Introduction.

A set of characters used by specialists in medieval European philology and linguistics is absent from the Universal Character Set. These characters differ in nature; some are original ligatures which acquired letter status due to their phonemic value; some are letterforms distinct from other letterforms innovated to distinguish sounds; some are combining diacritical letters used in abbreviations or suspensions of various kinds; and some are best described as “letters with syllabic content”. This is not a complete proposal, but a work in progress; all of the example Figures referred to will be added in the next version of this document. Note that the glyphs used in the code charts here may not be optimal.

Theoretical preliminaries.

Contemporary medievalist philologists and linguists want to be able to represent typographically (in printed format and on computer screens) the character sets which were in use for many centuries in several regions of medieval Europe. Those character sets derive from the common Latin script and contained many characters which simply disappeared with the development of contemporary printing conventions. Early printers made abundant use of “special” medieval characters, but eventually these fell out of use, with notable exceptions like $, ¶, &@, Ç, ˜, and the ¯ used in Ireland.

Contemporary philologists and linguists who want to study the graphemic conventions in use in medieval times—thereby drawing solid or grounded conclusions about the nature and structure of the language systems represented in writing—must rely on bona fide transcriptions of the texts. Bona fide transcriptions are only possible when the elemental character set used in the manuscripts is encoded uniquely and available for use in fonts.

What most philologists did in the 20th century was to publish transliterations, that is, editions which substitute modern characters (or sequences of modern characters) for the original medieval characters. Transliteration-based editions are virtually useless for those scholars who are interested in the study of medieval writing systems, phonology, and even textual structure. Transliterations (or “normalized editions”) and even translations may, of course, be required for editions aimed at students or the general public, but the base texts must result from transcribing the sources: the first step must always be a transcription.

Transcription is an editorial process which does not entail the replacement or the distortion of the original character set. A bona fide transcription of a Runic, Coptic, or Egyptian Hieroglyphic text can be no less than a close rendition of the original characters, using typographic versions of Runic letters, Coptic letters, and Hieroglyphic characters. The same practice must apply to European medieval texts.

The practice of expanding abbreviations common to medievalists throughout the 20th century is not transcription, but simply transliteration: abbreviations, or brachygraphemes, were special characters. (Brachygraphy is, according to the Oxford English Dictionary, “The art or practice of writing with abbreviations or with abbreviated characters; shorthand, stenography.”) Many of these graphemes were
polyvalent—that is, they could be transliterated into different sequences of “normal letters”, according to textual context, country, region, time period, and even individual scribal practices. Polyvalence is not a uncommon feature of alphabetic writing systems, as our own modern spelling systems show; for instance, in European Portuguese orthography the letter E can have the values [e], [e], [i], [i], [i], [i], and Ø; this letter combined with -m or -n can further represent [ê] and [êa]. Graphemic polyvalence results in many instances from language change and from the conservative nature of spelling systems.

Another common practice in the 20th century was to eliminate the original punctuation and to add modern punctuation—many scholars believed that medieval punctuation served no discernable purpose. “Modernization” of the use of capital letters was also standard practice.

A medieval character set makes it possible to shed many “chronocentric” biases and prejudices which tainted many editorial efforts in the 19th and 20th centuries in many countries, rendering the ensuing editions virtually useless for at least some kinds of contemporary research. This does not mean that medievalist scholars wish or even need to represent in print every minutia that handwritten sources present—that is palaeography proper. In medieval texts this is a particularly delicate issue, because scholars have to deal with a considerable amount of regional or individual stylistic variation. The rationale behind encoding medieval characters and designing medieval fonts is not to capture in print every single glyph variation (a task which is virtually impossible and also meaningless), but to capture the character set used in the manuscripts under scrutiny. We understand the character_glyph model and how it applies to the medieval character set.

Accurate transcriptions of medieval texts allow scholars to quote medieval texts without distorting their graphemic content, and allow the texts to be studied by means of computer applications such as concordance generators and wordlist generators. Accurate transcriptions which make use of a medieval character set are a means to preserve—and interchange—all the relevant graphemic, textual, and linguistic information contained in a text; they are also indirectly a means to contribute to the preservation of Europe’s early heritage.

Case-pairing.

Most of the casing pairs shown below are attested in the examples. Those which are not, fall into two categories: those for which no capital can be constructed (such as LONG S) and those for which natural capitals can be easily formed. In an early version of this document we had proposed a single lower-case character <µ> LATIN SMALL LETTER Y WITH STROKE in use by some Welsh medievalists to indicate an epenthetic schwa sound (Figure 8). Subsequently we discovered that this character and its capital are ballotting in FPDAM2, as a character used in the Lubuagan Kalinga language of the Philippines. Because of the general structural feature of the Latin script (from a theoretical point of view), and in order to facilitate modern casing operations for these letters, we have judged it appropriate to supply case-pairs for all the letters which admit of them. In a scholarly publication, for instance, an article title at the top of a journal page might be set in all caps; it would be nonsensical for all but one or two of the medievalist Latin letters to be able to be cased with an all caps command. (This precedent was set with the encoding of the archaic Coptic extensions.)

Discussion.

1. Letters used for medieval Welsh.

<Il> The voiceless lateral fricative written <Il> in modern Welsh may be written in medieval Welsh as a joined ligature of <Il> LATIN SMALL LETTER MIDDLE-WELSH LL. Its capital form <Il> LATIN CAPITAL LETTER MIDDLE-WELSH LL has been used as an abbreviation in Welsh scholarship (Figures 1, 2, 3, 4, 5, 52).

<d> The voiced dental fricative written <dd> in modern Welsh is written <d> LATIN SMALL LETTER INSULAR D distinctly from <d> in some editions of medieval Welsh texts; Nordic medievalists also make use of this letter (Figures 1, 5, 21, 29, 39, 40, 53, 73).

<d> Some Welsh medievalists (and other Indo-Europeans of a certain era) also use <d> LATIN SMALL LETTER SCRIPT D to write this sound in transcription. While this letter may sometimes have
been represented in print by using a DELTA from a Greek lead-type font, it derives from the handwritten Latin d, and behaves like a Latin letter in ordering and is found alongside Greek text proper (Figures 6, 7, 8, 41).

A unique LATIN SMALL LETTER MIDDLE-WELSH V is used distinctly from <u>, <v>, and <w>, though it is true to say that the phonetic value of all four of these letters is polyvalent in medieval Welsh (Figures 3, 9, 10, 11, 12, 13).

Some Welsh medievalists use LATIN SMALL LETTER Y WITH LOOP to indicate the schwa sound of <y> (Figure 14).

As in many medieval traditions, LATIN SMALL LETTER R ROTUNDA is distinguished from LATIN SMALL LETTER R. This named character is derived from a positional variant of following <o> in the South Italian Beneventan, though in medieval Welsh and Nordic it is not limited to this position. In any case, Welsh and Nordic medievalists distinguish R from R ROTUNDA in their printed editions; the letter is also common in early printed texts throughout Europe (Figures 1, 3, 16, 17, 21, 24, 35, 36, 38, 42, 70, 71, 73). The case-pairing LATIN CAPITAL LETTER R ROTUNDA is attested in printed texts from the 15th century (Figure XXX).

2. Letters used for medieval Nordic vowels. Medieval Nordic orthographies innovated a number of letters out of original Latin script ligated letters. Some of these letters are well known today, as the letters Æ æ ÆŒ œ. In medieval Nordic:

LATIN LETTER AA is used for phonemic /a:/ (Figures 18, 19, 23, 28, 34).

LATIN LETTER AO is used for phonemic /ø:/ (Figures 15, 16, 17, 20, 66).

LATIN LETTER AU is used for phonemic /au/, /ɔː/ and /øː/ (Figure 38, 72).

LATIN LETTER AV is used for phonemic /au/, /ɔː/, /øː/, and /øː/ (Figures 18, 21, 24, 42, 70, 72).

LATIN LETTER AV WITH HORIZONTAL BAR is used for phonemic /ɔː/, /øː/, /øː/, and /ɛː:/ (Figure 24).

LATIN LETTER AV is used for phonemic /ɔː/, /øː/, and /ɛː:/ (Figure 69).

LATIN LETTER OO is used for phonemic /oː/, (Figure 28).

LATIN LETTER O WITH LOOP is used for phonemic /ɔː/, /øː/, and /ɛː:/ (Figures 19, 25, 35, 36).

LATIN LETTER VY is used for phonemic /yː:/ (Figure 28). This collection of characters is a superset of the letters found in the medieval Nordic corpus; no single manuscript contains all of them. Note also that none of these characters is a “ligature” that can be broken; indeed, all of them are known to bear diacritical marks.

3. Letters used for medieval Nordic consonants.

LATIN LETTER BROKEN L was used for phonemic /lː/ (Figure 70).

LATIN LETTER VEND (ultimately derived from the English letter WYNN) for phonemic /v/ or /u/. Some editions use u, v, and y in the same text (Figures 21, 26, 35, 36, 42).

The Icelandic First Grammarian’s orthography made use of small capital letters to indicate gemination of consonant sounds, as Uralic linguists did centuries later. Between letters encoded already for Uralicist and IPA use, most of the Latin alphabet is already encoded as small capitals; while SMALL CAPITAL F and SMALL CAPITAL S are yet missing from the UCS (Figure 22). It should be noted that of the traditional Latin alphabet, if these two are added, only *SMALL CAPITAL Q and *SMALL CAPITAL X will remain unencoded.

4. Letters used for medieval Ibero-Romance.

The Latin alphabet we use today is only one of several variants. Our own lowercase “Roman” type is derived from the Carolingian variant of the Latin script; the Insular and Germanic variants are fairly familiar to us, having enjoyed a period of typographic development as Gaelic and Fraktur, and a handful of letters from the Insular tradition have been adopted by the Carolingian tradition for one purpose or another (WYNN (as VEND) was used in Old Icelandic and Old Norwegain until ca. 1300; THORN and ETH are still used in Icelandic; Insular g and d have
been resurrected by linguists). The Visigothic variant of the Latin script, however, was replaced before the advent of typography, and its unique letterforms were simply lost to the Carolingian, apart from the <§> LATIN LETTER VISIGOTHIC Z. The Carolingian script was introduced in Northern Iberia in the 11th century—in Catalonia, the Spanish Mark of the Carolingian Empire, it was introduced earlier—but it only gained widespread use in the course of the 12th century; the Visigothic script was extinct in the second half of the 12th century (1172 is the date for the last known original Portuguese document). The Visigothic <§> was employed alongside the Carolingian <z>, and came to be used mainly to represent the voiceless alveolar affricate [ts], while <z> was used mainly for the voiced alveolar affricate [dz] in Old Portuguese, Old Leonese, and Old Castilian. In time, as Carolingian practices replaced all memory of the Visigothic, the head of the <§> was reanalyzed, its tail reduced, resulting in a new letter <ç>. While in modern analysis the tail <ç> is known as a cedilla/zedilla ‘little z’, in fact the whole letter is, in origin, a <z>. Documents exist in which <ç> and <§> are distinct (see Figures 43, 44).

5. Other letters of the Insular tradition. One of the letters of the Insular tradition has already been encoded at U+1D79. (A large number of letters in the Fraktur tradition have been encoded for use in mathematics.) The set of Insular letters which differ significantly enough from Carolingian to warrant distinction is small; medievalists have used them in typeset editions of Germanic and Celtic languages since the 16th century. Modern Germanic and Celtic languages do not use these letters, and modern Germanic and Celtic fonts which use Fraktur and Insular letterforms employ them as glyph variants pertaining to the entire font. The Insular letters proposed here are only to facilitate the specific need of historical linguistic specialists to differentiate the Insular letters from the Carolingian.

Insular and Carolingian letters coexisted but were often used in different contexts in Britain and Ireland in the Early Middle Ages, for example Insular letters being used for writing English and Carolingian for Latin. They were also mixed to varying degrees, and this unique variant of the Latin alphabet was exported to the Nordic countries in the 11th century; due to its dual inheritance it has often been termed Carolingian-Insular. Since the letter shapes of Insular and Carolingian script ultimately derive from Uncial script, the majority of Insular and Carolingian letters are basically identical, but a handful letters had quite distinct shapes and usage in Insular script, as we know it from English and Nordic writings. Four of these letters are already in the Standard; THORN (from the Runic alphabet), ETH, WYNN (also from the Runic alphabet), and INSULAR G. The letters THORN and ETH are still used in Icelandic, while WYNN was accepted by the Standard due to its usage in early English sources and INSULAR G on foot of its usage as a phonetic character. We now propose to add five distinct letter forms to the Standard, i.e. INSULAR D, INSULAR F, INSULAR R, INSULAR S, and INSULAR T. It should be underlined that it is not a question of adding Insular variants of every Latin character; it is a short list of distinctive letters that have been recognised as separate characters for several centuries in Medieval English and Nordic writing, and which have been used alongside and in contrast to their Carolingian counterparts. In Medieval Nordic editorial practice, these letters are rendered as separate characters in great many editions and distinguished from their Carolingian-based counterparts <d>, <f>, <r>, <s> and <t>. This is in part because the presence of these letters is used as a dating criterion (for example, INSULAR R fell out of use around 1200, while INSULAR F continued to be used well into the 14th century), and in part because they are used in contrast to their Carolingian counterparts. In Ælfric’s Old English grammar, the scribe, and the modern medievalist, distinguishes between <ç> and <f>, between <ç> and <s>, and between <ç> and <t> (Figure 39). In the sample from the edition of AM 645 4to <ç> and <f> are distinguished (Figure 38).

<Ç> LATIN CAPITAL LETTER INSULAR F (Figure 70), LATIN SMALL LETTER INSULAR F (Figures 29, 30, 32, 33, 35, 36, 37, 38, 39, 40, 42, 70, 71, 73)
<ç> LATIN SMALL LETTER INSULAR R (Figures 29, 30, 37, 39, 40)
<ç> LATIN SMALL LETTER INSULAR S (Figures 16, 29, 30, 37, 39, 40)
6. Letters used for medieval abbreviations. Medieval manuscripts, in both Latin and vernacular languages, use abbreviations extensively. Many of these are abbreviations for whole words, created by omitting letters, such as spīs for spiritus; often a line is placed over the letter(s) as an abbreviation marker, as shown here. Such “logographic” or “lexical” abbreviations can usually be represented through characters already encoded in the UCS. In other cases, however, only a part of a word is abbreviated; for example, the prefix con- is represented with the letter ꞏ. A number of such syllabic abbreviations, well-documented and commonly used in several languages, require letters or combining marks that are not in the UCS. A range of Latin letters, modified by strokes or hooks, is used to represent a variety of words, syllables, or quasi-syllabic letter sequences. That they are polyvalent is a chief indicator for the requirement to encode these “abbreviation letters” as characters, since they cannot be composed of any specific string of other characters; neither can they be decomposed into a single string. Of these abbreviations:

<K> LATIN LETTER K WITH STROKE is used for karta, kartula, kalenda (Figures 33, 45, 74)
<P> LATIN LETTER P WITH STROKE THROUGH DESCENDER is used for Latin and Romance per, par, por, and for pri in Cornish pvechter privechter ‘privacy’ (Figures 46, 47, 50, 58, 60, 63)
<P> LATIN LETTER P WITH FLOURISH is used for pro, por (Figures 46, 50, 52, 54, 55, 58, 60, 61, 63)
<Q> LATIN LETTER Q WITH STROKE THROUGH DESCENDER is used for quam, que, quan- (qndo quando, qtum quantum), qui- (qI3 quilibet, qdem quidem) and in Irish for ar (Figures 16, 17, 46, 49, 51, 53, 59, 63).
<K> LATIN LETTER K WITH DIAGONAL STROKE is used for kalendas and karta (Figure 45)
<K> LATIN LETTER K WITH STROKE AND DIAGONAL STROKE is used for kara(m) ‘document, writ’ when
<K> LATIN SMALL LETTER K WITH STROKE is used for karta (Figures 45)
<L> LATIN LETTER L WITH HIGH STROKE is used for Latin el, ul, vel, for Irish nó ‘or’, for Norse eða ‘or’, for el in vl vel ‘well’, for æl in mlti mælти ‘spoke’, for al in skål skal ‘shall’ (Figure 32, 51, 52, 60, 73)
<Ω> LATIN LETTER O WITH LONG STROKE OVERLAY is used for Latin obiit ‘he died’ (Figures XXX)
<P> LATIN LETTER P WITH SQUIRREL TAIL is used for Latin prae ‘before, in front of’ (Figures XXX)
<Q> LATIN LETTER Q WITH DIAGONAL STROKE is used for Latin quod ‘what’, qui ‘who’, que ‘that’, Portuguese q quem ‘who’, Irish ar ‘on’ (Figures 58, 59, 60, 61, 63, 64)
<f> LATIN SMALL LETTER LONG S WITH DIAGONAL STROKE is used for ser (fíir servir ‘to serve’), sere (fno sereno ‘serene’), siр, and by itself it also stands for Latin solidi, sed, sunt, secundum, etc., and for Portuguese soldo(s) (Figures 46, 48, 58)
<f> LATIN SMALL LETTER LONG S WITH STROKE is used in Norse with <d> for ñ skál ‘shall’ (Figures XXX)
<v> LATIN LETTER V WITH DIAGONAL STROKE is used for Portuguese ver ‘to see’, conversa ‘conversation’, for vere in vador vereador ‘town councillor’, uirgo ‘virgin’, and vir ‘to come’ (Figure 65)
<P> LATIN LETTER THORN WITH STROKE is used for pat, pess, por-, pat (Figures 29, 32, 33, 73)
<P> LATIN LETTER THORN WITH STROKE THROUGH DESCENDER is used for þeim, þeir (Figures XXX).

7. Letters with syllabic content. This set of characters are also abbreviations, but might be better considered as “letters with syllabic content”, because their reading tends to be less polyvalent than those of the abbreviation characters discussed above.

<ʒ> LATIN SMALL LETTER ET is used for -et in viʒ videlicet (whence “viz.”), habʒ habet ‘has’, for -m in abląnez ablationem ‘ablation’, for -ue in usąʒ usque ‘until, till, up to’, for -que in quiąʒ quicumque ‘whoever’, for -us in áquąʒ aliqubus ‘to someone’, for -est in Latin potest potů ‘is
able’, and for medial and final ed in Norse mære ‘with’, mæzan medan ‘while’ (Figures 58, 59, 60, 61, 63, 64).

LATIN SMALL LETTER REVERSED C is used for con and com in ̣ō2 convenientior ‘more suitable, in agreement’, ̣pa comparavit ‘bought’, for -us and -os in Latin in totō totus ‘all’, tūo tuos ‘your’, in Portuguese in todo todos ‘all’, dōo dous ‘two’, and in Old Icelandic ʻa kona ‘woman’ (Figures 46, 50, 66). This is the same character as the Claudian letter discussed in §10 below.

LATIN SMALL LETTER REVERSED C WITH DOT is used for con- and com- in ̣fmas confirmans ‘confirming, witnessing’, -us and -os in soldās soldos (for solidos) ‘a unit of currency’, maladōs malados ‘serfs’ (Figures XXX)

LATIN SMALL LETTER IS is used for -is in dtg dictis ‘from having said’, impq imperatoris ‘ruler, emperor’, and for ys and es in Cornish manuscripts: godg godys ‘god’s’, servantg servantes ‘servants’, mettq mettys ‘met’ (Figure 57).

LATIN SMALL LETTER CON is used for con and cum and co (Figures 46, 47, 48, 53, 61)

MODIFIER LETTER US is used for -us in manus ‘hand’, idus ‘ides’ (and thousands of other words) and final os in Latin, Portuguese, and Castilian: olyr olyros ‘potters’, nos nos ‘we’, uos ‘you’, and for us in Norse: hūus ‘house’ (Figures 50, 52, 59, 66)

LATIN SMALL LETTER DUM is used by itself for Latin dum ‘while, whilst’, die ‘day’, Portuguese dia pl. dias ‘day’

LATIN SMALL LETTER MUM is used for -num in prim, primum ‘first’ (Figures XXX)

LATIN SMALL LETTER NUM is used for -num in aetern, aeternum ‘eternal’, unum ‘one’ (Figures XXX)

LATIN SMALL LETTER RUM is used for -rum in martirum ‘martyr’, integrum ‘intact, whole, undivided’ (Figures XXX)

LATIN SMALL LETTER RUM ROTUNDA is used for -rum and -rom in Latin nō2n, nostrorum ‘of our’, Portuguese fo2n, forom ‘they went, they were’ (Figures 33, 53, 61, 64, 66)

LATIN SMALL CAPITAL RUM is used for -rum and -rom in Latin quos quorum (Figures 59, 63)

LATIN SMALL LETTER TUM is used for -tum in tantum ‘so much’, quantum ‘how much?’ (Figures 58)

LATIN SMALL LETTER UM is used for um and us in Latin ‘ductib2 aquar2’ ‘ductibus aqurum’ ‘to the water streams’, for os in Latin-Portuguese cub2 cubus, cubos ‘cubic measuring container’, nept2 neptos ‘grandson’, and for un in Latin vol2tas voluntas ‘will’, m2dum mundum ‘world’ (Figures XXX)

Note that the REVERSED C is not the turned c U+0254 SMALL LETTER OPEN O used in the IPA and African orthographies; it is ultimately derived from one of the Claudian letters (see §11 below). It should also be noted that these letters were widely used over a long period throughout Europe. As far west as Ireland, these conventional letters were used, sometimes for purposes quite different from their original use. The phrase nó ro-fetatar connacht a ‘or the Connachtmen found out’ could be written t rofratq f̄a, where t Latin uel ‘or’ is used for Irish nó ‘or’, where the Tironian sign ̄ is used for et, where q is used for ar ‘on’, where ̄ is used for conn (= con), and where ̄ Latin sed ‘but’ is used for Irish acht ‘but’. Old Icelandic manuscripts were among the most abbreviated of all vernacular European manuscripts; in some cases almost every word in a line was abbreviated (Figure 36)

8. Combining characters. Thirteen combining superscript letters are already encoded to represent medieval Germanic manuscripts. These comprise half of the basic Latin alphabet, shown in bold type here: abcdefghijklmnopqrstuvwxyz. We propose to add ten more basic superscript letters attested in medieval manuscripts which will bring the repertoire to 23 of the 26 letters: abcdefghijklmnopqrstuvwxyz. (It should be noted that of the traditional Latin alphabet, if these ten are added, only *COMBINING LATIN SMALL LETTER J, *COMBINING LATIN SMALL LETTER Q, and *COMBINING
LING SMALL LETTER W will remain unencoded.) We also propose to encode superscripted æ, œ, a', c, d, 
ؤ, K, L, M, N, R, 2, f, and t.

It should be noted explicitly that the combining “capitals” in Old Norse are considered as combining small capitals. Thus a COMBINING SMALL CAPITAL G would be an abbreviation for <gg>, in the same manner as a LATIN SMALL CAPITAL G (on the base line, that is) would be understood as equivalent to <gg>. The relative x-height of the COMBINING SMALL CAPITALS is the same as that of the x-height of the COMBINING SMALL LETTERS. The reason Old Norse added a few small capitals as superscript characters—
in addition to the inventory of ordinary small characters—is the peculiar Old Icelandic custom of using small capitals for geminates; this practice was transferred to the practice of abbreviation by way of superscript characters.

<">

COMBINING LATIN SMALL LETTER AE is used for atgðamikill atqvaðamikill ‘resolute’ (Figures XXX)

<=">

COMBINING LATIN SMALL LETTER AO is used for heiðmo heimqvaomo ‘return home’ (Figures XXX)

<"> COMBINING LATIN SMALL LETTER AV is used for b pf fozla, brauðsfözla ‘feeding with bread’ (Figures XXX)

<">

COMBINING LATIN SMALL LETTER B is used for XXXX (Figures XXX)

<">

COMBINING LATIN SMALL LETTER SMALL CAPITAL B is used for XXXX (Figures XXX)

<"> COMBINING LATIN SMALL LETTER C CEDILLA is used for Portuguese ç ‘concelho ‘municipality’ (Figures XXX)

<">

COMBINING LATIN SMALL LETTER CAPITAL D is used for XXXX (Figures XXX)

<">

COMBINING LATIN SMALL LETTER ETH is used for fpolt spiotid ‘the spear’ (Figures XXX)

<">

COMBINING LATIN SMALL LETTER F is used for XXXX (Figures XXX)

<">

COMBINING LATIN SMALL LETTER G is used for XXXX (Figures XXX)

<">

COMBINING LATIN SMALL LETTER SMALL CAPITAL G is used for xx ‘thirtieth’ (Figures XXX)

<">

COMBINING LATIN SMALL LETTER K is used for m mik ‘me’ (Figures XXX)

<">

COMBINING LATIN SMALL LETTER SMALL CAPITAL K is used for XXXX (Figures XXX)

<">

COMBINING LATIN SMALL LETTER L is used for Latin n nihil ‘nothing’, Portuguese g geral ‘general’. Old Norse til ‘to’ (Figures XXX)

<">

COMBINING LATIN SMALL LETTER SMALL CAPITAL L is used for mik mikill ‘great, tall’ (Figures XXX)

<">

COMBINING LATIN SMALL LETTER M is used for XXXX (Figures XXX)

<">

COMBINING LATIN SMALL LETTER N is used for in Latin unde ‘from’, ante ‘before’, quando ‘when’, Old Norse sid ‘since’ (Figures XXX)

<">

COMBINING LATIN SMALL LETTER SMALL CAPITAL N is used for m menn ‘men’ (Figures XXX)

<">

COMBINING LATIN SMALL LETTER P is used for XXXX (Figures XXX)

<">

COMBINING LATIN SMALL LETTER SMALL CAPITAL R is used for Gunh Gunnarr ‘Gunnar’ (Figures XXX)

<">

COMBINING LATIN SMALL LETTER R ROTUNDA is used for Latin uuó quatuor ‘four’, Portuguese pto porto ‘harbour’, M Martin ‘Martin’, Old Norse spð spordi ‘asked’ (Figures XXX)

<">

COMBINING LATIN SMALL LETTER S is used for Old Norse þ pes ‘this’, hans ‘his’ (Figures XXX)

<">

COMBINING LATIN SMALL LETTER LONG S is used for Latin uuó duos ‘two’, uué tres ‘three’ (Figures XXX)

<">

COMBINING LATIN SMALL LETTER SMALL CAPITAL T is used for XXXX (Figures XXX)

<">

COMBINING LATIN SMALL LETTER Y is used for þr fyrr ‘before’ (Figures XXX)

<">

COMBINING LATIN SMALL LETTER Z is used for q quað ‘said’ (Figures XXX)

In addition to these, seven other combining marks are proposed here.

<">

COMBINING DOUBLE CIRCUMFLEX ABOVE is used to denote the two diphthongs [ea] and [au] in the first Faroese orthography by Jens Christian Svabo (1746–1824)—it is also used in editions of
Old English poetry to indicate disyllabic pronunciation of a diphthong that is normally monosyllabic (Figures 27, 31)

Combining Ogonek Above is used for marking vowel-length in Norse or to indicate vowel affection—so ᄴ represents i-mutated ø (Figures 19, 23, 25, 28, 42)

Combining Zigzag Below is used for ⟨⟨⟩⟩ per ‘they f.’ together with Combining Zigzag Above (Figures XXX)

Combining Is Below is used in Visigothic script for is in nobis ‘to us’, script scriptis ‘written’, dic dictis ‘said’ (Figures XXX)

Combining Ur Above is used for ⟨⟨⟩⟩ in dicit dicitur ‘is said’, uocat uocatur ‘is called’ (Figures 32, 60, 61)

Combining Us Above is used for medial and final us in man manus ‘hand’, medial os in pt post ‘after’, ṕpitus praepositus ‘prelate, leader, governor, prevost’ (Figures 32, 33, 39, 51, 52)


9. Punctuation. Although many editors substitute modern for medieval punctuation, a growing number of medievalists insist that modern punctuation cannot adequately represent the syntactical features of medieval texts. In any case, at present, the reproduction of the punctuation of most medieval manuscripts using currently defined UCS characters is impossible, because many marks (e.g. Punctus Elevatus) are simply missing. More subtly, however, even those medieval marks that have modern reflexes may differ in usage as well as in shape. For example, the modern question mark (U+003F) is descended from the Punctus Interrogativus, but unlike the modern question mark, the Punctus Interrogativus may be contrasted with the Punctus Percontativus—the former marking questions that require an answer while the latter marks rhetorical questions. These medieval punctuation marks are proposed below. Note that where a number of characters are described as “variants of Positura”, this does not mean that they are freely-exchangeable glyph variants. Like the Norse ligated letters discussed in §2 above, this collection of punctuation marks is a superset of the letters found in the medieval corpus; no single manuscript contains all of them.

MEDIEVAL COMMA is used to mark a pause at the end of the rhetorical unit called a comma.

PARAGRAPHUS is used to mark the beginning of a section of text (Figures 53, 59).

POSITURA is used to mark the end of a section of text.

Colon with Comma Elevata Positura, an alternative form of Positura, was “employed by insular scribes at the end of a paragraph in a series of paragraphs or texts (e.g. annals) to imply that some continuation was to be expected to complete this series” (Parkes 306).

Colon with Middle Comma Positura is a variant form of POSITURA.

TWO DOTS OVER COMMA POSITURA is a variant form of POSITURA.

THREE DOTS WITH COMMA POSITURA is a variant form of POSITURA.

Punctus Elevatus is used in many medieval texts to signal a wide variety of sentence-medial pauses (Figures 53, 63, 70).

Punctus Exclamativus (or punctus admirativus) is used in late medieval texts to mark the end of an exclamation.

Punctus Flexus is used to mark sentence-medial pauses, but especially in liturgical texts where the pitch of the voice drops.

Punctus Interrogativus is used to mark the end of a question that requires an answer.

Punctus Percontativus is used in late texts to mark the end of a question that does not require an answer—a rhetorical question.

Punctus Versus is used to mark the end of a sentence or period, or the end of a verse of a psalm.

Low Punctus Versus is a variant form of Punctus Versus.
SIGN DE RENVOI is used to connect a marginal note to a place in the text. Its size is typically larger than that of other.

SQUARED FOUR DOT PUNCTUATION.

TWO DOTS OVER ONE DOT PUNCTUATION. (Figure 62).

ONE DOT OVER TWO DOTS PUNCTUATION.

VIRGULA SUSPENSIVA is “used to mark the briefest pause or hesitation in a text” (Parkes 307).

Three of the characters listed above are similar to three characters used in mathematics. The reason that these are proposed is to finally complete the set of generic punctuation encoded in the UCS. Most of these are based around the shape and size of the generic COLON and SEMICOLON: The UCS currently contains nine of such punctuation characters to which we propose to add the thirteen here: There are also tall characters like the existing seven to which we propose to add five . Therefore, U+2235 therefore, U+2235: proportion are likely not to be based on the typical proportions of the modern COLON (indeed RATIO is disunified from COLON); further research needs to be done into their linebreaking and other properties.

10. Claudian letters. The Roman emperor Claudius introduced three letters to the alphabet to indicate sounds he felt could not be represented otherwise. These letters originally occurred only in capital form, but the lower-case form has been adduced to meet Unicode casing requirements as well as those of normal scholarship. An inscription may read OLMIPICUS, but the scholar may prefer to transcribe this as Olγympicus (as in Figure XXX) without the letter, or as Olḥmpicus with it.

REVERSED C represented the consonant groups ḫs and ḫps. This is the same character as the letter used for con, com, us, and os discussed in §7 below.

INVERTED F represented consonantal q (as opposed to u)

HALF H represented the Greek (Figures XXX).

11. Editorial punctuation. The Nordic editorial tradition uses quilled bars to mark deletions in the manuscript and (in reverse order) to mark repetitions (Figures 26, 38, 68). The venerable Corpus Inscriptionum Latinarum of Berlin makes use of a large number of paired punctuation, most of which are encoded in the General Punctuation or one of the Mathematics blocks. Two pairs are not yet encoded: the SIDEWAYS U BRACKETS of Berlin makes use of a large number of paired punctuation, most of which are encoded in the General Punctuation or one of the Mathematics blocks. Two pairs are not yet encoded: the SIDEWAYS U BRACKETS are used to indicate Claudian letters in a transcription, to notate inverted letters, or to notate the insertion of an image into the line of text (such as to indicate a cross). The (double parentheses) are used by some other Latinists for similar purposes. They differ from ((double parentheses)) quite obviously, and are not joined at the top like the WHITE PARENTHESES are. The FULLWIDTH WHITE PARENTHESES used in Asia are unusual in having either a double or a joined presentation, but that glyph variation is not used by Latinists (Figures XXX).

Bibliography.


Bartoli, Francesco de. 1470. Historia quomodo beatus Franciscus petivit a Christo indulgentiam pro Ecclesia S. Mariae de Angelis. Trevi: Johannes Rothmann.

Everson et al.


Humphreys Henry Noel. [1868]. “The origin and progress of the art of writing; a connected narrative on the development of the art, in its primeval phases in Egypt, China, and Mexico; its middle state in the Cuneatic systems of Ninevah and Persepolis; its introduction to Europe through the medium of the Hebrew, Phœnician, and Greek systems; and its subsequent progress to the present day”, in *Webster's Improved Dictionary of the English language, exhibiting the origin, orthography, pronunciation, & definition of words; embracing all the principal terms used in literature, science & art, according to the best authorities; and likewise giving the synonymous terms for nearly all the words explained*. 2 vols. London, Glasgow, & Edinburgh: William MacKenzie.

Johnson, Samuel. 1828. A dictionary of the English language in which the words are deduced from their originals, and illustrated in their different significations by examples from the best writers, to which are prefixed a history of the language and an English grammar. London: Joseph Ogle Robinson.


Pacheco, José. 1988. A Divina Arte Negra e o Livro Português (séculos XV e XVI), Lisboa: Vega, Fig. 3 — Folha do Prólogo do Sacramental de Clemente Sanchez, Chaves, Autor desconhecido, [1488]., Aquele que se pressupõe ter sido, em Portugal, o primeiro livro impresso em português., p. 87


Tertullian, Quintus Septimus Florens. [1493]. Apologeticus adversus gentes. Venetiis: B. Benalius.


Acknowledgements

This project was made possible in part by a grant from Menota (the Medieval Nordic Text Archive) to the Script Encoding Initiative at UC Berkeley.
### N2957: Preliminary proposal to add medievalist characters to the UCS

#### Combining Diacritical Marks Supplement

<table>
<thead>
<tr>
<th></th>
<th>xDC</th>
<th>xDD</th>
<th>xDE</th>
<th>xDF</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
<td>L</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>N</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>p</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>r</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>t</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>T</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td></td>
<td>y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td></td>
<td>z</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>K</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Note to reviewers: This is a preliminary proposal. The code points shown are merely suggestions for a possible placement.

Medievalist and Latinist additions

- COMBINING DOUBLE CIRCUMFLEX ABOVE
- COMBINING OGONEK ABOVE
- COMBINING ZIGZAG BELOW
- COMBINING IS BELOW
- COMBINING UR ABOVE
- COMBINING US ABOVE
- \( \text{Æ} \) modifier letter us

Superscript Latin letter diacritics

- COMBINING LATIN SMALL LETTER FLATTENED OPEN A
- COMBINING LATIN SMALL LETTER AE
- COMBINING LATIN SMALL LETTER AO
- COMBINING LATIN SMALL LETTER AV
- COMBINING LATIN SMALL LETTER B
- COMBINING LATIN LETTER SMALL CAPITAL B
- COMBINING LATIN SMALL LETTER C CEDILLA
- COMBINING LATIN LETTER SMALL CAPITAL D
- COMBINING LATIN SMALL LETTER ETH
- COMBINING LATIN SMALL LETTER F
- COMBINING LATIN SMALL LETTER G
- COMBINING LATIN SMALL LETTER H
- COMBINING LATIN SMALL LETTER K
- COMBINING LATIN SMALL LETTER L
- COMBINING LATIN SMALL LETTER M
- COMBINING LATIN SMALL LETTER N
- COMBINING LATIN SMALL LETTER P
- COMBINING LATIN LETTER SMALL CAPITAL R
- COMBINING LATIN SMALL LETTER S
- COMBINING LATIN SMALL LETTER S LONG S
- \( \text{š} \) latin small letter long s
- COMBINING LATIN LETTER SMALL CAPITAL T
- COMBINING LATIN SMALL LETTER Y
- OMBINING LATIN SMALL LETTER Z
Note to reviewers: This is a preliminary proposal. The code points shown are merely suggestions for a possible placement.

**Latinist and medievalist characters**

<table>
<thead>
<tr>
<th>Code</th>
<th>Character</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1E9C</td>
<td>( f )</td>
<td>LATIN SMALL LETTER LONG S WITH DIAGONAL STROKE</td>
</tr>
<tr>
<td>017F</td>
<td>( f )</td>
<td>latin small letter long s</td>
</tr>
<tr>
<td>1E9D</td>
<td>( ſ )</td>
<td>LATIN SMALL LETTER LONG S WITH HIGH STROKE</td>
</tr>
</tbody>
</table>

**Latinist and medievalist characters**

<table>
<thead>
<tr>
<th>Code</th>
<th>Character</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1EFA</td>
<td>( IL )</td>
<td>LATIN CAPITAL LETTER MIDDLE-WELSH LL</td>
</tr>
<tr>
<td>1EFB</td>
<td>( ï )</td>
<td>LATIN SMALL LETTER MIDDLE-WELSH LL</td>
</tr>
<tr>
<td>1EFC</td>
<td>( ù )</td>
<td>LATIN CAPITAL LETTER MIDDLE-WELSH V</td>
</tr>
<tr>
<td>01B2</td>
<td>( ù )</td>
<td>latin capital letter v with hook</td>
</tr>
<tr>
<td>1EFD</td>
<td>( û )</td>
<td>LATIN SMALL LETTER MIDDLE-WELSH V</td>
</tr>
<tr>
<td>028B</td>
<td>( û )</td>
<td>latin small letter v with hook</td>
</tr>
<tr>
<td>1EFF</td>
<td>( Y )</td>
<td>LATIN CAPITAL LETTER Y WITH LOOP</td>
</tr>
<tr>
<td>1EFF</td>
<td>( y )</td>
<td>LATIN SMALL LETTER Y WITH LOOP</td>
</tr>
</tbody>
</table>
N2957: Preliminary proposal to add medievalist characters to the UCS
Note to reviewers: This is a preliminary proposal. The code points shown are merely suggestions for a possible placement.

Latinist and medievalist character

2C6B δ LATIN SMALL LETTER SCRIPT D
   → 018D ṝ latin small letter turned delta
   → 03B4 δ greek small letter delta
N2957: Preliminary proposal to add medievalist characters to the UCS

<table>
<thead>
<tr>
<th>0x00</th>
<th>0x01</th>
<th>0x02</th>
<th>0x03</th>
<th>0x04</th>
<th>0x05</th>
<th>0x06</th>
<th>0x07</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0x10</td>
<td>0x11</td>
<td>0x12</td>
<td>0x13</td>
<td>0x14</td>
<td>0x15</td>
<td>0x16</td>
<td>0x17</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0x20</td>
<td>0x21</td>
<td>0x22</td>
<td>0x23</td>
<td>0x24</td>
<td>0x25</td>
<td>0x26</td>
<td>0x27</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0x30</td>
<td>0x31</td>
<td>0x32</td>
<td>0x33</td>
<td>0x34</td>
<td>0x35</td>
<td>0x36</td>
<td>0x37</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0x40</td>
<td>0x41</td>
<td>0x42</td>
<td>0x43</td>
<td>0x44</td>
<td>0x45</td>
<td>0x46</td>
<td>0x47</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0x50</td>
<td>0x51</td>
<td>0x52</td>
<td>0x53</td>
<td>0x54</td>
<td>0x55</td>
<td>0x56</td>
<td>0x57</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0x60</td>
<td>0x61</td>
<td>0x62</td>
<td>0x63</td>
<td>0x64</td>
<td>0x65</td>
<td>0x66</td>
<td>0x67</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0x70</td>
<td>0x71</td>
<td>0x72</td>
<td>0x73</td>
<td>0x74</td>
<td>0x75</td>
<td>0x76</td>
<td>0x77</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Supplemental Punctuation

<table>
<thead>
<tr>
<th>0x00</th>
<th>0x01</th>
<th>0x02</th>
<th>0x03</th>
<th>0x04</th>
<th>0x05</th>
<th>0x06</th>
<th>0x07</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0x10</td>
<td>0x11</td>
<td>0x12</td>
<td>0x13</td>
<td>0x14</td>
<td>0x15</td>
<td>0x16</td>
<td>0x17</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0x20</td>
<td>0x21</td>
<td>0x22</td>
<td>0x23</td>
<td>0x24</td>
<td>0x25</td>
<td>0x26</td>
<td>0x27</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0x30</td>
<td>0x31</td>
<td>0x32</td>
<td>0x33</td>
<td>0x34</td>
<td>0x35</td>
<td>0x36</td>
<td>0x37</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0x40</td>
<td>0x41</td>
<td>0x42</td>
<td>0x43</td>
<td>0x44</td>
<td>0x45</td>
<td>0x46</td>
<td>0x47</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0x50</td>
<td>0x51</td>
<td>0x52</td>
<td>0x53</td>
<td>0x54</td>
<td>0x55</td>
<td>0x56</td>
<td>0x57</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0x60</td>
<td>0x61</td>
<td>0x62</td>
<td>0x63</td>
<td>0x64</td>
<td>0x65</td>
<td>0x66</td>
<td>0x67</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0x70</td>
<td>0x71</td>
<td>0x72</td>
<td>0x73</td>
<td>0x74</td>
<td>0x75</td>
<td>0x76</td>
<td>0x77</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Note to reviewers: This is a preliminary proposal. The code points shown are merely suggestions for a possible placement.

Latinist and medievalist editorial punctuation

2E1E ° MIDDLE RING
  ° 00B0 ° degree sign
  ° 02DA ° ring above
  ° 1424 ° candian syllabics final ring
  ° 2218 ° ring operator
  ° 25E6 ° white bullet
  ° 26AC ° medium small white circle

2E1F ° PALM BRANCH
  ° 205A ° two dot punctuation
  ° 205B ° four dot mark
  ° 2237 ° proportional

2E20 ° LEFT VERTICAL BAR WITH QUILL
2E21 ° RIGHT VERTICAL BAR WITH QUILL
2E22 ° LEFT SIDEWAYS U BRACKET
  ° 2292 ° subset of
2E23 ⇒ RIGHT SIDEWAYS U BRACKET
  ° 2283 ° superset of

2E24 ° LEFT DOUBLE PARENTHESIS
  ° 2985 ° left white parenthesis
  ° FFSF ° fullwidth left white parenthesis
2E25 ° RIGHT DOUBLE PARENTHESIS
  ° 2986 ° left right parenthesis
  ° FF60 ° fullwidth right white parenthesis

2E26 ° TOP LEFT HALF SQUARE BRACKET
  ° 230B ° left ceiling
  ° 231C ° top left corner
  ° 2E00 ° right angle substitution marker
  ° 300C ° left corner bracket
2E27 ° TOP RIGHT HALF SQUARE BRACKET
  ° 2309 ° right ceiling
  ° 231D ° top right corner

2E28 ° BOTTOM LEFT HALF SQUARE BRACKET
  ° 2301 ° left floor
  ° 231E ° bottom left corner
2E29 ° BOTTOM RIGHT HALF SQUARE BRACKET
  ° 230F ° right floor
  ° 231F ° bottom right corner
  ° 300D ° right corner bracket

Medieval punctuation

2E30 ° MEDIEVAL COMMA
2E31 ° PARAGRAPHUS
  ° 00B6 ° pilcrow sign
  ° 204B ° reversed pilcrow sign
  ° 2761 ° curved stem paragraph sign ornament
  ° 2E0F ° paragraphos
2E32 ° POSITURA
2E33 ° COLON WITH COMMA ELEVATA POSITURA
2E34 ° COLON WITH MIDDLE COMMA POSITURA
2E35 ° TWO DOTS OVER COMMA POSITURA
2E36 ° THREE DOTS WITH COMMA POSITURA
2E37 ° PUNCTUS ELEVATUS
2E38 ° PUNCTUS EXCLAMATIVUS
2E39 ° PUNCTUS FLEXUS
2E3A ° PUNCTUS INTERROGATIVUS
2E3B ° PUNCTUS PERCONTATIVUS
  ° 2424 ° symbol for substitute form two
  ° 061F ° arabic question mark
2E3C ° PUNCTUS VERSUS
2E3D ° LOW PUNCTUS VERSUS
Note to reviewers: This is a preliminary proposal. The code points shown are merely suggestions for a possible placement.

Medievalist additions

A720 Ä LATIN CAPITAL LETTER AA
A721 ä LATIN SMALL LETTER AA
A722 Å LATIN CAPITAL LETTER AO
A723 å LATIN SMALL LETTER AO
A724 Å LATIN CAPITAL LETTER AU
A725 å LATIN SMALL LETTER AU
A726 Å LATIN CAPITAL LETTER AV
A727 æ LATIN SMALL LETTER AV
A728 À LATIN CAPITAL LETTER AV WITH HORIZONTAL BAR
A729 æ LATIN SMALL LETTER AV WITH HORIZONTAL BAR
A72A À LATIN CAPITAL LETTERAY
A72B a LATIN SMALL LETTERAY
A72C Æ LATIN CAPITAL LETTER REVERSED C
A72D æ LATIN SMALL LETTER REVERSED C
A72E Æ LATIN CAPITAL LETTER REVERSED C WITH DOT
A72F ë LATIN SMALL LETTER REVERSED C WITH DOT
A730 Ø LATIN CAPITAL LETTER TURNED F
A731 ð LATIN SMALL LETTER TURNED F
A732 ñ LATIN CAPITAL LETTER HALF H
A733 ñ LATIN SMALL LETTER HALF H
A734 ë LATIN CAPITAL LETTER K WITH STROKE
A735 ë LATIN SMALL LETTER K WITH STROKE
A736 ë LATIN CAPITAL LETTER K WITH DIAGONAL STROKE
A737 ë LATIN SMALL LETTER K WITH DIAGONAL STROKE
A738 ë LATIN CAPITAL LETTER K WITH STROKE AND DIAGONAL STROKE
A739 ë LATIN SMALL LETTER K WITH STROKE AND DIAGONAL STROKE
A73A ë LATIN CAPITAL LETTER BROKEN L
A73B ë LATIN SMALL LETTER BROKEN L
A73C ë LATIN CAPITAL LETTER L WITH HIGH STROKE
A73D ë LATIN SMALL LETTER L WITH HIGH STROKE
A73E ë LATIN CAPITAL LETTER O WITH LONG STROKE OVERLAY
A73F ë LATIN SMALL LETTER O WITH LONG STROKE OVERLAY
A740 ñ LATIN CAPITAL LETTER O WITH LOOP
A741 ñ LATIN SMALL LETTER O WITH LOOP
A742 ñ LATIN CAPITAL LETTER OO
A743 ñ LATIN SMALL LETTER OO
A744 ñ LATIN CAPITAL LETTER P WITH STROKE THROUGH DESCENDER
A745 p LATIN SMALL LETTER P WITH STROKE THROUGH DESCENDER
A746 P LATIN CAPITAL LETTER P WITH FLOURISH
A747 p LATIN SMALL LETTER P WITH FLOURISH
A748 ß LATIN CAPITAL LETTER P WITH SQUIRREL TAIL
A749 ß LATIN SMALL LETTER P WITH SQUIRREL TAIL
A74A Q LATIN CAPITAL LETTER Q WITH STROKE THROUGH DESCENDER
A74B q LATIN SMALL LETTER Q WITH STROKE THROUGH DESCENDER
A74C Q LATIN CAPITAL LETTER Q WITH DIAGONAL STROKE
A74D q LATIN SMALL LETTER Q WITH DIAGONAL STROKE
A74E 2 LATIN CAPITAL LETTER R ROTUNDA
A74F 2 LATIN SMALL LETTER R ROTUNDA
A750 2 LATIN CAPITAL LETTER RUM ROTUNDA
A751 2 LATIN SMALL LETTER RUM ROTUNDA
A752 V LATIN CAPITAL LETTER V WITH DIAGONAL STROKE
A753 V LATIN SMALL LETTER V WITH DIAGONAL STROKE
A754 W LATIN CAPITAL LETTER VY
A755 w LATIN SMALL LETTER VY
A756 ℳ LATIN CAPITAL LETTER VISIGOthic Z
A757 ℳ LATIN SMALL LETTER VISIGOthic Z
A758 ℳ LATIN CAPITAL LETTER THORN WITH STROKE
A759 ℳ LATIN SMALL LETTER THORN WITH STROKE
A75A ℳ LATIN CAPITAL LETTER THORN WITH STROKE THROUGH DESCENDER
A75B ℳ LATIN SMALL LETTER THORN WITH STROKE THROUGH DESCENDER
A75C ℳ LATIN CAPITAL LETTER VEND
A75D ℳ LATIN SMALL LETTER VEND
A75E ℳ LATIN CAPITAL LETTER ET
A75F ℳ LATIN SMALL LETTER ET
A760 ℳ LATIN CAPITAL LETTER IS
A761 ℳ LATIN SMALL LETTER IS
A762 ℳ LATIN CAPITAL LETTER CON
A763 ℳ LATIN SMALL LETTER CON
A764 ℳ MODIFIER LETTER US

Insular tradition

Other characters for the insular tradition are coded elsewhere.

A765 õ LATIN SMALL LETTER INSULAR D
A766 ò LATIN CAPITAL LETTER INSULAR F
A767 ò LATIN SMALL LETTER INSULAR F
A768 ò LATIN SMALL LETTER INSULAR R
A769 ò LATIN SMALL LETTER INSULAR S
A76A ò LATIN SMALL LETTER INSULAR T

Other additions

Other small capital letters for phonetic transcription are encoded elsewhere.

A76B f LATIN LETTER SMALL CAPITAL F
Latin Extended-D

Latinist abbreviations

- A76C  ` LATIN LETTER SMALL CAPITAL S
- A76D  ç LATIN SMALL LETTER DUM
- A76E  ṭ LATIN SMALL LETTER MUM
- A76F  ņ LATIN SMALL LETTER NUM
- A770  ŕ LATIN SMALL LETTER RUM
- A771  ν Latin small letter long s
- A772  ō LATIN SMALL LETTER TUM
- A773  ş LATIN SMALL LETTER UM