

Title: When is a descender not a descender: Kildin Sámi voiceless consonants
Source: Michael Everson
Status: Expert Contribution
Date: 1999-04-16

I have been the past few days at Universitetet i Tromsø, where the Senter for samiske studier has invited me to participate in a project to provide the Saamskij sektor of the Murmansk branch of the Russian Academy of Sciences with a number of computers enabled with Kildin Sámi fonts, keyboards, sorting routines, and other system-level features. Rimma Kuruč, director of the Sámi program in Murmansk, Trond Trosterud of the linguistics department here, and I have spent a good deal of time working on font requirements, especially following a very interesting seminar Rimma gave today on the history of the alphabetization of Kildin Sámi, from 1895 to today. Real linguistic analysis of Kildin Sámi was only done during this century, especially by Rimma and her colleagues in the past twenty years.

Rimma brought with her a number of books printed in a number of typefaces which were not available to Trond and me a year ago when we proposed the addition of the Cyrillic Sámi characters to the UCS (N1590, N1744). The sources we had at the time were typewritten and handwritten sources from Murmansk (a fax of which was sent to the UTC, as I recall), and nicely-typeset children's books from Davvi Girji press of Norway using fonts made by Juhani Lehtiranta of Finland. The problem is that Juhani's nice fonts led us to misinterpret the typewriter face and the handwritten evidence with regard to character identity – the fact of the matter is that, nice as they are, they do not reflect the shapes of the characters as used by the Kildin Sámi themselves.

As a character/glyph question, this has a number of implications, the chief of which involves three of the six letters entering the FPDAM ballot. CYRILLIC LETTER EL WITH DESCENDER, CYRILLIC LETTER EM WITH DESCENDER, and CYRILLIC LETTER SHORT I WITH DESCENDER, are all misnamed, because the distinguishing mark is not a DESCENDER. It is a TAIL. (It may be useful to note that these three characters, in addition to a similarly modified CYRILLIC LETTER EN and the CYRILLIC LETTER ER WITH TICK, are used as voiceless vowels.)

The descender is a common and meaningless part of several basic Cyrillic characters, namely, DE, TSE, and SHCHA. It has a particular simple form, generally being a simple vertical dropping stroke. In Latinskij, Bastion, and Everson Mono:

Дд Цц Щщ Дд Цц Щщ Дд Цц Щщ

In some fancy Russian fonts, the descenders of TSE and SHCHA may swing back under the character in a long reversed-tilde swash.

Ц ц Щ щ

This swash-descender practice, however, is unique to these two letters; it does not occur with DE. In some other languages, the meaningless descender was attached to other letters to differentiate them from other base letters in the alphabet. Examples are such letters as ZHE WITH DESCENDER and EN WITH DESCENDER, used in Tatar, and HA WITH DESCENDER, used in Tajik. The swash form of the descender does not occur with these letters.

Жж Нн Хх

Assuming that the descender was, as it is in Tatar, used in Kildin Sámi, I produced the following set of glyphs, subsequently used in N1590 and N1744. (ER WITH TICK is also shown.)

Йй Лл Мм Нн Рр Йй Лл Мм Нн Рр

Йй Лл Мм Нн Рр

However, Rimma Kuruč maintains that she does not recognize the Bastion and Everson Mono glyphs to be Kildin Sámi letters (indeed she was surprised to learn that they were supposed to be). She does recognize the Latinskij glyphs as being useable, though she finds them extremely ugly. What does it mean when she says that Rimma can recognize these Latinskij glyphs with descender as ugly but useable examples of Kildin Sámi letters? It is exactly analogous to the situation on NKR-Teletekst in Norway. Because of the limited character set available, Swedish letters Ö and Ä are used in place of Norwegian letters Ø and Æ. Norwegian written with Swedish letters is incorrect. It is readable, but ugly.

Analysis of the books Rimma brought with her to Tromsø indicates that the proper form of these letters has not a TSE-based descender, but a very different TAIL.

Йй Лл Мм Нн Рр Йй Лл Мм Нн Рр

Йй Лл Мм Нн Рр

Note especially the important feature, namely, that unserifed fonts like Bastion and Everson Mono have a two-stroke diacritic TAIL attached to the letter, not a single DESCENDER falling off of a serifless terminal. That horizontal bar is not a serif. The reason Rimma recognizes the Latinskij DESCENDER letters as ugly variants of the Latinskij TAIL letters is because she sees the serif as being the top bar of the diacritic (though the bottom part is incorrectly drawn and is not the diacritic used in Kildin Sámi). Indeed, in serifed fonts it is customary to leave the letter's bottom-left serif as a serif, though the bottom-right is obliterated and replaced by the thicker horizontal bar.

What this indicates is that EN WITH DESCENDER and EN WITH TAIL are not the same characters, if they can be so easily distinguished in non-serif fonts. Indeed, they can be easily distinguished in serifed fonts, though there is a certain similarity; but this similarity is no greater than the similarity one has with LATIN SMALL LETTER A and GREEK SMALL LETTER ALPHA when mixed in plain and italic styles in serifed and nonserifed fonts.

In other publications the shape of the TAIL is different. Kert's 1986 Sámi dictionary and Barmič and Vèllo's 1994 Forest Nenets dictionary use the following style:

Йй Лл Мм Нн Рр

Note that there is simply no question that this TAIL can be interchanged with the DESCENDER. Consider the following ludicrous examples:

Дд Цц Щщ

Kildin Sámi and Forest Nenets use letters with TAIL, not with DESCENDER. It might not matter if this were the end of the story, but it is not. The problem is that if we do not, as we should, disunify TAIL from DESCENDER, but instead only give font guidance for the Kildin-Sámi-specific letters, as Ken Whistler suggested, we will *invariably* end up with fonts containing the following repertoire:

Йй Лл Мм Нн Рр

This is obviously *not* what we should do to support Kildin Sámi in the UCS. We should now, by UTC agreement and National Body comment on FPDAM 30, rename 04C5, 04C6, 04C9, 04CA, 04CD, and 04CE to CYRILLIC LETTERS WITH TAIL, correct their glyphs, and prepare for the eventual addition of CYRILLIC LETTER EN WITH TAIL, missing from FPDAM 30 (whether now or in a future amendment). The evidence here is strong enough to show that the TAIL and the DESCENDER are derived differently: one is a systematically-applied diacritic performing a specific function and the other is a simple graphic appendage derived from a meaningless mark on some other letters; their similarity ends with the place they are attached to the baseform.

If making this name change (which can be proposed by a number of interested National Bodies, such as Norway, Finland, and Ireland) is impossible, the only reasonable way forward that I can see is to remove the Kildin Sámi letters (though SEMISOFT SIGN, ER WITH TICK, and E WITH DIAERESIS are non-controversial) from the FPDAM pending further study. And that is just as troublesome a delta for Ken Whistler to handle as is fixing these three names.

(Postscript. It is not the intention here to suggest disunification of every Cyrillic letter that has some kind of whatsit hanging off of it. For example, in Yuit (Siberian Eskimo) the following letters are used:

Гг Кк (Кк) Нн Хх

Whatever may be said about the unification of this LOOP with DESCENDER (and I have no idea if it has been addressed), it is clear that the 6-COMMA is no more than a glyph variant of the LOOP, since there are hundreds and hundreds of examples in the book where they are used interchangeably (though only with the KA). That said, it is doubtful that a font containing the LOOP could be used with any measure of acceptance by Kazakhs, Tatars, or Tajiks. Which suggests that a descender may not always be a descender.)