Author(s):	Eerola, Jari
Title:	Examination of stylistic traits in sound production of the Veps lühüd pajo songs using computer-aided music analysis
Main work:	Perspectives on the song of the indigenous peoples of northern Eurasia : performance, genres, musical syntax, sound
Editor(s):	Niemi, Jarkko
Year:	2009
Pages:	160-197
ISBN:	978-951-44-8015-7
Publisher:	Tampere University Press
Discipline:	Theatre, dance, music, other performing arts
Item Type:	Article in Compiled Work
Language:	en
URN:	URN:NBN:fi:uta-201406021552

Examination of stylistic traits in sound production of the Veps *lühüd pajo* songs using computer-aided music analysis

JARI EEROLA
(University of Tampere)

Introduction

Song and speech are based on man's ability to use his voice according to specific, learned cultural models. For many reasons, studying various models of sound production is not an easy task. First, every human being has a unique, personal voice quality, due to individual physiological properties of sound production. In a way, nature has thus provided for the identifiability of individual voices. Furthermore, it is often possible to identify sex and age group by the voice. Men tend to have lower voices than women, due among other things to men's longer vocal folds. Childrens' voices are often even twice as high as men's. Second, every individual uses his (particularly speaking) voice in a different way, even in the same culture. There are likewise differences in the fluctuation of pitches, in the loudness, in the use of different vocal registers, in speech articulation and so on. Yet people living in the same culture can understand each others' speech. This peculiar phenomenon is largely based on the emphases of certain partials – resonating pitches – which are also called *formants* (see e. g. Sundberg 1987, 1–3; 19–20; Laukkanen & Leino 1999, 52). The identifiability of different musical instrument sounds is also based on the variation of the formants in their

sound spectra, in other words, on the emphases of specific different vibrational pitches (Dodge & Jerse 1997, 52).

In the human voice, the emphases of the partial tones reveal how we have learned to use our speech organs and vocal tract for sound production. Because speech and song are based on learned cultural models, it is possible to learn to use the voice according to a certain style, in an entirely new way. It is also possible to learn new languages, singing techniques or styles or to imitate other individuals. In this connection, the concept of style describes the totality of various models of sound production. Style is an important factor, especially in the study of songs. It is possible for the members of a culture to recognize their own songs on the basis of a certain singing style. It is likewise feasible for the representatives of the same culture to know when a song, for example, is *not* performed according to the stylistic conventions (Lomax 1968, 12).

Although speech and song are related to each other, they tend to differ to different extents depending often on the genre of a speech or a song. It is possible to say that in the song, the organs of sound production function as the instrument of the singer. Consequently, it becomes important to have the ability to match one's (vocal) instrument to the musical style in question. In general, it can be said that the most important difference between song and speech is the control of the pitch level and the loudness, as well as the production of a certain sound colour. These often require the control of the air below the vocal folds (= the subglottal pressure). This is one of the reasons why singing more often requires practice than speaking. (Sundberg 1987, 48.) In general, singing can be thought of as a more specialized use of the voice than speaking. It is often also more formally organized, including its own use of speech – vocabulary, special effects etc. (Lomax 1968, 3).

The sound production of both singing and speaking has long been studied and with similar methods. The most important of all is the hearing analysis, due to the irreplaceability of the human ear. However, the presentation of such analyses becomes easier if the results can be represented in a visual form. There are plenty of computer programs

available today, with the aid of which it is possible to produce different visual representations. Among other things, visual comparison of these representations may be of help in identifying similarities in different kinds of sound production. Also, conducting hearing analysis simultaneously with graphs may reveal details possibly left unnoticed in the hearing analysis.

In this article, I shall examine the singing style of the Veps song genre called *lühüd pajo* (Veps., 'short song'). First of all, I shall compare two different *lühüd pajo* styles and the characteristics of their sound production and use of sound (the latter with reference to the relationship with the melody). My analyses are based especially on computer-aided methods and thus result in graphs. It is my purpose to identify the general stylistic characteristics of the *lühüd pajo* revealed in this kind of analysis. The object of the research is music in oral tradition and especially song. In my research, the musical analysis of the songs is divided into three parts: melodic structures, metrical structures and sound production. In the present article, I shall concentrate only on the problems of sound production.

My purpose is to describe and analyse the stylistical characteristic of the *pajo* songs of the Veps, an ancient group of Baltic Finns living in northern Russia. My research began on a comparative song style project of Eurasian peoples (EULA) in 2001. It is largely based on materials I collected during my field work in Veps villages (six trips 2000–2003). The tradition I am studying is a still living one. My fieldwork trips have oriented towards the regions of central and southern Veps living in Leningrad and Vologda regions in Russia. On the other hand, I rely on historical recordings I have had the chance to use. These materials are from the archives of the Research Institute for the Languages of Finland (KOTUS), the Finnish Literary Society (SKS), The Linguistic Institute of the University of Petrozavodsk and the Karelian Radio (Petrozavodsk, Republic of Karelia, Russia) and the Universities of Joensuu and Oulu (Finland). In addition, I have at my disposal field recordings by other reseachers visiting Veps villages (Markku Nieminen, Juminkeko Foundation, Kuhmo, Finland) and Viktor Lapin (Russian Institute of Art History, St. Petersburg).

On the concepts of style and singing style in studies past and present

As a concept, style has many meanings and is used in several different contexts, as in literature, art and music. The factor common to all styles is that they are products of a specific culture or social group. By its content, a style represents regularities, which are norms approved and adopted by a certain social group (a linguistically defined social group, for example). The emergence and progress of such norms are largely based on choices made by a social group or individual. The eventual style begins to form on the basis of the environment in which the language is used, in other words with the intentions and the function of the linguistic expression of the speaker or writer. The relationship between performer[1] and audience has an effect on the perspective that the performer of the message chooses. There are numerous options for this. The eventual expressive option merges both an internal, individual perspective and an external perspective of appropriateness. Style can be said to be a sum of or a compromise between these aspects. (Saukkonen 1984, 9–15.)

It is not an easy task to identify clear-cut stylistic genres from a spoken linguistic expression (Saukkonen 1984, 26). What is the case, then, with musical styles? Is it possible to define singing styles of specific cultures or societies? Why is it necessary to define styles in the first place? Especially in the culture tradition of Western Europe there has long been a need to define different musical styles – both familiar and exotic. I feel that analysing styles is a means to understand the culture studied. Thus it can be argued that the styles of singing or playing should be easier to define than speech styles, since music is a more specific product of culture. There is, however, a possibility of an intellectual illusion in this kind of thinking and there are numerous examples of this in the history of ethnomusicology, when various means have been adopted to define musical styles.

1. Saukkonen uses here the term "speaker", but it can be replaced by "singer", for example, because this role essentially concerns performance.

One of the orientative studies on singing style was the so-called Cantometric model of analysis by Alan Lomax (1968). It was also defined as a tool for comparing musics (Nettl 1983, 92). The purpose of this study was to understand the singing styles of different cultures. The idea of Cantometric analysis was to create a fast and coherent method for the definition of various singing styles, which was feasible in co-operation with different researchers. To accomplish this task, the method included comparative materials, on the basis of which it became possible for different researchers to formulate coherent parameters for the singing styles studied. The whole system consisted of 37 parameters, designed for defining singing styles. (Lomax 1968, 3–37.) The cornerstone of the project, as was later noted, was the Cantometric system and its parameters (Saha 1996, 56; Nettl 1983, 94), which came in for harsh criticism. The criticism probably originated in the active discussion in ethnomusicology about the fundamental problems of intercultural comparison of musical styles. However, I feel that comparative research can be very illustrative, although it is often thought to be loaded with value judgments. Thus, the aim of this article is a non-normative description of sound production in different styles of singing.

In ethnomusicological studies, the concept of style has been contextualized through periodical, generic, individual and geographical dimensions (Saha 1996, 39; 75). However, these are highly abstract definitions and they do not necessarily tell much about the style itself. Although the definitions of singing style by Lomax (1968) are on a very general level – as also noted by Saha (1999, 56) – there are some interesting and converging notions about style in the Cantometric method, also for the present study. The Lomax research group (1968, 34) discovered that in each culture there are certain favoured ways of controlling paramusical means of expression (i. e. sound qualities of the singing voice) and that they may be comparable with similar phenomena in other cultures. In addition, in the study of musical styles it is noteworthy that very few people can master several musical styles. This accords well with the study of singing styles. Usually it is left for the researcher to identify these paramusical means of expression.

Points of departure in the description of style

The problem in the description of singing styles of cultures other than that of the researcher is that it is not always so easy to understand which characteristics are relevant for a song to belong to a certain style. The traditional method of stylistic description in comparative musicology was to analyse only musical scores. The music belonging to an oral tradition, like Veps song, is not necessarily involved with musical notation. On the other hand, if it were so, even then the score alone would not be sufficient to express the core of the performance style. For this reason, the definition of musical styles cannot be based only on musical scores, if the goal is an all-inclusive stylistic description. Furthermore, the music of oral tradition often contains very delicate expressive nuances, which are difficult, at times impossible to describe in a musical score. And in addition, the musical score is a reflection of its transcriber's own experience, education etc. (Jouste & Niemi 2003, 170–171; 188–189.) In the music analysis of Western musicology, the role of a music transcription is, nevertheless, important, in spite of its limits as a means of analysis (cf. e. g. Salavuo et al. 2003, 330). However, the musical score continues to be an indispensable part of the description, comparison and analysis of different kinds of music. On the other hand, it is easier to understand the musical score if it is made from a musical style familiar to the transcriber.

At a general level, it can be said that style is all that is perceptible – the rest is abstraction (Saukkonen 1984, 91). According to this thought, a song as a concept is an abstraction and the way it is produced represents the style. What, then, is the perceptible part of a song? As a song is largely a phenomenon of sound, its perceptibility centres on audibility. This audible part of the song represents the surface level and style. Quite often this surface level allows variation and this is more likely if the song tradition is governed by loose social stylistic norms. Consequently, these norms condition individual styles (Saukkonen 1984, 10). Studying a song style involves observing this variation. However, there may be also visible domains involved in a singing style,

like those reflected in singing technique, i. e. voice production and the physical context of a performance. Usually the whole body participates in voice production. For this reason, other than audible sense domains may be of additional help in the analysis of voice production.

I shall concentrate on the properties of voice production in the description of a singing style, in other words on vocal sound and timbre. It is of great relevance from the perspective of a singing style how the song is performed. The members of a culture will recognize immediately if a change in the vocal quality leads to a violation of stylistical norms. On the other hand, a song sung with a different sound quality may be associated with another stylistic genre (Mantere 2002, 7; Lomax 1968, 12). This way, a song having the same melody and content may be used in different performance contexts. Often the purpose and performance context are connected to the voice production. In these cases the different emotional states are substantial, being attached to the performance situations of the songs and having their effect on the use of the voice (Sundberg 1987, 146–156).

My own approach is based, in a way, on the functional examination of style. So-called functional stylistics examines alternative modes of expression of content in varying communication situations, where important factors are perspective, contextual properties associated with content and the situational system (e. g. Saukkonen 1984, 111; 163). In the present article I shall not concentrate on content so much as on the situational system of the songs.

In the examination of style it is also relevant to consider the historical factors in the formation of a style, if only possible. The point of departure in the examination is that singing and speaking are learned behaviour and they conform to certain patterns (see e. g. Laukkanen & Leino 1999, 55). Seen in a historical perspective it is conceivable that in the course of time certain local styles or models have developed in a society. However, these styles are not stable and eventually they begin to evolve into new forms. A new style appears when members of a social group begin to imitate the innovative behaviour of an individual (see Saha 1996, 81).

Quality and timbre of the voice and various graph forms

In different fields of science there are different concepts in use for the description of vocal characteristics. Sometimes the meanings of these concepts are almost identical or they have a similar semantic flavour. There have been many studies of singing, for example, in the fields of speech studies and in medicine. In several publications of these disciplines the concept of *tone quality* has been used to describe voice qualities. As such, "quality" also has connotations of an ideal, desirable voice. True, in speech studies the aim is hygienic voice production, which should be minimally strenuous and economical. Here the concept of voice quality is associated with the problems of description of voice: pressed, nasal, raspy etc. The aim in these studies is to look for reasons for the voice to have different qualities. Consequently these concepts have a normative, value-oriented meaning. In music studies and in singing pedagogy the quality of the voice refers to the nature of the voice, which can be seen as more connected with performance interpretation. Thus, a voice can be described as "light" or "dramatic". Nevertheless, the ultimate aim also here is to give a verbal description of how a voice quality is thought to be produced (Slawson 1985, 19; Laukkanen & Leino 1999, 56; Hemsley 1998, 59–66).

It can be said that all three properties of voice – pitch, loudness and register – are included in the concept of voice quality. Voice quality is, in essence, an auditory impression, which is created by the amplitudinal relationship of the partial tones of a voice. In turn, these relationships vary according to pitch, loudness, register and relative degree of adduction[2]. With the concept of voice quality is associated the concept sound colour or timbre, which is affected by the voice quality, as well as resonance (see e. g. Laukkanen & Leino 1999, 202–206). Voice timbre is an abstract concept and refers more to the general "sound"

2. Degree of adduction refers to tension of the adductor muscles pulling the vocal folds towards each other. Consequently, the abductor muscles pull the vocal folds from each other, thus opening the glottis (Laukkanen & Leino 1999, 37).

of a voice. The sound of a voice may, for example, be croony, metallic, "rear one" etc. (Laukkanen & Leino 1999, 56).

I shall examine the quality of voice with different graphic representations and with auditory impression (timbre). Two kinds of graphs are used. The first is a *long term average spectrum* (LTAS) and the second is a *spectrogram*. Both types of graphs have been used in speech and song analysis. The purpose of the LTAS graph is to present some general features of voice production or voice quality present in singers' performances. When reading the average spectrum, attention is turned towards *formants*, i. e. emphases, which are connected with amplitudinal relationships in different pitch sections. A crucial factor in the LTAS graph is the general shape of the graph, instead of the single "peaks".

LTAS analysis yields information about average pitches or resonance, with which it is possible to discern differences in voice qualities. The first analyses with LTAS graphs were made by Jansson and Sundberg (1974) and by Leino (1975). LTAS graphs have been used to study, for example, the voice production of opera singers, professional actors and orators.

How appropriate is the average spectrum for the analysis of song? One of the problems is that in traditional singing the basic pitch level tends to fluctuate all the time. Consequently, when this happens, the complexes of partial tones, that is the resonating pitch complexes also tend to change. An example associated with speech will clarify this. In speech, too, the basic pitch level tends to vary, according to speaker, style, situation etc. Usually the alteration ranges approximately an interval of a fifth below or above the basic pitch level. However, a speaker's individual voice sound remains almost the same and it is thus possible for the listener to identify the speaker from the voice. There is a substantial amount of sound energy in speech at the lower pitch regions and less at higher pitches. The intensity of the decrease in sound energy is meaningful for the sound quality (Laukkanen & Leino 1999, 170–171). Hence, single peaks in the LTAS graph are not themselves as meaningful as the overall shape of the graph and the energy concentrations reflected in it.

A crucial factor in the examination of voice quality is the phenomenon of resonance[3]. It is the very resonance on which the individual sound timbre of human voice and musical instruments depend. Resonance is co-vibration. If, for example, the sound cavity does not receive sound corresponding to the specific pitch of the cavity, there will be no resonance. In the context of singing, this means that singing a specific pitch in a specific vocal style, there will be no resonance. Consequently, using another singing style or technique the vibrating air can be channelled to such a location within singer's body where the resonance can be realized. This kind of case is associated with the use of different vocal registers. It is noteworthy that it often happens to singers that they cannot raise the pitch level of their singing above a certain limit without having to change their vocal register (Laukkanen & Leino 1999, 44).

Singers can also take advantage of resonance. If a singer is able to utilize the specific pitch levels of his sound cavity, he will be able to produce a loud sound with less energy, because of the resonances strengthening the basic sound (cf. opera singing)[4]. This is also heard in the voice timbre or sound. This phenomenon is taken into account in the structures of acoustic musical instruments. Consequently, those partial tones that are not supported by resonance are bound to weaken. This is the reason why there are always peaks and valleys in the spectrum (ibid. 74–75). Using a specific technique, the sound cavity also assumes a specific form, which amplifies only those partial tones which are supported by the phenomenon of resonance. These partial tones and sonoric emphases are presented in the LTAS graph.

3. If a singer does not produce pitches that do not converge to his sound cavity, resonating pitches are not realized. This also happens when pronouncing unvoiced consonants (e. g. obstruents), because the glottis is open. Consequently, the sound cavity does not receive any air, which would be vibrating because of the glottal activity (Laukkanen & Leino 1999, 38; 87) and no resonance is produced. This is why it is usual in the analysis of singing and speech to concentrate on vowels or voiced consonants.
4. The resonance of the voice is utilized in many ways. The fundamental method is to "locate" sound in different parts of the body or to change the vocal register. This makes the voice sound or resonate, for example, in the head or chest.

About earlier studies

During the last 30 years there has been increasing interest in the acoustic research of singing. Primarily the research object has been singing in the European classical tradition, but lately other genres of vocal styles have been objects of sound research, like, for example, pop music (Fig. 1), jazz, blues (see e. g. Thalén & Sundberg 2001) and country (see Cleveland et al. 2001). Quite often these studies have adopted methods of speech analysis, using average spectra and spectrograms. The same methods have also been applied to studies in the vocal production of traditional singing (see Ross 1992; Lindestad et al. 2001; Mantere 2002; Kovačić et al. 2003).

In these studies it has been observed that there are certain properties in common in the graphs taken from the voices of professional vocalists, which are particularly associated with formants. Formants can be defined as strengthened zones of partial tones (Suomi 1990, 98) or realized resonating pitch levels of the sound cavity (Laukkanen & Leino 1999, 76). In Sundberg's (1980; 1987) studies there emerged the so-called singer's formant, which distinguishes itself as a strong emphasis in the average spectrum at about 3 kHz (Sundberg 1980, 86–89; 1987, 118–119). A so-called actor's formant has also been identified to approximate pitch level of 3.5 kHz (Laukkanen & Leino 1999, 171–175).

There are slight problems with the usage of the concept of formant. It is often used inaccurately also to signify the single peaks of the spectrum of the sound produced (Laukkanen & Leino 1999, 76). It is possible that this kind of usage can be explained by the fact that the studies in question belong in the discipline of voice research, which, in turn, is influenced by linguistic studies. In point of fact such single peaks represent rather a kind of emphasis in the spectrum than a formant. However, the concept is established in this context. (Sundberg 2004.) In phonetics the formants are presented beginning from the lowest one (F1, F2, F3 etc.), while the basic pitch is marked as F0 (Suomi 1990, 98).

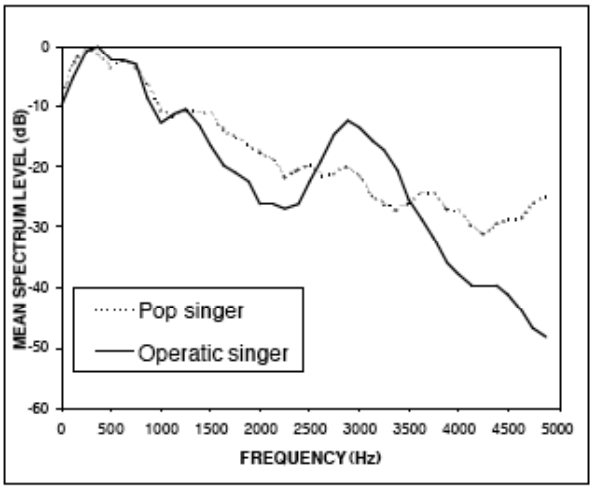


Fig. 1. LTAS of a pop singer and an operatic tenor who sang the same excerpt in the same key (Borch & Sundberg 2002, 33).

It is worth mentioning that all these emphases – formants – have their reason. First, they have appeared as a result of training and learning. Second, the training has had a distinct aim: to result in a voice production which is best suitable in a specific context of voice usage.

However, a substantially less studied area is the voice production of non-professional singers, especially folk singers. Studying traditional singing poses problems in the sense that the variety of different styles, substyles, cultures, subcultures etc. is enormous. On the other hand, it seems that the research of singing styles is still primarily associated with traditional methods of analysis, based on notated music materials, where the main interest is attached to analysing single tones. However, many singing styles are characterized by a specific overall timbre, which is usually left largely unnoticed in music analysis.

The quality and selection of the source material

The songs examined in the present article are from field materials recorded by the present author. The selection of these specific songs is a result of many things. First, analysing songs entails certain requirements of quality from the material and these are not always met in my materials. This has been one reason for selecting these specific songs. For example, it is often not possible to choose the best possible environment for recordings in the field. The overall timbre of the recordings is also influenced by the presence of other people in the recording situation and for this reason many recordings had to be excluded from the computer-aided analysis, since the computer cannot distinguish the analysed voice from others present. For this reason I made a preliminary qualitative analysis of the material and after that made the selection of materials best suited for the analysis. Thus, the qualitative prerequisites effect the selection of the material analysed. For example, this is the reason why the analysed material consists of unaccompanied singing, because the computer cannot distinguish the human voice from the accompanying instrument. On the other hand, it is possible to apply auditory analysis to the excluded material and compare the results to those made by computer analysis. This comparison is important throughout.

I started to study Veps songs from the perspective of individual performance. This is due to the prominent performance conventions: the songs are mostly sung alone. This is also one of the factors defining a singing style. On the other hand, even if the songs were sung ensemble, it would not be possible to analyse the voice production by examining the vocal sound of the group, at least this kind of analysis would be not very reliable. Yet the results of analyses of solo performances may be biased due to various psychological factors in the performance situation and they must be also considered in the final conclusions. First, some singers are inspired by the presence of other people: it is often the case that it is easier to sing with friends present. There were occasions in southern Veps when performers insisted on having their friends present during the singing.

When our fieldwork team visited the village of Haragl, we met a singer who agreed to sing in a private home during a village festivity. We asked her to sing the songs again at her home, in a more peaceful atmosphere suitable for successful recording. However, she had lost her mood for singing and did not want to perform her songs again. The presence of an audience may well have the effect that the performer needs to show she is able to sing well. Consequently, those criteria that are valued in a certain style are more clearly present, as for example singing with a great volume. From the standpoint of the quality of recording, the best situation is without extra people, but this may change the voice production of the performer substantially.

An important question in the selection of song data for analysis is whether the songs studied are supposed to be sung alone or in a group. The songs analysed here are usually performed alone, but I have heard them performed in a group as well. Studying unaccompanied soloistic singing I also have tried to take into account psychological factors effecting voice production. For example, if a song is usually performed *with* accompaniment, it may sound different *without* accompaniment. It seems to be likely for the singers performing in a conventional situation with accompanying instruments that their voice production is very pressed and the voice quality could be defined as shouting. The reason for this is often that in such performance situations the singers have to compete with the loudness of the accompanying instrument. When the singer performs the same song without an accompaniment, there is usually a change in the voice production. This may possibly be due to a learned, conventional and stylistically conforming way to perform certain kinds of songs. The accompanying instrument is usually an accordion or its Russian version, the *tal'yanka*, with its very loud sound, with which it is easy to cover the human voice.

Another notable factor in the field recordings is the recording environment, which usually cannot be optimised in all field work situations. Different indoor acoustics and the presence of other people have an affect on the overall performance sound and thus on sound production. These factors have psychological effects on the singers as well.

In my field recordings further problems are posed by the fact that I was not always able to control the distances of the microphones. It would have been an ideal situation if all the singers had sung from a standard distance (approx. 40 cm) into a standard microphone. For this reason, too, my field materials will not yield a fully reliable result, at least concerning the fluctuations in voice pressure.

I have chosen the sample songs for the analysis on the basis of subjective hearing analysis, which conforms to the criteria presented above: an unaccompanied performance, a recording of good quality, song performance in a conventional, traditional style. In addition, I wanted to select the sample so that it would contain samples from different Veps regions.

All the songs are sung by women. Their age range is between 60 and 80 years. This has an effect on the overall sound of the performances and it has to be taken into account in the final conclusion of the analysis results. On the other hand, the fairly homogenous age group of the performers may give an advantage in identifying the general characteristics of performance style.

Veps *lühüd pajo* songs

According to Rüütel (1990) and Salve (1998, 127), the Veps define their songs either as *pajo* 'songs proper' or as *voik* 'lament'. In my experience the Veps mostly define their songs as *lühüd* ('short') or *pitkä* ('long') *pajo*[5]. Salve adds that in modern Veps *pajo* means, in fact, *chastushka*[6]. In my field experience this has not always been entirely clear. In interviews with the singers I have encountered different definitions of the style genres and often these definitions have been

5. Note the etymological connection with the Veps *pajo* (pronounced *páyo*) 'song' to Russian *pet'* 'to sing'; *poyot* (pronounced *payót*) 'sings S3'.
6. *Chastushka* 'a small piece' (Russ.) improvised song form in Russian folk music (Editor's note).

confusing for me because of the semantic ambiguity and parallel use of the terms *lühüd pajo* and *chastushka*, especially when in reality the *lühüd pajo* is strongly reminiscent of the originally Russian *chastushka* in its auditive structure.

There is much variation in Veps *lühüd pajo* songs, because these songs are an oral tradition. This variation is not always stylistically intentional. A very important factor defining the style is the emotional content of the songs, which tends to change according to the performance situation. Because of the loose stylistic norms the songs are multifunctional and can be performed in various situations. The *lühüd pajo* songs are performed as children's lullabies, work songs or during festivities. The songs often have the same texts, but they are performed in different singing style, rhythm, tempo and volume. According to Sundberg (1987, 152), in a song, tempo and volume are controlled by emotion. Fear makes tempo slow down and voice softer. Anger raises tempo and volume.

I gained new insight with the question of style when I visited Mariya Semënovna Trishkina (b. 1927) with researchers from Petrozavodsk in Päzhjärvi, Vologda region in July 2003. She sang us a *lühüd pajo* which was used as a song for driving away bears. We discussed the song genre and asked her whether she would label the song as a *pitkä* or *lühüd pajo*. She replied it was a *lühüd pajo*. In addition, I asked whether it could have been called a *chastushka*. Trishkina, and also her son, joining in the discussion, did not agree with me, replying it was definitely a *lühüd pajo*, because of the length and tempo of the song. Thus these structural elements marked it as different from *chastushka*. Trishkina also identified the Veps language of the song particularly with the genre of *lühüd pajo*, but not *chastushka*. She also said that this song was performed during work (see Eerola 2003b, 104–105).

Improvisation is also often mentioned as a stylistic characteristic of the Veps *pajo* (Hakamies 1994, 80–81). In my opinion, the concept of improvisation is possibly not the point here, in which I agree with Saha (1996, 75), who speaks about improvisation in terms of idiosyncratic variation in the surface level of the musical structure.

For example, Makar'ev (1931, 33) has a description of the song performances of Veps men and women. According to him, during the *beseda* gatherings[7] usually "the boys and girls danced in silence, because they considered dancing with the accompaniment of singing as degrading". I have heard my informants telling, though, that men also used to sing in the old times. Some of these informants (women) even remembered the songs the men used to sing.

Still in the 1960s and in 1970s, singing was a part of everyday life in Veps villages. Songs were sung in fields during work, in the village street, at weddings and at various festivities. Nowadays it is only the oldest generation who can perform songs in Veps, especially the women. Most of my informants were able to sing *lühüd pajo* songs. However, very few could remember *pitkä pajo* songs. This is one reason why I have concentrated here on *lühüd pajo*. They are quite multifunctional songs and this property may in fact have been a crucial factor in their survival. Usually the singers have performed songs both in Russian and in Veps. Sometimes they can change the language even during the same song, sometimes the singers have intentionally sung the Veps couplets first and the Russian ones at the end. Sometimes Veps and Russian merge in a song into a mixed language, although this possibility often is a reflection of performers' difficulties in remembering Veps texts.

The textual contents of the songs are mostly associated with a woman's life. The song texts may tell about how a girl (usually the singer herself) was married to an unknown groom in an unknown family. Very often the textual themes touch upon humorous teasing songs in a dialogue between girls and boys. As an example, these songs may boast about the boys of the singers' own village and compare them with the "inferior" boys of the neighbouring village. In some songs there are often maxims for a young girl or suggestions about the arriving groom.

Maybe the most known *lühüd pajo* (Ex. 1.) tells about a boy and a girl. It is spring time because the cuckoo sings in a tree. The girl (*nietšukaine*) cries because her boyfriend or husband has to leave the home village. It is very ordinary among Veps that men have to leave the

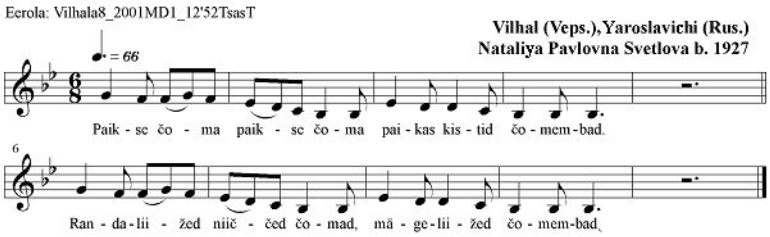
7. *Beseda* 'conversation' (Russ.) – an informal or festive gathering of people of the neighbourhood in traditional Russian village life. (Editor's note.)

villages and go for work in bigger towns. There may be other reasons too why the girl is crying and one is that the boy is leaving to join the army. So the song can be sung in different situations too. There are many variations of this song and here is one which was collected by Lauri Kettunen in 1935.

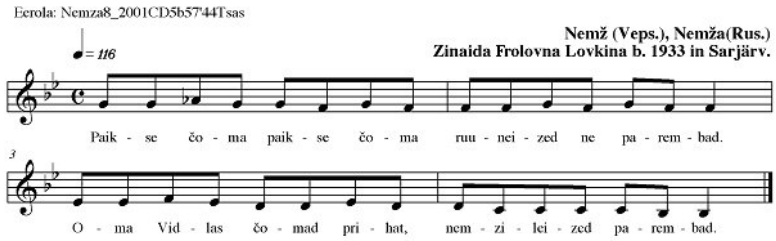
Kukku kukku kägoihut,	*Cuckoo, cuckoo little cuckoo,*
sures kuze ladvaižes,	*at the top of a tall spruce tree,*
voika voika neitšukaine,	*cry, cry little girl,*
tšomal prihal kaglaižes.	*on the neck of a handsome boy.*

Ex. 1. Song text example of a *lühüd pajo*. Village of Noidal/Noitala. Kettunen & Siro 1935, 140.

By their melodies, *lühüd pajo* songs divide into two groups: those of faster ones and those of slower, which are usually work songs. In songs of the faster type, the tempo varies (♩. = 70–140) and the meter is usually 4/4. In the work song type, the tempo is slower (♩. = 50–80) and the meter usually 3/4 or occasionally 6/8. See the following examples of these stylistic variants of *lühüd pajo* songs (Ex. 2. and Ex. 3.):



Ex. 2. Work song.



Ex. 3. Faster *lühud pajo.*

Prerequisites of the analysis

On the basis of the auditory analysis only it is possible to say that the songs of the faster *lühüd pajo* and work song style differ from each other. The voice production is also clearly different. In faster style the singing is much louder than in work song style, where the voice production is reminiscent of that of speech.

The auditory analysis is an obligatory initial phase for the computer analysis. However, it is not possible to present more detailed analytic results only by hearing, but on the other hand, the computer-aided analysis supports and possibly helps to explain the results of the hearing analysis. It is possible that the analyst's interpretative framework changes after seeing graphs from the computer analysis (see e. g. Mantere 2002, 75) or examining them simultaneously with hearing. For example, a small and subtle vibrato may be perceptible only in the graphs made with the computer.

With the analysis I shall search for explanations for differences in the voice production in these two singing styles. How do the singers arrive at the tones or how do they perform them? What effects are used in the songs? I have defined no sound features in advance to be analysed from the material. Instead I shall examine the voice production of the songs directly with the LTAS graphs (see e. g. Kovačić et al. 2003, 5).

The range for the lengths of the songs is from 30 seconds to two minutes. Making an LTAS graph requires a coherent passage of recording for at least one minute in order to disperse the effect of single vocal sounds (Laukkanen & Leino 1999, 170). For this, some examples are slightly too short. However, I believe that the short duration does not substantially affect the results of these songs. I have added an additional trend line to the graphs, showing the curves in the graphs. In these I have used settings of 600–6000 Hz. I have used the computer program called Praat, which calculates the audio signal with the method of Fast Fourier Transform.

All the graphs were produced with the program PRAAT v5.0.03, meant for speech studies and developed by Paul Boersma and David Weenink (see http://www.praat.com). In the LTAS graphs I have used a pitch band of 200 Hz. If the value had been greater, the graphs would have assumed a "rounder" form and the single peaks would have been flattened. As I tested different values I found that with values greater than 300 Hz the graph result seemed too "round". On the other hand, values below 100 Hz would have been too detailed. LTAS with a ratio of 1:1 was substantially slower to count and draw and I deemed it unsuitable in the absence of any relevant information. In speech analysis a (large scale) pitch band of 0–5000 Hz is common, because partial tones exceeding the upper limit are usually already too weak. The partial tones are usually considered to reflect singer's individual voice timbre. As I wanted to examine this also, I decided to make the (large scale) pitch band graphs with the limits of 0–8000 Hz, after several preliminary experimentations, as in many songs there were pitches and emphases in many songs that exceeded 5000 Hz.

The LTAS and spectrogram graphs of the songs

I made a preliminary grouping for the songs according to the description or definition of the informants. On this basis I selected 13 *lühüd pajo* songs. Of these, nine are faster *lühüd pajo* and four work songs. The examples were chosen to represent different Veps regions: the central and southern Veps (Leningrad region) and eastern Veps (Vologda region) (see Heikkinen & Mullova 1994, 10–12).

I chose a representative sample of the songs (Fig. 1–13) so that the energy concentrations, i. e. formants would emerge in the graphs more clearly. Namely, the formants are not necessarily seen directly from the spectra, but their locations and pitches must be estimated. The formants are more easily discernible in a spectrogram with a wide pitch range. However, the overall shape of the graph together with the points of emphasis are the most relevant features.

The LTAS graphs are on the left and the wide range spectrograms on the right.

Faster lühüd pajos:

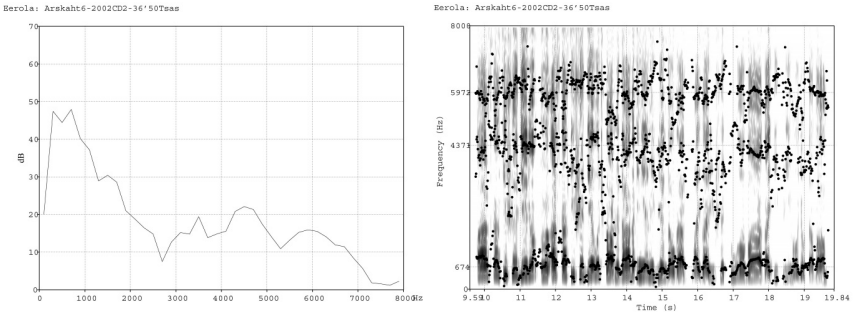


Fig. 1. Roza Nikolaevna Veselova, b. 1935, Arshkaht' (Russ. Radogoshcha). Recorded in Arshkaht' 3.6.2002 by Jari Eerola. Eerola: Arskaht6_2002CD2_36'50Tsa

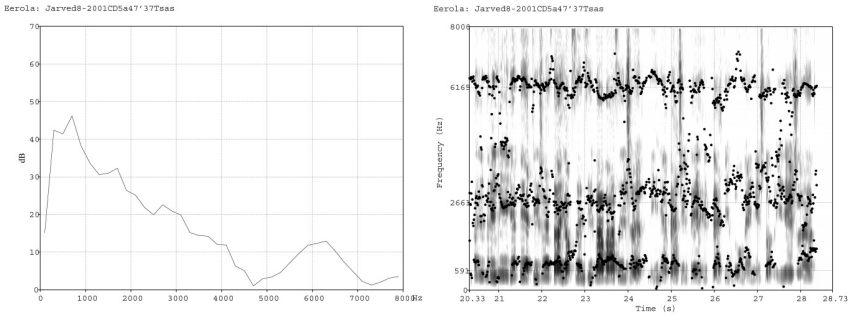


Fig. 2. Evgeniya Stepanovna Maksimova, b. 1941, Järved (Russ. Ozëra). Recorded in Järved 23.08.2001 by Jari Eerola. Eerola: Järved8_2001 CD5a47'37Tsas.

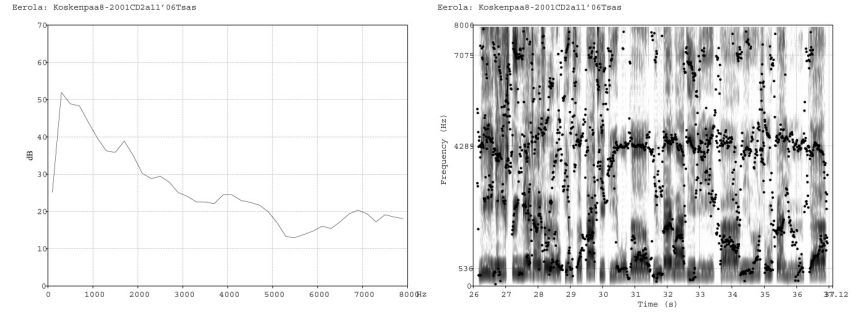


Fig. 3. Anna Aref'evna Shul'gina, b. 1941, Koskenpää (Russ. Nadporozh'e). Recorded in Koskenpää 9.8.2000 by Jari Eerola. Eerola: Koskenpää8-2000CD2a11'06Tsas.

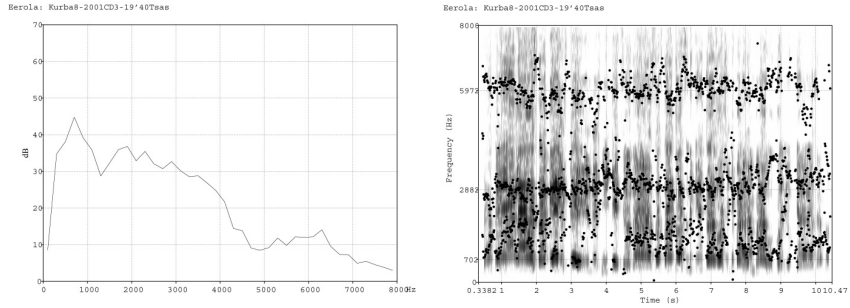


Fig. 4. Anna Ivanovna Tsaretskaya, b. 1931, Kurba (Russ. Kurba). Recorded in Kurba 22.08.2001 by Jari Eerola. Eerola: Kurba8-2001CD3-19'40Tsas.

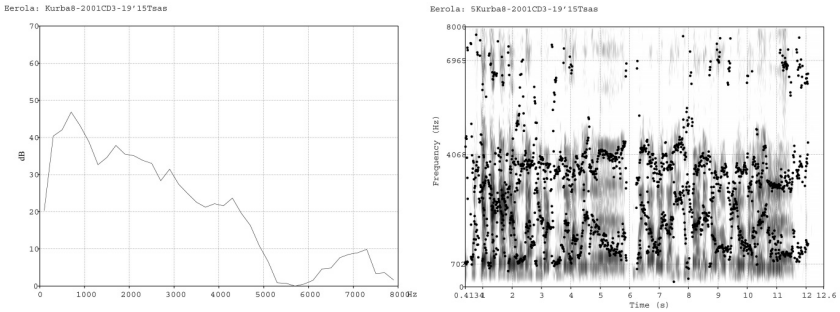


Fig. 5. Mariya Petrovna Andreeva, b. 1916, Kurba (Russ. Kurba). Recorded in Kurba 22.08.2001 by Jari Eerola. Eerola: Kurba8-2001CD3-19'15Tsas.

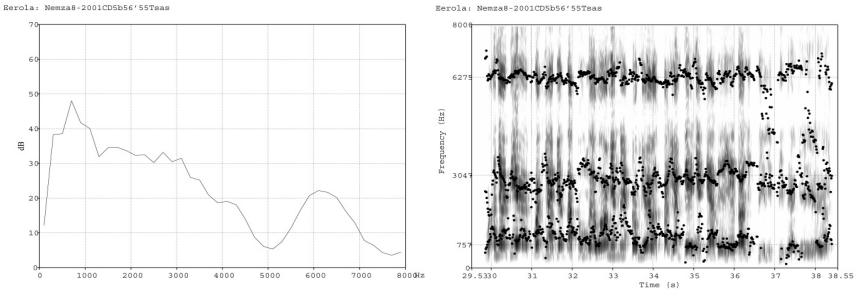


Fig. 6. Zinaida Frolovna Lovkina, b. 1933, Nemzh (Russ. Nemzha). Recorded in Nemzh 25.08.2001 by Jari Eerola. Eerola: Nemza8-2001CD5b56'55Tsas.

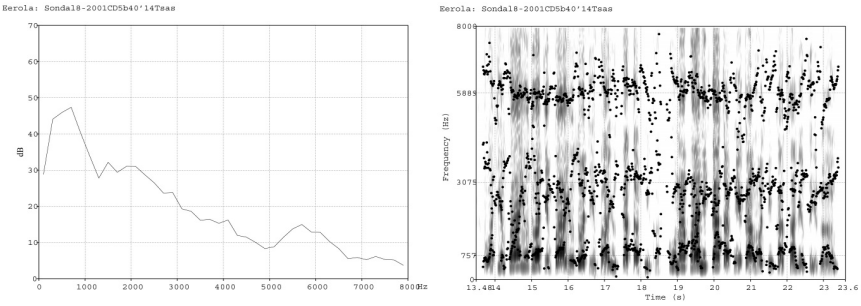


Fig. 7. Lidiya Nikolaevna Lukina, b.1934, Sondal (Russ. Shondovichi). Recorded in Sondal 25.08.2001 by Jari Eerola. Eerola: Sondal8-2001CD5b40'14Tsas.

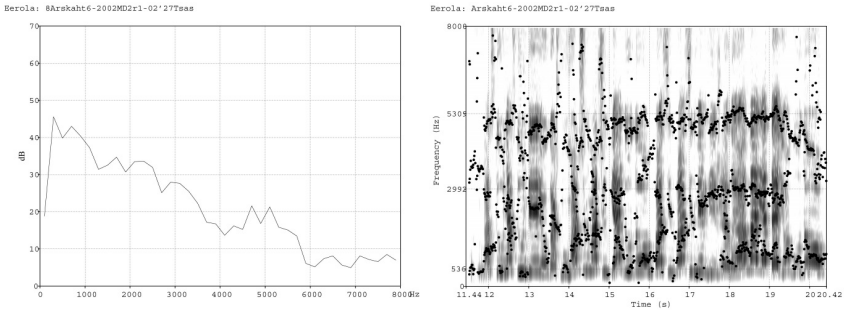


Fig. 8. Lyudmila Vasil'evna Semechkina, b. 1952, Arshkaht' (Russ. Radogoshcha). Recorded in Arshkaht' 3.6.2002 by Jari Eerola. Eerola: Arskaht6_2002MD2r1_02'27Tsas.

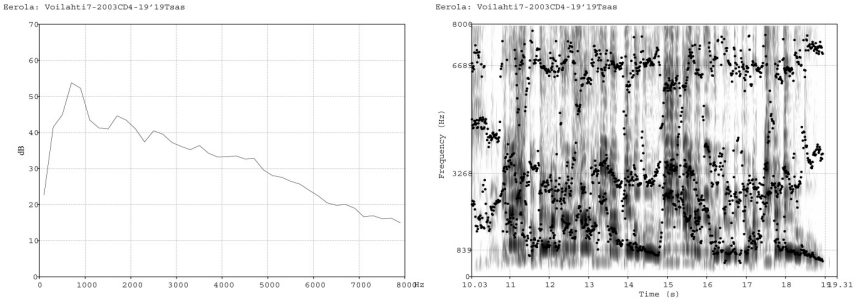


Fig. 9. Valentina Pavlovna Eraticheva, b. 1927, Voilaht (Russ. Voylakhta). Recorded in Voilaht 5.7.2003 by Jari Eerola. Eerola: Voilahti7-2003CD4-19'19Tsas.aif

Lühüd pajo songs of work song type:

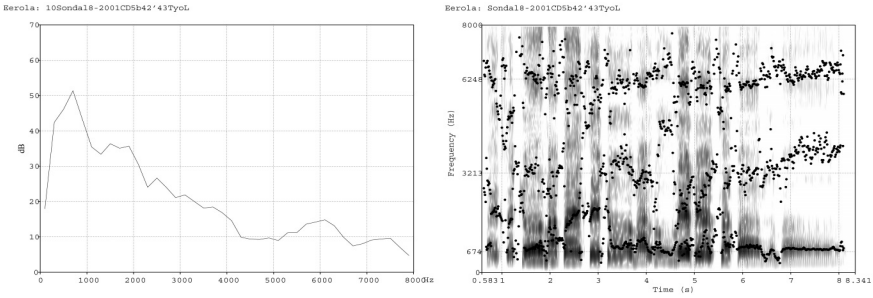


Fig. 10. Lidiya Nikolaevna Lukina, b. 1934, Sondal (Russ. Shondovichi). Recorded in Sondal 25.08.2001 by Jari Eerola. Eerola: 10Sondal8-2001CD5b42'43Työ.

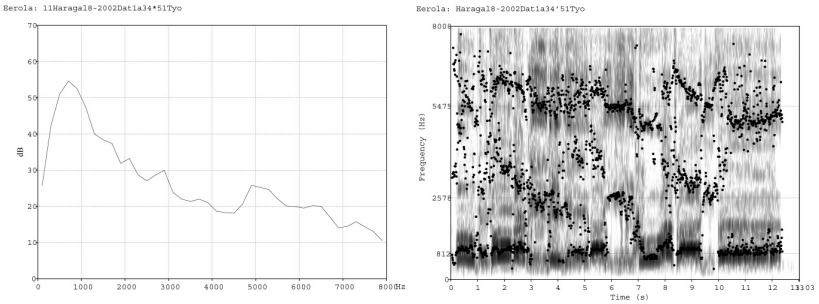


Fig. 11. Antonina Osipovna Bogdanova, b. 1932, Haragl (Russ. Kharagenichi). Recorded in Haragl 27.8.2002 by Jari Eerola. Eerola: Haragal8-2002 Dat1a34'51Tyo.

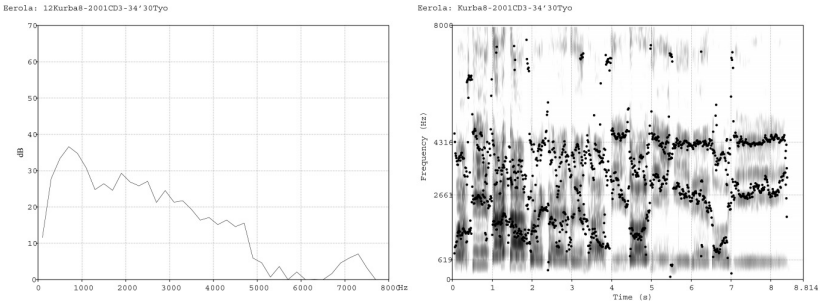


Fig. 12. Mariya Petrovna Andreeva, b. 1916, Kurba (Russ. Kurba). Recorded in Kurba 22.08.2001 by Jari Eerola. Eerola: Kurba8-2001CD3-34'30Työ.

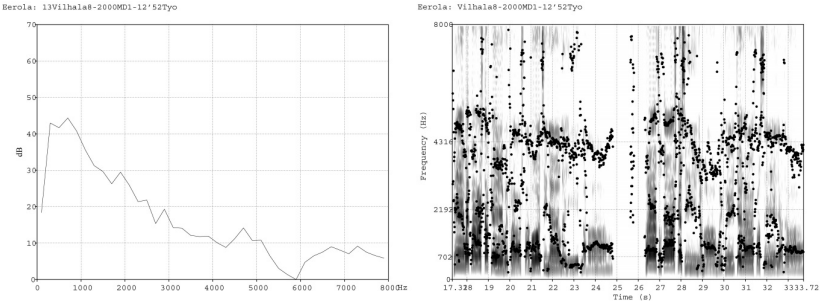


Fig. 13. Nataliya Pavlovna Svetlova, b. 1927, Vilhal (Russ. Yaroslavichi). Recorded in Vilhal 6.8.2000 by Jari Eerola. Eerola: Vilhala8-2000MD1-12'52Tyo.

In both song styles the strongest peak in the spectra is located in around 400–1000 Hz. In the faster *lühüd pajo* style, pitches following this pitch range die away more slowly and the graph is not so steep as in the work songs. In the faster *lühüd pajos* there are also a clear energy concentration at about 500 Hz to 3000–4000 Hz. This can be seen in the spectrogram as a dark region indicating that the lower partial tones gain more energy – most clearly in Fig 4, 5, 6, 8 and 9. It is possible that the stronger voice production and pressure raises the pitches of the formants (Mantere 2002, 54; see also the spectrograms).

According to the auditory analysis, there is less "metallic" and more "creaky" sound in the timbre of the faster *lühüd pajos*. These timbres can also be described as tense and pressed and are probably due to raising of the larynx. In the graphs this is reflected in the relationship of the basic pitch level (the first peak in the graph) to the following peaks.

Singing with open mouth tends to halve the distance between formants, which, in turn, increases loudness. In the faster *lühüd pajo* style, the singers sang comparatively loudly, they had a clear articulation with the mouth fully open. In the graphs this is shown in the emphasis of the first pitch zones.

The curve of the spectrum is also connected with the relationship of the voice and its sonoric background. A steep curve is a reflection of the merging of the voice into its sonoric background (Laukkanen & Leino 1999, 176). In all songs of the faster *lühüd pajo* style, the curve is weaker than in the work songs. It is the purpose for a faster *lühüd pajo* to stand out from its background, whether it consists of sound from the accompanying instrument or from the voices of the audience. This kind of vocal sound is specifically intended to catch the attention of the audience. It is therefore natural that the faster *lühüd pajos* are rather performed alone, whereas the work songs are often performed in a group.

There is a marked emphasis in the spectra around 6000–7000 Hz. This emphasis is most marked in the faster *lühüd pajo* style in Fig. 6, but also in Fig. 2 and Fig. 5. There were similar emphases in Mantere's (2002, 53–54) analysis materials on the traditional singing

of the western Russians. Mantere concluded that this was due to the pressed vocal quality which lead to emphasis of pitch level above 5000 Hz. On the basis of earlier studies, the highest formants of the spectrum result rather from the individual physical formation of the performer's sound cavity and this is why the highest formants are interpreted as having more effect on the vocal timbre than on the vocal quality. In any case, this emphasis is discernible in the spectra of every singer analysed here. Furthermore, the sample of these singers was also homogenous regarding the factors of social environment, age group, sex and singing styles used. There was crack or rasp in many singers' voices due to exhaustion of the voice (especially in Fig. 6, sung by Z. F. Lovkina). The cracky vocal sound is not, in my opinion, intentional, and thus it has to be excluded from the description of the vocal style.

For the sake of comparison, I made a sum average spectrum from both singing styles. This spectrum combines average values of all the LTAS graphs from both styles. The spectra show how the summed average styles differ from each other (Fig. 14).

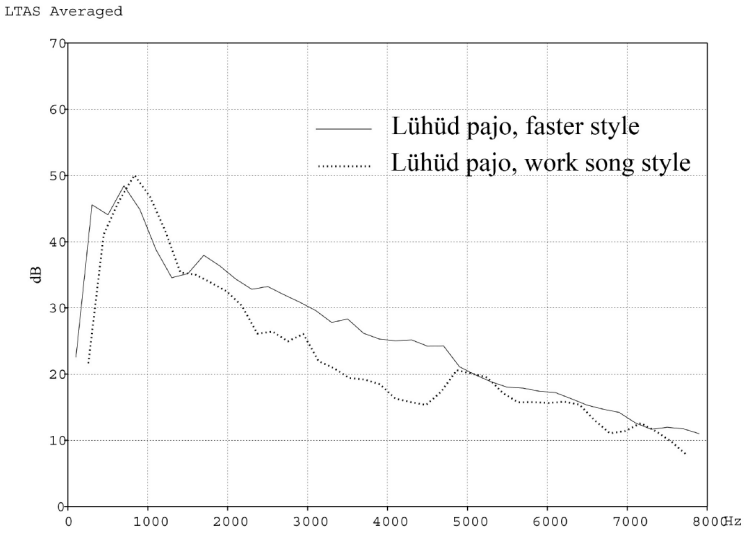


Fig. 14. LTAS averaged, sum of LTAS spectra of both *lühüd pajo* styles.

The differences between these two singing styles are most clearly seen between 1500 and 5000 Hz. In the faster *lühüd pajo* style, the pitch region of the basic pitch zone gains less energy than in the work songs. This may be caused by more pressed quality of the vocal production, as also indicated by relationship between the energy peak around 1000–2000 Hz and the highest peak. The smaller the inclination of the peaks, the less pressing in the voice (Mantere 2002, 65).

Comparing song and speech leads to similar conclusions: in a song the curve of the spectrum is less steep and the partial tones are stronger (Laukkanen & Leino 1999, 174). The style of the work songs is, indeed, closer to speech sound.

Discussion on the results of the spectral analyses

It was interesting to observe that the graphs were quite similar in Mantere's (2002, 52–62) analysis. The graphs of Veps faster *lühüd pajo* style are reminiscent of those made from cattle calls of Swedish Dalarna, whereas the graphs of Veps work songs are reminiscent of traditional songs from western Russia. Similar graphs can be found in other studies as well. For example, Kovačić, Boersman and Domitrović (2003) compared two singing styles, the Croatian *klapa* and *dozivački*. In the *dozivački* style, where voice production is reminiscent of shouting, there were similar emphases in the spectra than in the Veps faster *lühüd pajo* style. Furthermore, the curves and emphases were similar. They likewise found no singers' (3000 Hz), speakers' (3000–5000 Hz) or actors' (3500 Hz) formants, as were found in earlier studies of speech and song (Laukkanen & Leino 1999, 170–176). Consequently, they suggest that in the acoustic studies of speech a kind of shouters' formant should also be included. They consider it possible that the voice production of traditional singing style is based largely on other factors than on those associated with the ways professional singers or orators control their voices.

A characteristic in common in the faster *lühüd pajo* style and the phenomenon of the singer's formant is that both are motivated by a strong sounding and distinct voice quality. The singer's formant is, thus, evolved for the purpose of voice production audible over the sound of the accompanying musical instruments.

The differences between the graphs of the *same* singing style are due to various reasons. One factor is connected with the performer's individual competence within the style in question. This competence is naturally defined and evaluated by the norms of the society in question. Another factor is how the members of the society define the competence and reputation of a singer (Virtanen 1968, 8–15). This has an effect on the formation of a style as normative and an object for imitation. Most of the singers presented here represent that category of singers. This is also reflected in the spectra, where there are many features in common within a singing style.

The description of the voice timbre in a singing style contains some factors causing bias. The resonating pitches are dependent on individual physical structures of sound cavities. The voice timbre and quality are dependent on how a performer uses his/her voice. Aging brings loosening of the muscles as well as loosening of the vocal folds. This results in extra resonances, which are possibly those reflected in the emphasis at 6000 Hz (see Fig. 6). The deterioration of muscular and overall physical condition accompanies aging. There may be songs, which require great physical exertion and the faster *lühüd pajos* seem to be such songs. In my fieldwork I have observed how the singers have been exhausted after their performances. Furthermore, some of them may not be able to perform in the way they used to do, the way in which they still remember the songs were performed. As regards a singing style, it is important to consider these things in the analysis of the voice timbre. If, for example, a sample consists of singers of various age groups, the results may be biased.

Examination of the songs using spectrograms

For the spectrogram analysis I chose two songs from each of the style groups mentioned above. The first song, a *lühüd pajo* reminiscent of the faster *lühüd pajo* type was recorded in the village of Voilahti (Vologda region), among the eastern Veps (Fig. 15). The song was performed by Valentina Pavlovna Erotiseva (b. 1927). Her voice production results in a strong, pressed voice, reminiscent of shouting. In a narrow-band spectrogram it is possible to discern the partial tones from the songs and examine the actual way the voice is used there.

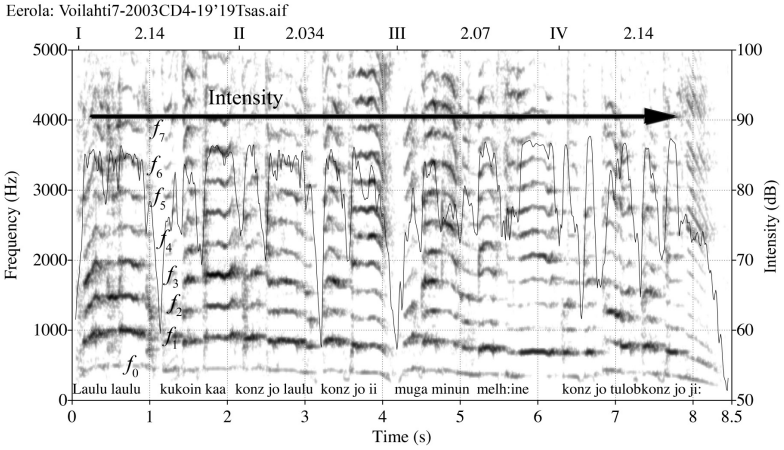


Fig. 15. Faster *lühüd pajo* style. Performed by V. P. Eraticheva, b. 1927, village of Voilaht (Russ. Voylakhta). Recorded in Voilaht, 5.7.2003 by Jari Eerola. Eerola: Voilahti7-2003CD4-19'19Tsas.aif.

The vertical axis of the graph shows pitches (Hz, left), intensity[8] (dB, right), the horizontal axis (below) shows duration (seconds) and the location of every other barline (of 4/4 metre). The value between the

8. Intensity is the acoustic analogy for the concept of stress, defined in everyday usage as loudness. Intensity reflects the strength of the signal, while amplitude shows the amount of air pressure in a given time unit.

Roman numbers there shows the duration of two bars. The graph shows how the basic pitch level (*f*0) gains only a small amount of energy (the lowest line in the graph). By contrast, the first, second and third partials are very well emphasized, likewise, to some extent, the fifth, sixth and seventh partials[9]. The graph also shows a strong region around 1.8 sec and at 2000 Hz, where the singer prononunces a syllable "*-kaa*".

The singer moreover sang with exceptional force, with almost a physical effect on my ears. This may have been due to her technique, where the resonating pitches coincided well with the resonance pitches of her voice cavity, thus strenghtening her voice. This way, the singer can be thought to have achieved an ideal way of voice production. The weak basic tone level reflects a pressed voice. The strong pressure of the voice is also reflected in the overall stability of the intensity level of the voice. Examining the performance on the video demonstrates how the singer uses her abdomen in order to have a good muscular support for her voice. This also results in the pressure well heard in the singing voice. Perhaps this is why the performance was a physically demanding task for the singer, although the duration of the verses was only some ten seconds.

Professional singers and teachers of singing use the concept of "support", referring to characteristics of "good" singing voice and its prerequisites. Support is defined as "intentional slowing down of exhalation" and is thus associated with control of breath during pronunciation with coordination of the breathing muscles and throat. (Laukkanen & Leino 1999, 30).

In the second example (Fig. 16) there is no vibrato, or no other voice effects. The words are pronounced with a strong separation, which is reflected in sharp little drops in the graph showing intensity. At the end there is a sharp decrease in the intensity. At this point in

9. There are differences in registering the partial tones. Usually a pitch is symbolized with *f*, but there are different conventions in marking the subindices of the basic pitch level. Sometimes the basic pitch is marked as *f*1 (Rossing 1990, 28) and especially in the field of phonetics and vocology with *f*0. I have conformed with the latter usage and I shall mark the basic pitch level as *f*0 and correspondingly the partial tones as *f*1, *f*2 etc. (see, e. g. Fig. 4.).

the song there is a kind of noise mark, which corresponds to a kind of sighing sound. Possibly this is a reflection of a local style, because I have not observed it among the central or southern Veps. The conclusion of the graph also shows how the basic tonal level bends together with the other partial tones. At this point it is impossible to identify the melodic progression.

The third example is a work song recorded in the village of Haragl in 2002 (Fig. 17). The performer was Antonina Bogdanova, who was in her sixties. She was from the neighbouring village of Jogenz (Russ. Ust'-Kapsa). This example is interesting in the sense that she sang the first two verses in Russian but then changed into Veps. She planned to sing the whole song in Veps, but accidentally began the song in Russian.

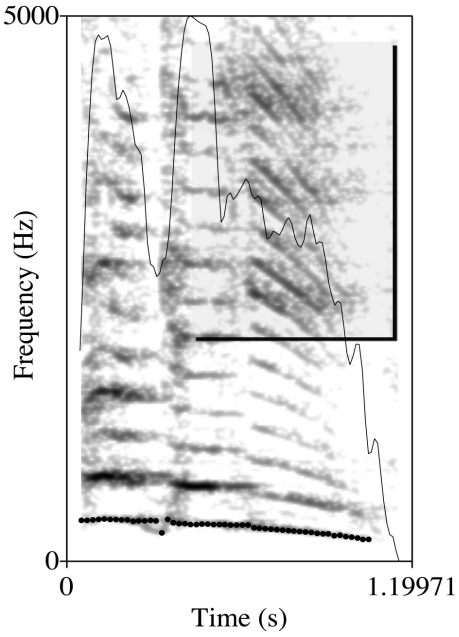


Fig. 16. A local way of concluding a faster *lühüd pajo*.

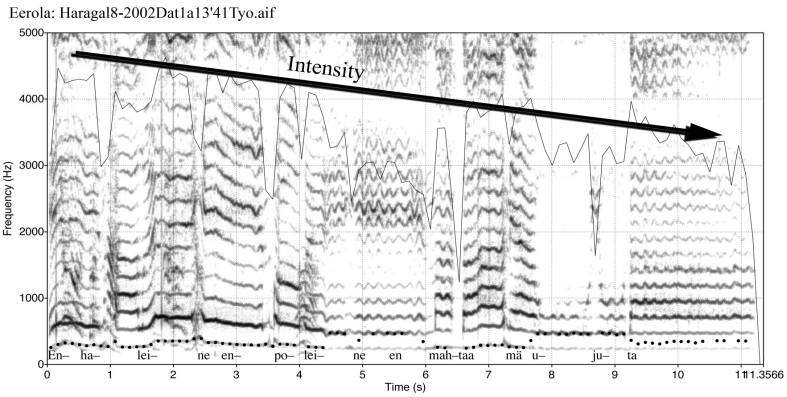


Fig. 17. Work song. Performed by Antonina Osipovna Bogdanova, b. 1932, village of Haragl (Russ. Kharagenichi). Recorded in Haragl 27.8.2002 by Jari Eerola. Eerola: Haragal8-2002Dat1a34'51Tyo.

The figure shows the spectrogram made from the work song. There is a remarkable difference from the previous song. The intensity decreases steadily towards the end of the song. The intensity line does not show such dramatic drops as in the faster *lühüd pajo* style. In the work song the tones are tightly woven together and a vibrato is also discernible in the voice. There were also slightly more little nuances in the melodic movement of the song. The intense binding of the words or syllables also has its effect on the overall vocal sound, because usually it is possible to stretch only vowels or voiced consonants. The vibrato moreover adds to the original flavour of the sound. The vibrato is modest here, but persists throughout the song in places where it is possible to sing with vibrato. These sections can be found especially in the border areas of the melodic lines (around 5 sec and 8–10 sec) and at the end. One of the stylistical peculiarities of this singing style is the maximal elongation of the final syllable (depending on the air capacity of the lungs). This elongation often makes the border area of the melodic lines of undefinable length. This kind of durational freedom is possible in the absence of accompanying instruments that may have had a controlling effect to the pulse of the song.

In this example, too, the region of the basic tonal level gains only a small amount of energy. The second and the third partials are most strongly emphasized. The higher partials weaken steadily. This was also shown as a slow descent of the spectral graphs.

Conclusion

In almost all the faster *lühüd pajos*, the voice quality can be described as pressed. Otherwise, timbre description is more difficult. The pressed voice is reflected in the graph as the weak position of the basic tone and in the emphasis of the nearest higher partials, whereas in a normal (speech) sound the basic tone and the upper tone are relatively strong. (Laukkanen & Leino 1999, 163.) The pharynx of a non-professional singer easily rises together with the pitch and the vocal folds collide with excessive force, which increases the pressure on the voice. On the other hand, this can be seen as a natural way of producing higher pitches. (Mantere 2002, 20–21.) According to Laukkanen and Leino (1999, 56) a metallic timbre is also associated with pressed voice. However, I did not observe this in these examples. Perhaps the singers could no longer achieve the pressure they could formerly achieve. I would describe the quality of their pressed sound as some kind of piercing hollow sound.

There was some sort of rasp, crack or hoarseness in the voices of all the singers, which may be due to their high age. Only two singers sang in a standing position (examples from Voilahti and Nemzha), while the others sang sitting in a chair. This may also have had an effect on the result, since those who sat could not achieve the full extent of voice production. Thus it is difficult to estimate how the different registers used by the singers differ from each other. With some examples it is possible to clarify this by examining the video recording of the performance.

The work songs are clearly calmer in nature. It is as if they reflect the overall nature of hard work. The voice production reminiscent of speech seems to be a reflection of the fact that it is not essential for the individual to be discerned or heard in these songs. On the other hand, this tells about the context of use or performance. There were no accompanying instruments available during the performance and everybody joined in the singing.

Singing, like other learned behaviour, contains internal references about how to do it "right". Different societies have different criteria for singing well and for who can be deemed as a good singer. These criteria are the fundament on which evolve different ways and stylistical norms of singing.

I divided the songs into stylistic genres largely depending on the use of the songs. It is also possible that in the past there was only one style or way of singing the *lühüd pajo* songs. Both the faster *lühüd pajo* and the work song style of singing are still called *lühüd pajo* songs. They have in common only the song texts and metrical basic structures (see Eerola 2003b). The social group may also have adopted new foreign influences and they may have started to perform their songs in a new manner and gradually new stylistical forms evolved according to different contexts of performance.

Literal sources

Cleveland, Thomas F., Sundberg, Johan & Stone, R. E. 2001. Long-term average spectrum characteristics of country singers during speaking and singing. *Journal of Voice*, v*ol 15, no. 1*, pp. 54–60.

Eerola, Jari 2003a. Laulujen kuvaajat. Spektrianalyysiohjelmat kansanlaulujen tutkimisen apuvälineinä [Song graphs. Programs for spectral analysis as an aid in folk song research]. *Musiikin suunta 3/2003*, pp. 33–44.

Eerola, Jari 2003b. Vepsäläisten lyhyiden pajojen metrinen malli ja sen toteutus. [The metrical model of the Veps lühüd pajo and its realisation] *Etnomusikologian vuosikirja 14*, ed. Antti-Ville Kärjä. Helsinki: Suomen etnomusikologinen seura. Pp. 101–130.

Dodge, Charles and Jerse, Thomas A. 1997. *Computer music. Synthesis, composition, and performance*. Second edition. Macmillan: New York.

Hakamies, Pekka 1994. Vepsäläisten suullisesta perinteestä [About Veps oral folklore]. *Vepsäläiset. Vepsäläiset tutuiksi. Kirjoituksia vepsäläisten kulttuurista [The Veps. Getting to know the Veps. Writing on the culture of the Veps]*, eds. Kaija Heikkinen & Irma Mullonen. Joensuu: Joensuun yliopisto. Pp. 73–86.

Heikkinen, Kaija & Mullonen, Irma 1994. Esipuhe. *Vepsäläiset [Foreword. The Veps]. Vepsäläiset tutuiksi. Kirjoituksia vepsäläisten kulttuurista [The Veps. Getting to know the Veps. Writing on the culture of the Veps]*, eds. Kaija Heikkinen & Irma Mullonen. Joensuu: Joensuun yliopisto. Pp. 10–12.

Hemsley, Thomas 1998. *Singing and imagination. A human approach to a great musical tradition*. New York: Oxford University Press.

Karlsson, Fred 2003. *Yleinen kielitiede [General linguistics]*. Revised edition. Helsinki: Yliopistopaino.

Kettunen, Lauri & Siro, Paavo 1935. *Näytteitä vepsän murteista [Samples of Veps dialects]*. Helsinki: SKS.

Kovačić, Gordana & Boersma, Paul & Domitrović, Hrvoje 2003. Long-term average spectra in professional folk singing voices: a comparison of the Klapa and Dozivački styles (ks. http://www.fon.hum.uva.nl/paul/papers/CroatianFolk_IFA25.pdf, downloaded 16.4.2004). *IFA Proceedings* 25, ed. E. H. Berkman. Amsterdam: University of Amsterdam. Pp. 53–64.

Laukkanen, Anne-Maria & Leino, Timo 1999. *Ihmeellinen ihmisääni. Äänenkäytön ja puhetekniikan perusteet, arviointi, mittaaminen ja kehittäminen [The wonderful human voice. The basics, assessment, measurement and development of voice usage and speech technique]*. Helsinki: Gaudeamus.

Lindestad, Per-Åke, Södersten, Maria, Merker, Björn & Granqvist, Svante 2001. Voice source characteristics in Mongolian 'throat singing' studied with high-speed imaging technique, acoustic spectra, and inverse filtering. *Journal of Voice, vol. 15, no. 1*, pp. 78–85.

Lomax, Alan 1968. *Folk song style and culture*. New Brunswick, NJ: Transaction publishers.

Makarjev, S. 1931. *Vepsäläiset. Lyhyt kansatieteellinen kuvaus [The Veps. A concise ethnographic description]*. Leningrad.

Mantere, Elina 2002. *Äänenlaatu neljässä perinteisessä laulutyylissä [Voice quality in four traditional singing styles]*. Unpublished master's thesis. Tampereen yliopisto, musiikintutkimuksen laitos.

Niemi, Jarkko & Jouste, Marko 2003. Teesejä pohjoisen laulun analyysiin: tarkastelussa nenetsit ja saamelaiset [Theses on the analysis of northern singing: the Nenets and the Sami under scrutiny]. *Etnomusikologian vuosikirja 14*, ed. Antti-Ville Kärjä. Helsinki: Suomen etnomusikologinen seura. Pp. 161–209.

Ross, Jaan 1992. Formant frequencies in Estonian folk singing. *The Journal of the Acoustical Society of America, vol. 91, no. 6*, pp. 3532–3539.

Rossing, D. Thomas 1990. *The science of sound*. Second edition. Reading (Mass.):Addison-Wesley.

Rüütel, Ingrid 1990. Vepsäläinen kansanmusiikki [Veps traditional music]. *Kansanmusiikki 3/1990*, pp. 30–32.

Saha, Hannu 1996. *Kansanmusiikin tyyli ja muuntelu [The style and variation in traditional music]*. Kaustinen: Kansanmusiikki-instituutti.

Salavuo, Miikka, Toiviainen, Petri & Välimäki, Vesa 2003. Teknologia musiikintutkimuksessa [Technology in music research]. *Johdatus musiikintutkimukseen [Introduction to music research]*, eds. Tuomas Eerola, Jukka Louhivuori & Pirkko Moisala. Helsinki: Suomen Musiikkitieteellinen Seura. Pp. 327–336.

Salve, Kristi 1998. Vepsäläisten kansallisesta identiteetistä [On the Veps national identity]. *Ison karhun jälkeläiset. Perinne ja etninen identiteetti yhteiskunnallisessa murroksessa [The descendants of the Great Bear. Tradition and ethnic identity in social transformation]*. Ed. Pekka Hakamies. Helsinki: SKS. Pp. 119–140.

Saukkonen, Pauli 1984. *Mistä tyyli syntyy [Where does style come from]?* Juva: WSOY.

Slawson, Wayne 1985. *Sound color*. Berkeley: The University of California Press.

Sundberg, Johan 1980. *Röstlära. Fakta om rösten i tal och sång*. Stockholm: Proprius.

Sundberg, Johan 1987. *The science of the singing voice*. DeKalb, IL: Northern Illinois University Press.

Sundberg, Johan 2004. *My research on the singing voice from a rear-view-mirror perspective*. Voice Research Centre. Department of Speech Music Hearing, KTH, Stockholm (ks. http://www.med.rug.nl/pas/conf_contrib/Sundberg/Sundberg/Sundberg_bio_touch_Groningen+figures.html, luettu 18.4.2004).

Suomi, Kari 1990. *Johdatusta puheen akustiikkaan [Introduction to speech acoustics]*. Publications of the Dept. of Logopedics and Phonetics no. 4. Oulu: University of Oulu.

Thalén, Margareta & Sundberg, Johan 2001. Describing different styles of singing: a comparison of a female singer's voice source in 'Classical', 'Pop', 'Jazz' and 'Blues'. *Logopedics Phoniatrics Vocology 26: 2*, pp. 82–93.

Virtanen, Leea 1968. *Kalevalainen laulutapa Karjalassa [The Kalevalaic singing style in Karelia]*. Helsinki: SKS.

Zangger, Borch Daniel & Sundberg, Johan 2002. Spectral distribution of solo voice and accompaniment in pop music. *Speech, Music and Hearing, vol. 43*. Stockholm: KTH. Pp. 31–35.