede the word "LATIN". The general terms "CAPITAL" and "SMALL", which signify the two major letter series, immediately follow the label "LATIN", whereas the more specific terms for special letter styles, such as "UNCIAL" or "ENLARGED SMALL", will follow the word "LETTER". This last convention is not very accurate paleographically, because the uncial letters were originally majuscules which co-existed for centuries with the capitals, and were not regionally located; i.e they were not a regionally circumscribed Latin script and coexisted in several regions with both the capital and the minuscule alphabets, so they should be treated as a major series.

Examples:

M LATIN CAPITAL LETTER M

m LATIN SMALL LETTER M

m LATIN **SMALL** LETTER M WITH LONG RIGHT LEG

ᵹ LATIN LETTER **UNCIAL** M

ᵹ LATIN LETTER **SMALL UNCIAL** M

ᵹ LATIN LETTER **ENLARGED SMALL** M

m
 COMBINING LATIN **SMALL** LETTER M

ᵹ
 COMBINING LATIN **LETTER UNCIAL** M

The so-called national scripts, like Carolingian, Visigothic, Lombardic, Beneventan, Insular, etc., should be regarded as local variations of basically the same minuscule alphabet. Thus, if a Medieval Unicode character set requires that a special lettershape from a specific script be recognised (i.e.encoded) – as is the case of several Visigothic and Insular script letter-shapes – the label for the script should follow the word "letter".

Examples:

E LATIN **CAPITAL** LETTER E

- € LATIN LETTER **UNCIAL** E
- e LATIN **SMALL** LETTER **CAROLINGIAN** E (or perhaps better LATIN **SMALL** LETTER E, and also LATIN **SMALL** LETTER E WITH LONG SHAFT)
- Ɔ LATIN LETTER **ENLARGED SMALL VISIGOTHIC** E (or perhaps better LATIN CAPITAL LETTER OPEN E)
- ɛ LATIN SMALL LETTER **VISIGOTHIC** E (or perhaps better LATIN SMALL LETTER OPEN E)

This is just a first (and tentative) approach to this subject, because we feel that regional or national labels should be avoided whenever possible, and general descriptive labels should be used instead. That is to say that the MUFI proposal should be based on a comparative analysis of the main Medieval scripts, to make sure that there are no unnecessary character repetitions (some of the pre-Carolingian scripts had very similar characteres, and these should receive a general name, instead of being associated with a particular region, country or script).

In the character tables below we use the labels Carolingian and Visigothic in some cases, although we are ready to concede that this is not an optimal solution, and that this is a matter that warrants further discussion.

The MUFI subranges

The Medieval Unicode Font Initiative comprises 12 groupings of characters, each assigned to a Unicode character subrange. We agree with the proposed subranges, but we feel that the names of the character groups must be more accurate paleographically.

The MUFI proposal is as follows (quoted from the website ¹):

¹ <http://www.hit.uib.no/mufi/proposal/proposal-v1.html>

MUFI - Proposed subranges

No.	Name of range	Inventory	Allocated span
1	Mixed script characters	19	E000 - E0FF
2	Precomposed diacritical characters	183	E100 - E1FF
3	Small capitals	19	E200 - E2FF
4	Enlarged minuscules	28	E300 - E3FF
5	Ligatures	15	E400 - E4FF
6	Punctuation marks	4	E500 - E5FF
7	Base line abbreviation marks	15	E600 - E6FF
8	Combining abbreviation marks	11	E700 - E7FF
9	Precomposed abbreviated characters	8	E800 - E8FF
10	Superscript (interlinear) characters	22	E900 - E9FF
11	Metrical symbols	12	EA00 - EAFF
12	Critical and epigraphical signs	4	EB00 - EBFF

Total number of characters included in this proposal 340

We feel that the names of some of the subranges could be improved, and that the contents of one (#08) should be altered.

The use of the words "sign" and "mark" could perhaps be made more consistent. "Sign" could refer to lettersize characters on the base line and to superscript spacing abbreviation characters, whereas "mark" could refer to combining characters and to punctuation characters.

Subrange 08 should include both spacing and combining diacritics (supralinear and infralinear).

Subrange 10 (superscript letters) could perhaps come after subrange 04 ("Enlarged minuscules").

In the table below we give the proposed changes to the names of the MUFI subranges:

No.	Name of MUFI range	change to
07	Base line abbreviation marks	Base line abbreviation signs
08	Combining abbreviation marks	Diacritics (combining and spacing)
09	Precomposed abbreviated	Precomposed abbreviational characters

	characters	
10	Superscript (interlinear) characters	Combining superscript letters [renumber as 05?]

Regarding the names of the individual medieval characteres, we also feel that they can be in some cases more paleographically accurate, and so we propose a few changes to the MIFI proposal (which we give in the character tables below).

The character tables that follow below present the several series of characters contained in the PMF. In some cases a note explains briefly (with examples when needed) the use of a specific character.

A few general remarks are in order regarding the names of characters.

- (1) Character names, especially names of letter-characters, should be descriptive and not functional, whenever possible. They should refer to the shape of the character, or to a specific trait (compared to a more basic letter-shape), rather than to its value, use or distribution. Examples:

m LATIN SMALL LETTER **M WITH LONG RIGHT LEG** (&mlrleg;)

is better than

m LATIN SMALL LETTER **FINAL M** (&mfin;)

- (2) Abbreviations resulting from modified letters or the addition of elements (strokes, flourishes, etc) to letters, should be described as letter-characters rather than as abbreviation-characters. Thus:

d LATIN SMALL LETTER **D WITH STROKE** (&dstr;)

is better than

đ LATIN **ABBREVIATION SIGN "DE"** (&de;)

and

þ LATIN SMALL LETTER **P WITH STROKE** (&pstr;)

is better than

þ LATIN **ABBREVIATION SIGN "PER"** (&per;)

(this abbreviation sign also had the values "par" and "pre" in Medieval Portuguese)

(3) Generic names for characters should be used whenever possible, instead of names containing a reference to a specific script or style:

ſ LATIN SMALL LIGATURE "IT" (&itlig;)

is better than

ſ LATIN SMALL LIGATURE **VISIGOTHIC** "IT" (&itvislig;)

and

Ʒ LATIN SMALL LETTER **LONG Z** (&zlong;)

is better than

3 LATIN SMALL LETTER **CAROLINGIAN** Z (&zcarol;)

- (4) If the Unicode standard already contains a character that corresponds without ambiguity to a special medieval character (even if it is originally a mathematical or a phonetic character or whatever), there should be no duplication in the MIFI proposal. Examples:

ſ



UNICODE COMBINING VERTICAL TILDE (U+033E)

vs. MIFI COMBINING SUPRALINEAR ABBREVIATION MARK "ER/RE" (M+E700)²

≈

UNICODE ALMOST EQUAL TO (U+2248)

vs. MIFI LATIN ABBREVIATION SIGN "ESSE" (M+E600)

÷

UNICODE DIVISION SIGN (U+00F7)

vs. MIFI LATIN ABBREVIATION SIGN "EST" (M+E600)

∴

UNICODE THEREFORE (U+2234)

vs. MIFI PUNCTUATION MARK TRIANGULAR DOTS (M+E500)

² The MIFI character codes are given with the prefix "M+", instead of the Unicode prefix "U+".

PORTUGUESE MEDIEVAL FONT: CHARACTER SET

Contents:

1. Non-letterlike editorial and critical symbols
2. Letters and letterlike characters
 - 2.1 Basic letter characters: capitals and minuscules
 - 2.2 Special letter characters
 - 2.3 Combining superscript letters
3. Diacritics
 - 3.1 Combining non-abbreviational diacritics
 - 3.2 Combining (superscript and subscript) abbreviational diacritics
 - 3.3 Spacing superscript abbreviational diacritics
4. Abbreviation signs on the base line
5. Punctuation signs and additional medieval signs
6. Combining character sequences

Abbreviations:

Med. Ptg. Medieval Portuguese
 n.a. not available
 t.b.s. to be supplied

1. Non-letterlike editorial and critical symbols ³

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY
!	#033	EXCLAMATION MARK	U+0020	

³ For the proposed use for these characteres see http://www.fcsh.unl.pt/clunl/Linha4_Ficheiros/Convencoes_de_transcricao.pdf

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY
←	#034	LEFTWARDS ARROW	U+2190	←
↑	#035	UPWARDS ARROW	U+2191	↑
→	#036	RIGHTWARDS ARROW	U+2192	→
↓	#037	DOWNWARDS ARROW	U+2193	↓
(#040	LEFT PARENTHESIS = OPENING PARENTHESIS	U+0028	
)	#041	RIGHT PARENTHESIS = CLOSING PARENTHESIS	U+0029	
*	#042	ASTERISK	U+002A	
+	#043	PLUS SIGN	U+002B	
-	#045	HYPHEN-MINUS	U+002D	

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY
/	#047	SOLIDUS = SLASH	U+002F	&slash;
0	#048	DIGIT ZERO	U+0030	
1	#049	DIGIT ONE	U+0031	
2	#050	DIGIT TWO	U+0032	
3	#051	DIGIT THREE	U+0033	
4	#052	DIGIT FOUR	U+0034	
5	#053	DIGIT FIVE	U+0035	
6	#054	DIGIT SIX	U+0036	
7	#055	DIGIT SEVEN	U+0037	

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY
8	#056	DIGIT EIGHT	U+0038	
9	#057	DIGIT NINE	U+0039	
<	#060	LESS-THAN SIGN	U+003C	<
=	#061	EQUALS SIGN	U+003D	
>	#062	GREATER-THAN SIGN	U+003E	>
?	#063	QUESTION MARK	U+003F	
□	#064	WHITE SQUARE	U+25A1	
[#091	LEFT SQUARE BRACKET = OPENING SQUARE BRACKET	U+005B	
\	#092	REVERSE SOLIDUS = BACKSLASH	U+005C	

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY
]	#093	RIGHT SQUARE BRACKET = CLOSING SQUARE BRACKET	U+005D	
<	#094	LEFT-POINTING ANGLE BRACKET LEFT ANGLE BRACKET MATHEMATICAL LEFT ANGLE BRACKET	U+2329 U+3008 U+27E8	⟨
—	#095	LOW LINE = SPACING UNDERSCORE	U+005F	&underscore;
>	#096	RIGHT-POINTING ANGLE BRACKET RIGHT ANGLE BRACKET MATHEMATICAL RIGHT ANGLE BRACKET	U+2324 U+3009 U+27E9	⟩
	#124	VERTICAL LINE = VERTICAL BAR	U+007C	

2. Letters and letterlike characters

2.1 Basic letter characters: capitals and minuscules

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY
A	#065	LATIN CAPITAL LETTER A	U+0041	
B	#066	LATIN CAPITAL LETTER B	U+0042	

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY
C	#067	LATIN CAPITAL LETTER C	U+0043	
D	#068	LATIN CAPITAL LETTER D	U+0044	
E	#069	LATIN CAPITAL LETTER E	U+0045	
F	#070	LATIN CAPITAL LETTER F	U+0046	
G	#071	LATIN CAPITAL LETTER G	U+0047	
h	#072	LATIN LETTER UNCIAL H	n.a.	&hunc;
I	#073	LATIN CAPITAL LETTER I	U+0049	
J	#074	LATIN CAPITAL LETTER J	U+004A	
k	#075	LATIN CAPITAL LETTER K WITH LONG ASCENDER	n.a.	&Kasc;

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY
L	#076	LATIN CAPITAL LETTER L	U+004C	
M	#077	LATIN CAPITAL LETTER M	U+004D	
N	#078	LATIN CAPITAL LETTER N	U+004E	
O	#079	LATIN CAPITAL LETTER O	U+004F	
P	#080	LATIN CAPITAL LETTER P	U+0050	
Q	#081	LATIN CAPITAL LETTER Q	U+0051	
R	#082	LATIN CAPITAL LETTER R	U+0052	
S	#083	LATIN CAPITAL LETTER S	U+0053	
T	#084	LATIN CAPITAL LETTER T	U+0054	

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY
U	#085	LATIN LETTER ENLARGED SMALL U	M+E300	&uenl;
V	#086	LATIN CAPITAL LETTER V	U+0056	
X	#088	LATIN CAPITAL LETTER X	U+0058	
Y	#089	LATIN CAPITAL LETTER Y	U+0059	
Z	#090	LATIN CAPITAL LETTER Z	U+005A	
a	#097	LATIN SMALL LETTER A	U+0061	
b	#098	LATIN SMALL LETTER B	U+0062	
c	#099	LATIN SMALL LETTER C	U+0063	
ð	#100	LATIN SMALL LETTER SLANTED D	n.a.	

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY
e	#101	LATIN SMALL LETTER E WITH LONG SHAFT or LATIN SMALL LETTER E or LATIN SMALL LETTER CAROLINGIAN E	n.a. U+0065 n.a.	&elshft; &ecarol;
f	#102	LATIN SMALL LETTER F	U+0066	
g	#103	LATIN SMALL LETTER G	U+0067	
h	#104	LATIN SMALL LETTER H	U+0068	
i	#105	LATIN SMALL LETTER DOTLESS I	U+0131	&idotless;
j	#106	LATIN SMALL LETTER DOTLESS J	n.a.	&jdotless;
k	#107	LATIN SMALL LETTER K	U+006B	
l	#108	LATIN SMALL LETTER L	U+006C	
m	#109	LATIN SMALL LETTER M	U+006D	

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY
n	#110	LATIN SMALL LETTER N	U+006E	
o	#111	LATIN SMALL LETTER O	U+006F	
p	#112	LATIN SMALL LETTER P	U+0070	
q	#113	LATIN SMALL LETTER Q	U+0071	
r	#114	LATIN SMALL LETTER R	U+0072	
ŕ	#115	LATIN SMALL LETTER LONG S	U+017F	
t	#116	LATIN SMALL LETTER T	U+0074	
u	#117	LATIN SMALL LETTER U	U+0075	
v	#118	LATIN SMALL LETTER V	U+0076	

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY
x	#120	LATIN SMALL LETTER X	U+0078	
y	#121	LATIN SMALL LETTER Y	U+0079	
Ʒ	#122	LATIN SMALL LETTER LONG Z or LATIN SMALL LETTER CAROLINGIAN Z	n.a.	&zlong; &zcarol;

2.2 Special letter characters

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY	NOTES
ǎ	#130	LATIN LETTER ENLARGED SMALL A	M+E300	&aenl;	
ɑ	#131	LATIN SMALL LETTER ALPHA = LATIN SMALL LETTER SCRIPT A	U+0251	𝒶	
Ɽ	#132	LATIN LETTER ENLARGED SMALL TALL OPEN A	n.a.	&aenlop;	special lettershape of the Visigothic script
ⱥ	#133	LATIN SMALL LETTER TALL OPEN A	n.a.	&atallop;	special lettershape of the Visigothic script

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY	NOTES
Ꝑ	#134	LATIN SMALL LETTER OPEN SCRIPT A	n.a.	&ascrop;	special lettershape of the Visigothic script
Ð	#135	LATIN CAPITAL LETTER D WITH BAR or LATIN ABBREVIATION SIGN CAPITAL "DE"	n.a.	&Dbar; &De;	= D(e)
Ɔ	#136	LATIN LETTER UNCIAL D	M+E000	&dunc;	
ḏ	#137	LATIN SMALL LETTER D	U+0064		
ḏ̄	#138	LATIN SMALL LETTER D WITH STROKE or LATIN ABBREVIATION SIGN "DE"	U+0111	&dstr; &de;	= d + general abbreviation mark = d(e)
ð	#139	LATIN SMALL LETTER ETH	U+00F0	ð	= d + general abbreviation mark = d(e) = numeral d
Ǝ	#140	LATIN LETTER UNCIAL E	M+E000	&eunc;	
Ǝ̄	#141	LATIN LETTER ENLARGED SMALL E	M+E30.	&eenl;	
Ɔ̄	#142	LATIN CAPITAL LETTER OPEN E or LATIN LETTER ENLARGED SMALL VISIGOTHIC E	U+0190	&Eopen; &evisenl;	special lettershape of the Visigothic script

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY	NOTES
Ɔ	#143	LATIN SMALL LETTER OPEN E = LATIN SMALL LETTER EPSILON or LATIN SMALL LETTER VISIGOTHIC E	U+025B	&eopen; &evis;	special lettershape of the Visigothic script
Ɔ	#144	LATIN SMALL LIGATURE "FI"	n.a.	ﬁ	= fi ligature of the Visigothic script
Ɔ	#145	LATIN LETTER UNCIAL G	n.a.	&gunc;	
Ɔ	#146	LATIN SMALL LETTER VISIGOTHIC G	n.a.	&gvis;	special lettershape of the Visigothic script
Ɔ	#147	LATIN SMALL LETTER TALL I	n.a.	&itall;	special lettershape of the Visigothic script
Ɔ	#148	LATIN SMALL LIGATURE "IT"	n.a.	&itlig;	= -it ligature of the Visigothic script
Ɔ	#149	LATIN CAPITAL LETTER J WITH STROKE	n.a.	&Jstr;	= I(n)
Ɔ	#150	LATIN SMALL LETTER K WITH STROKE THROUGH LEG	n.a.	&kstrleg;	= karta, kartam
Ɔ	#151	LATIN LETTER UNCIAL L	n.a.	&lunc;	

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY	NOTES
ℓ	#190	LATIN LETTER SMALL UNCIAL L	n.a.	&lsunc;	
ᵐ	#152	LATIN LETTER ENLARGED SMALL M	M+E300	&menl;	
ᵑ	#153	LATIN LETTER UNCIAL M	M+E000	&munc;	
ᵓ	#154	LATIN LETTER SMALL UNCIAL M	n.a.	&msunc;	
ᵔ	#155	LATIN SMALL LETTER M WITH LONG RIGHT LEG or LATIN SMALL LETTER FINAL M	n.a.	&mlrleg; &mfin;	
ᵕ	#156	LATIN SMALL LETTER M WITH VERTICAL BAR or LATIN ABBREVIATION SIGN "MUM"	n.a.	&mbar; &mum;	= -m(um)
ᵎ	#157	LATIN LETTER ENLARGED SMALL N	M+E300	&nenl;	
ᵏ	#158	LATIN SMALL LETTER N WITH LONG RIGHT LEG or LATIN SMALL LETTER FINAL N	U+019E	&nlrleg; &nfin;	
ᵌ	#159	LATIN SMALL LETTER N WITH VERTICAL BAR or LATIN ABBREVIATION SIGN "NUM"	n.a.	&nbar; #	= -n(um)

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY	NOTES
Ɔ	#160	LATIN CAPITAL LETTER P WITH STROKE	M+E800	&Pstr;	= P(er) = P(ar) = P(re)
Ɔ	#161	LATIN CAPITAL LETTER P WITH FLOURISH	M+E800	&Pflour;	= P(ro) = P(or)
Ɔ	#162	LATIN SMALL LETTER P WITH STROKE	M+E800	&pstr;	= p(er) = p(ar) = p(re)
Ɔ	#163	LATIN SMALL LETTER P WITH FLOURISH	M+E800	&pflour;	= p(ro) = p(or)
Ɔ	#164	LATIN SMALL LETTER Q WITH STROKE	M+E800	&qstr;	= q(ue)
Ɔ	#165	LATIN SMALL LETTER Q WITH SLANTED STROKE	n.a.	&qslstr;	= q(ui) = q(ue)
Ɔ	#166	LATIN SMALL LETTER Q WITH 2-SHAPED APPENDIX or LATIN ABBREVIATION SIGN "QUIA"	n.a.	&q2; &quia;	= quia
R	#191	LATIN LETTER SMALL CAPITAL R	U+0280	&Rscap;	
2	#167	LATIN SMALL LETTER ROUND R	M+E000	&rround;	

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY	NOTES
ꝛ	#168	LATIN SMALL LETTER ROUND R WITH VERTICAL BAR or LATIN ABBREVIATION SIGN ROUND "RUM"	M+E600	&rumrnd; &rndvbar;	= -r(um)
ꝛ	#169	LATIN LETTER ENLARGED SMALL R WITH LONG LEG	n.a.	&renlleg;	special lettershape of the Visigothic script
ꝛ	#170	LATIN SMALL LETTER R WITH LONG LEG	U+027C	&rlleg;	
ꝛ	#171	LATIN SMALL LETTER R WITH VERTICAL BAR or LATIN ABBREVIATION SIGN "RUM"	n.a.	&rbar; &rum;	= -r(um) special lettershape of the Visigothic script
s	#172	LATIN SMALL LETTER S	U+0073		
ſ	#173	LATIN SMALL LETTER S WITH DESCENDER or LATIN SMALL LETTER FINAL S	n.a.	&sdesc; &sfin;	
ꝛ	#174	LATIN SMALL LETTER LONG S WITH LONG LEG	M+E00.	&slngll;	
ꝛ	#175	LATIN SMALL LETTER LONG S WITH WITH STROKE	M+E60.	&slngstr;	= scilicet = Med. Ptg. seer
ꝛ	#176	LATIN LETTER UNCIAL T	M+E00.	&tunc;	






CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY	NOTES
Ƨ	#177	LATIN SMALL LETTER VISIGOTHIC T	n.a.	&tvis;	special lettershape of the Visigothic script
Ƨ̄	#178	LATIN SMALL LETTER VISIGOTHIC T WITH VERTICAL BAR or LATIN ABBREVIATION SIGN VISIGOTHIC "TUM"	n.a.	&tvisbar; &tumvis;	= -t(um) special lettershape of the Visigothic script
Ƨ̅	#179	LATIN SMALL LETTER T WITH VERTICAL BAR or LATIN ABBREVIATION SIGN "TUM"	n.a.	&tbar; &tum;	= -t(um)
Ƨ̅	#180	LATIN SMALL LETTER VISIGOTHIC TALL T or LATIN SMALL LETTER VISIGOTHIC FINAL T	n.a.	&tvistall; &tvisfin;	
Ƨ̅	#181	LATIN SMALL LETTER LONG V	n.a.	&vlong;	
Ƨ̅	#182	LATIN SMALL LETTER VISIGOTHIC X	n.a.	&xvis;	special lettershape of the Visigothic script
XL	#183	LATIN CAPITAL LIGATURE "XL"	n.a.	&XLlig;	= numeral XL
x̅	#184	LATIN SMALL LIGATURE "XL"	n.a.	&xllig;	= numeral xl
Z	#185	LATIN SMALL LETTER Z	U+007A		

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY	NOTES
Ʒ	#186	LATIN SMALL LETTER VISIGOTHIC Z	n.a.	&zvis;	special lettershape of the Visigothic script

2.3 Combining superscript letters

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY	NOTES
ⱱ □	#192	COMBINING LATIN SMALL LETTER OPEN A	n.a.	&aopsup;	
Ⱳ □	#193	COMBINING LATIN SMALL LETTER A	U+0363	&asup;	
ⱳ □	#194	COMBINING LATIN SMALL LETTER ALPHA	n.a.	&alphasup;	
ⱴ □	#195	COMBINING LATIN SMALL LETTER B	M+E900	&bsup;	
Ⱶ □	#196	COMBINING LATIN SMALL LETTER C	U+0368	⫐	
ⱶ □	#197	COMBINING LATIN SMALL LETTER D	U+0369	&dsup;	

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY	NOTES
e ◻	#198	COMBINING LATIN SMALL LETTER E	U+0364	&esup;	
i ◻	#199	COMBINING LATIN SMALL LETTER I	U+0365	&isup;	
l ◻	#200	COMBINING LATIN SMALL LETTER L	M+E900	&lsup;	
ꝛ ◻	#201	COMBINING LATIN LETTER UNCIAL M	n.a.	&luncsup;	
ꝛ ◻	#203	COMBINING LATIN SMALL LETTER M	U+036B	&msup;	
ꝛ ◻	#204	COMBINING LATIN SMALL LETTER M WITH LONG RIGHT LEG or COMBINING LATIN SMALL LETTER FINAL M	n.a.	&mlrlsup; &mfinsup;	
n ◻	#205	COMBINING LATIN SMALL LETTER N	M+E900	⊅	
o ◻	#206	COMBINING LATIN SMALL LETTER O	U+0366	&osup;	
r ◻	#207	COMBINING LATIN SMALL LETTER R	U+036C	&rsup;	

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY	NOTES
$\overset{2}{r}$ 	#208	COMBINING LATIN SMALL LETTER ROUND R or COMBINING SUPRALINEAR ABBREVIATION MARK "UR"	M+E700	&ur; &rroundsup;	= r = ur, or/ro $\overset{o}{\overset{2}{.}}\text{.iiij.}$ = quattuor $\overset{2}{P}oto$ = Porto $\overset{2}{p}caçam$ = proc(ur)açam $\overset{2}{p}curadores$ = p(ro)curadores $\overset{2}{p}$ = p(or)
$\overset{s}{s}$ 	#209	COMBINING LATIN SMALL LETTER S	M+E900	&ssup;	
	#210	COMBINING LATIN SMALL LETTER S WITH DESCENDER or COMBINING LATIN SMALL LETTER FINAL S	n.a.	&sdescsup; &sfinsup;	= s = ue = us $un\overset{s}{\delta}$ = unos $\overset{s}{q}$ = q(ue) $hered\overset{s}{i}\overset{b}{\delta}$ = heredib(us) (special abbreviation mark of the Visigothic script)
	#211	COMBINING LATIN SMALL LETTER LONG S	M+E900	&slongsup;	
$\overset{t}{t}$ 	#212	COMBINING LATIN SMALL LETTER T	U+036D	&tsup;	





CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY	NOTES
u □	#213	COMBINING LATIN SMALL LETTER U	U+0367	&usup;	
v □	#214	COMBINING LATIN SMALL LETTER V	U+036E	&vsup;	

3. Diacritics








3.1 Combining non-abbreviational diacritics









CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY
□ ₃	#187	COMBINING CEDILLA	U+0327	&cmbcedil;
□ ₂	#188	COMBINING OGONEK	U+0328 M+E700	&cmbogon;
□́	#234	COMBINING ACUTE ACCENT	U+0301	&cmbacute;
□ [•]	#235	COMBINING DOT HIGH ABOVE	M+E700	&cmbdothiab;
□ [•]	#236	COMBINING DOT ABOVE	U+0307	&cmbdotab;

3.2 Combining (superscript and subscript) abbreviational diacritics


CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY	NOTES
	#189	COMBINING SUPRALINEAR ABBREVIATION MARK STROKE LEFT	n.a.	&cmbstrleft;	= general abbreviation mark (combining diacritic for crossing left-positioned ascenders)
	#216	COMBINING SUPRALINEAR ABBREVIATION MARK "US/UE"	M+E700	&cmbusue;	= us = ue = os dam [?] = dam(us), Med. Ptg. dam(os) p [?] t = p(os)t q [?] = q(ue)
	#218	COMBINING SUPRALINEAR ABBREVIATION MARK CONVEX "US/UE"	n.a.	&cmbconvusue;	= us = ue dam ^ξ = dam(us) q ^ξ = q(ue) (special abbreviation mark of the Visigothic script)
	#219	COMBINING ABBREVIATION MARK OVERLINE WITH ONSET	n.a.	&cmbolon;	= us = general abbreviation mark b [←] = -b(us) d [←] = d(e)/-d(us) cā [←] = sancta

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY	NOTES
◌ [?] ◻	#220	COMBINING COMMA or COMBINING COMMA ABOVE RIGHT	U+0313 U+0315	&cmbcom;	= e = er = re = ue ð = d(e) pp [?] r = p(ro)pt(e)r a [?] n [?] i [?] u [?] f [?] a [?] r [?] i [?] o = an(n)iu(er)sario f [?] c [?] u [?] y = sc(re)uy q [?] = q(ue)
◌ [˘] ◻	#221	COMBINING VERTICAL TILDE or COMBINING SUPRALINEAR ABBREVIATION MARK "ER/RE"	U+033E M+E700	&cmbtilver; &cmberre;	= er/re a [˘] u = au(er) p [˘] dicte = p(re)dicte ap [˘] f = ap(re)s
◌ [~] ◻	#222	COMBINING SUPRALINEAR ABBREVIATION MARK UR	n.a.	&cmbur;	= ur dicit [~] = dicit(ur) f [~] c [~] pta = sc(r)ipt(ur)a
◌ [—] ◻	#223	COMBINING OVERLINE or COMBINING ABBREVIATION MARK OVERLINE	U+0305 M+E700	&cmbol;	= general abbreviation mark f [—] co = sancto f [—] co = sancto
◌ [—] ◻	#224	COMBINING MACRON or COMBINING SUPRALINEAR ABBREVIATION MARK STROKE	U+0304	&cmbmacron; &cmbstr;	= general abbreviation mark f [—] co = sancto

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY	NOTES
	#225	COMBINING SUPRALINEAR ABBREVIATION MARK OVERLINE WITH OFFSET	n.a.	&cmboloff;	= general abbreviation mark ꝛcō = sancto
	#226	COMBINING SUPRALINEAR ABBREVIATION MARK STROKE WITH OFFSET	n.a.	&cmbstroff;	= general abbreviation mark ꝛcō = sancto
	#227	COMBINING SUPRALINEAR ABBREVIATION MARK SHORT STROKE WITH OFFSET	n.a.	&cmbshstroff;	= general abbreviation mark ꝛcō = sancto
	#228	COMBINING SUPRALINEAR ABBREVIATION MARK STROKE WITH ONSET AND OFFSET	n.a.	&cmbstronoff;	= general abbreviation mark ꝛ = cum
	#229	COMBINING SUPRALINEAR ABBREVIATION MARK SHORT STROKE WITH ONSET AND OFFSET	n.a.	&cmbshstronoff;	= general abbreviation mark = er/re ꝛ = in ꝛmini = t(er)mini ꝛ = p(re)
	#231	COMBINING SUPRALINEAR ABBREVIATION MARK STROKE WITH HUMP	n.a.	&cmbstrhump;	= general abbreviation mark ꝛcō = sancto (special abbreviation mark of the Visigothic script)
	#232	COMBINING SUPRALINEAR ABBREVIATION MARK FLOURISH	n.a.	&cmbflour;	= general abbreviation mark ꝛcō = sancto (special abbreviation mark of the Visigothic script)

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY	NOTES
	#233	COMBINING SUPRALINEAR ABBREVIATION MARK STROKE WITH LOOP	n.a.	&cmbstrloop;	= general abbreviation mark  co = sancto  co = sancto (special abbreviation mark of the Visigothic script)
	#237	COMBINING SUPRALINEAR ABBREVIATION MARK LONG TILDE or COMBINING GREEK PERISPOMENI?	n.a. U+0342?	&cmbtilde;	= general abbreviation mark = -m  cō = sancto  tātū = tantu(m)
	#242	COMBINING INFRALINEAR ABBREVIATION MARK STROKE or COMBINING INFRALINEAR ABBREVIATION MARK "IS"	n.a.	&cmbinfstr; &cmbis;	= is  nob ₂ = nob(is) (special abbreviation mark of the Visigothic script)

3.3 Spacing superscript abbreviational diacritics

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY	NOTES
	#039	APOSTROPHE = APOSTROPHE-QUOTE or MODIFIER LETTER APOSTROPHE	U+2019 U+02BC		= e d' = d(e) u ' = u(e)l

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY	NOTES
9̇	#215	LATIN ABBREVIATION SUPRALINEAR SIGN "US/UE"	n.a.	&usue;	= us = ue ḍam ^{9̇} = dam(us), Med.Ptg. dam(os) q ^{9̇} = q(ue)
9̂	#217	LATIN ABBREVIATION SUPRALINEAR SIGN CONVEX "US/UE"	n.a.	&convusue;	= us = ue ḍum ^{9̂} = dam(us) q ^{9̂} = q(ue) (special abbreviation sign of the Visigothic script)

4. Abbreviation signs on the base line

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY	NOTES
&	#038	AMPERSAND	U+0026	&	

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY	NOTES
;	#059	SEMICOLON or LATIN ABBREVIATION SIGN "ED/US/UE"	U+003B M+E600	&edusueet;	= ed = us = ue = et ſ; = s(ed) t̄sb; = testib(us) q;cūq; = quecumque q̄b;cūq; = quib(us) cumq(ue) ſc̄t̄c; = scilic(et)
9	#238	LATIN ABBREVIATION SIGN "CON"	M+E600	&con;	= con, com
9̇	#239	LATIN ABBREVIATION SIGN "CON" WITH DOT	n.a.	&condot;	= con = com 9̇f̄ = confirmat, confirmans (special abbreviation sign of the Visigothic script)

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY	NOTES
ſ	#240	LATIN ABBREVIATION SIGN "US/UM/UN"	n.a.	&usumun;	= us = um = un dumſ = dam(us) tērminſ = termin(um) mſdum = m(un)dum (special abbreviation sign of the Visigothic script)
·̇	#241	LATIN ABBREVIATION SIGN TILDE WITH DOT ABOVE	n.a.	&tildedot;	= ue q̇ = q(ue)
≈	#243	ALMOST EQUAL TO = ASYMPTOTIC TO or LATIN ABBREVIATION SIGN "ESSE"	U+2248	≈ &esse;	= esse
÷	#244	DIVISION SIGN or LATIN ABBREVIATION SIGN "EST"	U+00F7 M+E600	÷ &est;	= est id÷ = id est
+	#245	LATIN ABBREVIATION SIGN CROSS	M+E600	✗	
7	#246	TIRONIAN SIGN ET or LATIN ABBREVIATION SIGN "ET"	U+204A M+E600	&et;	= et

5. Punctuation signs and additional medieval signs ⁴

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY	NOTES
7	#044	PUNCTUATION MARK COMMA (SIMPLEX DUCTUS)	U+002C		
.	#046	PUNCTUATION MARK FULL STOP (SUBDISTINCTIO)	U+002E		
:	#058	PUNCTUATION MARK COLON	U+003A		= colon = ue q: = q(ue)
¶	#087	PILCROW SIGN = PARAGRAPH SIGN	U+00B6	¶	
Ɔ	#119	PUNCTUATION MARK PARAGRAPHUS SIGN	n.a.	¶g;	= paragraphus, beginning of section (cf. Parkes 1992: 12, 43, 305)
—	#123	HORIZONTAL BAR	U+2015		

⁴ For the meaning and use of medieval punctuation marks see Parkes, M. B. 1992. *Pause and Effect. An introduction to the history of punctuation in the West*, Aldershot, Scolar Press.

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY	NOTES
☐ —	#125	COMBINING LOW LINE	U+0332 M+E700	&cmbllol;	= scribal underline
/	#126	PUNCTUATION MARK END OF LINE	n.a.	&eol;	= end of line, translineation
☐ •	#128	COMBINING DOT BELOW	U+0323		= scribal deletion ιρϝϛ̣οϝ = ips<ud>os
☐ ☐	#129	COMBINING LONG STROKE OVERLAY	U+0336		= scribal deletion
/	#247	FRACTION SLASH or DIVISION SLASH or BIG SOLIDUS	U+2044 U+2215 U+29F8	&fracsl; &divsl; &bigsol;	= virgule
::	#248	PROPORTION or PUNCTUATION MARK FOUR DOTS or QUADRANGULAR DOTS	U+2237 n.a.	&propor; &4dots; &quaddots;	
∴	#249	THEREFORE PUNCTUATION MARK TRIANGULAR DOTS	U+2234 M+E500	∴ &tridots;	
∵	#250	BECAUSE or PUNCTUATION MARK INVERTED TRIANGULAR DOTS	U+2235	∵ &invtridots;	

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY	NOTES
⋮	#251	PUNCTUATION MARK LOSENGE DOTS	n.a.	&losdots;	
∴	#252	PUNCTUATION MARK COLON PLUS POSITURA	n.a.	&colposit;	
⋱	#253	PUNCTUATION MARK RAISED COMMA WITH DOT BELOW	n.a.	&comdot;	cf. Parkes 1992: 301
⋮	#254	PUNCTUATION MARK INVERTED SEMICOLON (PUNCTUS ELEVATUS)	M+E500	&seminv; &punctel;	
7	#255	PUNCTUATION MARK POSITURA	n.a.	&posit;	

6. Combining character sequences [to be completed]

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY	NOTES
ᵇ	#098 + #189	LATIN SMALL LETTER B WITH STROKE	U+0180	&bstr;	= b + general abbreviation mark

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY	NOTES
ḥ	#072 + #189	LATIN LETTER UNCIAL H WITH STROKE	n.a.	&huncstr;	= H + general abbreviation mark = Hec = H(e) = H(er) ḥreditas = H(e)reditas ḥeditate = H(er)editate
h̄	#104 + #189	LATIN SMALL LETTER H WITH STROKE	U+0127	&hstr;	= h + general abbreviation mark = hec = h(e) = h(er) h̄reditas = h(e)reditas h̄editate = h(er)editate
ƀ	#075 + #189	LATIN CAPITAL LETTER K WITH STROKE	n.a.	&Kstr;	= Kalendas = Karta, Kartam
ƀ̄	#107 + #189	LATIN SMALL LETTER K WITH STROKE	n.a.	&kstr;	= kalendas = karta, kartam
ƀ̄	#150 + #189	LATIN SMALL LETTER K WITH STOKES THROUGH THE ASCENDER AND LEG	n.a.	&kstrleg;	= karta, kartam
†	#147 + #223	LATIN SMALL LETTER TALL I WITH STROKE	n.a.	&itallstr;	= I(n)

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY	NOTES
†	#108 + #223	LATIN SMALL LETTER L WITH LONG STROKE THRU THE ASCENDER or LATIN SMALL LETTER L WITH BAR?	n.a. U+019A?	&lstr; &lbar;	= l + general abbreviation mark
Ɔ	#164 + #192	LATIN ABBREVIATION SIGN "QUAM/QUAN" WITH SUPERSCRIPT SMALL OPEN A	n.a.	&opquam;	= quam = quan- ^u q = quam ^u qta = quanta
Ç	#067 + #187	LATIN CAPITAL LETTER C WITH CEDILLA	U+00C7	Ç	
ç	#099 + #187	LATIN SMALL LETTER C WITH CEDILLA	U+00E7	ç	
Ě	#069 + #188	LATIN CAPITAL LETTER E WITH OGONEK	U+0118	Ę	
ě	#101 + #188	LATIN SMALL LETTER CAROLINGIAN E WITH OGONEK	U+0119	&ecarogon;	
Ě	#140 + #188	LATIN LETTER UNCIAL E WITH OGONEK	n.a.	&euncogon;	
Ě	#141 + #188	LATIN LETTER ENLARGED SMALL E WITH OGONEK	n.a.	&eenlogon;	

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY	NOTES
Ɔ	#142 + #188	LATIN LETTER ENLARGED SMALL VISIGOTHIC E WITH OGONEK	n.a.	&evisenlogon;	
Ɔ	#143 + #188	LATIN SMALL LETTER VISIGOTHIC E WITH OGONEK	n.a.	&evisogon;	
ȳ	#121 + #236	LATIN SMALL LETTER Y WITH DOT ABOVE	n.a.	&ydot;	
ĥ	#068 + #236	LATIN SMALL LETTER H WITH DOT ABOVE	n.a.	&hdot;	= hoc
;	#039 + #235	APOSTROPHE WITH COMBINING DOT HIGH ABOVE	n.a.	&apodot;	= us = ue man [̇] = man(us) q [̇] = q(ue)
◌̇	#224 + #236	COMBINING SUPRALINEAR ABBREVIATION MARK MACRON WITH DOT	n.a.	&macdot;	= m]nṙuṗenḋū = Inru(m)pendu(m) (special abbreviation sign of the Visigothic script)
t.b.s.	n.a.	LATIN SMALL LIGATURE VISIGOTHIC "ET"	n.a.	&etvislig;	= & (ligature of the Visigothic script)
t.b.s.	n.a.	LATIN SMALL LIGATURE "BB"	n.a.	&bblig;	= bb (ligature of the Gothic script)
t.b.s.	n.a.	LATIN SMALL LIGATURE "FF"	n.a.	ﬀ	= ff (ligature of the Gothic

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY	NOTES
					script)
t.b.s.	n.a.	LATIN SMALL LIGATURE "PP"	n.a.	&pplig;	= pp (ligature of the Gothic script)
t.b.s.	n.a.	LATIN SMALL LIGATURE "PROP"	n.a.	&proplig;	= p(ro)p (ligature of the Gothic script)
t.b.s.	n.a.	LATIN SMALL LIGATURE "SS"	n.a.	&sslig;	= ss (ligature of the Gothic script)
ā	#097 + #224	LATIN SMALL LETTER A WITH MACRON	U+0101	&amacron;	
ē	#101 + #224	LATIN SMALL LETTER E WITH MACRON	U+0113	&emacron;	
ī	#105 + #224	LATIN SMALL LETTER I WITH MACRON	U+012B	&imacron;	
ĵ	#106 + #224	LATIN SMALL LETTER J WITH MACRON	n.a.	&jmacron;	
ō	#111 + #224	LATIN SMALL LETTER O WITH MACRON	U+014D	&omacron;	
ū	#117 + #224	LATIN SMALL LETTER U WITH MACRON	U+016B	&umacron;	

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY	NOTES
ç	#099 + #224	LATIN SMALL LETTER C WITH MACRON	n.a.	&cmacron;	
ɸ	#109 + #224	LATIN SMALL LETTER M WITH MACRON	n.a.	&mmacron;	
ɹ	#110 + #224	LATIN SMALL LETTER N WITH MACRON	n.a.	&nmacron;	
ɸ	#112 + #224	LATIN SMALL LETTER P WITH MACRON	n.a.	&pmacron;	
ɸ	#113 + #224	LATIN SMALL LETTER Q WITH MACRON	n.a.	&qmacron;	
ɹ	#114 + #224	LATIN SMALL LETTER R WITH MACRON	n.a.	&rmacron;	
ɹ	#170 + #224	LATIN SMALL LETTER LONG R WITH MACRON	n.a.	&rlngmacron;	
ɹ	#115 + #224	LATIN SMALL LETTER LONG S WITH MACRON	n.a.	&slngmacron;	
ɸ	#121 + #224	LATIN SMALL LETTER LONG Y WITH MACRON	U+0233	&ymacron;	

CHARACTER	CHR#	UNICODE NAME	UCS#	ENTITY	NOTES
á	#097 + #234	LATIN SMALL LETTER A WITH ACUTE	U+00E1	á	
é	#101 + #234	LATIN SMALL LETTER E WITH ACUTE	U+00E9	é	
í	#105 + #234	LATIN SMALL LETTER I WITH ACUTE	U+00ED	í	
í	#106 + #234	LATIN SMALL LETTER J WITH ACUTE	n.a.	&jacute;	
ó	#111 + #234	LATIN SMALL LETTER O WITH ACUTE	U+00F3	ó	
ú	#117 + #234	LATIN SMALL LETTER U WITH ACUTE	U+00FA	ú	

etc.