

Copyright
by
sally ransom
2022

becoming lyrics: how word prosody and musical meter
negotiate the rhythmic terms of prominence

by

sally ransom, B.A., Linguistics

THESIS

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF ARTS

THE UNIVERSITY OF TEXAS AT AUSTIN

December 2022

**becoming lyrics: how word prosody and musical meter
negotiate the rhythmic terms of prominence**

APPROVED BY

SUPERVISING COMMITTEE:

Scott Myers, Supervisor

Katrin Erk, Supervisor

I think that when linguists discuss and dispute sound length and stress among themselves they would definitely benefit from inviting ethno-musicologists to join them. They could discuss the issues together, and not only based on written records but also sung records. That what is written is fiction. Only that what is sung is truth.

Tormis (2007)

**becoming lyrics: how word prosody and musical meter
negotiate the rhythmic terms of prominence**

sally ransom, M.A.

The University of Texas at Austin, 2022

Supervisors: Scott Myers
Katrin Erk

Table of Contents

Abstract	5
List of Tables	7
List of Figures	8
Chapter 1. Introduction	1
1.1 On words becoming lyrics	1
1.1.1 Phonetics of Estonian Prosody	2
1.1.2 Metrical Structure in Music	2
1.1.3 Metrical Principles of Estonian folksong	3
1.2 Previous Studies with <i>regilaul</i>	5
1.3 The present study	8
1.3.1 Hypotheses	9
Vita	14

List of Tables

1.1	ternary syllable weight contrast	2
-----	--	---

List of Figures

1.1	“Millal saame sinna maale”	4
1.2	notation of “Loomine” performed by Liisu Orik in 1965	4
1.3	music notation of “The King Game” as performed by Liisa Kümmel	5
1.4	The swinging song ‘Kiik tahab kindaid’ analyzed in Ross (1989, 1992)	6

Chapter 1

Introduction

1.1 On words becoming lyrics

To join song and become lyrics, meaningful linguistic material must be modified to fit the strict temporal structure of music, while both poet and performer are tasked with preserving intelligibility sufficient for semantic interpretation of the whole. Thus the rhythm of language must fit into the song's rhythm, but enough of the language's own rhythm must remain in order for the lyrics to have meaning. The study of metrical prosody often focuses on the patterns in the texts independent of the songs. In this paper, I analyze the acoustic-phonetic correlates of linguistic rhythm *within* the context of the song. That is, I aim to see how the text-setting of metrical verses is acoustically realized by performers of an oral folksong tradition. I define metrical as the mapping of the pattern on a frame formed of equal time intervals: such patterns have been demonstrated to be more easily replicated by humans, something that is necessary for both the synchronization of musicians performing together and for the transmission of an oral tradition of songs (Essens & Povel, 1985). Previous studies have compared acoustic measurements with transcriptions of previously not annotated music. The problem with these transcriptions is that they “cannot be wrong,” in some cases, the transcriber

was noncommittant to a measure’s value; in others, the subjectivity of human perception is the obstacle. In the present study, I use the annotations as a starting point for each song, but introduce the use of beat-tracking algorithms to provide a rigorous definition of the location of beats, while simultaneously bringing both replicability and falsifiability along.

1.1.1 Phonetics of Estonian Prosody

Q1	sada <i>‘hundred’</i>	kabi <i>‘hoof’</i>
Q2	saada <i>‘send’</i>	kapi <i>‘of the cupboard’</i>
Q3	saada <i>‘recieve’</i>	kappi <i>‘into the cupboard’</i>

Table 1.1: ternary syllable weight contrast

Estonian is famous for its ternary quantity distinction: both vowels and consonants have three distinct lengths or degrees. As can be seen in the examples in 1.1, the length of a given segment can indicate lexical contrasts as in sada vs saada (‘hundred’ vs ‘send’) (root minimal pairs) and also indicate case: the difference between ‘receive’ and ‘send’ is in the quantity of the first syllable’ or directional case marking ‘of’ versus ‘into’ the cupboard.

1.1.2 Metrical Structure in Music

In music, the smallest prosodic constituent is an individual note event whose relationship to the other notes in the song are indicated by the time signature, i.e., $\frac{3}{4}$ and $\frac{4}{4}$. The denominator corresponds to the number of di-

visible beats of a “whole” note (♩), while the numerator refers to the number of “beats” in a single measure. In $\frac{4}{4}$, a whole note is sustained for the same duration as four quarter notes (♩) in the same measure, and each measure must culminate in enough notes and rests to equal a whole note.

1.1.3 Metrical Principles of Estonian folksong

Estonian *regilaul* is part of the Finnic runosong tradition shared by several other members of the Finnic language family: Finnish, Karelian, Votic, Ingrian, and Livonian (Ross & Lehiste, 2001). The metrical basis of the tradition is a trochaic tetrameter often referred to as the Kalevala meter (Oras, 2019), which is realized in Estonian 20th century work as syllabic-accentual trochaic tetrameter (Lotman & Lotman, 2013).

Each verse line has four beats with eight syllable-note positions which can also be occupied by rests or, in the case of trisyllables, two sixteenth syllable-notes. The “beat” or “ictus” position falls on the note corresponding to the first beat of a measure, and on every other following syllable-note in the invariant form of eight eighth notes. Runic songs that follow quantity rules for trochaic meter oppose metrically strong and weak positions by means of syllable quantity and stress. Ictus position, or on the beat, prefers syllables that are both long and stressed but avoids short stressed syllables: these can occur off the beat, while this position is avoided by long syllables. These “singable songs” (Tormis, 1985) follow a metrical pattern such that a given *regilaul* text can be sung to any of the numerous *regilaul* melodies (Ross &

Lehiste, 2001).



Figure 1.1: “Millal saame sinna maale”

1.1 illustrates the invariant pattern of eight syllables notes each occupying one eighth of a $\frac{4}{4}$ measure.



Figure 1.2: notation of “Loomine” performed by Liisu Orik in 1965

1.2 illustrates a melody variation with seven syllables: in the first verse, the last (heavy) syllable is extended to fill a quarter note, in the second the last syllable is sung as an eighth note and followed by a quarter rest (7).

How do the word-prosodic requirements negotiate with the imposed prosodic hierarchy of music? The rhythmic organization of song is said to integrate the prosodic structure of the language with musical rhythmic principles (Palmer & Kelly, 1992). However, earlier studies of *regilaul* explored

temporal aspects of the songs and found that duration characteristics that would usually indicate important semantic differences lost their distinctions partially or entirely. Assuming that the intention of the singer is for the lyrics to be understood, I hypothesize that if some durational correlates of contrasting word-prosodic constituents are made less distinct in the process of compromising with the song that some other acoustic correlate of the relevant contrast at the word-prosodic level will be present, if not enhanced.

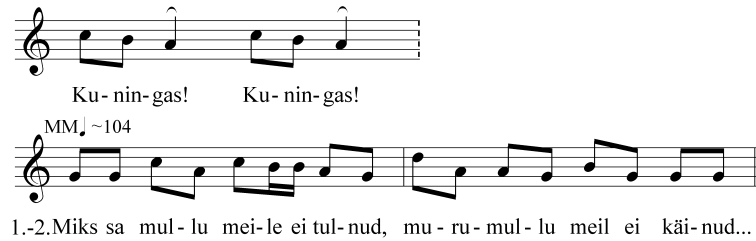


Figure 1.3: music notation of “The King Game” as performed by Liisa Kummel

1.2 Previous Studies with *regilaul*

The intuitions of those who study the runosong tradition is that the burden of upholding the temporal structure of the song is the result of symbiosis between the musical rhythm and the natural prosodic features of the lyrical text (Ross, 1992; Tampere, 1934): the song’s melody a musical abstraction of the natural prosody of spoken runic verse inspiring decades of research at the interface of metrical phonetics and computational musicology Rüütel (1999).

Swinging songs are characterized by a swinging rhythm with alternating



Figure 1.4: The swinging song ‘Kiik tahab kindaid’ analyzed in Ross (1989, 1992)

long and short notes, differing from the main body of *regilaul*’s iconic isochrony. Jaan Ross, a musicologist and native speaker of Estonian analyzed a 1936 recording of the swing song ‘Kiik tahab kindaid’ 1.4, publishing results on syllable-note duration Ross (1989) and later the vowel quality of odd-numbered syllables (Ross, 1992). In the study on syllable-note durations,

In the second, Ross measured formant frequencies f1 and f2, finding a reduced vowel space in song compared to measurements in spoken Estonian. However, upon examination of the song, it is clear that all the vowel space measurements are from syllable-notes in non-initial positions of Estonian words: that is, the sample of vowels taken from the song were all unstressed, and compared to a mixed sample of spoken Estonian. Thus, the conclusion needs to be evaluated again with comparable samples.

In 1992, Ilse Lehiste asked “Whether there is a correlation between poetic metre and the prosodic structure of a language. (Lehiste, 1992) by means of measuring the acoustic-phonetic realizations of so-called trochaic metrical po-

etic patterns across several languages including Estonian and Finnish. While these languages share in the more general Balto-Finnic tradition of *runosong* utilizing what is called the *Kalevala* meter, there were significant differences in the phonetic realizations of trochees in each language.

In 1994, their collaboration begins. Ross and Lehiste published several papers examining the temporal dimensions of Estonian word prosody and metrical prominence in *regilaul* folksongs. In (Ross & Lehiste, 1994), they conclude that duration differences ordinarily present at the word level (stressed-unstressed) are “lost” to the temporal restrictions of the song. In another paper, syllable-notes are again measured, this time examining the role of syllabic quantity in the song. They likewise conclude that the duration of syllable-notes in *regilaul* match more closely with the metrical structure of the songs, that is, the durations of syllable-notes are best predicted by their beat position in the song: on the beat, syllables are longer, off the beat shorter. They extend this finding to conclude that the song “dominates” the metrical status of the words, claiming that the intelligibility of lyrics is enhanced more strongly by “top-down” processes (i.e., semantic context).

(Ross & Lehiste, 1996).

(Ross & Lehiste, 1998) Analyzed and concluded that the *regilaul* lyrics are the result of an interaction between word and song prosodic hierarchies. This conclusion relied critically on measuring the durations of syllable-notes, where they found being on or off the beat was the better predictor for duration.

Finally, in (Ross & Lehiste, 2001) summarize their body of work until that point and extend their findings with fine-grained acoustic phonetic measurements of segment durations within syllable-notes. In their discussion, they mention that some of the durational differences not present at the syllable-note level are present at the segmental level, in complex codas.

1.3 The present study

To extend the findings of this body of work, the present study examines two acoustic correlates of prominence at the segmental level: namely, the syllable nucleus. Because the vowel is the most sonorant part of a syllable, it is the most acoustically and perceptually salient portion of the syllable. Measuring the vowel duration will also offer indirect information about the rhyme as a whole: the presence of codas and complex codas should have an effect on the vowel duration in isochronous syllable-note sequences. In the cases where quantity is distinguished by coda consonant length, the syllable nucleus (measured by vowel duration) will be necessarily shorter to accommodate the geminate or complex coda. In cases where quantity is indicated by the length of the vowel, the opposite should be true. This study therefore examines the ternary quantity distinction in the context of its syllable shape: no coda (CV), single coda (CVC, CVVC), and complex coda (CVCC, CVCCC) rather than collapsing all syllable shapes according to their quantity. This allows a closer look at the microprosodic features at the segmental level.

In addition to vowel duration, I also include vowel dispersion measured by the euclidean distance from the center of each singer’s vowel space on the (f1, f2) plane. A larger vowel perimeter generally corresponds to hyper-articulation or clear speech, while a smaller perimeter with hypo-articulation or reduction Lindblom (1990); Smiljanić & Bradlow (2005). In natural Estonian speech, reduction is only allowed on unstressed syllables, with /i/ being the most resistant (Eek & Meister, 1998). This measurement is included for two reasons: one, the acoustic correlates of prominence in both language and music are almost always not a single cue but a convergence of several cues, and two, to see if the findings of (Ross, 1992) extend to the style of *regilaul* songs that make up the largest portion of the body of work (i.e., non-swinging songs).

More recently, syllabic-accentual trochaic tetrameter has received some statistical attention to its textual structure (Lotman & Lotman, 2013)

and to rhythmic variation in singing (Oras, 2019).

1.3.1 Hypotheses

Null hypothesis: on or off-beat position is best predictor for vowel duration of syllables in all categories: stressed, unstressed, short, long, and overlong. This would mean that the syllable nucleus is proportional to the total syllable duration, and that the duration contrast is indeed “lost.”

H₁: duration contrasts for syllable quantity will be evident in the vowel duration, with vowel duration *decreasing* as syllable weight increases. Given the findings of (Ross & Lehiste, 2001), isochronous syllable-notes would result

in heavier syllables having shorter nuclei to accommodate for the coda and complex codas that distinguish the syllable weights from each other.

H₂: stress/unstress contrasts will be evident in the nucleus in terms of hypo and hyper articulation. For this I measure both nucleus duration and vowel space dispersion.

Bibliography

- Eek, A., & Meister, E. (1998). Quality of standard Estonian vowels in stressed and unstressed syllables of the feet in three distinctive Quantity degrees. In *Proceedings of the Finnic Phonetics Symposium*, Linguistica Uralica. Tallinn.
- Essens, P., & Povel, D.-J. (1985). Metrical and Nonmetrical representations of temporal patterns”. *Perception & Psychophysics*, 37, 1–7.
- Lehiste, I. (1992). The Phonetics of Metrics. *Empirical Studies of the Arts*, 10(2), 95–120.
- Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In W. J. Hardcastle, & A. Marchal (Eds.) *Speech Production and Speech Modelling*, (pp. 403–439). Dordrecht: Springer Netherlands.
- Lotman, M.-K., & Lotman, M. (2013). The Quantitative structure of Estonian syllabic-accentual trochaic tetrameter. *TRAMES*, 3, 243–272.
- Oras, J. (2019). Individual rhythmic variation in oral poetry: the runosong performances of Seto singers. *Open Access Linguistics*.
- Palmer, C., & Kelly, M. H. (1992). Linguistic Prosody and Musical Meter in Song. *Journal of Memory and Language*, 31(4), 525–542.

- Ross, J. (1989). A study of timing in an Estonian runic song. *The Journal of the Acoustical Society of America*, 86(5), 1671–1677.
- Ross, J. (1992). Formant frequencies in Estonian folk singing. *Journal of the Acoustical Society of America*.
- Ross, J., & Lehiste, I. (1994). Lost Prosodic Oppositions: A Study of Contrastive Duration in Estonian Funeral Laments. *Language and Speech*, 37(4), 407–424.
- Ross, J., & Lehiste, I. (1996). Trade-off between quantity and stress in Estonian folksong performance? *Folklore: Electronic Journal of Folklore*, 02, 116–123.
- Ross, J., & Lehiste, I. (1998). Timing in Estonian Folk Songs as Interaction between Speech Prosody, Meter, and Musical Rhythm. *Music Perception*, 15(4), 319–333.
- Ross, J., & Lehiste, I. (2001). The temporal structure of Estonian runic songs / by Jaan Ross, Ilse Lehiste. In *The Temporal Structure of Estonian Runic Songs*, Phonology and Phonetics ; 1. Berlin, [Germany] ;: Mouton de Gruyter, reprint 2015 ed.
- Rüütel, I. (1999). Some results of a computerized comparative analysis of the Balto-Finnic runotunes. *Etnomusikologian vuosikirja*, 11, 27–45.

- Smiljanić, R., & Bradlow, A. R. (2005). Production and perception of clear speech in Croatian and English. *The Journal of the Acoustical Society of America*, 118(3 Pt 1), 1677–1688.
- Tampere, H. (1934). Mõningaid mõtteid Eesti rahvaviisist ja selle uurimismetodist. *Eesti Muusika Almanak I*, (pp. 30–38).
- Tormis, V. (1985). Kalevala – the Estonian perspective. *Finnish Music Quarterly*.
- Tormis, V. (2007). Some problems with that *regilaul*. In *proceedings of the international RING conference*.

Vita

Permanent address: 305 E. 23rd Street STOP B5100
Austin, Texas 78712

This thesis was typeset with \LaTeX^\dagger by the author.

[†] \LaTeX is a document preparation system developed by Leslie Lamport as a special version of Donald Knuth's \TeX Program.