Basic Latin brevigraphs listed in *Polonia Typographica Saeculi Sedecimi* — progress report

Janusz S. Bień

1 Introduction

The fonts of several 16th century printers active in Poland, namely Aleksander Augezdecki, Jan Haller, Kasper Hochfeder, Florian Ungler (the first and second printing house) and Maciej Wirzbięta, have been described in the series of 12 fascicules entitled *Polonia Typographica Saeculi Sedecimi* published in years 1959–1981.[1] Almost all of them are digitized but available only for "digital lending" in ACADEMICA[2] because, I surmise, it's not clear who owns the copyright as this was a collaborative effort of several persons (only one contributor is still alive) and institutions.

In the fascicules every font is illustrated by an excerpt of a text and sometimes additionally by a table of the character set; an example of such a table is presented in Fig. 1. Most of the tables have been prepared by Maria Błońska with some help from Anna Wolińska and Henryk Bułhak (the editor of several fascicules); some tables were prepared by Anna Śliwa, Alodia Kawecka Gryczowa (also the editor of several fascicules) and Paulina Buchwald-Pelcowa (the editor of the whole series). The number of font tables is over seventy and the total number of glyphs in the tables is over six thousand.

Unfortunately I missed the opportunity to get first-hand information on how the tables were prepared when talking by phone with Paulina Buchwald-Pelcowa in 2022 (she died two years later). I got some rudimentary information from Henryk Bułhak, also in phone calls, but he was able to provide me only with rather general information: the glyphs were cut out with a razor blade from photocopies and pasted together. This information seems relevant because it shows what kind of mistakes can be expected in the tables: if a character occurs in a table then it can be displaced or misassigned (see sec. 7), but definitely exists in a text; on the other hand, some omissions are possible. For example, the glyph in Fig. 2 is not listed in the table in Fig. 1, but can be found in the texts printed reportedly with this font; according to Peter Baker, the meaning is *cis*.[3]

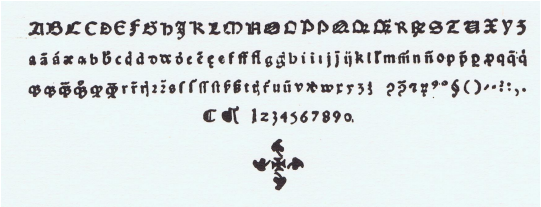


Figure 1: Ungler's second printing house font no 16 (a fragment of Plate 359)

Figure 2: The letter or the ligature *cis*? Cf. Fig. 40.

Figure 3: A fragment of Plate 168. The fourth glyph is interpreted as *h*, not a *b*, because of its position in the font table. See sec. 8

Figure 4: A fragment of Plate 357. The last glyph is interpreted as *e* with ogonek because of its position in the font table.

Figure 5: A fragment of Plate 411. The third glyph is interpreted as *h* because of its position in the font table.

No comments to the tables are provided, but the order of glyphs is sometimes relevant for their interpretation (see Figures 3, 4, 5).

The quality of the glyph images is sometimes quite low; I understand no better samples were available.

Early prints used many abbreviations which were the descendants of the abbreviations used in manuscripts; we discuss here a subset of them, called brevigraphs. Quite often they consisted of a regular non-modified letter supplemented by a diacritical mark, usually a macron or a similar glyph, placed above. It is natural to call them composed brevigraphs. On the other hand there are abbreviations in a shape of a modified letter or a letter-like symbol; we call them basic brevigraphs even if they are accompanied by a diacritical mark.

The work described here consisted primarily of creating computer indexes to allow comparison of similar characters from the same or different fonts

[1] The first two fascicules were published in 1936 and 1937, but we are interested in their second revised editions because these were supplemented by the character set tables.

[2] *Interlibrary loan system of books and scientific publications*: https://academica.edu.pl/

[3] github.com/psb1558/Junicode-font/discussions/255

doi.org/10.47397/tb/45-1/tb139bien-brevigraphspt

Figure 6: Comparing characters with djview4poliqarp

(Fig. 6). The indexes and other resources are available in a public repository;[4] the repository site provides *Issues* and *Discussions* tabs for reporting mistakes and giving comments.

The paper is considered a progress report for two reasons. First, the character indexes should be supplemented by indexes showing their meaning and their use in texts, in a similar way as described in [5]. Secondly, I omitted some interesting glyphs because I was not sure how to interpret them. An example is presented in Fig. 7: is the last but one glyph a modification of the letter *h*, or is this just the letter *h* with a diacritic mark which happens to touch the letter?

Figure 7: A fragment of Plate 21. Is the last but one glyph a basic brevigraph?

An important question for every basic brevigraph is whether it has been assigned a codepoint in the Unicode standard.[5] Checking the character charts is unfortunately not sufficient for two reasons. First, the character name is not intended to provide the full information about the character, it should be treated as a more or less arbitrary label. Secondly, in principle (the practise is sometimes questionable) the standard defines characters, not glyphs, and the glyph in a chart is only one of the possible representations of the character (an example is given in sec. 9). In consequence it is useful to also look up the character proposals and related documents in the Unicode Technical Committee Document Registry[6] (a similar resource is the ISO/IEC JTC1/SC2/WG2 register[7]). It is also useful to look for alternative glyphs in specialized fonts (see sec. 12).

[4] github.com/jsbien/early_fonts_inventory
[5] home.unicode.org/
[6] unicode.org/L2/
[7] unicode.org/wg2/

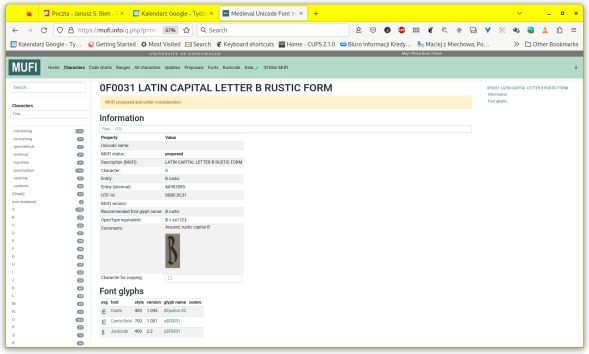


Figure 8: MUFI LATIN CAPITAL LETTER B RUSTIC FORM assignment proposal as of 2024-01-01

Another interesting question is whether the brevigraph has been assigned a codepoint in the Unicode Private Use Area by the Medieval Unicode Font Initiative.[8] The MUFI assignments at first, up to version 4.0, were published as recommendations in the form of PDF documents [11]. They list over 1500 pure Unicode and Private Use Area characters in the Latin alphabet of potential use for the encoding of old text sources.

Nowadays, the recommendations have the form of a database which can be browsed online (Fig. 8). For some time a subset of the data content is also available for download, under the Creative Commons Attribution-ShareAlike 4.0 license, in the form of a CSV or JSON file.

We will use here also the resources of *Projet d'Inventaire des Caractères Anciens*[9] created and maintained by Jacques André.

Last but not least, it is important whether a brevigraph can be rendered adequately by a font. Our primary focus is on Peter Baker's Junicode Two font, as it is available under a free license[10] and contains some characters not available elsewhere [7]. We also use George Douros' Symbola font[11] for some characters not available in Junicode.

It is worth mentioning that some of the brevigraphs discussed here were used in Gutenberg's bibles. The character set of these books has been the subject of several publications; they are referenced in [2] and [6]. I also found very useful the unpublished text [4] kindly provided to me by the author. (It is attached by the Printing Museum in Lyon to digital copies of a folio of the Gutenberg bible purchased by the visitors.)

[8] mufi.info/
[9] jacques-andre.fr/PICA/
[10] github.com/psb1558/Junicode-font
[11] dn-works.com/ufas/

2 The workflow

It is an old idea of mine to use the fact that for compression purposes the identical or similar shapes are collected into shape dictionaries. It seems that this approach, named mixed raster content, is used now in JPEG2000, but the first format to use it successfully was DjVu.

I designed two tools which are based on this approach. The first one was a quick and dirty modification of a standard DjVu viewer (Fig. 9). It was originally implemented by Michał Rudolf twelve years ago, with important contributions made later by Alexander Trufanov.[12] It is quite good for getting a quick overview of the shapes in a document, but it is not convenient enough for analysing them in detail.[13]

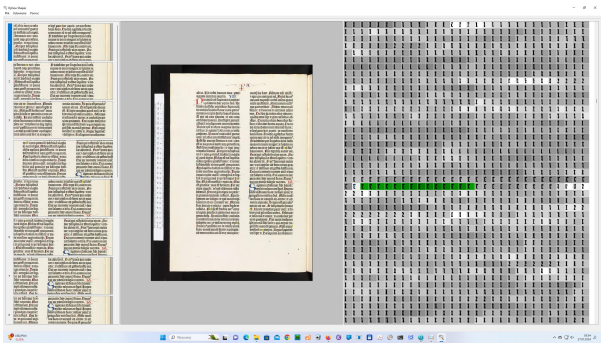


Figure 9: djview-shapes and Gutenberg's bible

The second tool was a sophisticated client-server system. The idea was that a database will store shapes from different documents provided by different persons or institutions and accessed remotely by interested users. The shapes were exported to a MySQL database. Unfortunately the client[14] was a complete failure, since due to some wrong coding decisions it was prohibitively slow. There was neither opportunity nor sufficient motivation to reimplement it in a better way.

So when working on the present paper my main tool was Alexander Trufanov's `djvudict` program[15] which dumps a DjVu shape dictionary in an almost human-readable form, despite the fact that the program does not seem to be fully reliable (e.g., for some not yet known reasons some shapes are skipped).

The first step was accessing the scans of *Polonia Typographica* and preparing (with Gimp) the images of the relevant tables. Then the images were converted to DjVu (with Friedric Foebel's Python 3 fork of didjvu[16]) and supplemented by appropriate metadata. Next they were processed by djvudict. The primary DjVu file names have a form like `Augezdecki-01a_PT08_403.djvu`, where `Augezdecki` is the name of the printer, `01a` is the font number sometimes supplemented by a letter, `PT08` is the identification of the *Polonia Typographica* fascicule and `403` is the plate number (they are numbered continuously in the whole series; many contain woodcuts of no interest to us). The djvudict output is placed in the directories with shorter names, as in `Augezdecki-01a`.

A quick and dirty Python program (written, or rather put together from pieces of code found on the Internet, by myself) converts the djvudict output to an index for the djview4poliqarp program (described already in [5] and [8]); the shape identifiers are preserved. The index contains also the results of OCR processing done with *Tesseract*, but at present, due to the lack of appropriate training, they are of essentially no use. The file names are in the form `Augezdecki-01a.csv`.

The indexes unfortunately require some hand editing with djview4poliqarp. The first stage is to create an index named like `Augezdecki-01a_tmp.csv`, where the interesting elements are marked with '`+`' in the so-called comment field. Entries are marked with '`#`' when there is a need to adjust the bounding box; it is not yet clear why this is sometimes needed. Entries marked with '`^`' also require adjusting the bounding box, but for a different reason: the shapes recognized by the DjVu compression algorithm are just connected components, so diacritics are usually separate objects.

The files `*_tmp.csv` are processed with grep to put the marked entries into the indexes named `*_work.csv`, where the bounding boxes are adjusted if needed. The files form the basis for the intermediate brevigraph indexes named `*_workbr.csv` where the entries are supplemented by the brevigraph names.

The brevigraph names serve a purely technical goal: they allow grouping similar brevigraphs together in the djview4poliqarp program. However, the choice of the names is not obvious. The official Unicode names and the Unicode-like MUFI names are cumbersome because of their length, e.g., LATIN CAPITAL LETTER V WITH DIAGONAL STROKE. As an alternative, I was considering using names derived from the XML entity names provided by MUFI, also for selected pure Unicode characters. Some are

[12] `github.com/jsbien/djview4shapes`

[13] When the present paper was almost finished, some changes were made to the program which make it much more useful.

[14] `github.com/jsbien/ndt/wiki/z_shapes`

[15] `github.com/trufanov-nok/djvudict`

[16] `github.com/FriedrichFroebel/didjvu`

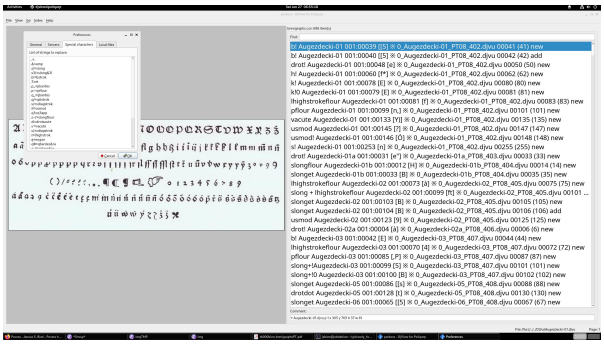


Figure 10: Keyboard shortcuts in djview4poliqarp

short and mnemotechnical, e.g. `&pbardes;` (U+A751 latin small letter P with stroke, here BAR, through DEScender). Some are short and not mnemotechnical, e.g., `&q3app;` (U+E8B4 LATIN SMALL LETTER Q LIGATED WITH FINAL ET; 3 may suggest the shape of the final *et*, but I have no association to suggest for app), some are mnemotechnical but rather long, e.g. `&lhighstrok;` (U+A749 latin small letter L with HIGH STROKe).

Nevertheless, after some hesitation, I decided to use those names (with '`&`' and '`;`' stripped) for my purposes. The crucial factor in making this decision was the fact that djview4poliqarp has a kind of macro facility (Fig. 10). The configuration file[17] has an `[edit]` section which can contain appropriate settings.

For characters which are present neither in Unicode nor in MUFI I use *ad hoc* names. Characters which are difficult to identify I handle in an analogical way. Such names often don't identify the glyphs uniquely, but merely point to similar glyphs.

For technical reasons the names are placed first in the comment field and then moved to the entry field with a program. The final indexes have the names in the form of `*_br.csv`. An aggregated index is also created which is named simply `brevigraphs.csv`.

The figures in the present paper were prepared in a way similar to that used for [8]: a program converts the index of the selected glyphs into the expex[18] code and creates a set of graphic snippets from the DjVu documents.

The glyphs in the figures are numbered for reference purposes and accompanied by the self-explanatory abbreviations of printing house names and font numbers.

[17] `~/.config/djview-poliqarp/djview-poliqarp.conf`
[18] `ctan.org/pkg/expex`

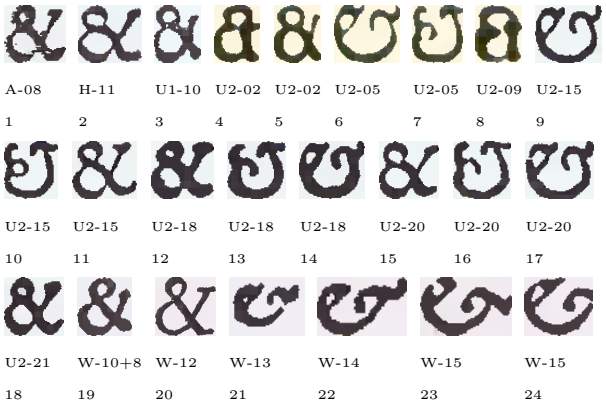


Figure 11: Ampersand

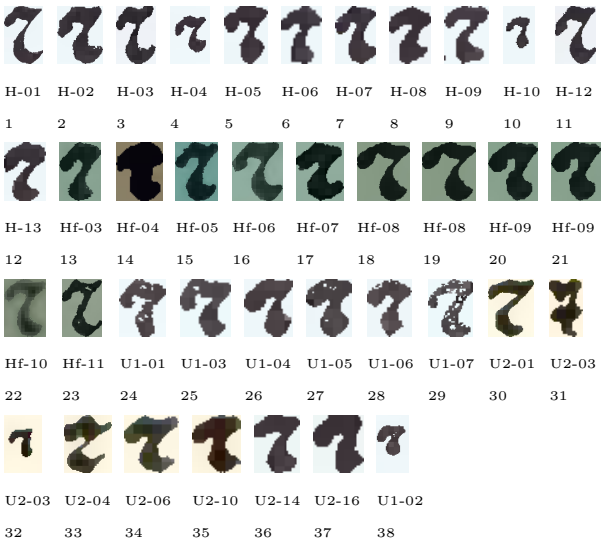


Figure 12: Tironian note *et*

3 Non-alphabetic brevigraphs

Figure 11 shows ampersand, the brevigraph which in one of its forms has survived to the present time; it is a very old abbreviation of the word *et*. It has two forms, both of them are available in the Junicode family of fonts: '&' (Junicode-Regular) and '*&*' (Junicode-Italic). In computer code the first form has been available at least since ASCII (*American Standard Code for Information Interchange*), which was created in 1963. The Unicode charts also show only the first form.

The brevigraph presented in Fig. 12 is without any doubt the Tironian note *et* (Tironian notes are named after Tiro, the secretary of Cicero, who is credited with inventing them), used always as a separate word. The brevigraph is present in Unicode since version 3.0 (published in 1999) as U+204A `TIRONIAN SIGN ET` with the canonical glyph '⁊'. The Junicode

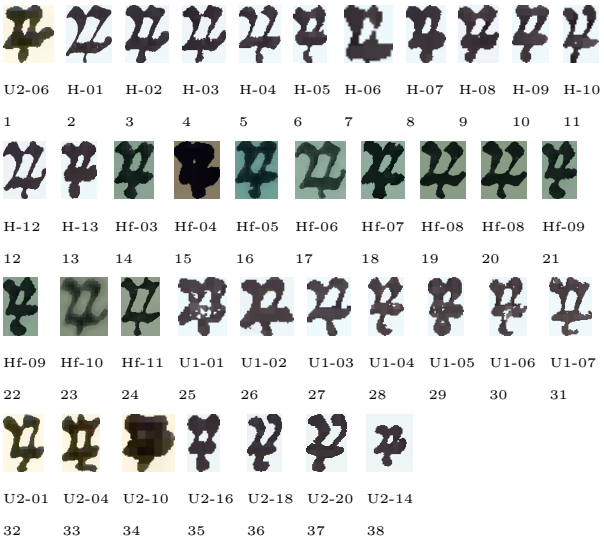


Figure 13: The letter *rum* rotunda

font has also another variant of the glyph, namely '' (accessed with OpenType character variant feature or just with the code U+F001D), which is quite close to the shape of most glyphs in Fig. 12. Item 31 in the figure is yet another variant, called in the font manual *Tironian et sign later form with bar*; it is available in Junicode with OpenType feature *ss10* and the tags '[a][ss]'[1][ss]': ''.

The brevigraph presented in Fig. 13 is present in Unicode since version 5.1.0 (published in 2000) as U+A75D LATIN SMALL LETTER RUM ROTUNDA with the canonical glyph '', which is quite close to the glyphs in the figure. It can mean *-rum* or *-rom*. Although the name may suggest that this is a variant of the letter *rum*, their shapes have little in common (see sec. 12).

As noted in [1, p. 130], Unicode has additionally two similar symbols: U+0264 JUPITER, U+1F729 ALCHEMICAL SYMBOL FOR TIN ORE. In the Symbola font these three characters look like, respectively: ♃, ♃ and ♃. They look different, but this is the decision of the contemporary font designer. The glyphs listed in Fig. 13 could probably represent any of those three characters; this has to be checked in the original texts.

All the glyphs in Fig. 14 are in my opinion variants of the character U+A76D LATIN SMALL LETTER IS, present in Unicode since version 5.1.0 (published in 2000) with the canonical glyph ''. Of course their usage should be verified in the original texts.

The glyphs in Fig. 15 looking like the digit *9* are instances of the well-known brevigraph present in Unicode as U+A770 MODIFIER LETTER US with representative glyph '℘'; it was introduced in version

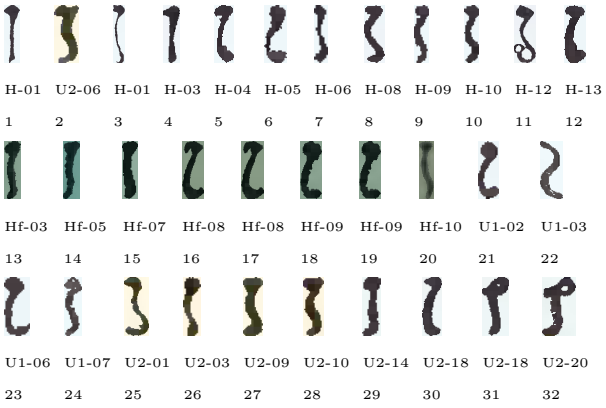


Figure 14: The letter *is*

Figure 15: The letter *us*

5.1.0 (published in 2008), and *modifier* means it is not on the baseline. It is used always at the end of words. The meaning of the glyphs similar to a circle, like item 43, is to be checked in the texts, as it can be just a raised small letter *o* (in Unicode, U+1D52 MODIFIER LETTER SMALL O).

The base glyphs in Fig. 16, despite a slightly different shape, can be identified with the character called by MUFI LATIN ABBREVIATION SIGN SMALL CON [11, s. 29] and in Unicode unified with U+2184

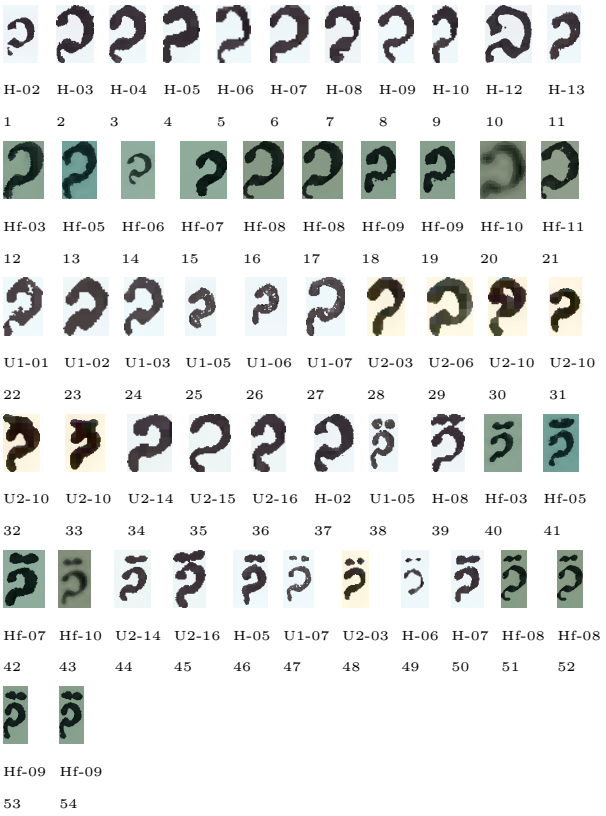


Figure 16: The letter small *con*

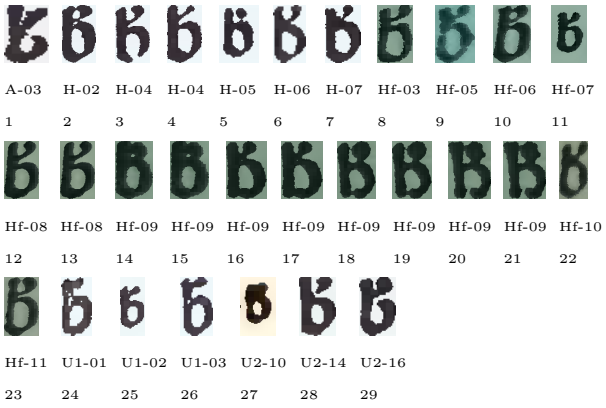


Figure 17: The variants of the letter *b*

Figure 18: Modified letter *b* in Gutenberg's bible: *sublime* and *substantia* (Bodleian Library copy, page 21 and 64 of volume II)

LATIN SMALL LETTER REVERSED C with representative glyph 'ɔ'. As the name suggests, it meaning is just *con* (perhaps with some exceptions).

As for the letter with diacritics, the situation is much more complicated. I have not yet found their occurrences in the texts, so I don't know their meaning. Another problem is the form of the diacritics. Besides a diaresis, we have the diacritic which seems to be the same as the one described in [10] as *jagged horizontal line* which is encoded in Unicode as U+1DD3 COMBINING LATIN SMALL LETTER FLATTENED OPEN A ABOVE but rendered differently: in Junicode it is '◌ᷓ' and in Symbola it is ◌ᷓ. Moreover there is an open question what kind of diacritics, if any, are used in Ungler's font 10 (items 32 and 33).

4 Modifications of the letter *b*

Old texts used many variations of the letter *b*, many of which are assigned code points by MUFI. Many variants of the letter *b* are also listed in *Polonia Typographica*. Fig. 17 presents those instances which are directly relevant to our purposes here.

We will focus on item 3 (Haller's font no 4) and those from Hochfeder's fonts, items 8–17, as their shapes seem to be carefully designed while other items seem to be just more or less crude variations of those.

A character with an almost identical shape appeared already in Gutenberg's 42-line bible. Despite this, it seems it still has no name and even no generally accepted description. In [2, p. 12] Jacques André proposes the name *latin small letter b with flourish* (he considers also an alternative *latin small letter b ligated with arm of latin small r*, but cf. sec. 8).

According to Gerald Bettridge [4], it means *bis* and, after the long *s*, *ub* (see Fig. 18, also [2, p. 12] and [3, p. 12]). It can be ligated with long *s*; see sec. 13.

It seems this was not always a brevigraph, sometimes it is just equivalent to a normal *b* [6, p. 8]. Or perhaps it was just a printer's mistake?

5 Modifications of the letter *d*

The similarity of items 6 and (e.g.) 19 to item 3 from Fig. 17 and items 5 and (e.g.) 8 from Fig. 22, all from respectively the same fonts, seems to be a design decision.

I think this is the brevigraph called *d with two ascenders* by Erin Blake in [10]; she states that the brevigraph stands for

> *de* and (depending on the language) also for *der*, *dis*, *dum* and other *d*-syllables

Peter Baker suggested[19] treating the character as MUFI U+F159 LATIN ABBREVIATION SIGN SMALL

[19] github.com/psb1558/Junicode-font/discussions/133

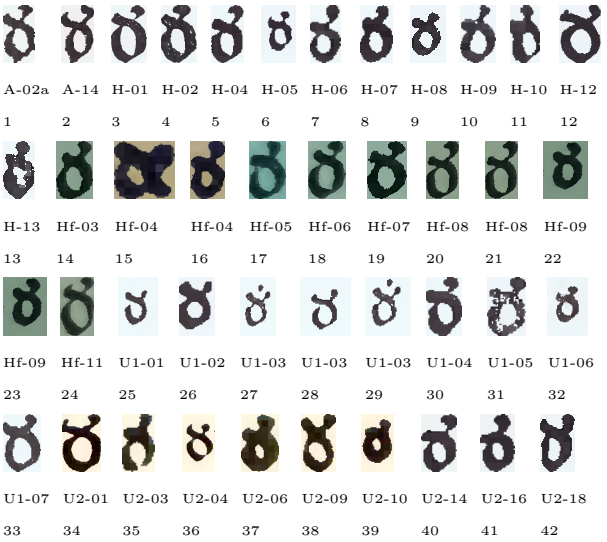


Figure 19: The variants of the letter *d*

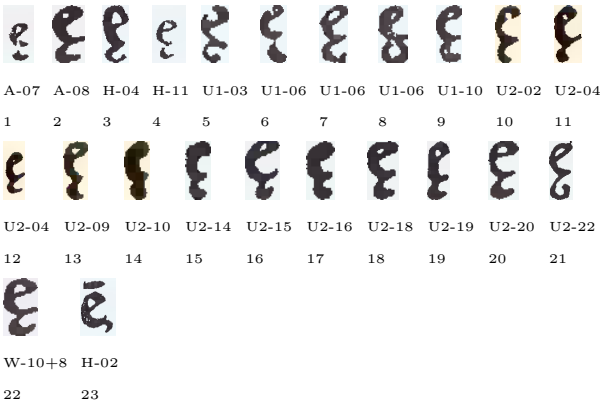


Figure 20: The variants of the letter *e*

DE ('ȸ'), called also LATIN SMALL LETTER D ROTUNDA WITH BAR.[20] He also points to another similar MUFI character, namely U+EBB2 LATIN SMALL LETTER D ROTUNDA WITH ACUTE 'ȸ'.

The meaning of the brevigraph with a dot above is yet to be checked in the texts.

6 Modifications of the letter *e*

Fig. 20 presents the well-known *e caudata*, meaning *ae*.

It is an open controversy whether *e caudata* and the contemporary U+0119 LATIN SMALL LETTER E WITH OGONEK should be considered the same character. Peter Baker wrote[21]

> Perhaps it is time to admit that the Latinate cauda and the ogonek used by Polish and other languages are different beasts.

and provided an OpenType feature (*ss15*) to distinguish them in the Junicode font. I have no opinion on this matter.

The meaning of the letter with a bar above (item 19) is yet to be checked in the texts.

Figure 21: The variants of the letter *g*

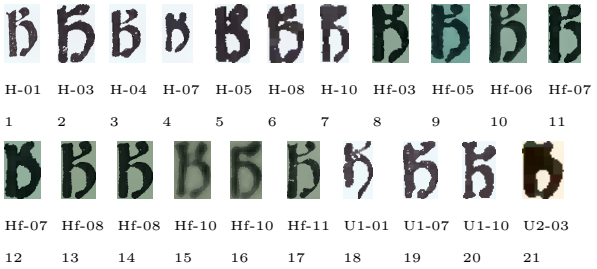


Figure 22: The variants of the letter *h*

7 Modification of the letter *g*

I have little to say about the glyphs in Fig. 21, since I have not found any occurrence of them in a text. On one hand they resemble the letters *rum* (sec. 12), and *tum* (sec. 14). On the other hand it resemble also the glyph from Fig. 2. Moreover, Blake [10] says that *weird vertical line at end of word* means an *s* preceded by a vowel (typically *es* in English and *is* in Latin); in other words in Latin it can be perhaps considered as the ligature of *g* and the letter '*is*' which has been discussed already in sec. 3.

8 Modifications of the letter *h*

For some fonts there is an evident similarity of items among Figures 17, 19, and 22. It seems to be a design decision.

A character with an almost identical shape appeared in Gutenberg's 42-line bible. In the MUFI recommendation, it is identified as U+E8C3 LATIN SMALL LETTER H LIGATED WITH ARM OF LATIN SMALL LETTER R ('ℏ'). Jacques André [2, p. 17] notes that the name is strange and I agree with him.

The glyphs in Fig. 22 seem to be the same as those described as *h with a tick on top* by Blake, who states

> Stands for *h*-syllables like *han*, *het*, and *hic*

[20] `mufi.info/q.php?p=mufi/chars/unichar/61785`
[21] `junicode.sourceforge.net/ecaudata.html`

Figure 23: The variants of the letter *l*

Another interpretation was presented by Lisa Howarth on the Facebook The Paleography Society group:[22]

> When attached to an 'h', it usually means 'er' or 'ab' depending on the word

She considers the glyph to be composed from the letter *h* and an diacritical sign, similar to Peter Baker, who identifies[23] the diacritics as U+0335 COMBINING SHORT STROKE OVERLAY ('◌̵').

9 The modification of the letter *l*

Reportedly the glyphs in Fig. 23 have the same meaning as U+A749 LATIN SMALL LETTER L WITH HIGH STROKE ('ꝉ') and therefore a separate code point has not been assigned.[24] However the Junicode font contains at code point U+F000F the glyph *l with high stroke ending with flourish* ('ﰏ'), accessible also as *l* with the OpenType feature ss10 and the tags '[ss10]'[ss10]'.

A character with an almost identical shape appeared in Gutenberg's 42-line bible; cf. Fig. 24.

Figure 24: Modified letter *l* in Gutenberg's bible: according to [4] *iherusalem* (Bodleian Library copy, page 574 of volume II)

10 Modifications of the letter *p*

Fig. 25 contains the variants of a well-known brevigraph, available in Unicode since version 5.1.0 (published in 2008) as U+A751 LATIN LETTER P WITH STROKE THROUGH DESCENDER with representative glyph 'ꝑ'. The brevigraph is ambiguous; the most popular meanings are *per*, *par* and *por*. It can be used as an individual word or as a prefix.

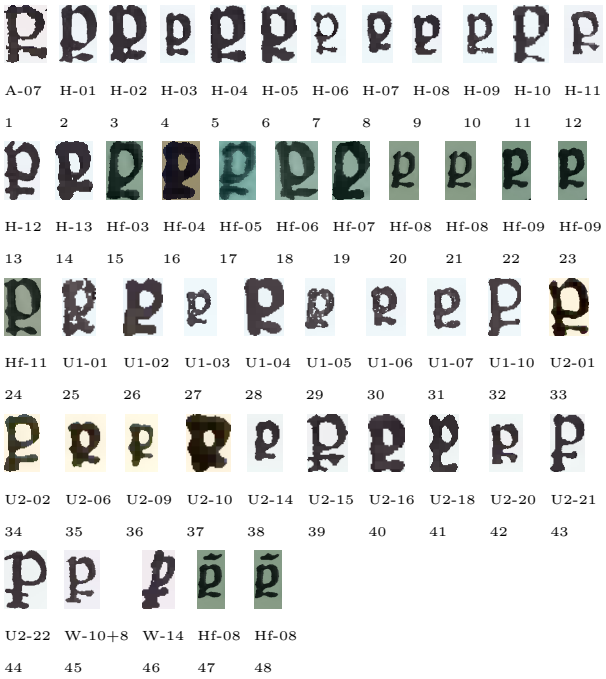


Figure 25: The letter *p* with stroke

The base characters in Fig. 26 are also the variants of a well-known brevigraph, available in Unicode since version 5.1.0 (published in 2008) as U+A753 LATIN LETTER P WITH FLOURISH with representative glyph 'ꝓ'. The brevigraph is ambiguous; the most popular meanings are *pro* and *por*. It too can be used as an individual word or as a prefix.

The last characters are included in the MUFI recommendation as U+EED7 LATIN SMALL LIGATURE PP WITH FLOURISH with the glyph 'ﰗ'; the meaning is *prop-*.

The meaning of the characters with diacritics is yet to be checked in the texts.

11 Modifications of the letter *q*

The glyphs in Fig. 27 represent a well-known brevigraph, included in Unicode since version 5.1.0 (published in 2008) as U+A757 LATIN SMALL LETTER Q

[22] facebook.com/groups/7687162686/permalink/10158299890607687

[23] github.com/psb1558/Junicode-font/discussions/134

[24] github.com/psb1558/Junicode-font/issues/4

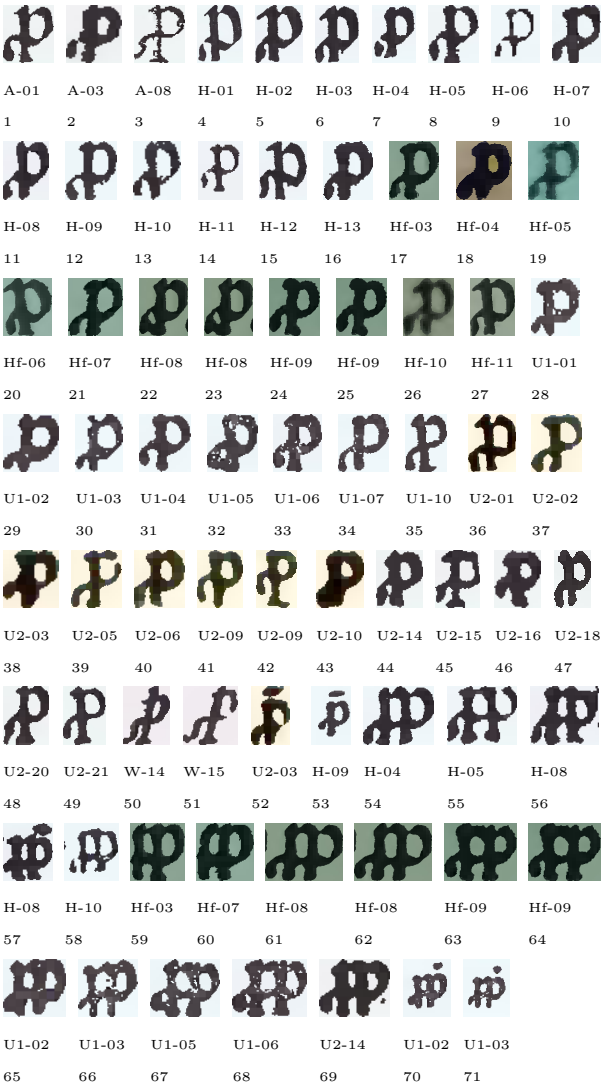


Figure 26: The letter *p* with flourish

Figure 27: The letter *q* with stroke through descender

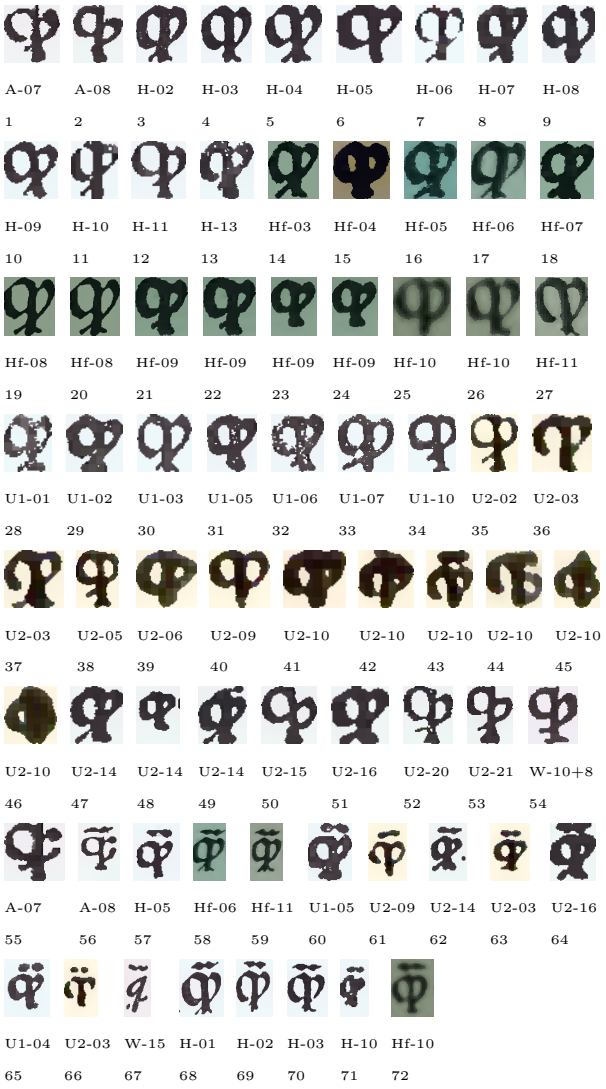


Figure 28: The letter *q* with diagonal stroke

WITH STROKE THROUGH DESCENDER with representative glyph 'ꝗ'. It can be used as an individual word or as part of it and is quite ambiguous; the reported meanings are *quam*, *que*, *quan-* and *qui-*.

Fig. 27 demonstrates also various kinds of diacritical marks which can be used with this brevigraph. The meaning of modified brevigraphs is not clear and requires further research.

The characters in Fig. 28 are in my opinion variants of the brevigraph introduced to Unicode in version 5.1.0 (published 2008) as U+A759 LATIN SMALL LETTER Q WITH DIAGONAL STROKE with representative glyph 'ꝙ', although such a classification of some shapes is questionable. The Unicode name is not very adequate; in [3, p. 70] the name LATIN SMALL LETTER Q WITH SWASH is proposed. The

brevigraph has three meanings: *quod*, *qui* and *que*; it can be used as an individual word or as a part of it.

The meaning of modified brevigraphs with diacritical marks is not clear. Here we will mention only that in [3, p. 71] a glyph similar to those from Fig. 28 is classified as LATIN SMALL LETTER Q WITH SWASH AND LATIN SMALL LETTER FLATTENED OPEN A ABOVE (an alternative name LATIN SMALL LETTER Q WITH FLOURISH ... is also considered), and an example is given where the brevigraph means *quan-*.

The glyphs in Fig. 29 represent the brevigraph assigned the Private Use Area code U+E8BF and the name LATIN SMALL LETTER Q LIGATED WITH FINAL ET by MUFI in version 4 of the recommendation [11, p. 81]. In the Junicode font, it is rendered as '𝔮'.

Finding the meaning of the brevigraph with a diacritical mark requires additional research, but we will note that according to [10] some of the glyphs from Fig. 29 mean *quam* or *quan*.

12 Modifications of the letter *r*

The first glyph in Fig. 30 is an interesting and rather little-known character. Although this is far from obvious, it is U+A775 LATIN SMALL LETTER RUM despite the fact that the Unicode representative glyph is 'ꝛ', as in Junicode we have 'ꝛ' — the glyph is practically identical to that on the figure. The character was added to Unicode in version 5.1 (published in 2008), along with some similar characters (see sec. 14). I assume the second glyph in the figure is just a variant of the first one.

The glyphs in Fig. 31 are ambiguous. They can represent U+A776 LATIN LETTER SMALL CAPITAL RUM (ꝛ), but they can also be interpreted as U+211E PRESCRIPTION TAKE ('℞') and, last but not least, U+211F RESPONSE ('℟') which in prayer books can be paired with the *versicle* character (see sec. 15).

13 Modifications of the letter long *s*

The glyphs in Fig. 32 are noted in the MUFI recommendations as M+E8B7 LATIN SMALL LETTER LONG S WITH FLOURISH ('ſ').

The glyphs in Fig. 33 are present neither in Unicode nor in the MUFI recommendation, but they are obviously the long *s* (U+017F) ligated with the final *et* (U+A76B). As far as I know, this ligature is available only in the Junicode font[25] with the Historic Ligature (*hlig*) feature: 'ſ&'. The meaning is *sed*, as exemplified in [9, example (68)].

Fig. 34 shows a problematic glyph which I'm not sure how to interpret.

One of the component characters of the ligatures presented in Fig. 35 has been already mentioned in

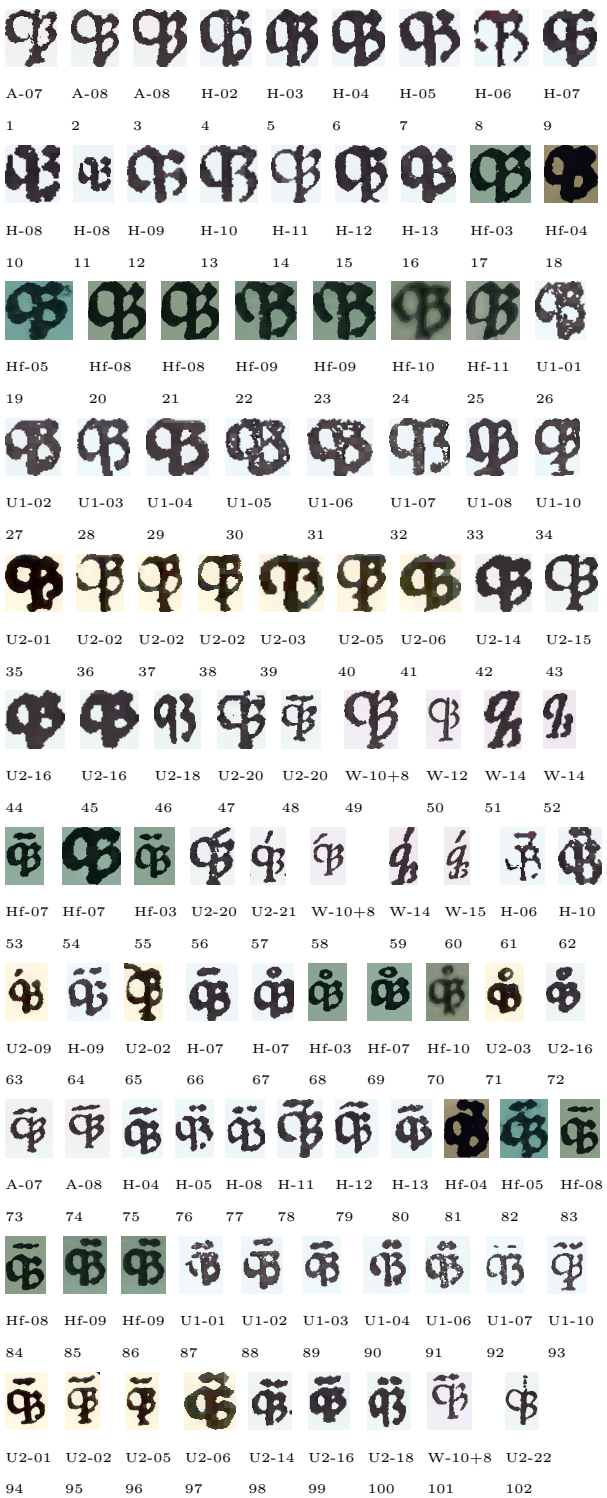


Figure 29: The letter *q* with final *et*

[25] github.com/psb1558/Junicode-font/discussions/140

Figure 30: The alternative glyphs of the letter *rum*

Figure 31: The alternative glyphs of the letter *response*

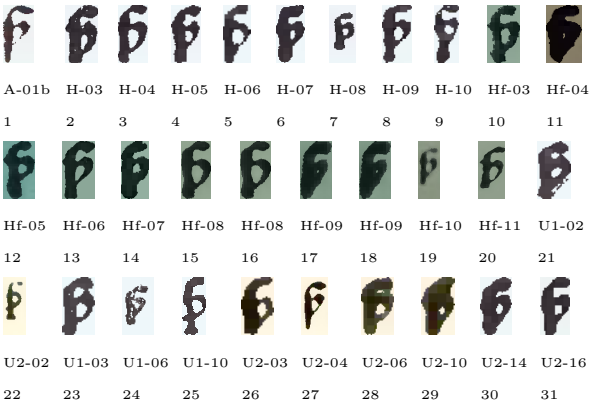


Figure 32: Long *s* with flourish

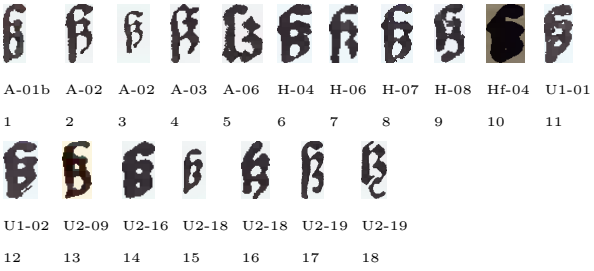


Figure 33: Long *s* with final *et*

Figure 34: Long *s* with final *et*?

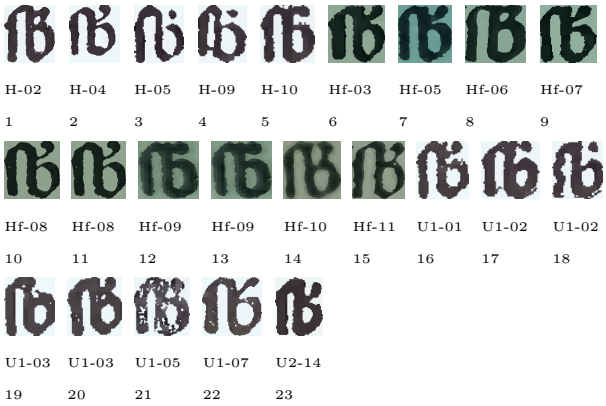


Figure 35: The ligature of long *s* with the letter *b* and its modifications

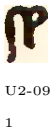


Figure 36: The ligature of long *s* with the letter *l* with flourish?

Figure 37: The letter *tum*

sec. 4. We see also the letter *b* with a dot above; the meaning of the letter, ligated or not, is yet to be investigated.

It is worth noting that in item 18 instead of a normal long *s* we have a LONG FUNNY S proposed to be included in the MUFI recommendation.[26]

In Fig. 36 we have a ligature which can be perhaps treated as a variant of MUFI U+E8AF LATIN SMALL LIGATURE LONG S L WITH DIAGONAL STROKE ('ﬄ').

14 Modifications of the letter *t*

Fig. 37 shows the rather rare brevigraph U+A777 LATIN SMALL LETTER TUM with representative glyph 'ꝧ'. It was introduced in version 5.1 (published in 2008), together with some related letters such as U+A775 LATIN SMALL LETTER RUM (see sec. 12).

15 Modification of the letter v

The primary interpretation of the glyphs in Fig. 38 seems to be U+A75F LATIN SMALL LETTER V WITH

[26] `mufi.info/q.php?p=mufi/chars/unichar/1048876`

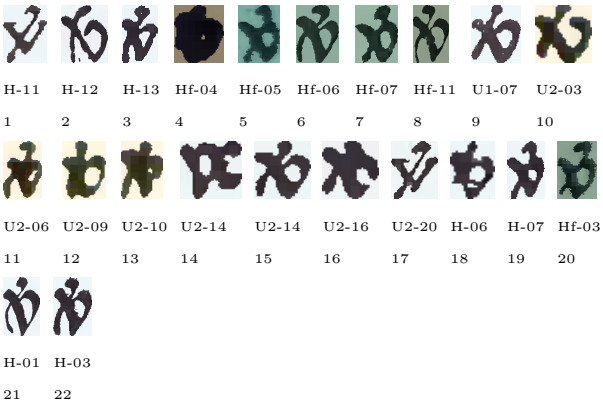


Figure 38: The letter *v* with diagonal stroke

DIAGONAL STROKE ('℣') added to Unicode in version 5.1.0 (published in 2008); in item 14 it is the other arm which is crossed. It means *ver* or *vir*.

The glyphs can stand also for U+2123 VERSICLE ('℣'), used to mark in the prayer books the beginning of a versicle, i.e., a short sentence said or sung by the minister in a church service, to which the congregation gives a response.[27]

16 Final remarks

As has already been mentioned, the next step should be to find the occurrences of the discussed brevigraphs in the texts and in this way find or verify their meaning. For this, we don't need the full transcriptions of the texts. What is sufficient for our purposes is *glyph* or *character spotting*. These tasks are discussed in some publications, but there seems to be no tool available directly for use. With some limitations, a variant of a workflow described earlier can be used for this purpose. The djvudict output can be converted to a PDF document (created with TEX) with enlarged glyphs which make it relatively easy to search for interesting items and to note their identifiers (Fig. 39). Additionally, a djview4poliqarp index can be created, which uses the shape identifier as the searchable entry fields (Fig. 40). The identifiers are not unique, nevertheless it is possible with some effort to find the context of a shape in the document, as illustrated in Figures 39 and 40 (note the shape 01344).

At present my programs supporting this approach are too primitive to be used conveniently, but I will try to improve them. Any help from Python and/or QT programmers (QT was used to implement djview4poliqarp and djview4shapes) would be welcomed and appreciated.

[27] Definition provided by Google in a non-linkable form.

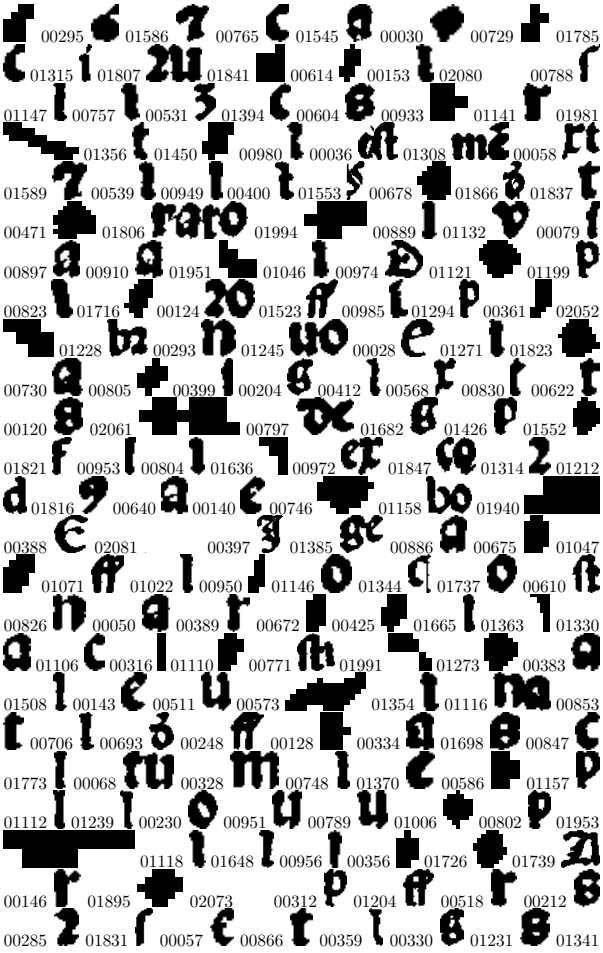


Figure 39: The djvudict output in the form of a PDF document

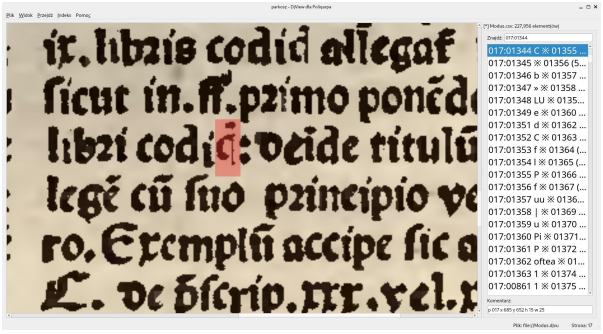


Figure 40: The djvudict output in the form of a djview4poliqarp index

Acknowledgments. Thanks to Barbara Beeton and Karl Berry for their careful reading of the manuscript and improving the presentation and English wording.

References

[1] J. André, R. Jimenes. Transcription et codage des imprimés de la Renaissance. *Revue des Sciences et Technologies de l'Information* —

Série Document Numérique, 16(3):113–139, 2013. doi.org/10.3166/DN.16.3.113-139

[2] J. André. Inventaire des typèmes de la B42. Note de travail NT-1, Projet d'Inventaire des Caractères Anciens, 2015. jacques-andre.fr/PICA/B42-typemes.pdf

[3] J. André. Inventaire des typèmes latins et français existant dans Unicode/MUFI ou à y faire entrer. Note de travail NT-2, Projet d'Inventaire des Caractères Anciens, 2022. jacques-andre.fr/PICA/SIGMA-PICA.pdf

[4] G. Bettridge. How to read the Gutenberg Bible. Text prepared for the Lyon Printing Museum.

[5] J.S. Bień. 16th century Latin brevigraphs in Unicode — a computer resource. In *Grapholinguistics in the 21st Century 2022. Proceedings. Grapholinguistics and Its Applications*, Y. Haralambous, ed., pp. 31–46. Fluxus Editions, Brest, 2023 (to appear). doi.org/10.36824/2022-graf-bien

[6] J.S. Bień. Repertuar znaków pisma nr 1 pierwszej drukarni Unglera (1510–1516) na podstawie Polonia Typographica. *Acta Poligraphica*, pp. 1–20, 2021. www.cobrpp.com.pl/actapoligraphica/uploads/pdf/AP2021_Bien.pdf

[7] J.S. Bień. Representing Parkosz's alphabet in the Junicode font. *TUGboat* 43(3):247–251, 2022. doi.org/10.47397/tb/43-3/tb135bien-parkosz

[8] J.S. Bień. Towards an inventory of old print characters: Ungler's *Rubricella* , a case study. *TUGboat* 44(3):364–375, 2023. doi.org/10.47397/tb/44-3/tb138bien-rubricella

[9] J.S. Bień. Towards an inventory of old print characters: Ungler's *Rubricella*, a case study – errata. *TUGboat* 45(1):44, 2024. doi.org/10.47397/tb/45-1/tb139bien-rubricella-errata

[10] E. Blake. A briefing on brevigraphs, those strange shapes in early printed texts, 2021. folger.edu/blogs/collation/brevigraphs/

[11] O.E. Haugen. *MUFI character recommendation version 4.0.* Medieval Unicode Font Initiative, 2015. hdl.handle.net/1956/10699

◇ Janusz S. Bień
Warsaw, Poland
jsbien (at) uw.edu.pl
sites.google.com/view/jsbien
ORCID 0000-0001-5006-8183

Towards an inventory of old print characters: Ungler's *Rubricella*, a case study — Errata

Janusz S. Bień

Abstract

Errata for the article in *TUGboat* issue 44:3, pp. 364–375 (tug.org/TUGboat/tb44-3/tb138bien-rubricella.pdf): 1) "ꝉ" and "ꝗ" were typeset where 'ꝛ' and 'ꝓ' were intended; 2) 'ꝑ' was confused with 'ꝯ'; 3) an *h* should have been *b*.

4.5 Brevigraphs

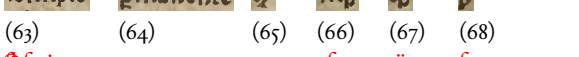
From the first paragraph of this section in the original article:

> . . . in Junicode also "'ꝉ" . . .
> . . . in the Junicode font also "ꝗ'' . . .

The wrong characters were typeset; these should have been 'ꝛ' and 'ꝓ'. This was because of a problem with the default font renderer in Lua(LA)TEX. Switching to the HarfBuzz renderer solves it. Thanks to Marcel Krueger and Luigi Scarso.

A second error was towards the end of the section. The figure (below, corrected from the original) lists two brevigraphs based on the long *s*. The first one in example (65) is 'ꝑ' (M+E8B7 LATIN SMALL LETTER LONG S WITH FLOURISH [MUFI 4.0]). The second, in example (68), is similar and in the paper was confused with it. Actually it is the long *s* (U+017F) ligated with the final *et* (U+A76B); it is present neither in Unicode nor in the MUFI recommendation. As far as I know, this ligature is available only in the Junicode font[1] with the Historic Ligature (*hlig*) feature: 'ꝯ'. It can be seen in example (68) where it is used as a separate word (its interpretations were suggested in the Facebook Paleography Society group by Gionata Brusa and Carolus Hrachowiczensis;[2] unfortunately neither I nor none of the group noted that I had confused 'ꝑ' with 'ꝯ').

The first brevigraph, example (63) with the ligature with the letter *b* (the original article incorrectly wrote *h* here) and a diacritic mark, was discussed above.



(63)	(64)	(65)	(66)	(67)	(68)
ſ̄bſcripto	ꝑmanente	qꝑ	vſqꝫ	qꝫ	ꝯ
subscripto	permanente	quod	usque	quam	scilitet? sed?

(69)	(70)
ɔcluſa	lcōibꝰ
conclusa	lectionibus

◇ Janusz S. Bień
Warsaw, Poland
jsbien (at) uw.edu.pl
sites.google.com/view/jsbien
ORCID 0000-0001-5006-8183

[1] github.com/psb1558/Junicode-font/discussions/140
[2] facebook.com/groups/7687162686/posts/10159250228377687

doi.org/10.47397/tb/45-1/tb139bien-rubricella-errata