Unicode 2C1A – Glagolitic “Pe”: Fact or Fiction?

Sebastian Kempgen (Bamberg, Germany)

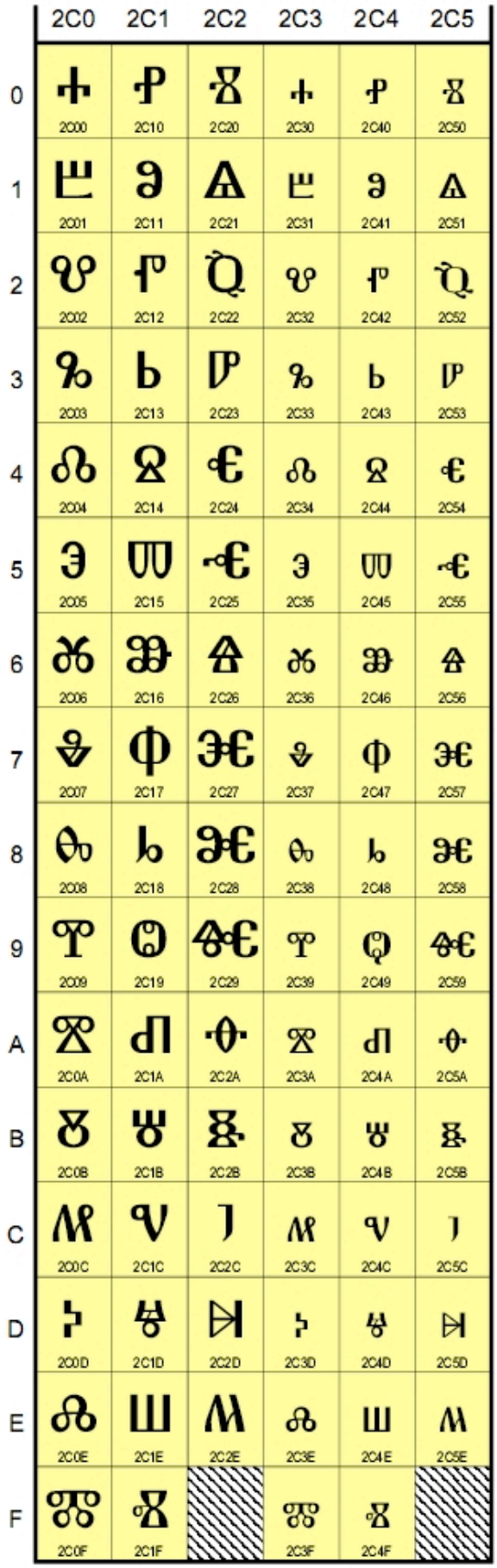
1. Introduction. In recent years, Unicode has become a buzz-word within the information technology industry. As an encoding standard for all the world’s characters, it can be said to have fulfilled its promise to serve as a platform-independent open standard for assigning each and every character a unique (numerical) code, so that a file containing any given character can be exchanged between users (scholars, publishers, printers, etc.), and all characters will always be displayed correctly—assuming of course, that an adequate set of fonts is available to each user.[1] To make this possible, the OpenType TrueType and OpenType PostScript font formats have been developed. Fonts whose suffix ends in .ttf or in .otf most likely are Unicode fonts and follow this standard.[2] Today, the same .ttf and .otf files can be used on all computing platforms, irrespective of operating system, and because the font files are the same, users can be sure that files containing characters from these fonts will be displayed correctly.[3]

---

[1] Because this is not necessarily always the case, operating system vendors can and should implement strategies to handle situations where a character must be displayed although the original font is not available. A good choice is to have a “fall back” or “last resort” font in the operating system; under Mac OS X, for example, Lucida Grande, the system font, fulfills that role. If such a mechanism has been put into place, each character that comes from a currently installed font will be correctly displayed using that font, and characters for which the correct original font is missing will be displayed using the last resort font. An additional consideration is the question of how to interact with the user in such situations: should the last resort font be used tacitly, without notifying the user that some characters could not be displayed in their original fonts and a substitute font has been used instead? Or should the user be notified and given a chance to react to the situation? Surely the latter is better, especially because, if the last resort font and the actual font are very similar, the substitution might go unnoticed.

[2] In practice, this may not always be the case. When users speak of Unicode fonts, they assume them to be fonts where each character has been put into the character slot allocated to it in the Unicode standard, but nothing prevents a font designer from putting a “wrong” character into a slot, either by mistake or deliberately. Inadvertent errors in assigning characters to slots are not uncommon, and deliberately using a wrong slot is a common strategy for giving users a character they need that does not have a proper place assigned in the Unicode inventory.

[3] Using Unicode fonts, however, does not prevent the user from making mistakes in selecting the correct character. One example is the transliteration of the soft and the hard sign in Cyrillic; both have their own counterparts in Unicode, but out of tradition (and because these characters are seldom present in standard fonts like Times New Roman) most users will use the curly quote instead, which may look better than the actual character, but which is nonetheless incorrect from an encoding perspective.

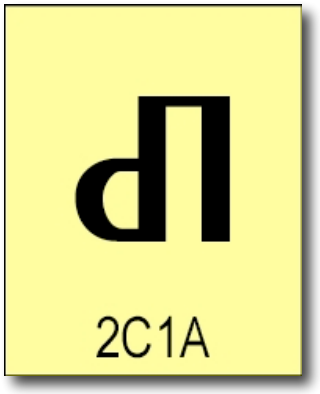


2. Unicode is not a fixed standard. It is a standard under continuous development, to which new characters are added with every revision. For an overview of characters of interest to Slavists that were missing from version 4.1 of Unicode, see two articles by the author on the subject (Kempgen 2006a, 2006b), while on the new Slavic characters in Unicode 5.1, see Kempgen (2009). One of advances in Unicode version 4.1 in 2005 was the addition of the Glagolitic script. The figure shows the Unicode block (table) containing the Glagolitic characters that were added then (the table has not changed since). In Unicode, characters are arranged in cells of a large table, with each cell having a unique number, seen here in fine print at the bottom of each cell. That hexadecimal cell number can be arrived at by combining the number of the column with the number of the row. For example, "2C0" (first column) with "0" added at the end (first row) is the number of the Glagolitic "Az," i.e., 2C00. All characters are available in uppercase and in lowercase forms even if such a concept is or was questionable for the given script at the time of its original use.

Now let us turn our attention to character number 2C1A (and 2C4A). This character is called "Glagolitic Pe." Because Unicode character charts such as this one are now widely accessible and also because they form the basis of the user interfaces designed by operating system vendors, this chart will most likely be the first occasion a user (or font-designer!) will ever see the Glagolitic "Pe." For most Slavists, it will also be the last encounter they have with this character, and more curious Slavists might be tempted to learn something more about it.

This is starting point of our contribution.

3. Glagolitic "Pe." Let us first stress a simple fact about the character,[4] which is called "Pe," or "Pě," depending on the language or the context. The most astonishing fact about this character is that *there are no actual texts in which it appears*! Yet here in the Unicode table for Glagolitica this character is included as part of the alphabet, it has a well-defined shape, and it even has uppercase and lowercase versions. So what exactly is the matter with this character and what is special about it?[5]



4. Abecedaria

Abecedaria have rightly been called "a key to the original Slavic alphabet" (Marti 1999). In this article, Marti discussed extensively all relevant questions concerning abecedaria that have to be solved in reconstructing the original Slavic alphabet as invented (or developed) by Konstantin-Kiril, and we will not repeat all of those arguments here. We will concentrate instead on letter names, as "Pe" obviously *is* a letter name. Marti calls letter names "optional" elements of alphabets in the sense that they are not required for an alphabet to fulfill its basic function, he argues that letter names "were considered to be indispensable elements of an alphabet by Constantine-Cyril" (1999, 177), and he also gives examples for several types of letter names (1999, 178): a logographic tradition, as in Hebrew (where a letter name is a word starting with the sound represented by the letter); a loan-word tradition, as in Greek (whose letter names have been borrowed and adapted along with the letters themselves); and a "minimal solution," as in the Latin tradition (where vowels represent themselves and a vowel sound is added (either before or after) consonants to facilitate their pronunciation (cf. German "em" in contrast to "ka"). It is obvious that the Slavic languages have changed the type of their alphabet names in historic times; they follow the Latin model today, but the original system followed the Hebrew

[4] We will treat the uppercase and the lowercase shapes as *one* character for the remaining of the article, which means that we will speak about one character (singular), not a character pair or characters (plural).

[5] This puzzling character has already been the subject of many articles. Space does not permit us to go into the details of all of them, but the most important contributions include Nahtigal (1923), Mareš (1971), Mošin (1973), and Velčeva (1973). At this point, we might point out that Unicode as an encoding system defines descriptive character names for all its entries. Unicode *fonts*, however, do not need to use these names to work correctly, since TrueType fonts rely only on the Unicode numbers for their characters. The default character name in TrueType fonts is simply "uni…" followed by the Unicode number of the character, for example "uni2C1A."

pattern, as is evidenced by the alphabet acrostics. As far as the Glagolitic alphabet is concerned, it is well known that we have several Glagolitic abecedaria, but none of them directly represents the original alphabet. All of them, even the oldest (the famous Glagolitic alphabet on the wall of the "Golden Church" in Preslav), are of a later origin, they may be only partially preserved, and mistakes and corruptions in the ordering and/or in the shapes are a matter of fact, especially when Cyrillic alphabets come into play as a source for reconstructing the original Glagolitic alphabet. For our further discussion it is important to recall what Marti has to say about the peculiar situation in Slavic: "some sounds (and consequently also the letters representing them) do not appear in initial position. In these cases, the alphabetic texts have two options: either the letters are left out completely or they are represented by a substitute. Regardless of the solution chosen, alphabetic texts do not provide reliable information in these cases" (1999, 180). In the following sections we will examine some abecedaria in more detail. The **Preslav Abecedarium** itself is worthless here because it is incomplete and misses exactly the portion of the alphabet that would be needed.

5. The Paris Abecedarium

This is one of the most famous early (eleventh- or twelfth-century) Slavic abecedaria. It is a Glagolitic alphabet with letter names written above the Glagolitic letters in Latin script. The spelling of the names clearly indicates a Romance provenience of the alphabet, and it is what Marti (1999, 181) calls a "xenographic" alphabet—not written by a native speaker and thus "susceptible to mistakes and distortions." This alphabet has a letter named "pe" in the sequence "fort • ot • pe • ſaraué • ſa" (see illustration). From the context, it is obvious that the "long s" denotes the sound [š], see for example the letter "ša." From the letter name "cſlóuo" (not pictured here) it can be deduced that the letter "u" stands for the sound [v]. Consequently, the letter name "ſaraué" is nothing but "červ" with the initial affricate simplified to a fricative. Because of the precise spacing of the letter names it is also reasonable to assume that the names were written first and the Glagolitic letters were then written below them.

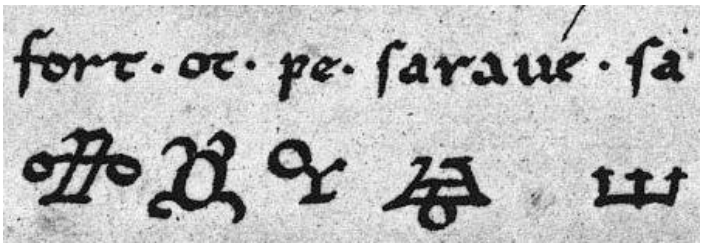


Fig 3: The "ot – pe" section from the Paris Abecedarium

It was pointed out long ago that the Paris Abecedarium contains errors, one of them pertaining to the sequence containing the "pe": the letter names do not

always match the Glagolitic letters they accompany. The relevant letters themselves are easily and clearly recognized: they are the so-called "spidery h" and the "ts." This means that the mistakes are in the letter names. If we correct the alphabet, it would look like this:

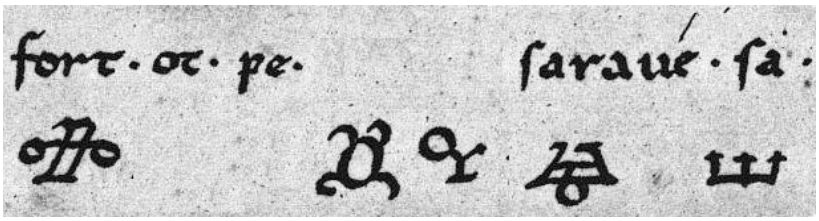


Fig 4: Corrected "ot – pe" section from the Paris Abecedarium

In other words, we have two letter names which do not have accompanying letters, and we have two letters that are missing their names. After this confusing part, the correctly matching sequence of letter names and letters starts again. Thus, the Paris Abecedarium does give us a letter name "pe" but does not give us a picture of that character.

Can the mistake in the sequence of letters and letter-names in the Paris Abecedarium be explained? We think it can. One just has to imagine a scribe familiar with Latin character shapes and perhaps superficially familiar with Cyrillic letters (although this latter assumption is not strictly necessary). At the time when the alphabet was written, the "spidery h" had long lost its use and was hardly known any more. If you "think Latin" (or Cyrillic), the shape of the "spidery h," i.e., ⱨ, is easy to describe as something like "an o with diacritical parts"—just as the Cyrillic and Glagolitic letters named "ot" (i.e., ѻ or ⱁ), which can be described in exactly the same way! Thus, the letter name "ot" may not be simply a strange mistake, but, rather, the result of a confusion based on optical similarity. Marti (1999, 184) advances this explanation, but he does not offer an explanation for the pairing of ⱁ and "pe".

6. The Munich Abecedarium

The Munich Abecedarium consists of two alphabets, one Cyrillic and one Glagolitic, and is to be found on the last page of a tenth-century Latin codex kept in the Munich State Library (sig. CLm. 14485) under the name of the "Hrotsvitha manuscript." The Slavic alphabets were added in the second half of the eleventh century or at the beginning of the twelfth century. The author (see Kempgen 2007) has published the abecedarium for the first time in color with comments on many individual letters, along with an overview of research related to these alphabets. As for their origin, Trubetzkoy (1930, 31) clearly states: "Die slavischen Alphabete am Ende der Hrotsvitha-Handschrift müssen von jemandem eingetragen worden sein, der diese Alphabete gar nicht kannte und die slavischen Buchstaben mechanisch von einer alten Vorlage abzeichnete" [The Slavic alphabets at the end of the Hrotsvitha manuscript must have been added by

someone who was not familiar with these alphabets and who copied them mechanically from an old sample.]. Durnovo (1930, 32) also confirms that judging from the mistakes in drawing the Cyrillic as well the Glagolitic characters, the scribe could not have been familiar with either alphabet.

With this in mind, lets us now turn to the relevant portion of the alphabets. The Cyrillic alphabet has a strange-looking shape in position 26 between two neighboring characters, which are clearly to be identified as **ѡ** and a horizontally flipped (or somewhat distorted) **ц**, see picture. The character in position 26 is normally perceived as a **п**, although it looks more like a ligature of П (or Г) and Ⅼ, and it is the second occurrence of the letter "p" in the alphabet. The two shapes are not identical; nevertheless, the character sequence is usually understood as denoting **ѡ п ц**.

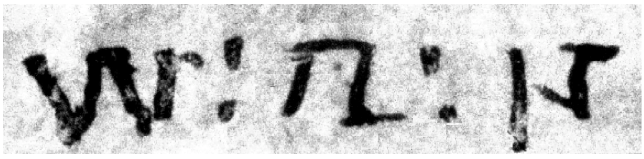


Fig 5: Part of the Cyrillic half of the Munich Abecedarium

In the Glagolitic alphabet there is also a strange character in position 26, again between two clearly identifiable neighbors, see illustration (left half).

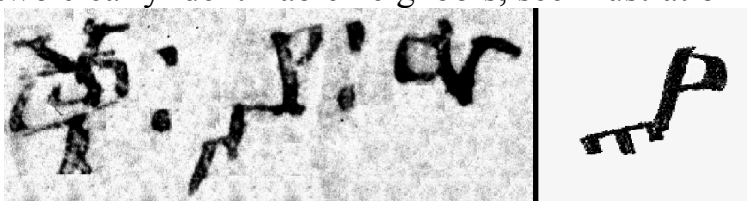


Fig 6: Part of the Glagolitic half of the Munich Abecedarium

The similarity of the mysterious Glagolitic letter no. 26 to ⰻ (i.e., **ѣ**) in the Prague fragments (see right half of illustration) was pointed out long ago, and they are, indeed, very similar. Reading the letter that way, the illustrated sequence in the Glagolitic alphabet has been reproduced as **Ⰺ : ⰻ : Ⱛ** by the author, where the **ⰻ** is considered a mistake by the scribe who did not seem to know what to make of the original before him. Therefore he is thought to have used a shape similar to a Latin "p," with added strokes, making it somewhat similar to the Latin abbreviation for "pro." Thus, the chosen letter could at least represent the initial sound (i.e., [p]) of the acrostic. As it is, in both alphabets the letter seems to be a "p"; in the Cyrillic alphabet, it is similar to a Cyrillic "p" with an added stroke to the right, and in the Glagolitic alphabet, it is similar to a Latin "p" with strokes added to the left. It should also be noted that both alphabets are mixed up in at least one position: there is a Glagolitic character in the Cyrillic alphabet and vice versa. Let us stress that the Munich Abecedarium itself does not have any letter names, so any names given to the two mysterious

letters are based on structural correspondences of the sequences these letters appear in to the sequences in other abecedaria. Hamm (1974), by the way, reads the mysterious letter in the Glagolitic alphabet as a Ⱈ.

7. Abecedaria Sinaitica

The newly discovered **Abecedaria Sinaitica**, edited and commented upon by Marti (1999), are of direct relevance to our question. They contain a Greek, a Latin, and a Glagolitic alphabet (SinG), all by the same scribe on the same page, and another, "better" Latin alphabet by another scribe on the next page. The Glagolitic alphabet is missing several parts due to damage to the parchment, but the sequence that is of interest to us has been preserved. This is what it looks like in context, with a little tweaking to enhance the contrast and remove the background:

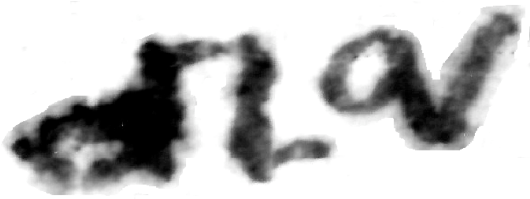


Fig 7: Glagolitic letters "Pě" and "Tse" from SinG

The letter preceding the "Pě" has been only partially preserved, but there is no reason *not* to assume that it is an "Ot." Thus, the sequence is ⰖⰔ [Pě] ⱝ.

Marti points out the importance of this alphabet, as it is an autochthonous one, and, furthermore, the same scribe has left additional notes in Glagolitic, so the shapes of the letters in the alphabet can be compared to the shapes used in writing running texts (as opposed to single letters). The mysterious "Pě," however, is among those characters that do *not* appear in the text. Marti stresses (1999, 192) that the Glagolitic Abecedarium Sinaiticum "confirms" the existence of the letter, which he denotes as "P_2." See also Miklas (2004, 394ff.) on the importance of SinG for all questions related to the "second p."

8. The graphical shape of the Sinai Pe

It is obvious that the shape of Glagolitic "Pě" in the Unicode tables has been based upon this very letter. Therefore, let us have a closer look at the letter and its shape.[6] Close inspection reveals that the basic part of the Sinai "Pě" is identical to a character shape occurring in the Munich Abecedarium – namely, the "second pe" in the *Cyrillic* half of the Abecedarium.[7] The similarity between them can hardly be a pure coincidence:

[6] A small (and flipped) picture of the alphabet is reproduced by Tarnanidis (1999, 167).

[7] We have independently reached the same conclusion here as Koch (2004, 439). He does not, however, represent the shape correctly, using a *dot* instead of a *line* for the right-most part (ⰔⰏ). This may be due to the low quality of the reproduction he had to work with.

Fig 8: Sinai, Munich-Cyrillic and Unicode "Pě"

There are several possible logical consequences: the "P_2" shown above (Fig. 8, left) may be a *Cyrillic* character, and thus is out of place in the Glagolitic alphabet (improbable); or it may be a *Glagolitic* character, in which case the corresponding glyph is out of place in the Cyrillic part of the Munich abecedarium; or the Munich character must be explained differently (possible). Alternatively, this character might have had an identical shape in *both* the Cyrillic and Glagolitic alphabets, which would require further explanation. What should be assumed cannot immediately be decided, but we are inclined to see this as a Glagolitic character. Obviously, though, it had enough relationship to a Cyrillic П from the perspective of the scribe of the Munich Abecedarium that he used it in that position.

Of course it is next to impossible to derive a prototypical form of any character from a single occurrence only. However, we can nevertheless take a closer look at the Sinai "Pě" in comparison to the shape of the Unicode "Pe" (Fig. 8, right). Closer inspection of the Sinai "Pě" in very large magnification reveals that it is composed of the following strokes:

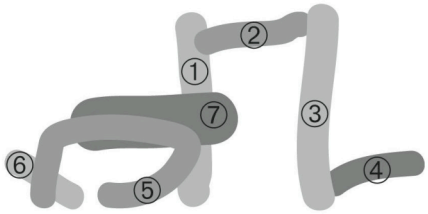


Fig 9: Strokes of the Sinai "Pě"

Obviously, the scribe of the Sinai letter was not familiar with this shape – and how could he have been when the letter never occurs in any texts? That the scribe had difficulty in writing the letter can be seen from stroke 7, which seems to be a correction made to the letter after its other parts had been drawn. Strokes 1 to 4 are the ones that together form the same shape as the "second p" of the *Cyrillic* character from the Munich Abecedarium. Strokes 5–7 together have been interpreted for the Unicode glyph representation as forming half a circle, directly connected to the stem (stroke 1). However, the strokes really seem to reveal another structure. Strokes 5 and 6 actually seem to form one shape together, a circle, which is broken up on its base point. This, however, seems to be simply a deficiency in the writing. Other characters in the alphabet show the same effect where "normally" there would be a continuous line. Stroke 5 does

not connect to the stem (stroke 1) in a horizontal line, as the Unicode shape makes us believe; rather, strokes 5 and 6 form a circle that touches the stem (stroke 1). Now what could be the reason the scribe found it necessary to correct the letter he had drawn so far? It seems plausible that he wanted to indicate that the connection between the circle and the stem was not a section of the curve but a straight line. This, in turn, makes sense only if we assume that the circle should really *not* touch the stem. It seems to us then that the underlying structure of the letter could be reconstructed as follows, i.e., in a structure that is partly identical to the lower half of a character like ѱ:

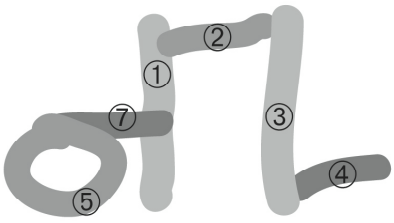


Fig 10: Strokes of the recontructed "Pe"

Compare this, once again, to the shape given to that character in Unicode. First, and not mentioned so far, is the fact that the Unicode shape is inexplicably missing the right-most stroke (4). The stroke is clearly present in the Munich Abecedarium as well (although in the mysterious glyph from the Cyrillic alphabet!), but if it is present twice in an unambiguous form, it should be assumed to be a part of the canonical shape. Second, in our opinion at least, the left part of the character could also be a circle that is connected at its top with the stem by a separate stroke, not a half-circle directly joining the stem.

9. The "Pě" in the Abecedaria

Let us sum up the situation. We have a letter name "Pe" in the Paris Abecedarium (which stands above the letter shape ᗐ – clearly a different character). We have *two* identifiable glyphs in the same slot between ᗅ __ ᗐ in position no. 26 in *Glagolitic* alphabets, one in the Munich Abecedarium and one in the Abecedarium Sinaiticum. Additionally, we have an unidentifiable glyph in the same position in the *Cyrillic* portion of the Munich Abecedarium. However, no-one has claimed that the mysterious twenty-sixth letter was ever part of the *Cyrillic* alphabet. Its occurrence in the Cyrillic part of the Munich Abecedarium can only be explained by the tendency of the scribe to have a one-to-one correspondence between the Cyrillic and the Glagolitic parts. The Glagolitic letter 26 in the Munich Abecedarium, however, clearly shows a different shape (cf. again Fig. 6). In other words: we have a name ("Pe") and we have two related mysterious shapes from Abecedaria, one in the Glagolitic alphabet of the Abecedaria Sinaitica, one in the Cyrillic half of the Munich Abecedarium. "Other than in these abecedaria it [the letter "Pe" – S.K.] does not seem to have been used at all" (Marti 1999, 193). The shape of the Glagolitic "Pe" in Unicode does not faithfully reproduce all unambiguous parts from these two instances of

the letter (the rightmost stroke is missing), and the leftmost part can be reconstructed differently.

In a footnote, Marti has this to add: "In general it must be said that the sound value of this letter and the rationale of its inclusion in the alphabet remain a matter of speculation. Its existence in this position, however, is now beyond doubt" (1999, 193, Fn. 44).

Marti (and other authors) do not undertake to reconstruct the "true" shape of the "Pe." Instead, Marti uses the notation "P_2," while Veder (2004) calls the letter simply by the name "Pe." Below is a section from Marti's table of the Glagolitic alphabet, where the third column represents the Paris Abecedarium, the fourth column the two Munich Abecedaria, and the last column the Abecedarium Sinaiticum:

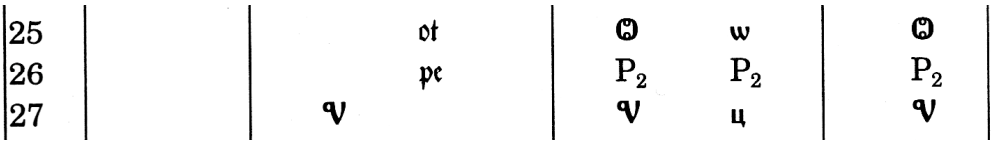


Fig 11: Glagolitic letters 25 to 27 from Marti (1999, 199)

The only purpose that the Unicode "Pe" really has at present is to reproduce *a single instance of a dubious shape.* There are no other occurrences besides the character from SinG that one would choose to represent by that glyph; obviously, it is not suited to reproducing either the Cyrillic or the Glagolitic letter from the Munich Abecedarium. The shape of the Unicode "Pe" omits one clearly visible stroke, with other parts of the character open to a slightly different reconstruction. In other words, the empirical basis for the reconstruction of the shape of this character is the smallest possible one, and its use is limited to the same single occurrence. Since all researchers actually refer to this character by its descriptive names, and not by using a shape, it would have been sufficient (and possibly better) for the time being *not* to include the "Pe" in the Glagolitic entity set.

That said, let us stress the consequences of our view regarding the status of the Unicode "Pe" and the shape of its glyph. It is a fact that the Abecedaria mentioned above do have a character in position no. 26; the problem is that it is not clear *which* character it is and how it should be *drawn*. The only thing that *is* clear is the fact that the name of the character is "Pě." With these preliminary results, let us now turn to the alphabet poems and to the famous treaty "About the letters".

10. Černorizec Xrabr, O pismenex

Some important evidence regarding the "Pě" can also be gleaned from the text about the Slavic alphabet commonly attributed to Xrabr. Let us cite the most important phrases from the text as reconstructed by Veder in his translation: "And he made them 30 letters and 8, some, indeed, according to the system of

the Greek letters, others, however, according to the Slavic speech" (1999, 98f.). If the author is talking about the Glagolitic alphabet, as it is commonly assumed, then it is impossible to assume that he is talking about the *design* of the characters, since Glagolitic letters certainly look different from Greek. Consequently, this passage can only mean that he had Greek *sounds* in mind, and that the expression "according" ("по чину") means the *order* of the Greek letters. Veder's reconstruction of the "Greek-influenced" letters themselves ends in the sequence "…R S T U F Ѡ, and Ps Ks Th" (1999, 108) in the transliterated variant. And because the shape of the "Psi" is not attested in Glagolitic, Veder substitutes the Greek letter for it in the reconstructed Glagolitic text. One of the sources cited by Veder (β), in general considered to be the most reliable, switches from letters to letter names for the last four items of the sequence: "**… п ρ ϲ т ѵ ф χ ѿ и пѣ χлъ тъ**" (1999, 109). All other sources either have **ρ ϲ т ѵ ф χ ѱ ѡ** or the sequence **ρ ϲ т ѵ ф χ ѱ ѡт ї п ѧ**. This last variant obviously restores the letters (instead of letter names) for the last two characters, resulting in the occurrence of a "second p" (not "pĕ") in the alphabet. What we can conclude from the name "pe" for "psi" is that the Greek affricate was not perceived in Slavic as such, but rather as a sequence of two consonants.

From these sources, we again have the letter name "pĕ," and a **п** as a representation of the letter (in Cyrillic). Two points merit being repeated: the author mentions the character "pĕ" among those that have been modeled after Greek, and he mentions it at the end of this list in a special group, denoting "alternate" sounds, digraphs or allophones in Greek.[8] This suggests that "pĕ" must also be "a variation" of some similar Greek letter (or sound). If the Slavic name begins with a [p], the letter should have some phonetic relationship with a "base [p]." There are only two possible candidates in Greek: "psi" [ps] and "phi" [p^{h}].

It has been noted before, and we support this assumption, that **ѡ** was not part of the original Glagolitic alphabet – *if and only if this means* **ѡ** *in the function of designating the sound sequence* [št].[9] This does not, however, imply that **ѡ** could not have been present *in another function*. Because of the obvious similarity in shape, we are inclined to consider **ѡ** to have originally been the Glagolitic reworking of the Greek "psi" (**ѱ**). It is entirely possible, though, that the **ѡ** might have originally been turned left, right or upside down, as is the case with several other characters. It is telling that *all three* Greek-derived letters mentioned in "O pismenex" at the end (i.e. **пѣ χлъ тъ**) have only slightly disguised Greek letters (**ѱ ξ θ** or **ѱ χ θ**) as the base of their Glagolitic graphic

[8] This would be our answer to the question Marti (2004, 413) put forth regarding the position of these letters at the end of the sequence.

[9] Marti (2004, 411) believes that **ѡ** *was* part of the original Glagolitic alphabet, but in position 31, and only later moved to position 26, i.e., "became 'pĕ,'" when the original character designated by that name went out of use. From this it would follow that **ѡ** would have changed places twice, later to the position after **ш**.

equivalents (ѡ ѯ ѳ).[10] In their function as digraphs, however, neither a Slavic "ps" nor a Slavic "ks" was really needed.

In the lists of "Slavic letters" in "O pismenex", we find the sequence **ц ч ш щ ы ь**; here, **щ** has already moved to its modern position in the alphabet, right after **ш**. But it has not only moved to a new position; **щ** has also acquired a new sound value, namely [št].

11. The Acrostics (Alphabet poems)

The acrostics have a well-known twenty-sixth line starting with words whose first letter is a "p": **Печаль моѭ на радость прѣложи** "Turn my sorrow into bliss" (Veder 1999, 73 and 154). Of course, the acrostics already mention a "normal" **п** some lines before this one. The existence of line 26 surely means that it was meant to represent a specific letter, and this letter was given the name "pě" in certain sources. Does this mean that this letter is an *unknown* letter? No! Veder represents the beginning of the line with a Cyrillic **ѱ**. As for the Glagolitic letter in question, which is supposed to have served as the base for this letter, he argues that "The evidence for **ⱁ** in this place is undoubtedly more convincing than that for any other letter" (1999, 73). Now any problems in the phonetic correspondence would be immediately resolved if one thinks of **ⱁ** as being the original Glagolitic equivalent to Greek **ψ**.

As we said, OCS really did not have very much use for [ps] or [ks] or a second [x], and even [f] and [th] were quickly confused, and so it seems that **ⱁ** was soon restricted to its established use as a number ("800") or simply given up in its original function. When Konstantin re-created the Glagolitic alphabet and created the Cyrillic alphabet in Preslav, he borrowed the Greek **ψ** to form a corresponding Cyrillic letter: **ѱ**. The Glagolitic letter **ⱁ** could now be re-used in a different function: because it was so similar to **ш** and looked like having the value of "[š] plus something," it was stripped of its circle and simplified to Cyrillic **щ**; it acquired its new sound value [št] and was given its new place in the alphabet, right behind **ш**. It is also possible that the obvious graphic similarity of **ⱁ** to **ш** led to confusion as to which sound **ⱁ** designated.[11] It is entirely possible that it was only under the influence of the newly-created Cyrillic alphabet with its letter **щ** = [št] that the Glagolitic letter **ⱁ**, originally

[10] Veder (1999, 109) sees strong support for the theory that the "spidery h" (ѯ) originally had the function of standing for Greek **ξ**. In contrast, Marti (2004, 406) asserts that the three Greek letters **ψ ξ θ** are those that were not given Glagolitic equivalents by Konstantin-Kiril. His paper also contains a broader perspective on the specific problems in the reconstruction of the consonantal subsystem of the Glagolitic alphabet in contrast to the vowel system. In the case of **ξ** vs. **χ** as the original function of ѯ, we follow the usual view (**χ**).

[11] See the remark by Miklas (2003, 176): "Weisen doch die von Konstantin gewählten Akronyme unmißverständlich darauf hin, daß sich Konstantin über die urglagolitische Lautkorrelation des **ⱁ** (26) nicht mehr im Klaren war…".

assumed to be a Greek ѱ (psi) in disguise, acquired the same sound value as in Cyrillic, i.e., [št]. This would explain why some Glagolitic texts use ⱎ as [št] while others regularly use ⱛⱏ; in our view, then, ⱎ in its new function [št] would be a character that did not spread (immediately or at all) to all regions where the Glagolitic alphabet was (still) used. Consequently, those texts that *do* use ⱎ would implement a *younger* function than those that do not use it. All preserved Glagolitic texts are younger than the *creation* of the Cyrillic alphabet, otherwise they could serve as a test for our hypothesis.[12]

It seems possible that for ⱎ to be used in its new function it was turned around, that is, that its original orientation was different. The *name* of the letter in position 26, however, seems to have stayed "pě," so a line like "Pečal′.... could still represent this letter. The only problem this seems to have created was for writers of acrostics or abecedaria. When the preserved abecedaria were written, Cyrillic ѱ and ѱ᷄ were firmly established parts of the alphabet with clear sound values, and the same is true for the then-current Glagolitic alphabets, with ⱎ also having the sound value [št]. The writers must have been confused now as to which character (shape) to use for position 26, "pě." They could not simply repeat a letter which, for them, clearly had a different sound value by now (ⱎ [št]), and whose connection to its origin in Greek "psi" was already lost. If the shape of the "pě" from the Sinai Glagolitic alphabet is any indication, the original form of the "pě" could have been a ⱎ that was turned sideways, and that later scribes reproduced in a distorted view with only certain elements preserved: a circle either directly joining or connected to a base line, and a base line with two parallel arms, creating a Π-like structure. In that respect, even the Glagolitic "pě" from the Munich Abecedarium and the Glagolitic "pě" from the Sinai Abecedarium can be seen as being really only two distorted reproductions of the same original letter (ⱎ or ⱍ, possibly turned sideways, judging from the position of the circle).

Fig 12: The two instances of the Glagolitic "Pě"

Thus, in our view the shape ⱎ was indeed present twice in the Glagolitic alphabet and in the abecedaria based upon it: once in its original position, 26, under its original name "pě" (and with its original sound value [ps]), and then again a

[12] For another picture of the developments that took place, see Miklas (2003), and (2004, 395). It is interesting to note, however, that certain elements we arrived at independently occur also in his description (cf. 2003, 178f.): that the new Kirilica influenced the reconstruction of the Glagolica in its second, Bulgarian, phase, and that several regional variants of the Glagolitic alphabet are supposed to have existed.

second time as its later re-interpretation as [št], under the influence of the Cyrillic alphabet and its letter **ⱎ**. If we look at both sources, alphabets and abecedaria, as representing—at the same time—different historical steps in the evolution of the alphabet and the sound values of their entities, it becomes apparent that facts that seem contradictory and illogical at first can suddenly be understood without any contradiction.[13] The scribes of the abecedaria no longer recognized the old letter in position 26 as being identical to the newer one in shape. Because they could not rely on any practice to write the original **ⱋ** in Glagolitic, they would have been unsure about its canonical form and orientation, which led them to reproduce certain easily recognizable parts of the character, broken up into three distinctive elements and re-used, each in a different orientation: a circle connected to a stem, and parallel lines, and straight lines joined at 90°.

For the question of how to encode this letter in Unicode this situation creates a certain dilemma. Because Unicode is about *shapes* and not about *sound values*, the two instances of **ⱋ** (i.e. Unicode 2C1A and 2C1B) are really simply one and the same thing, with one of them *maybe* having a different orientation originally. However, because we either have only one occurrence (SinaiG) or two conflicting occurrences (see Fig. 12), we cannot be certain as to what the correct orientation should be. In other words, because the question of the orientation cannot be solved, it is better not to create artificially a new Glagolitic "Pe" (moreover one whose Unicode shape does not completely correspond to *either* of the two known instances in Fig. 12), but to leave the question open for the time being.

Maybe the pairing of the letter name "pe" with the character shape below it in the Paris Abecedarium is really not simply a mistake, just as is the case with "ot." Again, one just has to "think Greek" or "think Cyrillic" to see the solution: the shape of the Glagolitic character next to the "Pě" in the Paris Abecedarium (see Fig. 4 again) can be described using the same structural elements just outlined for the Sinai "Pě": a circle connected to a stem (though not vertical in this case), two (nearly) parallel lines, and a line extending to the right. The similarity of these structures could have lead to confusion with **ⱛ** = ts.

12. Results

In our view, the missing or mysterious character called "pě" in some sources is a character shape that is well-known even today, but has acquired a new identity. We are convinced that this letter is the original **ⱋ** that presumably was derived from Greek "psi" in shape and function (i.e., with the pronunciation [ps]) and later served as the basis for developing the Cyrillic letter **ⱎ** with its sound value

[13] Our interpretation, we believe, is able to solve the problem Marti (2004, 412, fn. 66) correctly pointed out of having one letter occurring twice.

[št] and re-interpreting ⱋ (perhaps turned around now to its current orientation) with the same sound value in the Glagolitic alphabet.

As for the Unicode character, it has been shown in the present paper that the Sinai text could lead to a revised glyph. The basic problem, however, is not solved by redesigning the normal glyph: because all known glyphs of the Glagolitic "pe" differ from one another, no standardized glyph can represent them all faithfully. Either one has to choose a picture as the conventionally "true" representation or one could prefer to call the letter "P_2." Whichever glyph is being chosen as the Unicode representation, it will always represent a single instance only (i.e., itself), and the most reliable source, SinG, does not even have a glyph whose structure is completely clear. Under such circumstances, no harm would have been done, or so it seems to us at least, if the Glagolitic "Pe" would not have been added to Unicode immediately.

References

Durnovo, N.
1930 Das Münchner Abecedarium. *Byzantinoslavica* 2, 32–41.

Hamm, J.
1974 *Staroslavenska Gramatika. Četvrto izdanje*. Zagreb.

Kempgen, S.
2006a Unicode 4.1 and Slavic Philology – Problems and Perspectives (I). In: A. Miltenova, D. Radoslavova, E. Pancheva (eds.), *Computer Applications in Slavic Studies. Proceedings of Azbuky.net. International Conference and Workshop. 24–27 October 2005*, Sofia, Bulgaria. Sofia 2006, 131–159.

2006b Unicode 4.1 and Slavic Philology – Problems and Perspectives (II). In: T. Berger, J. Raecke, T. Reuther (Hgg.), *Slavistische Linguistik 2004/2005*, München 2006, 223–248.

2007 Das «Münchener Abecedarium» – ein neues Facsimile samt einigen neuen Beobachtungen. In: *Scripta & e-Scripta* vol. 5, Sofia 2007 (2008) 73–93 plus 2 figs. after p. 160.

2009 Unicode 5.1, Old Church Slavonic, Remaining Problems – and Solutions, including OpenType Features. In: *Slovo: Towards a Digital Library of South Slavic Manuscripts. Proceedings of the International Conference, 21–26 February 2008, Sofia, Bulgaria*. Sofia 2008, 200–219.

Koch, Chr.
2004 Das Bamberger glagolitische Alphabet. Zur Entstehungsgeschichte der Zeichenfolgen der kroatischen glagolitischen Alphabete. In: *Glagoljica i hrvatski glagolizam. Zbornik radova…*, Zagreb, 435–452.

Mareš, Fr. V.
1971 Hlaholice na Moravě a v Čechách. *Slovo* 21, 133–200.
Marti, R.
1999 Abecedaria – a Key to the Original Slavic Alphabet: The Contribution of the Abecedarium Sinaiticum Glagoliticum. In: *Thessaloniki – Magna Moravia*, Thessaloniki, 175–200.
2004 Die Bezeichnung der Konsonanten in der Glagolica. In: M.-A. Dürrigl, M. Mihaljević, Fr. Velčić (eds.), *Glagoljica i hrvatski glagolizam*, Zagreb–Krk, 401–417.
Miklas, H.
2003 Jesus-Abbreviatur und Verwandtes: zu einigen Rätseln der glagolitischen Schriftentwicklung am Material der Azbučnaja Molitva. In: Honselaar, W. et al. (eds.): *Times Flies. A Festschrift for William R. Veder on the occasion of his departure as Professor of Slavic linguistics at the University of Amsterdam*, Amsterdam 2003, 171–204.
2004 Zur Relevanz des neuen sinaitischen Materials für die Entwicklungsgeschichte der Glagolica. In: *Glagoljica i hrvatski glagolizam. Zbornik radova*…, Zagreb, 389–399.
Mošin, V.
1973 Još o Hrabru, slavenskim azbukama i azbučnim molitvama. *Slovo* 23, Zagreb, 5–71.
Nahtigal, R.
1923 Doneski k vprašanju o postanke glagolice. In: *Razprave* I, Ljubljana, 135–178.
Tarnanidis, I. X.
1999 To slaviko (gklagkoliko alfavitari tou Sina). In: *Thessaloniki – Magna Moravia*, Thessaloniki, 165–173.
Trubetzkoy, N.:
1930 Das «Münchner Slavische Abecedarium». *Byzantinoslavica* 2, 29–31.
Veder, W. R.
1999 *Utrum in alterum abiturum erat? A Study of the Beginnings of Text Transmission in Church Slavic*. Bloomington, Indiana.
2004 The Glagolitic Alphabet as a Text. In: *Glagoljica i hrvatski glagolizam. Zbornik radova*…, Zagreb, 375–387.
Velčeva, B.
1973 Văprosat za ѫ v glagoličeskata azbuka. *Izvestija na Instituta za bălgarskija ezik* XXII, 105–124.