



Arvo Krikmann

**Relations between vowels
in the alliteration
of Finnic runic folksongs**

**The conference "Across Borders IV"
Krosno, Poland, April 16–17, 2010**

The older folksongs of most Finnic peoples share the so-called Kalevala form, the main features of which are parallelism, alliteration (and assonance), and quadripartite trochaic rhythm.

Here are some fragments from the monograph "The Temporal Structure of Estonian Runic Songs" by Jaan Ross and Ilse Lehist:

meie kodu kauge'ella	our home is far away
viisi verstada vaheta	five versts from here
kuusi kuivada jõgeda	with six drive rivers
seitse sooda sitke'ada	seven sloppy swamps
kaheksa kala+mereda	eight seas of fish
üheksa hüva ojada	nine beautiful brooks
kümme külma allikada	ten cold springs in between

pere+naine naisukene

housewife, dear woman

muile annid muida tööda

you gave different tasks to others

sulasele suurta tööda

to the hired man big jobs

mulle tööda albusamba

to me worse work

annid mul aned ajada

you gave me to drive the geese

annid mul kanad kaitseada

to protect the chickens

lestas+jalad le pistada

to pacify the web-footed ones

varvas+jalad vaigistada

to calm down the toefooted ones

ajasin aned vesile

i drove the geese to the waters

kargutin kanad kesale

i drove the chicken to the fallow field

lestas+jalad lepikusse

the webfooted ones to the alder ground

varvas+jalad vainiulle

the toefooted ones to the meadow

tuli kuri, kurja lindu

an angry, evil bird came

ajas mo aned vesilta

drove my geese from the waters

karguti kanad kesalta

drove my chicken from the fallow field

lestas+jalad lepikusta

the webfooted ones from the alder ground

So, the alliteration works "from the left to the right" within a single line;

parallelism, in turn, operates "from the top to the bottom", paraphrasing the content set up in the first line through the following two or more lines. And, of course, it makes the unfolding of events or thoughts very slow and long-winded.

Finnic languages have strong natural premises to evolve the alliterative type of verse:

the stress is on the first syllable, before the invasion of late foreign words they had almost no consonant clusters at the beginning of words and so on.

The percentage of alliterative verses in Finnic runic songs is said to oscillate somewhere between 80 and 95 in average, depending on the age of the material, language and dialect, the kind of the material (whether epic or lyric), etc.

The alliterative chain usually consists of two links, but their number can also be bigger (up to five).

A verse usually includes one single alliterative chain, but sometimes also two or even more:

mure musta parre peale: mure musta + parre peale ('[I put] my concern on the black joist')

Alliteration is, of course, also a semantic phenomenon: it results from the choice of words and thus brings about the specific so-called "semantic mist" in the contents of the text. But primarily, though, it is considered as a euphonic phenomenon which helps to make music out of the lexical substance of the song as well.

The research of the euphonic (phonetic) aspect of alliteration readily suggests all kinds of quantitative, that is statistical, approaches.

There are some typical problems that were discussed in the existing literature already long ago and which one inevitably encounters when trying to build up the statistics of the topic.

The first question is, as Pentti Leino has put it, "are the instances of alliteration in the text under investigation sporadic or at least partially the result of a conscious effort to alliterate?". Leino himself is convinced that only the cases of conscious choices must be taken into consideration, leaving out, for example, spontaneous coincidences of first sounds of various "grammatical" or "auxiliary" words in a verse line.

However, it seems quite dubious how such a bivalent distinction could be made, considering the highly continuous character of the lexicon – be it in the runic song or elsewhere. In my own observations all cases of alliteration – from clearly conscious and poetic to clearly spontaneous and "grammatical" ones – have been included in the statistics.

Of course, the repetition of initial C-s, particularly stops, is the most reliable ground to speak about alliteration, therefore some authors tend to exclude "zero consonant" cases from alliteration.

Things are more problematic with the first syllable vowels. Some authors suggest to clearly distinguish ALLITERATION as a repetition of initial C-s from ASSONANCE as a repetition of first-syllable vowels (whether in post-consonant or "zero-consonant" position).

Matti Sadeniemi's typology of Finnic alliteration, however, includes all four possible combinations:

1) CV : CV – both consonants and (nuclear) vowels coincide:

kuusi kuivada jõgeda 'six drive rivers'

2) ØV : ØV – words begin with one and the same vowel (so-called "zero consonant" case):

ajasin aned vesile 'I drove the geese to the waters'

3) CV₁ : CV₂ – identical consonants are followed by different vowels:

viisi verstada vaheta 'five versts from here'

4) ØV₁ : ØV₂ – words begin with different vowels (another variant of the "zero consonant" case):

mitu orja ilma pealla 'how many slaves in the world'

The first two cases represent the "**strong**", and the last ones the "**weak**" alliteration.

Hence, the question arises whether it is justified to make a simple bivalent distinction between the assonance and non-asonance, or rather, perhaps, we have a continuous scale of vowel relationships where the "pure" CV-alliteration is just the highest level or "limit case" of the scale.

If, under the given lexical-semantic conditions, the perfect coincidence of postconsonant or word-initial vowels is impossible to achieve, the intuitive search for a subsequent word pair with maximally or at least a sufficiently good similarity of different first syllable vowels (CV1:CV2 or ØV1: ØV2) will follow.

This hypothesis is also supported by observations of early Finnish folklorists and linguists (Lönnrot, Europaeus, Ahlqvist, Genetz) who have mentioned that some vowel combinations like a-ä, o-u, y-ö seem to be more favoured than others.

My own preliminary tests with quite a limited material of Estonian runic songs in the middle of the 1960ies also indicated the existence of certain ranks of preference, certain system of rules that seemed to govern the vocalism of runic alliteration, and encouraged to continue the search for more valid evidence for and nature of these rules.

This leads us directly to the next cluster of questions: What kind of and how much empirical material should we use for that purpose in order to reach the necessary level of representativeness and reliability?

Runic songs evidently strive to achieve the "pure", or "strong", same-vowel alliteration (CV:CV or ØV: ØV), making these cases statistically prevailing.

In other words, the general frequency level of "weak" alliteration is quite low, therefore the total number of scoured texts should be notably larger (?).

Toimetanud Ülo Tedre

EESTI RAHVALAULUD
ANTOLOOGIA



freimidega / ilma

In 1966, I guess, I started to go through the manuscript of (then yet unpublished) "Anthology of Estonian folk songs" (1969–1974) that includes over 7000 song texts and is now also accessible on the Internet:

<http://www.folklore.ee/laulud/erla/>

On small sheets of paper I wrote out each verse where the "consonant-proved" CV1:CV2 alliteration occurred – perhaps 40 000 slips altogether.

I had great difficulties finding a suitable piece of Finnish runo songs.

Actually, the situation is paradoxical because in general and in principle the premises for studying any textual aspects of Finnic runic songs are honestly brilliant:

during the last ten years or so, practically all of the corpuses of Finnish-Karelian and Estonian runic songs have been digitised – perhaps 160 – 180 thousands of song texts from both sides are now at the disposal of folklorists in Helsinki and Tartu.

Finns have also published on the Internet the data base containing the huge collection of "Suomen kansan vanhat runot" (Ancient Poems of the Finnish People):

The data base of the "Suomen kansan vanhat runot" (Ancient Poems of the Finnish People) on the Internet

<http://dbgw.finlit.fi/skvr/>

SKS SEURA KANSANRUNOUSARKISTO KIRJALLISUUSARKISTO KIRJASTO SKS KIRJAT FILI YHTEYSTIEDOT SKVR

SUOMEN KANSAN VANHAT RUNOT SKVR TIETOKANTA

► Kalevalamittaiset runot ► Suomen Kansan Vanhat Runot (SKVR) ► ©

SKVR-tietokanta
Suomen Kansan Vanhat Runot sähköisessä muodossa

Paikkakunta

Perinnealue

Aunus q
Ei aluetietoa
Etelä-Karjala (Karjalan kannas) h
Etelä-Pohjanmaa k
Etelä-Savo f

Pitäjä

Ahlainen
Akaa
Alahärmä
Alajärvi
Alakiminki

Kerääjä on Sukunimen alkukirjain: Kaikki kerääjät

Keruu aika alaraja - yläaraja

Runon tekstissä tai valitse lisäehto esiinty sanat

Tulokset Näytä 40 tulosta / sivu Järjestä tulokset runon numeron mukaan

Hae Tyhjennä

However, in their present form, the Finnish texts are unfortunately "inedible" for any computer programs. They are dialectal, full of all kinds of diacritic marks and need a lot of time-consuming preparatory work.

Of course, nobody uses paper slips anymore.

The only Finnish(-Karelian) sources I find on the Internet were two books by Elias Lönnrot –

his epos "Kalevala" (the "New" Kalevala, 1849) and his edition of slightly edited and modified runic songs "Kanteletar" (1840) –

both about 22 000 verses.

In full awareness of their partly non-folkloric nature and a certain "Lönnrot's impact factor" in them, I nevertheless tested both of them for CV1:CV2 rules, recalling the old joke Estonian folklorists used to tell 40–50 years ago:

Question: What is the main difference between "Kalevala" and the Estonian epos "Kalevipoeg"?

Answer: "Kalevala" contains 95% of runic songs and 5% of Lönnrot's own poetry, whereas "Kalevipoeg" contains 5% of folkloric verses and 95% of Kreutzwald's personal creation.

It took almost a year to build up the statistics of the Estonian anthology.

It took almost a week to build up the statistics of "Kalevala".

The next question is: what and how one should count when encountered with CV₁:CV₂ problematics.

The first thing to come to one's mind is certainly: a representative enough random sample of verses.

But as that sample, under the above-mentioned conditions, must be voluminous enough – including, in the ideal case, each CV₁:CV₂ occurrence in the given corpus – the sample will inevitably include a considerable number of highly recurrent pairs or chains of alliterating word stems – spontaneous co-occurrences of "grammatical" words, various stereotypical verses and formulae, etc.

Evidently, they are frequent, first and foremost, not due to their euphonic beauty, but primarily for some semantical reasons.

Some highly frequent alliterative word pairs in "Anthology of Estonian folk songs"

neiu(ke)~neitsi(ke)+noor(uke)	maid ~ virgin+young	147
kuulma + kostma	hear + reply	81
peiukene + poisikene	bridegroom + boy	72
mis + meie	what + we	68
siis + saama	then + get ~ receive ~ ...	68
mina + mees	I + man ~ ~ ...	39
mis + ma	what + I	38
kui + kodu	when ~ if + home	37
piht(a) + peen(ike)	waist + thin ~ slim	33
mina + mõistma	I + understand	32
lehm + lüps(ma)	cow + milk(ing)	30
nutma + noor(ik)	cry + young ~ bride	30
see + saama	this~that + get~receive~...	30
naine + noor(uke)	woman ~ wife + young	27
mis + mu(l(le))	what + me ~ to me ~ my	25
meie + maa	our + land ~ soil ~ ...	24

Some highly frequent alliterative word pairs in "Kalevala"

vanha Väinämöinen	old Väinämöinen	321
noin nimesi	so mentioned	232
lieto Lemminkäinen	mild Lemminkäinen	187
en ~ et ~ ei ole ~...	not + be (misc.)	182
vaka vanha	old pious (...)	116
se seppo	that + smith	71
sanovi sanalla ~ ...	say + word (s.s.)	62
kaunis Kauko(mieli)	pretty Kaukomieli	57
kirjo ~ kirja kansi	mottled ~... + cover	55
Kullero, Kalervon poika	Kullervo, the son of Kalervo	52
iän ikuinen	age + old (s.s.)	50
on ~ oli Ilmarinen	be ~ was ~ ... + Ilmarinen	46
oli ~ on + hyvä	be ~ was ~ ... + good	41
nuori neito ~ ...	young maiden	39
lausu lausehella	say + sentence (s.s.)	35
naisten nauru ~ ...	laugh(ter) + woman ~-men	33

s.s. = 'the same word stem'

Therefore a better estimate could be the total number of different CV1:CV2 word stem pairs themselves, ignoring their individual recurrences. This method should indicate more adequately the pressure each CV-initial part of the vocabulary has undergone in the process of creating and recreating (and not just transmitting) folk songs.

Conspicuously enough, in some parts of the lexicon the alliterative pressure has been so strong that the supply of "normal", "meaningful" words appears to have been exhausted and the last "emergency" resorts have been activated, like proper names, descriptive words, etc.

Many more specific questions will arise, for example, with standardising of the multitude of first syllable vowel combinations. The phonetic and phonological systems of various Finnic languages – and of various dialects of the same language – differ quite substantially, so it may be hard to make the results of different observations comparable.

One particular question is, for example, how to proceed in our statistics with the great variety of first syllable diphthongs that can be of three different origins in Finnic languages:

Diphthongs in Finnic languages are of three different origins:

1) old diphthongs that existed already in the hypothetical Proto-Finnic language:

EST **naine**, FIN **nainen** < ***nainen** 'woman'

EST **teine**, FIN **toinen** < ***toinen** 'other, second'

EST **poeg**, FIN **poika** < ***poika** 'son'

EST **kõiv**, FIN **koivu** < ***koivu** 'birch'

2) as the result of the syncopation of weak stops on the border between the first and second syllables:

EST **viga** : **vea**, FIN **vika** : **vian** < ***vika** : **viyan** 'mistake; vice'

EST **tuba** : **toa**, FIN **tupa** : **tuvan** < ***tupa** : **tuβan** 'room; hut'

EST **rida** : **rea**, FIN **rita** : **ridan** < ***rita** : **riðan** 'row, line; trap'

EST **süsi** : **söe**, FIN **sysi** : **syden** < ***süsi** : **süðen** 'coal'

3) the late diphthongs that in some Finnic languages (like Finnish and Karelian), and also in North-Estonian dialects are the descendants of Proto-Finnic long vowels:

FIN, N-EST **tie** < ***tee** 'way, road'

FIN, N-EST **suo** < ***soo** 'swamp'

FIN, N-EST **työ**, **tüö** < ***töö** 'work'

In my statistics all diphthongs were registered according to their so-called nuclear vowels, which, in the general case, means their first components and, in the case of late diphthongs, the second component.

Still one question is:

How to calculate the strength of preferences for each pair of vowels preceded by this or that word-initial consonant?

One could guess that it might be sufficient just to take the representative dictionary of the given language, ascertain and square the "lengths" of each particular CV-section and compare them with corresponding frequencies in the alliterative matter.

This way would be, however, deceptive because the individual probabilities of actualising different constituents (words or stems) of the given CV-group are far not equal.

What is really needed seems to be something "between the vocabulary and the text".

I found the absolute frequencies of all CV1:CV2 word stem pairs, summarised "meetings" (intersections) of different nuclear vowels through all consonants (what is not statistically flawless).

To estimate of the density of connection between different vowels in the alliteration of Finnic folk songs I used the so-called **colligation coefficients (or "bivariate percentages") that are calculated from the formula**

$$\lambda_{AB} = \frac{A \cap B \times \Sigma_{tab}}{\Sigma A \times \Sigma B}$$

where $A \cap B$ is the number of "meetings" (intersections) of events A and B ,

ΣA and ΣB are the summary frequencies of these events in the given collection and

Σ_{tab} is the sum total of all numerical data in the table through all of its rows and columns.

Here are the preliminary results for the "Anthology of Estonian Folk Songs":

Raw data

A	X									
E	433	X								
I	407	945	X							
U	430	181	180	X						
O	339	161	144	780	X					
Õ	315	195	188	239	185	X				
Ä	408	201	183	105	78	91	X			
Ü	126	115	118	145	90	62	59	X		
Ö	32	48	40	27	20	26	15	60	X	
Σ	2490	2279	2205	2087	1797	1301	1140	775	268	14342
	A	E	I	U	O	Õ	Ä	Ü	Ö	Σ

λ-field

A	X									
E	0.88	X								
I	0.88	2.46	X							
U	0.94	0.46	0.50	X						
O	0.98	0.55	0.53	2.84	X					
Õ	1.21	0.90	0.92	1.15	1.18	X				
Ä	1.96	1.15	1.12	0.64	0.63	0.98	X			
Ü	0.62	0.67	0.74	0.90	0.73	0.69	0.81	X		
Ö	0.54	0.95	0.85	0.58	0.56	0.97	0.70	2.85	X	
	A	E	I	U	O	Õ	Ä	Ü	Ö	

The ranks of decreasing preferences in the complex of non-low vowels in Estonian runic songs appears to be the following (the vowel \tilde{o} being regarded as back mid illabial):

1) vowels differing only in height (high/mid), all the rest being the same: $o \leftrightarrow u$, $e \leftrightarrow i$, $\tilde{o} \leftrightarrow \tilde{u}$;

2) vowels differing only in labiality (labial/illabial), all the rest being the same: $o \leftrightarrow \tilde{o}$, $e \leftrightarrow \tilde{e}$, $i \leftrightarrow \tilde{i}$;

3) vowels differing in height and labiality, the gravity (back/front) being the same: $u \leftrightarrow \tilde{o}$, $i \leftrightarrow \tilde{e}$, $e \leftrightarrow \tilde{u}$;

4) vowels differing only in gravity (back/front), all the rest being the same: $e \leftrightarrow \tilde{o}$, $u \leftrightarrow \tilde{u}$, $o \leftrightarrow \tilde{e}$;

5) vowels differing in height and gravity (back/front), all the rest being the same: $i \leftrightarrow \tilde{o}$, $o \leftrightarrow \tilde{u}$, $u \leftrightarrow \tilde{e}$;

6) vowels differing in labiality and gravity (back/front), all the rest being the same: $\tilde{o} \leftrightarrow \tilde{e}$, $e \leftrightarrow o$, $i \leftrightarrow u$;

7) vowels differing in all three features: $\tilde{o} \leftrightarrow \tilde{u}$, $i \leftrightarrow o$, $e \leftrightarrow u$.

Group	Height	Labiality	Gravity	V_1V_2	λ	λ average
(1)	X			üö uo ei	2.85 2.84 2.46	2.72
(2)		X		oö eö iü	1.18 0.95 0.74	0.96
(3)	X	X		uö iö eü	1.15 0.85 0.57	0.89
(4)			X	eö uü oö	0.90 0.90 0.56	0.79
(5)	X		X	iö ou uö	0.92 0.73 0.58	0.74
(6)		X	X	öö eo iu	0.97 0.55 0.50	0.67
(7)	X	X	X	öü io eu	0.69 0.53 0.46	0.56

And here is the same for **low vowels**:

		λ			
a	aä	1.96	low		
	aõ	1.21	low/ not-low	back	illabial
	ao	0.98			illabial/ labial
	au	0.94			labial
	ae	0.88	not-low	back/ front	illabial
	ai	0.88			illabial/ labial
	aü	0.62		illabial/ labial	
	aö	0.54		labial	

		λ			
ä	äa	1.96	low		
	äe	1.15	low/ not-low	illabial	front
	äi	1.12			front/back
	äõ	0.98			front/back
	äu	0.81	not-low	illabial/ labial	front
	äö	0.70			front/ back
	äu	0.64		front/ back	
	äo	0.63		back	

As can be seen on the graph above, the low *a* and *ä* also reveal quite regular behaviour.

The most important marker is lowness.

On the background of all field of relations *a* continues with gravity (that is, prefers back partners); whereas *ä* seems to prefer illabial partners.

I have not yet managed to build up an integral quantitative estimate to describe the alliterative behaviour of all Estonian vowels together.

The same calculations for "Kalevala":

Raw data

A	X								
E	271	X							
O	271	138	X						
U	256	122	383	X					
I	206	342	118	105	X				
Ä	190	142	95	76	64	X			
Y	62	39	46	69	46	28	X		
Ö	19	41	15	20	19	12	17	X	
Σ	1275	1095	1066	1031	900	607	307	143	6424
	A	E	O	U	I	Ä	Y	Ö	Σ

λ -field

A	X							
E	125	X						
O	128	76	X					
U	125	69	224	X				
I	115	223	79	73	X			
Ä	158	137	94	78	75	X		
Y	102	75	90	140	107	97	X	
Ö	67	168	63	87	95	89	249	X
	A	E	O	U	I	Ä	Y	Ö

Ranks of preferences between non-low vowels in "Kalevala"

Group	Height	Labiality	Gravity	V ₁ V ₂	λ	λ average
(1)	X			yö ou ei	2,29 2,38 2,13	2,27
(2)		X		eö iy	1,50 1,06	1,28
(3)			X	uy oö	1,28 0,68	0,98
(4)	X		X	oy uö	0,88 0,87	0,88
(5)	X	X		iö ey	1,02 0,61	0,82
(6)		X	X	eo ui	0,72 0,78	≈ 0,75
(7)	X	X	X	oi eu	0,91 0,61	≈ 0,75

The same for **low vowels** in "Kalevala":

		λ			
a	aä	1.70	low		
	ao	1.31	low/ not-low	back	illabial/ labial
	au	1.20			illabial
	ai	1.19		back/ front	illabial/ labial
	ae	1.06			
	ay	0.89			
	aö	0.64			

		λ			
ä	äa	1.70	low		
	äe	1.37	low/ not-low	not-high	?
	äo	1.13			
	äö	1.00		high	?
	äy	1.00			
	äi	0.91			
	äu	0.88			

Preferences for pairing different vowels in Finnic alliteration neatly correlate with other types of phonetic relationships between these Finnic sounds, e.g. diachronic changes, synchronic differences between related languages and different dialects of one and the same language, etc.

THANK YOU!