EasyChair Preprint

№ 9177

Relics of the Finnovolgian Accent in the Seto Language

Yulia Normanskaya and Denis Dogadushkin

October 28, 2022

Normanskaya Y. V. Dogadushkin D. M.

Relics of the Finnovolgian accent in the Seto language

In 2008, the book "Reconstruction of the Proto-Finnovolgian accent" [Normanskaya 2008] was published. In it, we have addressed five problems of Finnovolgian comparative and historical linguistics, which at the first glance seem unrelated to each other: 1) the genesis of vocalism systems in modern Mordovian languages; 2) the origin of the two types of conjugation in the Mari language; 3) the origin of the third type of declensions of nouns (based on stressed vowels) in the Mari language; 4) the origin of reduced vowels in the Mari language; 5) the genesis of the vocalism system in the Saami language. These issues have been the subject of active discussions over the past hundred years. We were able to show that these phonological and grammatical phenomena, which at first glance were unrelated, developed depending on one very important and previously overlooked proto-linguistic factor: mobile stress, not fixed grammatically or phonetically, which should be reconstructed for the Proto-Finno-Volga language. The place of this stress (at the root or in the end) has been preserved unchanged in the a-bases in the modern Moksha language. In the following years, research was carried out looking for similar phenomena: a miscellaneous accent and its influence on the development of the vocalism system — in other Uralic languages. The material had to be collected bit by bit. At first, it was unclear whether such systems still exist, since most authoritative science works indicate that the stress in Uralic languages is fixed on the first syllable. Indeed, in most dictionaries published in the XX century, there is no indication of the stress placement. And in those works where it is marked, as in the Komi-Permian [Batalova, Krivoshchekova-Gantman 1985] and Komi-Yazvin [Lytkin 1961] languages, the stress is automatic and depends on the quality of the vowels. Another accent is noted in the dictionary of forest Nenets [Popova 1978] of the language, but when correlating this data with a more authoritative dictionary [Lehtisalo 1956], researchers previously assumed that in the Nenets forest longitude was marked, which is indicated in T. Lehtisalo's dictionary. Further, it was revealed that in many pre-revolutionary dictionaries, books, published and unpublished archival materials collected at different times in Komi, Khanty, Mansi, Nenets, Selkup languages, there was a diverse paradigmatic stress. A complete analysis of this data was published in the monograph [Normanskaya 2018].

A dictionary of the Seto language [Käsi 2016] has recently been published, and it also notes a different accent. This is the Baltic-Finnish language of the small ethnic group Seto, who lived in the south-eastern part of Estonia (the counties of Vyrumaa and Pylvamaa), the Pechora district of the Pskov region of Russia (from 1920 to 1940 — in Petseri County of the Republic of Estonia) and adjacent areas. As part of the work supported by the grant of the Russian National Science Foundation No. 20-18-00403 "Digital description of dialects of Uralic languages based on big data analysis" in the summer of 2021, an expedition was conducted to the areas of residence of Seto: in the village of Izborsk, in the Pechora district of the Pskov region of Russia, Obinnica Setumaa Parish, Vyrumaa County, Estonia. Audio dictionaries were collected from two Pechora Seto from D. Izborsk, and from two Estonian Seto from d. Obinitsa. They are available on the LingvoDoc platform lingvodoc.ispras.ru.

In this article we will focus on the description of the stress in the Seto language in names that have Finno-Ugric origin. The analysis showed that the stress system is not fundamentally different between Russian and Estonian setos. Examples used below are

taken from the dictionary:
http://lingvodoc.ispras.ru/dictionary/4056/45883/perspective/4056/45884/view, which was recorded in v. Obinitsa.
We have analysed the placement of stress in the names of Finno-Ugric origin from the point of view of experimental phonetics in the Praat program. It was found that the ratio of the length and intensity of vowels depend on the quality of the vowels of the first and second syllables. Words with narrow vowels in the first syllable and wide in the second syllable should be considered separately, since they always have L1<L2 and the relevant feature, which is perceived by native speakers of the Russian language as an accent, is only the intensity of the vowel. Interestingly, for the rest of the words, the relevant feature is only the duration of the vowel of the first or second syllable, since intensity always marks the first syllable. From our point of view, this system cannot be described as a full-fledged accent in comparison to, for example, the Russian language, since in most cases it is expressed only by a longitude contrast. But, considering the fact that in words with narrow vowels in the first syllable and wide in the second, the contrast is expressed by intensity, it can be assumed that earlier it was an accent, which later, probably under the influence of contacts with native speakers of the Estonian language, began to be rearranged and replaced by vowel lengthening with a fixed intensity on the first syllable, as is typical for most Baltic-Finnish languages. On the other hand, it can be noted that these longitude oppositions do not correspond to any etymological Baltic-Finnish longitude, which also confirms the hypothesis of their secondary formation.

The question arises, what historical reality is behind the disappearing "stress" in seto, which is reborn into longitude oppositions? Is this related to the Finnovolgian accent, which in turn is coming from Proto-Ural accent, which is described in [Normanskaya 2018].

Briefly here are the rules of stress reconstruction looked at in [Normanskaya 2018]:

1) if in the Mordovian languages the vowel *a, o, e, ä* is represented in the first syllable of a word, then the stress in the Finnovolgian language is reconstructed on the first syllable, if in the first syllable the vowel *u, ə, 0, i,* then the stress is reconstructed on the second syllable.

2) Wide Proto-Perm. the vowel of the first syllable indicates PP I and II a. p., and the narrow one indicates PP III and IV a. p. Also, for determining the Proto-permian place of stress, the diagnostic forms are primarily the Lower Neva Komi-Permian verb forms. Semusheva G.P. from the village of Timino coincided III a. p. and I a. p., but they differ from IV a. p. Therefore, we provide illustrative material on her dialect to indicate the difference between III and IV a. p.

3) For the PS language, we give the proto form from the dictionary [Janhunen 1977], if it is presented there. The place of stress in the PS is also determined by the quality of the vowels: a wide PS vowel occurs only in the stressed position, **ə*— only in the unstressed position. Narrow vowels can be both stressed and unstressed, the placement of stress can be determined by stress / longitude in the Nenets word form or by stress in the southern and central Selkup word form, in which suffixes with plus accent markings are presented.

4) For the Proto-Khanty language, the Irtysh forms according to M. A. Kastren's archive and Nizyam non-derived names from our field records are diagnostic for determining the place of stress.

5) The verb forms of the Kondinsky, Ob and Sosvinsky dialects from our field records are diagnostic for Pramansi.

The Ural accent paradigm is reconstructed based on the comparison of the place of stress in the daughter proto-languages:

PU I a. p.: FV I a. p. — PP I a. p. — PS I a. p. —PKhant. I a. p. — PMans. I or III a. p.

PU II a. p.: FV I a. p. — PP II a. p. — PS I a. p. — PKhant. I a. p. — PMans. II or IV a. p.

PU III a. p.: FV I a. p. — PP III a. p. — PS II a. p. — PKhant. II a. p. — PMans. II or IV a. p.

PU IV a. p.: FV II a. p. — PP IV a. p. — PS I a. p. — PKhant. I a. p. — PMans. I or III a. p.

Below are examples of Seto words that have a known proto-linguistic accent paradigm, see [Normanskaya 2008, Normanskaya 2018] for more details.

I The ratio of longitude and intensity for most words, except in cases with the first narrow, second wide vowel:

I.1. L1 < L2, I1 ≥ I2 correlates with FV I accent paradigm < PU I, III accent paradigm:

1) *lǽmmí* ‘warm’ (L1 0,09, I1 81.1; L2 0,13, I2 79.0) < PU **lämpV* ‘warm’ [UEW: 685] PU I-III a.p. > *FV I a.p. (*mord. erz. *Ílembe*, *Ílämbä*, moksh. *Ílämben* [Paasonen 1992: 1108], mar. *leváš* (2 conjugation);

2) *sátá* ‘hundred’ (L1 0,08, I1 84.7; L2 0,16, I2 75.3) < PFU**śata* ‘hundred’ [UEW: 467] PU I-III a.p. > mord. erz. *śado*, moksh. *śadă* (M) [Paasonen 1996: 2083] FV I a.p.;

3) *vesí* ‘water’ (L1 0,09, I1 81.6; D2 0,19, I2 74.3) < PU **wete* ‘water’ [UEW: 570] PU III a.p. > PS II a.p. **wit* (selkup. Par. *ÿdə́* ‘water, vodka’, *ÿdъgý* ‘to get wet’, *ÿmálǯçygy* ‘to give water, to get drunk’, *ÿdə́unygy* ‘drink’);

4) *sulá* ‘thawed’ (L1 0,05, I1 77.2; L2 0,13, I2 74.1) < PFU **sula* ‘melt’ [UEW: 450] PU III a.p. > FV I a. p. (mord. *sola-* ‘melt’) — PP III a. p. (komi-perm. tim. *śĭla* ‘I am melting’) — PKhant. II a. p. (khnat. niz. *tudáti*);

5) *kúsi* ‘urine’ (L1 0,08, I1 78,07; L2 0,11, I2 75.3) < PU **kuńće* ~ **kuće* ‘urine’ [UEW: 210] PU III a. p. > PP III a. p. (Komi *gi̯ž* ‘fingernail’) — PS II a. p. (**kə̂tå* ‘fingernail’ [Janhunen 1977: 155]) — PMans. II a. p. (obsk. *kosχáti* ‘scratching’, *kosχáteit* ‘are being scratched’);

8) *jokí* ‘river’ (L1 0,08, I1 82.09; L2 0,27, I2 78.3) < PU **joke ‘river’ [UEW: 99]* PU III a. p. > PP III a. p. (komi *ju*) — PS II a. p. (**jə̑kå*) — PKhant. II a. p. (khant. irt. *Jeagá* ‘river’ — khant. niz. *juχán* ‘river’);

6) *muná* ‘egg’ (L1 0,05, I1 76,1; L2 0,12, I2 72.9) < **muna ‘egg’ [UEW: 285]* PU III a.p. > FV I a. p. (mord. *mona* ‘egg’ [Paasonen 1992: 1282]) — PS II a. p. (**mə̑nå* ‘egg’ [Janhunen 1977: 86]);

7) *teräe* ‘blade’ (L1 0,09, I1 82,0; L2 0,17, I2 80,2) < FU **terä ‘blade’ [UEW: 522]* PU I a.p. > udm. *tir* ‘axe’ PP I a.p.;

9) *pʉ́mmæ* ‘dark’ (L1 0,10, I1 81.6; L2 0,15, I2 81.0) < PU **piĺmV* ‘dark’ [UEW: 381] PU I or II a. p. > PP I a. p. (komi *pemi̯d* ‘dark’) — PS I a. p. (nenets. *päewde̊*);
10) *varə́s* ‘raven’ (L1 0,07, I1 76.9; L2 0,10, I2 76.5) < PU **warV* ‘raven’ [UEW: 559] PU III a.p. > FV I a. p. (mord. erz. *varaka*, *varśej*, *vaŕćeŋ*, *vaŕćej*, *varkśij*, *varksij*, мокш. *varśi*, *varći*) — PS II a. p. (**wə̑r-* ‘crow’ [Janhunen 1977: 170]) — PKhant. II a. p. (khant. niz. *wurŋá*);

11) *palá* ‘piece’ (L1 0,07, I1 68; L2 0,15, I2 68) < PU **pala* ‘piece’ [UEW: 350] PU I or II a. p. > FV I a. p. (mord. *pal* ‘piece (about food), meat’ [Paasonen 1994: 1510]) — PP I a. p. (komi *palak* ‘formation; lump (earth, manure, snow)’) — PS I a. p. (**pålä-* ‘swallow’ [Janhunen 1977: 116]);

12) *kalá* ‘fish’ (L1 0,07, I1 76; L2 0,12, I2 76) < PU **kala* ‘fish’ [UEW: 119] PU I или II a.p. > FV I a. p. (mord. *kal* [Paasonen 1992: 574]) — PS I a. p. (**kålä* ‘fish’ [Janhunen 1977: 59]);

I.2. L1 ≥ L2, I1 ≥ I2 correlates with FV II accent paradigm < IV accent paradigm:
1) *órav* ‘squirrel’ (L1 0,11, I1 75,7; L2 0,11, I2 70,2) < FP **ora* ‘squirrel’ [UEW: 343] PU IV a. p. > FV II a. p. (mord. erz. *ur*, moksh. *ura* ‘squirrel; a penny’ [Paasonen 1996: 2461]) — PP IV a. p. (*ur* ‘squirrel’);
2) *kórga* ‘tall’ (L1 0,12, I1 75.4; L2 0,12, I2 71.7) < FP **korkV* ‘tall’ [UEW: 672] PU IV a.p. > FV II a. p. (mord. moksh. *kuRka* ‘deep’);
3) *póiskone* ‘boy’ (L1 0,07, I1 86.2; L2 0,06, I2 83.6) < ФП **pojka ‘boy’ [UEW: 339]* PU IV a.p. > FV II a. p. (mord. erz. *bujo, pijo* ‘grandson, granddaughter’ [Paasonen 1990: 147]) — PP IV a. p. (komi *pi*);

II. For words with vowels of the first syllable *i, ɨ* there are also two types of the ratio of longitude and intensity:

II.1. L1 < L2, I1 ≤ I2 correlates with FV I accent paradigm < PU I, III accent paradigm:

1) *inǽp* 'more' (L1 0,08, I1 74.8; L2 0,11, I2 78.9) < PU **enä* 'big, lots' [UEW: 74] > FV I a. p. (mord. erz. *iṅe*, moksh. *iṅä* 'big' [Paasonen 1990: 463]) — PP III a. p. (komi *una* 'lots') — PS I a. p. (**inä* 'elder brother' [Janhunen 1977: 27]);

2) *imǽ* 'mother' (L1 0,04, I1 71.8; L2 0,12, I2 76.2) < PU **imV* 'old woman, grandmother' [UEW: 83] > PS II a. p. (**imɜ*- 'old woman, grandmother' [Janhunen 1977: 27]) — PKhant. II a. p. (*Íma* Castrén, khant. niz. *imî*);

3) *sisálik* 'lizrd' (L1 0,05, I1 64.6; L2 0,06, I2 66.4) < PU **sVŋćV* (**sVŋćV-lV*) ~ **sVčV* (**sVčV-lV*) 'lizard' [UEW: 454]: PP I a. p. (komi *ćoʒ́ul*) — PS I a. p. (**t1ånsə̂* 'lizard' [Janhunen 1977: 151]).

II.1. L1 < L2, I1 ≥ I2 correlates with FV II accent paradigm < PU IV accent paradigm

1) *mɨ́lá* 'paddle' (L1 0,09, I1 78,9; L2 0,18, I2 75,8) < FP **melä* 'paddle' [UEW: 701] > FV II a.p. (mord. erz. *miĺe* 'paddle', moksh. *məĺəṅá* 'pouring spoon'[Paasonen 1992: 1262]);

2) *ígǽ* 'age' (L1 0,11, I1 74,2; L2 0,23, I2 73,0) < PFU **jikä* (**ikä*) 'age, year' [UEW: 98] > FV II a.p. (mord. erz. *ije*, moksh. *ij* [Paasonen 1990: 441]).

So, we can see that there is a certain correlation between PU and FV a.p. and the ratio of the length and intensity of the vowels of the first and second syllables in the names in seto:

- if I a.p. is represented in the FV language (< PU I, III a.p.), then in SRT the longitude of the vowel of the second syllable will be greater than the longitude of the vowel of the first syllable;
- if II a.p. is represented in the FV language (<PU IV a.p.), then in seto the length of the vowel of the first syllable will be greater than or equal to the length of the vowel of the second syllable
- with the exception of words with the first narrow, second wide vowel, in which the second vowel will always be longer, but with FV I a.p. the maximum intensity will mark the vowel of the second syllable, and with FV II a.p. the vowel of the first syllable.

Of course, in the future it is planned to collect a complete audio dictionary of the primordial vocabulary of the Seto language and test the results obtained on it. It is also interesting to

examine verbs in the paradigm of which a heterogeneous paradigmatic stress has been identified, but for this it is necessary in the future to collect complete paradigms of verbs of Finno-Ugric origin.

At present, it can be noted that the Seto language probably has preserved relics of the Proto-Finnovolgian accent. So far, this is the only Baltic-Finnish language in which it has been possible to identify at least a partial preservation of this phenomenon.

List of abbreviations

Komi-Perm tim. — Temkinsky dialect of the Komi-Permian language

Mar. — Mari language

Mord. Moksh. — Moksha language

Mord. Erz. — Erzya language

Nenets — Nenets language

Obsk. — the Ob dialect of the Mansi language

PMans. — Proto mansi language

PP — Proto Permian language

PS — Proto Samodian language

PU — Proto Uralic language

PKhant. — Proto Khanty language

Selk. Par. — field materials on the Selkup language collected in 2009 in the village of Parabel, Tomsk region

Udm. — Udmurt language

FV — Finnovolgian язык

Khant. Irt. — the Irtysh dialect of the Khanty language based on the materials of M.A.Kastren

Khant. Niz. — Nizami dialect of the Khanty language

L1 - Length of 1st vowel

L2 - Length of 1st vowel

I1- Intensity of the first vowel

I2 - Intensity of the second vowel

List of references

Batalova, Krivoshchekova-Gantman 1985 - Batalova R. M., Krivoshchekova-Gantman A. S.

Komi—Permian-Russian dictionary. M.: Russian language, 1985.

Lytkin 1961 - Lytkin V. I. Komi-Yazva dialect. M., 1961.

Normanskaya 2008 — Normanskaya Yu. V. Reconstruction of the Proto-Volga accent. M., 2008.

Normanskaya 2018 — Normanskaya Yu. V. Reconstruction of the Proto-Ural paradigmatic stress and its influence on the development of the vocalism system. Moscow, 2018.

Popova 1978 — Popova Ya. N. Nenets-Russian dictionary. Forest dialect. Szeged, 1978.