

Proposal to include the Old Polish nasal vowel letter

(draft: 4 Jan 2021)

Daniel Bunčić

1 Introduction

Old Polish, written from the 12th to the early 16th century, had a nasal vowel /ã/, which manuscripts represented by a special letter, which is also used in modern transcriptions, e.g. in scholarly text editions. This letter in its ‘canonical’ form looks more or less like an *o* with vertical lines attached at the top and at the bottom (ø), although other forms are also frequent.

In Middle Polish, the Old Polish nasal vowel split into two nasal vowels: /ɔ̃/, which is spelled *ą* (U+0105 LATIN SMALL LETTER A WITH OGONEK), and /ɛ̃/, which is spelled *ę* (U+0119 LATIN SMALL LETTER E WITH OGONEK). Accordingly, the Old Polish letter ø sometimes corresponds to the one and sometimes to the other modern vowel letter, e.g. Old Polish *bødø* ‘they will be’ > Modern Polish *będq*. Consequently, it would be inappropriate to represent ø as either *ą* or *ę*.

2 Suitability for inclusion

The Old Polish nasal vowel character is clearly a character according to the definition in the Glossary. It is functionally distinct from all other characters used in Old Polish and also from the two Modern Polish nasal vowel characters *ą* (U+0105 LATIN SMALL LETTER A WITH OGONEK) and *ę* (U+0119 LATIN SMALL LETTER E WITH OGONEK). So far, this character is neither in the Pipeline nor in the Archive of Nonapproval Notices.

3 Evidence of use

3.1 Use in medieval manuscripts

The Old Polish nasal vowel letter occurs in a great number of Old Polish manuscripts. For example, in the [Holy Cross Sermons](#) (*Kazania świętokrzyskie*) from around 1300, it is written as an *o* with a straight line from the top of the letter through its bottom, see fig. 1. (In all facsimiles, the instances of the letter are emphasized by increased contrast.) This shape looks similar to some forms of Greek phi (φ; the third instance in fig. 1 shows the intended shape more clearly than the other two instances). However, in the [Sankt Florian Psalter](#) (*Psalterz floriański*) from around 1400, which is written in a beautiful textualis hand, the Old Polish nasal vowel letter always appears as an *o* with thin lines at the top and bottom, which do not extend into the interior of the *o*, see fig. 2. The same shape is also found in [Queen Sophia’s Bible](#) (*Biblia królowej Sofii*) from the middle of the 15th century, see fig. 3, where the lines are so thin that they are sometimes hardly visible, as well as, in a less calligraphic hand, in a 1407 manuscript of the famous hymn *Bogurodzica*, see fig. 4. In another, even more cursive, manuscript from the beginning of the 15th century, the [Gniezno Sermons](#) (*Kazania gnieźnieńskie*), see fig. 5, there sometimes seem to be two separate lines, which do not cross the *o* (as in the instances in the top right corner and in the bottom left corner of fig. 5), but sometimes they seem to be combined into one line, crossing the *o*, which in these cases resembles the Scandinavian letter ø, especially because in this cursive script all ‘vertical’ lines are considerably slanted to the right.

All these shapes are clearly allographs of a single grapheme, i.e. variant glyphs of the same character.

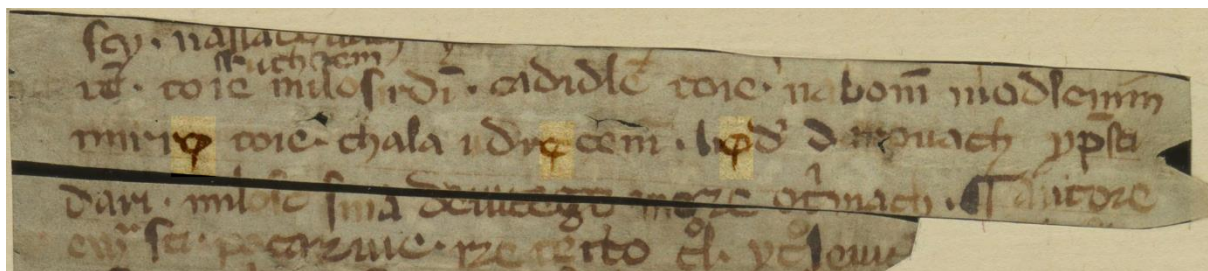


Fig. 1. Holy Cross Sermons (fol. 4r)

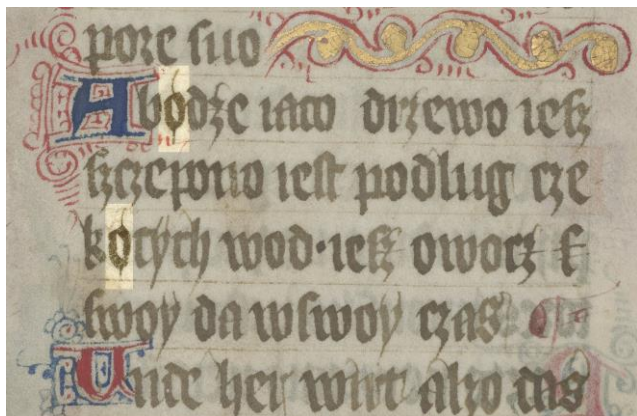


Fig. 2. Sankt Florian Psalter (fol. 3r)

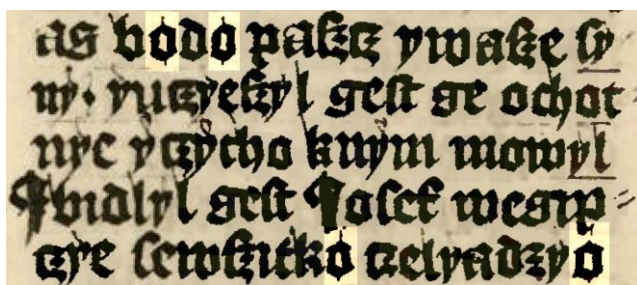


Fig. 3. Queen Sophia's Bible (fol. 24v; from Bernacki 1930)

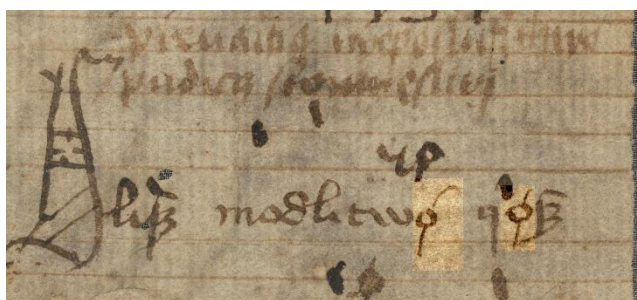


Fig. 4. Bogurodzica of 1407

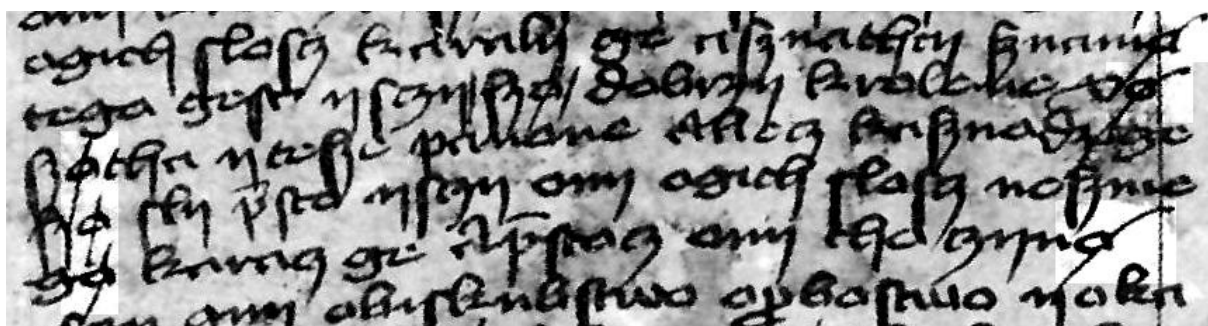


Fig. 5. The Gniezno Sermons (Siatkowska 1991: 278)

3.2 Use in modern scholarly texts

In the normative work *Principles of editing Old Polish texts* (*Zasady wydawania tekstów staropolskich*, Górski et al. 1955: 28), an *o* with two vertical lines attached at the top and bottom (ϕ) was chosen and recommended as the ‘canonical’ form of the Old Polish nasal vowel letter:

“We convey the different forms of the slashed *o* ($\phi \phi \phi \phi$) in the normalized form ϕ .”
 (“Różne postaci *o* przekreślonego ($\phi \phi \phi \phi$) sprowadzamy do znormalizowanej postaci ϕ .”)

This normalized form (in its non-italic form) can be seen in fig. 6 and 7 from the same book. Consequently, whenever modern editors have had access to a type foundry and have thus been able to have a special sort cast for the purpose, this has been the preferred shape, as fig. 8 (from 1965) and 9 (from 1984) show, although the previous practice of designing glyphs that more or less faithfully reproduce the specific shapes of the letter in the manuscripts can still be seen in older books, cf. fig. 10 (from 1950, but itself a reprint from 1930). Even before 1955, however, the most frequent form of the letter in scholarly editions was ϕ , as in fig. 11 (from 1883). The faithful representation of the variants of ϕ used in a specific manuscript can still be a legitimate aim, of course, but variant glyphs of the same character should be handled by fonts (e.g. by OpenType features), not by Unicode.

As Bień (2020: 2) points out, the norm introduced by Górski et al. (1955) advances a shape of the character that cannot be represented in Unicode and is consequently not supported by modern computer fonts. Therefore, since about the 1980s, when typesetting in scholarly publishing started to be increasingly done by the authors themselves rather than professional typesetters, such technical problems have led to a variety of makeshift solutions even in very respectable publications. Thus, fig. 12 (from 1999) shows an alternative shape of the character with a long vertical line through the *o* (ϕ), which was probably chosen for technical reasons, just like the integral symbol instead of long *l*. Similarly, fig. 13 (from the typewriter era, 1991) shows a manually slashed *o*; fig. 14 and 15 (from 1993) exhibit the Greek letter phi (ϕ) and the Scandinavian letter ϕ , respectively; and in fig. 16 from a current Wikipedia article the Unicode character U+03D5 GREEK PHI SYMBOL (ϕ) is used. This variety of characters employed for the same function shows all the more clearly that a unique Unicode character for the Old Polish nasal vowel is desperately needed.

Fig. 6. Model for a scholarly edition (Górski et al. 1955: 114) of the Gniezno Sermons (a different excerpt from the one in fig. 5)

Fig. 7. Model for a scholarly edition (Górski et al. 1955: 144) of the *Bogurodzica* (cf. fig. 4)

Fig. 8. Scholarly edition (Urbańczyk & Kyas 1965: 89) of Queen Sophia's Bible (cf. fig. 3)

Fig. 9. Scholarly anthology (Wydra & Rzepka 1984: 32); here: the Holy Cross Sermons (cf. fig. 1)

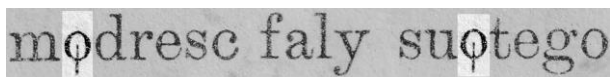


Fig. 10. Scholarly anthology (Vrtel-Wierczyński 1950: 12); here: the Holy Cross Sermons (a different excerpt from the one in fig. 1; scan by Janusz S. Bień)

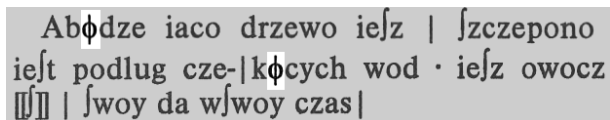


Fig. 12. University textbook (Gehrmann et al. 1999: 27f.); here: the Sankt Florian Psalter (cf. fig. 2)

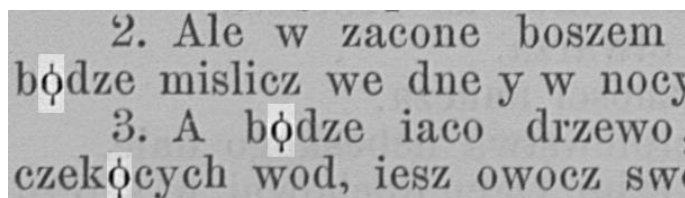


Fig. 11. Scholarly edition (Nehring 1883: 2) of the Sankt Florian Psalter (cf. fig. 2)

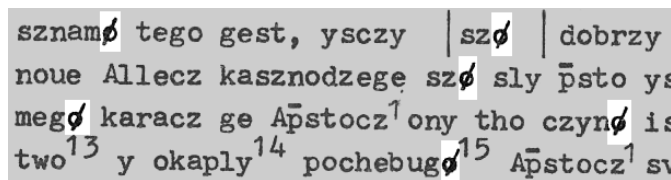


Fig. 13. Manually slashed o in a scholarly anthology (Siatkowska 1991: 275); here: the Gniezno Sermons (cf. fig. 5)

Phonem	Graphem	Beispiele
/e/	⟨ϕ⟩	ſϕ narodil (= się narodził),
/e/	⟨ϕ⟩	meſkajϕch (= mieszkając),

/e/	⟨ϕ⟩	bϕdze (= będzie),
/e/	⟨ϕ⟩	mϕsz (= mąż),

Fig. 14 (left) and **fig. 15** (right). Both phi and ø in one and the same linguistic text (Mazur 1993: 158, 160)

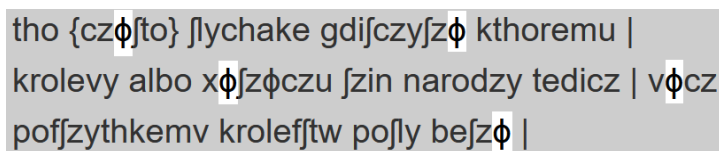


Fig. 16. Phi symbol in the article “Kazania gnieźnieńskie” in the Polish Wikipedia (another excerpt not in fig. 5; the article was created in 2005, but the text excerpt was only added to the article in 2018 by user Witia)

4 Proposed characters and their properties

I therefore propose to encode a new Latin small letter ϕ together with its upper-case counterpart Ŧ in the block “Latin Extended-D”:

A7C0	Ŧ	LATIN CAPITAL LETTER OLD POLISH NASAL VOWEL
A7C1	ϕ	LATIN SMALL LETTER OLD POLISH NASAL VOWEL

Since there are no words beginning with a nasal vowel in Old Polish, the **capital letter** does not seem necessary at first glance. However, it should be included for systematic reasons because in general Old Polish does have capital letters, as do the modern scholarly editions (cf. Górski et al. 1955: 28). The inclusion of the capital letter enables various styles of capitalization (including all-caps and having a capital Ŧ as the second letter of the first word of a paragraph starting with an ornamental initial). Moreover, capital Ą and ĕ have been included in Unicode for Modern Polish (as U+0104 and U+0118, respectively), although nasal vowels do not occur word-initially there either.

The characters ought to be added to a block of the Latin script in the **Basic Multilingual Plane** (BMP), so that editions of and scholarly papers about Old Polish texts, which otherwise exclusively use letters from the BMP, do not have to draw this single, very frequent character from a different plane.

As to **presentation**, since the shape ϕ with two vertical lines attached is the preferred form both in the scholarly editions (fig. 6–9, cf. Górski et al. 1955: 28) and in the most calligraphic Old Polish manuscripts (fig. 2 and 3), this is the shape that all fonts that are not designed to faithfully represent the layout of a particular manuscript should adopt. However, fonts might also include the shape ø

with the long vertical line as a stylistic alternate glyph, making sure that it looks sufficiently different from the characters listed in 5 below.

The **name** of the letter might alternatively also describe the shape rather than the function (e.g. LATIN CAPITAL/SMALL LETTER O WITH VERTICAL LINES AT TOP AND BOTTOM), but this description would exclude alternative shapes like ϕ , which is why I propose the ‘functional’ name given above. A further, shorter, alternative would be LATIN CAPITAL/SMALL LETTER HORNED O, referring to the letter’s name *o rogata* ‘horned o’ in colloquial use among Polish specialists (e.g. Bień 2020: 2).

The **character properties** should be the same as for other Latin letters, with the two characters case-folding to each other:

A7C0;LATIN CAPITAL LETTER OLD POLISH NASAL VOWEL;Lu;0;L;;;;N;

LATIN CAPITAL LETTER OLD POLISH NASAL VOWEL;;;A7C1;

A7C1;LATIN SMALL LETTER OLD POLISH NASAL VOWEL;Ll;0;L;;;;N;

LATIN SMALL LETTER OLD POLISH NASAL VOWEL;;;A7C0;;A7C0

In **collation**, ϕ ought to be sorted after \mathfrak{a} (U+0105 LATIN SMALL LETTER A WITH OGONEK), so that e.g. in dictionaries words spelled with the unique Old Polish nasal vowel will be found where words with the first of the two modern Polish nasal vowel letters would be. [Depending on the concrete word, modern users might expect to find it under either \mathfrak{a} or \mathfrak{e} , but of course in collation we have to decide for one of the alternatives. The choice of \mathfrak{a} is based not only on the fact that it happens to be first in the alphabet but also on the fact that on the basis of ϕ Stanisław Zaborowski in 1514 developed the shape \mathfrak{a} , which (via \mathfrak{a} in blackletter prints) ultimately resulted in \mathfrak{a} , whereas the shape of \mathfrak{e} (U+0119 LATIN SMALL LETTER E WITH OGONEK) was directly taken over from the Latin *e caudata* (cf. Bunčić 2012: 229, 235–236).]

A free OpenType **font** that contains the letter in the proposed position (“FreeSerif DB”, based on FreeSerif) as well as an “Old Polish” **keyboard layout** for Windows to simplify input of this and other characters needed for writing older forms of Polish are available from <https://slavistik.phil-fak.uni-koeln.de/buncic/tech#c197307>. (The font **JuniusX** by Peter S. Baker contains several variant glyphs of the character in the Supplementary Private Use Area. As of version 1.007, this includes at U+F0011 an *o* with short lines at the top and bottom, but the lines are oblique rather than vertical; a variant with vertical lines will be included in a later version, as [the author announced on 21 Dec 2020](#).)

5 Similar characters

A number of characters bear some resemblance to the proposed characters:

00D8	Ø	LATIN CAPITAL LETTER O WITH STROKE
00F8	ø	LATIN SMALL LETTER O WITH STROKE
0278	ϕ	LATIN SMALL LETTER PHI
03C6	φ	GREEK SMALL LETTER PHI
03D5	ϕ	GREEK PHI SYMBOL
233D	⊙	APL FUNCTIONAL SYMBOL CIRCLE STILE
23C0	⊙	DENTISTRY SYMBOL LIGHT VERTICAL WITH CIRCLE
29B6	⓪	CIRCLED VERTICAL BAR
2C65	ⱥ	LATIN SMALL LETTER A WITH STROKE

Another theoretical alternative would be a combination of *o* with U+20D2 COMBINING VERTICAL LONG LINE OVERLAY to form ϕ . However, U+20D2 is in the Unicode block “Combining diacritical marks *for symbols*”, so that font designers often do not provide for these combining marks to combine properly with *letters*. Moreover, the ‘canonical’ glyph ϕ of the proposed character (cf. Górski et al. 1955: 28) cannot be formed with existing combining marks, nor would it be efficient to introduce a

mark “combining short vertical strokes at top and bottom”, for which font designers would have to make sure that the lines merge with the outlines of the *o* when combined with it. (The existing diacritics U+030D COMBINING VERTICAL LINE ABOVE and U+0329 COMBINING VERTICAL LINE BELOW do *not* touch the letter by design: ȯ, ȱ).

While, as pointed out above, ȱ was the inspiration for the creation of the letter that turned out as ą in the 16th century, it cannot be treated as a variant of ą or of ȁ, because both ą and its forerunner ȁ always signify the Polish back nasal vowel /ɔ̃/ *in opposition to* the front nasal /ɛ̃/, which is represented as ę. ȱ, by contrast, always signifies the Old Polish *intermediate* nasal /ã/, which was the precursor of both Modern Polish nasal vowels. This is why it is linguistically important to distinguish ȱ as a ‘neutral’ nasal vowel letter from the modern letters ą and ę and why scholarly editions have always avoided using ą (or ę) to represent ȱ.

Although some of the characters above have already been used as makeshift solutions to represent the proposed character (see fig. 12–16), all these characters have a different semantics than ȱ, and none of these Unicode characters ought to be represented by the glyph ȱ in any font. That is, if the use of, say, ø for the Polish nasal vowel were encouraged because its shape has been considered more or less acceptable by some scholars (see fig. 13 and 15), font designers might want to improve the visual representation of Old Polish by giving U+00F8 LATIN SMALL LETTER O WITH STROKE the shape ȱ sanctioned by Górski et al. (1955: 28), which would result in unacceptable forms like Danish *København* or Faroese *Føroyar*. Only by distinguishing the Old Polish nasal vowel *functionally* from these similar-looking characters can its variety of glyph shapes, both in the original manuscripts and in modern scholarly literature, be adequately represented in Unicode.

6 Sponsor

The address for correspondence is:

Daniel Bunčić
Cologne University, Slavic Department
Weyertal 137
50931 Köln
Germany

Phone: +49 2232 1504280

E-mail: daniel.buncic@uni-koeln.de

Homepage: <https://slavistik.phil-fak.uni-koeln.de/buncic>

References

- Bernacki, Ludwik (ed.). 1930. *Biblia szarospatacka: Podobizna kodeksu bibljoteki ref. gimnazjum w Szarospataku* [The Sárospatak Bible: Facsimile of the codex of the library of the secondary school in Sárospatak]. Kraków: Polska akademja umiejętności.
- Bień, Janusz S. 2020. Standard Unicode i dawne teksty polskie: Uaktualniony postprint [The Unicode standard and older Polish texts: An updated postprint]. *Zenodo*. <https://doi.org/10.5281/zenodo.4058702>.
- Bunčić, Daniel. 2012. The standardization of Polish orthography in the 16th century. In Baddeley, Susan & Voeste, Anja (eds.), *Orthographies in Early Modern Europe*, 219–257. Berlin: de Gruyter. <https://doi.org/10.1515/9783110288179.219>.
- Gehrmann, Maria & Hentschel, Gerd & Menzel, Thomas. 1999. *Textlehrbuch zum Altpolnischen: Eine exemplarische Einführung in das altpolnische Schrifttum bis 1543 für die universitäre Lehre* [Textbook on Old Polish: An exemplary introduction to the Old Polish writings until 1543 for universities]. Oldenburg: Bibliotheks- und Informationssystem der Universität Oldenburg.
- Górski, Konrad & Kuraszkiewicz, Władysław & Peplowski, Franciszek & Saski, Stefan & Taszycki, Witold & Urbanczyk, Stanisław & Wierczyński, Stefan & Woronczak, Jerzy. 1955. *Zasady wydawania tekstów staropolskich: Projekt* [Principles of editing Old Polish texts: A project]. Wrocław: Zakład imienia Ossolińskich.

- Mazur, Jan. 1993. *Geschichte der polnischen Sprache* [History of the Polish language]. Frankfurt am Main etc.: Peter Lang.
- Nehring, Władysław. 1883. *Psalterii florianensis partem polonicam ad fidem codicis recensuit, apparatu critico indice locupletissimo instruxit Wladislaus Nehring* [The Polish part of the Sankt Florian Psalter, faithfully edited according to the codex and endowed with a critical apparatus and a most comprehensive word index by Władysław Nehring]. Posnaniae: apud J. K. Żupański.
- Siatkowska, Ewa (ed.). 1991. *Wybór tekstów zachodniosłowiańskich w ujęciu porównawczym* [A selection of West Slavic texts from a comparative perspective]. Warszawa: Wydawnictwo Uniwersytetu warszawskiego.
- Urbańczyk, Stanisław & Kyas, Vladimír (eds.). *Biblia królowej Zofii (szarospatacka) wraz ze staroczeskim przekładem Biblii* [Queen Sophia's (Sárospatak) Bible with the Old Czech translation of the Bible]. Wrocław etc.: Zakład narodowy imienia Ossolińskich.
- Vrtel-Wierczyński, Stefan. 1950. *Wybór tekstów staropolskich: Czasy najdawniejsze do roku 1543* [Selection of Old Polish texts: From the most ancient times to 1543]. 2nd edn. Warszawa: Państwowe Zakłady Wydawnictw Szkolnych. [1st edn. Lwów 1930, 3rd edn. Warszawa 1963, 4th edn. Warszawa 1969.]
- Wydra, Wiesław & Rzepka, Wojciech Ryszard (eds.). 1984. *Chrestomatia staropolska: Teksty do roku 1543* [Old Polish anthology: Texts before 1543]. Wrocław etc.: Zakład narodowy im. Ossolińskich.