Timing in Estonian Folk Songs as Interaction between Speech Prosody, Meter, and Musical Rhythm

Author(s): Jaan Ross and Ilse Lehiste

Source: *Music Perception: An Interdisciplinary Journal*, Summer, 1998, Vol. 15, No. 4 (Summer, 1998), pp. 319-333

Published by: University of California Press

Stable URL: https://www.jstor.org/stable/40300861

Music Perception
Summer 1998, Vol. 15, No. 4, 319–333

Timing in Estonian Folk Songs As Interaction Between Speech Prosody, Meter, and Musical Rhythm

JAAN ROSS
University of Tartu, Estonia

ILSE LEHISTE
Ohio State University

Durations of acoustical segments were measured in four Estonian folk songs sung by a single performer, consisting of 152 verse lines, eight syllables each, with one note in the melody normally corresponding to one syllable in the text. The results were analyzed with regard to three aspects: notation, meter, and speech prosody. Three songs out of four are notated as isochronous sequences of 8 eighth notes per each verse line; in one song, certain pairs of eighth notes are replaced by a dotted eighth note plus a sixteenth note. The results revealed a complex interaction between meter, musical rhythm, and speech prosody. Variations in durations of sound events reflect the Kalevala meter on which the songs are based, with average rises in a foot being acoustically longer than falls. The duration differences between rises and falls are reduced in the so-called broken lines, which contain monosyllabic and trisyllabic words and allow for accommodation of short stressed syllables at a fall of a foot as required by the meter. Semantically relevant oppositions of word-initial short-long and long-short disyllabic units in speech are not kept completely intact in folk songs. Short-long disyllables are treated in a different manner by the performer, depending on whether their initial syllable occurs at a rise or at a fall in a foot.

Introduction

Music is different from the other arts (such as painting or literature) in a sense that a musical composition normally requires the mediation of musicians other than its creator, called performers, in order to be experienced by the listeners. Music performance is an increasingly popular object of studies in music psychology because it is believed to provide a rich domain

Address correspondence to Jaan Ross, Faculty of Philosophy, University of Tartu, Ülikooli 18, EE-2400 Tartu, Estonia. (e-mail ross@psych.ut.ee)

for the investigation of basic psychological constructs underlying such topics as sequence production, skill acquisition, individual differences, and emotional response (Palmer, 1997).

Within the framework of contemporary European tradition, the process of musical communication is commonly viewed as a series of largely independent actions by different individuals. The basic structure of this communicative process may be separated into three basic units: (1) the musical composition as such in the form of a score, which is prepared by the composer in advance, (2) performance of the score by musicians, and (3) perception and cognition of the performance by listeners (Ingarden, 1986). Thus the musical score is thought to be a more or less abstract representation of the musical work that must be elaborated by an instrumentalist or a singer in order for a listener to enjoy it. The performer's contribution to a musical composition is understood to comprise the musical expression.

Most music performance studies so far have focused on those compositions for which notation is available (i.e., on compositions done in the European tradition), in particular on piano performance (e.g., Palmer, 1989; Repp, 1992a, 1992b; Sloboda, 1983, 1985). When the investigator's aim is to study musical expression, the notation provides a good reference for interpretation of the data. The data must be separated into the information that originates from the musical composition proper (or the score) and the information that originates from the expressive intentions of the performer. This separation seems to be not too complicated: for example, deviations from the normative duration values in the score during the course of musical performance provide information for studying how the time domain can be used by the performer to convey his or her musical expressive ideas to listeners (e.g., Palmer, 1989).

Such a tripartite interaction of composer, performer, and listener in the musical communication, however, is not applicable to the vast majority of music performed in the world, such as popular and traditional music. In the kind of music that constitutes the so-called oral tradition, the functions of composer and performer are often not well separated and, as a consequence, such music is usually notated only partially or not at all. The performer is thought to recreate the musical work anew from memory each time the composition is performed.

An example of such oral musical tradition is the old Estonian folk song (Tampere, 1983). Although some portions of old Estonian folk songs exist also in the form of written notation, the relationship between the notation and the performance in this case is rather different from that prevailing in the classical European tradition. In this European tradition, the system of notation was devised essentially for the music of Western civilization, and its purpose is to help a performer carry out the composer's intentions rather than to describe the musical actions of the performer (Nettl, 1973, p. 31).

Seeger (1958) has called the European notation prescriptive. Its opposite, descriptive notation is intended to convey to a reader the characteristics and the details of a musical composition that the reader does not already know. Descriptive notation is compiled post factum by an ethnomusicologist for the purpose of information storage (a large part of old folk song collections date from a time when sound recording was not available) and, because of the abstractions and categorizations involved in notation, it cannot capture the full detail of a performance. In addition, the musical idiom of the old folk songs is in many aspects (scale and temporal organization) different from the Western tradition, which makes the European notation system to a certain extent inadequate even for information storage.

Consequently, the reference system to which nonnotated folk music performance is to be compared, should, instead of or in addition to the score, include components that are thought to influence the acoustical shape of the performance. In the case of folk singing, such first-order additional components appear to be the phonetics of the sung text (which determines the intelligibility of the performed words) and the verse meter (which determines the structure of the composition). The purpose of this study is to investigate the relationship between word-level prosody, the meter, and the notation in the timing of the Estonian folk song performance.

Old Estonian Folk Song: An Overview

The old Estonian folk song[1] has deep historical roots. It is believed to have stabilized in its present form during the first millennium B.C. (Leino, 1986). The tradition is relatively well documented since the end of the 19th century when it was still alive (e.g., Helmholtz, 1877).[2] The old folk songs are now sung only in a few remote regions of the country. Estonian archives contain about 50,000 records of field notations of old folk songs and about 25,000 tape-recorded performances of those songs (Anonymous, 1987).

The two primary components in folk songs are the text and the melody. In spoken Estonian, contrastive duration (linguistic quantity) is used to

1. We use the expression "old folk songs" to refer to folk songs in the Kalevala meter. During the 19th century, the old songs were gradually replaced by so-called "new folk songs," which are based on the Western European musical idiom.

2. Helmholtz (1877, p. 465) writes: "*Noch jetzt übrigens sträuben sich nach A. v. Oettingen's Bericht die Esthen in Chorälen der Molltonart den Leitton zu singen, selbst wenn er ihnen deutlich durch die Orgel angegeben wird.*" A high leading tone is not typical of the minor scale in old Estonian folk songs. Oettingen's [who was professor of physics at the University of Tartu in 1863 to 1893] observation gives evidence that the musical idiom of old folk songs was more familiar to the congregation of that time than was the Western European functional tonality on which the Protestant hymns are based.

convey distinctions in meaning (Lehiste, 1997). Estonian words can be broken down into pairs of consecutive syllables, or disyllabic feet. Standard Estonian uses three types of feet that are durationally contrastive. In the short quantity (Q1), the ratio of the duration of the first syllable (S1) to that of the second syllable (S2), the S1/S2 ratio, is approximately 2/3 (0.66), which means that a Q1 foot consists of a short(er) syllable followed by a long(er) syllable. In the long quantity (Q2), the S1/S2 ratio equals approximately 3/2 (1.5), and in the overlong quantity (Q3),the S1/S2 ratio is equal to 2/1 (2.0). Thus both a Q2 foot and a Q3 foot consist of a long(er) syllable followed by a short(er) syllable. We may, as an approximation, ignore the precise differences in duration between individual syllables and consider short and long syllables only.

The rhythm in folk song melodies is not very variable. According to the taxonomy of folk songs (Kolk, 1989), the basic pattern (the length of which is determined by the verse meter, the four-foot trochaic line) is an isochronous string of notes of duration approximately 350 ms each, denoted as eighth notes in the notation. Certain deviations from this pattern may occur: a pair of eighth notes may be replaced by a dotted eighth note plus a sixteenth note, or a single eighth note may be replaced by a pair of sixteenth notes. A characteristic exception from the generally isochronous sequence of note durations is the rhythm of swing songs, which imitates the movement of a swing and consists of alternations of a short and a long unit in the triple meter.

Two corpora are at the disposal of a folk song performer, the texts and the tunes, which are largely independent of each other. This means that almost any arbitrary text from one corpus may be combined with almost any arbitrary melody from the other corpus (Tampere, 1983). Certain restrictions did apply, however, associating particular melodies with specific groups of texts and vice versa (e.g., in swinging songs).

The framework within which the free combination of texts and melodies in folk songs could take place is provided by the meter of the folk songs, where the basic unit is a line consisting of four trochaic feet comprising eight syllables (or eight notes, as there is normally a one-to-one correspondence between the syllables and the notes in the melody). It is called the Kalevala meter, which is shared by Finnish and Estonian old folk songs. It is not strictly trochaic in the sense that each metric foot should consist of a long and short syllable; the contrast between ictus and off-ictus positions ("ictus" or "rise" referring to metrical stress) in the Kalevala meter is realized by means of both stress and quantity, in a manner which Leino (1986, p. 29) describes as follows: "The only metrically relevant syllable is the first one [which is always stressed in Estonian]. If this is long, it is placed on a rise, but if it is short it must occur on a fall; however, these restrictions do not affect the first rise and fall of a line."

The preceding discussion shows that acoustical syllable/note durations in folk song performance must satisfy requirements that originate from at least three different sources. First, the contrastive short-long and long-short patterns characteristic of spoken Estonian are expected to be preserved as far as possible in singing, in order to convey the meaning of the words. Second, these disyllabic patterns must be adapted to the requirements imposed by the (partly) duration-based metrical structure of the Kalevala verse. Third, the rhythm of folk song melodies, albeit not too variable, does involve expressive patterns that must be accommodated within the musical communication as well.

Materials and Methods

In the northeastern Estonian coastal dialect, the ternary system of quantity oppositions, characteristic of standard Estonian, has been replaced by a binary system more similar to that of Finnish—perhaps because of widespread communication with speakers of Finnish (Must, 1987). LK, the performer of the analyzed folk songs, originates from the Haljala district, where the Northeastern coastal dialect was spoken. She was considered to be the favorite singer of the Haljala district during the first decades of this century, when the majority of folk songs was collected. Four songs from her repertoire, recorded in 1937 and preserved in the Estonian Folklore Archives, were chosen for analysis. The songs had been notated by an ethnomusicologist (Udo Kolk) and published in Laugaste (1989, Vol. I: pp. 259–260, 329–331, 359–360, and Vol. II: 197–198; indexed as 36o, 55c, 60i, and 649k, respectively). The total number of verse lines in the four songs is 152.

The typology of folk song melodies (Kolk, 1989) is based on the following parameters: (1) the length of the basic structural unit repeated during the song, expressed in the number of lines, with a line equaling four trochaic feet, (2) the melody contour, (3) the rhythmic pattern, (4) the scale, and (5) the pitch range. The length of the basic structural unit is two lines in three of the songs and four lines in the other song (55c). The melody contour is different in all four cases. The rhythmic pattern is notated as an isochronic sequence in three songs; in 60i, certain pairs of eighth notes (feet) are replaced by a dotted eighth note plus a sixteenth note. The scale is major in all four cases; the pitch range varies from 4 to 5 semitones (649k) and 8 to 9 semitones (60i) to an octave (36o and 55c). Two randomly chosen lines from each of the four songs are presented in Figure 1.

Durations of sound events (interonset time intervals) were measured in these songs. Measurements were performed by using two parallel, wideband and narrowband spectrographic representations of the sound signal produced by a Kay CSL Workstation Model 4300.

Results

METER

Difference between Ictus and Off-Ictus Positions

Ictus positions coincide with the odd-numbered sound events in a line (1, 3, 5, 7); off-ictus positions coincide with the even-numbered sound events

in a line (2, 4, 6, 8). A trochaic metric foot consists, by definition, of a long and a short syllable. In a four-foot trochaic line, it is expected that the syllables that fall in ictus position (the first syllables of trochaic feet) are longer than the syllables falling in off-ictus position (second syllables of trochaic feet). In the Kalevala line, ictus and off-ictus positions are differentiated from each other on the basis of both stress and quantity (Leino, 1986); thus it may be expected that the average durations of syllables in these positions differ from one another. This turned out to be the case: analysis of variance shows that in all four songs, the sound events at ictus positions were systematically longer than those at off-ictus positions. The average durations for ictus and off-ictus positions are shown in Table 1.

It is evident from Table 1 that all four songs were performed at slightly different tempi. The differences between average durations of sound events (i.e., ictus and off-ictus positions pooled) are highly significant at $p < .001$, $F(3,1177) = 21.34$. A monotonous relationship from slower to faster is apparent among the first three songs (36o, 55c and 60i), where the increase in average duration at ictus positions is accompanied by an increase in average duration at off-ictus positions. The ictus/off-ictus duration ratio is 1.15 for song 36o, 1.19 for song 55c, and 1.17 for song 60i. The song 649k, however, deviates from this succession with smaller durational contrast (1.07) between ictus and off-ictus positions.

Trochaic and Broken Verse Lines

Roughly two types of verse lines can be distinguished in old folk songs. The first type, referred to as trochaic lines, contains only words with an even number of syllables (2, 4, 6, or a theoretically possible maximum of 8) or pairs of monosyllables. Such lines can be segmented into disyllabic units

TABLE 1
Average Durations of Sound Events at Ictus and Off-Ictus Positions of the Melody in Four Folk Songs

	Duration of Sound Event (ms)		
Song	Ictus	Off Ictus	Significance of Difference
36o	306 (40.4)	265 (40.2)	$F(1,297) = 80.13$, $p < .001$
55c	321 (49.8)	270 (49.5)	$F(1,357) = 95.00$, $p < .001$
60i	342 (49.5)	292 (52.7)	$F(1,316) = 77.54$, $p < .001$
649k	303 (32.1)	283 (40.8)	$F(1,203) = 15.99$, $p < .001$
Overall	320 (47.2)	277 (47.9)	$F(1,1179) = 243.14$, $p < .001$

Standard deviation is enclosed in parentheses.

or feet. In Estonian, stress is carried by the first syllable of a foot, which implies that in trochaic lines, stresses fall on odd-numbered positions, coinciding with the metrically strong ictus positions.

Forty of the 152 lines belonged to the second type, referred to as broken lines. Examples of broken lines may be found in Figure 1, in songs 36o ("kol-man-da ka-la jä-re-le") and 649k ("sin-na mo sõr-med su-re-vad"). These lines contain one or two trisyllabic words whose presence destroys the regular stress pattern of trochaic lines. Trisyllabic words are likewise stressed on the first syllable and constitute dactylic metric feet. Their quantity is determined by the first two syllables; as is the case with disyllabic (trochaic) metric feet, the stressed first syllable may be either short or long, and if the first syllable is short, the second syllable of Estonian words is longer than the first. As the meter takes into account both stress and quantity, the dactylic lines make it possible to accommodate the short-quantity words to the meter, so that their short stressed syllables may be kept from falling in ictus positions.

No significant differences were found in average durations of sound events between the trochaic and the broken lines. However, two-way analysis of variance shows significant interaction between the metrical position (ictus/off ictus) and the line type (trochaic/broken) in two songs (60i and 55c)

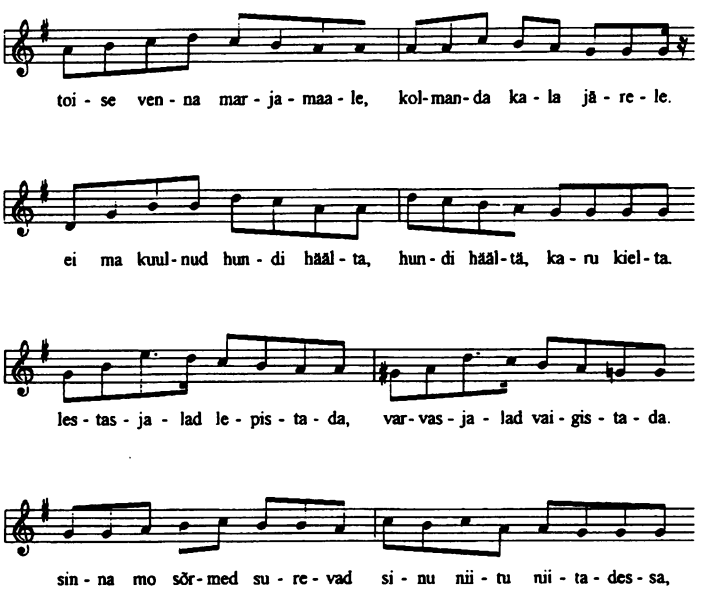

Fig. 1. Two randomly chosen verse lines, notated by Udo Kolk, from the four analyzed folk songs, indexed (from top to bottom) as 36o, 55i, 60i, and 649k by Laugaste (1989).

and marginally significant interaction in the other two (649k and 36o). The average values of sound event durations at ictus and off-ictus positions are given separately for trochaic and for broken lines in Table 2.

Table 2 shows that in broken lines, the durational contrast between ictus and off-ictus positions is smaller than in the regular trochaic lines: the difference between the average durations of syllables falling in ictus positions and in off-ictus positions is 50 ms in trochaic lines, whereas in broken lines the difference is only 22 ms. It is worth noting that the average durations of metric feet in the two types are not too different from each other: 594 ms in trochaic lines and 608 ms in broken lines.

NOTATION

According to the notation (Laugaste, 1989), the rhythmic pattern in three analyzed folk songs (36o, 55c, and 649k) is an isochronous sequence of eighth notes. Only a few occasional deviations from this pattern can be found in the score. The fourth song, 60i, belongs to a different category in which certain feet are notated as consisting of a dotted eighth note followed by a sixteenth note, instead of a pair of isochronous eighth notes. The average durations of sound events in the song 60i are as follows (standard deviation is given in parentheses): for "normal" eighth notes, 318 ms (47.5); for dotted and fermatted eighths, 395 ms (46.2); for notes shorter than eighths (mostly sixteenths), 276 ms (61.3). Analysis of variance shows significant differences among the three groups [$F(2,315) = 51.33, p < .001$]. The difference in duration between the eighth notes and the sixteenth notes (13%) and between the "normal" eighth notes and the dotted eighth notes (24%), however, is not nearly as large as suggested by the notation (50% in both cases). It may be thought that durational categories are more liberally defined in one-voiced music than in the Western multivoiced musical idiom,

TABLE 2
Average Durations of Sound Events at Ictus and Off-Ictus Positions of the Melody, Presented Separately for Trochaic and Broken Lines

Song	Duration of Sound Event (ms)				Significance of Interaction
	Trochaic Lines		Broken Lines		
	Ictus	Off Ictus	Ictus	Off Ictus	
36o	307	263	298	280	$F(1,295) = 3.09, p = .080$
55c	327	268	298	280	$F(1,355) = 9.28, p = .002$
60i	352	285	332	300	$F(1,314) = 9.60, p = .002$
649k	305	279	298	295	$F(1,201) = 3.77, p = .053$
Overall	322	272	315	293	$F(1,1177) = 18.85, p < .001$

where acoustical coordination between voices is strongly needed and that dotted eighth notes and sixteenth notes were used by the transcriber to denote durations "longer than" or "shorter than" the normative eighth note, respectively.

Even though the notation suggests that the eighth notes should all have the same duration, as Table 2 shows, in trochaic lines, the notes occurring in odd-numbered (ictus) positions of the line are noticeably longer.

WORD-LEVEL PROSODY

Quantity oppositions in Estonian words are realized on the disyllabic (foot) level rather than on the syllabic level (Lehiste, 1997). Tables 3 and 4 present comparisons of both short- and long-quantity disyllabic and trisyllabic words starting at ictus and off-ictus metrical positions.

Table 3 shows that short-quantity disyllabic words may start both at ictus (n = 64) and off-ictus (n = 13) positions; short-quantity trisyllabic

TABLE 3

Durations of Sound Events and S1/S2 Ratios in Word-Initial Disyllables in Disyllabic and Trisyllabic Short-Quantity Words Starting at Ictus and Off-Ictus Positions

		Duration of Sound Event (ms)		
	No. of Instances	First Syllable	Second Syllable	S1/S2 Ratio
Disyllabic at ictus	64	293	297	.987
Disyllabic at off ictus	13	266	345	.771
Trisyllabic at ictus	2	270	353	.765
Trisyllabic at off ictus	29	285	323	.882

TABLE 4

Durations of Sound Events and S1/S2 Ratios in Word-Initial Disyllables in Disyllabic and Trisyllabic Long-Quantity Words Starting at Ictus and Off-Ictus Positions

		Duration of Sound Event (ms)		
	No. of Instances	First Syllable	Second Syllable	S1/S2 Ratio
Disyllabic at ictus				
Assumed long	100	320	280	1.143
Assumed overlong	182	328	269	1.219
Disyllabic at off ictus	2	330	308	1.071
Trisyllabic at ictus	14	331	289	1.145
Trisyllabic at off ictus	1	315	290	1.086

words begin almost exclusively at off-ictus positions (in 29 of 31 cases). Recall that the characteristic S1/S2 ratio of short-quantity words is 2/3; if a short-quantity word begins at an off-ictus position, the longer second syllable falls into ictus position, which is expected to be longer in a trochaic meter. Short-quantity words starting in off-ictus position have S1/S2 ratios that are closer to the values found in the spoken language; disyllabic short-quantity words starting at ictus position have a shorter than expected second syllable, resulting in almost equal duration of the two syllables—a pattern that is not found in normal speech.

Table 4 shows that in the case of words beginning with long syllables, the texts of the analyzed songs follow the metrical rules of the Kalevala verse system almost without exception. Of a total of 299 such words, the initial long syllable was placed in ictus position in 296 instances. Although the Northeastern coastal dialect spoken in the area from which the singer LK originates does not differentiate between long and overlong quantity (Must, 1987), we attempted to separate these categories according to what the quantity of the words would be in the standard dialect. The result is shown in the first two lines of Table 4.

The S1/S2 ratio of words in the long quantity is expected to be 3/2, or 1.5; the corresponding ratio for words in the overlong quantity is expected to be 2/1, or 2.0. With a longer first syllable and a shorter second syllable, the words that are overlong in the standard language here have a larger S1/S2 ratio, which brings them closer to the expected value; the differences between the S1/S2 values found in the present materials, however, are not significant.

SOUND-EVENT DURATIONS AND THE MELODY CONTOUR

Generative performance rules (Friberg, 1991) indicate that durations of musical events may be correlated with the direction of leaps in the melody: in a rising melodic interval, making the lower note shorter than the upper one is thought to contribute to the musical expression. We checked whether such a relationship can be detected in the folk song performance. Analysis of variance showed that durations of sound events were significantly different in three conditions for all songs (Table 5): (1) after a rising melodic interval, (2) after a falling melodic interval, and (3) after an interval in which the pitch remains the same. The duration pattern, however, varies across the songs. In all four songs, the second note in rising melodic intervals was longer than (or, in 36o, equal to) the second note in falling intervals or in unison. In two songs, 36o and 55c, the duration of a repeated pitch in the melody is considerably shorter than the second note in a rising or falling interval, which is suggested also by the overall duration pattern shown in Table 5.

TABLE 5
Durations of Sound Events (ms) After an Upward or Downward Pitch Jump or Repetition of the Same Pitch

	Change from Preceding Melodic Interval			
Song	Up	Down	Same	Significance of Differences
36o	294	294	270	$F(2,296) = 9.72, p < .001$
55c	305	301	281	$F(2,356) = 6.26, p = .002$
60i	331	309	317	$F(2,315) = 4.38, p = .013$
649k	303	280	295	$F(2,202) = 6.08, p = .003$
Overall	309	300	287	$F(2,1178) = 15.62, p < .001$

Discussion

Results of this study confirm the two earlier findings of Ross and Lehiste (1994) about microdurational variations in old Estonian folk songs: (1) in the trochaic tetrameter, rises, or ictus positions, are performed systematically longer than falls, or off-ictus positions (Table 1), and (2) differences between semantically relevant duration ratios in disyllabic units are much reduced (but still evident) in singing as compared with those ratios in speech (Tables 3 and 4). This finding suggests that quantity-related features of speech prosody are subordinated to the meter in folk song performance, which may be explained by the need for free combination of melodies and texts from the two independent corpora.

The classical Kalevala meter expects short-quantity words, that is, words starting with the short-long disyllabic pattern, to begin at off-ictus positions. In the analyzed songs, this rule was indeed maintained in the case of trisyllabic words, where 29 of 31 start at a fall (Table 3). In the case of disyllabic words, however, the rule is more often violated than followed: 64 of 77 short-quantity disyllabic words start at a rise, and only 13 start at a fall.

Abandoning the rule with disyllabic words may be correlated with the reduction in the number of so-called broken verse lines in a certain portion of the Kalevala songs. In the present case, the broken lines are replaced by normal trochaic lines, which (contrary to the old Kalevala metrical principle) may also include short stressed syllables in ictus positions. Within the classical Kalevala meter the proportion of broken lines is nearly 50%, but in the northern Estonian folk poetry (to which the four songs analyzed in this study belong), the proportion is estimated to have decreased to approximately 25% (Leino, 1986). In the analyzed set, the proportion is in fact 26%.

One song (60i) that systematically explores the dotted rhythm contains 22 instances of dotted eighth notes followed by a sixteenth note. From those 22 instances, 13 words are in the long quantity (i.e., following the long-short disyllabic pattern) and 9 words are in the short quantity (i.e., following the short-long disyllabic pattern). The performer shows no preference for combining words in long quantity with the dotted rhythm, as it might be reasonable to expect. The proportion of short-quantity and long-quantity words in feet with the dotted rhythm suggests that the occurrence of either type of word is simply random.

Palmer and Kelly (1992) and Palmer, Walker, and Kelly (1998) investigated the role of meter in language (English) and in common metrical patterns of Western tonal music. Their analysis of song compositions indicated that musical and poetic meters were aligned more often than not, both for binary and ternary meters. Experiments were done that addressed listeners' perception of binary and ternary meters in combined speech and music. The listeners performed better when the musical and linguistic meters matched, that is, when word stresses coincided with musical stresses.

Repp (1995) reports a psychoacoustic experiment in which subjects had to detect temporal perturbations in an isochronous melody played on a computer-controlled piano. He found that detectability of lengthening of tones by the listeners was negatively correlated with lengthening in the expert performance and that false-alarm frequencies correlated positively with detection accuracy. He interpreted those results as supporting the so-called top-down hypothesis, according to which microtemporal variations reflect listeners' expectations of expressive performance microstructure, rather than the so-called bottom-up hypothesis, according to which the microtemporal variations are caused by psychoacoustic, or pitch-related stimulus factors. Repp (1995) found generally no effects of metrical structure on the timing of the performance.

The results obtained in the present study are quite different from those obtained by Repp (1995), Palmer and Kelly (1992), and Palmer et al. (1998). There are many possible reasons for the differences: Repp (1995) used purely instrumental music played on a keyboard, the interonset time intervals were considerably longer in his study (which perhaps weakened the metrical structure), and the music was more melodic than metric in character. In this study, we did find effects of metrical structure on the timing of the performance: these effects are perhaps the strongest among all those checked for during this study. Other results, such as fitting foot-level speech prosody patterns to durational pattern of the melody, seem to be of the bottom-up rather than the top-down type.

As far as the findings of Palmer and Kelly (1992) and Palmer et al. (1998) are concerned, the language with which these experimenters were working was English. It is possible that their results apply only to stress-timed lan-

guages such as English. Our results indicate that in Estonian folk songs, the music tends to accommodate to those speech characteristics that are related to quantity but not to those speech characteristics that are related to stress. This is in accordance with the primary role of quantity instead of stress in Estonian prosody.

We do not think that the timing in songs in the old Kalevala meter is governed by the demands of expressive performance microstructure. Instead, it seems to be determined by the different needs of the three systems interacting with each other in those songs: poetic meter, musical rhythm and melody, and speech rhythm. This judgment, if true, points once more to the truth of the observation that the rules that govern the Western musical idiom need not be applicable to music from other cultures, such as music associated with the old Kalevala tradition.

Conclusions

This study reveals a complex interaction between meter, word-level prosody, and musical rhythm in old Estonian folk songs. The classical Kalevala verse is a trochaic tetrameter, in which the opposition between rises and falls is accomplished by means of durational contrast. In Estonian speech, the disyllabic sequence either constituting or starting a word may have a short-long or long-short structure, and the difference between them is semantically relevant (the long-short pattern in turn providing two possible oppositions in the standard language). In the Kalevala verse, therefore, words in long quantity (i.e., those starting with the long-short disyllabic sequence) fit the metrical structure better when their initial syllable falls in the ictus position. Words in short quantity (i.e., those starting with the short-long disyllabic sequence), on the contrary, fit the metrical structure better when their initial syllable falls in the off-ictus position.

In the analyzed songs, this schema is working with trisyllabic words, which tend to begin at rises when they are of the long quantity and at falls when they are of the short quantity. Disyllabic short-quantity words, however, may begin equally at falls and rises, which suggests that elements of stress-based verse are to some extent starting to replace the quantity-based classical Kalevala meter.

No regularity was observed in combining the dotted rhythm within the otherwise isochronous rhythmic pattern of the songs with appropriate words. Both short- and long-quantity words appeared to fit the dotted rhythm, despite the fact that in the case of long-quantity words, the correlation between the musical timing and the prosody-prescribed duration ratio of the two syllables is better.

Acoustically, sound events at ictus positions were found to be longer than those at off-ictus positions. This difference, however, was reduced in the broken lines, in which trisyllabic short-quantity words had to be accommodated to the metrical pattern. The reduction may be a part of the general neutralization of short and long sound events in the songs, possibly because of the need to combine texts and tunes from the two independent corpora.[3]

References

Anonymous (1987). Eesti rahvamuusika [Estonian folk music]. In G. Naan (Ed.), *Eesti nõukogude entsüklopeedia,* Vol. 2 (pp. 444–445). Tallinn: Valgus.

Friberg, A. (1991). Generative rules for music performance: A formal description of a rule system. *Computer Music Journal, 15*(2), 56–71.

Helmholtz, H. von (1877). *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik* [On the sensations of tone]. Braunschweig: Vieweg und Sohn.

Ingarden, R. (1986). *The work of music and the problem of its identity.* Berkeley: University of California Press.

Kolk, U. (1989). Haljala regiviiside tüpoloogia [Typology of Kalevala melodies from Haljala]. In E. Laugaste (Ed.), *Vana kannel VI: 1. Haljala regilaulud* (pp. 45–69). Tallinn: Eesti Raamat.

Laugaste, E. (1989). *Vana kannel VI: Haljala regilaulud* [Complete Kalevala songs from Haljala], 2 vols. Tallinn: Eesti Raamat.

Lehiste, I. (1997). Search for phonetic correlates in Estonian prosody. In I. Lehiste & J. Ross (Eds.), *Estonian prosody: Papers from a symposium* (pp. 11–35). Tallinn: Institute of Estonian Language.

Leino, P. (1986). *Language and metre: Metrics and the metrical system of Finnish.* Helsinki: Suomalaisen Kirjallisuuden Seura.

Must, M. (1987). *Kirderannikumurre: Häälikuline ja grammatiline ülevaade* [Northeastern coastal dialect: Survey of phonology and grammar]. Tallinn: Valgus.

Nettl, B. (1973). *Folk and traditional music of the Western continents.* Englewood Cliffs, NJ: Prentice-Hall.

Palmer, C. (1989). Mapping musical thought to musical performance. *Journal of Experimental Psychology: Human Perception and Performance, 15*(12), 331–348.

Palmer, C. (1997). Music performance. *Annual Review of Psychology, 48,* 115–138.

Palmer, C., & Kelly, M. H. (1992). Linguistic prosody and musical meter in song. *Journal of Memory and Language, 31,* 525–542.

Palmer, C., Walker, T., & Kelly, M. H. (1998). The role of meter in language and music. Manuscript in preparation.

Repp, B. (1992a). Diversity and commonality in music performance: An analysis of timing microstructure in Schumann's "Träumerei." *Journal of the Acoustical Society of America, 92*(5), 2546–2568.

Repp, B. H. (1992b). Probing the cognitive representation of musical time: Structural constraints on the perception of timing perturbations. *Cognition, 44,* 241–281.

Repp, B. (1995). Detectability of duration and intensity increments in melody tones: A partial connection between music perception and performance. *Perception and Psychophysics, 57*(8), 1217–1232.

3. Parts of this paper were presented at the Third Triennial Conference of the European Society for the Cognitive Sciences of Music, held in Uppsala, Sweden, June 7–12, 1997. We thank Bruno Repp for comments on an earlier draft.

Ross, J., & Lehiste, I. (1994). Lost prosodic oppositions: A study of contrastive duration in Estonian funeral laments. *Language and Speech, 37*(4), 407–424.
Seeger, C. (1958). Prescriptive and descriptive music writing. *Musical Quarterly, 44,* 184–195.
Sloboda, J. A. (1983). The communication of musical metre in piano performance. *Quarterly Journal of Experimental Psychology, 35A,* 377–396.
Sloboda, J. A. (1985). Expressive skill in two pianists: Metrical communication in real and simulated performances. *Canadian Journal of Psychology, 39,* 273–293.
Tampere, H. (1983). *Estonskaya narodnaya pesnya* [Estonian folk song]. Leningrad: Muzyka.