

Kimmo Granqvist

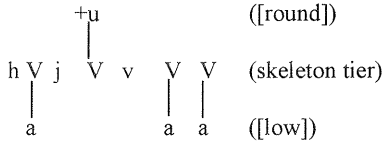
Vowel Harmony in Finnish and Finnish Romani

1. Introduction

The purpose of this paper is to discuss some aspects of the Finnish Romani vowel harmony by presenting a contrastive analysis of Finnish and Finnish Romani types of vowel harmony. The questions discussed in this paper include (i) the description of the Finnish Romani vowel system and (ii) the description of the Finnish Romani vowel harmony, in particular the distinction between ‘internal harmony’ and ‘suffix harmony’, and the treatment of compounds and disharmonic stems.

The theoretical framework chosen for the present paper is the autosegmental phonology approach (Goldsmith 1976). It has proven very suitable in description of features whose scope is beyond the segment, such as tonal phenomena and vowel harmony (for vowel harmony, see Goldsmith (1985), for other approaches, see van der Hulst & van de Weijer (1996)). Autosegmental phonology is a representation of generative phonology that allows (in contrast to strictly segmental theories) features to belong to one or more segments. The description comprises several tiers; each tier consists of segments that are linearly arranged. In the autosegmental approach to vowel harmony, vowel features are placed on separate autosegmental tiers (Goldsmith 1985: 254). As in (1), the segments are linked together with association lines, which indicate how they are coarticulated.

(1) *hajuva* ‘to know’



The vowel features used are either binary or unary (bivalent or monovalent/equipollent or privative). As in Chomsky & Halle (1968), the feature [round] is binary as it is always specified as either [+] or [-]. Following Goldsmith (1985), the feature [low] is regarded as unary, as it is specified only for its presence. As illustrated in (1), the symbols [u] and [a] are used to abbreviate the feature names [round] and [low]. V denotes a V-slot.

Additionally, some empirical notes are presented in the paper. The material used for this study consists of eight computerized corpora. The corpora are available at the Research Institute for the Languages of Finland. Five of the corpora used are SGML-coded dictionaries. The Jalkio and Kronqvist corpora are based on manuscripts of word-lists, owned by Mustalaislähetys in Helsinki. The manuscript of a Finnish–Romani word-list, by Yrjö Temo, was given to the Research Institute for Languages of Finland in 1984. The MNS and Thesleff corpora are based on printed books (Mustalaiskielen ortografiakomitea 1972; Thesleff 1901). Three of the corpora are passages from the Bible, translated into Finnish Romani by Yrjö Temo (Bible1) (available at the Research Institute for the Languages of Finland), Viljo Koivisto (1971) (Bible2) and Pertti Valtonen (1970) (Bible3). The dictionaries provide 25,289 lexical entries; naturally, the same lexeme may occur several times in the material. The overall size of the text corpora is 26,043 words. By means of normal UNIX commands and Awk programs, a subcorpus of 1,050 words (~ 2 %) was extracted from the original corpora. The subcorpus contains only those items in which there are front harmony vowels /y, ä, ö/. The composition of the

used is presented in table (1). The size of the material provided varies a lot from corpus to corpus, the smallest amount of data being provided by the three text corpora. Due to the small size of the data, I did not compile separate statistics for the dictionaries and the texts.

Corpus	Overall size	Number of words containing /y, ö, ä/	%
Thesleff	7,577	169	2.2
Kronqvist	2,679	67	2.5
Jalkio	7,563	263	3.5
Temo	4,478	284	6.4
MNS	2,992	107	3.6
Bible1	4,086	56	1.1
Bible2	17,295	68	0.4
Bible3	4,662	36	0.8
Total	51,332	1,050	2.0

Table 1. The composition of the corpus used.

2. Finnish Romani Vowel System

Presumably, the Romani language had originally a five-vowel-system, which comprised the vowel phonemes /a, e, i, o, u/ (according to some scholars, the occurrence of /ə/ was also possible) (Valtonen 1968: 93). This five-vowel-system is a subset of the vowel system of Sanskrit, which Romani (as an Indo-Aryan language) often has been compared with in the research tradition, especially in historical linguistics (e.g. Bloch 1921; Kochanowski 1963; Miklosich 1872/1881; Pott 1844/1845). Sanskrit has the vowels /a, a:, i, i:, u, u:/, the syllabic consonants /l, r, r:/, and the

diphthongs /ai, e:, o:, au/ (Mishra 1972; Valtonen 1968: 93). The five vowels /a, e, i, o, u/ are common to all dialects of Romani (Cortiade 1989: 14). In dialects of Albanian Romani, the vowels /ɨ, ə, ü/ and a syllabic /ɾ/ are found in addition to the five basic vowels (Cortiade 1989: 14). Both Finnish and Latvian dialects of Romani have the additional front vowels /y, ö, ä/ (Manuš & Neilands & Rudevičs 1997; Valtonen 1968: 93).

According to Valtonen (1968: 93), /y, ö/ originate from Hungarian. /ä/ was borrowed later (from Scandinavian languages?). It is not exactly known, when the vowels /y, ö, ä/ were adopted. In Finnish Romani, the vowels /y, ö, ä/ are found mainly in a group of Scandinavian and Germanic loans, such as *byykä* 'laundry' < Germ. *büke*, Sw. *byk*, *bäri* 'hill' < Sw. *berg*, *hyög* 'high' < Sw. *hög*, *lyördä* 'Saturday' < Scand. *lördag*, Sw. *lördag*, *stykös* 'piece' < Germ. *stykke*, Sw. *stycke* etc. In older layers of the Romani vocabulary they are extremely rare (Valtonen 1968: 93). Some examples, given by Valtonen (1968: 93), are: *čäj* 'girl', *däj* 'mother', *gäaji* 'housewife', *phyyli* 'widow', *räj* 'lord', *thyöoli* 'tobacco'. Note that also the forms *čaj, daj, gaaji, raj* are used.

Due to their distribution in loan items only, the vowels /y, ö, ä/ are rare in Finnish Romani. Their respective phoneme frequencies (computed on the basis of the Romani corpora at the Research Institute for the Languages of Finland) are 0.8 % (y), 0.5 % (ä) and 0.8 % (ö) of all vowel phonemes. In Finnish, the corresponding phoneme frequencies are higher except for /ö/: 1.8 %, 4.7 % and 0.5 % of all vowel phonemes (Karlsson 1982: 75).

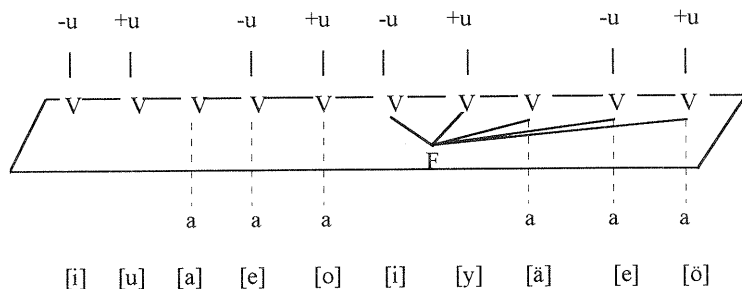
From the autosegmental point of view, Finnish Romani has (exactly like Finnish) a vowel system that has (i) an unary feature [low] on one tier and (ii) a binary feature [round] on another tier. In autosegmental phonology, segments that are considered autonomous and represented on their own tier are

autosegmentalized unary feature [front] (hereafter referred to as F). The phonetically fronted vowels are associated with the feature F. Furthermore, both Finnish and Finnish Romani have a Front Specification rule (Goldsmith 1985: 261):

- (2) Associate the feature F with any [-round] V-slot

The rule acts on /i, e/; /a/ is not specified for the feature [round], see (3). The phonetically fronted vowels, thus, have two sources. Following Goldsmith (1985), we end up to the following description of the vowel system (note that the neutral vowels [i] and [e] are listed twice, as both back and front vowels).

- (3)



3. Vowel Harmony

Valtonen (1968: 94) argues that around the same time when these vowels entered into Finnish Romani, phenomena that resemble the Finnish vowel harmony began to occur in the language, probably influenced by the Finnish vowel harmony. He (1968: 94) points out that allomorphs of suffixes, containing /y, ö, ä/ instead of the back

vowels /u, a, o/, were first found in the notes of Reinholm (1819-1883).

Finnish Romani has same kind of front/back harmony as the Finnish language. As in Finnish, the vowel harmony acts linearly from left to right (for Finnish, see Karlsson 1982: 100). If the first vowel is associated with the feature F, the vowel harmony rule (the principle is shown in (4)) associates all subsequent V-positions with the feature F. As Goldsmith (1985: 258) points out, it does not matter whether some of these V-positions already are associated with the feature F.

(4)



The point up to which the vowel harmony rule can spread associations, is defined morphologically (Goldsmith 1985: 258). In Finnish Romani, the feature F can never spread across a word boundary. Morpheme boundaries may stop the spreading of the feature F, too (see 3.2 and 3.3).

In the same way as in Finnish, we must distinguish between ‘internal harmony’ that takes place in the stems and ‘suffix harmony’. While internal harmony is present at the lexical entry level, suffix harmony is not. Instead, suffix harmony is subordinate to internal harmony.

3.1. Internal Harmony

Like the Finnish vowels, the Romani vowels are divided into three sets:

- (i) front harmony vowels: {y, ö, ä}
- (ii) back harmony vowels: {u, o, a}
- (iii) neutral vowels: {i, e}.

The harmony vowels constitute three pairs:

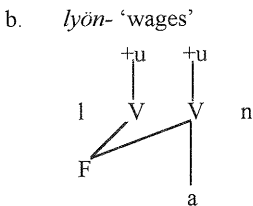
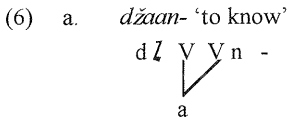
- (5) y - u
- ö - o
- ä - a

In these pairs both members share the rounding feature (if present), but contrast as for the feature F. Like in Finnish (Karlsson 1982: 99; Kiparsky 1982: 115), the neutral front vowels {i, e} remain with their back equivalents, since the language lack the non-rounded back vowels *[ɨ] and *[ɯ].

In the same stem, it is only either front or back harmony vowels that may co-occur. The neutral vowels may co-occur with both front and back harmony vowels.

The feature F is present in harmonic stems at the stage of word-level phonology. There are two kinds of harmonic stems:

- (i) those with no feature F in their lexical entry, as in (6a)
- (ii) those with the feature F in their lexical entry; the feature F spreads over all V-positions in the stem. Consider example (6b). (Goldsmith 1985: 269.)



Internal harmony takes place fairly regularly in Finnish Romani. As much as 94 % of all stems studied consist of front harmony vowels or /and neutral vowels.

3.1.1. Violations of the Vowel Harmony Rule

Some stems, especially loan words, violate the vowel harmony. The following disharmonic stems are found in the subcorpus used for this study:

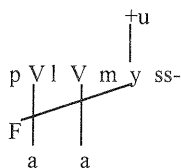
- (7) faarilyj- ‘road’
 hambys- ‘docker’
 hamyör- ‘picture’
 kostymm- ‘suite’
 martyyr- ‘martyr’
 palamyss- ‘story’

psykiatr- ‘psychiatrist’
 synagoog- ‘synagogue’

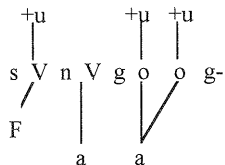
hyov- ‘to need’
 lyoon- ‘wages’

The two last stems may be mistyped in the corpora. In disharmonic stems, the feature F is associated with a subset of V-positions only (Goldsmith 1985: 267). Consider the examples in (8):

- (8) a. *palamyss-* ‘story’



- b. *synagoog-* ‘synagogue’



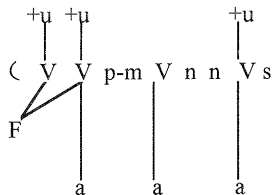
3.1.2. Treatment of Compounds

A few evident compounds were found in the material used for this study:

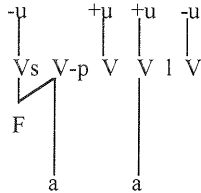
- (9) *čyöp-mannos* ‘shopkeeper’
kyöp-mannos ‘shopkeeper’
tsyöp-mannos ‘shopkeeper’
myörda-mannos ‘peculiarity’

These compounds indicate that the feature F cannot spread across the (word)boundary between the two parts of the compounds. Thus, Finnish Romani treats the compounds in the same way as the Finnish language does (Karlsson 1982: 104):

- (10) a. *čyöpmannos* ‘shopkeeper’



b. *isäpuoli* ‘stepfather’

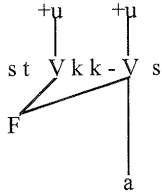


3.2. Suffix Harmony

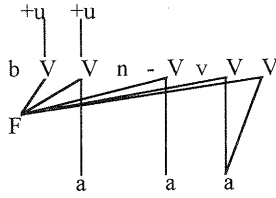
Like in Finnish, none of the suffixes is underlyingly associated with the feature F; thus, underlyingly there are five possible suffix vowels. These are the original Romani vowels {a, e, i, o, u}. The only possible exception is the very rare suffix *-ys* that occurs three times in the corpora; in the Thesleff corpus, we find the word *balamys* ‘story’, in which the front vowel *-y* cannot be result of suffix harmony. Polymorphemic items such as *miriki+ä* ‘pearl’ *sniidr+ä* ‘line (of a fishing-rod)’ and *seng+ös* ‘bed’ indicate that the suffix harmony follows somewhat different principles than the internal harmony. All the five front vowels {i, y, e, ö, ä} can act as front harmony vowels triggering the spread of the feature F from stem to suffixal V-positions.

Finnish Romani has never fully adopted the Finnish type of suffix harmony. The spreading of the feature F is quite often blocked within the stem. Suffix harmony takes place even partially in only about 45 % of the instances where it could be expected to function. This is an important difference compared with the Finnish suffix harmony, which takes place very regularly (Karlsson 1982: 99). Consider the examples presented in (11):

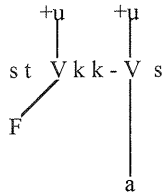
(11) a. *stykös* 'piece'



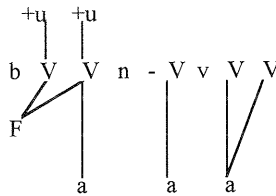
b. *byönävää* 'to pray'



c. *stykkos* 'piece'



d. *byönavaa* 'to pray'



In (11a) and (11b), we see that the suffixes *-Os* and *-AvAA* harmonise in frontness with the stems, while in (11c) and (11d) the suffixes do not obey vowel harmony.

Table (2) illustrates the tendencies of the suffix harmony to take place in different phonological contexts. The statistics presented include all 506 items in which (i) the stem that ends in a consonant and (ii) there is at least one front harmony vowel in the stem. The suffix harmony here is considered to have taken place if at least one of the suffixal harmony vowels is fronted. Thus, full and partial effects of the suffix harmony are treated together. Table (2) indicates that the tendency of the vowel harmony to apply may decrease as the sonority of the last segment of the stem increases.

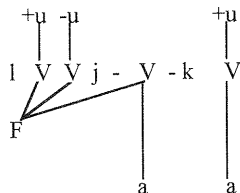
	Front		Back	
	N	%	N	%
Stop / ____	163	67.71	80	32.92
Fricative / ____	33	62.26	20	37.74
Sonorant / ____	86	54.41	73	49.01
Semivowel / ____	26	50.98	25	45.91

Table 2. Suffix harmony according to the end of the stem.

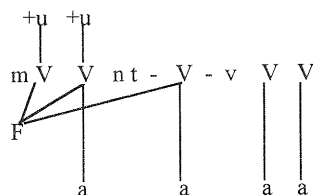
Suffix harmony is fully completed (throughout all suffixes) in about 37 % of the items studied, mostly in forms with one short - (C)V(V)C suffix (about 20 %), such as *-A* or *-Os*. In forms with more than one suffix, the feature F usually spreads only up to the first suffix (the first V-position); after the first suffix, the spreading of the feature F is normally blocked by the morpheme boundary that follows. This is the case in many inflected forms and derivatives of nouns with oblique stem and in many verbal forms. The ACC.SG. morpheme that forms the oblique stem often still obeys the vowel harmony, but that is not the case with the suffixes

that follow it (12a). Likewise in verbal forms in *-AvAA*, the first vowel alone is affected much more commonly than all three (12b).

(12) a. *lyijäko* 'voice + GEN.SG.'



b. *myöntävää* 'to admit'



As for the verbal endings like in *-AvAA*, it is open to dispute whether the first vowel belongs to the stem or to the suffix. Hedman (1996) provides verbal paradigms in which the first vowel remains unchanged throughout the inflection and might therefore be considered a part of the stem. However, other accounts such as Valtonen (1968: 132) and Koivisto (1987) give paradigms with contracted forms in 3rd pers. SG. and 3rd pers. PL.:

- (13) 1. *tenk-avaa* 'to think'
 2. *tenk-aveha*
 3. *tenk-avela*, *tenk-ela*, *tenk-ina*
 1. *tenk-avaha*, *tenk-aveha*
 2. *tenk-avena*, *tenk-ena*
 3. *tenk-avena*, *tenk-ena*

Also derivatives such as participles in *-imen*, e.g. *tenk-imen* ‘thought’ imply that the first vowel is part of the suffix rather than part of the stem.

Due to the reasons stated above, not all suffixes are intent on obeying vowel harmony. Table (3) shows the tendencies of some Finnish Romani suffixes to obey vowel harmony. Unfortunately, many of the suffixes are very rare in the material used. Change, thus, may play a role in the results. The tendencies are presented according to the distance from the stem.

The results found here seem to coincide quite well with Valtonen (1968). Valtonen (1968: 94) writes that the suffixes *-ös*, *-ö*, *-kö*, *-äl* and *-tä* are occasionally found in the ‘lower’ (informal) style. Valtonen (1968: 95) also correctly points out that the variants **-elä* and **-inä* do not occur. However, the results found here do not give support to Valtonen’s opinion that the verbal ending *-avaa* does not obey vowel harmony.

3.3.1. Suffix Harmony and Disharmonic Stems

In Finnish, the feature F usually spreads to a suffix vowel in forms where the final vowel of the disharmonic stem is associated with the feature F (cf. Goldsmith 1985: 267). As Karlsson (1982: 101) points out, there are, however, a few items with [y] and a back harmony vowel, in which the suffix harmony may fail to operate (e.g. *analyysi* ‘analysis’, *dynamiitti* ‘dynamite’, *fysiikka* ‘physics’). In Finnish Romani, the situation is somewhat different: while the suffix harmony occasionally takes place, mostly a final front harmony vowel of a disharmonic stem does not enforce the spreading of the feature F to the suffix vowels. Consider the following examples:

- (14) a. hamyör-ä ‘picture’
 palamyss-ös ‘story’
- b. hambys-os ‘docked’
 kostymm-a ‘costume’

Stem type	Suffix				%	Freq.	Example	Gloss	
	I	II	III	IV					
Noun	-A				56,1	82	työöm-ä	rein	
	-Os				56	50	lyön-ös	wages	
	-A-				69,17	43	lyij-ä	voice	
	-As-				44,44	9	symn-äs-	thimble	
	-O-				0	2	tyyg-o-	tissue	
	-Os-				75	20	syön-ös-	sea	
	-U-				100	1	fäärd-y-	car	
			-hA			33,33	3	lyij-ä-hä	voice +ABL
			-jA			100	1	fäärd-y-jä	car+PL
			-k-	-O		27,91	43	byööv-ä-kö	gallows+ GEN
			-kier-	-O		0	17	symn-äs-kier-o	thimble
		-ib-	-A			3,33	57	gryyn-ib-ä	clarity
			-Os-			0	57	byöv-ib-os-	hanging
			-k-	-O	0	9	byöv-ib-os-k-o	hanging + GEN	
			-kier-	-O	0	42	snygg-ib-os-kier-o	cleaner	
Adjective	-O				22,22	9	rönsk-ö	dissolute	
	-ik-	-O			0	1	yyl-ik-o	weak	
	-itik-	-O			0	1	byg-itik-o	of barley	
			-vitik-	-O		0	5	hyöst-ä-vitik-o	autumn-like
Verb	-A-				48	75	byön-ä-	to need	
	-U-				60	46	byr-y-	to begin	
			-vAA			3,33	61	myönt-ä-vää	to admit
		-AA				0	3	syy-vaa	to sew
		-eIA				0	2	symm-ela	sew + SG 3
		-inA				0	3	byöv-inä	hand + PL 3
		-UIA				50	4	flyyg-ylä	fly + SG 3
	Adverb	-Al				25	4	tryst-äl	around
-Al-					0	2	tryst-äl-	around	
			-O			0	1	tryst-äl-ö	around
-Om					100	1	tryst-öm	around	

Table 3. Tendencies of different suffixes to obey vowel harmony.

martyyr-o ‘martyr’
 palamyss-os ‘story’

Thus, also here suffix harmony is optional in Finnish Romani.

4. Conclusions

This paper indicates that today’s Finnish Romani has a vowel harmony system that in many respects resembles its origin, the Finnish front/back vowel harmony. However, as for the suffix harmony in particular, Finnish Romani has not fully adopted the Finnish system. There are a few interesting differences between the two languages. These include the following facts:

- (i) the front vowels in Finnish Romani are distributed in a small set of lexical entries, which remarkably limits the scope of the internal harmony.
- (ii) unlike the internal harmony, the suffix harmony is not obligatory in Finnish Romani. Instead, quite often the spreading of the feature F is blocked within the stem.
- (iii) not all suffixes tend to obey suffix harmony in Finnish Romani. Typically, in forms with more than one suffix, the suffix harmony affects only the first suffix (the first V-position) in Finnish Romani; after the first suffix, the spreading of the feature F is blocked by the morpheme boundary that follows.

Valtonen (1968: 95) points out that the suffix harmony characterizes the ‘lower’ (informal) register. According to him, the ‘upper’ (i.e. formal) register lacks suffix harmony. On the basis of the corpora available, this cannot be verified, as the corpora are mostly dictionaries or word-lists, or texts written in formal style.

The study of the differences between the registers would require material based on tape-recordings with authentic spoken language.

References

- Bloch, Jules (1921) *L'indo-Aryan du Vēda aux temps modernes*. Paris.
- Chomsky, Noam & Halle, Morris (1968) *The Sound Pattern of English*. New York: Academic Press.
- Cortiade, Marcel (1989) *Romany Grammar. Volume 1: General Information, Phonology and Morphology*. (Unpublished manuscript; available at the Research Institute for the Languages of Finland.)
- Goldsmith, John (1976) *Autosegmental Phonology*. Bloomington.
- Goldsmith, John (1985) Vowel Harmony in Khalka, Mongolian, Yaka, Finnish and Hungarian. In C. Even & J. Anderson (eds.). *Phonology Yearbook 2*. Cambridge: Cambridge University Press.
- Hedman, Henry (1996) *Romanikielen kielioppiopas. Sar me sikjavaa romanes*. Jyväskylä: Opetushallitus.
- van der Hulst, Harry & van de Weijer, Jeroen (1996) Vowel Harmony. In John A. Goldsmith (ed.) *The Handbook of Phonological Theory*. Cornwall: Blackwell.
- Karlsson, Fred (1982) *Suomen kielen äänne- ja muotorakenne*. Juva: WSOY.
- Kiparsky, Paul (1982) Lexical Morphology and Phonology. In I.-S. Yang (ed.). *Linguistics in the Morning Calm*. Seoul: Hanshim.
- Kochanowski, Jan (1963) *Gypsy Studies I-II*. New Delhi.
- Koivisto, Viljo (1971) *Johannesko evankeliumos*. Risadas Viljo Koivisto. Turku: Suomen Piiliseura.
- Koivisto, Viljo (1987) *Rakkavaha romanes. Kaalengo tsimbako sikjibosko liin*. Helsinki: Ammatikasvatusthallitus & Valtion painatuskeskus.
- Manuš, Leksa, Neilands, Jānis & Rudevičs, Kārlis (1997) *Ķigānu-Latviešu-Angļu- Latviešu-Ķigānu Vārdnīca*. Rīgā: Zvaigzne ABC.
- Miklosich, Franz (1872/81) *Über die Mundarten und Wanderungen der Zigeuner Eupora's*. Denkschriften der Kaiserlichen Akademie der Wissenschaften, philosophisch-historische Classe, Bd. 21-31. Wien.
- Mishra, Vidhata (1972) *A Critical Study of Sanskrit Phonetics*. The Chowkhamba Sanskrit Studies Vol. LXXXIII. The Chowkhamba Sanskrit Series Office.
- Mustalaiskielen ortografiakomitea (1972) *Mustalaiskielen normatiivi sanasto*. Helsinki.

- Pott, August Friedrich (1844/45) *Die Zigeuner in Europa und Asien I-II*. Halle.
- Thesleff, Arthur (1901) *Wörterbuch der finnländischen Zigeuner*. Acta Societatis Scient. Fenn. 29. Helsinki.
- Valtonen, Pertti (1968) *Suomen mustalaiskielen kehitys eri aikoina tehtyjen muistiinpanojen valossa*. (Unpublished licentiate thesis; available at the Research Institute for the Languages of Finland.)
- Valtonen, Pertti (1970) *Markusesko evankeliumos. Kaalengo tsibbaha*. Riitba Pertti Valtonen. Tikkurila: Kristillisen kirjallisuuden seura.

Contact address:

Kimmo Granqvist

Research Institute for the Languages of Finland

Sörnäisten rantatie 25

FIN-00500 HELSINKI

Finland

E-mail: kimmo.granqvist@domlang.fi