Published in the Journal of New Music Research, Feb 2021

RESEARCH ARTICLE

A Study of Variability in Raga Motifs in Performance Contexts

Kaustuv Kanti Ganguli and Preeti Rao

Department of Electrical Engineering
Indian Institute of Technology Bombay, Mumbai 400076, India.

ARTICLE HISTORY
Compiled December 14, 2020

ABSTRACT
Corpus based studies of Indian art music are relatively recent. Encouraged by the availability of computational methods for audio MIR, studies of raga performance have been attempted on the available and growing corpora of concert recordings. A particularly interesting aspect of this genre is the prominent place of improvisation in any performance. The spontaneous creation of musical ideas has been of interest to musicologists across music genres, and Indian art music provides fertile ground for the investigation of this phenomenon in all its different dimensions. In this paper, we explore a specific aspect of performance related to the influence of structural constraints of the music on raga motifs in the course of improvisation. Audio recordings of North Indian vocal concerts are analysed to extract measurements of the defining parameters of the recurrent melodic phrases that characterize the raga in performance. While improvisation relates to choosing the sequence of phrases at the larger time scales, we show that the categorical nature of the musical melodies is preserved through the interaction of the phrases with the changing melodic and rhythmic contexts.

KEYWORDS
Indian art music, Raga phrase, Improvisation, Melodic-Rhythmic structure.

1. Introduction

Indian art music is based on the melodic framework of ragas. A raga specifies the tonal material, tonal hierarchies and a set of characteristic melodic motifs or phrases. While phrases are nominally labeled by the sequence of notes, it is the melodic shape including the continuous pitch movement between notes and the ornaments that fully characterises the raga motif. In a raga performance, the melody is loosely constrained by the pre-composed piece chosen for the occasion but otherwise improvised in keeping with the raga grammar. That is, the distinctiveness of the raga is maintained at all times while the performer enjoys considerable freedom to display creativity. This creation of music in the course of a performance is an example of the phenomenon of musical improvisation, where the improviser uses implicit knowledge of the musical structure familiar to listeners "to make their in-the-moment compositions coherent and stylistically appropriate" (Ashley, 2009). As such, Indian art music offers a rich domain for a study of musical improvisation in all the varied dimensions of interest including physiological, cognitive, cultural, and structural (Nettl & Russell, 1998). A

CONTACT Preeti Rao. Email: prao@ee.iitb.ac.in

study of performances by acclaimed musicians of the genre can serve to address interesting questions on improvisation such as, for instance, what does an individual performer do, how does this differ from concert to concert and how is the required musical structure maintained. In this work, we explore, through the empirical analyses of available concert recordings, how improvisation interacts with the structural constraints imposed by raga grammar and practice in the context of melodic phrase shapes in Hindustani (North Indian) vocal music.

There have been a limited number of empirical studies of Hindustani classical music performance. Due to the absence of notated music, and the fact that every performance of a raga evolves in a unique way, the few prominent available studies have involved the analytical listening of a recorded concert possibly in collaboration with the performer himself (Clayton, 2008; van der Meer, 1980; Widdess, 1994, 2011). The analyses derived from the collaborative transcription have yielded important insights about the nature of the improvisation. More recently, computational techniques originating in the field of music information retrieval (MIR) have been applied to Indian art music performance audio to visualize and subsequently model melodic and rhythmic aspects. These approaches typically employ signal processing and machine learning techniques and offer the prospect of large scale comparisons and generalizations. Implemented on corpora of commercial recordings, the propounded MIR tasks such as raga recognition, melodic motif detection, structural segmentation and detection of rhythmic landmarks, have served as the bases for the development of computational models for the genre-specific musical attributes (Chordia & Şentürk, 2013; Dutta, Krishnaraj, & Murthy, 2015; Ganguli et al., 2017; Ganguli & Rao, 2015; Ganguli, Rastogi, Pandit, Kantan, & Rao, 2015; Gulati, Serra, Ganguli, & Serra, 2014; Gulati, Serrà, Ishwar, Şentürk, & Serra, 2016; Gulati, Serrà, Ishwar, & Serra, 2014; Gulati, Serrà, & Serra, 2015; MA, TP, & Rao, 2020; P. Rao et al., 2014; P. Rao, Vinutha, & Rohit, 2020; S. Rao & Rao, 2014; Ross, Vinutha, & Rao, 2012; Srinivasamurthy, Holzapfel, Ganguli, & Serra, 2017).

More recently, Ganguli and Rao (2017, 2018) optimized the parameters of a distributional representation (capturing tonal hierarchy) for a concert based on discriminating closely related ragas. The computed distribution showed strong variations in the melodic content across concerts in the same raga by the same performer at the time scale of a rhythmic cycle. The distributions computed at larger time scales, on the other hand, are more similar in shape across concerts and characteristic of the raga itself. This confirms that the improvisation involves the choice of notes and phrases at specific instances while conforming overall to the tonal hierarchy mandated by the raga grammar (Deva, 1974; Nooshin & Widdess, 2006; van der Meer, 1980). In the current paper, we extend the work to examine more closely how structural constraints are met by the performer given this spontaneous decision to use a particular melodic phrase at a particular instance in time. We take concerts by well-known artists in chosen ragas and present the computational analyses of melodic motifs across occurrences in the different contexts within and across concerts. Apart from generating musicologically important insights, we expect the study of phrase shape variability to aid the computational modeling of melodic similarity in MIR tasks.

In the next section we provide an overview of the relevant music theory and structure of the Hindustani vocal music concert. The scope of the presented work is outlined next followed by a description of the dataset and audio processing methods developed for the melodic phrase analyses. Our observations of phrase variability and dependence on context are presented next. Finally, we discuss the findings in relation to available musicological studies.

2. Music Background and Scope

A Hindustani music concert is predominantly improvised and occurs within a chosen *raga* (melodic) and *tala* (rhythmic) framework. The rules of the raga grammar are manifested at different time scales, at different levels of abstraction, and demand different degrees of conformity. While some of the elements of raga grammar are explicit, others are implicit and can take years of musical training to master. A number of textbooks and scholarly studies exist that describe different improvisatory aspects of melody in Indian art music (Bagchee, 1998; Chakrabarty, 2002; Danielou, 2010; Ganguli, 2013; Jairazbhoy, 2011; S. Rao, Bor, van der Meer, & Harvey, 1999; van der Meer, 1980). A majority of these are musicological in nature and typically involve either a compilation of expert domain knowledge or the qualitative analysis of a handful of chosen musical excerpts. In this section, we provide a concise description of raga music, both in theory and in performance, followed by a discussion about the scope of the proposed research.

2.1. Grammar and Concert Structure

A raga is characterized by a set of melodic features that includes a set of notes (*svara*), the ascending and descending melodic progression (*arohana-avrohana*), and a set of characteristic phrases (*pakad*). van der Meer (1980) comments that technically a raga is a musical entity in which the intonation of svaras, as well as their relative duration and order, is defined. The svaras have different importances in a tonal hierarchy, and play different roles in the melodic elaboration via their durations and frequencies of occurrence (Bagchee, 1998; Danielou, 2010; Raja, 2016; S. Rao et al., 1999; van der Meer, 1980). The two predominant notes in decreasing order are the *vadi* and *samvadi*; these are best understood in terms of their roles in the improvisation with their higher frequencies of occurrence and duration. Further, a raga is characterised by its phrases, or melodic motifs, in which the svaras have their proper relative duration. This does not mean that the duration, the recurrence and the order of svaras are fixed in a raga; they are fixed only within a context (van der Meer, 1980). The melodic phrase shape, unique to a raga, is not necessarily represented in terms of a sequence of svaras of discrete-pitch and -time intervals. The totality of the musical characteristics are thus captured in the set of phrases (*'calana'*) or skeleton of the raga, with a phrase itself acting as a gestalt that is immediately identified with the raga by an informed listener.

While the description so far clearly leaves room for flexibility in the delivery of the raga, it is useful to realise that the technical boundary of a raga in performance is not explicitly specified by the grammar. Instead, it is implied by the notion that a performer can freely explore the tonal space as long as they are "not treading on another raga", potentially a closely related or allied raga (Kulkarni, 2011; Raja, 2016; Vijaykrishnan, 2007). "Allied ragas" share the same grammar in major attributes while differing in a few. For example, the pentatonic allied ragas, Deshkar and Bhupali, have the same set of svaras $\mathrm{S}, \mathrm{R}, \mathrm{G}, \mathrm{P}, \mathrm{D}$ (corresponding to 0, 200, 400, 700, and 900 cents respectively) and several common phrases in terms of svara sequences. Learners are typically introduced to the two ragas together and warned against confusing them (Kulkarni, 2011; S. Rao et al., 1999; van der Meer, 1980). A study of phrase variability would therefore be particularly meaningful in the context of allied ragas. Recently, subjective experiments on perceived similarity of synthetically manipulated raga phrases by musicians trained in the genre clearly demonstrated the existence

of a sharp boundary between valid variations of a given raga phrase from variations of the same phrase (in terms of the sequence of svaras) corresponding to an allied raga (Ganguli & Rao, 2019).

Table 1 presents a comparison of the melodic attributes corresponding to the grammars of the allied ragas as compiled from musicology texts by way of the Deshkar and Bhupali allied pair example. These cover the aspects of duration and intonation of the tonal material that includes ascending (arohana) and descending (avarohana) scales, dominant (vadi) and subdominant (samvadi) svara, and characteristic phrases. 'Natural *shruti*' (last row) refers to the Just Intonation tuning, but there is no quantification of the term 'higher'. Also, there is some indication of a duration constraint on R (as a short or passing svara relative to its neighbours) in the form of parantheses (e.g. G(R)S in raga Deshkar). The characteristic phrases of a raga (pakad) are typically referred to in terms of notation (i.e. as a sequence of svaras) but can be fully described only via the acoustic realization. The performer uses his knowledge of the raga grammar to interpret the notation when it appears in a written composition in the specified raga. The melodic-rhythmic form of a recurring melodic motif, in terms of continuous pitch versus time, within and across performances of the raga shows variability in terms of one or more of the following aspects: pitch intervals, relative note durations, and shape of *alankars* (ornaments), if any, within the phrase (P. Rao, Ross, & Ganguli, 2013).

A Hindustani vocal concert in the modern *khayal* style begins with an unmetered improvised section known as the *alap* which introduces the musical features of the raga through its phrases. Next, the chosen composition, or *bandish*, is identified by its lyrics and melody corresponding to a specific raga, tala and *laya* (tempo). There are typically three types of improvisatory movements, known together as *vistar*, that occur as the concert progresses, viz. the free-flowing *bol-alap*, and the faster and more rhythmic *bol-banth* and *taan* (Raja, 2016). In the earliest improvised section, the bol-alap, the singer elaborates within each rhythmic cycle of the tala using the words of the bandish interspersed with solfege and melisma (*akara*), purposefully reaching the strongly accented first beat (the *sam*) of the next rhythmic cycle on a fixed syllable of the signature phrase of the bandish. This recurring phrase, known as the *mukhda*, is the title phrase of the bandish that acts like a refrain throughout the exposition whereas the other components of the bandish can undergo extensive variation in the course of improvisation. In the improvisation, there is an increase in the tempo as the concert progresses. This is linked to a decrease in the tala cycle duration which is characterized by a fixed number of beats for a given tala. Also, there is a pattern of melodic development, involving gradually widening melodic range, from the lower octave, through the middle and then to the upper octave. The improvisation lasting several cycles of the tala is interspersed with the *sthayi* (first stanza of the composition) and later, when the improvisation reaches the higher (octave) $\overline{\text{S}}$, the *antara* (second stanza of the composition). The raga identity as embedded in the phrases is most distinct in the alap and vistar sections while the taan sections with their higher note density are more a display of musician virtuosity (Raja, 2005, 2016).

2.2. Musicological Perspective on Improvisation

Musicologists have attempted to formalize a performance model around a universal structure of musical improvisation. Nettl (1974) suggested that all music consists of a sequence of fixed structural points; if these are close together, the music is pre-composed; but if they are further apart, the performer must navigate from one to

Deshkar	Bhupali
Tonal material: SRGPD	Tonal material: SRGPD
Arohana: SGPD, $\overline{\text{S}}$PD$\overline{\text{S}}$ Avarohana: $\overline{\text{S}}$, PDGP, DPG(R)S	Arohana: SRG, PD$\overline{\text{S}}$ Avarohana: $\overline{\text{S}}$DP, GDP, GRS
Vadi: D, Samvadi: G	Vadi: G, Samvadi: D
Phrases: SG, G(P)DPD, P(D)$\overline{\text{S}}$P, DGP, DPG(R)S	Phrases: R$\underline{\text{D}}$S, RPG, PD$\overline{\text{S}}$, $\overline{\text{S}}$DP, GDP, GRS
Higher shrutis of R, G, D	Natural shrutis of R, G, D

Table 1. Specification of raga grammar for the two pentatonic allied ragas of the present study (Autrim-NCPA, 2017; ITC-SRA, 2017; Oak, 2017; S. Rao et al., 1999). The svaras S, R, G, P, D correspond to C, D, E, G, A respectively (with C as an arbitrary reference) of the Western chromatic scale. Overline/underline indicates higher/lower octave.

the next by interpolating blocks of material, whether spontaneously created or selected from a memory bank. Widdess (2013) considers this model of improvisation to fit Indian music well in many respects: the metrical cycle of tala provides fixed temporal points for action, especially the first beat of the cycle, called sam; a pre-composed block, one of the verses of the composition (bandish), alternates with improvised episodes; the melodic grammar of the raga provides pathways from note to note and motivic material for filling up space within the tala cycle. But this oversimplified form seems a rather one-dimensional concept of performance (Widdess, 2013), as if the performer needs only to think about one thing at a time and needs only to make decisions about how to get to the next structural point at the beginning of each "block". van der Meer (2008) describes improvisation as a range of different processes in Indian art music in which the line between 'fixed and 'free is extremely subtle. Many commentators of Indian music have observed that improvised music is based on memorized material and procedures, consistent with the oral tradition, and what is spontaneous is the decision to use a particular phrase or sequence of phrases at a particular instant of time (Nooshin & Widdess, 2006; Widdess, 2013). Nettl (1974) remarks that every performance must embed some of the 'points of reference' of the raga, else it will be considered "ignorance" (rather than creativity). That is, the chosen raga provides the building blocks in terms of the tonal material and representative melodic motifs in the form of the raga phrases.

Nooshin and Widdess (2006) and Widdess (2013) propose that, in Indian music, improvisation can be partly understood in terms of a small number of fundamental processes of development. The structuring of the large sections in performance are based on melodic expansion which describes gradual widening of melodic range to include successively higher, and/or lower, pitches. The reverse process, a gradual contraction of melodic range, can also occur, e.g. in the final descent of an alap or the descending phase of a tan. The next process, rhythmic intensification, defines on the large scale, existence of a gradual increase in tempo or rhythmic density, with different technical procedures becoming available at each new tempo. On the small scale, an individual motif can be progressively reduced in length at successive repetitions. Another important process is the development of individual pitches which says the a single pitch may, for a time, be treated as the focus of attention or "subject of discussion"; such a pitch may be repeated, emphasized, prolonged, and/or taken as the concluding note of successive phrases. Typically, the next higher pitch is hinted at dur-

ing the development of the previous pitch, before becoming the focus of attention in its turn. At each instance, of course, there are many possibilities and the "great musician chooses one while giving the feeling that every step was inevitable" (van der Meer, 1980). Given that the characteristic phrases are among the important reference points in a raga performance, we attempt to study the variation, if any, in the realizations of the phrases in the course of the improvisation.

2.3. Scope of this Work

Tallotte (2017) remarks, in the context of South Indian (Carnatic) raga music (which has much in common with Hindustani music), that creativity in the course of improvised sections of a performance lies in the emergence, within a specific development, of moving melodic-rhythmic variations of well-established phrases rather than the exhibition of new phrases. The modified phrases may be pre-composed and memorised, or imagined in the course of performance. Of the different processes of development, listed in the previous subsection, the two that are directly relevant to our study on raga phrase variation are the (i) rhythmic intensification or gradual increase in tempo across the performance, and (ii) shifting melodic focus by treating a single pitch as the focus of attention for a while. The musical cues to the latter are the introduction of the note in the gradually widening melodic range that slowly includes successively higher and/or lower pitches and the perceived emphasis. The melodic-rhythmic variation of the phrase is expected to be influenced then by the local tempo and the status of its constituent notes in terms of local melodic focus. Apart from this, the structural constraints provided by the tala cycle and the chosen composition could play a role.

Our approach will be to carry out the empirical analyses of audio recordings for suitably chosen examples that can help us to understand the influence, if any, of the mentioned improvisation processes on the acoustic realization of the phrase. The analyses involve measurements relating to selected pitch and duration attributes estimated from signal segments corresponding to the phrases of interest extracted from concert recordings. As mentioned before, the significance of variation in a given raga phrase can be appreciated better if we consider it with reference to its homonymous counterpart phrase in an allied raga. We therefore choose a pair of allied ragas, Deshkar and Bhupali, and phrases that occur in both ragas for our study. As seen in Table 1, the phrases, or melodic motifs, GRS and PD$\overline{\text{S}}$ are suitable candidates for our study. We create a dataset of concerts and study the variability of the phrases in the different measured contexts within and across concerts. The specific questions we seek to address are (i) what is the nature and extent of variation in a given phrase and whether the distinctiveness of the raga is maintained (in terms of discriminating it from the allied raga), (ii) how and to what extent is the variability predictable from the context. To study the latter question, we consider the following specific contexts derived from what is known about the broad structure of a performance.

- Temporal location in the course of the concert (i.e. the tala cycle index): the cycle index loosely serves as a marker in time for the attained speed as well as the stage of melodic development in the concert.
- Location with respect to structural points such as the sam (tala cycle boundary): the position of a phrase within the tala cycle is significant in terms of its distance from the upcoming sam.
- Whether and which svara is focal in the corresponding cycle, whether the phrase

is realised in a higher or lower octave, etc.

3. Dataset and Annotation

Serra (2014) stresses that the fundamental attribute of a research corpus lies in its ability to capture the essence of a particular music culture. The paper advocated the relevance of five criteria that were accounted for in the compilation of the CompMusic [1] collection (Serra, 2011, 2014), namely purpose, coverage, completeness, quality, and reusability. We have carefully designed our dataset ensuring the representativeness and diversity of the Hindustani repertoire. The scope of this work has been limited to vocal music concerts, considered the most spontaneous expression of the raga nuances – even instrumentalists are taught ragas by vocalizing them. While different datasets of Hindustani music have been used as training corpora for mainstream retrieval tasks, the present dataset of concert sections is curated for our focused empirical ethnomusicological study, and serves something between between a corpus based study and close listening in context. The concerts used in this study come from either the Hindustani music dataset from the Dunya corpus [2] or augmented by personal collections. The editorial metadata for all the concerts is publicly available in MusicBrainz [3]. The artists are acclaimed musicians of the Hindustani vocal (khayal) genre from the past 7 or 8 decades, hailing from different *gharana*s (lit. school of thought). The accompanying instruments are the drone (tanpura tuned to the singer's tonic), tabla and possibly also a melodic instrument such as the harmonium or sarangi. A comprehensive account of the concert metadata is presented in Table 2. The pitch range of any performance spans approximately two octaves (octave of the tonic and about half of each of the lower and upper octaves). All of the concerts comprise elaborations based on a selected bandish. The mentioned duration is that of the usable segment (corresponding to the alap and the vistar of the bandish) and the time-stamps as labeled by a trained Hindustani musician. For reproducibility, the available links to the audio material, which are either part of Dunya or acquired from artist(s) under a creative common license, as well as the annotations can be found on GitHub [4].

We select the melodic motifs, GRS and PD$\overline{\text{S}}$, that are common to the ragas Deshkar and Bhupali. Thus the next set of annotations corresponds to the occurrences of the phrases. The key phrases of a raga are readily identified by the corresponding sequence of svara categories although the melodic realization includes specific intonations and transitions to/from neighboring svaras apart from context-dependent ornamentation (Kulkarni, 2011). With reference to allied ragas, he states that delineating the phrases becomes quite complex when the ragas in question comprise the same notes. The GRS phrase is common to the chosen two ragas and it happens to mark the end of the descending line of the scale. To bring in contrast, we include the ascending phrase PD$\overline{\text{S}}$, also common to the two ragas. A phrase which ends with $\overline{\text{S}}$ helps in terms of its higher frequency of occurrence as it is realized both in higher and middle octaves, apart from the fact that S is a *nyas* svara or resting note to return to. The main distinctive musicological features of PD$\overline{\text{S}}$ are the slowness of the D $-$ $\overline{\text{S}}$ glide and intonation of D. Kulkarni (2011) uses the notation "P(D)$\overline{\text{S}}$" to indicate a fast transition through D and the Autrim-NCPA (2017) website mentions the slow D $-$ $\overline{\text{S}}$ glides in

[1] https://compmusic.upf.edu/
[2] https://dunya.compmusic.upf.edu/Hindustani
[3] https://musicbrainz.org/
[4] https://github.com/kaustuvkanti/hindustani_raga_phrase_dataset

Bhupali. The GRS phrases are distributed across three octaves (upper, middle, and lower octaves), although lower octave instances are fewer. The reverse holds for PD$\overline{S}$. An account on the segmental and contexual measurements on the phrases used in this study is presented per concert in Table 3.

Raga	Audio ID	Artist	Tonic in Hz	Tala (Laya)	Bandish	Duration in min
Deshkar	DK_AC-1	Ajoy Chakrabarty	146.8	Ikwai (Madh)	Aai Ri Badariya	21:25
	DK_AC-2	Ajoy Chakrabarty	138.5	Teental (Madh)	Aai Ri Badariya	16:22
	DK_KA	Kishori Amonkar	233.0	Teental (Madh)	Piya Jaag	29:55
	DK_RK	Rashid Khan	155.5	Teental (Madh)	Hoon To Tore	21:54
	DK_VK	Venkatesh Kumar	155.5	Jhaptal (Madh)	Chidiya Chutava	20:45
	DK_UK	Ulhas Kashalkar	155.5	Teental (Madh)	Ab Na Sahe	16:29
Bhupali	BP_AC-1	Ajoy Chakrabarty	130.7	Ektal (Vil)	Prabhu Rang	33:46
	BP_AC-2	Ajoy Chakrabarty	146.8	Ektal (Vil)	Prabhu Rang	37:57
	BP_RK	Rashid Khan	155.5	Teental (Madh)	Karoge Tum	22:21
	BP_OD	Omkar Dadarkar	155.5	Ektal (Vil)	Jab Hi Sab	55:46
	BP_DVP	D V Paluskar	158.7	Ektal (Vil)	Jab Hi Sab	28:11
	BP_BGAK	B Ghulam Ali Khan	140.7	Ektal (Vil)	Prabhu Rang	23:59

Table 2. Metadata of concerts in our dataset. We list the artist, tonic, tala, laya (Vil: *vilambit* or slow tempo: $<$ 80 bpm, Madh: *madhyalaya* or medium tempo: 80-180 bpm), and analysed duration of each concert.

Audio ID	# Tala cycles	Cycle dur. range: sec	# GRS/PD$\overline{S}$ phr. (in alap)	Octave: H, M, L	Mukhda phr.: Yes, No	Mukhda appr.: Yes, No
DK_AC-1	30	24 – 19	18 (1) / 16 (1)	2, 16, 0 / 0, 13, 3	0, 17 / 10, 5	11, 6 / 3, 12
DK_AC-2	24	31 – 27	14 (1) / 13 (1)	1, 13, 0 / 0, 11, 2	0, 13 / 9, 3	9, 4 / 2, 10
DK_KA	45	23 – 18	20 (1) / 11 (1)	2, 18, 0 / 0, 9, 2	0, 19 / 0, 10	9, 10 / 5, 5
DK_RK	38	14 – 8	22 (2) / 19 (3)	3, 19, 0 / 0, 13, 6	0, 20 / 12, 4	12, 8 / 5, 11
DK_VK	28	9 – 7	18 (3) / 12 (1)	2, 16, 0 / 0, 10, 2	0, 15 / 0, 11	11, 4 / 6, 5
DK_UK	44	10 – 6	18 (1) / 15 (2)	3, 15, 0 / 0, 12, 3	0, 17 / 10, 3	13, 4 / 6, 7
BP_AC-1	26	59 – 53	25 (2) / 19 (2)	6, 18, 1 / 0, 14, 5	16, 7 / 0, 17	4, 19 / 9, 8
BP_AC-2	31	45 – 39	28 (1) / 15 (1)	4, 22, 2 / 0, 12, 3	16, 11 / 0, 14	5, 22 / 7, 7
BP_RK	45	11 – 9	54 (4) / 25 (3)	2, 52, 0 / 0, 20, 5	18, 32 / 0, 22	8, 42 / 14, 8
BP_OD	31	56 – 52	33 (2) / 24 (2)	4, 29, 0 / 0, 18, 6	24, 7 / 0, 22	5, 26 / 16, 6
BP_DVP	21	47 – 42	25 (1) / 18 (1)	3, 22, 0 / 0, 15, 3	15, 9 / 0, 17	6, 18 / 13, 4
BP_BGAK	25	39 – 36	23 (2) / 14 (1)	3, 18, 2 / 0, 9, 5	15, 6 / 0, 13	4, 17 / 6, 7

Table 3. Information about the number of (#) tala cycles, cycle durations in the vistar, and GRS/PD$\overline{S}$ phrases across concerts. We see the total count of instances (in alap and combined alap+vistar sections), octave belongingness (high (H), middle (M), and low (L)), whether the instance is a (non-)mukhda, and whether it is the penultimate phrase before an approaching mukhda of the given bandish.

4. Audio Processing

The concert sections of interest are segmented from the full audio recordings and processed to obtain the melodic-rhythmic features of the desired phrases. The steps involved in the process, as presented here, are predominant melody detection corresponding to the singing voice, and the segmentation of the raga phrases from the overall melody contour. This is followed by specified measurements of the sub-events comprising each phrase while also noting its location within the concert and other context information relevant to our study.

4.1. Predominant F0 Detection

Predominant-F0 detection is implemented using a method presented by V. Rao and Rao (2010) based on the harmonic grouping of sinusoidal partials detected in the poly-

phonic audio, via short-time Fourier analysis, to estimate one or more F0 candidates. Next, the spectral and temporal properties of the singing voice are exploited to discriminate its partials from those of any melodic accompaniment. Further, temporal smoothness constraints minimize pitch octave errors. The melodic pitch is detected at 10 ms intervals throughout the audio recording with no pitch assigned to the detected purely instrumental regions. The algorithm also estimates the vocal intensity (energy of the vocal harmonics) at each time step. The predominant-F0 detection algorithm is based on a small set of tunable parameters which can be selected for pitch tracking with high accuracy from preset values using available knowledge such as singer gender and whether the expected pitch variation is rapid or slow (V. Rao, Gaddipati, & Rao, 2011).

As our dataset comprises performances by a number of artists, male and female, the detected vocal melody must be normalized with respect to the tonic pitch. The fundamental frequency (F0) is converted to cents by normalizing with respect to the concert tonic determined using a classifier based multi-pitch approach to tonic detection (Gulati, Bellur, et al., 2014). With an accuracy of over 90 percent, any gross errors are easily corrected based on raga (or rather, allied raga group) information. The final preprocessing step is to interpolate short silence regions below a threshold of 250 ms, indicating musically irrelevant breath pauses or unvoiced consonants, by cubic spline interpolation, to ensure the continuity of the melodic contour (Ganguli, Gulati, Serra, & Rao, 2016). The threshold value 250 ms was chosen empirically from the distribution of the duration of the unvoiced segments in the pitch contour. This was observed to be bimodal in shape with a dip near 250 ms indicating a clear boundary between unvoiced segments during singing and musician-intended breath pauses. Median filtering with a 50 ms window is performed to get rid of spurious pitch fluctuations. Eventually, we obtain a continuous time series of pitch values in cents representing the melody line throughout the vocal regions of the concert.

4.2. Event Segmentation

The melodic contour can be viewed as comprised of three broad categories of segments: stable svara regions, transitory regions and pauses. The svara are the notes of the raga and play various roles such as that of the predominant note (vadi), resting note (nyas), etc. The transitions between notes can also be categorised into one of a set melodic ornaments (*alankar*) like *meend, andolan, kan, khatka* etc. (van der Meer, 1980). The third important melodic event is the pause. The underlying scale interval locations (i.e. the raga svara) are detected from the prominent peaks of the long-term tonic-normalized pitch histogram across the concert (Ganguli & Rao, 2018). The pitch histogram is computed by binning the singing voice F0 values estimated every 10 ms into 1 cent bins across the octave folded range of 12 semitones. Next, regions of melodic contour lying within a fixed neighbourhood of a svara location are segmented as corresponding to stable notes. The neighbourhood about the detected svara location, T_{tol}, is chosen to be ± 35 cents, fixed experimentally determined to optimize the recognition of the svara sequence from the melodic contour corresponding to a raga phrase across many different instances of the same phrase extracted from different concerts (P. Rao et al., 2014).

The above step retains relatively flat segments of the pitch time-series that approximate the scale notes, while omitting the transition regions. Next, a lower threshold duration of T_{dur} is applied to discard fragments that are considered too short to be

perceptually meaningful as held svaras (P. Rao et al., 2013). T_{dur} is empirically set to 250 ms, supported by previously reported subjective listening tests (Vidwans, Ganguli, & Rao, 2012), where 250 ms was observed to be the minimum required duration in musicians' annotation of a segment as "khada svara" (standing or stable note). Finally, we have a string of labeled svara, each of duration over 250 ms, separated by pauses. Segments with the same note (svara) value that are separated by gaps less than 100 ms are merged. The svara sequence information (i.e. scale degree and absolute duration of each steady pitch segment) across the concert recording is stored and then searched for the raga phrases of interest e.g. the sequence $\mathrm{G}, \mathrm{R}, \mathrm{S}$ could indicate the occurrence of the raga Bhupali motif GRS. Once the coarse location of a potential raga phrase is confirmed by the musician annotator, the continuous pitch contour corresponding to the phrase is automatically extracted from the concert pitch contour. We also label the sam (first beat or downbeat) of each tala cycle with a semi-automatic tala tracker based on information about the mean tempo and a few hand-marked downbeats (Srinivasamurthy, Holzapfel, & Serra, 2017).

4.3. Acoustic Measurements

A phrase has a melodic-rhythmic realization comprising specific intonations and durations of its constituent svaras, together with the transitions to/from neighboring svaras (Kulkarni, 2011). Figure 1 shows a representative GRS phrase from each of the ragas. Distinctive features suggested by the comparison are: (i) durations of each of the stable svara regions, (ii) the durations of the glides connecting the svaras, and (iii) the pitch interval of the svara G.

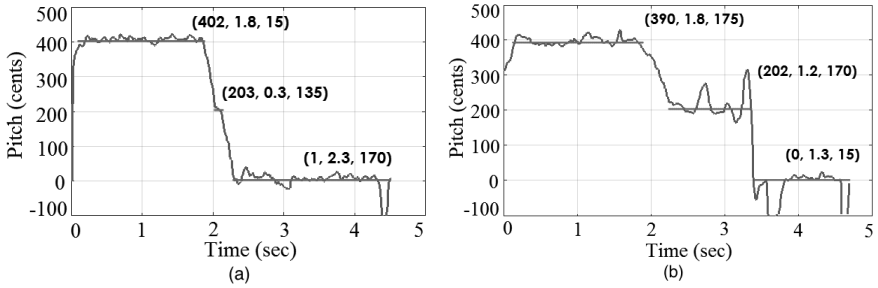

Figure 1. Pitch contours (F0 versus time) for arbitrarily chosen GRS phrases from (a) raga Deshkar and (b) raga Bhupali. The horizontal lines indicate the segmented svaras with the features corresponding to each svara (*Intonation*, *Duration*, *Slope*).

We describe a phrase as a sequence of 'events' that can each be described by the chosen features. For the GRS phrase in question, we consider the following five events, i.e. the individual svaras $\mathrm{G}, \mathrm{R}, \mathrm{S}$, and the $\mathrm{G}-\mathrm{R}$ and $\mathrm{R}-\mathrm{S}$ transitions. The selected features are: (i) *Start_time* : onset of an event, (ii) *Duration*, (iii) *Intonation* : precise pitch interval location of a stable svara in the octave obtained as the median pitch value over the duration of the svara, and (iv) *Slope* : gradient between the mean of last 20% and the first 20% pitch samples of a stable svara segment. Figure 2 shows the matching measurements for the $\mathrm{PD}\overline{\mathrm{S}}$ phrase.

The measurements can be categorized as follows.

(1) Segmental measurements
  (a) Duration based measurements
    - Overall duration of the phrase: We define this as the duration extending

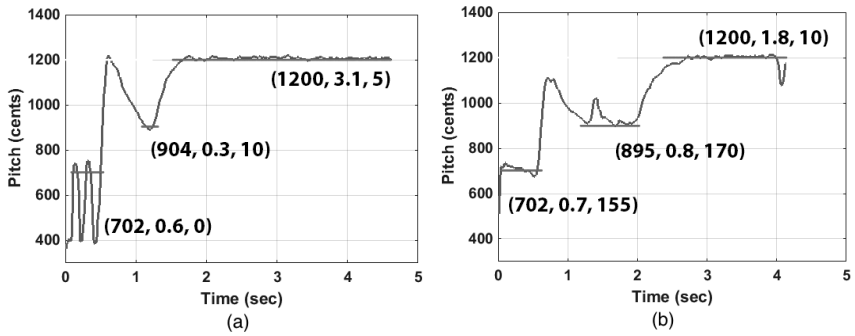

Figure 2. Pitch contours (F0 versus time) for arbitrarily chosen PD$\overline{\text{S}}$ phrases from (a) raga Deshkar and (b) raga Bhupali. The horizontal lines indicate the segmented svaras with the features corresponding to each svara (*Intonation*, *Duration*, *Slope*). The raga Deshkar PD$\overline{\text{S}}$ phrase displays gamak/kan and an overshoot toward $\overline{\text{S}}$ on the P − D glide.

from G onset to the S onset. This includes the two svaras G and R, and the transients G − R and R − S. Similarly for the PD$\overline{\text{S}}$ phrase. The final event (S/$\overline{\text{S}}$) is a resting note and there are no structural considerations expected in how long it is held.
- Duration of each constituent segment (svaras and transients)
- Relative duration of the events within a phrase

(b) Intonation based measurements
- Intonation of the segmented stable svaras in terms of median pitch interval and slope of the pitch contour. Median is chosen over mean/mode as this is less affected by presence of kan svaras (touch notes) within a stable note segment.

(c) Intensity based measurements

(2) Measurements related to context

(a) Temporal context
- Whether belonging to the unmetered (alap) or metered composition (vistar)
- Tala (in terms of beat count) and local tempo (as represented by the cycle duration)
- Location of cycle within concert: cycle index (absolute and normalized with respect to cycle count)
- Location within cycle: proximity of the approaching sam (both in terms of beat count and absolute time)

(b) Octave location information (lower, middle, or higher)

(c) Metadata-based measurements
- Whether the phrase in question is a mukhda or a non-mukhda instance
- The local "focal note"

5. Observations and Discussion

We intend to use the acoustic signal analyses to validate some of the previously reviewed musicological assumptions. Further, we expect the analyses to inform us about other, possibly implicit, principles followed by musicians during improvisation. In this section, we discuss our observations with references to the musicology where applica-

ble.

5.1. Occurrence of the Raga Characteristic Phrases

Improvisation during performance draws from the motivic material of the raga although the decision to use a particular phrase at a particular time is a part of the extemporaneousness (Nettl, 1974). Signature raga-characteristic phrases serve to reinforce a listener's melodic expectancy (Powers & Widdess, 2001; Widdess, 2013). We expect therefore that the phrases, as markers of raga identity, recur often through the performance. In Table 4, we provide estimates of the frequency of occurrence of each of our motifs in the concerts of our dataset. The frequency in terms of phrases per minute comes from the number of occurrences of the phrase divided by the duration of the concert section as provided in Table 3. The phrases per rhythmic cycle is more easily interpreted. We observe that a given motif (GRS) occurs as often as once in two cycles in Deshlar and in every cycle in Bhupali (with its longer cycle durations). While the mukhda motifs occur in the same position within a cycle, terminating on the first beat (sam) of following cycle, the remaining phrases can occur anywhere within a cycle depending on the melodic improvisation or, even span across cycle boundaries.

Measure of frequency of occurrence (average over the dataset)	Deshkar		Bhupali	
	GRS	PD$\overline{\text{S}}$	GRS	PD$\overline{\text{S}}$
A: phrases/min	0.84	0.66	0.90	0.54
B: phrases/cycle	0.48	0.36	0.98	0.58

Table 4. Frequency of occurrence of each of the raga-phrases in the analysed sections of the concerts.

5.2. Rhythmic Intensification

We note a wide range for tala cycle duration within a concert in Table 3. The duration of a cycle is the time-interval between its sam (downbeat) and that of the next cycle. Given that the number of beats per cycle is fixed by the chosen tala, this implies variations in the local tempo. Informal discussions with musicians revealed that they visualize the tala cycle more directly (rather than the local bpm) as the "available space" for the improvisation during a performance. Within a concert, of course, the cycle duration is inversely related to the local tempo. Clayton (2008) remarks, in the context of instrumental concerts, that tempo increases as the concert progresses. Wade (1979) mentions the increasing speed or "acceleration" of the vocals in both dhrupad and khayal concerts. However we do not find specific comments on the rate of change (whether gradual or step-wise).

Assuming tempo increase with concert progression, we expect the cycle duration to decrease with time into the concert. Figure 3 shows plots of measured cycle duration vs cycle index (both normalised to 0-1 range) for each of the concerts along with the corresponding linear fit. We see the expected trend of decreasing cycle duration, with concert-dependent slope, and note that the $\mathcal{R}^2$ values (goodness-of-fit) are relatively high. The concert BP_RK has the worst fit (at $\mathcal{R}^2 = 0.59$) because the artist adopts a step-wise, rather than gradual, increase in tempo.

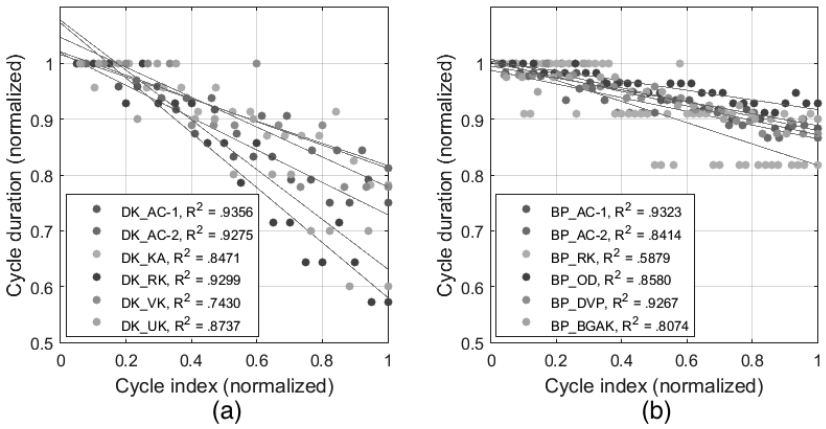

Figure 3. Scatter plot of cycle duration (normalised with respect to maximum cycle duration for each concert) versus normalized cycle index for concerts in (a) Deshkar, (b) Bhupali. The $\mathcal{R}^2$ value indicates the goodness of linear regression model fit. All values are found to be significant ($p < 0.001$).

5.3. Dependence of Phrase Duration on Local Tempo

Widdess (2013) remarks that on the small scale, an individual motif can be progressively reduced in length at successive repetitions. In the typical musical context of a composed piece, we expect the duration of a given phrase to change linearly with the inter-beat duration (inverse of tempo). The variation in cycle duration across a concert ranges between 10 % to 50 % as seen in Table 3. We examine the distribution of phrase durations in Figure 4, separately for the mukhda and non-mukhda instances given that the former can be considered precomposed.

We see that the non-mukhda phrases show larger spread but even this is no more than 10% of the mean absolute duration in seconds. This indicates that phrase durations change very slowly, if at all, with changing local tempo within a concert. Figure 5 shows a scatter plot of the purely non-mukhda phrases (i.e. Deshkar GRS and Bhupali PD$\overline{\text{S}}$) durations versus cycle duration for the different concerts. Superposed is the best linear fit for each concert. While we see the expected positive slope, the fits are poor indicating that phrase duration variability is not quite well explained by the tempo variation from the rhythmic intensification across the concert.

Turning to musicological explanations for alternate factors, we note that the tala cycle boundaries serve to delimit improvised utterances to within a cycle. Thus it is common for the artist to complete within the same cycle, a phrase started at some instant during the cycle while leaving room for the mukhda towards the end of the cycle. To validate this explanation, we show in Figure 6 phrase duration versus its proximity (distance from the phrase onset in seconds) to the approaching sam. We observe a dense clustering of instances but also several outliers that reveal that the artist extends the phrase duration at both very low and high distances from the approaching sam. While the latter is expected given that there is more room for the phrase to stretch, it is observed that an extension of phrase duration also occurs when the sam is in close vicinity based on the artist's decision to do away with the mukhda at that point. In this case, we have a long phrase going across the cycle boundary and, consequently, a missed mukhda.

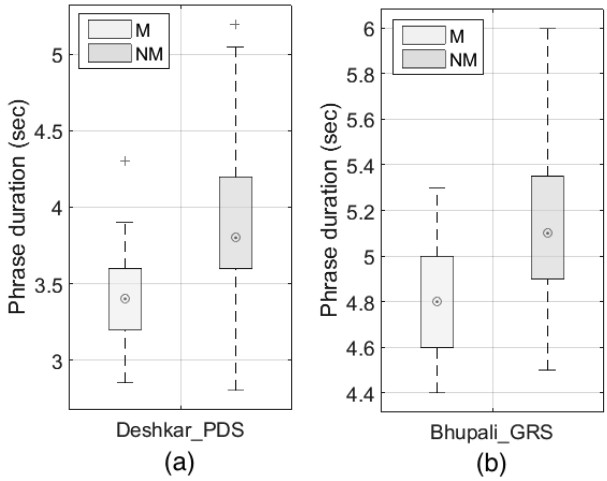

Figure 4. Distributions of phrase duration (sec) for mukhda (M) and non-mukhda (NM) instances: (a) PD$\overline{\text{S}}$ phrase in raga Deshkar, and (b) GRS phrase in raga Bhupali.

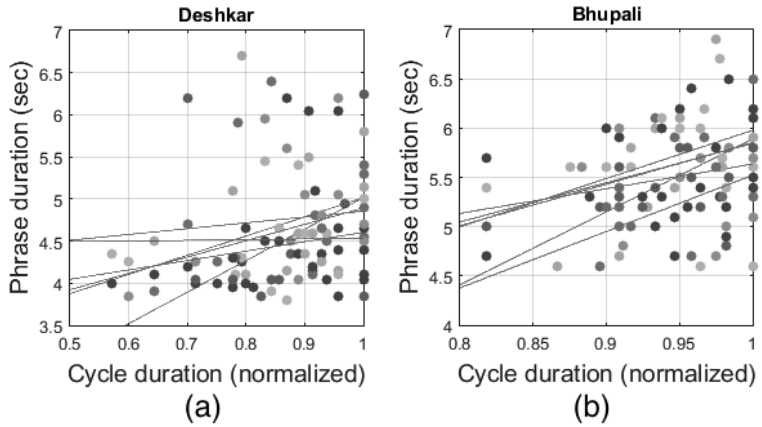

Figure 5. Scatter plots of phrase duration versus the cycle duration (normalised with respect to maximum cycle duration for each concert) for non-mukhda instances of (a) GRS phrase in raga Deshkar and (b) PD$\overline{\text{S}}$ phrase in raga Bhupali. All the goodness-of-fit $\mathcal{R}^2$ values are in the range [0.17, 0.32] with $p < 0.01$.

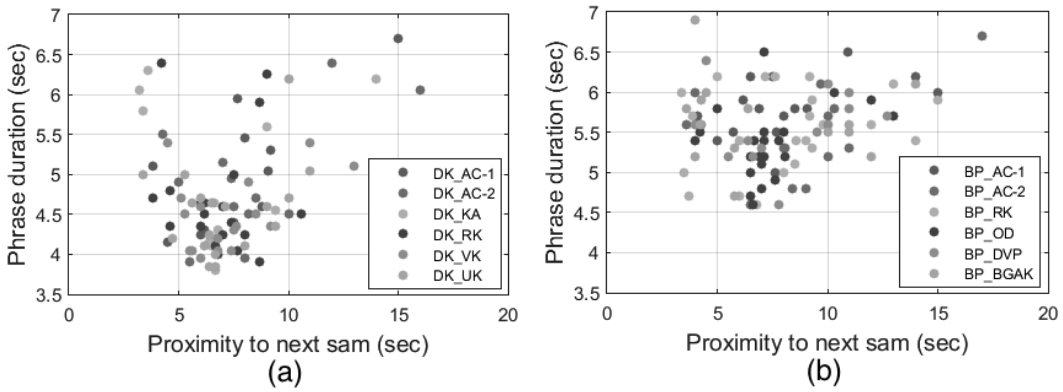

Figure 6. Dependence of phrase duration on proximity of the approaching sam for non-mukhda instances of (a) GRS phrase in raga Deshkar and (b) PD$\overline{\text{S}}$ phrase in raga Bhupali.

5.4. Variations in the Melodic-Rhythmic Structure of Phrases

van der Meer (1980, 2008) remarks that in a typical Hindustani music concert, an artist executes variations of the raga characteristic phrases that represent the raga identity in the course of the processes related to improvisation. From the many possibilities, "the great musician chooses one while giving the feeling that every step was inevitable." We have just seen that phrase duration varies in a complex manner with tempo and other factors. Deeper insights are likely to be obtained by examining variations in the phrase shape. The melodic contour of a segmented phrase can be described by the relative durations of its constituent notes and transitions, and by the intonations of the notes. The raga grammar of Table 1 specifies particularly the notes R, G and D for their unusual intonation (shruti or pitch interval). We show in Figure 7 the distribution of svara median pitch for G and D segmented from the corresponding raga phrases for each of the two ragas. The note R was not included due to its short duration in Deshkar. We note that the Deshkar svara are indeed pitched slightly higher than the corresponding Bhupali svara (which are closer to just intonation intervals). Further, the distributions of the same svara are non-overlapping across the two ragas validating the musicological assumption that artists carefully maintain the distinction between ragas in the realization of the common phrases.

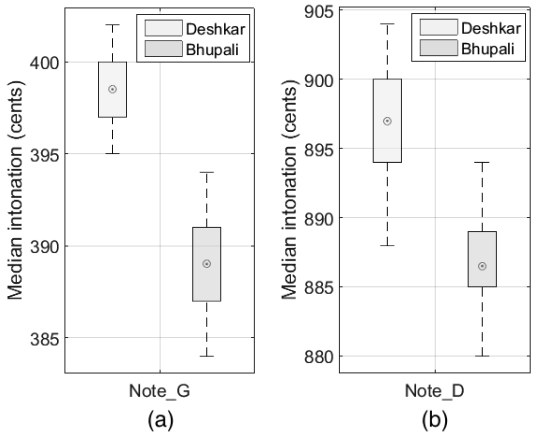

Figure 7. Distribution of median intonation of the (a) G svara in GRS phrase and (b) D svara in PD$\overline{S}$ phrase in the two ragas.

The relative durations of the notes comprising the phrase contribute to its rhythmic structure. Figure 8 shows distributions of the absolute duration of each of the constituent events of the phrase across ragas and phrases. These constituents contribute to the phrase duration variation reported in the previous section. We see that the extent of variation is highly non-uniform across segments with the transition segments remaining nearly fixed duration. That is, when a phrase is stretched or compressed, the melodic transitions remain unchanged while the stable svara regions are modified to different extents. While the Bhupali svara durations have similar spreads, the R and D notes of Deshkar are narrowly confined. In Table 1, we see that the Deshkar R svara is marked by parantheses, and designated as *alpatva* ("little" or weak). The R and D svara are considered to be non-emphasized in the given contexts and give the associated Deshkar phrases a distinct rhythmic structure that serves as an important cue to differentiation from the homonymous phrases of Bhupali (Ganguli & Rao, 2019).

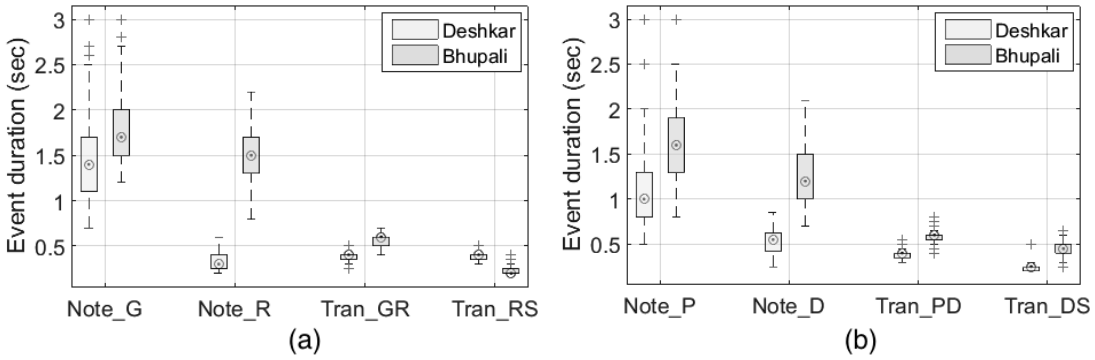

Figure 8. Distributions of event durations across the (a) GRS and (b) PD$\overline{\text{S}}$ phrase instances in ragas Deshkar and Bhupali across concerts.

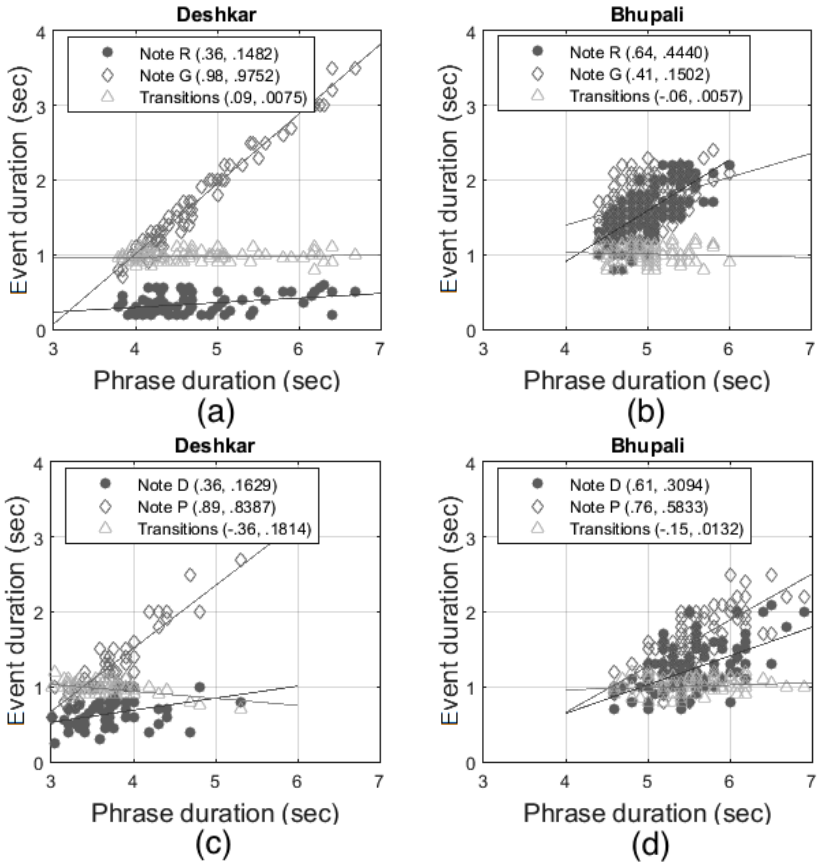

Figure 9. Correlating event durations with overall phrase duration for (a) Deshkar GRS, (b) Bhupali GRS, (c) Deshkar PD$\overline{\text{S}}$, and (d) raga Bhupali PD$\overline{\text{S}}$. In brackets (correlation coefficient, $\mathcal{R}^2$ goodness of fit). All the values are significant ($p < 0.01$).

In Figure 9, we plot the variation of the constituent segment durations versus the overall phrase duration. We observe, as expected, that the durations of the transition segments and the constrained svara of the Deshkar phrases (R, D) are weakly correlated with phrase duration (as seen in the low $\mathcal{R}^2$). As a consequence, all Deshkar phrase duration changes are absorbed by the G (in GRS) and the P (in PD$\overline{\text{S}}$). On the other hand, with no grammar-specified constraints on svara durations in the Bhupali phrases, the duration changes in principle affect both svara in each of GRS and PD$\overline{\text{S}}$ phrases. We observe moderate correlations with phrase duration for the Bhupali svara R, G, P and D. It can be assumed that there are additional contextual factors, among those listed in Section 2.3, at play that influence the artist's choice of segment durations.

In Figure 10, we show the durations of R and G for each instance of GRS in Bhupali, as they vary across each concert. We see a cyclic variation in the duration of R and the roughly complementary pattern in the variation of G. This is explained by the phenomenon of the melodic focus shifting to and past R, and recurring in the higher octave towards the end of the concert. The regions of different melodic focus were labeled by a musician. The G duration evolves in a roughly complementary manner to R as expected for constant overall phrase duration. The boxplots in Figure 10 compares the duration distributions of the same svara in focal and non-focal contexts. The focal notes have higher mean duration.

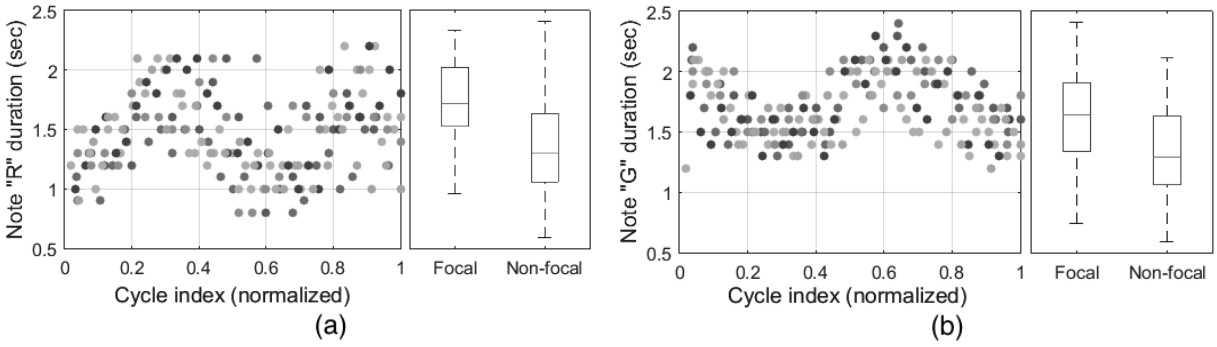

Figure 10. Svara *Durations* for (a) R and (b) G in GRS phrase instances in raga Bhupali versus cycle index. The subplots (right) show the boxplot for the same in focal and non-focal contexts. The precise locations of melodic focus are R (0.3 and 0.9) and G(0.6) in units along the normalized cycle index. The different shades represent different concerts.

5.5. Other Observations

This section presents a set of observations relating to other potential predictors of variation in raga phrase shape. Due to either the lack of a sufficient number of labeled instances in our phrases dataset or the difficulty in obtaining reliable estimates of certain parameters, we content ourselves with qualitative descriptions supported to the extent possible by musicological explanations.

- Tala dependence: The phrase shape variability seems to be independent of the tala as long as the cycle durations are similar. We have an example of the same bandish being performed in two different talas of 16-beat cycle (concerts DK_AC-1 and DK_AC-2) of comparable cycle duration. There was no detectable pattern indicating dependence on tala. It is known that the coupling between the melody and metrical accent is loose in khayal. There is no known relationship between the

emphasized svara and metrical accent. While the tala comprises a fixed number of beats and provides the regular division of time, melodic improvisations depart freely from the internal divisions (van der Meer, 1980).

- Intensity variation with emphasis: Given the occurrence of emphasized notes in the course of melodic development, it is of interest to investigate the acoustic correlates of the melodic focus. Apart from the already discussed frequency and duration, a potential cue is the relative intensity of the note. The presence of tanpura and tabla along with the vocals makes the estimation of local intensity difficult. Given that we reliably detect the fundamental frequency (pitch) of the singing voice, an estimate of the intensity is obtained by summing the energy across the spectral regions of the signal corresponding to the vocal harmonics. We obtain an energy contour sampled at 10 ms intervals, similar to the pitch contour. The mean intensity is computed across the voiced region of each svara segment. In the case of notes with focus, it is observed that the duration and intensity tend to roughly trade off with each other. Further it is observed that the intensity of the focal notes is higher in the higher octave, which phenomenon may have more to do with the physiology of the singing voice.
- Ornamentation: While we have noted the influence on phrase duration of its proximity to the approaching tala cycle boundary, it is also seen that larger pitch modulations can occur on the svara near the boundary. As we have seen, the approaching sam constrains the phrase duration. The gamak-like oscillatory motion in the melodic realization then serves to make the reduction less obvious and soften its impact.
- Octave dependence: The octave location of the phrase introduces a change in its immediate preceding context, e.g. the melodic motif in the higher octave is more likely to be near the boundary of the singer's range and hence may be approached by 'coming from below', i.e. via lower pitched svara. The different ways of approaching the phrase thus contribute to some of the variability in realizing the same phrase in different octaves.

6. Summary

Computational analyses of audio recordings can provide insights into music performance practice of potential value to musicological, cultural, and historical studies. This is especially useful in the case of improvisatory traditions with weak notation systems such as Indian art music. Using the alap and vistar improvised sections of a dataset of medium and slow tempo concerts across a pair of allied ragas, we applied melody extraction, phrase detection and event segmentation techniques to study raga phrase shape variation. Various local melodic and rhythmic contexts motivated by the musicological literature were considered.

A raga phrase is notated by the svara sequence, but realised and recognised by a specific melodic-rhythmic shape defined by the relative durations of its constitutent notes and the connecting pitch transitions including ornamentation. We observed variation in phrase realisation in terms of its pitch contour given the melodic constraints of raga grammar and the rhythmic constraints of the tala cycle. The raga grammar requires the characteristic phrases to represent the raga unambiguously, in particular, with reference to its allied raga counterpart. The tala cycle specified by the number of beats and relative accenting plays the role of segmenting the improvisation into cycle-length measures, typically also delimited by the melodic refrain phrase (mukhda).

Our findings are that the tempo increases with concert progression typically gradually, and sometimes in a step-like manner corroborating what is well known to those familiar with khayal. The phrase duration, contrary to expectation, does not scale linearly with the beat duration. With beat duration explaining very little of the phrase duration variation, a dependence is observed instead on distance from the approaching sam (or tala cycle boundary). A reduction occurs when the phrase onset is relatively close to the sam and the refrain is intended to be sung. Further, in the reduction or elongation of the phrase, the constituent events (notes and transitions) do not scale uniformly. Instead, the raga-specific characteristics (transitions and special notes) change very little maintaining thus the melodic categorization of the raga. The remaining svara undergo duration changes that seem to further depend on whether the note is the object of melodic focus. In summary, our findings support the observation of van der Meer (1980), viz. "The only concern is to return to the sam. For continuity and to avoid the feel of undecidedness, he (the performer) has a number of stereotyped reductions and enlargements of the phrase."

Applications of such a study include digital tools and interfaces for music appreciation and pedagogy. Further, the notion of similarity in the context of raga music is refined by the understanding of within-raga variation in the tonal space and the sharp between-raga phrase shape distinctions. This was seen to aid the choice of computational features in a raga discrimination MIR task (Ganguli & Rao, 2017). Finally, with the growing interest in 'music artificial intelligence', such work can serve in computer music synthesis with rules for genre-specific aesthetics such as adapting raga phrases to different speeds, similar to the work of Subramanian, Wyse, and McGee (2011) on modeling gamakas in Carnatic Veena music.

Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author Contributions

KG and PR contributed to the design of the study. KG contributed to music concepts and data preparation. KG and PR contributed to the analyses and writing the manuscript.

Funding

This work received partial funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ ERC grant agreement 267583 (CompMusic).

Acknowledgments

The authors would like to thank Pt. Ajoy Chakrabarty for the musicological expertise provided.

References

Ashley, R. (2009). *Musical improvisation.* The Oxford handbook of music psychology.

Autrim-NCPA. (2017). *Music in Motion: The automated transcription for Indian music (AUTRIM) project by NCPA and UvA.* url: https://autrimncpa.wordpress.com/. (Last accessed: April 26, 2017)

Bagchee, S. (1998). *Nād: Understanding raga music.* Business Publications Inc.

Chakrabarty, A. (2002). *Shrutinandan: Towards universal music.* Macmillan.

Chordia, P., & Şentürk, S. (2013). Joint recognition of raag and tonic in north Indian music. *Computer Music Journal*, *37*(3), 82–98.

Clayton, M. (2008). *Time in indian music: rhythm, metre, and form in north indian rag performance.* Oxford University Press.

Danielou, A. (2010). *The ragas of Northern Indian music.* Munshiram Manoharlal Publishers.

Deva, B. (1974). Music and science. *Journal of the Indian Musicological Society*, *5*(1), 19.

Dutta, S., Krishnaraj, S. P., & Murthy, H. A. (2015). Raga verification in Carnatic music using longest common segment set. In *Int. soc. for music information retrieval conf. (ismir)* (pp. 605–611).

Ganguli, K. K. (2013, October). How do we 'See' & 'Say' a raga: A Perspective Canvas. *Samakalika Sangeetham*, *4*(2), 112–119.

Ganguli, K. K., Gulati, S., Serra, X., & Rao, P. (2016, August). Data-driven exploration of melodic structures in Hindustani music. In *Proc. of the international society for music information retrieval (ismir)* (pp. 605–611). (New York, USA)

Ganguli, K. K., Lele, A., Pinjani, S., Rao, P., Srinivasamurthy, A., & Gulati, S. (2017). Melodic shape stylization for robust and efficient motif detection in hindustani vocal music. In *National conference on communications (ncc).*

Ganguli, K. K., & Rao, P. (2015, February). Discrimination of melodic patterns in Indian classical music. In *Proc. of national conference on communications (ncc).*

Ganguli, K. K., & Rao, P. (2017, October). Towards computational modeling of the ungrammatical in a raga performance. In *Proc. of the international society for music information retrieval (ismir).* (Suzhou, China)

Ganguli, K. K., & Rao, P. (2018). On the distributional representation of ragas: experiments with allied raga pairs. *Transactions of the International Society for Music Information Retrieval (TISMIR)*, *1*(1), 79–95.

Ganguli, K. K., & Rao, P. (2019). On the perception of raga motifs by trained musicians. *The Journal of the Acoustical Society of America*, *145*(4), 2418-2434. Retrieved from `https://doi.org/10.1121/1.5097588`

Ganguli, K. K., Rastogi, A., Pandit, V., Kantan, P., & Rao, P. (2015, October). Efficient melodic query based audio search for Hindustani vocal compositions. In *Proc. of the international society for music information retrieval (ismir)* (pp. 591–597). (Malaga, Spain)

Gulati, S., Bellur, A., Salamon, J., Ranjani, H. G., Ishwar, V., Murthy, H. A., & Serra, X. (2014). Automatic tonic identification in Indian art music: Approaches and Evaluation. *Journal of New Music Research (JNMR)*, *43*(1), 53–71.

Gulati, S., Serra, J., Ganguli, K. K., & Serra, X. (2014). Landmark detection in Hindustani music melodies. In *Proc. of int. computer music, sound and music computing* (pp. 1062–1068). (Athens, Greece)

Gulati, S., Serrà, J., Ishwar, V., Şentürk, S., & Serra, X. (2016). Phrase-based rāga recognition using vector space modeling. In *Ieee int. conf. on acoustics, speech and signal processing (icassp)* (pp. 66–70).

Gulati, S., Serrà, J., Ishwar, V., & Serra, X. (2014). Mining melodic patterns in large audio collections of Indian art music. In *Int. conf. on signal image technology & internet based systems (sitis-mira)* (pp. 264–271).

Gulati, S., Serrà, J., & Serra, X. (2015). Improving melodic similarity in Indian art music using culture-specific melodic characteristics. In *Int. soc. for music information retrieval conf. (ismir)* (pp. 680–686).

ITC-SRA. (2017). *Distinguishing between similar ragas.* url: http://www.itcsra.org/Distinguishing-between-Similar-Ragas. (Last accessed: April 26, 2017)

Jairazbhoy, N. A. (2011). *The rags of north indian music: Their structure & evolution* (second ed.). Popular Prakashan.

Kulkarni, S. (2011). *Shyamrao gharana* (Vol. 1). Prism Books Pvt. Ltd.

MA, R., TP, V., & Rao, P. (2020, October). Structural segmentation of dhrupad vocal bandish audio based on tempo. In *Proc. of int. soc. for music information retrieval conf. (ismir).* (Montreal, Canada)

Nettl, B. (1974). Thoughts on improvisation: A comparative approach. *The Musical Quarterly*, *60*(1), 1-19.

Nettl, B., & Russell, M. (1998). *In the course of performance: Studies in the world of musical improvisation.* University of Chicago Press.

Nooshin, L., & Widdess, R. (2006). Improvisation in iranian and indian music. *Journal of the Indian Musicological Society*, *36*, 104–119.

Oak, V. (2017). *22 shruti.* url: http://22shruti.com/. (Lat accessed: April 26, 2017)

Powers, H. S., & Widdess, R. (2001). India, subcontinent of. In (2nd ed., chap. III: Theory and practice of classical music). Macmillan, London. (Contributions to S. Sadie (ed.))

Raja, D. (2005). Hindustani music: A tradition in transition. D. K. Printworld.

Raja, D. (2016). *The raga-ness of ragas: Ragas beyond the grammar.* D.K. Print World Ltd.

Rao, P., Ross, J. C., & Ganguli, K. K. (2013, December). Distinguishing raga-specific intonation of phrases with audio analysis. *Ninaad*, *26-27*(1), 59–68.

Rao, P., Ross, J. C., Ganguli, K. K., Pandit, V., Ishwar, V., Bellur, A., & Murthy, H. A. (2014, April). Classification of melodic motifs in raga music with time-series matching. *Journal of New Music Research (JNMR)*, *43*(1), 115–131.

Rao, P., Vinutha, T. P., & Rohit, M. A. (2020). Structural segmentation of alap in dhrupad vocal concerts. *Transactions of the International Society for Music Information Retrieval*, *3*(1).

Rao, S., Bor, J., van der Meer, W., & Harvey, J. (1999). The Raga Guide: A survey of 74 Hindustani ragas. Nimbus Records with Rotterdam Conservatory of Music.

Rao, S., & Rao, P. (2014, April). An overview of Hindustani music in the context of Computational Musicology. *Journal of New Music Research (JNMR)*, *43*(1).

Rao, V., Gaddipati, P., & Rao, P. (2011). Signal-driven window-length adaptation for sinusoid detection in polyphonic music. *IEEE transactions on audio, speech, and language processing*, *20*(1), 342–348.

Rao, V., & Rao, P. (2010). Vocal melody extraction in the presence of pitched accompaniment in polyphonic music. *IEEE Trans. on Audio, Speech & Language Processing*, *18*(8).

Ross, J. C., Vinutha, T. P., & Rao, P. (2012, October). Detecting melodic motifs from audio for Hindustani classical music. In *Proc. of int. soc. for music information retrieval (ismir).*

Serra, X. (2011). A multicultural approach to music information research. In *Proc. of int. conf. on music information retrieval (ismir)* (pp. 151–156).

Serra, X. (2014). Creating research corpora for the computational study of music: the case of the Compmusic project. In *Proc. of the 53rd aes int. conf. on semantic audio.* London.

Srinivasamurthy, A., Holzapfel, A., Ganguli, K. K., & Serra, X. (2017). Aspects of tempo and rhythmic elaboration in hindustani music: A corpus study. *Frontiers in Digital Humanities*, *4*, 20.

Srinivasamurthy, A., Holzapfel, A., & Serra, X. (2017). Informed automatic meter analysis of music recordings. In *18th international society for music information retrieval conference, ismir 2017, 23 october 2017 through 27 october 2017, suzhou, china* (pp. 679–685).

Subramanian, S., Wyse, L., & McGee, K. (2011). Modeling speed doubling in carnatic music. In *International computer music conference (icmc).*

Tallotte, W. (2017). Improvisation, creativity, and agency in south indian temple rāga performance. *Asian Music*, *48*(2), 24–61.

van der Meer, W. (1980). *Hindustani music in the 20th century.* Martinus Nijhoff Publishers.

van der Meer, W. (2008). Improvisation versus reproduction, india and the world. *New Sound: International Magazine for Music*(32).

Vidwans, A., Ganguli, K. K., & Rao, P. (2012, July). Classification of indian classical vocal styles from melodic contours. In *Proc. of the 2nd compmusic workshop.*

Vijaykrishnan, K. G. (2007). *The grammar of carnatic music.* De Gruyter Mouton.

Wade, B. C. (1979). *Music in india: The classical traditions.* Prentice-Hall Englewood Cliffs, NJ.

Widdess, R. (1994). Involving the performers in transcription and analysis: a collaborative approach to Dhrupad. *Ethnomusicology*, *38*(1), 59–79.

Widdess, R. (2011). Dynamics of melodic discourse in Indian music: Budhaditya mukherjees alap in rag Puriya-Kalyan. *Analytical and Cross-cultural Studies in World Music, ed. Michael Tenzer and John Roeder. Oxford University Press, New York*, 187–224.

Widdess, R. (2013). Schemas and improvisation in Indian music. *Language, Music and Interaction. London: College Publications*.