Heini Arjava & Gerrit Kentner

Alignment of prosodic weight and musical length in Finnish vocal music textsetting

Abstract: In the study of textsetting, i.e., the alignment of lyrics to music in song writing, the rhythmic treatment of prosodic syllable weight in relation to note length has so far received less attention than other areas of prosody, such as stress prominence and pitch. This study offers an exploratory analysis of textsetting in a Finnish song corpus, focusing on the alignment of note length in music with the prosodic weight and segmental filling of the corresponding syllables. Music is rhythmically exceptionally versatile, and the present perspective on Finnish is original in that it allows to study the relationship of prosodic weight and note length in textsetting while controlling for lexical stress and musical prominence. Here, note length is conceived not only in terms of absolute note values but relative to neighbouring note values as well, with long notes considered particularly long when followed by short notes. The results of the corpus analysis suggest that song writers align prosodic weight and musical length, with light syllables particularly avoided on notes that are long both in absolute as well as in relative terms. Over and above prosodic weight, sonority and certain segmental features are shown to affect the alignment of syllables with notes. In sum, this study regarding the phonological choices in textsetting sheds light on how song writers align the durational prosody of lyrics with musical length in vocal music.

Keywords: prosody, vocal music, textsetting, quantity, length, weight, metrics, lengthening, asymmetries

1 Introduction

Consider this opening phrase of the Gloria suite of Mass C by Ludwig van Beethoven in Figure 1, the first syllable spanning the musical duration of no less than two to three seconds. This musical phrase, where notes differ greatly in length, illustrates the peculiarity of our traditions of singing: contrary to regular

Heini Arjava, Department of Languages, University of Helsinki, Finland
Gerrit Kentner, Department of Linguistics, Goethe-University Frankfurt, Germany, and Max Planck Institute for Empirical Aesthetics, Frankfurt, Germany

https://doi.org/10.1515/9783110770186-007

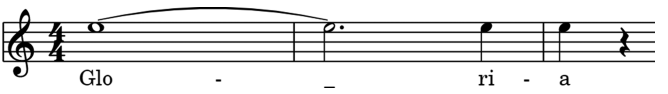


Figure 1: *L. v. Beethoven*: Gloria, Mass in C.

speech, sound segments, and especially vowels, can be held in singing for what phonetically counts as an eternity.

This article aims to examine how song writers align verbal material in the lyrics with note length in vocal music. In this task, we specifically focus on environments where length is asymmetrically distributed over regularly occurring musical boundaries, also presented by the Beethoven passage above. At the outset, a terminological clarification is in order: "length" refers to abstract duration. In the linguistic domain, length refers to the quantity of phonological units like segments or moras; in music, note length equals the note value in the notational system. Hence, with the term length, we abstract away from actual duration in speech or note durations in the musical performance.

Singing and speech are both essential for the human condition, and they share a few conspicuous characteristics of production and organization: Processes of speech and singing draw on the same neurological substrate (cf. Patel, 2003, 2010), and both domains are based on fine-grained phonetic production in the human vocal tract. Most importantly for the present discussion, they tend to share similar principles of rhythmic organization, the basis of which is a hierarchically graded accentual prominence, manifesting itself as *meters* in musical and poetic traditions. However, singing and speech also differ in some key aspects of rhythmic and vocal production, most notably in their use of length and pitch variation.

Musical and linguistic vocal production interact most clearly when the two are combined in a shared rhythmic and temporally regulated frame, a song. *Textsetting*, that is, the arrangement of lyrics into music, offers insights into the principles that govern the interaction of language and music in song composition. Here, we focus on the alignment in textsetting of musical note length with one of its linguistic counterparts, namely *prosodic weight* that is determined by the quantity of phones in the syllable, whereas *stress* refers to the increased emphasis of a syllable. One challenge regarding this length alignment lies in the fact that prosodic weight is represented on an ordinal scale with two or at most three categories (light vs heavy vs super-heavy); music, in contrast, measures note length on an interval scale, i.e., there is no a priori limit to the number of categories that can be distinguished. Therefore, the alignment cannot be fully deterministic. Still, our study reveals clear correlations suggesting that

song writers consider musical length and prosodic weight together when aligning text and tune, aiming to avoid conflicts between them.

In the remainder of the introduction, we briefly review previous musico-linguistic work on textsetting and then introduce critical phonological features of the language under study, Finnish. Section 2 provides background on rhythm in metrical poetry and music with a focus on the role of length or quantity in the two systems. In Section 3, we describe the construction of the Finnish song corpus which serves as the basis for the analysis presented in Section 4. Section 5 concludes the paper.

A cornerstone in the field of rhythmic musico-linguistics was laid by Lerdahl & Jackendoff (1983), who adopted generative grammar in the metrical analysis of Western (instrumental) music. Following this tradition, the alignment of musical prominence and linguistic stress or accent has been studied in various song corpora (e.g., Palmer & Kelly, 1992, Kiparsky, 2006, Dell & Halle, 2009, Rodríguez-Vázquez, 2010a, 2010b, Temperley & Temperley, 2013, Proto & Dell, 2013, deCastro Arrazola, 2018, and Girardi & Plag, 2022). Broader literature-based reviews on musico-linguistic case-studies are provided, for instance, by Proto (2015) and McPherson (2019). Only few studies have examined the phonological length or phonetic duration of syllables in language in relation to note length in music: Temperley & Temperley (2011) investigated syllable and note durations comparing speech and corpora of music. Hayes & Kaun (1996) formulated a general preference rule for the alignment of musical and syllabic length in the vicinity of prosodic phrase boundaries; and Rodríguez-Vázquez (2010a) expanded the discussion by making important qualitative remarks on musical length in her wide-ranging volume on rhythm in English and Spanish. However, these studies investigate length as either a correlate of linguistic stress/accent or prosodic phrasing, and mostly disregard the inherent phonological length of the syllable, i.e., the prosodic weight.

The present study is one of the first empirical studies systematically exploring phonological length in textsetting in terms of prosodic weight while controlling for the influence of linguistic stress or accent.[1] In order to disentangle effects of linguistic stress/accent and prosodic weight, we study textsetting in Finnish. Finnish is a quantity language in which length or prosodic weight is relatively independent of stress. Several stress-independent properties make

1 In this study, we do not discuss the role of musical pitch (melody), which is not a prosodically salient feature in Finnish (Suomi et al., 2008:76). Generally, the less important role of pitch compared to the other prosodic domains in popular tunes has been noted by Hayes & Kaun (1996) and Girardi & Plag (2022). Pitch peaks in textsetting have been treated by, for instance, Domene Moreno & Kabak (2022).

phonological quantity a clearly discernible part of the Finnish prosodic system (see, i.e., Suomi et al., 2008).

i) Quantity of both vowels and consonants is phonemic, that is, it creates semantic distinctions (cf. *tuli* 'fire', *tuuli* 'wind', *tulli* 'customs') (id: 76). Beyond underlying geminates, long consonants also arise as a common sandhi process, making geminates a particularly salient feature of the language.[2]
ii) Syllable weight is phonologically independent of stress: the primary stress is weight-independent and fixed on the first syllable (cf. '*tu.lee*, '(s)he comes') (id: 75).
iii) All segments in the syllable rhyme (i.e., nucleus vowel and coda consonant) are *moraic* (μ) and contribute equally to the prosodic weight of a syllable. Prosodic weight in turn determines wordhood: A minimal word, for instance, consists of at least two voiced moras (id: 71).
iv) The phonetic salience of sound segments varies relatively little, as unstressed segments are not heavily reduced in speech (Suomi, 2007), contrary to some weight-sensitive languages, such as English or German (Goedemans & van der Hulst, 2013).

However, prosodic weight is not wholly independent of stress in Finnish either, as evidenced by phonological quantity-based regulations on secondary stress, or the phonetic lengthening of the primary stress placed on the first two moras of a polysyllabic word, irrespective of syllabic association (Suomi et al., 2003, Suomi & Ylitalo, 2004). Length and stress also interact in metrical verse: long syllables with two or more moras are preferred on metrically strong positions in classical Finnish poetry, and there are stricter length-based rules on heavy positions and primary-stressed syllables in the archaic Kalevala meter (Ryan, 2017, 2019:147). Altogether, however, quantity is an audible and important feature in the prosodic system of Finnish, which, as the hypothesis goes, makes also the listeners of Finnish vocal music sensitive to misalignments between musical and linguistic length.

In this paper, we set out to study the alignment between musical note length and the prosodic weight of the corresponding syllable and were particularly interested in how it manifested in the comparison of durationally symmetric and asymmetric musical contexts (to be discussed in Section 2.2.). The analysis of a corpus of popular Finnish vocal music demonstrates that song writers align musical length and prosodic weight when assigning syllables in the lyrics to notes

2 Note that in Finnish, a geminate actually consists of two adjacent identical phonemes straddling a syllable boundary (e.g., *kuk.ka* 'flower') (Suomi et al., 2008:40). The syllable boundary inside the geminate is perceptually arbitrary, although its phonetic nature is still partly an open question (cf. Lehiste, 1970:44).

in the tune. We also find certain segmental features of the syllables to affect the assignment to notes, although their interpretation remains more ambiguous.

2 Rhythmic conflicts between language and music

2.1 Music and spoken verse: Differences between the durational systems

The rhythmic organization of both spoken verse and vocal music of the Western tradition is built on similar metrical templates in which prominent and less prominent elements alternate systematically and combine to form metrical hierarchies of various depths (see the classic discussion in Lerdahl & Jackendoff, 1983, among others). In poetry, metrical building-blocks are traditionally called *weak* and *strong* positions, which build up the higher units of poetic feet Figure 2.[3]

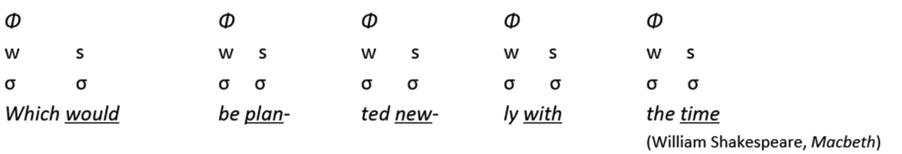


Figure 2: Positions and hierarchies of prominence in a poetic line in iambic pentameter (σ = syllable, w=weak position, s=strong position, Φ=poetic foot).

In traditional musical terminology, on the other hand, the analogous basic strong positions of the metrical template are called *beats*, which are grouped into *tacti*, i.e., the higher level in the prominence hierarchy. The tactus (see the note examples in 3 below) refers to the most prominent recurring beat of a musical composition, which the listener, with or without musical education, can feel as the basic 'ticking', or the *pulse* of the piece (cf. Lerdahl & Jackendoff,

3 Accentual poetry is based on the alignment of metrically prominent positions and linguistic stresses, with some less strict preferences regarding syllable weight (cf. Leino, 1982:278–287). Since the prototypical Western art song and its lyrics first and foremost follow a similar accentual meter based on stress prominence, we do not expand here on the other typical determiners of poetic organization, such as syllable count, as in the French *Alexandrine*, or phonological quantity such as in the Classical *hexameter* or the Finnish *Kalevala meter*.

1983:70–74 and passim). In other words, tactus could be described as the basic perceptual unit that constitutes the rhythmic structure of a musical passage.

Determining the perceptual distinction between the basic beat unit and the tactus is notoriously challenging (cf. id.:73), but the distinction is important to make if one needs to find a common metrical ground to compare length variation within or between songs, which is the aim of this study. Musical lengths are measured in beats, but not all beats are the same because of the song-specific variation of the tactus. The *time signature* of a piece is a good but not a definite indicator of the tactus level, as will be illustrated in the examples Figure 3 below. The time signatures are expressed as fractions, the most typical ones of which include 4/4, 3/4, and 6/8. The time signature implies how many beats are included in a bar (a further musical unit above the tactus-level, visually separated by vertical lines). In a 3/4 time signature, for instance, a bar includes three quarter-note-long beats (cf. Figure 3c below), and the 6/8 time signature denotes six eighth-note beats in a bar.

In ternary time signatures, the tactus level is especially challenging to determine, as the asymmetric prominence structure does not mathematically indicate a clear secondary prominence level besides the primary musical stress of the first note. Based on the overall rhythmic structure of a song, it is sometimes necessary in ternary compositions to qualitatively determine if the tactus equals one beat as in Figure 3c, or, as for instance in the case of many waltzes, a whole three-beat bar.

As for length, unlike spoken poetry where a syllable usually fills only one or two metrical positions, syllables in songs can be lengthened to cover any number of metrical positions in music. In Figure 3a, the first syllable of the Finnish lyrics (*kuu-*) takes up three short 1/16 notes (see level 0, the first two 1/16s not marked in the grid but implied on level 1), whereas in Figure 3b, the syllable *ai* in the second bar takes up two long quarter-note beats (level 1), and one whole tactus (level 2). The metrical depiction below follows the convention of earlier musico-linguistic studies, where metrical positions have typically been presented as x-grids following Liberman (1975) and introduced to music by Lerdahl & Jackendoff (id.).

As indicated above, both music and poetry operate on positional templates. But there are differences regarding the typical principles of how these abstract musical and linguistic meters generate rhythm. This poses challenges for song writers when aligning text and tune in actual textsetting, as well as for the theoretical comparison of the durational differences between language and music. A key difficulty is that, while rules of accentual/stress prominence of verse and music tend to be similar, there are more fine-grained differences when it comes to length variation, both within and between the two domains.

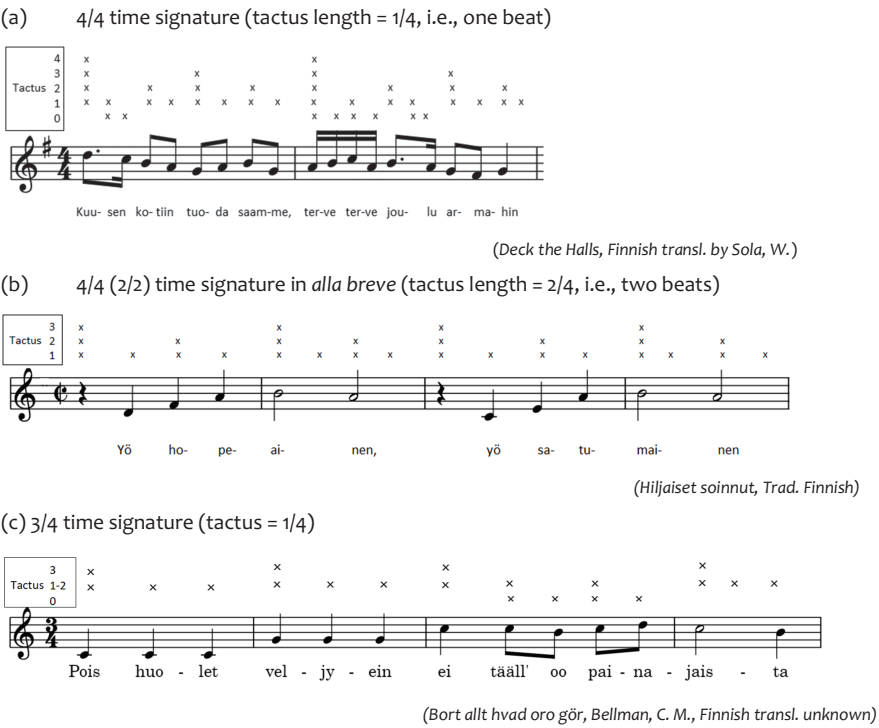


Figure 3: Levels of musical prominence in a song line. (a) 4/4 time signature (tactus length = 1/4, i.e., one beat), (b) 4/4 (2/2) time signature in *alla breve* (tactus length = 2/4, i.e., two beats), (c) 3/4 time signature (tactus = 1/4).

The first general challenge of rhythm alignment comes from the requirement to take *both* the prominence and the length of notes into account in songs. In music, both musical prominence and length are regulated at the same time. In contrast, the rules of poetic meter tend to regulate only one rhythmic aspect (e.g., either stress or length, depending on the language or poetic tradition). As writers and composers of vocal music are well known to aim for both accentual alignment and, to some extent, length alignment (Hayes & Kaun, 1996, Rodriguez-Vazquez 2010a), the requirement of considering both makes textsetting particularly demanding. This area is still understudied.

What is more, poetic systems and similarly organized musical meters are not necessarily geared towards the language used in that system. For example, accentual stress-based poetry, illustrated by the Shakespeare line in Figure 2 above, is the dominant modern verse tradition of many European languages. Crucially, it is not only employed in accent-driven languages with positionally free and weight-

dependent stress (such as English, German, or Russian), but also in quantity-based languages with fixed and weight-independent stress such as Finnish. The classical Finnish poetry is accent-based despite the archaic Finnish Kalevala meter being primarily based on prosodic weight (Leino, 1982).

A further complication for the comparison of musico-linguistic durations is that, in the majority of languages studied, length is dependent on, and a phonetic correlate of, linguistic stress or accent. Western music is considerably more variable. Note length may be a cue to note prominence, and it is hence often difficult to separate musical length and prominence. However, length in music may also be entirely independent of musical prominence, as exemplified in Figure 4. There are no particular musical constraints against the last long note being completely non-prominent (indicated by a single x in the grid) – if the composer so wishes.

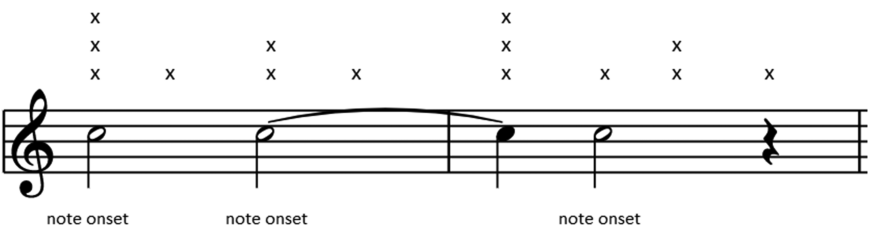


Figure 4: Musical durations unaligned with musical prominences.

Music and language also differ in their treatment of length distinctions and categories themselves. The typical prosodic weight classification in language trades in the unit of *mora*: every segmental position in the syllable rhyme (the nucleus vowel(s) and the coda consonant(s)) corresponds to one mora. In languages that allow for fairly complex syllable structure like Finnish or English, there is a minimally ternary distinction between *monomoraic short syllables*, *bimoraic heavy syllables*, and *trimoraic* (or longer) *superheavy syllables* (Table 1).

Quantitative meters of spoken verse (which are based on these or similar syllabic length contrasts) usually regulate phonological weight even more heavily, contrasting only metrically short and long syllables (Ryan, 2019:137 ff.). In contrast to this confinement of linguistic length to binary oppositions in verse, musical notes cover a much wider scale of durational variation, as illustrated by the note examples and discussion in Figure 3. The durational tools of vocal music are thus generally much more diverse than those of poetry. All this variation needs to be considered when studying the relation between musical note length and syllabic or prosodic weight in the textsetting of vocal music. In the following,

Table 1: Some basic syllable types in English and Finnish.

Prosodic weight	Examples	English	Finnish
Light/short (1μ)	(C)V	*to*	*jo*
Heavy (2μ)	(C)VC	*can*	*jos*
	(C)V:	*be*	*tee*
	(C)VV	*may*	*voi*
Superheavy (3μ)	(C)VCC	*cost*	*kilt.ti*
	(C)V:C	*seen*	*kuut.ti*
	(C)VVC	*fine*	*kuit.ti*

we consider specific environments in which the length distinction can be studied while controlling for musical prominence or linguistic stress/accent.

2.2 Durational regulation in asymmetric musical environments

How, then, should the durations between language and music be compared, if one aims to systematically calculate linguistic and musical length correspondences? A basic comparison can naturally be made between short and long notes, and the properties of syllables assigned to them. Our special focus in this study is on the long notes because we expect these to be the rhythmic environments which can more easily trigger linguistically non-neutral prosodic choices.

Most authors have a good intuitive conception of a long note that is lengthened beyond what can be considered neutral, but it is not always clear where the line between 'short' and 'long' notes should be drawn. Various definitions and approaches used for the definition of musical duration, such as by Hayes & Kaun (1996), or Temperley & Temperley (2011), differ in precision and method, but typically determine musical length in beats, the smallest prominent unit in music (see 2.1.).

For practical reasons stated in 2.1., we define note length based on the tactus. Long notes, in our conception, are those that exceed half a tactus in tacti with an even number of beats, or one third of a tactus in ternary musical meters. All other notes are considered durationally neutral or 'short'.

Still, defining note length as a simple function of the absolute number of beats per tactus disregards that length, by its nature, is a relative concept. A long

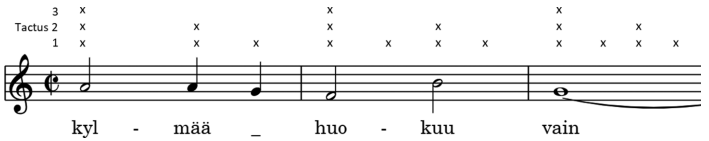


Figure 5: Puusydän/Muss i denn, *Trad. Ger*.

note is not felt as being particularly long when it is surrounded by other long notes (as on the words *huokuu vain*, in the second and third bars of Figure 5).

Conversely, note length is particularly salient when a long note is surrounded by shorter notes. Such a context is found in the asymmetrically distributed rhythms, illustrated by the first note of the first bar of Figure 5 above, or our first example, the Beethoven opening in Figure 1, revisited below in (?) Figure 6. In these examples the long note is followed by a shorter note within the same metrical unit and it is therefore perceived as very long. This lengthening does not depend on the note's absolute length. That is why it is important to look at the length relations within the whole group or pair of notes.

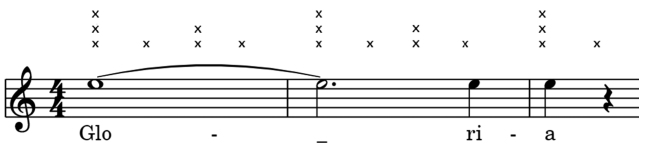


Figure 6: [1] *L. v. Beethoven*: Gloria, Mass in C.

We hypothesize that, in the context of textsetting, the linguistic relevance of the asymmetric durational ratios manifests itself particularly clearly in beat- and tactus-internal asymmetries, as beats and tacti are the most closely comparable musical units to the metrical feet or larger units such as hemistitches (half-lines) in poetic verse.

We call any two notes in asymmetrically distributed tacti *Asymmetric Durational Pairs* (ADP), modifying and expanding on the term used for the traditional note pair with a 3:1 length ratio, the *Regular dotted pair* (RDP) by Temperley & Temperley (2011). The widely used term "dotted" comes from musical notation but overlooks the very common similar asymmetries in ternary meters, which is why we refrain from its use in this more general concept of asymmetry. The ADP, and its symmetric counterpart, the *Symmetric Durational Pair* (SDP) can be defined as follows:

– SDP (Symmetric Durational Pair) = Pair of notes within a beat or tactus, starting with a musical beat with equal or greater metrical prominence than the following note, and with durations distributed in a 1:1 ratio, as in Figure 7.

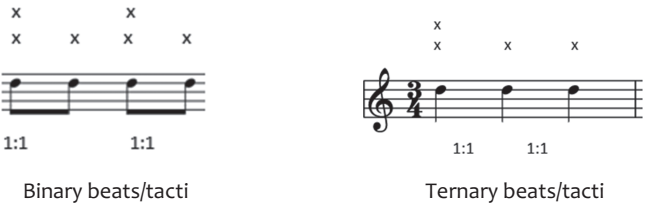


Figure 7: Symmetric Durational Pairs (SDP).

– ADP (Asymmetric Durational Pair) = Pair of notes within a beat or tactus, starting with a prominent beat and with durations distributed in an asymmetric 3:1 or 2:1 ratio, as in Figure 8, depending on the binarity (Figure 8a) or ternarity (Figure 8b) of the musical meter.

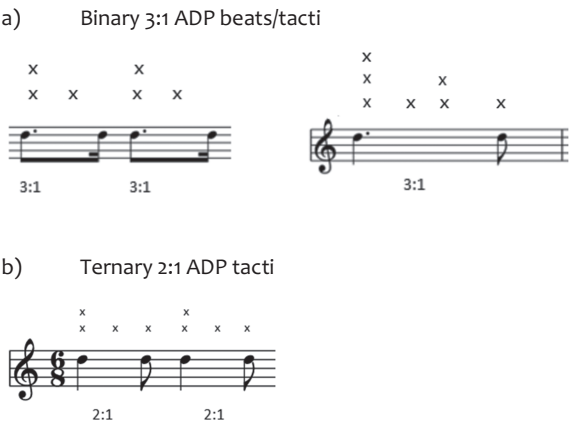


Figure 8: Asymmetric Durational Pairs (ADP).

A long first (i.e., prominent) note in an ADP is not only long but lengthened, as it were. The perceptual duration of ADPs is particularly strictly connected to the following note, as both notes participate in the realization of the asymmetric durational distribution within the same metrical unit. This interdependence and extra lengthening make the long note of the ADP different from other long notes, which, as pointed out by Hayes & Kaun (1996:245), tend to serve for articulation and phrasing rather than rhythmic purposes. In this study, we therefore examined

how syllable quantity in terms of prosodic weight relates to the durationally regulated musical environment by comparing the prosodic weight and structure of syllables on the long notes of Asymmetric Durational Pairs to corresponding long notes of Symmetric Durational Pairs.

To conclude, the main hypotheses of this study are:

- In Finnish textsetting, musical length is correlated with prosodic weight (a feature that is functionally loaded in the phonological system of the language): long notes are preferably assigned to heavy syllables (and vice versa), and short notes to light syllables.
- The correlation between prosodic weight and note length is stronger for long notes in asymmetric contexts, reflecting their especially lengthened nature.

Apart from examining these hypotheses concerning the relation of prosodic weight and note length, we explore whether the durational asymmetry engenders segmental length-related effects over and above prosodic weight, such as the nature of the nucleus (short or long vowel, diphthong), the sonority profile of the syllable, and gemination of coda consonants.

3 Methods of sampling and corpus construction

The songs used in this study represent original and translated vocal music in the Finnish language, and we focused on music that can be considered mainstream, prototypical and generally accepted as metrically and rhythmically neutral by Finnish listeners. For these purposes, we used as the basis of our song corpus the iconic 23-volume Finnish song collection *Suuri toivelaulukirja* ('The Great Song Book of Popular Tunes'). The sample consisted of 27 songs mainly from the volumes 2, 3, 4, 6, and 7; the beginning in the series contains many of the most well-known songs in Finnish, and presumably represent prosodic choices that are generally appreciated. In addition, a few similar songs were randomly selected from an earlier pilot corpus.

It should be noted that our sample does not make a strict distinction between the three main variants of the text-to-tune relationship in the musical creation process. First, in many cases, the lyric writer worked on an existing melody. This includes the pieces where a foreign song was translated or otherwise provided with Finnish lyrics. There is also at least one clear reverse case in the sample of a composer writing music to an existing poem. Finally, it would also be possible that the music and text were created simultaneously. In such cases, such as in traditional Finnish songs, it is impossible to reconstruct the creation process.

Although the direction of the process is a potentially relevant factor in the analysis, a systematic comparison was beyond the scope of this study. Different scholars have approached this question with varying sampling methods (cf. e.g., Temperley & Temperley, 2013, and Girardi & Plag, 2022), and different theoretical perspectives on whether song lyrics are inherently set to linguistic or musical metrical templates (see Dell, 2015:183–185). It is therefore not always easy to compare the results of such studies.

The songs were selected according to some parameters that aimed for an adequate stylistic and rhythmic representation of Finnish popular vocal music: they include variation in genres (e.g., popular songs, folk music, hymns, film music, and Christmas songs), source languages, Finnish lyric writers, and the rhythmic organization of songs, including different time signatures and the presence or absence of ADPs. The data was collected in a spreadsheet database with each syllable annotated for musical and linguistic features related to rhythm (e.g., musical prominence on a scale from 0–2, 2=tactus, and note lengths in ratios of beats and tacti). Of every song, the first 70 to 100 syllables were included (approximately following strophic boundaries), refrains and other longer repetitions were excluded, and ADPs and SDPs were coded in the corpus. The key features of the corpus are summarized in Table 2.

To examine the effect of asymmetric length relations while controlling for effects of linguistic stress and musical prominence, we compared syllable features on the long first[4] (metrically prominent) notes of ADPs with the first (metrically prominent) notes of SDPs, distinguishing between long SDPs and durationally neutral, or 'short' SDPs. Melismas, the distribution of a syllable on several notes, were extremely rare in the data, and if included, were counted for the total duration of the notes.

Another reason for the focus being on the first notes of the SDPs was the need to exclude phrase-final long notes from the comparison. In both language and music, there is a well-known overarching tendency of constituent-final elements to be treated differently from other positions in a phrase or a meter. (e.g., Hayes & Kaun, 1996:259, Ryan, 2019:139). Phrase-final notes are also more prone to repetition of previous syllabic features because of rhyming.

Linguistic parameters compared included syllable lengths in moras, as well as their segmental rhyme types (e.g., V:, VV, and VC), which will be discussed in more detail in the following analysis section. In addition to phonological

4 It should be noted that ADPs of the opposite relationship with short-long ratios of 1:3 and 1:2 (the so-called 'Scotch Snap' pattern, see Temperley & Temperley 2011) were marginal in the vocal music of the corpus and are ignored in this study.

Table 2: Details of the song corpus.

Parameter	Examples in corpus
Time signatures	2/4, 4/4, 4/4 (*alla breve*), 3/4, 6/8
Original languages	Finnish, English, French, German, Italian, Russian, Swedish
Genres	hymns, Christmas songs, film music, trad., var. popular tunes
Songs	27
Writers	25
Prominent notes	914
ADPs	347
SDPs (long)	304
SDPs (short)	263

weight classification, we also took into account the so-called boundary lengthening, a morphophonetic rule in Finnish (Karlsson, 1983:349, Suomi et al., 2008:44–46), which is triggered by some final vowels and morphemes and adds a pre-boundary lengthening to the following onset (e.g., /*tule tänne*/ [tulet:änne]; /*anna se*/ [annas:e]). In the song data, light syllables that triggered the boundary lengthening were thus analysed as bimoraic, and the syllable boundary as including a geminate. The inclusion of the phonetic lengthening is justified as it is perceptually prominent in Finnish and plays a key role in the flow of singing, in which notes are usually bound together without breaks by continuously resonant phonation, or in musical terms, *legato*.

4 Results and discussion

4.1 Results

4.1.1 Syllable weight and the assignment of syllables to prominent first members of note pairs

In a first step, we compare the assignment of prosodic weights to the prominent first notes of all the SDPs (long and short) and ADPs (inherently long) identified in the corpus. These notes are linguistically comparable owing to their typical alignment with word-initial and hence stressed syllables.

The plot in Figure 9 shows the distribution of prosodic weights of the syllables on the prominent note categories (long ADP notes and long as well as short SDPs; only notes that are not phrase-final were considered although including them did not affect the significance and interpretation of the results). The left panel shows the raw numbers from the corpus (ADP-L, n=347; SDP-L, n=304; SDP-S, n=263, total n=914). Conspicuously, the percentage of monomoraic or light syllables increases as the relative note length decreases, i.e., from ADP-L (18%) through SDP-L (32%) to SDP-S notes (50%). Conversely, the percentage of both heavy and super-heavy syllables decreases along the same scale (for bimoraic syllables: 41% on ADP-L, 33% on SDP-L, and 26% on SDP-S; for super-heavy syllables: 39% on ADP-L, 36% on SDP-L, and 26% on SDP-S). Correspondingly, a Chi square test leads us to reject the assumption of independence of note category and prosodic weight (Chi squared=34.11, df=4, $p<0.001$), conforming instead that note category and prosodic weight are significantly correlated. The second panel of Figure 9 shows the standardized residuals from this Chi square distribution. Positive residuals indicate overrepresentation, i.e., more observed cases than expected under the null hypothesis (which would predict a uniform distribution), and negative residuals indicate underrepresentation.[5] The residuals show that monomoraic or light syllables are significantly underrepresented on long ADP notes, and, conversely, significantly overrepresented on short SDP notes.

The comparison underlying this comprehensive Chi square test, however, confounds two factors of interest, namely note length and symmetric vs asymmetric context. To disentangle these factors, we considered two subset comparisons. First, we isolated the note length contrast by comparing the distribution of moras to only prominent long SDP notes vs prominent short SDP notes (ADP-L notes discarded). The Chi square test again confirms a deviance from a homogeneous distribution (Chi squared=9.39, df=2, p=0.009). The third panel in Figure 9 shows the corresponding Chi square residuals, attesting to underrepresentation of monomoraic syllables on long SDP notes and, conversely, overrepresentation of these syllables on short SDP notes.

Secondly, we pitted long ADP notes against long SDP notes (Chi squared=8.08, df=2, p=0.0176), ignoring the short SDP-S notes. The significant deviance from a homogeneous distribution suggests that the assignment of prosodic weight differs between asymmetrically long and symmetrically long notes. The corresponding

5 The standardized residuals can be interpreted like standard scores: residuals exceeding a value of |2| can be considered significant contributors to the deviance from the even distribution that would be expected under the null hypothesis.

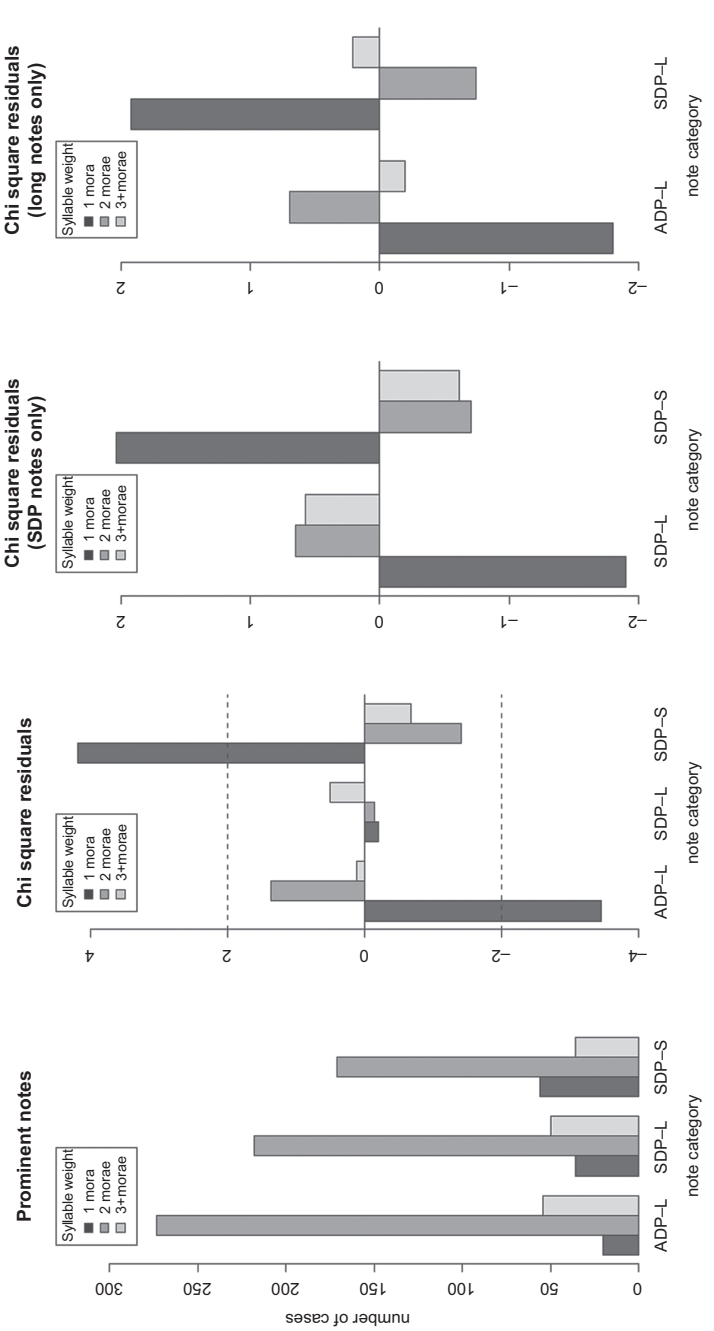


Figure 9: Left panel: Distribution of syllable types to three categories of prominent notes (ADP-L, SDP-L, SPD-S, see main text for details). Second panel: Chi square residuals showing underrepresentation and overrepresentation of syllable types based on the distribution in the first panel. Third panel: Chi square residuals on a subset regarding the length contrast (long vs short notes, only symmetric durational pairs). Fourth panel: Chi square residuals on a subset regarding the symmetry contrast (asymmetric vs symmetric durational pairs; long notes only).

residuals are plotted in the fourth panel of Figure 9. In this comparison, monomoraic syllables are overrepresented on long SDP notes when compared to long ADP notes. That is, irrespective of length, the symmetry distinction has a role to play in the assignment of syllable weights to note types.

In sum, this analysis suggests that both note length and the symmetry of the relationship to the following note in a note pair contribute to the assignment of the different syllable weights to the note categories.

4.1.2 Segmental makeup of the syllables

The distribution of the heavy syllables (grey and light grey bars in Figure 9) does not immediately suggest a significant difference regarding their assignment to the different note categories. However, there are various types of heavy syllables with differences regarding the distribution of consonantal and vocalic segments and their respective lengthening. We therefore explore whether the distribution of segmental content is contingent on the note type the syllable is set to in music (long vs short note, asymmetric long vs symmetric long). We were especially interested in whether segmental length (short vs long nuclei vs diphthongs; geminate vs non-geminate codas), is related to length and symmetry/asymmetry in music. We also explored whether sonorant versus obstruent codas differ regarding their assignment to the different note categories.

Figure 10 juxtaposes nine different types of syllable rhyme with short, long, and diphthongized nuclei and different types of codas (no coda, non-geminate coda consonants, geminate coda consonants), broken down by affiliation to asymmetrically long (ADP-L, black), symmetrically long (SDP-L, dark grey) and symmetrically short (SDP-S, light grey) prominent notes.

First, the numbers in Figure 10 reproduce the results of the first part of this study, that is the particular avoidance of monomoraic syllables on long ADP notes and the overrepresentation of the monomoraic type on the short SDP notes (left triplet of bars). Conversely, the numbers suggest that long ADP notes are preferably filled with simple -VC syllable rhymes featuring a non-geminate coda consonant after short vowels (second triplet of bars). Short SDP notes are underrepresented on all syllable types with bimoraic vowels (4th to 9th triplet of bars), irrespective of coda type. The two types of long notes (ADP-L vs SDP-L) do not appear to differ greatly in their assignment to syllables with bimoraic nuclei, apart from one observation regarding the geminate codas: Syllables with geminates after long vowels are rare, and they seem to be especially avoided on long SDP notes 6th pair of bars). Syllables with geminates after diphthongs, on the other hand, are slightly preferred on long SDP notes (last or

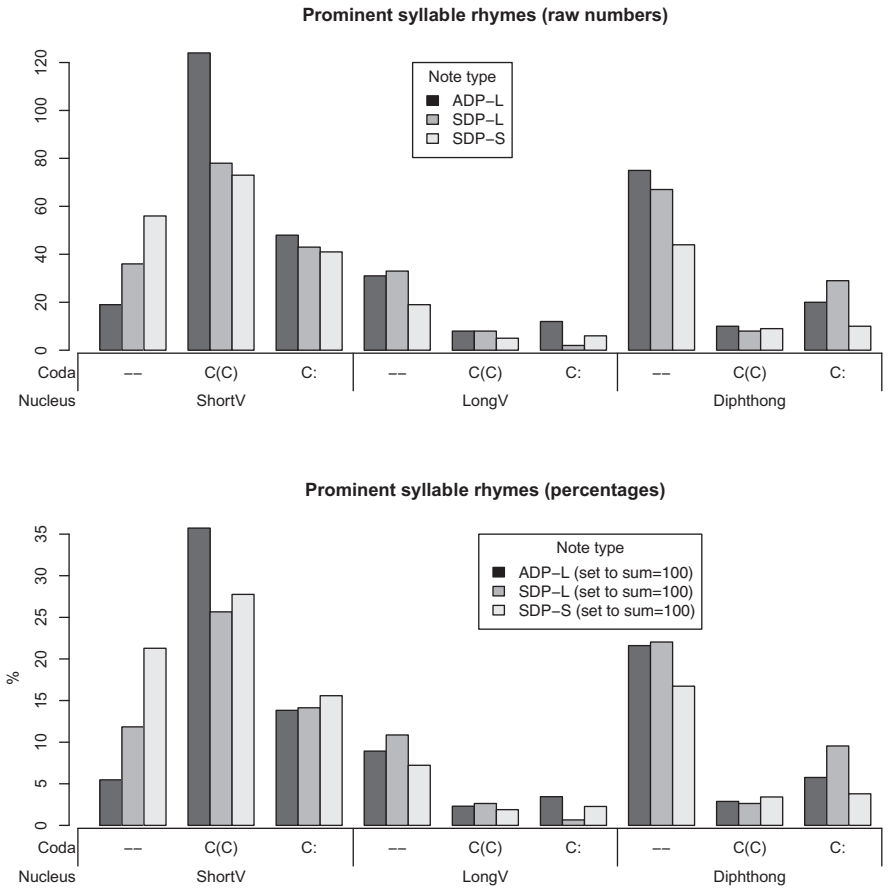


Figure 10: Distribution of nine different syllable types (broken down by type of nucleus und type of coda) over the prominent notes of asymmetric (ADP, black bars) and symmetric durational pairs (SDP-L and SDP-S, dark and light grey bars, respectively). The upper panel depicts the raw number of syllables, with percentages per vowel category for each note type shown in the bottom panel.

9th pair of bars). This suggests that the type of nucleus (short, long, diphthong) and the type of coda status (no coda, geminate or non-geminate) interact regarding the assignment of the syllables to the prominent notes of symmetric or asymmetric durational pairs.

A second explorative analysis concerns the distinction between sonorant versus obstruent coda consonants. Comparing the raw number of sonorant and obstruent coda consonants shows that sonorant codas are generally more frequent (298 codas with sonorant consonants vs. 236 obstruent codas). At first sight,

there does not appear to be a significant difference regarding the distribution of sonorant vs obstruent codas among the three prominent note categories (ADP-L: 101 obstruents, 121 sonorants; SDP-L: 71 obstruents, 97 sonorants; SDP-S: 64 obstruents, 80 sonorants; Chi-squared=0.41, df=2, p-value=0.815). However, closer inspection reveals that, in the subset of songs that feature both short and long SDP notes (13 of 27 songs), the ratio of sonorant codas is in most songs higher on long notes than on short notes (see Figure 11). In fact, in only two of the 13 songs (*Kertokaa se hänelle* and *Pohjolan maa*), the ratio of sonorant codas on short notes is higher than on long notes, and this reverse tendency has relatively little weight as it is due to only very few observations in these songs (three syllables).

A comparable analysis for the symmetry contrast does not reveal a systematic relationship between symmetric versus asymmetric long notes on the one hand, and the ratio of sonorant versus obstruent codas on the other (see Figure 1 in the appendix).

To ascertain the general validity of these observations, we used R statistical software (R Core Team, 2020) to fit two hierarchical generalized linear models (Bates et al., 2014) to the data set.

To operationalize the effects of syllabic structure on syllable-to-note association, various syllabic features were determined as independent variables. These features were coded as binary orthogonal sum contrasts as follows:

1. Nucleus weight contrast: short monomoraic nuclei (–1) vs bimoraic nuclei (+1)
2. Bimoraic vowel type contrast: Diphthongs (–1) vs long vowels (+1); short vowels (irrelevant for this contrast) were set to 0
3. Coda contrast: open syllables (–1) vs closed syllables (+1)
4. Gemination: Non-Geminate (–1) vs Geminate (+1) coda; open syllables set to 0
5. Coda sonority: Obstruent (–1) vs Sonorant (+1) coda; open syllables set to 0

Along with these main effects, the full model included the set of all possible two-way interactions, namely

6. Nucleus weight : Coda contrast
7. Nucleus weight : Geminate
8. Nucleus weight : Coda sonority
9. Bimoraic vowel type : Coda contrast
10. Bimoraic vowel type : Geminate
11. Bimoraic vowel type : Coda sonority
12. Geminate : Coda sonority

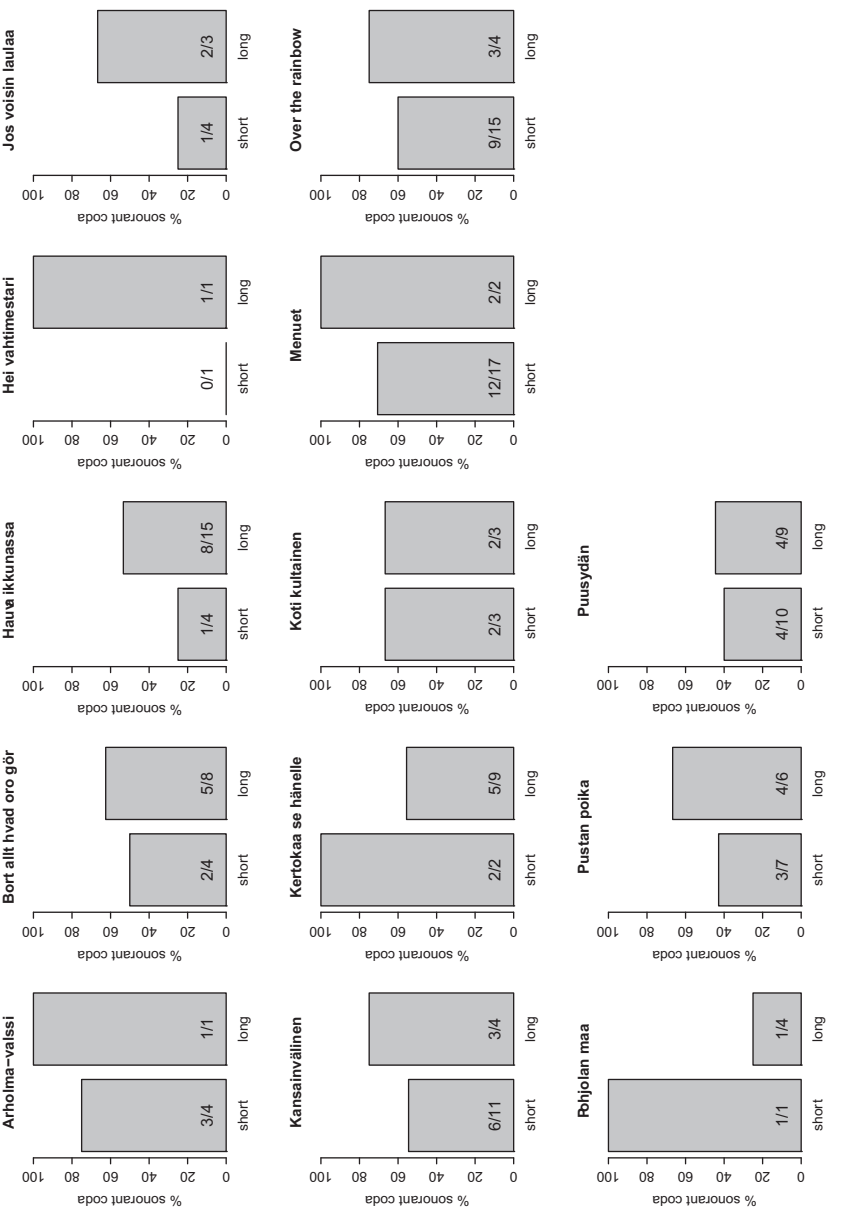


Figure 11: Percentage of sonorant coda consonants by note length (left bar: short SDP notes, right bar: long SDP notes) in syllables featuring coda consonants. Values are broken down by songs (with the ratio sonorant / all coda consonants superimposed on each bar). Only those songs that exhibit the length contrast, i.e., feature both long and short SDPs, are considered (ADP-L notes disregarded).

We included Song (n=27) as the grouping variable (random intercept only, as models with slopes did not converge, probably due to a lack of statistical power).

Given that note length is confounded with note symmetry in our data set, we fit separate models on two different subsets of the data. The first subset model considers the binary distinction between long symmetric (SDP-L) vs short symmetric notes (SDP-S) as binary dependent variable. That is, it tests whether note length per se is correlated with the syllabic features of interest. For the purpose of this model, long ADP notes were discarded. For model selection, we reduced the full model (which includes all possible two-way interaction terms), successively discarding the non-significant interaction terms with the smallest effect size unless and until further model reduction leads to deterioration of model fit, as determined by a log-likelihood test. We take the model determined this way to be the model with the best fit and report it below in Table 3.

Table 3: Summary of best fitting generalized linear mixed effects model with note length as binomial dependent variable (data from SDP notes only, ADP notes discarded).

	Estimate	Std. Error	z value	p value
Nucleus Weight	1.18	0.32	3.706	0.0002*
Coda	0.16	0.29	0.54	0.59
Coda Sonority	0.81	0.37	2.161	0.03*
Gemin-NonGemin	0.25	0.38	0.67	0.505
Diphthong-LongV	−0.005	0.41	−0.013	0.989

This model returns two significant main effects, first an effect of nucleus weight, and second, an effect of coda sonority. The coefficient estimate for the effect of nucleus weight (1.18 log odds) translates to roughly 3.2 in odds ratios and suggests that a syllable with a bimoraic vowel is about 3 times as likely to be set to a long as opposed to a short note. The effect of coda sonority (0.81 log odds, 2.24 odds ratio) suggests that it is roughly twice as likely for a syllable featuring a coda sonorant to be set to a long note compared to a syllable with a coda obstruent (cf. Figure 11). Interestingly, the effect of coda presence alone is very small and non-significant (0.16 log odds or 1.17 odds ratio). That is, the effect of syllable weight on the assignment to long vs short notes appears to be mainly an effect of the nucleus, not of the coda.

The second model tests whether the same syllabic features have an effect on the assignment of syllables to symmetrically versus asymmetrically long

Table 4: Summary of best fitting generalized linear mixed effects model with note symmetry as binomial dependent variable (data from long notes only, short SDP notes discarded).

	Estimate	Std. Error	z value	p value
Nucleus Weight	0.08	0.27	0.319	0.75
Coda	0.34	0.25	1.331	0.183
Coda Sonority	−0.16	0.26	−0.625	0.532
Gemin-NonGemin	−0.09	0.27	−0.343	0.7313
Diphthong-LongV	0.206	0.31	0.663	0.507
Gemin: DiphLongV	2.71	1.1	2.462	0.014*

notes. Consequently, for this subset, short SDP notes were discarded. Again, we report the model with the best fit, according to the above selection procedure. The output of this second model is shown in Table 4.

The second subset model yields a significant interaction between coda gemination (geminate vs non-geminate coda) and the type of bimoraic vowel (long vowel vs diphthong). This interaction reflects preferences regarding the assignment of super-heavy syllables (3+ moras, which make up roughly 15% of the data) to the prominent notes of symmetric/asymmetric durational pairs (long vowel+geminate avoided in SDP-L notes as compared to ADP-L notes, while diphthong+geminate syllables are preferably set to SDP-L, see Figure 10). This interaction suggests that over and above prosodic weight, the segmental filling of the moras has an influence on the assignment of these syllables to symmetric or asymmetric durational pairs.

4.2 Discussion

4.2.1 Prosodic weight and stress-related lengthening

This study has produced two kinds of results: those connected with the prosodic weight and those linked to the segmental makeup of the syllable rhyme. The interpretation of the former results is relatively straightforward: The statistical analysis shows that monomoraic or light syllables are preferably set to short notes and, concomitantly, particularly avoided on all prominent notes that exceed the neutral half-tactus length. A similar avoidance of light syllables on metrically strong positions has been reported by Leino for classical Finnish

poems (Leino, 1982). However, while a difference between light and heavy syllables regarding note assignment is apparent in the present data, we did not find a clear difference between the assignment of heavy and super-heavy syllables to note categories (in contrast to Ryan [2019] who reports such a difference for certain strong positions in the archaic chanted Finnish Kalevala verse). Nevertheless, the present results confirm that in Finnish songs, syllable weight or quantity is a salient prosodic feature, as in Finnish speech prosody.

Beyond the contrast between musically long and short notes, the proportions of light syllables also show interesting differences between the two types of long notes (ADP-L vs SDP-L). Light syllables appear to be especially avoided on the long notes of the ADPs. Since the note values of long SDPs are often even longer than the long ADPs, which themselves are rhythmically very heterogenous, this suggests that it is not only the absolute length of a given note value but the length relationship between the adjacent notes within a prominence unit which guides the prosodic choices of the songwriters.

It has been noted that alignment in textsetting in general is mainly driven by phonological constraints (deCastro Arrazola, 2015:172), and the present results demonstrate that the inherent prosodic weight of syllables is a decisive factor in Finnish textsetting as well, as evidenced by the avoidance of light syllables on long ADP notes. Over and above the low prosodic weight, we propose that the particular avoidance of monomoraic syllables on long ADP notes may also be, in part, related to the *phonetic* consequences of the mora-based nature of word-initial stress in Finnish. Recall that stress in Finnish affects the first two moras of a word regardless of syllabic association of the moras. Therefore, if the first syllable of a word is monomoraic, stress-related lengthening extends into the second syllable (Suomi & Ylitalo, 2004). This linguistic configuration, however, is particularly misaligned in the musical context of an ADP, which emphasizes a contrast between a lengthened first note and a short second note. Instead, the musical asymmetry is rendered naturally when both of the two lengthened moras are assigned to the first syllable, making it a heavy one and, at the same time, keeping any effect of stress-related lengthening away from the short second note of this pair.

Conversely, light initial syllables appear to be tolerated on surprisingly long tacti in Symmetrical Durational Pairs. With a light first syllable, the concomitant stress-related lengthening of the following second syllable does not pose a problem in this musical context, as the second note is as long as the first: the second long note is therefore an appropriate exponent of the stress-related lengthening.

4.2.2 Analysis of the segmental syllabic features

The results concerning the segmental makeup of the syllables in our data are less clear-cut and more complicated to interpret than the results concerning prosodic weight. The segmental makeup, however, is equally relevant for the consideration of the special rhythmic and phonetic characteristics of vocal music. As discussed in the introductory sections, the lengthening potential of notes exceeds linguistic weight limits, and moreover, the continuously resonant method of *legato* singing affects the transitions between sound segments in a way that is different from speech.

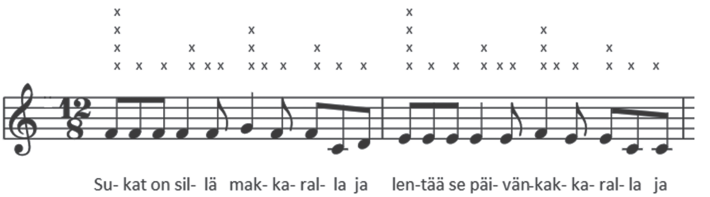
Our results showed that in the basic length comparison between long and short SDP notes, it is particularly the weight of the nucleus that is decisive. This result seems natural, considering that singing long notes is aided by as resonant articulation as possible and "we sing with the vowels", as succinctly formulated by Ophaug (2015:295). We also note an effect of coda sonority: syllables with little sonority (obstruent coda) are preferably set to short SDP notes and, conversely long SDP notes attract sonorant codas (cf. Figure 11); this is probably motivated by singability as well.

As for particular segmental effects concerning the comparison between symmetric and asymmetric contexts, the picture is less clear. It seems that in ADPs, not only nucleus weight but also coda consonants may interactively contribute to the particular rhythmic effect. Among the superheavy syllables, rhymes with long vowels and a geminate coda are preferred on the particularly long and lengthened ADP notes. This might be another indication of an alignment of linguistic lengthening (concerning both nucleus and coda) with musical length.

On the other hand, as shown in Figure 10, ADP notes generally favor the simpler -VC syllables rhymes, i.e., those with a short nucleus followed by a non-geminate coda. At first sight, the potential preference of the -VC rhyme on the lengthened ADP notes seems counterintuitive. This heavy syllable is minimal on three prosodic levels: its total mora count, the separate mora counts of its nucleus and coda, and the lack of lengthening on each of its segments (i.e., lacking long or diphthongized vowel segments or a geminate). However, the aforementioned features also make this syllable type particularly clearly defined, both in terms of its acoustics and in terms of articulation. Such syllables might therefore more properly stand out as prominent syllables on the particularly prominent ADP notes.

However, there appear to be song-specific effects that blur the picture: Consider the excessive use of geminate stops after short vowels in the Finnish children's song *Magdaleena* in Figure 12. This song systematically employs ternary

ADPs combined with a salient high peak note to enforce the rhythmic salience of the geminate stops of the original poem, as illustrated by the words *makkaralla* and *päivänkakkaralla*.[6] These kinds of individual effects are evident on close inspection but are less likely to emerge as clear statistical results in the data of the present kind.



(Magdaleena, Chydenius, K., composed to a poem by Kaarina Helakisa)

Figure 12: Geminate sonorants and stops on ADPs.

Thus, although more complex nucleus-coda combinations clearly also interact with, and emphasize, musical lengthening – as in the case of *Magdaleena* – it is nevertheless possible that the most neutral way of filling an asymmetric ADP rhythm in songs could favor a simple heavy -VC syllable described above, with clear segmental transits and an unambiguous syllable boundary.

All in all, our results do not give straightforward evidence on the behavior of any particular segmental type, but do indicate that different segmental combinations play a role in the realization of length relationships in singing. The clearest sign of length alignment in Finnish textsetting based on this study is the prosodic weight of the syllables.

5 Conclusions

The analysis of the present corpus of popular vocal music of Finnish suggests that Finnish song writers are sensitive to, and consider prosodic weight and certain segmental features when setting text to musically prominent notes of varying lengths and with different length relationships to adjacent following

6 Note that segmental processes such as consonant lengthening connected with metrical quantitative prominence also feature in spoken verse traditions, such as Ancient Greek, where light metrically accented syllables are often connected to the following onset by gemination (Revithiadou, 2004:40).

notes. The results clearly show that song writers strive for the alignment of musical note length with prosodic weight and the overall sonority of the syllables. Effects of other segmental syllabic features on note alignment in textsetting are detectable but less clear regarding their interpretation.

The present study on Finnish demonstrates that durationally regulated musical contexts and their alignment with prosodic weight can be studied independently of linguistic stress or accent. Inspecting the interaction between musical length and prosodic weight is especially fruitful in the asymmetric durational contexts (ADPs), in which length relations between the notes concerned are more strictly regulated and which clearly behave as a particular rhythmic unit in vocal music, possibly differing from the symmetric environments also in their make-up of segmental syllabic features.

In sum, our approach extends the perspective on language-music relationships in textsetting beyond the more widely studied stress-prominence connection by focusing on the role of the prosodic weight of syllables and their alignment with musical note length in Finnish. We emphasize that the specifics of the phonological and phonetic systems of the song language and the musical context of the notes under consideration are crucial for our understanding of how phonology is put to use in the expression of rhythm in vocal music.

Appendix

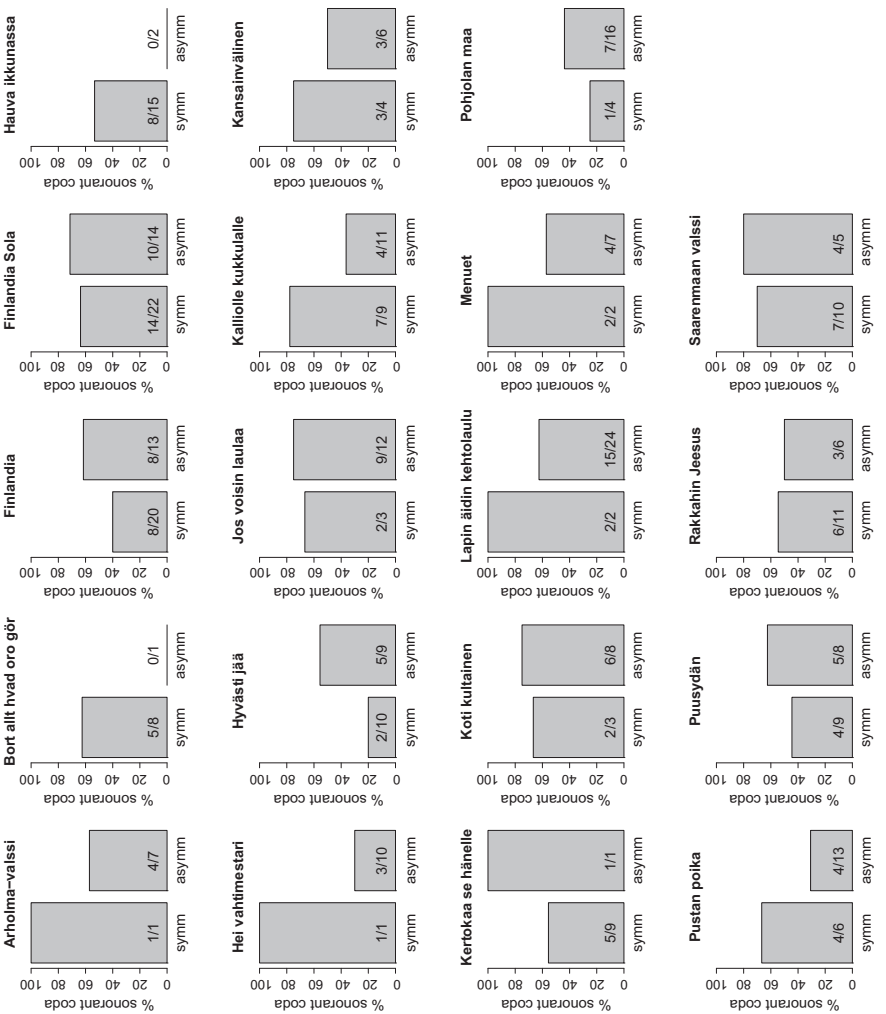


Figure 1: Percentage of sonorant coda consonants by note symmetry (left bar: symmetric SDP-L notes, right bar: asymmetric ADP-L notes) for syllables featuring coda consonants. Values are broken down by songs (with ratio sonorant / all coda consonants superimposed on each bar). Only those songs that exhibit the symmetry contrast, i.e., feature both symmetric and asymmetric long notes, are considered (SDP-S notes disregarded).

References

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2014. Fitting Linear Mixed-Effects models using lme4. *arXiv preprint arXiv:1406.5823*.

deCastro Arrazola, Varun. 2015. The prosody of Basque songs. A methodological proposal. In: Teresa Proto, Paolo Canettieri, Gianluca Valenti, editors. *Text and tune: On the association of lyrics and tune in sung verse*, 171–182. Bern: Peter Lang.

deCastro Arrazola, Varun. 2018. Typological tendencies in verse and their cognitive grounding [Doctoral Dissertation]: Leiden University.

Dell, François. 2015. Text-to-tune alignment and lineation in traditional French songs. In: Teresa Proto, Paolo Canettieri, Gianluca Valenti, editors. *Text and tune: On the association of lyrics and tune in sung verse*, 183–234. Bern: Peter Lang.

Dell, François & John Halle. 2009. Comparing musical textsetting in French and in English songs. In: Jean-Louis Aroui, Andy Arleo, editors. *Towards a typology of poetic forms: From language to metrics and beyond*, 63–78. Amsterdam: John Benjamins Publishing Company.

Domene Moreno, Christina & Bariş Kabak. 2022. Prominence alignment in English and Turkish songs: Implications for word prosodic typology. In: Mathias Scharinger, Richard Wiese, editors. *How language speaks to music: Prosody from a cross-domain perspective*, 223–258. Berlin: De Gruyter.

Girardi, Elena & Ingo Plag. 2022. Metrical mapping in text-setting: Empirical analysis and grammatical implementation. In: Mathias Scharinger, Richard Wiese, editors. *How language speaks to music: Prosody from a cross-domain perspective*, 191–221. Berlin: De Gruyter.

Goedemans, Rob & Harry van der Hulst. 2013. Weight-sensitive stress. In: Matthew S. Dryer, Martin Haspelmath, editors. *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.

Hayes, Bruce & Abigail Kaun. 1996. The role of phonological phrasing in sung and chanted verse. *The Linguistic Review* 13(3–4). 243–303.

Karlsson, Fred. 1983. *Suomen kielen äänne-ja muotorakenne [The phonological and morphological structure of Finnish]*. Juva: WSOY.

Kiparsky, Paul. 2006. A modular metrics for folk verse. In: B. Elan Dresher, Nila Friedberg, editors. *Formal approaches to poetry*, 7–49. Berlin: De Gruyter.

Lehiste, Ilse. 1970. *Suprasegmentals*. Cambridge, MA: The MIT Press.

Leino, Pentti. 1982. *Kieli, runo ja mitta: suomen kielen metriikka [Language, poem, and meter: the metrics of the Finnish language]*. Pieksämäki: Suomalaisen kirjallisuuden seura.

Lerdahl, Fred & Ray Jackendoff. 1983. *A generative theory of tonal music*. Cambridge, MA: The MIT Press.

Liberman, Mark. 1975. *The intonational system of English*. New York: Garland Publishing Inc.

McPherson, Laura. 2019. Musical adaptation as phonological evidence: Case studies from textsetting, rhyme, and musical surrogates. *Language and Linguistics Compass* 13(12).

Ophaug, Wencke. 2015. The challenge of identifying vowel phonemes in singing. In: Teresa Proto, Paolo Canettieri, Gianluca Valenti, editors. *Text and tune: On the association of lyrics and tune in sung verse*, 287–312. Bern: Peter Lang.

Palmer, Caroline & Michael H. Kelly. 1992. Linguistic prosody and musical meter in song. *Journal of Memory and Language* 31(4). 525–542.

Patel, Aniruddh D. 2003. Rhythm in language and music: Parallels and differences. *Annals of the New York Academy of Sciences* 999(1). 140–143.
Patel, Aniruddh D. 2010. *Music, language, and the brain*. Oxford: Oxford University Press.
Proto, Teresa. 2015. Prosody, melody and rhythm in vocal music: The problem of textsetting in a linguistic perspective. *Linguistics in the Netherlands* 32. 116–129.
Proto, Teresa & François Dell. 2013. The structure of metrical patterns in tunes and in literary verse. Evidence from discrepancies between musical and linguistic rhythm in Italian songs. *Probus* 25(1). 105–138.
Revithiadou, Anthi. 2004. The Iambic/Trochaic Law revisited: Lengthening and shortening in trochaic systems. *Leiden Papers in Linguistics* 1(1). 37–62.
Rodríguez-Vázquez, Rosalía. 2010a. *The rhythm of speech, verse and vocal music: A new theory*. Bern: Peter Lang.
Rodríguez-Vázquez, Rosalía. 2010b. Text-setting constraints: A comparative perspective. *Australian Journal of Linguistics* 30(1). 19–34.
Ryan, Kevin M. 2017. The stress–weight interface in metre. *Phonology* 34(3). 581–613.
Ryan, Kevin M. 2019. *Prosodic weight: Categories and continua*. Oxford: Oxford University Press.
Suomi, Kari. 2007. On the tonal and temporal domains of accent in Finnish. *Journal of Phonetics* 35(1). 40–55.
Suomi, Kari, Juhani Toivanen & Riikka Ylitalo. 2003. Durational and tonal correlates of accent in Finnish. *Journal of Phonetics* 31(1). 113–138.
Suomi, Kari, Juhani Toivanen & Riikka Ylitalo. 2008. *Finnish sound structure: Phonetics, phonology, phonotactics and prosody*. Oulu: University of Oulu.
Suomi, Kari & Riikka Ylitalo. 2004. On durational correlates of word stress in Finnish. *Journal of Phonetics* 32(1). 35–63.
Team, R Core. 2020. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.
Temperley, Nicholas & David Temperley. 2011. Music-language correlations and the "Scotch Snap". *Music Perception* 29(1). 51–63.
Temperley, Nicholas & David Temperley. 2013. Stress-meter alignment in French vocal music. *The Journal of the Acoustical Society of America* 134(1). 520–527.