

Time, gesture and attention in a *khyāl* performance

Martin Clayton

Open University

Revised submission for Asian Music, Dec 2006

Additional diacritics needed:

Subscript dot under **ḍ** in adacautal

Subscript dot under **ṭ** in baithak, theka, tarab

Subscript dot under **ṛ** in jor, mukhra

Subscript dot under both **ṇ** and **ṭ** in bant

~ over last **ṇ** in Nainana men

Substitute <sharp> with sharp sign

Introduction

North Indian *rāg* performance, especially as practised in intimate and informal settings, is often distinguished by a lively interaction involving both musicians and listeners, mediated by gestures and vocal interjections. Performers gesture to each other, to the audience, and expressively with the music, and audiences become part of that process. The premise of this study is that observing the behaviour of audience members, as well as that of performers, should provide a valuable window onto the ways in which *rāg* performance is experienced by all of its participants. The main questions I aim to elucidate in this paper are: What does observable behaviour tell us about the way people experience the metrical and formal structures of a *rāg* performance? When and how do listeners become involved in the performance gesturally and/or verbally? I shall address these and related questions through an analysis of a *khyāl* performance by Vijay Koparkar recorded in Mumbai in 2005. Detailed analysis of this performance indicates that these questions can indeed be answered using observational methods, and also suggests other important issues that may not have been raised had this approach not been adopted.

Observing the behaviour of listeners alongside that of performers can yield vital clues about the relationships between all participants in a performance event. This behaviour – including hand and head movements – informs us only about certain aspects of the participants' experiences, of course. It would be incorrect to assume that the extent of the audience's verbal and gestural involvement indicates directly the strength of their emotional response, or that such involvement is a straightforward response to the music unaffected by other aspects of social relationships. Audiences respond, to some extent, because they perceive it be expected of them: as Goffman would have it, they perform their role as audience members in this particular form of social encounter (Goffman 1969 [1959]). Nonetheless the evidence of audience behaviour remains a vital source of information that has been often remarked upon but rarely investigated: it can elucidate listeners' role in the performance, and – complementing the study of performers' behaviour – help us to draw a more complete picture of performance dynamics. The methods employed here draw both on ethnomusicological precedents and on work in gesture studies, cognitive psychology and elsewhere, and the results are relevant to each of these disciplines.

The study in context

Despite the existence of a sophisticated theory of *tāl* (metre), research on the rhythmic organisation of *rāg* music remains weak in several areas, for example in the interpretation of the temporal organisation of *ālāp* or of the perceptual significance of the long *tāl* cycles in *vilambit khyāl* – both instances where, in the absence of detailed theories elsewhere, theories of time organisation and experience have to be developed from a very low base in the context of Indian music studies (see Clayton 2000; for an introduction to *khyāl* singing see Clayton and Sahasrabuddhe 1998). This paper study forms part of a large-scale project, “Experience and meaning in music performance”, which aims – amongst other things – to develop new theoretical models of musical time that take account of entrainment theory (see Clayton, Sager and Will 2005) and of the significance of physical gesture in musical performance. These theoretical models are applied and tested through empirical analysis, exploiting the potential of digital video and audio recording, as well as ethnographic interviews. Some of these ideas and methods are introduced below, alongside reference to some of the most relevant literature in ethnomusicology, cognitive psychology and gesture studies, and illustrated in this analysis.

Slawek has argued that understanding Indian *rāg* music involves understanding its performance, and that this depends on attending to audience reactions (1990): this route may offer hope of advances in problematic areas of rhythmic analysis too. His observations suggest that vocal interjections from the audience occur not only at structurally important points such as the arrival at *sam* (‘beat one’, the temporal focus of most extemporisation), but also in response to dramatic musical gestures or technical feats. For Slawek, an important element in such study should focus on performer intention. It is striking, however, that detailed study of the behaviour of audiences has not been carried out in the sphere of *rāg* performance, and no-one has yet followed up Slawek’s suggestion with a detailed study of performers’ intentions in their deployment of gesture and their interaction with audiences. Racy has gone somewhat further in his study of Arabic *tarab* performance (2003), which seems to have some features in common with my topic here, but even the fascinating observations he presents would not be sufficiently precise or empirically based to answer the sorts of questions posed in this paper (although interestingly, many of his comments on audience involvement, like mine below, stress the importance of the *timing* of audience contributions).

The single most important methodological precursor of this study is in fact Qureshi's 'contextual input model', which uses video transcriptions to study interactions between participants in Sufi assemblies in India and Pakistan. Qureshi's approach was constructed in order to overcome two challenges: that of describing interactions between two domains – 'music' and its 'context' – and that of analysing process without reducing it irreversibly to structure (Qureshi 1995 [1986], 135). The contextual input model views the event from the performer's perspective, and analyses context only in so far as it observably impacts on the music (136). Qureshi's analyses are based on video recordings, and two novel graphical tools are introduced: the *videograph*, in which a conventional musical transcription is placed in parallel with transcriptions of each participant's behaviour (each person, or group of people, being allotted one row in the transcription); and the *videochart*, in which the musical transcription is broken down into chunks and juxtaposed with the most pertinent behavioural observations so that evidence of interaction between the two can be extracted.

The approach taken here owes a lot to Qureshi's model, with some differences of emphasis. The first difference is that, rather than analysing music and context as separate domains – music as the aural product of the performers, context the behaviour of members of the assembly – my analysis treats performers and audience as one group of participants constructing the event together. (Of course, at times it remains necessary to abstract elements of the music in the same way that Qureshi does, and it would be ridiculous to ignore the performers' privileged status as the primary sound producers: I am not proposing a radical break from Qureshi's model). My graphical representations are also somewhat different to Qureshi's, reflecting the different performance dynamics: Qureshi plots audience *responses to* the music and their *influence on* the subsequent performance, but in this case the audience is mainly displaying their *participation in* the performance.

This participation is felt by many to be a significant part of intimate *rāg* performance, similarly to the *tarab* performance described by Racy: "practitioners tend to view direct and continued interaction between performer and listener as a prime condition for good entertaining" (2003: 65). Many Indian performers will insist that since the audience cannot affect what they sing (the *rāg*, compositions and so forth), the quality of their music is not dependent on the audience: nonetheless they will acknowledge, somewhat like the *tarab* performers, that an attentive and knowledgeable audience can significantly

enhance a performance event. Enhancement or otherwise, audience participation of the kind described below certainly forms a significant part of the performance as experienced by all present.

This study refers to the concept of attention, and in doing so is influenced by psychologist Mari Riess Jones's theory of *attentional periodicity*. According to Jones our attentional resources are rhythmical – quasi-periodic – and many aspects of temporal behaviour, including music and speech, involve the entrainment of these attentional rhythms (see e.g. Large & Jones 1999, Barnes & Jones 2000, Drake, Jones and Baruch 2000). Thus, for example, since speech has distinctive rhythms, people listening to speech can most easily follow the speaker's meaning by locking into the speech rhythm and focussing their attentional energy at the most salient points in the speaker's discourse. Aspects of temporal structure in music, such as metre or cadential formulae, likewise help the listener to organise her attentional resources. Jones's theory is an integral part of London's theory of musical metre (2004), and was a significant influence on my interpretation of north Indian *tāl* (2000; see Clayton, Sager and Will 2005 for a more detailed exposition of this topic). One of the questions explored in this study is whether it is possible to track listeners' patterns of attention, and to discover their relation to *tāl* and other aspects of musical time.

The present study also draws on the insights of gesture studies – an interdisciplinary field allied to linguistics in which researchers have interpreted the relationship between verbal utterances and physical gesture in conversation. Two key figures here are David McNeill and Adam Kendon. McNeill argued (1992) that gesture and language are two complementary aspects of a single system, and presented analyses of the relationships between the two (including their temporal relationships). In *Gesture and Thought* (2005) he has developed this theory further, drawing inspiration from the Russian psychologist Lev Vygotsky, suggesting that speech and gesture emerge together from a single 'growth point'. Kendon's work shares many features with McNeill's, although he expresses scepticism over the growth point idea and places more emphasis on the role of gesture in social interaction (Kendon 2004). Both use similar systems of analysis and representation, involving first the transcription of speech and its parsing into phrases, and then the analysis of gestures on a phrase by phrase basis. Gestures are categorised according to their function in relation to speech (the way in which they illustrate or otherwise complement the linguistic meaning), and the flow of movement can be parsed into 'Gesture Units' (the

longest meaningful units of gesture, between which hands will occupy a rest position), and ‘Gesture Phrases’ (subdivisions of the gesture unit) (see McNeill 1992: 83f). Gesture Phrases can be broken down temporally into three main phases: preparation, stroke and retraction. In this paper, analysis of the gestural phrase structure – units, phrases, and their constituent parts – can be juxtaposed with that of the musical (melodic) phrase structure so that their relationship is determined: this could also be glossed as a study of the relationship between ‘intensity contours’ in different modalities (see Eitan and Gronot 2006 for a summary of interdisciplinary work on this idea and its application to music analysis).

Although there has been increasing interest in applying the principles of gesture studies to musical performance, there are relatively few detailed studies in press: the work of Jane Davidson stands out in particular (Davidson 1993, 2001, 2005; Williamon & Davidson 2002). I have outlined a number of proposals concerning the most appropriate scheme for classifying gestures, with reference to Indian music performance in particular, and written up a couple of short examples illustrating how a singer’s gestures change character between *ālāp* and *jor*, and how the relationship between musical phrases and gesture phrases might be analysed (2005, in press). The three categories of gesture most relevant to the particular performance discussed below can be described briefly as follows (terms in square brackets are from Rimé and Schiaratura 1991).

- **Markers** [nondepictive gestures] of musical process or structure, include marking focal moments such as the *mukbrā* in *ālāp*, or beating out a regular pulse or the *tāl* structure.
- **Illustrators** [depictive gestures] are tied to the content of the singing, appearing analogous to the melodic flow or “motion”.
- **Emblems** [symbolic gestures] have verbal equivalents: “Well done”, “Take a solo”, and so on. This type of gesture is often used by musicians to instruct subordinate musicians (e.g. telling *tānpūrā* players to play louder, or *tablā* players to play faster), to offer approval, or to invite the audience or fellow musicians to share appreciation of the music.

In my gestural analysis I have followed an analogous procedure to McNeill and Kendon’s gesture studies: parse both the music and gesture into units and then see how the two relate to each other. I find detailed analysis of this kind most economical using multitrack recordings and Praat software.¹ The vocal track can be analysed to produce both a pitch plot and an intensity plot, and the *tablā* track can be used to

mark up the *tāl* or metrical structure: this is fairly straightforward in *khyāl* where for the most part *tablā* players repeat a *theḱā* (a more or less standard drum pattern identifying the *tāl*). Pitch and intensity plots of the voice form the basis of an analysis of phrase structure, and can be presented in parallel with the mark up of the *tāl* cycles. Where appropriate, transcriptions of text and/or pitch (in *sargam* or solfège) can be appended to each phrase (see Figure 2, below).

Gestures are studied with the help of video recordings and observational analysis software. This involves the coding of specific aspects of behaviour (such as changes of hand position or direction of gaze, or the striking of instruments) in relation to the video time code (see Clayton 2007).

Physical, and particularly manual gesture does, of course, have an important historical role not only in musical behaviour but also in musicological discourse in India. The best known, and perhaps the most persistent, function of hand gestures is the marking of metrical structure: in modern North Indian practice two principal gestures are employed – a clap (sometimes substituted by a slap on the thigh) and a silent wave – sometimes augmented or substituted by the counting of beats on the fingers (Clayton 2000, 61ff). In earlier times the system used in elite musical traditions was more intricate, using up to eight different hand gestures to indicate the metrical structure. Moreover, early cheironomic practice in Samavedic recitation included indications not only of metre but also of pitch and other features (Rowell 1992, 65). Indian classical dance traditions have developed their own systems of hand gestures or *mudrās*, many of which are symbols deployed in representational or pantomimic performance. Important as these traditions are, however, the gestures considered here are largely of a different order: a proportion can be regarded as marking the metrical structure of the music, albeit not in a highly codified manner; the rest are described by musicians as ‘natural’, ‘unconscious’ or ‘automatic’. In the words of the singer considered in this article, Vijay Koparkar:

Whenever we are performing, whatever the body language is, it is very natural, there is no artificial thing. Because whenever the *sam* [beat one] is coming we have some body language, whenever we are extending the *sūr* [pitch] [...] there is some expression. It is different for everyone. It is not one and the same, [like] this one [gesture] is this one note. So that is a very natural process and it should be giving pleasure to the audience. (Interview with Vijay Koparkar, 20 May 2005.)

In summary, the present study applies insights from ethnomusicology, cognitive psychology and gesture studies, together with digital video and audio recording and associated analysis software, to the study of a *khyāl* performance, focussing on metrical structure and audience involvement. In presenting a detailed account of performance interactions, I also present evidence supporting Koparkar's own contention that his use of gesture is largely 'natural' and not codified.

The performance

This study focuses on the analysis of a *khyāl* performance in Rāg Multānī by Vijay Koparkar, recorded on 20 May 2005 at IIT Powai, Mumbai.² This performance was part of a *baithak*-style event: that is, audience as well as performers sat on rugs on the floor in close proximity to each other (the performers were only slightly raised). This type of setting is a common one for the performance of Hindustani music, one that is appreciated by many listeners as well as performers for its intimacy and informality. (Notwithstanding this intimacy the event was amplified, as is the case even in most such events). Vijay Koparkar's performances of Rāgs Multānī and Swānandī comprised the first part of the performance, beginning at around 5.40pm; he was followed by Veena Sahasrabuddhe, and both singers were accompanied by Viswanath Shirodkar (*tabla*) and Seema Shirodkar (harmonium). Veena Sahasrabuddhe's students provided *tānpūrā* support – in Vijay Koparkar's case, Surashree Ulhas Joshi (left of Figure 1, see below) and Bageshree Vaze (right of Figure 1). The description and analysis below concentrates on Vijay Koparkar's performance of Rāg Multānī.

Method

Vijay Koparkar's performance was recorded using four digital video cameras: recordings were made in miniDVCAM format (progressive scan PAL at 15fps),³ and miniDV (interlaced PAL 25 fps). The four cameras were all locked down onto tripods and took static shots, since zooming and panning can complicate the movement analysis. The four shots can be briefly described as:

Camera 1 (miniDVCAM): Wide frontal master shot of all performers

Camera 2 (miniDVCAM): Vijay Koparkar and Seema Shirodkar (harmonium) from stage left

Camera 3 (miniDVCAM): Audience from stage left

Camera 4 (miniDV): Closer frontal shot of performers (Vijay Koparkar and the two *tanpura* players)

A separate multitrack audio recording was made in Pro Tools (24 bit, 96kHz) using balanced feeds from the live mixer. The preparatory stage of analysis involved synchronising the four camera views in Avid Express Pro, and creating a composite video image (the view of Camera 4 with the other 3 shots placed as ‘picture in picture’ windows in the upper part of the screen – see Figure 1), using a stereo audio mix from Pro Tools as the sound track. Using this file on DVD I made outline notes of the contents of Vijay Koparkar’s performance of Rāg Multānī.

<<Figure 1 near here>>

Vijay Koparkar’s (hereafter VK) performance of Rāg Multānī lasted a little over 49 minutes, and comprised a short *alap* (c. 5 mins) followed by a *vilambit khyāl* (a slow vocal composition lasting c. 35 mins) and a *dhrut khyāl* (a faster vocal composition lasting c. 9 mins). The *bandishes* performed by Vijay Koparkar (hereafter VK) were the following.⁴

1. Rāg Multānī, vilambit ektāl

Sthāyī

Gokula gāva kā chora re Barasāne kī nār re

The boy from Gokul village, the girl from Barasānā [Krishna and Radha]

Antarā

Uno dou man mohaliyo man kahe Sadārang bāta re

These two have enchanted my mind, so Sadārang says

2. Rāg Multānī, drut ektāl

Sthāyī

Nainana men āna bāna kauna sī parī

What has fallen into my eyes?

Antarā

Bāra bāra jovata palakana lāgata jita dekho uta Shyāme sī parī

Time and again I try but cannot fall asleep – I see Shyām [Krishna] wherever I look

I selected 3 clips for detailed analysis, all from the slow section (*vilambit*) – I concentrated on this section because of my interest in studying the temporal structures in this slow section and their physical expression. For these clips I prepared individual video files for each camera view, as well as individual audio files for the singer, *tablā* and harmonium tracks. Praat was used to generate pitch and intensity

tracks of voice and harmonium, and a wave form display and intensity plot of the *tablā* track. The Observer (observational analysis software) was used to log details of hand and eye movements of the participants, by watching each camera view in turn and then combining the results.⁵

Results

1. Summary description of performance

The processes of *rāg* performance in *khyāl* are not completely standardised, but the details of this presentation are well within the normal range of possibilities. In the *vilambit* VK starts by presenting the *sthāyī* (first section of the composition), ‘Gokula gāva kā chora’. The *tāl* is *ektāl*, comprising 12 *mātrās* or time units. (Each *mātrā* is subdivided into four pulses that determine the tempo: in the following analysis therefore I will stick to the term ‘*mātrā*’ for the longer time units and use ‘beat’ only for the subdivisions.) VK then elaborates with *bol alāp* (melodic development similar to *alāp* in style, but using the text and accompanied by *tablā*) interspersed with episodes of *sargam* (solmization) before switching to *tāns* (more rapid vocalisation either to the text, the vowel ‘aah’ – *ākār* – or *sargam*). The *antarā* (second section of the composition) is introduced around 29’. At 32’ the musicians accelerate, from what has been a very steady tempo just over 15 *mātrās* per minute (cycle = 48 secs) to just over 20 *mātrās* per minute (cycle = 35 secs),⁶ the density of the *tablā* accompaniment increases and VK sings with the text in a more rhythmically defined, syllabic style (*bol bānt*), later switching back to *tāns* until the end of the *vilambit* section at c. 40’.

He then introduces the *drut khyāl*, ‘Nainana men āna bāna’, in fast *ektāl*: first presenting the *sthāyī* and then inviting Viswanath Shirodkar [VS] to play his opening *tablā* solo before resuming with further *tāns*. The *antarā* of the *drut khyāl* (the second of the two sections, focussing on the upper tonic or Sa) is heard from c. 42’30”, and the remainder of the presentation features more *tāns*, a couple of accelerations, and another *tablā* solo. The global tendency to increase tempo and rhythmic definition, and the episodic improvisation punctuated by the *mukhārā* (opening phrase) of each composition, are standard features of *khyāl* presentation. VS maintains his *thekā* for the most part, while harmonium player Seema Shirodkar [SS] shadows the singer’s melodic line, filling in the many gaps while VK rests: in the *vilambit khyāl*

presentation he typically rests for 10-15 secs at the start of each cycle before beginning a new episode of improvisation.

VK's gestural style is consistent with a generally sober demeanor. In the *ālāp* he rests his right hand on his knee while gesturing gently with his left, maintaining an open hand position. As the *vilambit khyāl* picks up he starts to use both hands, sometimes resting his right hand on his left in front of his chest, sometimes moving his two hands in and out together (as in a mirror image), and sometimes moving his hands in parallel from side to side in a kind of swaying motion. Almost every *sam* (beat one) in the *vilambit* section is indicated gesturally by an upward, rhythmically marked movement of his left hand, and a simultaneous tap of his right hand on his thigh. As the speed of his singing increases so, on the whole, do his gestures: an interesting feature, however, is that he sometimes begins *tāns* with his hands completely still in his lap, starting to move them only after a few seconds – and sometimes marking the end of a particularly vivid *tān* with a sharp upward motion of both hands. He occasionally makes a lateral shaking motion with both hands while singing *gamak* (a rapid oscillation), but this is quite restrained: likewise, his adoption of a more closed hand position is not dramatic – he does not use pinching or grasping hand shapes. Most of VK's movements can be categorised as Illustrators – apparently moving 'along with' the melody – but with short rhythmic movements (Markers) included in them. More definite Markers can be observed in the counting of the *tāl* on his knee or in front of his chest, which he does occasionally. Most instructions to his accompanists are achieved simply by making eye contact and making subtle head movements – only once does he make a definite instructional movement, beating the time of a new tempo when asking VS to speed up (Emblem). He makes no obviously depictive gestures – he hints at 'stretching' occasionally as he pulls his two hands apart, or 'shaping' a spherical body in front of his chest, but in neither case unambiguously – and it is very difficult to read his gestures as obviously 'expressive' in any direct way – of the words, or of the mood of the *rāg*.

VK keeps his eyes closed, or looks down, in the opening stages: he makes little or no direct eye contact with anybody. (This is consistent with his and other performers' descriptions of *ālāp* as the most introverted portion of a performance.) As the *vilambit khyāl* progresses he makes eye contact principally with his two accompanists – first with VS just before beginning the *vilambit khyāl*, then usually with both in turn, before and after the *sam* (the precise pattern varies, but he usually looks at SS first, then swivels

his head to look at VS). He makes very little eye contact with audience members, principally at *sam* towards the end of the *vilambit* and during the *drut*.

As for the other musicians, their gestural communication can be distinguished. Harmonium player SS concentrates most of her attention on VK, occasionally glancing briefly at the audience or *tablā* player VS. She makes few overt gestures, occasionally making short head or hand movements either in time with the beat (especially the *sam*) or in response to a vocal phrase: apart from the need to concentrate on what the soloist is doing, she is of course restricted physically by the need to keep her hands on her instrument. *Tablā* player VS, in contrast, is considerably more mobile, frequently shifting his visual attention between his fellow musicians and audience members, showing his approval, making eye contact and smiling sympathetically, and emphatically marking certain beats with extended arm movements or by swaying his body forward and/or to the right. This difference in manner was described by the two accompanists in an interview carried out the day before this performance:

SS [on harmonium accompaniment] My attention is [on the singer] – I can't look here and there. I should give my full attention, without that it doesn't work.

VS [on *tablā* accompaniment] But my role [is that] whenever I'm playing the *theḱā*, it is also my duty to help [the singer] convey [his] message to the audience. So whenever it is rhythmic I always interact with the audience... and show them that "this is the *sam*, and this is the way she is going to come". So I also interact with the audience. [...]

SS But I can't. I can't look like that, because [...] I have to support [the singer] for each note. Whichever note [he] sings I have to reach there: I should not be behind. If I am stuck in Sa [when the singer is on Ga], then she will think "What's wrong with you, come with me!"... That is why I can't look here and there while playing. My full attention should be with the main artist.

Interview with Seema and Viswanath Shirodkar, Mumbai, 19 May 2005 (SS mostly translated from Hindi).

Besides confirming their understanding of their different roles in the group's interaction – harmonium player fixed on the singer, *tablā* player interacting more widely – this also confirms that the time for such interaction is in the 'rhythmic' part (i.e. not in the *ālāp*), and that its focus is on the *āmad*, the approach to *sam*. Both of these tendencies can be observed in this performance. The soloist's intention may be displayed by subtle movements, and accompanists have to remain aware in order to pick up subtle hints, as Sadanand Naimpalli suggests to students in his manual of *tablā* playing: "Be very alert and observe

every movement or nuance of the main artist. Many times instructions are conveyed to the accompanist by small gestures of the hands, head or eyes” (Naimpalli 2005, 69).

The two *tānpūrā* players, Bageshree Vaze and Surashree Ulhas Joshi, mostly fix their gaze either straight ahead or downwards: they occasionally make eye contact with VS or with an audience member, or show their enjoyment of the music with restrained head movements, but not in a conspicuous way. This is consistent with their supporting role in the performance, in which they are required to provide a consistent drone (and, perhaps just as importantly in these days of the electronic *tānpūrā*, its visual analogue), without assuming a more proactive role in the musical performance. (Had they been performing with their own teacher, Veena Sahasrabuddhe, they may well have been asked to sing occasional phrases, but they would not expect to do so for another singer.)

The audience appears to be disposed in two groups, with two rows at the front of the auditorium (roughly 6-8 feet away) and another group at the back. This may well have been partly the result of our presence, those at the back indicating an unwillingness to sit within range of our cameras. All audience members, in both groups, display a highly concentrated attention to the performance over the 50 minute duration: there is very little coming and going, or individual communication between audience members, as is sometimes quite prevalent: it is impossible to judge to what extent this is due to the quality of the performance, to what extent to our presence. At the start of the performance audience members are very static, showing very little physical movement (mostly shifting position as if to achieve more comfortable sitting positions, rather than obviously responding to the music). Once the *tabla* begins, a significant proportion of the audience marks each *sam* (and occasionally other stressed beats), with head and/or hand movements. This begins as a single ‘beat’ – typically a downward stroke of palm on thigh, or a sharp sideways movement of the head – but within a few minutes it takes on a more elaborate character, with a marked ‘upbeat’ (typically an upward hand or head movement) preceding the beat. In terms of gestural analysis, this can be described as an extended preparation phase before the ‘stroke’. Apart from marking the time structure, audience members occasionally make responsive gestures of approval, for instance immediately after an impressive *tān*, sometimes accompanied by vocal interjections. The energy level of the audience noticeably builds following the acceleration in the slow section (32’) through to the early stages of the fast *khyāl*, where many count out the full cycle of *drut ektāl* with hand gestures. This activity soon subsides however, so although the energy expended by the musicians increases more or less

linearly through the performance, that of the audience peaks between 36 and 42 minutes and then subsides in the latter stages of the *drut khyāl*.

The singer himself states that he is well aware of the degree of attention and appreciation the audience is showing:

VK One of my concerts was at Sawantwari [near] Goa. It is just like a village, not a city, and the audience is the layman audience. Ninety percent of the people, they don't know the theory of the classical [music]. [...] I have started with Shree [*rāg*] and *jhūmrā* [*tāl*]. Both of the things were difficult to understand, but from that day I never underestimate the audience, because they were enjoying every beat and every *sam* and every *sur* [note]. [...] So they are coming with me to the *sam* in *jhūmrā tāl*. It is very difficult, *jhūmrā*. [...]

MC When you say the audience is feeling the *sam* [...] how do you know that?

VK From their expression: and if we are coming [to *sam*] with some design, and aesthetically, there is “Vāh”. That is the expression to encourage, and that is the appreciation: “Vāh, kyā bāt hai”.⁷ Or [...] with their heads, [through] their body language we come to know that they're understanding all the things. Every *sam*, they are coming with me.

(Interview with Vijay Koparkar, 20 May 2005)

2. Detailed transcription of clips

The clips chosen for more detailed analysis are as follows.⁸

No	Timing	Duration (secs)	Content
1	04.39-06.04	85	Transition from <i>ālāp</i> to <i>vilambit khyāl</i> , including one cycle of the latter. This clip was chosen in order to cover the moment of transition from unmetred to metred rhythm. It also provides footage of VK's gestures, and of the behaviour of accompanists and audience, in the <i>ālāp</i> phase.
2	18.34-19.36	62	Full cycle of <i>vilambit ektāl</i> featuring <i>sargam</i> . This clip was

			chosen as a sample of the middle portion of the <i>vilambit</i> , and as a section with prominent interaction between VK and his accompanists.
3	36.01-36.49	48	Full cycle of <i>vilambit ektāl</i> featuring <i>ākār tāns</i> , at faster tempo. The clip was selected as a sample of the latter stages of the <i>vilambit</i> , and because it contains the most emphatic (and audible) audience response.

In each case, four separate video clips were prepared for observational analysis, and five different audio clips (Camera 1 sound, a stereo mix, and the separate tracks for voice, harmonium and *tablā*). An outline gestural transcription was prepared from the four video files, using The Observer. The three individual audio tracks were analysed in Praat, producing pitch tracks for voice and harmonium and intensity tracks for all three. An example of the result can be seen in Figure 2, a summary of the first 40 seconds of Clip 1.

Clip 1: Transition from *ālāp* to *vilambit khyāl* (04:39-06:04)

This clip was edited to include the final phrase of *ālāp* and the first cycle of *vilambit ektāl*, in which VK presents the *sthāyī* of the *bandīś* (composition) ‘Gokula gāva kā chora’. An obvious point of interest is how this transition is effected, and how the tempo of the composition is agreed between VK and his *tablā* accompanist.

VK begins the clip with eyes half-closed, making no eye contact; he gestures with left hand only, resting his right on his thigh. As he reaches the end of his final *ālāp* phrase, a descent from Pa to Sa (\wedge^5 to \wedge^1), he turns to VS with eyes open: this clearly indicates that he is about to begin his composition. VS actually turns to look at VK around 26” into the clip, two seconds *before* the latter turns to him to give his signal. This suggests that VS is anticipating the very signal he then receives, and indeed the two musicians acknowledge each other in synchrony: the mutual gesture is a confirmation that they both understand what is to come next, rather than a signal from one to the other. VK immediately starts to use his right hand: following a dramatic rise the downward sweep of this hand on ‘go-’ can be seen, retrospectively at

least, to mark the 12th *mātrā* of *ektāl*, and is followed by two finger taps (marking the 3rd and 4th quarters of the 12th *mātrā*) before a full palm stroke on *sam* (33.56"). For the next two *mātrās* he taps out each quarter-*mātrā* with his right palm, in time with the *tablā* strokes.

As he continues the *bandīś*, VK's first gestures involve both hands moving in contrary motion (as in a mirror); he then reverts to tapping with his right index finger, continuing the Illustrator gestures with left hand only. He then curls his fingers in slightly and moves his hand diagonally, almost as if imitating a sitarist (although moving his hand in the opposite direction – upwards as his melody ascends to the upper tonic and beyond, whereas a sitarist would move his hand towards the bridge of his instrument). Overall his gestures include Markers, Illustrators, and Emblems, shifting between and sometimes combining these functions. A good example of this is at the start of the *bandīś* as noted above – where having glanced at VS (Emblem) he moves his right hand up and then sharply down, both accompanying the melodic movement (Illustrator) and then marking a beat (Marker); he then continues to count time with his right hand (Marker) while raising his left hand, palm upwards (Illustrator).

VS begins the clip looking mainly ahead, with a quick glance at the audience. The timing of his glance towards the singer clearly suggests that he expects the *bandīś* to begin, which is predictable from the development of the melody in the *ālāp* – which, having focussed on Ma and Pa (^4<sharp> and ^5) now descends in a relaxed manner towards the Sa (^1). He acknowledges VK's look as noted above (28.20"), moves his hands to his drums to signal that he is ready to play (29.48"), and then marks *sam* emphatically – right on the beat. His eyes are more mobile than VK's – he alternates looking down, looking at VK, and looking at the audience.

SS has two glances at the audience (36.64-38.04"/57.52-59.04"), but otherwise focuses on VK. She marks the second *sam* with a sharp movement of her left hand (33.84", a fraction of a second after VK), which she follows with a sequence of taps with her left index finger on the *mātrā* subdivisions (beats). The audience is not at all animated: we can observe a few approving shakes of the head in the *ālāp* and nods on each *sam* (more noticeable on the second than on the first), but the bigger movements seem to be positional adjustments.

Some of these features can be seen in more detail in Figure 2. Figure 2 includes, from top to bottom, (a) a pitch plot, calibrated in semitones relative to the Sa at 138.8 Hz, roughly C<sharp>;⁹ (b) an intensity plot, which clearly shows the alternation between his sung portions and his pauses or breaths; and (c) a

text grid. The text grid is based on one prepared within Praat, which includes an outline *sargam* transcription,¹⁰ transcription of the text or other syllables articulated by VK, and the main *tablā* beats (indicating both the time of the intensity peak of the *tablā*, and the beat to which it refers – X for *sam*, 2 for the second *mātrā*, + for the subdivisions of the *mātrā*). Into this I have interpolated a schematic representation of VK's hand movements ('Gesture'), with the vertical axis representing relative effort (i.e. the higher the line, the higher or further from his body are his hands); periods of rest – when both hands are in resting positions – are shaded. I have also interpolated a transcription of his eye movements, from which it can be seen that VK's eyes are either closed or looking down for the greater part of the clip, the exception being when he looks at VS at the start of the composition (shaded).

It can be seen from the gestural transcription that the only times when VK rests his hands are at either end of this extract: the Gesture Unit can be said to run from 1.6 to 38.9 secs (a duration of 37.3 secs). This unit can be broken down into two main Gesture Phrases, with the boundary falling at the transition from *ālāp* to composition. It is noteworthy here that, as VK's left hand assumes a resting position, his right hand immediately picks up in a dramatic gesture marking the start of the *mukhbrā* – thus, his hands mark both the transition and the continuity of the musical flow across this boundary. There is quite a close correlation at times between the raising of his hands and of the pitch of a melodic phrase, especially at the start of each Gesture Phrase: in the case of the start of the composition (28.9") the rapid rise and fall of the melody is matched almost exactly by the contour of his right hand movement. Within each Gesture Phrase there are periods (the horizontal lines) where he is moving his hands little or not at all, punctuated by more obvious movements: these periods when his movement pauses do not always correlate with either pauses or held notes in the sung line, although they can do so. Most significantly, he does not return to his rest position when he momentarily takes a breath (e.g. around 11 secs and around 20 secs): it seems that his gestures are indicating the continuity of melodic flow *across* the inevitable boundaries that occur when he has to take a breath (cf. Clayton, in press). This interpretation could also be glossed in terms of the relationship between two different intensity contours – of voice and hands – showing the former to be effectively nested within the latter.

<<Insert Figure 2 – covers two pages>>

In conclusion there are two main points to be observed in this clip. First, there is very little visual interaction at this stage: the singer is introverted, accompanists largely concentrated on the singer, and

the audience likewise – and although they recognise the *sam*, they are not yet ready to move with it. Secondly, VK's gestures mostly illustrate the melodic phrases – with left hand leading – and indeed clarify the grouping or phrase structure of the music. These Illustrators are accompanied by Markers, which have an instructional intent (guiding the *tablā* player and audience), and refer to the music's metrical structure.

In terms of gestural analysis this exemplifies a somewhat different temporal patterning of gesture to that found in most speech. In speech the normal pattern is described as preparation-stroke-retraction: positions can be held momentarily to extend a phase, or a Gesture Phrase may contain more than one stroke, but the basic pattern seems to be consistent. In VK's first Gesture Phrase here we see a short preparation (moving his hands away from their rest position), after which he holds his hands steady (pause), moves them sharply a short distance further out (stroke, marking the Ma-Pa move), pauses again, strokes again, and so on... followed by a fairly quick retraction as he returns his hands to the rest position. In other words, the gesture phrase is much longer than those usually encountered in speech, and marked by multiple strokes and significant pauses or prolongations.

The laterality or asymmetry in VK's gestures is also of interest: he appears to lead with his left hand most often when making Illustrator gestures, and with his right hand when performing Markers. While this one performance would be too little evidence on which to make sweeping generalisations, other factors suggest that this asymmetry is in fact highly significant. Trevarthen reports that in proto-conversations between infants and their carers, the infants display such an asymmetry from birth:

The evidence favours the conclusion that assertive or demonstrative activity concentrates in the left side of the brain, moving the right arm and hand, often at the same time as apprehensive self-regulatory withdrawal is more active on the right side of the brain, moving the left limb. (Trevarthen 1996)

McNeill and Pedelty argue along similar lines – on the basis of studies of speech-accompanying gesture in subjects with damage to the right brain hemisphere – that the left brain alone, working with the right hand, “produces a type of narrative in which there is linear form, but form deficient in imagistic content”; they continue, “When visuo-spatial input from the right hemisphere is lacking, narrative is incomplete: [displaying] lack of coherence, inability to match physical and abstract content, and inconsistent treatment of structural boundaries.” (1995, 83-84).

Interestingly, Rowell also notes a marked asymmetry in the early codified systems of *mudrās* in Indian performance traditions, although the relationship to asymmetries discussed up to this point is not obvious: “the function of the right hand is primarily tonal, and the function of the left hand, temporal, in that the left hand is used more for counting and indicating special durations” (1992, 66). VK’s use of gesture seems to show a greater coherence with the distinctions displayed by neonates, or by adults in normal conversation, than it does with early Indian codified systems: in other words, the asymmetry supports his (and other *khyāl* singers’) contention that their gestures in performance are largely spontaneous and are not codified.

Clip 2: *Sargam tāns* (18:34-19:36)

This clip was selected as an example of visible interaction between singer and *tablā* player: of several instances in which the two exchange glances and gestures of approval this was one of the more dramatic. It is also an episode in which VK’s attention is focused to a great extent on his harmonium accompanist SS: since he is also singing *sargam* syllables (the second such episode), we might deduce that he is focusing on SS’s imitation of his melody more explicitly than elsewhere. The clip begins with the *mukhrā* of the *vilambit* composition, ‘Gokula gā(va)’: his attention is on SS, then he swings round slightly to face the audience as he marks *sam* with an upward flick of his left hand (3”). He rests for about 10 secs, hands in lap, while SS plays a few phrases: when she reaches a sustained Ni (^7), he begins again (14”) – this move on her part feels like a clear (aural) invitation to him to recommence. He gestures with both hands in a balancing motion (one moves up as the other moves down), moving without a pause into the *mukhrā* and marking *sam* (50”) with a downward slap of his right hand and a sharp upward movement of his left hand. His head has been turned towards SS for most of this passage, occasionally glancing at both her and the audience. He immediately looks around at VS and nods in approval, appearing to mouth words of approval.

SS marks the first *sam* with a flick of her left hand. When after a short harmonium solo VK begins to sing again, she focuses her gaze on him, moving her head and shoulders with him as she follows his improvisation. She catches brief glances at VS (25”) and at the audience (35”), making eye contact as if in approval of VK’s performance, but then returns her gaze to the singer, marking the second *sam* once again with a flick of the fingers of her left hand. VS continues his role of engaging with other participants

in turn, with more animation than in the earlier clip. He nods his head vigorously on the first *sam*, then nods repeatedly as the second *mukhrā* approaches – again as if in approval – before marking the second *sam* most emphatically, smiling and nodding towards the audience so that his whole upper body leans far over to his right (towards the audience).

Several audience members mark the first *sam* and then relax; their nods of approval become more noticeable, then a few mark the start of the *mukhrā* and the final *sam*. A couple of factors become more noticeable in this Clip: first, that those who mark the *sam* gesturally prepare these movements well in advance, raising their hands in time with the preceding beat; second, that Veena Sahasrabuddhe – not coincidentally, the senior musician present – is the most active audience member, and others occasionally glance at her as if to take a lead from her.

What becomes clearer in this Clip is the nature of the interaction between participants, and the different roles they assume. VK is obviously the musical leader, but he maintains his introverted stance: he clearly focuses his attention on SS for much of his *sargam* singing, attending to her accompaniment, and exchanges glances with both SS and VS before and after *sam*, but only shoots quick glances at the audience (probably focusing on Veena Sahasrabuddhe, who as the senior musician is the most animated listener, assuming the role of ‘lead auditor’). SS has to concentrate mainly on VK in order to accompany him, but when she feels able she shoots glances at audience or at VS. VS takes on the most dynamic role, fixing his fellow performers as well as audience members in his gaze, and obviously showing his appreciation: this is clearly consistent with his own comments on his role, quoted above. The audience, led by Veena Sahasrabuddhe, show their approval periodically with subtle head gestures, but their most notable movement is at the *sam*. Here it is noteworthy that listeners start to prepare for their *sam*-marking gestures more than a second before *sam* itself: this is a demonstration of the fact that they are following the music closely and sharing the sense of release that the *mukhrā* signals and the *sam* confirms. The flow from *sam* to *sam* has established a clear attentional rhythm with a period equal to that of the cycle (c. 47 secs). The significance of this is that one would not normally observe such regular attentional rhythms of this duration in other kinds of interaction (such as conversation): musical performance is an efficient means of coordinating different individuals’ attentional rhythms, and slow *khyāl* appears to be particularly effective in establishing shared attentional periods of over 40 seconds.

Clip 3: *ākār tāns* (36:01-36:49)

Following an acceleration at 32' the energy levels of both performers and audience increase markedly: this clip was selected as an example in which VK's gestures are particularly animated, in which he appears to make eye contact with audience members (albeit briefly), and in which several audience members interject vocal expressions of approval.

This excerpt begins in a similar fashion to Clip 2, with VK focusing on SS as he reaches *sam* (3'') with a slap of his right hand and a flick of his left hand. He then sways away from SS and makes a brief shake of the head, clearly indicating to her that she should play solo for a while (Emblem). This time he does not wait for her invitation to restart, however, as he overlaps her improvisation with his own (11''). He first sings four rapid bursts of *ākār tāns* (rapid passages sung to 'aah'), with SS following closely on the harmonium: remarkably, he does this with no noticeable hand movement whatever, his hands resting on his lap. After the second of these bursts, the audience's (and SS's) calls of approval are audible (23''). The next *ākār* is accompanied by hand gestures, his hands rising together upward and rightward, the right hand then taking the lead with his left hand remaining a few inches above his lap; he then swings both hands to the left, and reverts to his normal '*sam*' position, right hand slapping his thigh and left hand flicking upwards (49'').

SS nods with VK on the first *sam*: her attention is then mostly on the singer, apart from a brief glance at the audience. As the second *sam* approaches, perhaps influenced by the high energy level of the audience, she turns to her left and looks toward the audience as she marks *sam* with a downward nod. VS once again seems to be interacting with fellow musicians and audience in turn, making exaggerated body movements towards *sam*.

Apart from the vocal interjection noted above, there is considerable audience involvement, both responsive and participatory. Not only the two *sams*, but also *mātrās* 8 (23''), after which the interjections are heard), 11 and 12 are marked by body sways and hand claps, before the second *sam* (37'').

The most notable feature of this clip is the response of accompanists and audience at 23''. The contrast here from the previous example is that this is a *response*, rather than an act of *participation*: where the gestures marking *sam* are clearly prepared well in advance, the motion of the responsive gestures begins only after the *tān* has been concluded. There are several approving glances during the preceding *tān* that indicate a shared appreciation of VK's improvisation and suggest that such a response was being

prepared, in some sense – the difference is that since no-one but VK could have predicted with any certainty that the *tan* would end on *mātrā* 8, they were unable to prepare the *timing* of their appreciative responses by focusing their attention at this moment, and thus the gestures of appreciation indicate a different focus of attention.

3. Discussion

The preceding section has described two phases of analysis: (1) a description of the gestural behaviour of all participants, based on viewings of the whole performance as a picture-in-picture image; and (2) detailed analyses of three clips. In the following discussion I will pick up on a few features of the performance that seem to be particularly noteworthy.

- i. The performance clearly must be understood as a multi-modal experience: although VK himself spends much of the time with eyes closed or averted, eye contact and bodily orientation do play an important role in the performance dynamics, for instance when the singer twists his body to the left when concentrating on SS's harmonium accompaniment, or when VS looks to the audience to share the experience of the *sam*.
- ii. Very little of the interaction between participants can be described as emblematic (in the gesture studies sense of “standing for verbal utterances”): in fact, even that which does occur, such as VK's signal that he is about to begin his *bandiś*, is almost certainly redundant. The more significant gestural functions seem to be the Illustrators (almost exclusively VK, as he moves his hands and upper body with his singing in a variety of ways); and the beat Markers (especially but not only on *sam*).
- iii. The last category takes on a particular importance: by marking the significant beats together, participants share these moments, and affirm that they share the experience of time and motion generated by the musical and gestural acts (something that VK confirms is significant for him as a performer). In this I include both performers and audience: there is no difference in principle between the way the musicians and the listeners mark the *sam*, and in this sense the audience clearly *participate in* the performance rather than *respond to* it.¹¹

- iv. Audience gestures of approval and vocal interjections are, in contrast, responses to the music, rather than the result of shared periodic attention: they can begin only after the event being responded to, and are therefore constrained by the time it takes to prepare and effect the response.
- v. The overall trend of the performance describes a linear intensification, matched by an increase in the energy expended by performers. This is matched by the audience, who display a steady increase in attention, involvement and physical movement. In the latter case, however, the audience's energy peaks around the end of the *vilambit khyāl* and relaxes after the early stages of the *drut* (fast) composition.
- vi. The logic of *tāl* and *bandīs* describes a periodic ebb and flow of attentional energy, rising through the latter part of the cycle to a peak on *sam* and then rapidly relaxing. In fact, the *mukhbrā* is also important here: once the improvised passage has flowed into this opening phrase there is an initial release of tension as the listeners appreciate how the *āmad* (approach) has been effected: the flow of the *mukhbrā* into *sam* is a familiar refrain and merely confirms the relaxation.
- vii. The gradual increase in attentional energy through the cycle is nurtured by the musicians' performance: we might suppose that the increase in the audience's attentional energy is a function not simply of the expenditure of energy by the performers, but of their producing the cues necessary to evoke a response. This interpretation is backed up by the evidence of the *drut khyāl* here, where an increase in speed and intensity is not matched by greater audience attention and involvement.
- viii. The audience's role seems to be one of continual feedback and affirmation, rather than input aimed at directing the course of the performance: by participating in the appropriate manner they encourage VK to maintain his concentration and affirm their own status as knowledgeable listeners. Therefore, in this case at least, it appears to be more productive to see the performance as an event constituted by all its participants, rather than to see the audience as the 'context' for the musicians' performance.

- ix. Vijay Koparkar's use of gesture displays a marked laterality, with the left hand dominating in what I have termed *Illustrator* gestures, the right hand leading in *Markers* of beat and *tāl*.

This is consistent with his description of his use of gesture as 'natural', i.e. not codified.

These observations tell us something about the experience of different time levels in the performance of the *vilambit khyāl*. The *tāl* cycle extends to 35 or 48 secs, divided into *mātrās* of 2.9 to 4 secs, which are further subdivided by the *tablā thekā* into 4 beats each lasting a second or less. The *mātrā* functions as a metrical unit – that is, each *mātrā* comprises one measure of a 4-beat metre. The *tāl* cycle functions as a regular hypermetric unit of 12 *mātrās*. This higher level is particularly interesting, for it appears to lie beyond the time scale within which metre can be modelled as neurological entrainment. It coincides with the level described by Trevarthen as a 'narrative' level – comprising durations within which an episode of a story can be presented and understood – which in speech would not usually be as clearly periodic as it is in this performance. The structure of the *thekā*, composition and other aspects of the performance support the kind of long-term attentional periodicity that allows performers and listeners to focus some of their attentional resources on this higher-level structure, to maintain its periodicity and to concentrate on a specific point in the period (the *sam*). This time scale has parallels in other musical repertoires, of course – the unusual feature here lies not in the time scale per se but in the participants' attention on the process of *āmad* and the release facilitated by the *mukhbrā*. It could be argued that one of the functions of the event itself is precisely to facilitate this sharing of attention and of temporal expectations within an intimate group.

The study has presented a series of observations on gesture, interaction, attention and temporal structure in a *khyāl* performance by Vijay Koparkar, considering both musicians and audience. I have demonstrated that given suitable audio-visual recordings and analytical methods, much can be deduced about all of these aspects of performance. Detailed observation of participants' movements demonstrates clearly, for instance, not only that listeners recognise the *sam* (beat one), but that they prepare their gestures marking this event well in advance: a significant part of their attentional resources is spent on tracking the long time cycles of the slow *ektāl*. The same observations indicate a clear distinction between gestures showing – to performers and other listeners – that one is tracking these durations correctly, and gestures and exclamations of approval, whose timing cannot be predicted with any certainty but whose occurrence indicates a different kind of attention, that which focuses on the

moment by moment events in the performance. Such exclamations are distributed unevenly through the performance, and there are too few in a single performance to generalise about when they are occur, except to say that this is likely to be when the singer has the full attention of the audience and is able to evoke an energetic response: in the case of this performance the audience's energy seems to peak towards the end of the *vilambit* phase, and this is where the approval gestures are concentrated. These are the main conclusions to be drawn with respect to my initial questions, although the analysis has also thrown up issues deserving further investigation, such as lateralism in singers' gestures, and the differing roles of *tablā* and harmonium accompanists in facilitating these socio-musical interactions. There can surely be no doubt that wherever such investigations lead, there is more for us to learn about Indian *rāg* music through the empirical study of its performance.

Figure 1. Picture-in-picture video image of Vijay Koparkar (VK, centre) with *tānpūrā* players Surashree Ulhas Joshi and Bageshree Vaze, Viswanath Shirodkar (VS, *tabla*), Seema Shirodkar (SS, harmonium), audience and author (back of centre top image). Main image Camera 4, small images from left to right Cameras 1, 2 and 3.



Figure 2. Vijay Koparkar, Rāg Multānī, Clip 1 (0-40 secs). From top to bottom: pitch plot (Praat); intensity Plot (Praat), schematic transcription of manual gesture (vertical scale representing effort/ distance from rest position); *sargam* transcription; text; *tal* (beats determined from *tablā* strokes); and gaze.

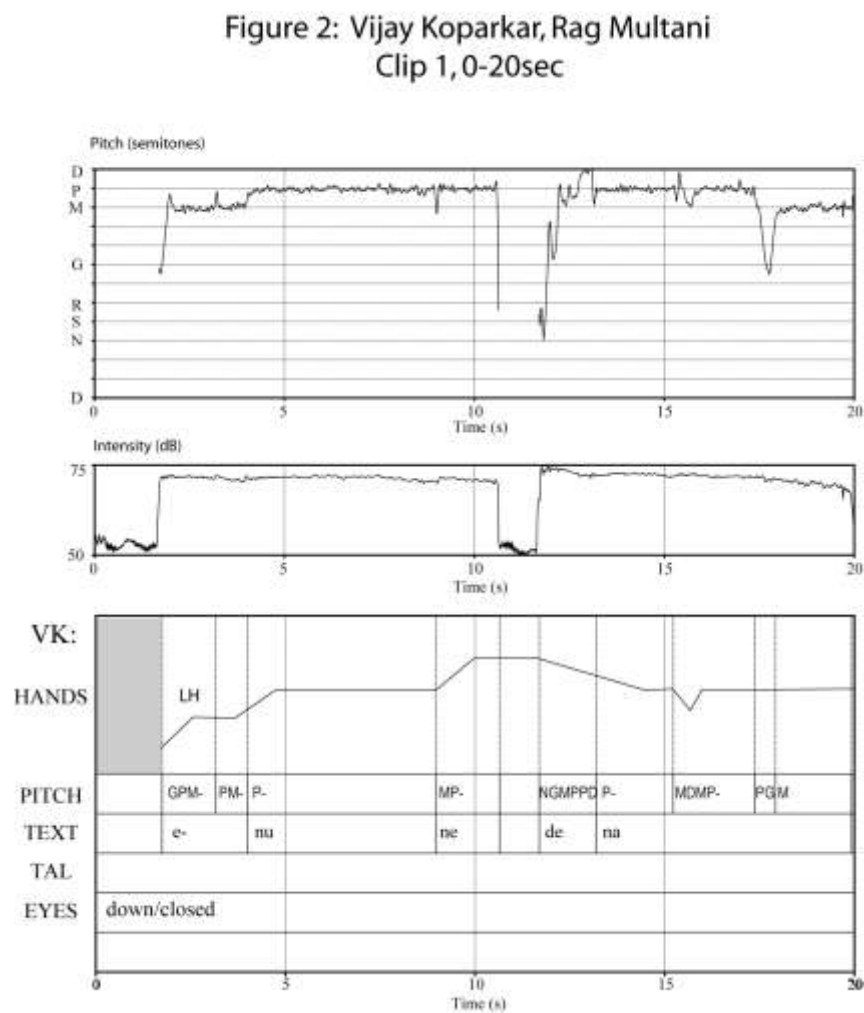
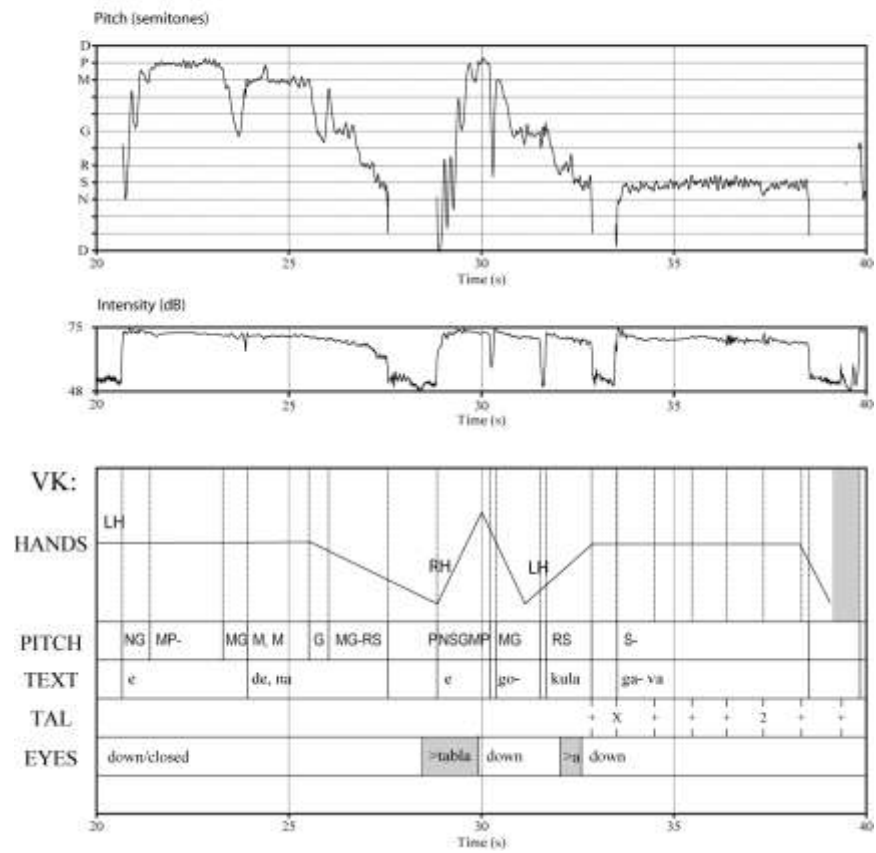


Figure 2 (ctd): Vijay Koparkar, Rag Multani
Clip 1, 20-40sec



References

- Baily, John
1985 Music structure and human movement. In *Musical structure and cognition* (Ed. P. Howell, I. Cross and R. West), pp. 237-258. London, Academic Press.
- Clayton, Martin
2000 *Time in Indian music: Rhythm, metre and form in North Indian rāg performance*. Oxford, Oxford University Press
2005 Communication in Indian raga performance. In *Musical Communication*, ed. Dorothy Miell, Raymond MacDonald and David Hargreaves. Oxford: Oxford University Press. (pp.361-381)
2007. Observing entrainment in Indian music performance: video-based observational analysis of tanpura playing and beat marking. *Musicae Scientiae*, 11(1).
in press. Toward an ethnomusicology of sound experience. In *The New (Ethno)musicologies*, ed. Henry Stobart. Scarecrow Press.
- Clayton, Martin., Rebecca Sager and Udo Will
2005 In time with the music: The concept of entrainment and its significance for ethnomusicology. *European Meetings in Ethnomusicology*, 11 (ESEM CounterPoint 1), 3-75. A pre-print version (dated 2004 with different pagination) is available from www.open.ac.uk/arts/music/mclayton.htm.
- Clayton, Martin and Veena Sahasrabudhe.
1998. *Khyāl: Classical singing of North India*. OU Worldwide ETHNO/VC01. VHS video, 45 mins, with accompanying booklet. Milton Keynes: Open University. The booklet is available from www.open.ac.uk/arts/music/mclayton.htm.
- Davidson, Jane W.
1993 Visual perception of performance manner in the movements of solo musicians. *Psychology of Music*, 21/2, 103-113.
2001 The role of the body in the production and perception of solo vocal performance: A case study of Annie Lennox. *Musicae Scientiae*, 5(2), 235-256.
2005 Bodily communication in musical performance. In *Musical Communication*, ed. Dorothy Miell, Raymond MacDonald and David Hargreaves. Oxford: Oxford University Press. (pp.215-237).
- Drake, Carolyn, Mari Riess Jones, and Clarisse Baruch
2000 The development of rhythmic attending in auditory sequences: attunement, referent period, focal attending. *Cognition* 77: 251-288.
- Goffman, Erving
1969 [1959] *The presentation of self in everyday life*. London: Allen Lane The Penguin Press.
- Kendon, Adam
2004 *Gesture. Visible action as utterance*. Cambridge: Cambridge University Press.
- Large, Edward W., and Mari Riess Jones
1999 The dynamics of attending: How people track time-varying events. *Psychological Review*, 106(1), 41.
- London, Justin
2004 *Hearing in time: Psychological aspects of musical meter*. New York and Oxford: Oxford University Press.
- McNeill, David
1992 *Hand and mind. What gestures reveal about thought*. Chicago, University of Chicago Press.
- McNeill, David and Laura L. Pedelty,
1995 Right brain and gesture, in Karen Emmorey and Judy S. Reilly, eds, *Language, Gesture and Space*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Naimpalli, Sadanand
2005 *Theory and practice of tabla*. Mumbai: Popular Prakashan.
- Qureshi, Regula B.
1995 [1986] *Sufi music of India and Pakistan*. Chicago, University of Chicago Press.
- Racy, Ali Jihad.
2003 *Making music in the Arab world: the culture and artistry of tarab*. Cambridge & New York: Cambridge University Press.

- Rimé, Bernard and Loris Schiaratura
 1991 Gesture and speech, In *Fundamentals of nonverbal behavior*, ed. R. S. Feldman and B. Rimé, pp. 239-281. Cambridge: Cambridge University Press.
- Rowell, Lewis
 1992 *Music and musical thought in early India*. Chicago: University of Chicago Press.
- Slawek, Stephen
 1990 Kaku-bhed, Raga-rasa, interpretive moves and musical intention: Parameters of musical meaning in North Indian music. Paper presented at Society for Ethnomusicology, Oakland, CA.
- Trevarthen, Colwyn
 1996 Lateral asymmetries in infancy: Implications for the development of the hemispheres. *Neuroscience and Biobehavioral Reviews*, 20 (4): 571-586.
- Williamon, Aaron and Davidson, Jane W.
 2002 Exploring co-performer communication. *Musicae Scientiae*, 6(1), 53-72.

¹For Praat see www.praat.org. Pitch plots were made using a procedure outlined by Wim van der Meer in his “Praat Manual for Musicologists” (www.musicology.nl > research, accessed 27 Sept 2005).

² The event was organised jointly by myself and Veena and Hari Sahasrabuddhe. It was recorded by myself, Laura Leante and Jaime Jones. The research trip was funded by Arts and Humanities Research Council grant no. 19110. I am grateful to Hari Sahasrabuddhe, Richard Widdess and Laura Leante for their comments on a draft of this paper.

³ This was my first experiment in using progressive scan video. Despite its advantages, the reduction of frame rate from 25 to 15 fps did result in some loss of temporal resolution in the analysis.

⁴ Since he wrote the words out for us before the concert, I know that he had prepared three, the third in fast *adācaṭal*, but in the event he only presented these two. Thanks to Hari Sahasrabuddhe for his help with the translations.

⁵ Although in most cases the timings logged from different views match within one frame (40 msec), in a few cases the difference is significantly greater (up to 360 msec). This is due to the limited depth of field of video recording: it is difficult to pick up movement towards and away from the camera. For this reason comprehensive gestural transcriptions should ideally be compiled from more than one view, although of course this is subject to practical limitations.

⁶ Since each *mātrā* is subdivided into 4 beats, the tempo in ranges from 60-80 bpm rather than 15-20 bpm.

⁷ “Vāh, kyā bāt hai” – literally “Wow, what a thing it is!” – is a conventional expression of approval in Hindi/Urdu, used also by those whose mother tongue is other than Hindi or Urdu.

⁸ Each of these video extracts can be viewed online: go to www.open.ac.uk/arts/experience.

⁹ This measurement was actually taken from the harmonium pitch track, where the Sa is a more stable pitch.

¹⁰ The letters used are as follows: S = Sa (¹), R = Re (², in this case flat), G = Ga (³, also flat), M = Ma (⁴, in this case sharpened), P = Pa (⁵), D = Dha (⁶, also flat) and N = Ni (⁷).

¹¹ In fact it is not unknown for members of intimate *baithak* audiences to join in with the singing of the *mukhra* (Hari Sahasrabuddhe, pers. comm.).