

## TRADE-OFF BETWEEN QUANTITY AND STRESS IN ESTONIAN FOLKSONG PERFORMANCE?

*Jaan Ross*, University of Tartu, Estonia

*Ilse Lehiste*, Ohio State University, USA

---

This paper is concerned with the issue of how speech and music prosodies are matched to each other in folksong performance. We consider speech and music to be two more or less autonomous structures which both are required to be accommodated in singing performance according to a certain *modus vivendi*. Under some conditions, this necessity for coexistence may lead to a conflict between the two media which, as we believe, is caused by the different nature in which speech and music employ the acoustical continua of pitch, duration, and timbre. In speech, contrastive sounds - phonemes - are primarily built up by exploiting the differences in sound spectrum (for example, vowels can be distinguished by differences in their lower formant frequencies). In a melody, however, the pitch dimension is of primary importance for its structure. The building-blocks for melodies are the scale steps which result from discretization of the frequency continuum into a number of separate levels.

It seems reasonable to suppose that a desirable result of the coexistence of speech and music prosodies in singing would be as perfect match as possible between the two prosodies, unless certain "existential« requirements from the side of one medium prevent reaching this. A number of studies about the relationship between the two prosodies in singing may be found in the literature which seem to justify this prediction. In tone languages like Chinese or Japanese, linguistically relevant tone patterns tend to be matched to melodic contours in music (Yung 1983). In languages like Indo-European where contrastive opposition exists between stressed and unstressed syllables, linguistic stress patterns tend to coincide with the metrical structure of stressed and unstressed patterns in music (Palmer & Kelly 1992). On the other hand, a classical example of lack of fit between the two prosodies is singing of high notes by female opera singers where it is physically impossible to keep the vowel formant frequencies at the vicinity of their etalon values from speech, because of the high fundamental frequency (Sundberg 1987).

In this paper, we will address the issue of the goodness-of-fit between metrical and word stress in old Estonian folksong repertoire. Estonian is a Finno-Ugric language which belongs to a larger group of the Uralic languages. Stress in Estonian words always falls on the first syllable. An important characteristic of Estonian phonology is use of contrastive duration. There are three contrastive quantity degrees in standard Estonian words, and the difference between them is semantically relevant. The three degrees are known as short, long, and overlong. The difference between long and overlong degrees in most cases is not indicated in written language. In spoken language, differences between the three quantity degrees are manifested in speech by means of the ratio of the duration of the initial syllable to the duration of the second syllable in a word (Lehiste 1968).

An important characteristic of the Estonian folksong repertoire is the relative independence of the text and the melody corpora. Almost every text from the former set may be combined with an arbitrary melody from the latter set. This implies that there must exist strong structural constraints which cause such an interplay between texts and melodies to be available to a performer. The necessary structural framework is provided by the metre which in Estonian folksongs is based on the contrast of long and short, rather than stressed and unstressed syllables. The standard so-called

Kalevala verse, which is used in old folksongs, consists of four trochaic feet, equalling eight syllables. Odd-numbered syllables bear metrical stress (or ictus).

It is evident that given the large freedom of combining melodies and texts with each other, the match between the speech and music prosodies cannot always be perfect. One conflict which may emerge between the two prosodies is the impossibility to match the word stress to the metrical accent in the performance. One requirement for the Kalevala metre is that a short stressed syllable be excluded from positions with metrical accent. In the case of words with a short first syllable, a stressed syllable thus is required to occur in a metrically weak position, and the following unstressed syllable occurs in a metrically strong position, resulting in a conflict between word stress and metrical accent (Fig. 1).

ha-	ned	hal-	jas-	ta	hõ-	be-	dat	
+	-	+	-	+	-	+	-	(metre)
+	-	+	-	-	+	-	-	(word stress)

**Figure 1.** Metrical stress (ictus) and word stress patterns in a Kalevala verse line. Pluses denote ictus (middle line) and stressed (bottom line) positions, minuses off-ictus and unstressed positions, respectively.

How is this acoustical conflict solved by a folksong performer? There have been extensive scholarly debates whether the word stress or the metrical accent predominated in original folksong performance (see Tampere, 1937, for a review). Initially, it was hypothesized on the basis of auditory examination of existing recordings that the metrical accent always overrides the expected word stress pattern. This phenomenon has been called scansion among ethnomusicologists. Later, a more liberal conclusion was reached that, in such cases, the two types of accentuation (based either on the metrical accent or the word stress) could be found side by side in different recordings by the same performer, or even in a single performance. No acoustical studies on this issue, however, have been conducted so far. Neither do we know why in those folksongs certain syllables are perceived as stressed while others are not.

In this paper we report the results of an acoustic analysis of three Estonian folksongs, recorded in the 1930s, using a PC-based Kay Elemetrics Computerized Speech Laboratory. The folksongs were performed by a female singer (L.K.) from a dialect area that has only two contrastive stressed-syllable durations, instead of three, which significantly reduces the complexity of the data. A structurally meaningful unit in those songs is a pair of two successive Kalevala verse lines of eight syllables each. In this analysis, we have treated all verse lines as metrically equivalent. The analysed recordings contained a total of 87 standard eight-syllable lines. From this amount, 10 lines had to be omitted because of inadequate quality of the recording. The remaining 77 lines, equalling 616 syllables, constitute the data for the present study.

Normally in such songs a syllable in the text corresponds to a sung tone in the melody, or a single note in the written score. We measured the duration of all syllables using spectrographic representations of the recorded sound signal. In this style of singing, tone/syllable durations are expected to be uniform. In Fig. 2, distribution of all measured tone/syllable durations is presented; the single-peaked distribution confirms this expectation, indicating also that the singer does not maintain the short-long opposition, characteristic to the phonology of her dialect.

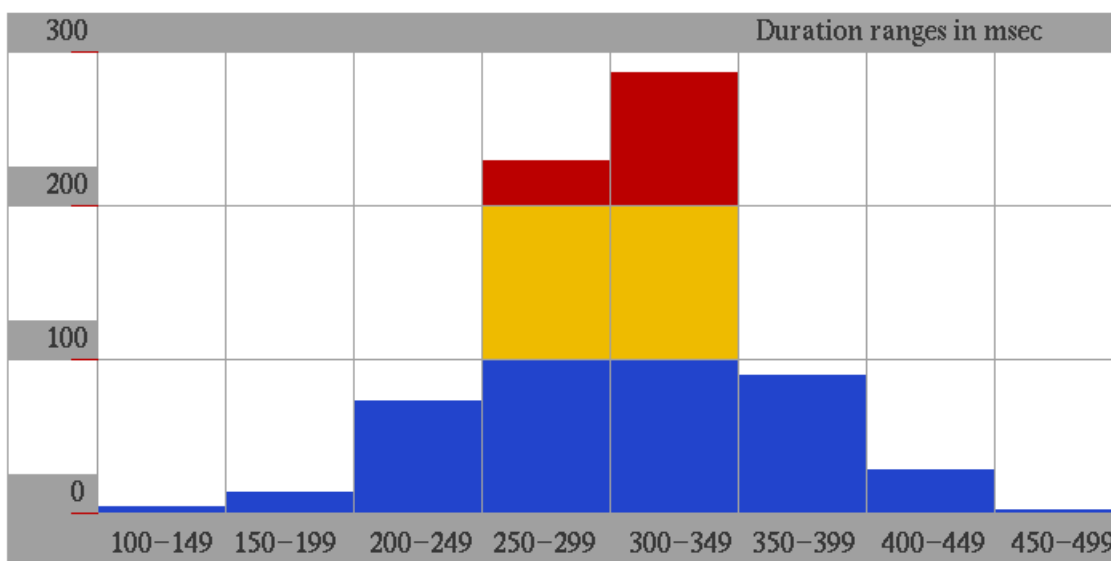


Figure 2. Distribution of syllable/tone durations in three analysed folksongs. Horizontal axis: duration ranges in msec.

It is known that word stress can be manifested in different ways, using the acoustic dimensions of duration, intensity and/or pitch (Lehiste 1970). Tampere (1937) has hypothesized that the conflict between word stress and metrical accent in folksongs may be solved by a performer using a trade-off between duration, intensity, and pitch. He proposed that in ambivalent metrical positions, the word stress may be realized by a singer using the intensity and/or frequency cues, i.e. by performing the syllable louder and/or higher in pitch. The metrical accent he thought was maintained by using the duration cue, i.e., making accented syllables longer than nonaccented ones.

We tried to test this hypothesis by studying whether syllables in ictus position (the odd-numbered syllables) are systematically longer than syllables in off-ictus positions (the even-numbered syllables). The results are presented in Table I. Indeed, the odd-numbered syllables are longer than the even-numbered syllables, the averages being 317 msec and 290 msec, respectively. This difference is significant at  $p < 0.001$ . Another possibility to look at the same issue is to investigate whether the words with their initial syllable falling in an odd numbered position are systematically performed differently from the words whose initial syllable falls in an even numbered position. It is possible to study this using the S1/S2 ratio, i.e., the ratio of the duration of the first syllable to the duration of the second syllable. The S1/S2 ratio has been demonstrated to be a good tool in order to measure the acoustic differences between the Estonian quantity degrees. It yields numerical values for the short, long, and overlong degree in standard Estonian which are significantly different from each other (Lehiste 1968).

Table I. Comparison of syllable durations (msec) and the S1/S2 ratio at ictus and off-ictus positions.

	Syllable duration		S1/S2 for Q1-words	
	ictus	off-ictus	ictus	off-ictus
$\bar{x}$	317.1	289.9	0.994	0.908
$\sigma$	46.5	44.8	0.191	0.161

We have counted the S1/S2 ratio separately for the two categories of short (Q1) words: for those with the initial syllable falling in the ictus position and for those with the initial syllable falling in

the off-ictus position. The results are presented in Table I. The S1/S2 ratio of Q1-words starting in the ictus position ( $N = 51$ ,  $S1/S2 = 0.99$ ) is greater than that of Q1-words starting in the off-ictus position ( $N = 36$ ,  $S1/S2 = 0.91$ ). The difference is significant at  $p < 0.04$ . The words starting with an odd-numbered metrical position were performed with the duration of two syllables almost equal to each other. The words starting with an even-numbered metrical position tend to have the first syllable shorter than the second one. It is to be remembered that the word stress falls on the first syllable in both cases; in the case where the first syllable falls on the even-numbered syllable, there is a conflict between word stress and metrical ictus.

We can conclude that it is possible quantitatively to trace acoustical differences between the two categories of metrical positions in old Estonian folksongs. Those differences show up as the durational contrast between odd- and even-numbered, or ictus and off-ictus positions of the Kalevala metre. This is in good agreement with the hypothesis that the contrast between syllables in ictus and off-ictus positions is realized by the means of duration. Whether or not this contrast should be classified as scansion, however, will depend on how we define the stress in folksongs, i.e., whether the prolonged syllables in ictus position are perceived as stressed (or accented) or not. No systematic study of this aspect has yet been carried through.

Figure 3. Excerpt from the transliteration of one analysed song (Laugaste 1989: 618).

Another factor which contributes to the perception of stress in speech by a listener is the fundamental frequency contour. In Estonian, a stressed syllable normally has a higher fundamental frequency than a following unstressed syllable. Observations on the transliteration of melodies in the analysed songs (see excerpt in Fig. 3) suggest that melody contours there may be shaped by the necessity to signal the word stress to a listener, in order to compensate the loss of durational cues in metrically weak positions. In the six verse lines presented in Fig. 3, no conflict occurs between the metrical accent and the word stress, except in the first line, where the word *tormis* starts in the off-ictus position. This word coincides with a deviation from the invariant melodic pattern: the first syllable of the word is performed at *b* in the melody, not *g*, as expected. It seems as if the performer has deviated from the melody in order to keep the stressed syllable *tor-* in the off-ictus position from being lower in pitch than the following unstressed *-mis* in the ictus position. The melody may be slightly adjusted to avoid placement of stressed syllables in those off-ictus positions where they would carry lower pitch than the following unstressed syllables in ictus positions. It seems that, in the case of a conflict between metrical accent and word stress, the first, i.e., linguistically stressed syllables are required to be accommodated at those melodic positions which are not lower in pitch than the consecutive syllables which are unstressed. No quantitative estimation of this hypothesis, however, is possible on the present material because of very limited number of such samples.

The present study seems to strengthen the conclusion that duration in the Kalevala-songs serves exclusively metrical functions and has lost its word-level functions. The linguistic opposition between short and long syllables is neutralized and subordinated to the metrical structure (Ross & Lehiste 1995). The melody, on the other hand, may be modified to conform to word-level pitch patterns. The role of intensity in establishing both ictus and stress is as yet unexplored. But the existing metrical conventions appear to function efficiently enough to allow for the observed free combination of elements from the text and melody corpora in the classical folksong repertoire.

*This study was partly supported by a travel grant to the first author by the Department of Speech and Hearing, Ohio State University.*

## References

- Laugaste, E. 1989. *Vana Kannel VI: 1. Haljala regilaulud (Old folksongs from Haljala)*. Tallinn: Eesti Raamat.
- Lehiste, I. 1968. Vowel quantity in word and utterance in Estonian. In *Congressus Secundus Internationalis Fenno-Ugristarum* (pp. 293-303). Helsinki: Societas Finno-Ugrica.
- Lehiste, I. 1970. *Suprasegmentals*. Cambridge: MIT Press.
- Palmer, C & Kelly, M. C. 1992. Linguistic prosody and musical meter in song. *Journal of Memory & Language*, 31, 525-542.
- Ross, J. & Lehiste, I. 1995. Lost prosodic oppositions: A study of contrastive duration in Estonian funeral laments. *Language & Speech*, 37, 407-424.
- Sundberg, J. 1987. *The Science of the Singing Voice*. DeKalb: Northern Illinois University Press.
- Tampere, H. 1937. Eesti vana rahvalaulu rütmiprobleemist (The problem of rhythm in the ancient Estonian folksong). *Looming*, 15, 190-198.
- Yung, B. 1983. Creative process in Cantonese opera I & II. *Ethnomusicology*, 27, 29- 47 & 297-318.