

# Usage of Unicode Plane 15 in JuniusX font

## (a Polish perspective)

Janusz S. Bień

January 3, 2021

## 1 Introduction

JuniusX font<sup>1</sup>, used to typeset this paper, is the sucesor to Junicode<sup>2</sup>, both fonts has been created by Peter Baker. It is a font featuring in particular full compliance with the Medieval Unicode Font Initiative specification (version 4.0) and including all medieval characters added since MUFI 4.0 (<https://mufi.info/>). The font is available on the principles of the SIL Open Font License (OFL) v. 1.1<sup>3</sup> (SIL International is the successor of the Summer Institute of Linguistics for training missionaries).

The fonts includes also character variants suggested by users, in particular several characters needed for old Polish texts and not available elsewhere.

Due to its coverage not limited to the Unicode amd MUFI characters it can be considered a small but important step towards the inventory of types proposed in (André 2003).

Regular Unicode characters have a lot of various properties, described in the so called Unicode Character Database. The properties are used directly or indirectly by applications, in particular by OpenType rendering engines.

---

<sup>1</sup><https://psb1558.github.io/Junicode-New/>

<sup>2</sup><https://junicode.sourceforge.io/>

<sup>3</sup>[https://scripts.sil.org/cms/scripts/page.php?site\\_id=nrsi&id=OFL](https://scripts.sil.org/cms/scripts/page.php?site_id=nrsi&id=OFL)

MUFI uses Private Use Area characters, but does not provide explicitly their properties. There was an attempt to provide them in a way mimicking The Unicode Character Database<sup>4</sup>, but applications are not capable to make use of such unofficial information; this of course can change in the future<sup>5</sup>.

One of the design goals of JuniusX was to supercede PUA characters by various features of OpenType fonts<sup>6</sup>, such as

- `calt` (Contextual Alternates)
- `cv` (Character Variant)

Despite the advantages of such approach<sup>7</sup> there are still some important applications unable to handle the features (at least without some modifications). One of them is `djview4poliqarp` (<https://bitbucket.org/mrudolf/djview-poliqarp/>) supporting indexes such as <https://github.com/jsbien/Zaborowski-index4djview>.

Using OpenType feature is also of no use for encoding texts to be included in linguistic corpora (it's theoretically possible to translate them into an appropriate markup such as that formulated by Text Encoding Initiative<sup>8</sup>, but it makes corpus processing much more complicated). For example, in the IMPACT corpus (Bień 2012) the letters `ą` and `ę` and their variants with the stroke instead of ogonek occurred with almost the same frequency. The stroke variants has been represented by `ą` ('2C65') and `ę` ('0247'), but this was considered a temporary solution as the glyphs differ significantly.

The solution is to duplicate in the Private Use Area the character variant accessible with OpenType features. As the primary Private use Area, i.d. U+E000-U+F8FF is rather crowded and already used for conflicting assignments, the decision was made to use plane 15 (U+F0000 – U+FFFFFD). It seems to be the very first usage of plane 15.

<sup>4</sup><http://www.kreativekorp.com/charset/PUADATA/PUBLIC/MUFI/>

<sup>5</sup>Cf. e.g. <https://lists.gnu.org/archive/html/emacs-devel/2021-01/msg00138.html>.

<sup>6</sup>Cf. e.g. <https://docs.microsoft.com/en-us/typography/opentype/spec/featurelist>

<sup>7</sup>Cf. e.g. <https://psb1558.github.io/Junicode-New/Searchability.html>

<sup>8</sup><https://tei-c.org/guidelines/>

One of the important Unicode character property is its name. In the present paper names, following the Unicode naming rules, are proposed for private use characters assigned in the font to plane 15. Additionally the names are provided in the format suitable to use it with `fntsample-with-comments`<sup>9</sup> to produce a Unicode-like font tables and for `Unihistext`<sup>10</sup> to produce character histograms.

Some principles of the naming policy are:

- Comments in brackets are considered as a part of the name by `Unihistext`, but not necessarily in other circumstances.
- Every proper name ends in `VARIANT`
- Names try to follow English Unicode usage. The term `terminal` comes from (Gaskell 1976).

Some characters presented here has been used in critical or didactic editions of old Polish manuscripts, such as (Vrtel-Wierczyński 1950). It is worth noting that the second edition is primarily a photo-offset reproduction of the 1930 first printing; one of the reasons was the difficulty to find a printing house, where the type needed for old Polish survived the WW II.

The best way to access links in the paper to DjVu documents is to use appropriately configured External Application Button extension (<https://github.com/andy-portmen/external-application-button/issues/50>) available for popular browsers.

## 2 List of characters

**0xF0000** LATIN CAPITAL LETTER A WITH STROKE THROUGH TERMINAL  
VARIANT [JuniusX]: A.

Cf. <https://github.com/psb1558/Junicode-New/issues/14>.

---

<sup>9</sup>Cf. <https://github.com/psb1558/Junicode-New/discussions/49>

<sup>10</sup><https://bitbucket.org/jsbien/unihistext>

In JuniusX font accessible also as Ą with cv02[1]; cf. 0xF001E below for a slightly different glyph.

May be considered also as a variant of LATIN CAPITAL LETTER A WITH STROKE (U+023A).

Occurs in particular in Nowe Ateny (1756, <http://www.wbc.poznan.pl/Content/6625/index.djvu?djvuopts=&page=2186.djvu>, cf. Fig. 1.

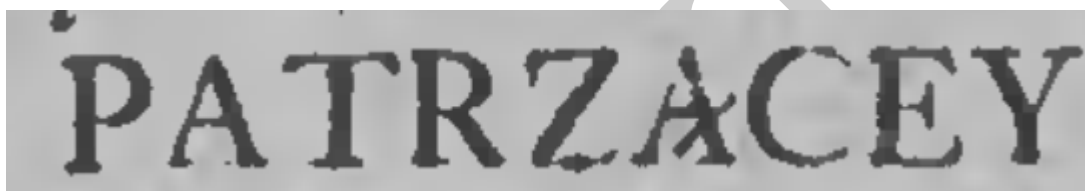


Figure 1: PATZRĄCEY (1756)

0xF0001 LATIN SMALL LETTER A WITH STROKE THROUGH TERMINAL VARIANT [JuniusX]: ą.

Cf. <https://github.com/psb1558/Junicode-New/issues/14>; cf. 0xF001F below for a slightly different glyph.

In JuniusX font accessible also as ą with cv02[1]; a slightly different glyph available as cv02[2], cf. (Baker 2020, p. 7).

May be considered also as a variant of LATIN SMALL LETTER A WITH STROKE (U+2C65).

Occurs in particular in Nowe Ateny (1756, <http://www.wbc.poznan.pl/Content/6625/index.djvu?djvuopts=&page=2355.djvu>, cf. Fig. 2.

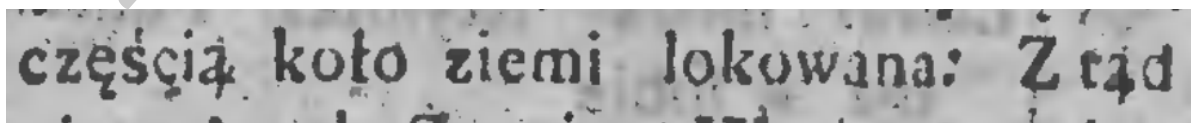


Figure 2: częścią [ldots] tą d (1756)

0xF0002 LATIN CAPITAL LETTER A WITH EXTENDED OGONEK VARIANT

[JuniusX]: Ą.

Cf. <https://github.com/psb1558/Junicode-New/issues/14>.

In JuniusX font accessible also as ą with cv02[3]; cf. (Baker 2020, p. 7).

Occurs at least in Zbiór rytmów duchownych Panegirycznych Moralnych i Światowych [...] Elżbiety z Kowalskich Drużbackiej (1752, [http://www.wbc.poznan.pl/Content/22431/directory.djvu?djvuopts=&page=0242\\_0001.djvu](http://www.wbc.poznan.pl/Content/22431/directory.djvu?djvuopts=&page=0242_0001.djvu)).  
cf. Fig. 3.



Figure 3: XIAŹĘCIA (1752)

0xF0003 LATIN SMALL LETTER C WITH MEDIUM HIGH OVERLINE VARIANT

[JuniusX]: Ĉ.

Cf. <https://github.com/psb1558/Junicode-New/discussions/44#discussioncomment-202860>.

In JuniusX font accessible with ss05; cf. (Baker 2020, p. 7).

0xF0004 LATIN SMALL LETTER C WITH HIGH OVERLINE VARIANT [JuniusX]:

Ċ.

Cf. <https://github.com/psb1558/Junicode-New/discussions/44#discussioncomment-202860>.

In JuniusX font accessible with ss04; cf. (Baker 2020, p. 6).

0xF0005 LATIN SMALL LETTER INSULAR D VARIANT [JuniusX]:

ḁ.

In JuniusX font accessible also as d with cv05[2]; cf. (Baker 2020, p. 7).

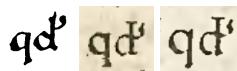


Figure 4: On the left Capelli's dictionary, next Zaborowski's treatise (1518)

May be considered also as a variant of LATIN SMALL LETTER INSULAR D (U+A77A).

0xF0006 LATIN SMALL LETTER ETH VARIANT [JuniusX]: ð.

Cf. <https://github.com/psb1558/Junicode-New/discussions/44#discussioncomment-202860>.

In JuniusX font accessible also as ð (U+00F0) with ss01 when loca enabled; cf.

(Baker 2020, p. 11) and (kenmcd 2020, p. 4 (30)).

0xF0007 LATIN SMALL LETTER D WITH STROKE ROUNDED VARIANT [JuniusX]:

ḑ.

Cf. <https://latin.stackexchange.com/questions/14740/encoding-abbreviated-quod-in-unicode>

In JuniusX font accessible also as ḑ (U+0111) with cv06[1]; cf. (Baker 2020, p.

7).

The character occurs twice in Zaborowski's treatise (one occurrence unclear), it is noted in Capelli's dictionary of (handwritten) abbreviations, cf. (Bień 2020), (Capelli 1928, s. 307) and Fig. 4.

0xF0008 LATIN SMALL LETTER D WITH HIGH OVERLINE VARIANT:

ḏ.

Cf. <https://github.com/psb1558/Junicode-New/discussions/44#discussioncomment-202860>.

In JuniusX font accessible with ss04; cf. (Baker 2020, p. 6).

0xF0009 LATIN CAPITAL LETTER E WITH STROKE THROUGH TERMINAL VARIANT [JuniusX]:

Ē.

Cf. <https://github.com/psb1558/Junicode-New/issues/13>.

In JuniusX font accessible also as Ē with cv08[2]; cf. (Baker 2020, p. 8).

May be considered also as a variant of LATIN CAPITAL LETTER E WITH STROKE (U+0246).

Occurs in particular in Nowe Ateny (Supplement 1774, <http://www.wbc.poznan.pl/Content/6663/index.djvu?djvuopts=&page=1783.djvu>, cf. Fig. 5.



Figure 5: WEZACH (1774)

0xF000A LATIN SMALL LETTER E WITH STROKE THROUGH TERMINAL  
VARIANT [JuniusX]:

ē.

Cf. <https://github.com/psb1558/Junicode-New/issues/13>.

In JuniusX font accessible also as ē with cv08[2]; cf. (Baker 2020, p. 7).

For completeness it should be noted that there is also the variant cv08[1] of ē intended for the Latin abbreviation of AE.

Occurs in particular in List o oblężeniu zamku Dyjamenckiego (1605, <https://cbdu.ijp.pan.pl/id/eprint/2910/>, cf. Fig. 6 on the next page.

May be considered also as a variant of LATIN SMALL LETTER E WITH STROKE (U+0247).

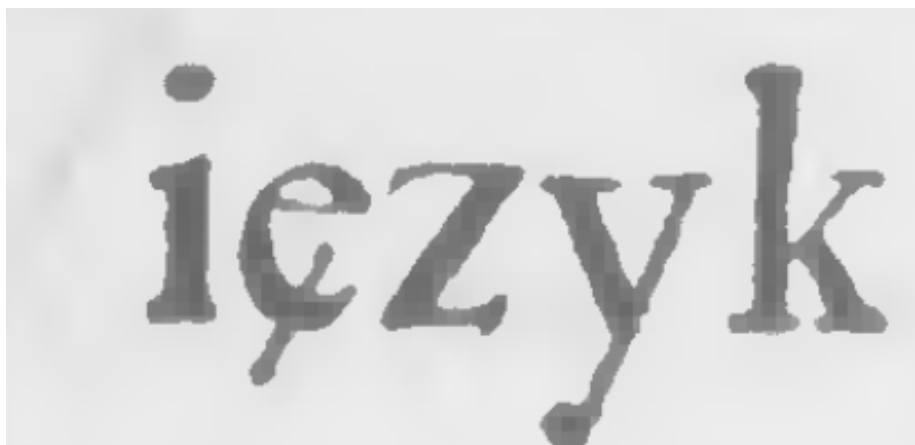


Figure 6: język (1605)

0xF000B LATIN SMALL LETTER F CONTEXTUAL VARIANT [JuniusX]:

f.

Used e.g. before vowels with dieresis, cf. <https://github.com/psb1558/Junicode-New/discussions/44#discussioncomment-208296>.

In JuniusX font accessible also as f with cv09 [5]; cf. (Baker 2020, p. 9), and used in proper contexts by calt (Contextual Alternates).

0xF000C LATIN SMALL LETTER I WITH HIGH OVERLINE VARIANT [JuniusX]:

ī.

Cf. <https://github.com/psb1558/Junicode-New/discussions/44#discussioncomment-202860>.

In JuniusX font accessible with ss04; cf. (Baker 2020, p. 6).

0xF000D LATIN SMALL LETTER J CONTEXTUAL VARIANT [JuniusX]:

j.

Used when the letter is preceded by g or certain vowels with ogonek, cf. <https://github.com/psb1558/Junicode-New/discussions/44#discussioncomment-208296>.

In JuniusX font used in proper contexts by calt (Contextual Alternates).

0xF000E LATIN SMALL LETTER J WITH HIGH OVERLINE VARIANT [JuniusX]:



j̄.

Cf. <https://github.com/psb1558/Junicode-New/discussions/44#discussioncomment-202860>.

In JuniusX font accessible with ss04; cf. (Baker 2020, p. 6).

#### 0xF000F LATIN SMALL LETTER L WITH HIGH ROUNDED STROKE VARIANT

[JuniusX]:

l̄.

Cf. <https://github.com/psb1558/Junicode-New/issues/4>.

In JuniusX font accessible also as ł (LATIN SMALL LETTER L WITH HIGH STROKE, U+A749) with cv17 [1]; cf. (Baker 2020, p. 9).

Used in Latin as an abbreviation, cf. (Balbi 1460), quoted after (Everson et al. 2006, s. 6).

The character occurs in particular in Zaborowski's treatise as the abbreviation, cf. 7, but also in the function of the present day ł.



Figure 7: From the left: vl (1514-1515) and 1518), regula (1514-1515), populus (1518)

#### 0xF0010 LATIN SMALL LETTER M WITH HIGH OVERLINE VARIANT [Ju-

nusX]:

m̄.

Cf. <https://github.com/psb1558/Junicode-New/discussions/44#discussioncomment-202860>.

In JuniusX font accessible with ss04; cf. (Baker 2020, p. 6).

#### 0xF0011 LATIN SMALL LETTER O WITH STROKE HORNED VARIANT [Ju-

nusX]

ō.

Cf. <https://github.com/psb1558/Junicode-New/issues/3>. In JuniusX font accessible also as ø (U+00F8) with cv21 [1]; cf. (Baker 2020, p. 9).

The principles of editing old Polish texts (Górski 1955) mention the character, but don't give specific examples. Nevertheless they recommend to use it in critical editions.

**0xF0012** LATIN SMALL LETTER O WITH SHIFTED STROKE VARIANT [JuniusX]

: .

Cf. <https://github.com/psb1558/Junicode-New/issues/3>.

In JuniusX font accessible also as ø (U+00F8) with cv21 [2]; cf. (Baker 2020, p. 9).

The principles of editing old Polish texts (Górski 1955) mention the character, but don't give specific examples.

**0xF0013** LATIN SMALL LETTER O WITH STROKE HORN DOWNWARDS VARIANT [JuniusX]

: .

Cf. <https://github.com/psb1558/Junicode-New/issues/3>.

In JuniusX font accessible also as ø (U+00F8) with cv21 [3]; cf. (Baker 2020, p. 9).

The principles of editing old Polish texts (Górski 1955) mention the character, but don't give specific examples.

**0xF0014** LATIN SMALL LETTER O WITH STROKE HORN UPWARDS VARIANT [JuniusX]

: .

Cf. <https://github.com/psb1558/Junicode-New/issues/3>.

In JuniusX font accessible also as ø (U+00F8) with cv21 [4]; cf. (Baker 2020, p. 9).

The principles of editing old Polish texts (Górski 1955) mention the character, but don't give specific examples.

**0xF0015 MEDIEVAL SUPERScript LETTER Q [JuniusX]**

: .

Cf. <https://github.com/psb1558/Junicode-New/discussions/44#discussioncomment-208296>.

In JuniusX font accessible also as sup (Superscripts).

**0xF0016 LATIN SMALL LETTER RUM ROTUNDA VARIANT [JuniusX]**

: .

In JuniusX font accessible also as ʀ (U+A75D) with cv41 [1]; cf. (Baker 2020, p. 12).

**0xF0017 LATIN SMALL LETTER V WITH MEDIUM HIGH OVERLINE VARIANT [JuniusX]:**

.

Cf. <https://github.com/psb1558/Junicode-New/discussions/44#discussioncomment-202860>.

In JuniusX font accessible with ss05; cf. (Baker 2020, p. 7).


**0xF0018 LATIN SMALL LETTER V WITH HIGH OVERLINE VARIANT [JuniusX]:**

.

Cf. <https://github.com/psb1558/Junicode-New/discussions/44#discussioncomment-202860>.


In JuniusX font accessible with ss04; cf. (Baker 2020, p. 6).

**0xF0019 LATIN SMALL LETTER X WITH MEDIUM HIGH OVERLINE VARIANT [JuniusX]:**

.

Cf. <https://github.com/psb1558/Junicode-New/discussions/44#discussioncomment-202860>.

In JuniusX font accessible with `ss05`; cf. (Baker 2020, p. 7).

**0xF001A** LATIN SMALL LETTER X WITH HIGH OVERLINE VARIANT [JuniusX]: .

Cf. <https://github.com/psb1558/Junicode-New/discussions/44#discussioncomment-202860>.

In JuniusX font accessible with `ss04`; cf. (Baker 2020, p. 6).

**0xF001B** LATIN LETTER GLOTTAL STOP VARIANT [JuniusX] .

In JuniusX font accessible as `?` (U+0294) with `cv50[1]`; cf. (Baker 2020, p. 11).

**0xF001C** TIRONIAN SIGN CAPITAL ET VARIANT [JuniusX] .

In JuniusX font accessible also as `]` (U+2E52) with `cv40[1]`; cf. (Baker 2020, p. 12).

**0xF001D** TIRONIAN SIGN ET VARIANT [JuniusX] .

In JuniusX font accessible also as `j` (U+204A) with `cv40[1]`; cf. (Baker 2020, p. 12).

**0xF001E** LATIN CAPITAL LETTER A WITH STROKE VARIANT [JuniusX]: .

Cf. <https://github.com/psb1558/Junicode-New/issues/14>.

In JuniusX font accessible also as `A` with `cv02[2]`; cf. (Baker 2020, p. 7).

May be considered also as a variant of LATIN CAPITAL LETTER A WITH STROKE (U+023A).

**0xF001F** LATIN SMALL LETTER A WITH STROKE VARIANT [JuniusX]: .

Cf. <https://github.com/psb1558/Junicode-New/issues/14>.

In JuniusX font accessible also as A with cv02[2]; cf. (Baker 2020, p. 7).

May be considered also as a variant of LATIN SMALL LETTER A WITH STROKE (U+2C65).

**dolne** Kazania świętokrzyskie s.12 kazanie III na dzień św. Mikołaja

Sankt Florian Psalter Holy Cross Sermons

Vrtel-Wierczyński 1950 photo-offset

(Wydra 2014)

Repository!!!

## References

André, Jacques (2003). “The Casetin Project – Towards an Inventory of Ancient Types and the Relate Standardised Encoding”. In: TUGboat 24.3, pp. 314–318. URL: <http://www.tug.org/TUGboat/tb24-3/andre.pdf> (visited on 12/28/2020) (cit. on p. 1).

Baker, Peter (2020). OpenType Features in JuniusX / JuniusVF. URL: <https://github.com/psb1558/Junicode-New/> (cit. on pp. 4–13).

Balbi, Giovanni (1460). Catholicon. Mainz. URL: <https://www.loc.gov/item/47043559/> (visited on 04/29/2020) (cit. on p. 9).

Bień, Janusz S. (Apr. 2012). “Delivering the IMPACT project Polish Ground-Truth texts with Poliqarp for DjVu”. Technical Report. URL: <http://bc.klf.uw.edu.pl/289/> (cit. on p. 2).

— (2020). Traktat Stanisława Zaborowskiego i skróty brachygraficzne. URL: <https://www.researchgate.net/publication/341930612> (cit. on p. 6).

- Cappelli, Adriano (1928). *Lexicon abbreviatarum Wörterbuch lateinischer und italienischer Abkürzungen*. Zweite verbesserte Auflage. Leipzig: J. J. Weber. URL: <https://archive.org/details/LexiconAbbreviatarum> (visited on 09/12/2020) (cit. on p. 6).
- Everson, Michael et al. (2006). Proposal to add medievalist characters to the UCS. Tech. rep. N3027. ISO/IEC JTC1/SC2/WG2. URL: <http://unicode.org/L2/L2006/06027-n3027-medieval.pdf> (visited on 05/10/2019) (cit. on p. 9).
- Gaskell, Philip (1976). "A Nomenclature for the Letterforms of Roman Type". In: *Visual Language X.1*, pp. 45–51. URL: <https://s3-us-west-2.amazonaws.com/visiblelanguage/pdf/10.1/nomenclature-of-the-letterforms-of-roman-type.pdf> (visited on 12/15/2020) (cit. on p. 3).
- Górski, Konrad, ed. (1955). *Zasady wydawania tekstów staropolskich: projekt*. Do wersji elektronicznej dostęp ograniczony. Wrocław: Zakład im. Ossolińskich - Wydaw. Polskiej Akademii Nauk. URL: <http://ebuw.uw.edu.pl/publication/1334> (cit. on pp. 10, 11).
- kenmcd (2020). *JuniusX User Guide - FIRST DRAFT*. URL: <https://github.com/psb1558/Junicode-New/issues/16> (cit. on p. 6).
- Vrtel-Wierczyński, Stefan (1950). *Wybór tekstów staropolskich*. Wydanie drugie. Warszawa: PZWS (cit. on pp. 3, 13).
- Wydra, Wiesław (2014). "The Oldest Extant Prose Text in the Polish language. The Phenomenon of the Holy Cross Sermons". In: *Polish Libraries Vol. 2*, pp. 133–159. ISSN: 2353-1835. URL: [https://polishlibraries.bn.org.pl/upload/pdf/93724\\_Wieslaw\\_Wydra\\_The\\_Oldest\\_Extant\\_Prose\\_Text\\_in\\_the\\_Polish\\_language.The\\_Phenomenon\\_of\\_the\\_Holy\\_Cross\\_Sermons\\_133-159.pdf](https://polishlibraries.bn.org.pl/upload/pdf/93724_Wieslaw_Wydra_The_Oldest_Extant_Prose_Text_in_the_Polish_language.The_Phenomenon_of_the_Holy_Cross_Sermons_133-159.pdf) (visited on 12/23/2020) (cit. on p. 13).