On the vowel euphony in Finnic alliterative folksongs

The folk poetry following the quadripartite trochaic metre has been common to most Baltic-Finnic peoples. The origin of the metre is hidden in the distant past, perhaps two millennia ago, and it was in active use in folk poetry until the 20th century. This folk poetry has included mythological epic songs, adventure warrior epic, incantations, wedding songs, ballads, historical songs, various lyric songs and proverbs and riddles – a broad spectrum of oral tradition. The same poetic language has been used also in the literary epics of the Finnish Kalevala and Estonian Kalevipoeg. In addition to meter, a number of various structural and stylistic means have formed the special register or poetic language, usually called Kalevala metre and Kalevala language, named after the most widely known product of the poetic tradition. The article of Arvo Krikmann is devoted to one very prominent quality of the Baltic-Finnic poetic register, alliteration.

Below are some excerpts from the monograph The Temporal Structure of Estonian Runic Songs by Jaan Ross and Ilse Lehiste (2001: 14, 139, 141):

meie koodu kaugõella
viisi verstada vaheta
kuusi kuvada jõgeda
seitsse sooda sitke’ada
kahiska kala+mereda
ühkiska hüva ojada
kümmne külmä allikada
. . .
pere+naine naisukene
muile annid muida tüöda
sulasele suurta tüöda

our home is far away
many miles from here
with six dry rivers
seven sloppy swamps
eight seas of fish
nine beautiful brooks
ten cold springs in between

housewife, dear woman
you gave different tasks to others
to the hired man you gave big jobs

mulle tüöda albusamba
annid mul aned ajada
annid mul kanad kaitsesta
leistas+jalad lepistada
varvas+jalad vaigistada

to me you gave worse work
you gave me the task to drive
the geese
you gave me to protect the chickens
to pacify the web-footed ones
to calm down the toe-footed ones

ajasin aned vesile
kargutin kanad kesale
leistas+jalad lepikusse
varvas+jalad vainiulle

I drove the geese to the waters
I drove the chickens to the fallow field
the web-footed ones to the alder grove
the toe-footed ones to the meadow

tuli kuri, kurja lindu
ajas mo aned vesilta
karguti kanad kesalta
leistas+jalad lepikusta

an angry, evil bird came
drove my geese from the waters
drove my chickens from the fallow field
the web-footed ones from the alder grove

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 two or more succeeding lines. And of course, it makes the development of events or other train of thought very slow and long-winded.

Finnic languages have strong natural bases to evolve the alliterative type of verse: the stress is on the first syllable (before the arrival of recent foreign words), they had almost no consonant clusters at the beginnings of words and so on.

The number of alliterative verses in Finnic runic songs is said to vary between 80 and 95 per cent on average, depending on the age of the material, language and
dialect, the kind of material (whether epic or lyric), etc. The alliterative chain usually consists of two links, but their number can also be higher (up to five).

A verse usually includes one single alliterative chain, but sometimes two or more: *mure musta parre peale: mure musta + parre peale* (‘I put] my concern on the black joist’).

Alliteration, 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 it is considered a euphonic phenomenon which helps to make music out of the lexical substance of the song.

Research on the euphonic (phonetic) aspect of alliteration readily suggests many quantitative, that is statistical, approaches.

There are some typical problems that were discussed in the literature long ago, and which one inevitably encounters when trying to build up the statistics. The first question is, as Pentti Leino (1970: 317) 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 instances of conscious choice 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 result of a conscious effort to alliterate? Leino himself is convinced that only the instances of conscious choice 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 result of a conscious effort to alliterate? Leino himself is convinced that only the instances of conscious choice must be taken into consideration, leaving out, for example, spontaneous coincidences of first sounds

1) **CV : CV** – both consonants and (nuclear) vowels coincide:
   - *Kuusi kuivada jõgeda*  ‘six dry rivers’

2) **OV : OV** – words begin with one and the same vowel (the so-called ‘zero consonant’ case):
   - *Ajasin ained vesile*  ‘I drove the geese to the waters’

3) **CV₁ : CV₂** – the identical consonants are followed by different vowels:
   - *Viisi verstada vaheta*  ‘five versts from here’

4) **OV₁ : OV₂** – words begin with different vowels (another variant of the ‘zero consonant’ case):
   - *Mitu orja lima pealla*  ‘how many slaves in the world’

The first two cases represent ‘strong’, and the last two ‘weak’ alliteration.

Hence, the question arises of whether it is justified to make a simple bivalent distinction between assonance and non-assonance, or whether, perhaps, we have to do with a continuous scale of vowel relationships where the ‘pure’ CV-alliteration is just the highest level or limit of the scale. If, under given lexical-semantic conditions, the perfect coincidence of postconsonantal or word-initial vowels is impossible to achieve, the intuitive search for a subsequent word pair with maximal or at least a sufficiently good similarity of different first-syllable vowels (CV₁ : CV₂ or ØV₁ : ØV₂) will follow. This hypothesis also finds support from investigators of early Finnish folklorists and linguists (Elias Lönnrot, D. E. D. Europaeus, August Ahlqvist, Arvid Genetz), who have observed that some vowel combinations like *a–ä, o–u, y–ö* seem to be more favoured than others.

My own preliminary tests with limited material from Estonian runic songs in the middle of the 1960s also indicated the existence of certain ranks of preference, and certain systems of rules that govern the vocalism of runic alliteration, and encouraged me to continue the search for further evidence for the nature of these rules.

This leads us directly to the next cluster of questions: what kind of and how much empirical material should we gather in order to reach the required degree 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 prevalent. In other words, the general frequency level of ‘weak’ alliteration is quite low, so should the total number of texts investigated be noticeably larger?

Fortunately, I had at my disposal the manuscript of (the then unpublished) *Eesti rahvalaulud. Antoloogia* (*Anthology of Estonian Folksongs*, Tedre *et al.* 1969–74) that includes over 7000 song texts and is now also accessible on the internet. In around 1966 I started to go through the manuscript. On small sheets of paper I wrote out each verse where the ‘consonant-proved’ CV₁ : CV₂ alliteration occurred, and thus arrived at a total of perhaps 40,000 slips.

I then had great difficulties in finding a suitable collection of Finnish *runo* songs.

Actually, the situation is paradoxical because in general and in principle the premises for studying any textual aspects of Finnish runic songs are promising: at the
The turn of the century, practically all major corpuses of Finnish-Karelian and Estonian runic songs have been digitised – perhaps from 160,000 to 180,000 song texts from both sides are now at the disposal of folklorists in Helsinki and Tartu.

Finns have also published on the internet the database containing the huge collection of Suomen kansan vanhat runot (‘Ancient Poems of the Finnish People’). However, in their present form, the Finnish texts are unfortunately ‘inedible’ for any computer programs. They are dialectal, full of all kinds of diacritical marks and need a lot of time-consuming preparatory work.

Of course, nobody uses paper slips any more. The only Finnish(-Karelian) sources I found on the internet were two books by Elias Lönnrot – his epic Kalevala (the ‘New’ Kalevala, 1849), and his book 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 the Kalevala and the Estonian epic, Kalevipoeg?

Answer: The Kalevala contains 95 per cent runic songs and 5 per cent Lönnrot’s own poetry, whereas Kalevipoeg contains 5 per cent folkloric verses and 95 per cent Kreutzwald’s personal creation.

It took almost a year to build up the statistics on the Estonian anthology. It took almost a week to build up the statistics for the Kalevala, as I did not need to write heaps of paper slips.

The next question was: what and how one should count when encountered with CV₁ : CV₂ problematics? The first thing to come to one’s mind is a representative 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 because of their euphonic beauty, but primarily for semantic reasons.

Here are some highly frequent alliterative word-stem pairs in the ‘Anthology of Estonian Folksongs’ (abbreviation ‘s.s.’ means ‘the same word stem’):

- neiu(ke) + noor(uke) maid ~ virgin + young 147
- kuulma + kostma hear + reply 81
- peiukene + poisikene bridegroom + boy 72

Runesinging in the Kalevala meter is one of the few European oral poetries to survive the long nineteenth century. In her comprehensive study of the poems collected in the Archangel Karelian parish of Vuokkiniemi, Lotte Tarkka places this tradition within historical and ethnographic realities, contexts of local and elite ideologies, and the system of folklore genres. The songs of the border people emerge as praxis, the communicative creation of individual and collective identities grounded in a mythic-historical view of the world. The bond between the songs and their singers is articulated through an intertextual analysis of key cultural themes and the textual strategies used in their elaboration. In performance, singers and their audiences could evoke alternative realms of experience and make sense of the everyday in dialogue with each other, supranormal agents, and tradition. The poems, as powerful representations and performatives, endowed those who voiced them with godlike creative capacities, as coined in the proverb ‘The things I put into words, I make real’.

Lotte Tarkka is Professor of Folklore Studies at the University of Helsinki. Her areas of expertise include Kalevala-meter poetry, Finnish mythology, oral poetries and textualization, genre, and intertextuality, especially in the context of archival sources.

Folklore Fellows’ Communications No. 305

50 euros (hardback)

Tiedekirja Bookstore, tiedekirja@tsv.fi
Shop online at <http://www.tsv.fi/tiedekirja>
The highly frequent alliterative word stem pairs in the *Kalevala* appeared to be the following:

<table>
<thead>
<tr>
<th>Vanha Väinämöinen</th>
<th>Old Väinämöinen</th>
<th>321</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noin nimesi</td>
<td>So mentioned</td>
<td>232</td>
</tr>
<tr>
<td>Lieto Lemminkäinen</td>
<td>Mild Lemminkäinen</td>
<td>187</td>
</tr>
<tr>
<td>En - et - ei ole - ...</td>
<td>Not + be (misc.)</td>
<td>182</td>
</tr>
<tr>
<td>Vaka vanha</td>
<td>Old pious (…)</td>
<td>116</td>
</tr>
<tr>
<td>Se seppo</td>
<td>That + smith</td>
<td>71</td>
</tr>
<tr>
<td>Sanovi sanalla - ...</td>
<td>Say + word (s.s.)</td>
<td>62</td>
</tr>
<tr>
<td>Kaunis Kauko(miel)</td>
<td>Pretty Kaukomieli</td>
<td>57</td>
</tr>
<tr>
<td>Kirjo - kirja kansi</td>
<td>Mottled - + cover</td>
<td>55</td>
</tr>
<tr>
<td>Kullero, Kalervon poika</td>
<td>Kullervo, the son of Kalervo</td>
<td>52</td>
</tr>
<tr>
<td>Iän ikuinen</td>
<td>Age + old (s.s.)</td>
<td>50</td>
</tr>
<tr>
<td>On - oli Ilmarinen</td>
<td>Be ~ was - ... + Ilmarinen</td>
<td>46</td>
</tr>
<tr>
<td>Oli - on + hyvä</td>
<td>Be ~ was - ... + good</td>
<td>41</td>
</tr>
<tr>
<td>Nuori neito - ...</td>
<td>Young maiden</td>
<td>39</td>
</tr>
<tr>
<td>Lausu lausehella</td>
<td>Say + sentence (s.s.)</td>
<td>35</td>
</tr>
<tr>
<td>Naisten naurus - ...</td>
<td>Laugh(ter) + woman ~ -men</td>
<td>33</td>
</tr>
</tbody>
</table>

Therefore, a better estimate could be the total number of different CV1 : CV2 word-stem pairs themselves, ignoring their individual repeatability. 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) folksongs. 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 and other ideophonic words.

Many more specific questions will arise, for example with standardising 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 cope in our statistics with the great variety of first-syllable diphthongs that can be of three different origins in Finnic languages:

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

   Est. naine, Fi. nainen < "nainen 'woman'
   Est. teine, Fi. toinen < "toinen 'other, second'
   Est. poika, Fi. poika < "poika 'son'

   Est. köiv, Fi. koivu < "koivu 'birch'

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

   Est. viga : vea, Fi. vika : vihan < "vika : vihan 'mistake; vice'

   Est. tuba : toa, Fi. tupa : tuvan < "tupa : tuhan 'room; hut'

   Est. rida : rea, Fi. rita : ridan < *rīta : rīdan 'row, line; a certain trap'

   Est. süsi : sōe, Fi. sysi : syden < *sīsi : sūden 'coal'

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

   Fi., North Est. tie < "tič 'way, road'

   Fi., North Est. suō < "sōo 'swamp'

   Fi., North Est. tyō ~ tiū < "tūo 'work'

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

The next question was 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-group and compare them with corresponding frequencies in the alliterative matter. This approach would, however, be deceptive because the individual probabilities of actualising different constituents (words or stems) of the given CV-group are far from equal. What is really needed seems to be something 'in between the vocabulary and the text.' To arrive at something like this, I checked the absolute frequencies of all CV1:CV2 word-stem pairs and summarised 'meetings' (intersections) of different nuclear vowels through all consonants (which is not statistically flawless). To estimate the density of connection between different vowels in the alliteration of Estonian folksongs and in the *Kalevala*, I used Zbigniew Pawłowski’s (1967: 38) so-called colligation coefficients (or λ- coefficients, 'bivariate percentages') that are calculated from the formula:

\[ \lambda_{AB} = \frac{A \cap B \times \Sigma_{AB}}{\Sigma A \times \Sigma B} \]
where $A \cap B$ is the number of ‘meetings’ (intersections) of events $A$ and $B$, $\Sigma A$ and $\Sigma B$ are the summary frequencies of these events in the given collection and $\Sigma_{tab}$ is the sum total of all numerical data in the table through all of its rows and columns. (For more discussion about calculations of Estonian data, see Krikmann 1994.)

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

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

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

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

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

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

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

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

Or, in the numerical expression:

<table>
<thead>
<tr>
<th>Group</th>
<th>Height</th>
<th>Labiality</th>
<th>Gravity</th>
<th>$V_1 \times V_2$</th>
<th>$\lambda$</th>
<th>$\lambda$ average</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>$X$</td>
<td></td>
<td></td>
<td>$\ddot{o}$ $u$</td>
<td>2.85</td>
<td>2.72</td>
</tr>
<tr>
<td>(2)</td>
<td></td>
<td>$X$</td>
<td></td>
<td>$e$ $\ddot{o}$</td>
<td>1.18</td>
<td>0.96</td>
</tr>
<tr>
<td>(3)</td>
<td>$X$</td>
<td>$X$</td>
<td></td>
<td>$i$ $\ddot{i}$</td>
<td>1.15</td>
<td>0.89</td>
</tr>
<tr>
<td>(4)</td>
<td></td>
<td></td>
<td>$X$</td>
<td>$u$ $\ddot{o}$</td>
<td>0.90</td>
<td>0.79</td>
</tr>
<tr>
<td>(5)</td>
<td>$X$</td>
<td></td>
<td>$X$</td>
<td>$e$ $\ddot{o}$</td>
<td>0.92</td>
<td>0.74</td>
</tr>
<tr>
<td>(6)</td>
<td>$X$</td>
<td>$X$</td>
<td></td>
<td>$i$ $\ddot{i}$</td>
<td>0.97</td>
<td>0.67</td>
</tr>
<tr>
<td>(7)</td>
<td>$X$</td>
<td>$X$</td>
<td>$X$</td>
<td>$u$ $\ddot{u}$</td>
<td>0.69</td>
<td>0.56</td>
</tr>
</tbody>
</table>

As can be seen on the graph below, the low $a$ and $\ddot{a}$ 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 $\ddot{a}$ seems to prefer illabial partners.
preliminary results allows us to see that preferences for pairing different vowels in Finnic alliteration neatly correlate with vowels’ phonetic parameters, as well as with other types of phonetic relationships between Finnic sounds, for example diachronic changes, synchronic differences between related languages and different dialects of one and the same language. Researchers have now around 185 megabytes of digitised Finnish and Karelian alliterative folksongs at their disposal, and my initial hypotheses above need to be tested on numerous larger samples. The results of the eventual forthcoming tests will probably be affected by differences in vowel systems in particular languages and dialects, by stylistic preferences of particular singers, and by the time period of recordings. But I would venture to predict that however large the empirical data we gather and however refined the calculation methods we apply, the basic patterns of preferences described above will persist.

ARVO KRIMANN is an Estonian folklorist living in Tartu and a member of the Estonian Academy of Science.

Sources and references

Leino, Pentti, 1970. Strukturaalinen alkusointu Suomessa. Folklore repellajainen tilastoanalyysi (Helsinki, SKS)
Pawłowski, Zbigniew, 1967. Vvedenie v matematicheskuyu statistiku (Moskva, Statistika)
Sadeniemi, Matti, 1951. Die Metrik des Kalevala-Verses, FF Communications 139 (Helsinki, Suomalainen Tiedekatemia)
Suomen kansan vanhat runot database, <http://dbgw.finlit.fi/skvr>