

CORPUS-BASED RESEARCH OF SEMANTIC ASPECTS OF *LAUL*-STEM WORDS IN ESTONIAN, FOCUSED ON PAST NEWSPAPER TEXTS AND FOLK SONG LYRICS

Taive Särg

Senior Research Fellow

Estonian Folklore Archives

Estonian Literary Museum, Estonia

taive.sarg@folklore.ee

Kaarel Veskis

Extraordinary Junior Research Fellow

Estonian Folklore Archives

Estonian Literary Museum, Estonia

PhD Student, Institute of Estonian and General Linguistics

University of Tartu, Estonia

kaarel.veskis@kirmus.ee

Abstract: This study examines conceptions of singing across various Estonian cultural spheres through corpus-based analysis of the singing-related vocabulary. The words with the stem *laul* ‘song’/‘sing’ are analysed in five linguistic corpora, with a detailed focus on two of them: the Estonian Media Corpus (EMC), comprising 1890s newspaper texts from the Corpus of Written Estonian (Eesti Kirjakeele Korpus: 1890ndad; see CWE 2018), and the Regilaul corpus from the Estonian Runosongs’ Database (Eesti Regilaulude Andmebaas, ERAB). Our methodology combines quantitative word analysis with qualitative examination of lexical categories and semantic associations. The analysis reveals differences in the usage and frequency of the singing-related vocabulary between the two main corpora, offers new insights into singing in Estonian society and shows how distinct conceptualisations of singing and modes of expression in oral and written traditions are reflected in the vocabulary.

Keywords: computational folkloristics, historical media analysis, singing concepts, ethnomusicology, corpus linguistics

1. INTRODUCTION

Scholarly discourse on singing traditionally provides insights into music history primarily from an academic perspective. A large proportion of the scholars receive Western art music education, which influences their analytical frameworks. The twentieth century saw the emergence of disciplines like ethnomusicology and popular music studies, broadening the scope of musical scholarship. However, to gain a more comprehensive understanding of the ways and roles of singing in society, it is valuable to explore alternative discourses that exist outside formal academic boundaries.

Various texts can offer alternative viewpoints and valuable insights into attitudes toward singing. Textual analysis, especially if combined with documented music examples, helps delineate concepts of singing across different social and cultural contexts. A group's musical concepts encompass their ideas about and attitudes towards music. Text corpora representing various social discourses, including those peripheral to music, can shed light on these diverse conceptions of singing. While Estonian folk and national music are quite well-studied, the scholarly writing on people's everyday music habits (e.g., Raud 1912; Vissel 2004; Lippus 2006, 2012; Särg 2009, 2023), subcultures (Davidjants 2022, 2024) and grassroot perspective remains more limited. In the case of the *regilaul* tradition, the “common” singers' perspective has received greater attention (Kõiva 1964; Kalkun 2014; Kalkun & Oras 2018; Oras 2008, 2017; Rüütel & Tiit 2005–2006; Rüütel 2013; Saarlo 2023). The semantics of the folk song vocabulary has been studied earlier in not numerous but thorough works (Peegel 2004; Labi 2006; Saarlo & Sarv & Mett 2022).

Large text corpora require the development of methodologies to examine their semantic content. Computational linguistics provides methods like vocabulary analysis, discourse analysis, and corpus analysis to extract information from texts. Teubert (2009: 53) articulates the goal of corpus linguistics: “to find out what a text, a text segment or a lexical item means in a specific context, at a given time”.

This study draws on the works of ethnomusicologist Allan Merriam (1964), scientific cognition theorist Keith Taber (2019–2024), and linguist Wolfgang Teubert (2009). Concepts are mental constructs that label and describe phenomena, simplifying the diversity of the world through generalisations (Rickert 1902 [1896]: IV, 32–33). Taber distinguishes between concepts, conceptions, and the words that designate them. A concept is a general idea shared by a group, while words serve to identify the concept. The words used more specifically within a discipline become terms. Taber uses the word ‘conception’ to refer to an individual's interpretation of a concept.¹ He also notes that concepts are

relational – our understanding of a concept involves its associations with other linked concepts² (Taber 2019–2024).

Merriam introduced the ‘concept of music’ in ethnomusicology, drawing on Vinigi Grottanelli’s idea of an “underlying system of conceptions” in visual arts (Merriam 1964: 34). Within the framework of cultural relativism, music is seen as a universal human activity that manifests in culturally specific ways, shaped by a group’s underlying concepts.³ Concepts, being mental entities, can be studied indirectly through their manifestations in sounds, speech, and writing. While the existence of song and singing concepts is evident in the words themselves, their content and meaning – how people perceive singing and relate it to other aspects of music and life – require further study.

Concept study is central to cognitive linguists like Richard Langacker (2002 [1990]) and is a key task in linguistic semantics (Teubert 2009: 36). Teubert stresses that meaning is social and emerges through communication, discoverable only in text collections: “An approach to language from a social perspective will analyse a discourse as a collective mind” (ibid.: 47). Our study explores singing concepts in Estonian text collections representing various social discourses.

Historical music in Estonia can be broadly categorised into two main parts reflecting societal structure: 1) Estonian peasants’ traditional music with Finnic cultural roots, and 2) art, church, and folk music used by the local upper class of Indo-European origin. These two categories interacted, with Western music influencing local Estonian music.

The Estonians’ ethnic groups have been inhabiting the Baltic shores for at least 2,000 years, and likely blended in with the earlier population of this area (Saag et al. 2019). Estonians consider themselves indigenous (*põlisrahvas*; *põline* ‘traditional, indigenous, age-old’) due to their long-standing presence and close relationship with the local environment, despite cultural and genetic influences over time.⁴ Estonian tribal territories were conquered by Northern Crusaders by 1227, leading to the baptism of the people, gradually depriving locals of their political and economic right to self-determination and culminating with the development of serfdom under foreign landlords by the eighteenth century. The Estonian territory was ruled by various imperia, including the Russian Empire (1710–1919) with Baltic German landlords maintaining privileges over Estonian peasants, and later the Soviet Union (1940–1941/1944–1991).⁵

For centuries, the population on the Estonian territory was divided primarily by ethnicity and language, with Estonians being a lower class. Estonians maintained a blend of pagan and Christian beliefs, along with the ancient Finnic *regilaul* tradition, into the early 1900s and partially to the present day. Nevertheless, local traditional music interacted with neighbouring musical traditions and Western folk, church, and art music, with Baltic German music

exerting a significant influence. In the nineteenth century, society in Estonia started to modernise, influenced by broader economic and social changes across Europe. Serfdom was gradually abolished with the reforms between 1816 and 1868, while national consolidation began in the 1860s. During this period, Estonians' ancient community practices and singing traditions diminished, while choral singing gained central importance in the formation of modern national identity. The late nineteenth and early twentieth centuries saw both extensive folklore collection and the establishment of modern Estonian-language culture. These processes culminated in the formation of the independent Republic of Estonia in 1918.

This article presents a collaborative study by ethnomusicologist and folklorist Taive Särg and language technology researcher Kaarel Veskis, examining the evolving concepts of singing throughout Estonia's cultural history. As native Estonian speakers with advanced philological training, we bring a nuanced understanding of both contemporary and historical language use to this research. Our complementary backgrounds in musicology further inform our analysis. The following chapters introduce Estonian singing-related vocabulary (Ch. 2), our research methodology and selected text corpora (Ch. 3–4), and present the results of a detailed analysis of two text corpora (Ch. 5–6) with a further comparison (Ch. 7).

2. THE CONCEPTS OF SONG AND SINGING IN ESTONIA

In many indigenous cultures, including Estonian, singing and instrumental music were distinct activities, with no generic term for 'music' until the nineteenth century (cf. Blacking 1995: 224). In Estonian colloquial language, the term *muusika* (music) still remains more closely associated with instrumental music, as illustrated by this media quote: "We believe that through music and singing, a child's soul can be made to sound" (Mamma Mia).⁶

The ancient Estonian verb *laulma* 'to sing' remains common and serves as the root for derivatives such as *laulmine* 'singing', *laul* 'song'⁷, and *laulja* (singer). Historically, *laulma* belongs to the Finno-Samic stems (Metsmägi & Sedrik & Soosaar 2013: 318). However, the practices and sounds denoted by *laul* vary considerably. Some archaic vocal styles (e.g., laments, herding calls) were not considered 'songs' or 'singing' even by practitioners (Rüütel 1998), while outsiders and even the Estonian intelligentsia sometimes questioned whether *regilaul* was a proper song (Jannsen 1857: 111), or occasionally using pejorative terms like 'screamy singing' (e.g. Schlegel 1831: 144).

Similar phenomena exist in other cultures, particularly for styles where lyrics predominate. For instance, early Welsh descriptions portrayed the singing style of hymns as pleasant and harmonious, while the strongly declamatory style of bards was disparagingly termed ‘jogging’ (Kinney 2015 [2011]).

While *laul* appears in all dictionaries and encyclopaedias, some studies on musical terminology focus exclusively on art music (Vahter 2003; Lock 2019). *Eesti keele seletav sõnaraamat* (EKSS; The Explanatory Dictionary of the Estonian Language, 2019) offers comprehensive entries for singing-related terms, reflecting their diverse uses and meanings. It defines *laul* as: 1) human vocal musical expression and characteristic bird sounds, as well as sounds from other living and non-living sources; 2) a piece of vocal music, singing, and (in older Estonian) poetic writing; 3) a component in various idiomatic expressions with metaphorical meanings beyond those mentioned above.

The historical derivative *laulik* can denote both a person (‘singer’ or ‘poet’) and an object (‘songbook’) (EKSS). In oral tradition, a *laulik* functioned as both the (re)creator and performer of songs. With modernisation, these roles diverged, and *laulik* came to denote only a poet or a collection of written songs. A new compound, *rahvalaulik*, was coined for ‘folk singer’, defined as “a mediator of folk song tradition, a talented improviser, a creator of simple folk songs and verses, and a beloved poet of the people” (EKSS). In poetic language, *laulik* can also mean ‘songbird’. The dictionary definitions reflect two distinct concepts of creation: “folk songs in a simple form” are “created”, while “folk song tradition” is “mediated” and “improvised” (EKSS).

Another derivative for a ‘singer’, an agent noun *laulja*, is defined without a reference to song creation. *Lauljanna*, a neologism denoting a female singer, incorporates the newer gender suffix *-nna* (EKSS).

Ingrid Rüütel has provided an overview of singing-related terms in Estonian traditional music, concluding: “The Balto-Finnic peoples regard as a song (*laulu*) in general a vocal genre which consists of a poetic text and a melody with a certain structure. The word *laulu* is common to all Baltic Finns who know the runo-song” (Rüütel 1998: 36).⁸ She brings out that several song genres, such as laments, spells, herding calls, recitative children’s songs and incantations were not referred to as songs, nor was their performance called singing, but they were designated by special, often descriptive linguistic expressions, e.g., *huikama*, *hellatama*, *itkema*, *lausuma*, etc.⁹ (ibid.). The sounds of certain birds and their imitations are also called songs (ibid.).

In conclusion, ‘song’ and related terms encompass a broad semantic field. We are interested in how singing-related words are used in their original contexts and what information they provide about songs and singing.

3. METHODOLOGY

Given that concepts are mental constructs, we employ indirect methods to study them, specifically analysing the meanings of song/sing-stem words in texts. In linguistics, ‘meaning of the word’ encompasses both the underlying idea or intention and the referent of a linguistic unit. Semantics examines the basic, literal meanings of words, while pragmatics focuses on their practical usage (Riemer 2010: 2, 22). Meaning is a broad category, related to the concept of lexical field as “a structure formed by lexemes”, and semantic field as “the underlying meaning which finds expression in lexemes” (Wylter 1992: 30; cf. Nordquist 2019b). For the concept of singing, the underlying meaning may manifest in sounds, as well as in synonyms, descriptions, and other linguistic elements. Meaning interpretation involves generalising from the specific meanings words acquire in unique contexts (Teubert 2009: 5). Words related to music (singing) represent cultural agreements within a language community. The same term, when used by different individuals or communities, may have varying relationships to its object and carry diverse aesthetic connotations or attitudes.

Text corpora are primarily used to investigate word frequencies, lexical and grammatical pattern distribution, neologisms, diachronic changes, meanings, collocations, lexical-semantic relationships, example sentences, translations, definitions, terms, and expressions (Koppel 2020). Research questions and methods for meaning analysis should incorporate semantic considerations. Teubert explains that according to principles of corpus linguistics “[t]he constitution of meaning is the result of linking together words in a given context, repeating what has been linked together time and again in the discourse, and thus in the corpus” (2009: 19). Bennett (2010: 1–4) notes that corpus linguistics is well suited for comparing different corpora, such as spoken versus written language, or for contrasting a specific text corpus with a standard language corpus.

Ulrike Oster (2010) highlights the advantages of applying specific corpus-based analysis procedures, such as combining lexical approaches with metaphorical pattern analysis and incorporating concepts like semantic preference and prosody, particularly for analysing emotion words. This approach enables exploration of a search word’s context within large text volumes to delineate a concept’s semantic field. Analysing meanings in text corpora seeks a balance between rapid automatic processing of big data and result interpretation, often requiring close reading, comparing, and reasoning.

The analysis’s effectiveness also depends on the specific words and text corpus, considering factors such as dialectical and spelling variations.

We will examine different text corpora to uncover how singing-related words were used, what they reveal about historical singing practices, and any

additional contextual or other information embedded in these words. We extract and quantify singing-related word classes, compounds, and collocations, compare their normalised frequencies across various corpora, and analyse the information revealed through word associations. Frequent primary words are often considered more common and “less marked” compared to “marked” terms such as compounds¹⁰ (Nordquist 2019a). An example of a marked term is the aforementioned archaic word for a (folk) singer, *laulik*, which evolved into the more specific ‘folk singer’ (*rahvalaulik*) in the era of literary culture.

The word *laul* was selected for analysis due to its musical significance and linguistic structure. *Laul* serves as 1) a fixed stem derived from the verb *laulma* (to sing), with *-ma* as the infinitive marker; 2) a productive element in compounds such as *rahvalaul* (folk song) or *laulupidu* (song festival); 3) a common word across all social strata; and 4) a historical Estonian word with minimal dialectal and historical variation.

The verb derivative *laulatama* (to wed in a Christian ceremony) and its noun form *laulatatus* (church wedding ceremony) were excluded, as their current meanings no longer primarily relate to singing. This word reveals the historical connection between singing and ecclesiastical marriage in Estonian. The word is likely derived from the historical verb *laulama*, with the addition of the suffix *-ta-*, which renders it transitive.

We focused on texts representing both written and oral culture, as well as the transitional period between Estonian oral and literary culture in the latter half of the nineteenth century. For our initial analysis of song/sing-stem words, we selected four distinct historical text corpora and counted the frequency of song/sing-stem words for them. The next steps were extracting sentences (or verse lines) containing the stem *laul*, categorising parts of speech, analysing derivatives and compounds, counting various word frequencies, and conducting comparative analysis across corpora. Words were analysed by four categories: 1) root word *laul*, 2) derivatives, 3) compounds of various structures, and 4) verb *laulma*.

4. THE STATISTICAL COMPARISON OF THE FOUR TEXT CORPORA

This article presents statistical data on the four corpora, with a more in-depth analysis of the corpus of Estonian media from the 1890s and the corpus of the Estonian *regilaul* (‘runosong’) texts from the Estonian Runosongs’ Database.

1. The Estonian Runosongs’ Database (ERAB) was developed in collaboration with the Finnic runosongs’ database SKVR-Tietokanta. This corpus reflects ancient Estonian poetic language, with the older parts of *regilaul* predating written Estonian.

2. *Vana Kirjakeele Korpus* (The Corpus of Old Literary Estonian, VAKK) encompasses written texts from the fifteenth to the nineteenth centuries. It includes all extant texts from the fifteenth and sixteenth centuries, most printed texts from the seventeenth century, and a selection from the eighteenth and nineteenth centuries. This corpus uniquely represents written texts from this period, reflecting the early stages of written Estonian, but also some aspects of folk life.

3. The 1890s Estonian Media Corpus (EMC) consists of selected Estonian newspaper texts and is a part of the larger Corpus of Estonian Literary Language (CELL 2018). It was compiled by researchers at the University of Tartu (Hennoste et al. 2001). We accessed the texts of the EMC through the KORP query interface,¹¹ which was managed by the Center of Estonian Language Resources (CELR).¹²

4. Folklorists from the Department of Folkloristics of the Estonian Literary Museum created a text corpus from calendars published between 1854 and 1938. These calendars served both educational and journalistic functions for rural populations. The texts were selected to the corpus based on folklorists' interests, including both folklore pieces and writings about folklore.¹³

5. For comparative purposes, a joint Estonian National Corpus (containing 1.5 billion words as of 2019) serves as a reference average corpus (Koppel & Kallas 2020).

Table 1. Frequency of laul-stems in five Estonian text corpora. Percentages are calculated based on total tokens and/or words, depending on corpus characteristics. Ø indicates unavailable data. M = million, B = billion.

	Regilaul corpus	VAKK	EMC	Old calendars	Estonian National Corpus 2019
Laul-stems	29,140	895	514	749	849,297
Total of tokens / words	8.9 M / 7.5 M	2.7 M / Ø	227,365 / 193,000	Ø / 961,419	1.83 B / 1.5 B
Percentage	0.33% / 0.39%	0.03% / Ø	0.22% / 0.27%	Ø / 0.08%	0.05% / 0.06%

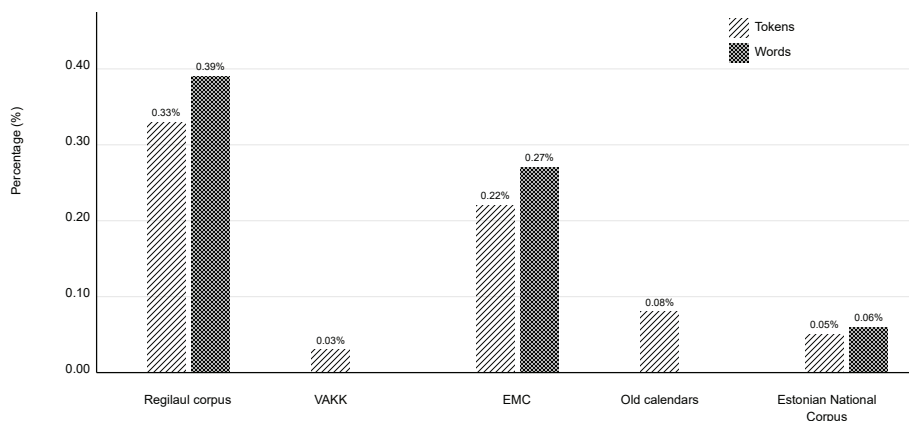


Figure 1. Frequency of laul-stems in five Estonian text corpora.

Laul-stems frequency in a corpus indicates the prevalence of singing-related topics and may reflect curators' selection criteria. The total volume of material in the corpora was reported differently by their maintainers, either as word count or as token count. Furthermore, the definition of what constitutes a token varied across corpora.¹⁴ Nevertheless, the data remain comparable as corpora reporting both metrics show minimal difference in the proportion of song/sing-words between token and word counts (Table 1, Fig. 1). *Laul* percentage in old literary texts (0.03% of tokens) and calendars (0.08% of words) approximates the Estonian National Corpus average (0.05% of tokens, 0.06% of words). As shown in Figure 1, the frequency of *laul*-stems varies across the five text corpora. The percentage of *laul* is approximately 6.5 times higher in the Regilaul corpus (0.33% of tokens, 0.39% of words) and 4.4 times higher in the EMC (0.22%) compared to the Estonian average (0.05% of tokens, 0.06% of words). Consequently, we selected these two corpora for a more detailed analysis.

5. THE 1890S ESTONIAN MEDIA CORPUS

5.1 Overview of the corpus and the 1890s media

Our study of the singing-related vocabulary in late nineteenth-century Estonian texts must be viewed within a broader context of Estonia's national awakening. This period saw efforts to build a modern society, written culture, and arts aligned with European standards. These developments were crucial for

improving the rights and economic conditions of native Estonians, who had long faced economic and cultural marginalisation under colonial rule.

The evolution of European musical culture was pivotal in this modernisation. Three all-Estonian song festivals in five years (1890, 1894, and 1896) underscore the role of choral singing in national identity formation (Kuutma 1996). Concurrently, Estonia, following a broader European trend, began the systematic collection of folklore, viewing it as a foundation for national culture. These activities occurred against intensifying Russification efforts by the Russian Empire in the 1890s.

The Estonian-language press played a key role in spreading democratic and enlightening ideas to the largely rural lower-class Estonian population. Anu Pallas wrote that by the late nineteenth century, the Estonian press became “the centre of promoting public education and social life. ... At the same time, showing loyalty to the tsar was completely natural in the circumstances at that time and characterised all the editors of the people’s newspapers” (Pallas 2000; see also Lauk & Pallas 2008).

The EMC reflects this multifaceted historical context with several local news about choir concerts and social life that Mart Laar (2001) claimed to have been crucial in promoting a shared Estonian identity.

The KORP web page’s KWIC (Keyword in Context) interface, which we used for the analysis, references a subset of the text sources that are available online.¹⁵ Through this resource, we were able to verify the inclusion of the following 1890s newspapers in the corpus: *Postimees* (est. 1857), *Eesti Postimees* (1864–1905), *Olewik* (1882–1915), *Valgus* (1880–1906), *Sakala* (est. 1878), *Ristirahwa Pühapäewaleht* (The Christian Sunday Newspaper) (1875–1919), *Wirmaline* (1887–1898). It is worth noting that, in line with the journalistic practices of the era, most newspaper articles from this period were published anonymously.

5.2 *Laul*-stem words in the EMC

Analysis of singing in newspaper texts began with creating concordances from the texts of the EMC, using the KORP interface, followed by automatic quantitative analysis and manual content analysis. The *laul*-stem words were classified into lexical categories. First, we examined words neighbouring *laul*, many of which expanded our list of compounds. In Estonian, the distinction between solid (combined) and separate writing is often subjective and variable. As a result, certain text units can be interpreted either as a noun with its attribute or as a compound word. Due to unstandardised orthography in the corpus, words with the same meaning appeared in different forms (e.g., *kirikulaul* and *kiriku laul*

for ‘church song’). We standardised them according to modern orthography.¹⁶ We ultimately identified six-word classes, which are presented in Tables 2 and 3.

The EMC contains 501 *laul*-stem words, after excluding *laulatama*-words and a typographical error from an initial count of 514.¹⁷

Table 2. *Laul-stem words in the Estonian Media Corpus.*

Word	Word class	Frequency	Examples	Meaning	Total	% of all <i>laul</i> -words	% of all <i>laul</i> -words
<i>laul</i>	primary word	164		song, poem, activity	164		32.7%
	compound, second - <i>laul</i>	73	<i>tänu laul, soolo-laul, naljalaul</i>	thanks-giving song, solo song, joke song	207	14.6%	41.3%
	compound, first <i>laulu</i> -	114	<i>laulu pidu, laulukoor, lauluhääl</i>	song festival, singing choir, singing voice		22.8%	
	3-part words of different structure	20	<i>laulu- ja käsiraamat</i>	song- and handbook		4.0%	
<i>lauluke</i>	diminutive	3		little song	3	(0.6%)	(0.6%)
<i>laulmine</i>	gerund	16		singing	18		3.6%
	compound, second	2	<i>kokku-laulmine</i>	singing together			
<i>laulja</i>	agent noun	38		singer	40		8.0%
	compound, second	2	<i>kunst-laulja</i>	classically trained singer			
<i>laulik</i>	agent noun	11		poet, bird, songbook	12		2.4%
	compound, second	1	<i>laste-laulik</i>	children’s songbook			
<i>lauljanna</i>	agent noun	2		female singer, songstress	2		(0.4%)
<i>laulma</i>	verb	54		to sing	54		10.8%
Total					501		

Laul denotes both poem lyrics and songs in the language of the time. The diminutive form *lauluke* occurs rarely (3) and is used exclusively to refer to poems. For example: “Only for the sake of good, he tied some little songs that had arisen in the course of time, into a bundle”¹⁸ (interpreted as: “The poet completed a book of poems”).

The term *laulik* in Estonian covers two English terms: ‘poet’ (*luuletaja*) and ‘singer’ (*laulja*), as discussed earlier (Ch. 2). When describing Estonian authors (11 times), and one Latvian author, *laulik* is used to mean ‘poet’. It also serves as a synonym for ‘poet’ in texts where the word *luuletaja* is used, regardless of the author’s nationality (10 times). However, in the four cases where only *luuletaja* stands for a ‘poet’, and *laulik* is not used, the texts were about non-Estonian authors. Additionally, *laulik* once used to denote a ‘bird’ as ‘nature’s singer’.

Laulik, when referring to a poet, carries emotive and romantic connotations. For example, a reviewer writes about the Estonian poet Jakob Tamm (1861–1907): “When reading the poem ‘Ärkvel’ [Awake], it’s hard to believe that the beautiful and poetic idea did not captivate the singer more than it did.”¹⁹

Laulja primarily refers to choir singers (34 times), rarely to singers in general (4). *Lauljanna* refers to a female classically trained solo singer (2). The suffix *-nna* was invented in the nineteenth century, influenced by Indo-European languages, where the female prominent person was linguistically distinguished. Feminine nouns with this suffix are rare in Estonian, while derivative suffixes for the masculine gender do not exist (Kasik 1996: 128; 2015: 243). Thus, *lauljanna* is a marked novel word in Estonian.

Words meaning ‘singer’ (54 times) comprise 10.8% of the studied vocabulary. Among the compounds, notable obsolete terms such as *kunstlaulja* ‘classically trained singer’ and *salmilaulja* ‘psalm chanter of the Orthodox Church’ were found.

The compounds and collocations (Table 3) mostly refer to various song genres and the phenomena related to choral singing: *laulupidu* ‘song festival’ (40), *laulukoor* ‘singing choir’ (30), and *lauluselts* ‘singing society’ (7). Similar longer word combinations are also associated with choral singing, such as *laulu- ja mängupidu* ‘song and music festival’ and *kokkulaulmine* ‘singing together’, among others.

The press primarily covers the social aspects of singing, such as choral song festivals and events in singing societies, often organised to raise funds for local purposes, such as obtaining instruments for an orchestra. Art music – solo songs, organ concerts, and opera – are described with respect, as if they were representing the highest branch of music. Multipart choral singing was promoted as it was associated with the development of local cultural standards: “We fervently wish that the awakening and guidance from the major song festivals

inspired learning and lead aspiration in our villages and homes” (Postimees 1897). Several compounds, such as *laulukevade* (song spring) and *lauluõpetus* (singing tuition) reflect Estonians’ growing alignment with Western music.

Other frequent singing-related compounds reflect dominant ideologies: *kirikulaul* ‘church song’ (12) and *lauluraamat* ‘hymnal’ (15) represent Christian influence, while *keisrilaul* ‘tsar’s song’ (16) and *Vene riigi laul* ‘song of the Russian state’ (2) denote the Russian Empire’s anthem. Both national movement leaders and ruling authorities deemed Christianity and its spiritual songs essential for rural populations, viewing religion as integral to European civilisation, order, and morality. A significant portion of the singing-related vocabulary derives from a lengthy article on the compilation of a new hymnal (*Uus Lauluraamat* 1900). The article’s author compares congregational and folk singing, stating: “Church song is the highest level of spiritual songs. ... A church song is to religious songs what a folk song is to secular songs”²⁰ (Olewik 1895a: 653). He argues that church songs are a fundamental need and should be comprehensible to “even the simplest person” (ibid.: 654). Realistically assessing church singing, he sees adapting songs as useful, noting that “everyone knows, how much physical strength it takes singing 8–11 stanzas in a row in a high voice” (ibid.).

The singing and playing of the Russian tsarist anthem are frequently mentioned in the late 1800s, particularly at state events, such as Nicholas II’s visit to Germany in 1896, tsarist family birthdays, and coronations. Given the period’s Russification efforts, both the performance and documentation of these anthems likely reflect attempts to appease authorities. Nevertheless, people may have been genuinely impressed by the uncommon sounds of large orchestras and choirs. Large musical ensembles are also praised for their impressive power in non-political contexts within the media. The anthem was also performed at anniversaries of various Estonian societies, including those of farmers and firefighters.

Table 3. Frequency of laul-stem compounds and similar word combinations occurring at least 3 times in the Estonian Media Corpus. Variant spellings have been consolidated under standardised forms.

Compound	Translation	Frequency	Longer word combinations with related meaning	Translation	Frequency
<i>laulupidu</i>	song festival	40	<i>laulu- ja mängupidu 1 / tänulaulupidu 2 / laulupiduline 2</i>	song and music festival / thanksgiving song festival / participant in the song festival	5

<i>laulukoor</i>	singing choir	30	<i>laulu- ja mängukoor</i>	singing choir and orchestra	7
<i>keisrilaul</i>	Tsar's song (Anthem)	16	<i>Veneriigi laul / Vene riigi laul</i>	Russian state song (Anthem)	2
<i>lauluraamat</i>	Hymnal	15	<i>laulu- ja käsiraamat</i>	song- and handbook (Hymnal with extra)	3
<i>vaimulik laul</i>	religious song	14			
<i>kirikulaul</i>	church song	12	<i>Ewangeliumi kiriku (koguduse) laul 1 / kiriklik ning ilmalik laul 1 / Jumalateenistuse (jumala teenistuse) laul 2</i>	Gospel / church song and secular song / song at the service	4
<i>lauluselts</i>	singing society	7	<i>laulu ja mängu selts 2 / muusika ja laulu selts 1 / laulu ja muusiku selts 1 / kasinuse ning laulu selts 1</i>	singing and playing society / music and singing society / singing and music society / chastity and singing society	5
<i>rahvalaul</i>	folk song	7			
<i>naljalaul</i>	humorous song	6	<i>naljakas laul</i>	funny song	1
<i>Eesti laul</i>	Estonian song	4	<i>Eesti algupäraline laul 1 / [eesti] algupäraline laul 1</i>	Estonian original song	2
<i>tundelaul</i>	lyrical poem	4			
<i>lugulaul</i>	epical song	4			
<i>lastelaul</i>	children's song	4			
<i>koorilaul</i>	choral song	3	<i>meestekoori laul</i>	male choral song	2
<i>lauluviis</i>	song melody	3			
<i>mõttelaul</i>	philosophical poem	3			

Estonia's future anthem was already known in the 1890s and sung "breath-takingly" at a choir's anniversary (Wirmaline 1892). Finland's national song "Maamme" (Our country) made news when sung at a Helsinki demonstration protesting the closure of *Päivälehti* newspaper (Postimees 1899b). The Latvian national anthem faced criticism for its German influence. A reviewer noted its similarity to German songs "Wenn ich ein Wöglein wär" and "Heil dir im Siegerkranz", suggesting Latvians create a new, more original melody based on their folk songs (Olewik 1895b).

The term *rahvalaul* 'folk song' (7 times), often represents a concept rather than actual singing. The sole direct reference to folk song performance describes it as a 'recital' at a gathering: "Mr. Orgusaar recited [*deklameeris*] several Estonian folk songs, eliciting enthusiastic applause and cheering from the audience" (Postimees 1896).²¹ It remains unclear to which performance style the news refers. While *deklameerima* 'to declaim' in Estonian typically refers to poetry recitation without melody, Orgusaar might have performed *regilaul* songs in their traditional recitative style, which the journalist may have perceived it as declamation rather than singing. Folk song collectors have also described the *regilaul* style as declaiming or reciting. For example, "It appeared that tunes did not feature in his [W. Servinski's] performance; he sang by reciting [*deklameerides*]" (Tampere 1935: 14; EÜS I 752; see also Ch. 6.1).

Performing Seto folk songs, likely on the stage, is mentioned in a description of a festive celebration at Navi school: "After several choral performances, the teacher humorously narrated Setu stories in dialect, interspersed with Setu songs" (Wirmaline 1892). This brief account does not clarify how (the traditionally multipart) Seto songs were actually performed.

Songs and folk songs were often characterised by their national or ethnic origin, with ethnonyms such as 'Estonian', 'Finnish', 'German', and 'Setu'. Notably, five out of six references to joke songs related to performances by a Russian theatre troupe, which highlighted ethnic tensions: a joke song about Jews performed by the troupe led to a lawsuit against the Jews who loudly protested (Olewik 1894).

The verb *laulma* 'to sing' appeared 54 times, the gerund *laulmine* 'singing' 16 times. Among these 70 references to singing, choral and spiritual contexts prevailed. An interesting obsolete verbal collocation, *ligi laulma* 'to sing close', meaning to sing together in a choir, was found. However, in a description of a *regilaul* performance a similar word, *ligilaulja* (lit. 'close singer', RKM II 94, 115 (2), not available in the ERAB), has been used for an echo singer – usually called *järellaulja*, who in the traditional way of singing repeats the lead singer's stiches.

In conclusion, the unqualified use of *laul* (song) in the EMC most frequently referred to choral songs or spiritual/church songs, each occurring about 50 times. *Laul* appears nearly 30 times in poetry reviews. However, the EMC does not cover all topics comprehensively; for instance, folk songs are rarely mentioned despite their prominence in oral tradition. This omission may reflect a lack of interest in folk songs among both journalists and corpus compilers.

6. THE CORPUS OF ESTONIAN *REGILAU* LYRICS

6.1 About the corpus

This study utilises the Estonian Runosongs' Database (ERAB) and analytical tools developed by the FILTER project for Finnic folk songs (Janicki & Kallio & Sarv 2023). For the purposes of our analysis, we treat the texts from the ERAB as a corpus, which we refer to as the Regilaul corpus. This corpus is more extensive than the media corpus (Table 1), comprising approximately two-thirds of all *regilaul* lyrics (primarily older texts) held in the Estonian Folklore Archives at the time of analysis, in 2021. It also occasionally includes some Estonian folk songs of other styles.²² *Regilaul* language is variable, with singers often using irregular forms that blend diverse regional and historical linguistic features. A large amount of song lyrics was transcribed by various individuals during fieldwork in the late nineteenth and early twentieth centuries, resulting in diverse orthography that resists complete standardisation through editing.

Estonian *regilaul* is part of the ancient Finnic runosong tradition, characterised by specific “Finnic” tetrameter, alliteration, parallelism, and content similarities. The primary structural unit is the verse line, sung successively without stanza formation, often alternating between a leader and a chorus (or echo singer). *Regilaul*, especially the leader's part, is often performed in a speech-like recitative style, with *ütleva* ‘to say’ used as an alternative term for singing (see, e.g., Sarv 2017; Tampere 1956; Rüütel 1998; Lippus 1995).

Several songs held ritual significance, accompanying activities such as farm work, weddings, and calendar traditions, while others were performed in various contexts. Song lyrics inhabit a poetic realm that reflects the real world through singers' perspective.

While folk songs about singing were not typically associated with specific activities, some appeared in ritual contexts, such as wedding singing contests. Folklorists classify *regilaul* lyrics into types based on content and function.²³ According to the preliminary statistics, some of the most prevalent song types

in the ERAB concern singing, including “Power of Song” (*Laulu võim*, 1000 variants), “From Where Songs Came” (*Kust laulud saadud*, 556), and “Mouth’s Quill” (*Suude sulg*, 480).²⁴ For comparison, other popular song types include calendar songs like St. Martin’s and St. Catherine’s songs (867 and 398 respectively), children’s songs like “Wirble-Warble, Small Lark” (*Liiri-lõõri*, 656), and narrative songs such as “Revenge to Manor Lords” (*Kättemaks sakstele*, 548). The prevalence of singing as a song topic suggests its popularity among both folk singers and collectors.

Folklore publications often prioritise songs about music, reflecting their symbolic importance. These songs start the volume of lyrical songs both of the Anthology of Estonian folk song lyrics (ERL 1970) and the collection of melodies (Tampere 1965).

6.2 *Laul*-stem words in the Regilaul corpus

The high linguistic variability of the Regilaul corpus has precluded automatic lemmatisation and grammatical unification. *Regilaul* lyrics preserve many archaic word forms due to their poetic meter, including the older two-syllable form *laulu*,²⁵ the stem of which varies dialectally (*laal*, *loul*, *laol*) and the final vowel appears as *a*, *e*, *i*, *o*, or *õ*. Our analysis comprised the dialectal forms, and below, when writing about *laul*-stems and words in the Regilaul corpus, we include dialectal variants by default. For this study, Taive Särg manually lemmatised and standardised all song/sing-stem words, aligning them with contemporary literary Estonian. Key terms used in this analysis are the following:

- *laul*-stem: any word containing a *laul*-stem letter combination (including dialectal variants);
- *laul*-word: song/sing-stem word in its basic form (e.g., *laul*, *laulja*);
- surface form: original letter combination in the text;
- standardised form: standardised grammatical form of a *laul*-stem word (e.g., historical *laulamasta* standardised to *laulmast*, the elative case of the verb infinitive).

We initially conducted an SQL search across the entire corpus for the stems *laul*, *laal*, *loul*, and *laol*. This yielded 29,143 words containing these letter combinations, including 1,562 unique surface forms. We transferred these words to an Excel spreadsheet for further analysis which revealed several challenges requiring manual intervention:

- 1) eliminating surface forms identical to *laul* or its dialectal forms unrelated to singing;
- 2) differentiating homonyms, with particular attention to the abundant morphological homonyms among *laul*-words, though some remained incompletely defined (e.g., *laulu* and *laul* representing various grammatical forms);
- 3) contextually determining the meaning (basic word and case) of numerous surface forms;
- 4) correcting inconsistencies in formatting, standardisation, and editing, as well as errors in the original data.

We excluded several types of words: compounds containing *laul* or *laal* as mid-part (e.g., *laulinad* ‘hay barn sheets’, *samblaalune* ‘under-moss’), unrelated homonyms (e.g., *laale* ‘to the yard’, *lault* ‘gently’), names (*Laala*), meaningless words and refrains (e.g., *lallaallaa*, *vaat laali*), and errors. We also manually excluded *laulatama*-stem words (‘to wed’), as their forms can be identical to *laulma* verb forms, e.g., *laalat* as both ‘wed’ (attribute) and ‘you sing’ (second person verb).

Through this cleaning process, we removed 427 “false” *laul*-stems, leaving 28,716 *laul*-stems represented by 1,414 distinct surface forms. We lemmatised the rest of the material, consolidating words with different spellings, dialectal features, and grammatical forms. The poetic verb form *laulemaie* (~*laulemahe*), characteristic of *regilaul* language, was treated as an independent historical word rather than being reduced to the main verb form *laulma* ‘to sing’.

The main procedures were 1) implementing syncope and apocope (e.g., *laulijaksi* > *lauljaks*), 2) adding reduced vowels (e.g., *lauls* > *laulis*), 3) unifying the vowels, especially different stem vowels (e.g., *laal* > *laul*) and vowel harmony (e.g., *laulõlõma* > *laulelema*), and 4) removing morphological endings. There occurred a high variety of forms, including many problematic cases. The examples in the tables represent words already processed, e.g., the standardised verb form *laulelema* stands for the surface forms *laulõlõma*, *laulel*, *laulelma*, *laulelda*, *laaleldes*.

Table 4. Frequency of laul-words in the Regilaul corpus.

Word	Lexical category, occurrences	Examples	Total	% of all laul-words	Category total	% of all laul-words
<i>laul</i> song	primary word 5,271		5,271	18.4%	noun <i>laul</i> 8,204	28.6%
	compound, second - <i>laul</i> 442 (1.5%)	<i>kukelaul</i> 81 cockcrow	2,921	10.2%		
		<i>sõjalaul</i> 44 war song				
		<i>lustilaul</i> 43 amusement song				
compound, first <i>laulu-</i> 2,479 (8.6%)	<i>laululeht</i> 652 song sheet	2,479	8.6%			
	<i>laululind</i> 326 song-bird					
	<i>laulukool</i> 219 singing school					
<i>lauluke(ne)</i>	diminutive 227		227	0.8%	noun <i>laul-</i> derivatives 2,433	8.4%
<i>laulja</i> singer	agent noun 958		970	3.4%		
	compound, second 11	<i>vastulaulja</i> 7 responsive singer				
	compound, first 1	<i>lauljarahvas</i> 1 singing folk				
<i>lauljanna</i> female singer	agent noun 1		1	0		
<i>laulik</i> singer	agent noun 1,072		1,085	3.8%		
	compound, second 3	<i>ilmalaulik</i> 3 heaven/world/great singer				
	compound, first 10	<i>lauliklind</i> 5 singer-bird <i>lauliku-soost</i> 5 from singer's line				
<i>laulmine</i> singing	gerund 149		150	0.5%		
	compound, second 1	<i>linnulaulmine</i> bird's singing				
<i>laululine</i> singing-apt, song-rich	adjective 151		151	0.5%	adjective <i>laululine</i> 151	0.5%

<i>laulma</i> to sing	verb 15,001		15,001	52.2%	<i>laulma</i> verb and derivatives 17,940	62.5%
<i>laulemaie</i> to sing	poetic form	<i>laulemaie</i> (about starting) 2,776 <i>laulemahe</i> (about starting) 116	2,892	10.1%		
<i>laulma</i> to sing –derivatives	change of state	<i>laulanema</i> 15	47	0.2%		
		<i>laulenema</i> 5				
	repetitive action	<i>laulaskema</i> 8				
		<i>lauliskelema</i> 2				
		<i>laulatelema</i> 1				
		<i>lauletelema</i> 2				
		<i>laulelema</i> 11				
		<i>laulatamaie</i> 1				
	reflexive	<i>lauldema</i> 1				
		<i>lauldumaie</i> 1				
Total			28,716			

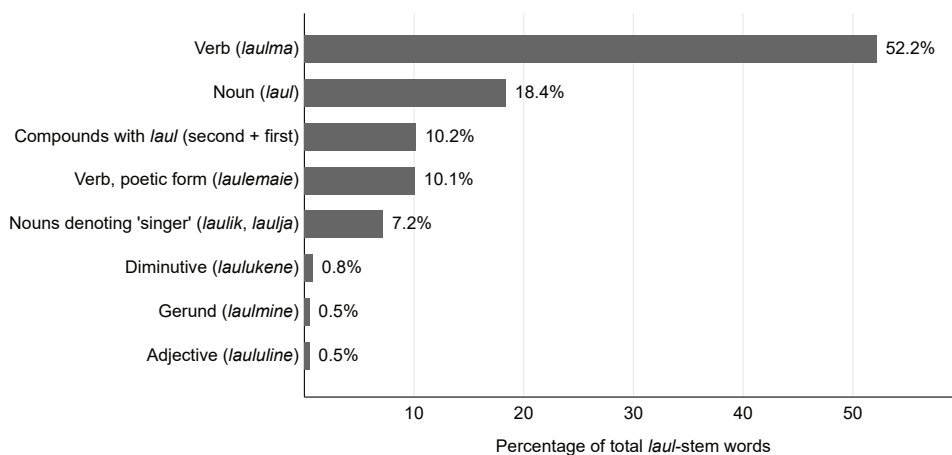


Figure 2. Frequency of laul-words in the Regilaul corpus.

Table 4 and Figure 4 summarise the results of the word-level analysis. The noun *laul* comprised 18.4% of all *laul*-words in the Regilaul corpus (5,271 instances). The most common *laul*-noun surface form (and the second most frequent among all *laul*-stems) was *laulu* (2,280). This form represents multiple grammatical cases (nominative, genitive, or partitive singular, and endingless nominative plural), which we did not differentiate. Including instances where the plural glottal stop is marked by an apostrophe (*laulu'*, 24), the total reaches 2,304. The next most frequent surface form is the regular plural *laulud* (1,285).

The two terms for 'singer' occur with similar frequency: *laulik* (1,072, and 1,085 including compounds) and *laulja* (958, and 970 including compounds), so 'singer' appeared 2,055 times (7.2% of *laul*-words total). Unlike in the EKSS dictionary and the EMC, these terms appear semantically indistinct in the Regilaul corpus. Despite their semantic similarity, these singer terms often associate with specific motifs in common song types. *Laulik* frequently appears in the "From the Line of Singers" (*Lauliku soost*) type, with typical verses like: "A singer [was] my father, a singer [was] my mother, a singer [was] my baby-swayer" (*Laulik isa, laulik ema, laulik lapse kiigutaja*) (EÜS VII 1682/3 (244) et alibi), "I come from the singer's line" (*Ma olen lauliku soosta*) (E 24031/2 (1) et alibi).²⁶

Laulja often appears in longer, pre-syncope surface forms retaining a historical stem vowel, most commonly as *laulija*, but also as *laulaja*, *laulõja*, etc. It features prominently in song types like "A Singer's Toil" (*Lauliku vaev*), "Pay to the Singer!" (*Laulikule palka!*), and "Drink for the Singer!" (*Laulikule juua!*), with characteristic verses such as "A singer has more toil [than]..." (*Enam on vaeva laulijalle; Inämb ol' l vaiva laulijal*), "Give money to the singer!" (*Andke raha laulijalle*), "The singer's palate is burning" (*Laulija lagi palasi*), and "Give vodka to the singer" (*Andke viina laulijale*).

In the poetic realm of *regilaul*, both *laulik* and *laulja* function as unmarked, primary words referring to ordinary singers. The potential association of *laulik* with a special blood lineage, as suggested by certain song types, warrants further investigation.

Traditionally, the gender of a singer appears from the context and from self-presentation, e.g., "this is a girl's singing voice" (*See on neiu lauluhääli*) or "[they are] listening to my words, a bad child's songs, a little boy's weavings" (*Kuulavad minu sõnuda, lapse halva laulusida, poisikese põime'eida*) (H II 25, 944 (3); H II 38, 189/90 (31)).

The gender of female singers is usually not marked. The modern term *lauljanna* 'female singer, songstress' appears only once in the corpus, specifically in a song from the newer end-rhymed layer of folk songs. Women constituted the majority of folk singers in recent memory (see, e.g., Kõiva 1964; Rüütel 2013; Oras 2008, 2017).

However, some male singers are also documented in *regilaul* tradition, suggesting that there may have been more in earlier times. The compounds for singers (Table 7) rather specify the male gender, such as *laulumees* ‘songman’ (25), or *lauluvennad* ‘song brothers’ (3), *lauluisa* ‘song father’ (1), and *lauluhärä* (1), *laulusaks* ‘singer of noble state’ (1); while the terms for ‘song child’ *laululaps* (8) and for female singers, *laulunaised* (4), are rarer. The singers’ traditional roles are denoted by terms such as *vastulaulja* ‘responsive singer’, *kaasalauljad* ‘co-singers’, and rather metaphorical *laululangukene* ‘singer-in-law’. In addition to people, birds are frequently mentioned as singers, with terms like *laululind* ‘songbird’ (326), *laulurästas* ‘song thrush’ (104), *laulukukk* ‘song rooster’ (8), and *ilmalaulik* ‘heaven/world/great singer’ (3).

Unlike the Media Corpus, the Regilaul corpus undergoes orthographic standardisation by its curators to align it with contemporary spelling rules. In relation to this standardisation, we analysed the existing compounds without searching for similar noun phrases.²⁷ Frequent compounds such as *laulukool* ‘singing school’ (219) and *laulukoor* ‘singing choir’ (13), refer to the modernising world and appear in similar contexts. By denying that they learned songs at school, in a choir, or at other modern literary institutions, the singers assert their connection with the local oral tradition, thereby affirming their singing power:

<p><i>See poiga õbissa olnud, seisand seppade pajassa, raamatmaakeri majassa, kirjatundija toassa, loalaulija eessa, laululaulija taganna. Mina aga kuulin vasta kostin: Ei ole õbissa olnud, [jne. ---]</i></p>	<p>This boy has got schooling, has been standing in the blacksmiths’ forge, in the bookmaker’s house, in the bookworm’s room, in front of the chanter, behind the cantor. I heard and replied: I have no schooling, [etc. ---]</p>
<p>(H II 38, 189/90 (31))</p>	

Some rare words, whose meanings were not immediately clear for oral tradition, were interpreted, using contextual clues, for instance, *lauluõpetaja* ‘song teacher’, *väljalaulja* ‘outsinger’, and *loalaulja* ‘chanter’[?]. It turned out that teaching was not connected with formal training, but the sister, from whom the songs had been learned, served as a ‘teacher’ here. Outsinger was a singer in the field.²⁸ *Loalaulija* ‘text chanter’, paired with *laululaulija* ‘cantor, song singer’ in the parallel verses refers to the church singers’ roles.

The diminutives *lauluke* or *laulukene* ‘a little song’, which are indistinguishable in inflected forms, occurred 227 times, while the gerund *laulmine* ‘singing’ appeared 150 times (0.8% and 0.5% respectively).

Verbs. The Regilaul corpus is rich in the number and variety of verbs, with verbs forming 62.5% of *laul*-words. In addition to the basic verb *laulma* (which appears in many archaic surface forms, such as *laulema*), the song lyrics also contain several verb derivatives. The forms *laule/le/ma* and *laule/ne/ma* might indicate: 1) singing as a continuous activity with *-le* as a frequentative suffix; 2) the state of singing with the subsistence suffixes *-le* and *-ne*; 3) the emergence of singing with the translative suffix *-ne*; 4) *laul/du/ma* and other verbs with suffixes *-u* and *-du* refer to self-reflexive activity (Kasik 2015: 116–142); and 5) *laulemaie* (*laulemahe*, *laulamaie*, etc.) is a special poetic form in *regilaul* language, used mostly in the phrase *hakkan laulemaie* (I begin singing). Grammatically, it represents a version of historical illative case (ending *-he*, *-je*) of the verb ‘to sing’ and seems to function as a crystallised poetic formula.

The verb *laulma* in successive parallel verses of *regilaul* was analysed in the dissertation of Kanni Labi, who found that *laulma* appeared more often in the first verse than in the second one within parallelism groups (66% versus 14%). The numerous parallel words (93 different words) had a wide range of meanings, such as ‘to poetise/versify/fantasise’, ‘let/shoot’, ‘roll’, and ‘tell’ (*luulema*, *laskma*, *veeretama*, *pajatama*). They include many onomatopoeic words for bird songs, especially ‘to cuckoo’ (*kukkuma*), and verbs that associate with moving (Labi 2006: 86–90).

The high proportion of verbs in *regilaul* texts emphasises singing as an activity, while “many songs” in lyrics refers to memorised motifs, plots, and improvisational skills. Some songs also refer to written notes (*laululeht* ‘song sheet’), which the singer could use as a memory aid (see Lintrop 2017).

Our analysis of neighbouring words for ‘song’ was limited to a smaller data set, focusing on attributes preceding the standardised diminutive form *lauluke(ne)* (227 times). The word pairs were identified and sorted automatically and then analysed. Adjectives for *lauluke(ne)* occur 45 times, the most frequent are *lahe* ‘cool, easy-going’ (28) and similar *lahke* (southern Estonian *lake*) ‘affable’ (7). Other attributes include ‘short’ (9), ‘long’ (1), ‘small’ (1), and *lõhmüksine* ‘made of linden’ (1).

The attribute can also refer to the “song’s owner”, i.e., the singer; “little songs” are most often sung by children – *lapse lauluke* (27). The analysis revealed that due to inversion, the adjective before ‘song’ might belong to the previous word, as in *lapse hullu laulukeisi* (silly child’s little songs).

Often, the attribute is a pronoun that emphasises the ownership of a song, such as ‘my’ or ‘our’ (each 8 times). The phrase *minu lauluke* ‘my little song’ is always paired with the antithesis *s(in)a laulad* ‘you sing’ and appears in the song type “Singer’s Abuse” (*Lauliku sõim*): “Why are you singing if you cannot? [Why are you] singing my songs, rolling my tunes?” (*Mis sa laulad, kui sa ei mõista, laulad minu lauluke, veeretad mu viisikesi?*) (E 7041).

Parallel words next to *lauluke(ne)* include a synonym *leelokene* (*leelokõnõ, lellokene*); similarly, after the conjunction *ja* (and) appears the word *louke* ‘little story’.

Before ‘little song’, the verb for singing *laulma* (14 times) appears, but also *laskma* ‘let, shoot’ (33), *lautama~laotama* ‘spread’ (5), *lahutama* ‘spread, release’ (3), *lautelema* ‘gradually spread’ (4), and *lööma* ‘strike’ (1).

The analysis of parallel words for ‘song’ was based on the parallelism in *regilaul* lyrics, where successive verse lines present and develop the content using similar words and syntax (see Sarv 2017). The semantic analysis of verbs in parallelism groups, conducted by Labi (2006), demonstrated the diverse semantic connections between parallel words. We compiled a frequency list of the words occurring in the “second” lines (i.e., those that followed the lines with the noun *laul*) and although the words were not lemmatised, the list still highlighted the most frequent parallel words for *laul*. At the top of the list was *viis* ‘tune, melody’, appearing at least 560 times.²⁹

Table 5. Examples of parallel verse pairs with words laul–viis ‘song–tune’.

<i>Küll mul laulusid kui ma laulan, küll mul viisisid kui ma viisin.</i>	I’ve got enough songs if I sing, I’ve got enough tunes if I care [to sing].
<i>Kui mina akkan laulema, viisi välja veeretama.</i>	If I start to sing, to roll out a tune.
<i>Laulaksin ma lugu kaksi, veeretaksin viisi kuusi.</i>	I’d sing two songs, I’d roll six tunes. (or: I’d roll five-six [tunes])
<i>Laulge, laulge, laisad neiud, virved, viisi veeretage!</i>	Sing, sing, lazy girls, sprouts, roll a tune!

<p><i>Otsas, otsas lauluke,</i> <i>otsas lauluviisike.</i> (E 60609 (5); H II 5, 1/2 (1); H II 35, 392/3 (236); ERA II 159, 239/40 (11) et alibi; H II 56, 350 (13); E A 790/2 (2) et alibi)</p>	<p>The little song is over, over, the little song's tune is over.</p>
---	--

The word *viis* is newer than *laul* and is not part of the old Finnic vocabulary. However, its frequent use in alliterative collocations, such as *viisi veeretama* 'roll a tune', and in pairs with words like the place name *Viru*, or verbs like *viitsima* 'care' and *viima* 'take', demonstrates its successful adaptation into *regilaul* (Table 5). *Viis* is likely a loanword from Low German.³⁰ The German *Weise* and its variant *wīs(e)* in Middle High and Middle Low German originally meant 'mode' and are related to the English word *wise*. The Old High German *wīsa* was used for melody as early as the tenth century. In Old Norse, the related stem *visa* referred to song lyrics, meaning 'strophe, verse'. Traditional German expressions include *Art und Weise* 'in a way' and *Wort und Weise* 'text and melody' (DWDS).

In Estonian, the word *viis* carries both meanings, 'mode' and 'melody'. Songs about singing in their current form could not have been composed before contact with Germanic peoples. In folk songs, *laul* and *viis* often appear as a pair of parallel words in successive lines, with *viis* signifying a melody (Table 5). *Viis* can also imply singing 'in a way', so *linnu viisi* can be interpreted as 'bird's wise', which closely resembles the meaning 'with bird's tune'.

<p><i>Siis ma laulaks lagle viisi,</i> <i>teeksin healta tedre viisi.</i> (H II 15, 430 (10))</p>	<p>Then I'd sing a barnacle goose's wise; make a sound a black grouse's wise.</p>
---	--

A variation of the preceding example includes expressions that clearly mean 'bird-like' or 'in birds' tongue':

<p><i>Siis ma laulas linnukeeli,</i> <i>tees ma hääli tedre muodi.</i> (TEM 4a, 6 (62))</p>	<p>Then I'd sing in birds' tongue, make a sound in a black grouse's way.</p>
---	---

Additionally, the word *viis* carries further ambiguity because it has a homonym meaning 'five', as well as morphological homonyms with the verbs *viitsima* 'care, bother' and *viima* 'bring'.³¹

Scholars note that *regilaul* practitioners used terms like *toon* 'tone', *hääli* 'voice' and *mõnu* 'delight/pleasure' for melody, not *viis* (Rüütel 1998). The word *toon* might also be a loan from Low German or similar languages (Metsmägi

& Sedrik & Soosaar 2013). Although the analysis of *regilaul* lyrics offers insights into this topic, we will set this question aside for now.

Other frequent parallel words to ‘song’ include place names like ‘village’ and ‘home’, as well as birds such as ‘cuckoo’ and ‘duck’. For example, “Oh, what I cuckoo, a sad bird, or sing, a bad child” (*Oi, mis kukun kurba lindu, või mis laulan halba lapsi*) (EÜS V 1088/9 (65)). Additionally, the analysis revealed verbs associated with *laul* and its parallel words, such as ‘to cuckoo’ and ‘to sing’ in the previous example (cf. Labi 2006).

Table 6. Most frequent compounds (occurring more than 10 times) with ‘song’ in the Regilaul corpus. The frequencies of the two types of compounds are presented in descending order on both sides of the table, with the columns arranged side by side according to the magnitude of the numbers.

Laul as the first part (modifier)	Translation	Occurrences	Laul as the second part (main word)	Translation	Occurrences
<i>laululeht</i> 652 / <i>laululeheke</i> 2	song sheet	654			
<i>laululind</i> 326 / <i>laululinnuke(ne)</i> 35	songbird	361			
<i>laulukool</i>	singing school	219			
<i>lauluhääl</i>	singing voice	189			
<i>lauluviis</i> 99 / <i>lauluviisike(ne)</i> 48	song’s tune	147			
<i>laululaud</i> 132 / <i>laululauake</i> 14	song’s board	146			
<i>laulurästas</i>	song thrush	104			
<i>laululugu</i> 79 / <i>laululookene</i> 3	song story	82	<i>kukelaul</i>	rooster’s crow	81
<i>laulukord</i>	singing turn	79	<i>sõjalaul</i>	war song	44
<i>laulusõna(d)</i>	song lyrics	56	<i>lustilaul</i>	amusement song	43
<i>laulupaelad</i>	singing ribbons	31	<i>linnulaul</i>	bird song	29
<i>lauluraamat</i>	hymnal	28	<i>kiigelaul</i>	swinging song	26
<i>laulumokad</i>	singing gobs	27	<i>rõõmulaul</i>	joy song	24
<i>laulumees</i>	songman	25	<i>kanalaul</i>	hen song	20
<i>laulusulg: -sule</i>	song quill ³²	16	<i>kirikulaul</i>	church song	17
<i>laulukoor</i>	singing choir	13	<i>leinalaul</i>	song of grief	15

<i>laulusuu</i>	singing mouth	11	<i>tänulaul</i>	song of gratitude	12
<i>laulupärg</i>	song wreath	10	<i>ilulaul</i>	beauty/joy song	10

Table 7. Compounds with laul-component, representing singers in the Regilaul corpus (frequency more than 4).

Word	Translation	Occurrences
<i>laululind</i> 326 / <i>laululinnukene</i> 35	songbird	361
<i>laulurästas</i>	song thrush	104
<i>laulumees</i>	song man	24
<i>laululaps</i>	song child	9
<i>lauluema</i>	song mother	7
<i>laulukukk</i>	song rooster	7
<i>vastulaulja</i>	responsive singer	7
<i>laulunaine</i>	song woman	4

Table 6 provides an overview of *laul*-compounds, while Table 7 lists the most frequent singer types. The compounds denoting vernacular song genres on the right side of Table 6 refer to various phenomena, such as *rooster crow* (81), *hen song* (20), *war song* (44), *amusement song* (43), and *bird song* (29). Roosters and chickens are frequently mentioned because the rooster's crow signifies morning. For example, "Thou shalt not hear the rooster song" (*Ära sina kuula kukelaulu*) (EÜS VI 1139/40 (19) et alibi) advised a maiden to rise early. The rooster's song also symbolised human habitation, so "Where thou do not hear the rooster song" (*Kus ei kuule kukelaulu*) (E 36877 (10) et alibi) indicated a remote uninhabited place. The hen's song '*kanalaul*' often appeared in a parallel verse.

There were 81 compounds of low frequency (less than 10 times), which included evocative types like oven-song, love-song, grove-song, silver-song, cook's song, gold song, blaming song, lament song, witch-song, whisk-song, etc. (*ahjulaul*, *armulaul*, *hiielaul*, *höbelaul*, *kokalaul*, *kuldalaul*, *laimulaul*, *nutulaul*, *nõialaul*, *vihalaul*).

Folk song and related terms appear infrequently in the corpus; single instances include *rahvalaul* 'folk song', *auk-laul* 'hole-song' or 'gap-song', *regilaul*, *leelulaul*, e.g., a controversial claim about the term 'regilaul': 'I have ancient verses, rhymed *regilauls*' (*Mul om muistese värssi, regilaulu riimilise*)³³ (H III 16, 355/6 (2)).

7. COMPARISON AND DISCUSSION

The corpus-based analysis of song/sing-stem words has revealed distinct conceptualisations of singing in the discourses of late nineteenth-century written media and traditional songs. Table 8 and Figure 3 below compare the proportions of different lexical categories among studied words in both corpora: the 1890s media corpus (EMC) and the Regilaul corpus. Given the differing character of texts and compilation methods of the corpora, the results may reflect these variations.

The noun *laul* ‘song’ constitutes 18.3% of all *laul*-words in the Regilaul corpus, about half the proportion found in the EMC, where it appears in 32.7% of the cases. The slightly less frequent use of terms for ‘singer’ in *regilaul* (7.2%) compared to the EMC (10.8%) may be attributed to the fact that the singer often serves as the poetic I.

Table 8. Comparison of the occurrence of different word classes with sing/song-stem in the Estonian Media Corpus and the Regilaul corpus.

Category	EMC (%)	Regilaul corpus (%)
Noun Forms		
<i>laul</i> as a root word	32.7	18.3
<i>laul</i> as main part of compound	14.6	2.7
<i>laul</i> as modifier in compound	22.8	7.5
<i>laul</i> in other compounds	4.0	-
<i>lauluke(ne)</i>	0.6	0.8
<i>laulja</i>	8.0	3.4
<i>lauljanna</i>	0.4	0
<i>laulik</i>	2.4	3.8
<i>laulmine</i>	3.6	0.5
Total Nouns	89.1	37.0
Adjective Forms		
<i>laululine</i>	-	0.5
Total Adjectives	-	0.5
Verb Forms		
<i>laulma</i> in various forms	10.8	52.2
<i>laulma</i> derivatives	-	0.2
<i>laulemaie (-mahe)</i>	-	10.1
Total Verbs	10.8	62.5

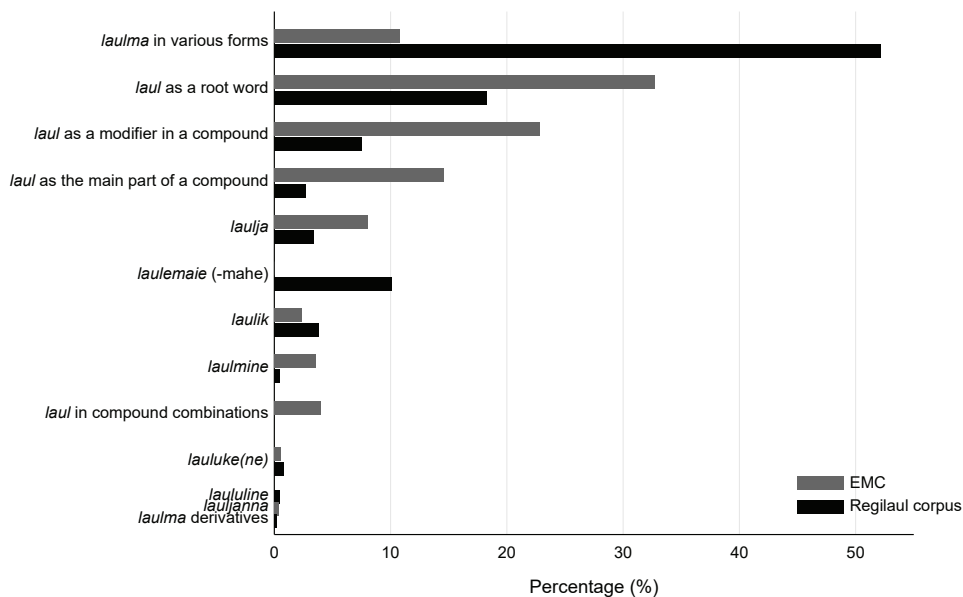


Figure 3. Comparison of the occurrence of different word classes with sing/song-stem in the Estonian Media Corpus and the Regilaul corpus.

Comparing the proportions of different word classes (Table 8), the most notable feature is the distribution of song/sing-stem nouns and verbs: nouns dominate in the EMC (89%), while verbs prevail in *regilaul* texts (62.5%). The noun ‘song’ appears significantly more frequently in the EMC compared to the Regilaul corpus (74.1% vs. 25.8%), both as a primary word (32.7% vs. 18.3%) and within compounds (41.4% vs. 10.2%).

The prevalence of the verbs among song/sing-stems in folk song lyrics emphasises singing as an activity, a process, with songs as objects appearing less frequently. The act of beginning holds particular significance, marked by the special introductory infinitive verb form *laulemaie (-mahe)*, used in the stereotypical verse ‘When I begin to sing’ (*Kui ma hakkam laulemaie*). This poetic form (excluding other inverted forms of the singing infinitive) accounts for approximately one-sixth of all singing infinitive verbs.

The abundance of singing verbs reflects oral tradition, where ideas manifest during performance, rather than as fixed songs. Mark Slobin characterises this practice, stating: „Folk music is not a set of songs and tunes; it is more a working practice. People take available musical resources and develop strategies to make good use of them” (Slobin 2011: 9; cf. Finnegan 1977; Honko 2000).

According to Labi, one more reason why the verb ‘to sing’ in *regilaul* is very frequent, is its referring to communication, with parallel words *üttelema*, *rääkima* ‘to say, to tell’ (Labi 2006: 86–90). The verb richness in *regilaul* is further evidenced by approximately a dozen derivations of ‘to sing’, such as *laulelema* and *laulanema*. The adjective *laululine* ‘songful’, absent in the EMC, appears in the *regilaul* texts.

The frequent compounds with a song-component in newspaper texts – song festival, singing choir, hymnal, ecclesiastical songs, and emperor’s song (anthem) – directly reflect the historical context and ruling ideologies. However, the scarcity of shared compounds of both corpora – only 14 – is notable (see Table 9, Fig. 4). The more frequent shared terms in both corpora relate to ecclesiastical music. Notably, terms for singing voice and melody are scarce in media texts but abundant in *regilaul*.

Table 9. The occurrence of joint compounds in the text corpora of the Estonian Media Corpus and the *Regilaul* corpus.

Word	Translation	Occurrences of <i>laul-</i> stems	
		EMC	Regilaul corpus
<i>laulukoor</i>	singing choir	30	13
<i>lauluraamat</i>	songbook, hymnal	15	28
<i>kirikulaul</i>	church song	12	17
<i>lauluviis(ike(ne))</i>	(a little) song melody	3	147
<i>lauluhääl</i>	singing voice	2	189
<i>laulupidu</i>	(choir) song festival	40	3
<i>lastelaul</i>	children’s song	4	1
<i>laulusalm</i>	song stanza	2	6
<i>laululõng</i>	song yarn	1	9
<i>kiidulaul</i>	song of praise	1	4
<i>laulumäng</i>	song game/dance	1	1
<i>laulumeister</i>	master of song	1	2
<i>söömalaul</i>	eating song	1	2
<i>tänulaul</i>	thanksgiving song	1	12

In the *Regilaul* corpus, ‘church song’ appears 17 times exclusively in the “Shepherd’s Sunday” (*Karjase pühapäev*) song type, metaphorically representing (pars pro toto) Sunday church services in contrast to the shepherd’s daily work in

the pasture: “I did not hear the church song, but the song of a gentle cow” (*Ei mina kuulnud kirikulaulu, vaid aga hella lehma laulu* (H, Mapp 1050/2 (4)).

Singing in a choir is frequently mentioned in the press and occurs in *regilaul* texts, albeit with differing attitudes. Journalistic sources value choral singing, whereas *regilaul* singers often proudly disavow choir participation, affirming their acquisition of the “true” oral tradition.

The most frequent song-related compound in *regilaul* – *laululeht* ‘song sheet’ – means a memory aid for a recreating folk singer and is possibly influenced by ecclesiastical song sheets. Its frequency is attributed to its presence in the wide-spread song type, “Mouth’s Quill” (*Suude sulg*), which was published already in the nineteenth century (Kreutzwald 1840). The second most frequent compound, *laululind* ‘songbird’, appears in general singing contexts, often as a positive synonym for the singer. The poetic imagery of birds also portrays a rich and joyful nature: *igas ladvas laululinnud* ‘songbirds in every treetop’ (E 8249/50 (32) et alibi).

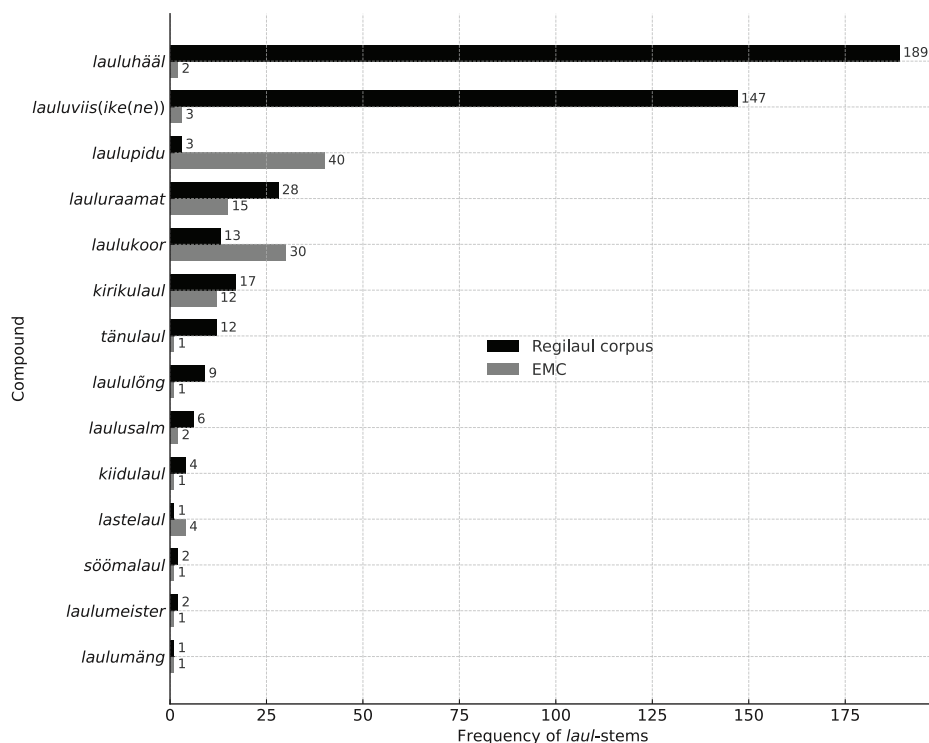


Figure 4. The occurrence of joint compounds in the Estonian Media Corpus and the Regilaul corpus.

The terms *laulja* and *laulik* ‘singer’ appear in both corpora. However, their usage differs: in *regilaul*, their meanings are hardly distinguishable, while written culture differentiates between *laulik* ‘folk singer, creator, poet’ and *laulja* ‘interpreter’.

Comparative analysis of words reveals how the Estonian Media Corpus and the Regilaul corpus represent distinct musical worlds. The EMC texts emphasise choral singing and its related social phenomena – societies, concerts, and festivals, while *regilaul* texts reflect the singer’s self-expression, communication in rural oral singing context. Both corpora show connections to Christianity and church culture.

CONCLUSION

This article employs text corpus analysis to explore Estonian concepts of singing across diverse social, cultural, and historical discourses. Focusing on the words with the stem *laul* ‘sing/song’, including its dialectal forms, across various Estonian corpora, we examined its frequency and semantic aspects, particularly in the 1890s media corpus (EMC) and *regilaul* (runosong) lyrics, which exhibit a significantly higher prevalence of *laul*-stem words compared to other corpora.

The EMC, while offering insights into choral and spiritual singing traditions, also provided a window into lesser-known communal singing practices, such as performances of the Russian Empire’s anthem and birthday songs at secular rituals. The analysis revealed how language patterns were related to national identity formation through culture.

The linguistically diverse Regilaul corpus required manual processing due to its complexity. The research involved the morphological unification of approximately 30,000 dialectal variants of *laul*-stem words, creating a list of nearly 1,500 standardised forms.

The verb-rich song corpus (63% of song/sing-stems) contrasts with noun-dominated (89%) newspaper texts. Our analysis also highlighted some Germanic loanwords in singing-related *regilaul* lyrics, especially *viis* ‘melody’, suggesting linguistic developments. The Regilaul corpus has a rich choice of compounds, while two corpora do not have many joint compounds, and among the existing ones the vocabulary related to church singing dominates.

This research shows the potential of corpus linguistics in folkloristics and ethnomusicology for analysing historical texts. It paves the way for further interdisciplinary studies in cultural and linguistic research.

ACKNOWLEDGMENTS

The study was supported by the Estonian Ministry of Education and Research (research project TK215) and the Estonian Research Council (PRG1288 “A corpus-based approach to folkloric variation: Regional styles, thematic networks, and communicative modes in runosong tradition”; EKKD-III2 “Patterns in the development of Estonian culture”; TK215 Estonian Roots: Centre of Excellence for Transdisciplinary Studies on Ethnogenesis and Cultural Diversity). We thank the Department of Folkloristics of the Estonian Literary Museum, especially Andres Kuperjanov, who found for us the necessary data from the database of calendar texts prepared by the Department of Folkloristics.

NOTES

- ¹ See <https://science-education-research.com/constructivism/>, last accessed on 29 October 2024.
- ² See <https://science-education-research.com/science-concepts/>, last accessed on 29 October 2024.
- ³ The idea that music is learned in a certain culture, together with cognitive and interpretive resources of music, emerged together with the field of comparative musicology (e.g., Ellis 1885; Hornbostel 1905), and has become a self-evident truth in today’s music philosophy (e.g., Levinson 1990; Ravasio n.d.).
- ⁴ The Estonian concept of *põlisrahvas* ‘indigenous people’ has evolved independently of the international definition of ‘indigenous’, which often refers to peoples living under foreign rule. This international definition does not fully capture the experience of Estonians, whose history has been characterised by long-standing inhabitation of their land and persistent struggles for independence against various foreign powers.
- ⁵ On Estonian history see, e.g., Kasekamp 2010.
- ⁶ The quotations from Estonian are translated by the authors of this article.
- ⁷ Verb-to-noun derivation or zero-derivation (Kasik 2012).
- ⁸ The word *loul* also occurs in Vepsian and Livonian, despite these languages lacking a documented runosong tradition. The same stem appears in Estonian dialects. A similar word also occurs in another Finno-Ugric language, Sámi: *lávlut* (ETY).
- ⁹ Approximate translations: to whoop, to call (special herding calls), to lament, to chant (a spell, an incantation).
- ¹⁰ The term “marked” has been contested (Haspelmath 2006), but it still seems useful in certain contexts.
- ¹¹ See <http://bit.ly/3zR8nJq>, last accessed on 29 October 2024.
- ¹² See <https://bit.ly/3BMsxEJ>, last accessed on 29 October 2024.
- ¹³ This corpus is not publicly accessible.
- ¹⁴ For example, only in the VAKK corpus the components of compounds are counted separately, e.g., *nelikümmend* ‘forty’ is two words *nelli*+ *+kümmend*.

- ¹⁵ See <https://www.cl.ut.ee/korpused/baaskorpus/1890/allikad/>, last accessed on 29 October 2024.
- ¹⁶ While the frequency of *laul* occurring independently versus in compounds is orthography-dependent and not central to our study, we provide these figures for both cases in Table 2 for reference.
- ¹⁷ However, this result was obtained through a word search conducted on the text file of the same corpus, because the KORP query interface does not support searches for partial words when variations in letters are present at the start or end of the word.
- ¹⁸ About the book of verses *Roosa kannel* (Pink Zither) by Rudolf Kallas in the editorial of *Postimees* (1899a).
- ¹⁹ Essay on the book of verses by Tamm, “Ärganud hääled II” (Olewik 1892).
- ²⁰ The author of the essay is likely Jaan Bergmann (1856–1916) – a pastor, poet, and folklore collector.
- ²¹ Jüri Orgusaar (1857–1912) was a playwright, teacher, and folk song collector.
- ²² The database contains about 6,000 transitional and end-rhymed songs.
- ²³ The type-index of the ERAB is incomplete. Many songs have preliminary type-names or lack classification data entirely. This should be considered when interpreting the statistics and categorisations of the ERAB. While the statistics of song types are not yet finalised, preliminary data demonstrate a significant number of songs that address the theme of singing.
- ²⁴ The song’s spread may have been aided by its publication in the popular book *Viina katk* (The Plague of Alcohol) (Kreutzwald 1840).
- ²⁵ Many Estonian words have lost their final vowel due to apocope, and word-internal vowel due to syncope (see Viitso 2003: esp. 183).
- ²⁶ Despite the data is based on observation, without statistics, the correlation between the words and song types is apparent.
- ²⁷ However, it is important to note that the editing process for the texts in the ERAB is still ongoing. Consequently, the number of compounds we identified may be lower than the actual count, as older transcriptions of folk song lyrics tended to favour separate word writing. Conversely, there is a possibility that modern editors, following the current trend towards solid writing, may have occasionally over-compounded words.
- ²⁸ *Välja* can function both as a noun (genitive form of *väli* ‘field’) and an adverb ‘outwards’.
- ²⁹ We counted only the most frequent forms, including *viis* (322), *viisid* (77), *viis* (41), *viisil* (22), and *viisisida* (19). There are also rarer forms, so the total count is higher.
- ³⁰ It is not always possible to distinguish between German, Low German, Old Norse, or Swedish loanwords in Estonian, coming into Estonian since about the ninth century (Metsmägi & Sedrik & Soosaar 2013). Numerous Low German loanwords are found in Estonian as well in *regilaul*, including *viis*, *kool* ‘school’, *raamatmaaker* ‘bookmaker’, and *saks* ‘noble, Saxon’.
- ³¹ Consider the verse “*Laulaksin ma lugu kaksi, veeretaksin viisi kuusi*”. Its primary interpretation, based on content logic and parallelism, is ‘I would sing two songs, would roll six tunes’. However, the choice of ‘six’ (*kuusi*) may be influenced by its proximity to ‘five’ (*viisi*). Consequently, an alternative interpretation of the second half could be ‘I would roll five-six [songs]’ (see Table 5).

³² *Laulusulg* – 1) presumably a small old-fashioned whistle (Oliver 2014; Lintrop 2017); 2) a long sickle-shaped feather on the tail of a cock (EKSS 2009).

³³ ‘Rhyme’ means firstly end-rhyme in Estonian, but in figurative and colloquial language it can probably also refer more generally to poetic means, similarly with the verb ‘to rhyme’, which can refer to creating poetry, and occasionally to improvising in *regilaul* style (see Kõiva 1964: 8).

ARCHIVAL SOURCES

Manuscript collections at the Estonian Folklore Archives of the Estonian Literary Museum

E – manuscripts of Matthias Johann Eisen

ERA – manuscripts of the Estonian Folklore Archives

EÜS – manuscripts of the Estonian Students’ Society

H – manuscripts of Jakob Hurt

RKM – manuscripts of the Folklore Department of the Estonian State Literary Museum, 1945–1996

TEM – manuscripts of the Estonian Museum in Tallinn

INTERNET SOURCES

CELL 2018 = The Corpus of Estonian Literary Language. Research Group of Computational Linguistics, University of Tartu. Available at <https://www.cl.ut.ee/korpused/baaskorpus/>, last accessed on 25 October 2024.

CWE 2018 = Newspaper Texts. Corpus of Written Estonian: the 1890s. Available at <https://www.cl.ut.ee/korpused/baaskorpus/1890/index.php?lang=en>, last accessed on 31 October 2024.

DWDS = “-weise”. *Digitales Wörterbuch der deutschen Sprache*. Available at <https://www.dwds.de/wb/-weise>, last accessed on 25 October 2024.

EKSS = *Eesti keele seletav sõnaraamat*. [The Explanatory Dictionary of the Estonian Language.] 2019. Available at <https://arhiiv.eki.ee/dict/eks/index.cgi?Q=laul&F=M&C06=et>, last accessed on 25 October 2024.

ERAB = *Eesti regilaulude andmebaas*. [Estonian Runosongs’ Database.] Comp. by Janika Oras & Liina Saarlo & Mari Sarv. Tartu: Eesti Kirjandusmuuseumi Eesti Rahvaluule Arhiiv. Available at <http://www.folklore.ee/regilaul/>, last accessed on 25 October 2024.

Koppel, Kristina & Kallas, Jelena 2020. *Eesti keele ühendkorpus 2019*. [Estonian National Corpus 2019.] <http://doi.org/10.1515/3-00-0000-0000-08565L>.

KORP = *Eesti Keeleressursside Keskuse päringusüsteem KORP*. [Query Interface KORP of the Center of Estonian Language Resources.] Available at <https://korp.keeleressurssid.ee>, last accessed on 25 October 2024.

- ETY = “laulma”. *Eesti etimoloogiasõnaraamat*. [Estonian Etymological Dictionary.] Available at <https://arhiiv.eki.ee/dict/ety>, last accessed on 25 October 2024.
- Mamma Mia = Ootame sind laulma. [Come and Sing.] *Mamma Mia Laulu- ja Huvistudio*. Available at https://www.mammamia.ee/ootame_sind_laulma/, last accessed on 25 October 2024.
- Lock, Gerhard 2019. Eestikeelse muusikaterminivara areng 19. ja 20. sajandil. [Development of Estonian Music Terminology in the 19th–20th Centuries.] *E-Koolikott*. Available at <https://e-koolikott.ee/kogumik/19540-14-Eestikeelse-muusikaterminivara-areng-19-ja-20-sajandil>, last accessed on 25 October 2024.
- SKVR-Tietokanta = *Karjalaisten, inkeriläisten ja suomalaisten kansanrunojen verkkopalvelu*. [Online Service of Karelian, Ingerian and Finnish Folk Poems.] Helsinki: SKS. Available at <https://aineistot.finlit.fi/exist/apps/skvr/index.html>, last accessed on 25 October 2024.
- VAKK = *Vana kirjakeele korpus*. [The Corpus of Old Literary Estonian.] University of Tartu. Available at <https://vakk.ut.ee/>, last accessed on 29 October 2024.

REFERENCES

- Bennett, Gena R. 2010. *Using Corpora in the Language Learning Classroom. Corpus Linguistics for Teachers*. Ann Arbor: University of Michigan Press. <http://dx.doi.org/10.3998/mpub.371534>.
- Blacking, John 1995. *Music, Culture, and Experience. Selected Papers of John Blacking*. Edited by Reginald Byron. Chicago: University of Chicago Press.
- Davidjants, Brigitta 2022. Women’s Experience in Estonian Punk Scenes during the Transition from Soviet to Post-Soviet Society. *Punk & Post-Punk*, Vol. 11, No. 3, pp. 307–330. https://doi.org/10.1386/punk_00145_1.
- Davidjants, Brigitta 2024. Contesting Narratives of Post-Soviet Estonian Fandoms: The Case of the Estonian Depeche Mode Fan Club. In: Jeffers Engelhardt & Kate Pukinskis (eds.) *Baltic Musics Beyond the Post-Soviet*. Tartu: University of Tartu Press, pp. 112–142. Available at <https://www.academia.edu/121104958/>, last accessed on 25 October 2024.
- Ellis, Alexander J. 1885. On the Musical Scales of Various Nations. *The Journal of the Society of Arts*, Vol. 33, No. 1688, pp. 485–527.
- ERL 1970 = *Eesti rahvalaulud: Antoloogia 2*. [Estonian Folksongs: Anthology.] Ed. by Ülo Tedre. Tallinn: Eesti Raamat.
- Finnegan, Ruth 1977. *Oral Poetry: Its Nature, Significance and Social Context*. Cambridge: Cambridge University Press.
- Haspelmath, Martin 2006. Against Markedness (and What to Replace It with). *Journal of Linguistics*, Vol. 42, No. 1, pp. 25–70. <http://dx.doi.org/10.1017/S0022226705003683>.
- Hennoste, Tiit & Kaalep, Heiki-Jaan & Muischnek, Kadri & Paldre, Leho & Vaino, Tarmo 2001. The Tartu University Corpus of Estonian Literary Language. In: Tõnu Seilenthal & Anu Nurk & Triinu Palo (eds.) *Congressus Nonus Internationalis Fenno-Ugristarum, IV: Congressus Nonus Internationalis Fenno-Ugristarum*, Tartu

- 7.–13.08.2000. Tartu: Eesti Fennougristide Komitee, pp. 337–344. Available at https://cl.ut.ee/yllitised/smugri_korpus_2001.pdf, last accessed on 5 November 2024.
- Honko, Lauri (ed.) 2000. *Thick Corpus, Organic Variation and Textuality in Oral Tradition*. Helsinki: Finnish Literature Society.
- Hornbostel, Erich Moritz von 1905. Die Probleme der vergleichenden Musikwissenschaft. In: Alfred Heuß (ed.) *Zeitschrift der Internationalen Musikgesellschaft*, Vol. 7, No. 3, pp. 85–97. Available at <https://archive.org/details/ZeitschriftDerInternationalenMusikgesellschaft071905-06/page/84/>, last accessed on 25 October 2024.
- Janicki, Maciej & Kallio, Kati & Sarv, Mari 2023. Exploring Finnic Written Oral Folk Poetry through String Similarity. *Digital Scholarship in the Humanities*, Vol. 38, No. 1, pp. 180–194. <https://doi.org/10.1093/lc/fqac034>.
- [Jannsen, Johann Woldemar] 1857. Veel nattuke Tallinna laulo piddust. [Some More on Tallinn Song Festival.] *Perno Postimees*, No. 14, 2 October, pp. 109–111. Available at <https://dea.digar.ee/?a=d&d=pernopostimees18571002.2.3&e>, last accessed on 25 October 2024.
- Kalkun, Andreas & Aabrams, Vahur (comps. & eds.) 2014. *Seto naisi elolaulu': Antoloogia*. [Biographical Songs of Setu Women.] Seto Kirävara 8. Tartu: Seto Instituut & Eesti Kirjandusmuuseum, pp. 11–16.
- Kalkun, Andreas & Oras, Janika 2018. “Stalin is a wise man, Lenin was a little bird.” On Creating Soviet Folklore in the Seto Region during the Stalin Era. *Res Musica*, Vol. 10, pp. 41–62. Available at https://resmusica.ee/wp-content/uploads/2018/11/rm10_2018_41-62_Kalkun_Oras.pdf, last accessed on 25 October 2024.
- Kasekamp, Andres 2010. *A History of the Baltic States*. Basingstoke & New York: Palgrave Macmillan.
- Kasik, Reet 1996. *Eesti keele sõnatuletus*. [Word Derivation in Estonian.] Tartu Ülikooli eesti keele õppetooli toimetised 3. Tartu: Tartu Ülikooli Kirjastus. Available at <https://dSPACE.ut.ee/server/api/core/bitstreams/c08678af-0ed6-46ec-998a-f4ed3efc96f2/content>, last accessed on 28 October 2024.
- Kasik, Reet 2012. Nulltuletus ja konversioon eesti keeles. [Zero-Derivation and Conversion in Estonian.] *Keel ja Kirjandus*, Vol. 11, pp. 793–806. <https://doi.org/10.54013/kk660a1>.
- Kasik, Reet 2015. *Sõnamoodustus*. [Word Formation.] Tartu: Tartu Ülikooli Kirjastus. Available at <https://dSPACE.ut.ee/server/api/core/bitstreams/1c23f4c2-9585-4bfc-b529-f92acab17c3c/content>, last accessed on 31 October 2024.
- Kinney, Phyllis 2015 [2011]. *Welsh Traditional Music*. Cardiff: University of Wales Press in association with Cymdeithas Alawon Gwerin Cymru.
- Kõiva, Otilie 1964. *Kihnu rahvalaulikutest*. [On Folksingers of Kihnu.] Tartu Riikliku Ülikooli Toimetised 159. Tartu: Tartu Riiklik Ülikool. Available at <https://dSPACE.ut.ee/server/api/core/bitstreams/472e9a4c-071f-4454-9235-c1b92ea950dd/content>, last accessed on 28 October 2024.
- Koppel, Kristina 2020. Näitelausete korpuspõhine automaattuvastus eesti keele õppesõnastikele. [Corpus-Based Automatic Detection of Example Sentences for Dictionaries for Estonian Learners.] Tartu: Tartu Ülikooli Kirjastus. Available at <https://dSPACE.ut.ee/items/a6a94cd7-97cd-4b5f-8bff-23e85ae375a3>, last accessed on 28 October 2024.

- Kreutzwald, Friedrich R. 1840. *Wina katk: Üks tähhele pannemise wäärt jut, mis wanna ja noore rahwale juhhatusseks ja kassuks wälja on antud*. [The Plague of Vodka. A Remarkable Story for Guidance and Benefit of Elderly and Young People.] Tartu: Lindfors'i pär. Available at <https://www.digar.ee/arhiiv/et/kollektsioonid/14567>, last accessed on 28 October 2024.
- Kuutma, Kristin 1996. Cultural Identity, Nationalism and Changes in Singing Traditions. *Folklore: Electronic Journal of Folklore*, Vol. 2, pp. 124–141. <http://dx.doi.org/10.7592/FEJF1996.02.ident>.
- Laar, Mart 2001. Oleviku esimene tegevusaasta. [The First Year of Olevik.] *Acta Historica Tallinnensia*, Vol. 5, No. 1, pp. 75–92. <https://doi.org/10.3176/hist.2001.1.04>.
- Labi, Kanni 2006. *Eesti regilaulude verbisemantika*. [The Verb Semantics of Estonian Runosongs.] Tartu: Tartu Ülikooli Kirjastus. Available at <https://dspace.ut.ee/items/8dd3028b-1b15-4e42-8842-ee0da6e0a712>, last accessed on 28 October 2024.
- Langacker, Ronald W. 2002 [1990]. *Concept, Image, and Symbol: The Cognitive Basis of Grammar*. 2nd edition. Berlin & New York: Mouton de Gruyter.
- Lauk, Epp & Pallas, Anu 2008. Early Development of Estonian Journalism as a Profession (from Late 19th Century to 1940). *Acta Historica Tallinnensia*, Vol. 13, No. 1, pp. 48–67. <http://dx.doi.org/10.3176/hist.2008.2.03>.
- Levinson, Jerrold 1990. The Concept of Music. In: Jerrold Levinson. *Music, Art, and Metaphysics: Essays in Philosophical Aesthetics*. New York: Cornell University Press, pp. 267–278.
- Lintrop, Aado 2017. Where Do Songs Come From? An Attempt to Explain Some Verses of *Regilaul*. *Folklore: Electronic Journal of Folklore*, Vol. 68, pp. 115–130. <http://dx.doi.org/10.7592/FEJF2017.68.lintrop>.
- Lippus, Urve 1995. *Linear Musical Thinking: A Theory of Musical Thinking and the Runic Song Tradition of Baltic-Finnish Peoples*. Studia Musicologica Universitatis Helsingiensis VII. Helsinki: University of Helsinki, Department of Musicology.
- Lippus, Urve 2006. Kirjeldused Ruthi muusikaharrastustest kui eesti muusikaloo allikas. [Descriptions of Ruth's Musical Pursuits as a Source of Estonian Music History.] In: Mirjam Hinrikus (ed.) *J. Randvere "Ruth" 19.–20. sajandi vahetuse kultuuris*. Tallinn: Underi ja Tuglase Kirjanduskeskus, pp. 40–54.
- Lippus, Urve 2012. Klaver eesti kodus. [The Arrival of the Piano in the Estonian Home.] *Res Musica*, Vol. 4, pp. 7–21. Available at <https://resmusica.ee/en/resmusica-4-2012/>, last accessed on 28 October 2024.
- Merriam, Alan P. 1964. *The Anthropology of Music*. Evanston, IL: Northwestern University Press.
- Metsmägi, Iris & Sedrik, Meeli & Soosaar, Sven-Erik 2013. Eesti kirjakeele tüvevara päritolu arvudes. [Origin of Literary Estonian Stemware in Numbers.] *Keel ja Kirjandus*, Vol. 5, pp. 313–332. <https://doi.org/10.54013/kk666a1>.
- Nordquist, Richard 2019a. What Is Markedness in Language? *ThoughtCo*, 3 July. Available at <https://www.thoughtco.com/markedness-language-term-1691302>, last accessed on 28 October 2024.
- Nordquist, Richard 2019b. What Is Semantic Field Analysis? *ThoughtCo*, 30 April. Available at <https://www.thoughtco.com/semantic-field-analysis-1691935>, last accessed on 28 October 2024.

- Olewik 1892 = Kirjandus. Seltsielu. Segasõnumed. [Literature. Social Life. Varia.] *Olewik*, No. 43, 26 October, pp. 879–882. Available at <https://dea.digar.ee/?a=d&d=olewik18921026.1.4&e>, last accessed on 28 October 2024.
- Olewik 1894 = Sõnumed. Valitsus. Kord. Kohus. [Messages. Government. Order. Court.] *Olewik*, No 43, 25 October, p. 916. Available at <https://dea.digar.ee/?a=d&d=olewik18941025&e>, last accessed on 28 October 2024.
- Olewik 1895a = Vaimulikud laulud ehk Tallinnakeelelise kiriku lauluraamatu prov. [Spiritual Songs or a Test of the Hymnal in North-Estonian.] *Olewik*, No. 28, 11 July, pp. 652–654. Available at <https://dea.digar.ee/?a=d&d=olewik18950711&e>, last accessed on 28 October 2024.
- Olewik 1895b = Teised lehed. [Other Newspapers.] *Olewik*, No. 28, 11 July, pp. 660–661. Available at <https://dea.digar.ee/?a=d&d=olewik18950711&e>, last accessed on 28 October 2024.
- Oliver, Paul 2014. Quills. *Grove Music Online*, 31 January. Available at <https://www.oxfordmusiconline.com/grovemusic/search?q=Quills&searchBtn=Search&isQuickSearch=true>, last accessed on 28 October 2024.
- Oras, Janika 2008. *Viie 20. sajandi naise regilaulumaailm. Arhiivitekstid, kogemused ja mälestused*. [The Regilaul World of Five 20th-Century Women: Archival Texts, Experiences and Memories.] Tartu: Eesti Kirjandusmuuseumi Teaduskirjastus.
- Oras, Janika 2017. Favourite Children and Stepchildren: Elite and Vernacular Views of Estonian Folk Song Styles. *Res Musica*, Vol. 9, pp. 27–44. Available at <https://resmusica.ee/en/res-musica-9-2017/>, last accessed on 28 October 2024.
- Oster, Ulrike 2010. Using Corpus Methodology for Semantic and Pragmatic Analyses: What Can Corpora Tell Us about the Linguistic Expression of Emotions? *Cognitive Linguistics*, Vol. 21, No. 4, pp. 727–763. <https://doi.org/10.1515/cogl.2010.023>.
- Pallas, Anu 2000. Ajalehe kirjutajast ajakirjanikuks: ajakirjanikuameti kujunemine Eestis aastatel 1878–1918. [From a Newspaper Writer to a Journalist: The Formation of a Journalistic Profession in Estonia.] In: Epp Lauk & Maret Mälk & Anu Pallas (eds.) *Peatükke Eesti ajakirjanduse ajaloost 1900–1940*. Tartu: Tartu Ülikooli Kirjastus, pp. 146–164.
- Peegel, Juhan 2004. *Nimisõna poeetilised sünonüümid eesti regivärssides*. [The Poetical Synonyms of Nouns in Estonian Regilaul.] Edited by Ülo Tedre. Tallinn: Eesti Keele Sihtasutus.
- Postimees 1896 = Lätimaalt. Valgast. [Latvia. Valka.] *Postimees*, No. 30, 6 February. Available at <https://dea.digar.ee/?a=d&d=postimeesew18960206&e>, last accessed on 28 October 2024.
- Postimees 1897 = Küllalt nüüd! [Enough!] *Postimees*, No. 27, 3 February. Available at <https://dea.digar.ee/?a=d&d=postimeesew18970203&e>, last accessed on 28 October 2024.
- Postimees 1899a = Tuulelipp. [Vane.] *Postimees*, No. 171, 7 August. Available at <https://dea.digar.ee/?a=d&d=postimeesew18990807.2.2&e>, last accessed on 28 October 2024.
- Postimees 1899b = Soomemaalt. Helsingist. [From Finland. From Helsinki.] *Postimees*, No. 171, 7 August. Available at <https://dea.digar.ee/?a=d&d=postimeesew18990807.2.2&e>, last accessed on 28 October 2024.

- [Raud, Märt] 1912 = Md. 1912. Üks vanameelne ettepanek laulusõpradele. [An Outdated Proposal to the Friends of Singing.] *Postimees*, No. 185, 13 August. Available at <https://dea.digar.ee/?a=d&d=postimeesew19120813.2.3&e>, last accessed on 28 October 2024.
- Ravasio, Matteo n.d. Analytic Perspectives in the Philosophy of Music. *The Internet Encyclopedia of Philosophy*. Available at <https://iep.utm.edu/music-an/>, last accessed on 28 October 2024.
- Rickert, Heinrich 1902 [1896]. *Die Grenzen der naturwissenschaftlichen Begriffsbildung: Eine logische Einleitung in die historischen Wissenschaften*. Fünfte Auflage. Tübingen: Verlag von J.C.B. Mohr.
- Riemer, Nick 2010. *Introducing Semantics*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511808883>.
- Rüütel, Ingrid 1998. Estonian Folk Music Layers in the Context of Ethnic Relations. *Folklore: Electronic Journal of Folklore*, Vol. 6, pp. 32–69. <http://dx.doi.org/10.7592/FEJF1998.06.ruutel>.
- Rüütel, Ingrid 2013. *Naised Kihnu kultuuris*. [Women in Kihnu Culture.] Tartu: EKM Teaduskirjastus.
- Rüütel, Ingrid & Tiit, Ene-Margit 2005–2006. *Pärimuskultuur Eestis – kellele ja milleks?* [Traditional Culture in Estonia – For Whom and for What?] Vols. 1–2. Tartu: Tartu Ülikooli Kirjastus.
- Saag, Lehti & Laneman, Margot & Varul, Liivi & Lang, Valter & Metspalu, Mait & Tambets, Kristiina 2019. The Arrival of Siberian Ancestry Connecting the Eastern Baltic to Uralic Speakers Further East. *Current Biology*, Vol. 29, No. 10, pp. 1701–1711. <http://dx.doi.org/10.1016/j.cub.2019.04.026>.
- Saarlo, Liina 2023. Need teised laulikud: Regilaulutraditsiooni vähem aktiivsetest osalistest. [The Other Singers: The Less Active Representatives of the Runosong Tradition.] *Keel ja Kirjandus*, Vol. 6, pp. 575–594. <https://doi.org/10.54013/kk786a2>.
- Saarlo, Liina & Sarv, Mari & Mett, Susanna 2022. The Language of Sorrow and Joy: Emotional Vocabulary and Formulae in Runosongs. Conference abstract. In: Janika Oras & Liina Saarlo (eds.) *From Desperation to Hope: The Meanings and Effect of Group Singing*. November 29 – December 1, 2022, Estonian Literary Museum, Tartu and MD Teams. Tartu: ELM Scholarly Press, pp. 48–49. Available at <https://www.folklore.ee/regilaul/konverents2022/abstracts.pdf>, last accessed on 28 October 2024.
- Särg, Taive 2009. Covers and Folklore: Contra's Songs. In: Mare Kõiva (ed.) *Media and Folklore: Contemporary Folklore IV*. Tartu: ELM Scholarly Press, pp. 347–389. Available at https://www.folklore.ee/ri/pubte/ee/cf/cf4/CF4_Sarg.pdf, last accessed on 28 October 2024.
- Särg, Taive 2023. The Institutionalisation of Participatory Singing Since the 1960s in Estonia. *Traditiones*, Vol. 52, No. 2, pp. 125–148. <http://dx.doi.org/10.3986/Traditio2023520207>.
- Sarv, Mari 2017. Towards a Typology of Parallelism in Estonian Poetic Folklore. *Folklore: Electronic Journal of Folklore*, Vol. 67, pp. 65–92. <http://dx.doi.org/10.7592/FEJF2017.67.sarv>.

- Schlegel, Christian Hieronymus Justus 1831. *Reisen in mehrere Russische Gouvernements in den Jahren 1801, 1807, 1815, 1830*. Sechstes Bändchen. Meiningen: Friedemann Keyssner. Available at <https://dspace.ut.ee/items/1ea49c46-6480-4108-aaf1-76711f52b9d3>, last accessed on 28 October 2024.
- Slobin, Mark 2011. *Folk Music: A Very Short Introduction*. Oxford University Press.
- Taber, Keith S. 2019–2024. *Science-Education-Research*. Available at <https://science-education-research.com/>, last accessed on 24 October 2024.
- Tampere, Herbert 1935. *Eesti rahvaviisiide antoloogia / Anthologia cantuum popularium Estonorum*. Vol. 1. Tartu: Eesti Rahvaluule Arhiiv. Available at <https://www.etera.ee/zoom/9514/view?page=1&p=separate&tool=info&view=0,0,1969,3197>, last accessed on 28 October 2024.
- Tampere, Herbert (comp.) 1956. *Eesti rahvalaule viisidega*. [Estonian Folksongs with Melodies.] Vol. 1. Töölaulud. Tallinn: Eesti Riiklik Kirjastus.
- Tampere, Herbert (comp.) 1965. *Eesti rahvalaule viisidega*. [Estonian Folksongs with Melodies.] Vol. 5. Tallinn: Valgus.
- Teubert, Wolfgang 2009. Corpus Linguistics: An Alternative. *Semen: Revue de sémiolinguistique des textes et discours*, Vol. 27. <https://doi.org/10.4000/semen.8912>.
- Uus Lauluraamat 1900 = *Uus Lauluraamat: kirikus, koolis ja kodus pruukida*. [New Hymnal: To Use at Church, School and Home.] Tallinn: Mickwitz.
- Vahter, Artur 2003. *Eestikeelse muusikaterminoloogia kujunemine ja areng*. [Formation and Development of Music Terminology in Estonian.] Tallinn: Scripta Musicalia.
- Viitso, Tiit-Rein 2003. Rise and Development of the Estonian Language. In: Mati Erelt (ed.) *Estonian Language, Linguistica Uralica: Supplementary Series*, Vol. 1, pp. 130–230. Available at <https://dspace.ut.ee/server/api/core/bitstreams/adbc35b-935e-4ef8-afd3-e3bf44f79040/content>, last accessed on 28 October 2024.
- Vissel, Anu 2004. *Lastepärimus muutuvast ühiskonnas*. [Child Tradition in a Changing Society.] Tartu: Eesti Kirjandusmuuseum, etnomusikoloogia osakond. Available at <https://dspace.ut.ee/items/b262c5f3-e8c6-4f1e-9a25-2ab56cc047ba>, last accessed on 28 October 2024.
- Wirmaline 1892 = Kodumaalt. Võrust. [From Homeland. From Võru.] *Wirmaline: St. Peeterburi Eesti ajaleht*, No. 23, 2 June. Available at <https://dea.digar.ee/?a=d&d=wirmalinepeterburi18920602&e>, last accessed on 28 October 2024.
- Wyler, Siegfried 1992. *Colour and Language: Colour Terms in English*. Tübingen: Gunter Narr Verlag.

Taive Särg, PhD in Estonian and comparative folklore (University of Tartu, 2005), is a Senior Research Fellow at the Estonian Folklore Archives of the Estonian Literary Museum, Estonia. Her main research interests lie in the field of ethnomusicology and include Estonian folk songs, their melodies and singing styles, the historical and contemporary contexts of singing, as well as the social processes and ideologies related to (folk) music and its study.

taive.sarg@folklore.ee

Taive Särg, Kaarel Veskis

Kaarel Veskis is a PhD student at the University of Tartu and Junior Research Fellow at the Estonian Folklore Archives of the Estonian Literary Museum. His research in digital humanities specializes in the computational analysis of the language of Estonian runosongs. He holds a BA in Estonian Literature and a MA in General Linguistics.

kaarel.veskis@kirmus.ee