

Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation Internationale de Normalisation
Международная организация по стандартизации

Doc Type: Working Group Document
Title: Proposal to add two SIGMOID S characters for mediaeval palaeography to the UCS
Source: Michael Everson
Status: Individual Contribution
Date: 2020-10-01

This proposal requests the encoding of two characters used in the medieval palaeography. If this proposal is accepted, the following characters will exist:

Œ A7D8 LATIN CAPITAL LETTER SIGMOID S

ſ A7D9 LATIN SMALL LETTER SIGMOID S

- used in Middle Cornish, Middle English, and Middle Scots

1. SIGMOID s. A letter common in Britain has long been identified, but is now proposed here for encoding because, for palaeographic purposes, its encoding can help identify and date hands. The proposal to encode this character is analogous to the proposal for encoding the ANGLICANA W: its identity has always been recognized, but it has only been in palaeographic context that its coexistence alongside Roman s and LONG S has been observed. An editor may transcribe $\mathfrak{w}/\mathfrak{W}$ or $\mathfrak{w}/\mathfrak{w}$ as w in some kinds of transcriptions, just as an editor may transcribe $\mathfrak{f}/\mathfrak{s}$ and $\mathfrak{f}/\mathfrak{r}$ as s, but in a palaeographic edition, the distinction between *w* and *is* or *f* and *s* (or *r* and *z*, etc) can be made. In the case of the SIGMOID s, a three-way distinction has been found in some manuscripts, so like the ANGLICANA W it should be distinguished at character level.

2. Ordering. These characters should be sorted as variants of the letter s. Note: The following changes to the UCA should also be made.

- The letters LATIN CAPITAL LETTER R ROTUNDA and LATIN SMALL LETTER R ROTUNDA should be sorted as variants of the letter r, not as separate letters between \mathfrak{r} and \mathfrak{r} .
- The letters LATIN CAPITAL LETTER ANGLICANA W and LATIN SMALL LETTER ANGLICANA W should be sorted as variants of the letter w, not as separate letters between (small-capital) w and w.

3. Security. As an historic character, it is expected that the SIGMOID s will not be required in identifiers.

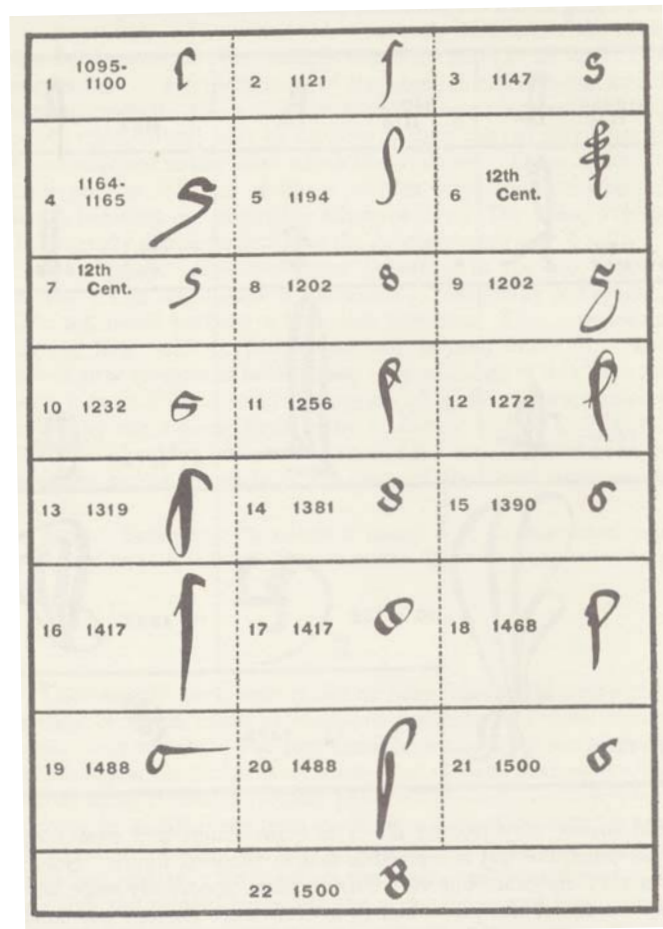
4. Unicode Character Properties. Character properties are proposed here.

A7D8;LATIN CAPITAL LETTER SIGMOID S;Lu;0;L;;;;N;;;A7D9;
A7D9;LATIN SMALL LETTER SIGMOID S;Ll;0;L;;;;N;;;A7D8;;A7D8

5. Bibliography

- Denholm-Young, N. 1964. *Handwriting in England and Wales*. Second edition. Cardiff: University of Wales Press.
- Johnson, Charles, and Hilary Jenkinson. 1915. *English Court Hand A.D. 1066 to 1500, illustrated chiefly from the public records*. Oxford: Oxford University Press. Reprinted New York: Friedrich Unger, 1967.
- Johnson, Samuel. 1773. *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 appended a history of the languages, and an English grammar*. Reprinted London: Joseph Ogle Robinson, 1828.
- Parkes, M. B. 1979. *English Cursive Book Hands 1250-1500*. London: Scolar Press. ISBN 0-85967-535-1
- Roberts, Jane. 2005. *Guide to Scripts used in English Writings up to 1500*. London: The British Library. ISBN 0-7123-4884-0
- Simpson, Grant S. 2009. *Scottish handwriting 1150-1650: An introduction to the reading of documents*. Edinburgh: John Donald. ISBN 978-1-906566-11-1.
- Wright, Andrew. 1879. *Court-Hand Restored, or the Student's Assistant in reading Old Deeds. Charters. Records, etc. neatly engraved on twenty-three copper plates, describing The Old Law Hands, with their Contractions and Abbreviations*. London: Reeves & Turner.
- Wright, C. E. 1960. *English vernacular hands from the twelfth to the fifteenth centuries*. (Oxford Palaeographical Handbooks.) Oxford: Clarendon Press.

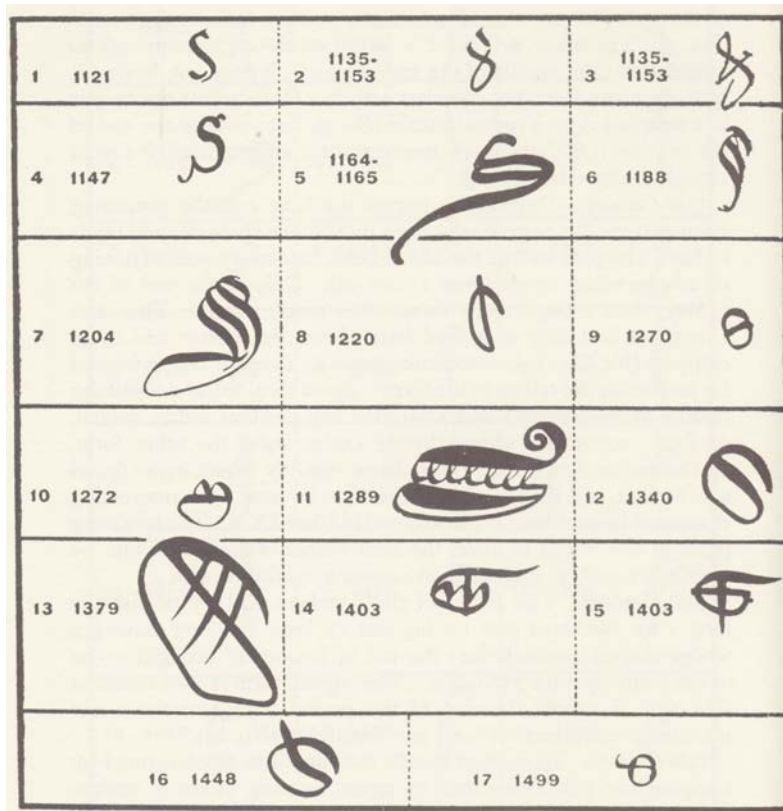
6. Figures.



13th Century. During this period the long *s* of the preceding century (type No. 2) gradually loses its sinuous character and tends to have a loop, as well as the added head, for convenience of joining on to preceding letters (Nos. 11 and 12). Only at the end of the century does it begin to be formed in a single stroke. The short *s*, used at first only as a final letter, becomes shorter and more compact (No. 8). It is sometimes made so that the top is formed by producing the tail as in the Greek *sigma* (No. 10), and about the middle of the century may be found in any position, initial, medial, or final; sometimes almost to the exclusion of the other form. At the end of the century it has almost entirely ceased to be found as a medial, and the long *s* has ceased to be final. An interesting decorated form of long *s* will be found in Plate IX, B. The tendency (seen in this Plate) to lower the serif almost to the line should be carefully noted, as it sometimes causes a confusion with *f*.

14th Century. The forms of the previous century persist, the long *s* for the most part having lost its loop and now having a wedge-shaped down-stroke: the tail is frequently brought up as in the contemporary *f* (No. 13). The *sigma* form is also common (No. 15). Towards the end of the period both varieties show increasing coarseness of form and angularity (No. 14).

Figure 1. Example from Johnson and Jenkinson 1915:44–45 showing a variety of lower-case *s*'s. The SIGMOID *s* appears in the thirteenth century and can be seen in Nos. 10, 15, 17, 19, and 21.



From the beginning of the 14th century till the end of our period, the usual form is (cf. one form of *s*) a modification of the normal *S*, very like a Greek σ , in which the lower bow is closed up into an *o* form; it is made, however, in one action of the pen, beginning from the centre of the letter (No. 12). Further modifications of this, in which the first stroke becomes horizontal instead of diagonal, are seen in Nos. 15 and 17.

Figure 2. Example from Johnson and Jenkinson 1915:46–47 showing a variety of upper-case *S*'s. The SIGMOID *s* appears in the thirteenth century and can be seen in Nos. 12–17.

the line (thus ζ). The characteristic insular *s* of the Anglo-Saxon hand (γ) is not exemplified in the facsimile (see no. 1) from the *Peterborough Chronicle MS.*; it had been used by the first scribe but disappeared entirely in the second scribe's work. The long *s* of the Anglo-Saxon hand (sometimes very exaggerated in form) persists, however, with some modification of its extravagance and towards the end of the period lost the little tag to the left. The modern form (*s*) appears by 1100 (e.g. in the *Textus Roffensis*) and in our examples is

already present in the *Ormulum MS.* (of *c.* 1200) (see no. 2). It is used normally in final position for the greater part of the period. Later there appears a third form, the Greek *s* (σ), at first finally (see no. 9). This derives from the court-hand, whose influence is at work on the vernacular book-hands throughout almost the whole of the Middle English period, some of the literary texts being actually written in it, e.g. the 'Harley Lyrics' (see no. 9). In writing the

Figure 3. Example from Wright 1960:xv-xvi describing the development of the SIGMOID *s*. Wright describes three forms: *f*, *s*, and *s*.

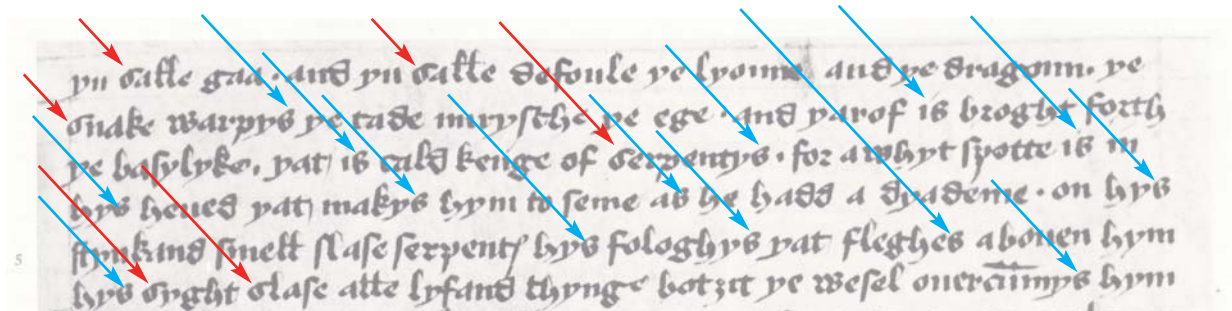


Figure 4. Example from Parkes 1979: Plate 20 (MS Bodley 467 (SC. 2487), f. 120r), in which ordinary s (with a B-shape; see Figure 8 below) is usual in final position, but in which LONG s and SIGMOID s both begin words; LONG s also appears medially. The text, a commentary by Richard Rolle on the Psalter, reads:

þu **s**alle gaa. and þu **s**alle defoule þe lyonne and þe dragonn. þe
snake wærpys þe tade nuryfch^s þe ege. and þarof **i**s bꝛoght foꝛth
þe basylke. þat **i**s cald kenge of **s**erpenty**s**. foꝛ a wꝛht **f**potte **i**s in
hys heved þat makys hym to **f**eme **a**s he hadd a dyademe. on hys
ftynkand **f**inell **f**lase **f**erpent^f hys **f**ologhys þat fleghes abouen hym
hys **s**yght **s**lase atte yfand thyng^s bot zit þe wæfel ouercūmys hym

From late in the thirteenth century scribes were trying to find a book hand that was not difficult to write on a small scale. At this time the tightly woven Gothic *textualis* was still the only alternative to a documentary hand, but English scribes seem to have happened upon a way of writing that soon became their ordinary book hand. Late in the twelfth century a semicursive Protogothic had been developing in England into a full-blown *cursiva* script, with its letters linked, loops, and other decorative embellishments, as a business hand (Gothic *littera cursiva anglicana documentaria*). From it emerges Anglicana,¹ first in England, but later as a script local to Britain and extending also into northern France.² Yet it was known as a distinctively English script and indeed a Joan Walkyngham, who died in 1346, was aware of its Englishness, for her will notes ‘quemdam librum scriptum littera Anglicana’.³ The dates for usage are from the thirteenth century to the sixteenth century,⁴ and even into the eighteenth century for some specialized purposes. About 1375 a new script, termed Secretary and originating in Italy, came from France into use in England, providing competition for Anglicana. Like Anglicana, this too was at first a documentary script (Gothic *littera cursiva Secretary documentaria*). Anglicana, like the incoming competing Secretary script, arises from within the Gothic system of scripts. Given their common ultimate origins, the two scripts share many letter-forms. A small number of features is particularly helpful for distinguishing an Anglicana from a Secretary hand: the two-compartment Anglicana **a**; both the pointed **e** and the cursive variety with reversed ductus; the tight **g**, sometimes described as shaped like the numeral 8, which looks rather like a pair of spectacles seen sideways on; the long **r**, descending below the line; the sigma-shaped **s** that looks a little like the numeral 6; **w** with its two long initial strokes completed by bows; and **x** made with two separate strokes. Also, the Tironian sign continued in use in Anglicana.

Figure 5. Example from Roberts 2005:161, the beginning of Section VII “The Gothic system of scripts: Anglicana” describing a variety of features including “the sigma-shaped s that looks a little like the numeral 6”.

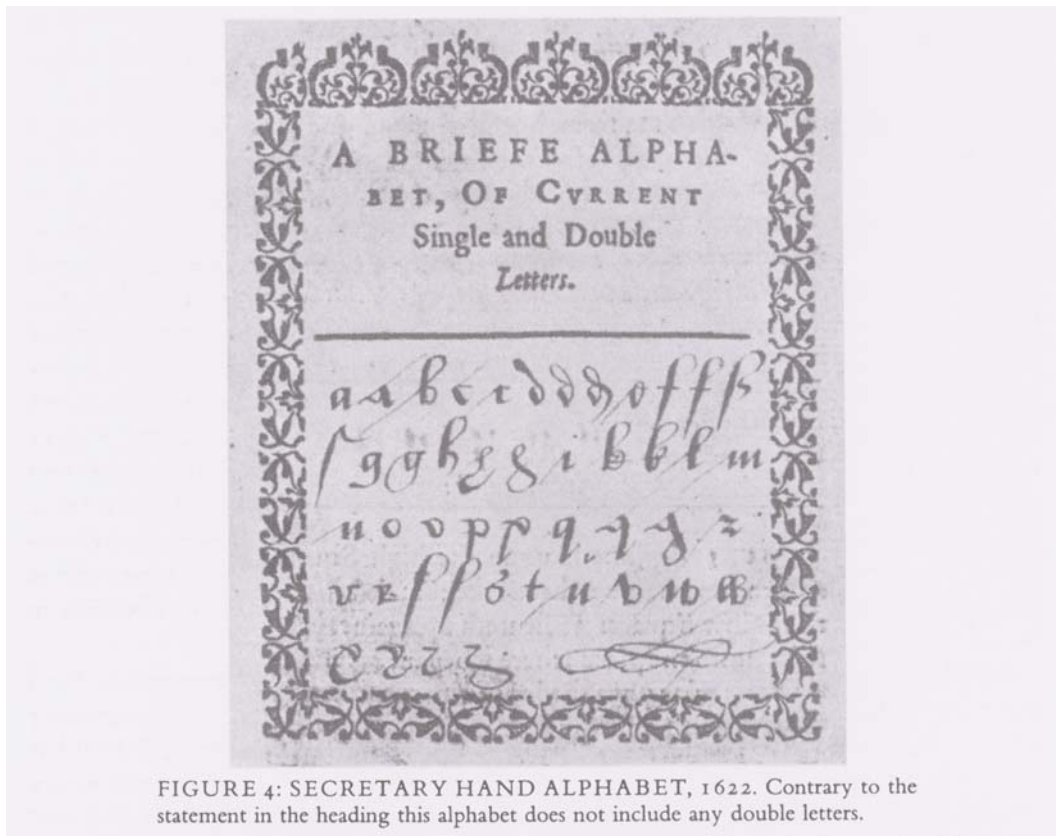


FIGURE 4: SECRETARY HAND ALPHABET, 1622. Contrary to the statement in the heading this alphabet does not include any double letters.

The alphabet shown in Fig. 4 displays some characteristic letters of the secretary hand. The letters *a* and *c* frequently carry an ‘attacking stroke’, that is, a long straight hair-stroke inclined to the right and rising high above the line (see line 1, letters 2 and 4).

For purposes of identification one of the key-letters of this script is *h*, which concludes with a steeply descending leftward curve, which sinks lower and lower as time goes on (line 2, letter 5). The *p* is usually written in one continuous action of the pen (line 3, letter 5) and can hardly be distinguished from *x*. Two varieties of *r* occur: one like an arabic-2 in appearance (line 3, letter 9), the other an open form with two stems (line 4, letter 1). The short-*s* which looks like a Greek sigma (σ) ends in a rather prominent curl above the line (line 4, letter 5). Although it is not so in this alphabet *d* and *e* may be liable to confusion, since both tend to fall apart and take the form of a curl below with a loop above.¹

Figure 6. Example from Simpson 1998:14–15 in which the secretary hand is described, along with the “short-*s* which looks like a Greek sigma (σ)”. Note the shape in which the tail of the sigma ascends; this is quite common in some hands.

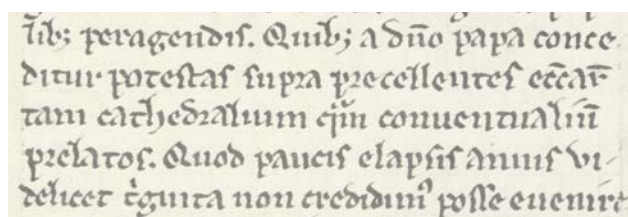


Figure 7. Example in a semi-Gothic minuscule book hand, from Denholm-Young 1964: Plate 12 (a text written by Matthew Paris, Brit. Mus. Royal MS 14C. vii, f. 150), showing LONG S used in all positions, as in *potestas supra* in line 2.

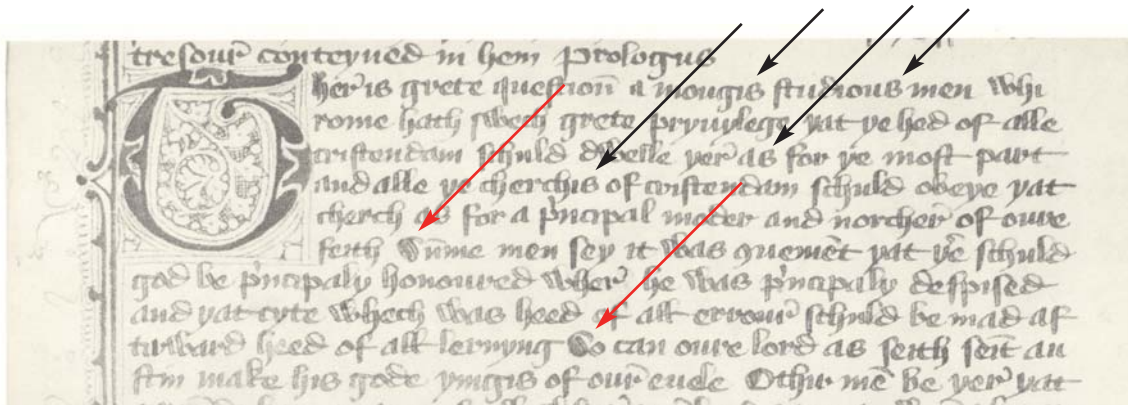


Figure 8. Example in an English vernacular hand c. 1430, from Denholm-Young 1964: Plate 22 (“The Solace of Pilgrims”, Brit. Mus. Royal MS 14C. vii, f. 150), showing Roman s with a B-like shape alongside LONG s. There are no examples of lower-case SIGMOID s in this text, though the capital S’s have the shape of No. 19 in Figure 2. Transcription:

Ther̄ is grete questiōū a mongis studios men v̄hirome hath f̄wech grete pryuylege þat
 þe hed of alle cristendum schule d̄welle þer̄ as for þe most part and alle þe cherchis of
 cristendum schule obeye þat cherch as for a p̄ncipla moder and norcher̄ of oure feith
 S̄ūme men fey it v̄as guenēt þat þe schuld god be p̄cipaly honoured v̄her̄ he v̄as
 p̄ncipaly despised and þat cyte v̄hech v̄as heed of all errour̄ schuld be mad af̄tirv̄ard
 heed of all lernyng So can oure lord as feith feit̄ austin make his gode þingis of oūr̄ ende.

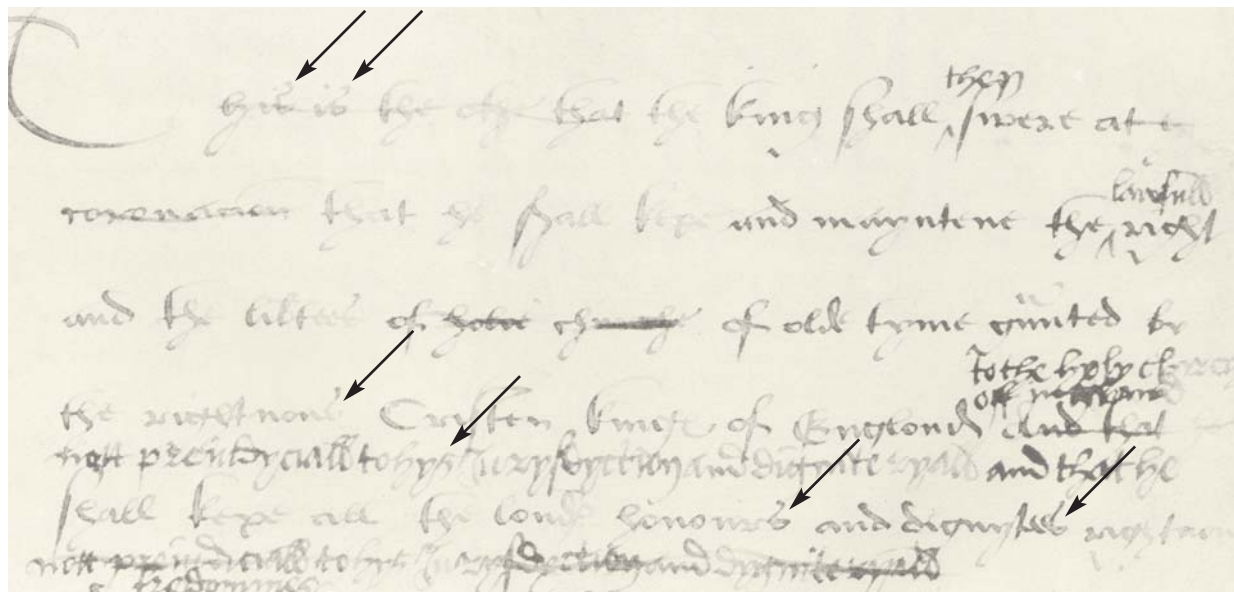


Figure 9. Example in a secretary hand c. 1509, from Denholm-Young 1964: Plate 26 (a draft of the Coronation Oath made for Henry VIII and corrected by his hand, Brit. Mus. Cott. MS Tib. E viii, f. 89), showing SIGMOID s which Denholm-Young describes as “flung s”. Transcription:

This is the othe that the king shall [then] swere at the coronacion that he shall kepe and
 mayntene the [lawfull] right and the libtees of holie church of old tyme gūnted by the
 rightuous Cristen kinḡ of Englonde And that he [to the holy church of ingland not
 preiydycciall to hys Iurydyccion and dignite ryall and that he] shall kepe all the lond̄
 honours and dignytys rightuou[s]. [nott preiudiciall to hys Iuryisdiction and dygnite ryall]

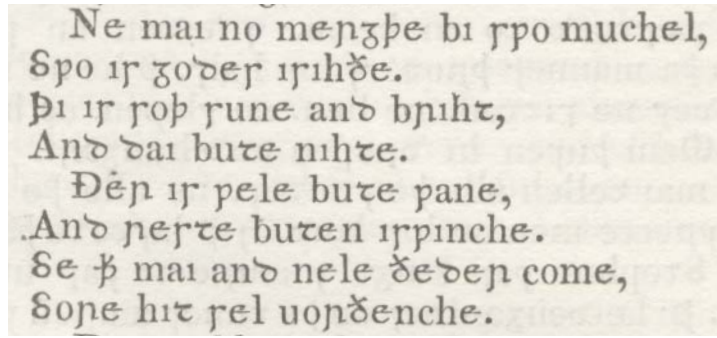


Figure 10. Example from Johnson 1773:11, showing the glyph for SIGMOID s in the “Saxon 2” font cut by Edmund Fry. Fry of course had cut insular letterforms and in his own context (very different from that of the UCS) he paired \mathcal{S} and \mathfrak{s} .

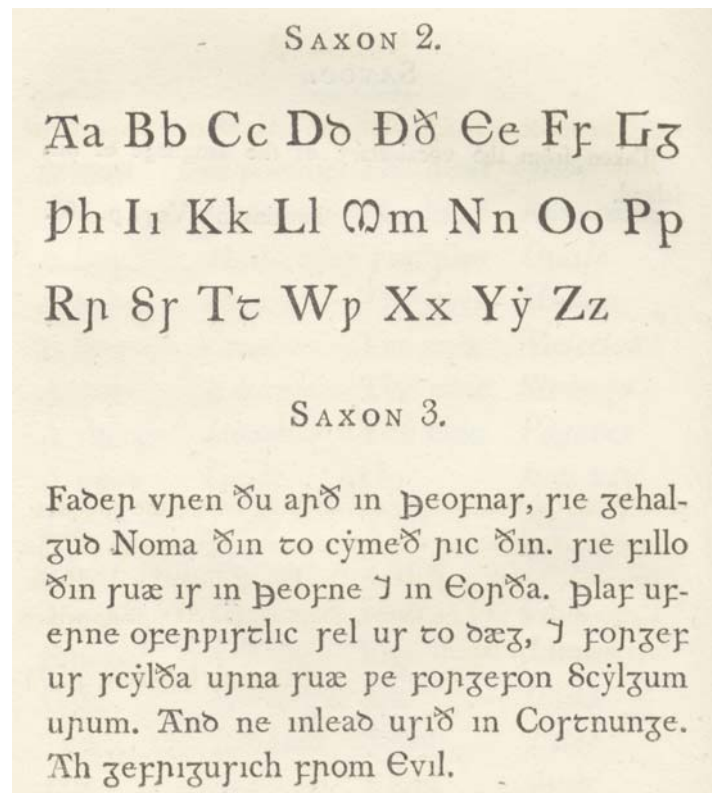


Figure 11. Example from Fry 1799:260, showing the glyph for SIGMOID s in the “Saxon 2” and “Saxon 3” fonts cut by Edmund Fry. Figures 10 11 are shown for the typographic form of the CAPITAL SIGMOID S.

\mathcal{S}	<code>&Sclose;</code>	F126	PUA-var	LATIN CAPITAL LETTER S CLOSED FORM
\mathfrak{s}	<code>&sclose;</code>	F128	PUA-var	LATIN SMALL LETTER S CLOSED FORM

Figure 12. Example from the MUFI Version 3.0, showing the characters for SIGMOID s given the name “s closed form”.

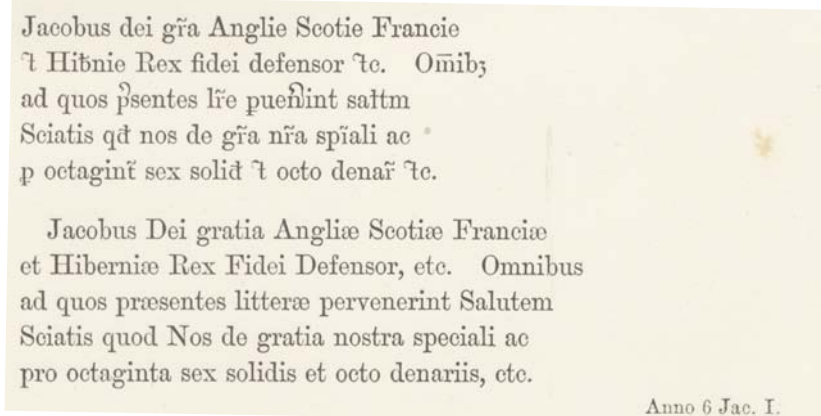
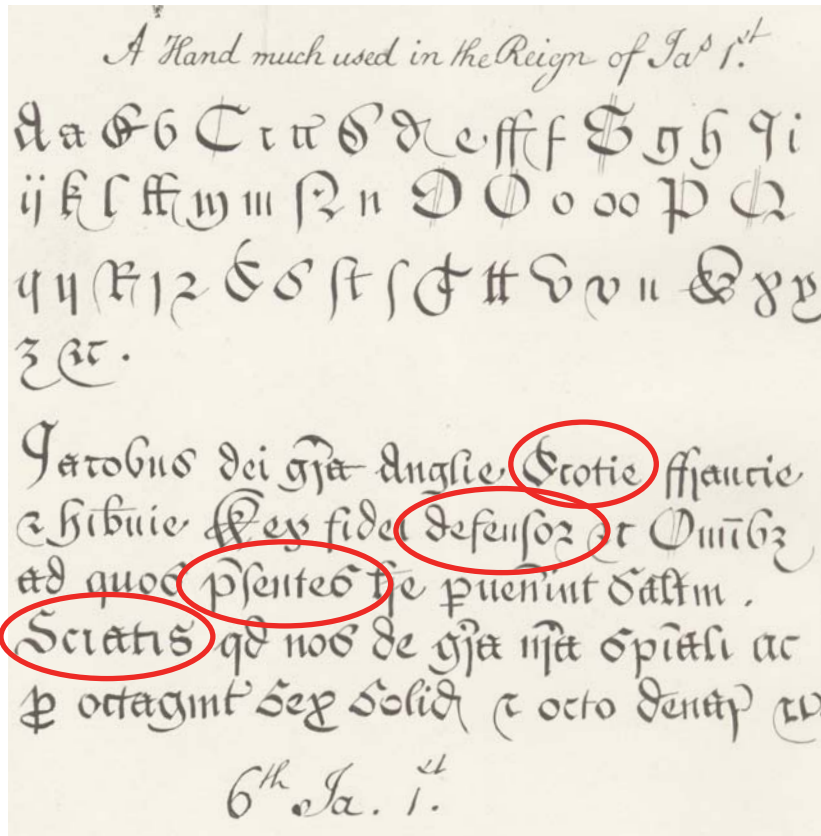


Figure 13. Example from Wright 1897: “A Hand much used in the Reign of James I” (of England, 1566-1625). The printed transcription is given in the “Record type” devised in the 1770s for the printing of the Domesday book. Of interest in palaeographic transcription is the distinctions found in “Scotie”, “defenfor”, “pŕsentes”, and “Sciatis”.

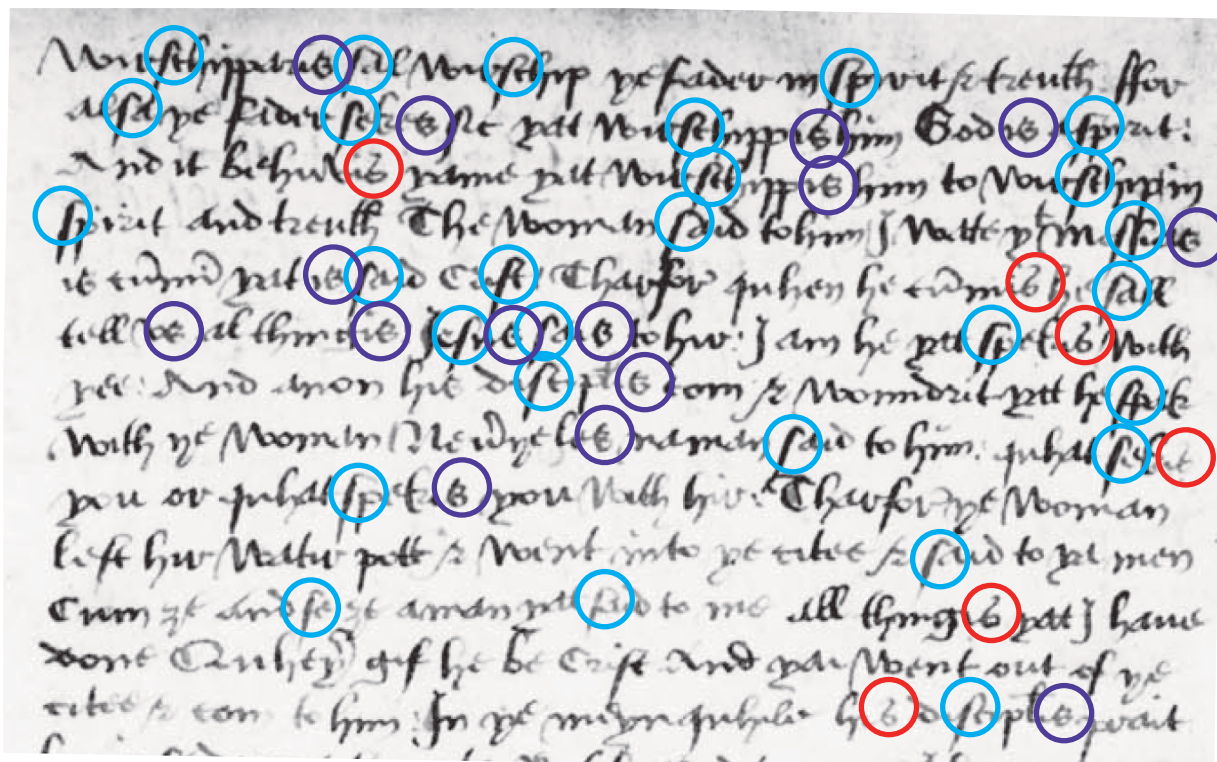


Figure 14. In the Gospel of John 4:23–31 (Egerton MS 2880, f. 91v). Here some final s's are written by Nisbet with (B-shaped) LATIN SMALL LETTER S and some are written with LATIN SMALL LETTER SIGMOID S, alongside LATIN SMALL LETTER LONG S as usual. The palaeographic transcription of this passage (with punctuation and quotation marks added) follows:

²³[Bot þe tyme is cūmīd 7 now it is quhen tzew] wirfchippazis fal wirfchip þe fader in fpirit 7 tzeuth: For alfa þe fader fekes fic þat wirfchippis him ⁰⁴⁻²⁴God is a fpizit: And it behuvis þame þat wirfchippis him to wirfchip in fpizit and tzeuth

⁰⁴⁻²⁵The woman faid to him, I wate þ^t Meffias is cūmīd þat is faid Czift: Tharfor quhen he cūmis he fall tell vs al thingis:

⁰⁴⁻²⁶Iefus fais to hir: I am he þat fpekis with þee:

⁰⁴⁻²⁷And anon his difciplis com 7 wonndzīt þat he fpak with þe woman; neūþeles na man faid to him: quhat fekis þou or quhat fpekis þou with hir:

⁰⁴⁻²⁸Tharfor þe woman left hir watir pott 7 went into þe citee 7 faid to þa men ⁰⁴⁻²⁹Cum ze and fe ze a man þat faid to me all thingis þat I haue done Quheþ gif he be Czift: ⁰⁴⁻³⁰And þai went out of þe citee 7 com to him:

⁰⁴⁻³¹In þe meynquhile his difciplis prait [him 7 faid: Maifter etc.]

A. Administrative

1. Title

Proposal to add two SIGMOID S characters for mediaeval palaeography to the UCS

2. Requester's name

Michael Everson

3. Requester type (Member body/Liaison/Individual contribution)

Individual contribution.

4. Submission date

2020-10-01

5. Requester's reference (if applicable)

6. Choose one of the following:

6a. This is a complete proposal

Yes.

6b. More information will be provided later

No.

B. Technical – General

1. Choose one of the following:

1a. This proposal is for a new script (set of characters)

No.

1b. Proposed name of script

1c. The proposal is for addition of character(s) to an existing block

Yes

1d. Name of the existing block

Latin Extended-D

2. Number of characters in proposal

2.

3. Proposed category (A-Contemporary; B.1-Specialized (small collection); B.2-Specialized (large collection); C-Major extinct; D-Attested extinct; E-Minor extinct; F-Archaic Hieroglyphic or Ideographic; G-Obscure or questionable usage symbols)

Category A.

4a. Is a repertoire including character names provided?

Yes.

4b. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?

Yes.

4c. Are the character shapes attached in a legible form suitable for review?

Yes.

5a. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?

Michael Everson.

5b. If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:

Michael Everson, Fontographer.

6a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?

Yes.

6b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?

Yes.

7. Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)?

Yes.

8. Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at <http://www.unicode.org> for such information on other scripts. Also see Unicode Character Database <http://www.unicode.org/Public/UNIDATA/UnicodeCharacterDatabase.html> and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

See above.

C. Technical – Justification

1. Has this proposal for addition of character(s) been submitted before? If YES, explain.

No.

2a. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?

No.

2b. If YES, with whom?

2c. If YES, available relevant documents

3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?

Germanicists, Anglicists, Celticists, dialectologists, lexicographers, and Scots.

4a. The context of use for the proposed characters (type of use; common or rare)

Common in Middle Cornish, Middle English, and Middle Scots.

4b. Reference

5a. Are the proposed characters in current use by the user community?

Yes.

5b. If YES, where?

Various publications.

6a. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?

Yes.

6b. If YES, is a rationale provided?

Yes.

6c. If YES, reference

Accordance with the Roadmap. Keep with other Latin characters.

7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?

No.

8a. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?

No.

8b. If YES, is a rationale for its inclusion provided?

8c. If YES, reference

9a. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?

No.

9b. If YES, is a rationale for its inclusion provided?

No.

9c. If YES, reference

10a. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?

Yes.

10b. If YES, is a rationale for its inclusion provided?

Yes.

10c. If YES, reference

Discussuon of the palaeographic usage is given above.

11a. Does the proposal include use of combining characters and/or use of composite sequences (see clauses 4.12 and 4.14 in ISO/IEC 10646-1: 2000)?

Yes.

11b. If YES, is a rationale for such use provided?

No.

11c. If YES, reference

11d. Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?

No.

11e. If YES, reference

12a. Does the proposal contain characters with any special properties such as control function or similar semantics?

No.

12b. If YES, describe in detail (include attachment if necessary)

13a. Does the proposal contain any Ideographic compatibility character(s)?

No.

13b. If YES, is the equivalent corresponding unified ideographic character(s) identified?