TOWARDS A VOCAL TRACT GEOMETRY-BASED PEDAGOGY OF SINGING

Melkote Krishnarao SHANKAR

National School of Drama and Centre for Public Understanding of Science, Bangalore, India
e-mail: shankarmelkote@gmail.com

Abstract
Recent music pedagogical research has begun to look at ways to apply emerging knowledge on the vocal tract for singing. In focus are aspects of articulation, measures of precision and accuracy, and the role of imagery in pitch production. This study proposes vocal tract geometry based heuristic schema for singing musical pitches, employing the idea of imagery.
The study is the result of the author's practice-based independent research. It has been found to work in the classroom as a practical pedagogic device and is amenable to public scrutiny and testing. This also constitutes an attempt at finding a theoretical context for the schema, a place within the overall discourse of music pedagogy.
Key words: *articulatory mechanism, vocal tract geometry, imagery in pitch production, music pedagogy, theoretical context*

Introduction

Traditional pedagogy of singing is based on imitation. Cox says: "*Humans understand other entities…and events in their environment in part via overt and covert imitation. We imitate gestures, facial expressions, postures, gaits, vocalizations, and other behaviors of those around us. This is part of how we learn to be human…*" (Cox, 2011, 4).

Imitation is associated with a neurobiologically based capacity for empathy and sociality (Iacoboni, 2009). This capacity is at the basis of all learning. Unlike as in visible actions - such as in acting, for instance - in singing, the series of stimuli is invisible and the imitator depends on hearing. The individual hearer does not 'see' how the sounds are shaped inside the vocal tract. Classical phonetics offers a static map of positions of articulation of consonants and vowels. In recent times there have been attempts to understand the aerodynamics of phonation (Khosla, Murgugappan, Paniello & Ying, 2009) and suggestions that the use of imagery may facilitate production of the right pitch. The current study proposes visual imagery that can be a useful pedagogical tool in teaching and learning singing. It is imagery that has evolved in the specific context of Indian Hindustani classical music.

One definition of singing describes it as consisting of complex multimodal multidimensional behavior that requires the integration of multiple motor, perceptual, and cognitive functions (Pfordresher, Halpern & Greenspon, 2015). This report will first propose the core findings of the study and then seek to locate a context for them. According to classical phonetics, an utterance begins with an abstract plan (Tatham & Morton, 2011). The plan is intentional and governed by a cultural context as well as laws regulating neuromuscular coordination. The utterance is executed with the help of mobile articulators - tongue, lips, uvula, larynx and the lower jaw. In practice, the resulting acoustic signal never quite matches the plan because of mechanical, acoustic and other artefacts introduced in the production process. This applies to the singing sounds as well. In translation of the utterance plan, the ideal mapping may elude the practitioner. This is where problems of teaching and learning may occur (Tatham & Morton, 2011). This mismatch between theory and reality constitutes a pedagogical gap.

Setting

A vocal utterance is the culmination of a series of physiological steps. Emerging from the lungs, air presses against the vocal folds of the larynx and the glottis. The sound wave - acoustic signal - issuing at the laryngeal cusp is further processed by the articulatory apparatus of the vocal tract to produce speech or song. Singing requires sustained vowels at the subglottal, glottal and supraglottal levels. It is the sustained vowel that the present study is interested in with the focus on the elongated standard English vowel /ɑ/.

Singing is a generative system that makes infinite use of finite media, creating something new, something that is not present per se in any of the constituents. In singing - as in speech - discrete parts are combined in a variety of ways to make new, expressive wholes. Merker (2002) proposed that humans do not exploit the continua of frequency (pitch) of their voice or of instruments – and of time, but create limited sets of discrete categories and combine them into complex patterns.

Curiously enough, the Indian musical aesthetics is inspired by the kind of continua that Merker is referring to. In the current study since we are going to focus on the steady vowel /a/, we shall not be speaking of voice onsets, pauses, terminations and other discretized aspects of pitch and duration. In the background of our concern is the recognition that mechanical properties of oral and pharyngeal mucosa and muscle tissues, among others, influence the transfer of the acoustical energy (Mainka et al, 2015); that specific control of these structures serves to tune the timbre characteristics of vowel sounds and; that the features change rapidly during speech and singing (Sundberg, 2019).

Of relevance to our study is the suggestion that production of pitch in singing is dependent on the appropriate sensory-motor translation of the utterance plan and a clear imagination of the expected conventions or rules. Singing, however, involves more than pitch: voice quality, intensity, and timbre. But pitch accuracy may be considered a first priority and its attainment is usually willed and planned. Intensity, too, can be willed and planned. Timbral quality is more complex business and depends on the default characteristics of the vocal tract and harmonics in the higher frequencies. A close study of steady state vowel spectrograms might help provide leads to the attainment of particular timbral quality.

Again, pitch accuracy is not always a requirement for reception of a song. Cultures tend to be willing to ignore minor departures from expected pitch executions and may be swayed by the emotional quality, timbre, and verbal import of the lyric. Even so, first of all, how well the singer translates the image of the plan is a function of how well she deploys the muscular conformations necessary for articulation of the desired pitch. The link between pitch image and articulatory action has been indicated as a contributing factor in inaccurate singing (Pfordresher, 2010).

The Multi Modal Imagery Association (MMIA) Model proposed by Pfordresher, Halpern & Greenspon (2015) on the sensorimotor translation of the utterance plan in general and imagery in particular is based on three sets of data: data suggesting a link between auditory imagery and singing accuracy; evidence for a link between imagery and the functioning of internal models for sensorimotor associations; and the use of imagery in singing pedagogy.

Articulation depends largely on the deployments of the tongue, the lips and the larynx. The tongue, through appropriate positioning, serves to increase or decrease resonating spaces and related timbres and is potentially capable of producing an infinite number of vowel sounds. It is recognized that shifting relative positions of tongue tip within the vocal tract, up, down, dorsal and back, can produce a continuum of vowel sounds to build a system of phonemes. These, with additional combinatorial rules, will allow construction of language specific syllables (combinations of vowels and consonants). The mandible, too, widened or collapsed to a default position, has the effect of changing the shape and volume of the vocal tract regions that affect the amplitude, pitch and timbre of the voice. In this manner, the mobile components of the vocal tract make it a variable resonator, with changes in their positioning causing changes in formant frequencies.

Further, in view of the fact that the frequency of the vibration of glottis, the source function - and the harmonics produced in the vocal tract - the filter - are largely independent of each other, by fixing the shape of the vocal tract it is possible to produce the same vowel repeatedly with widely different fundamental frequencies and harmonics. Or, keeping the source frequency - F_0- constant, it is possible to alter the shape of the vocal tract independently to change the vowel that is produced (Raphael, Borden & Harris, 2011).

This observation allows holding the shape of the vocal tract fixed in a particular conformation to produce a sustained vowel /ɑ/. Since the shape of the conformation plays a critical role in the production of the vowel, it may be fruitful to consider geometry of the vocal tract for articulating the same vowel as relevant to singing. Geometry is proposed in which the vocal tract is divided into upper and lower 'palate'. 'Upper palate' is the physical space extending from the upper pharyngeal region, the velum or soft palate, the hard palate, alveolar ridge, upper teeth and the region bound by the lips. The 'lower palate' is the region symmetric with the 'upper palate', and comprising the lower and inner part of the mouth, space under the tongue, and bound by the mandible. For all their apparent physicality, the two 'palates' are putative, virtual geometric spaces, conjuring up imagery of the vocal tract as a cavity resonator made up of two mutually interlocking hemispheres.

Another caveat at this point is in order. Practitioners of art have to contend with the subjective in judging the quality of art - the truth of art - as mediated by culture and tradition. In particular reference to music, its aesthetic is mediated by the idea, and perception, among others, of sensory consonance. The idea of consonance is mediated by culture so that what is 'consonant' to one tradition might not be consonant to another. Consonance is used here as tied up with the enjoyment that comes with listening (Zatorre, 2015). Cultural preferences are seen as modulating the perception and reception of music (Huron, 1994). It is in this backdrop that the pedagogic practices discussed and the schema proposed here must be viewed (see McDermott et al, 2016). The human species shares a universal neural disposition towards sensory consonance and comes with a default 'fine tuning' mechanism to allow for cultural preference. The fine tuning occurs through sympathetic oscillations in the sensory-motor regions; these are the neural modules that are also recognized as responsible for entrainment to rhythm (Large et al, 2015).

As will be seen, the Indian system allows for fair degree of deviations in the pitches. The idea of the sruti is an example of how this cultural tolerance is realised in musical practice.

A Note on Indian Musical Aesthetics

For the ancient Greeks proper sounds of music manifest the class of simple ratios by which the world itself is believed to be organized. This property - of simple ratios - allows the sounds to come into harmony with one another and with our organs of perception, which, too, are believed to share a natural affinity for the same ratios.

For the ancient Indians, reality was continuous and persistent as the background for all human actions and cognitions (recall Merker's observation above). It is through moments of singularity, of discontinuity, of ordinary experience, that the imperceptible, persistent continuum is to be inferred. In this scheme of things, musical reality is a continuum of sounds, and srutis are the resonant discontinuities that come to life by their sheer capacity to give pleasure. They become the svaras, or musical sounds. The loci of the svaras around which the srutis fructify are 12 in number - the same as in the western system of music. It is the positioning of these 12 notes, as actualizations of the srutis, however, that controversy has persisted. The srutis represent potential divisions of the pitch continuum scaling the octave into unspecified twenty two micro units and these are regions, intervals, frequency bands, rather than points or boundaries.

Previous exposure - literacy, if you wish to call it - plays a critical role in the acts of musical perception in the Indian context. An individual in this tradition is said to experience spontaneous perception (of musical sound) at the end of exposure to a series of individual sounds and syllables. The memory of qualities or attributes of experience acquired in a progressive and cumulative manner is the backdrop for and facilitator of recognition. The memory of these sonic qualities occurs within a gestalt called the raga. Operative within this raga gestalt is the functional concept of sonance of musical sounds or pitches. It may be defined as the degree of affinity or 'valence' between one note and another; it is the extent to which one pitch agrees or disagrees with another. In the idea of the raga, which is foundational to the classical musical system of India, a pitch or sound can enter into four kinds of musical possibilities with another pitch or sound. It can act as the sonant (vadi), consonant (samvadi), dissonant (vivadi), or assonant (anuvadi, that is, neutral or subordinate).

Analogous to the dominant note or pitch of the western musical system, the vadi is the voice; it 'speaks' for and on behalf of the other notes within the raga, entering into certain selected relationships with them. Once chosen, it determines the dynamic musical quality of a particular raga. The other musical notes, respectively, 'speak in concert with', 'speak in opposition to' and 'speak in a subordinate manner' with respect to the sonant or vadisvara. Several note pairs might be in consonant agreement, but only one note acts as the official samvadi, and consequently it cannot be dropped from the scale under any circumstances. It is the vadi and the samvadi in alignment - two strong and closely related svaras - that provide an axis of tonal stability and structural consonance (Rowell, 1990).

There are other rules governing the 'behaviour' of ragas. A raga may not have more than seven svaras, either in the ascendant or the descendant; it has a minimum of five svaras in the ascendant and descendant; it must always feature the tonic - represented as Sa; it must have at least one of the two notes - perfect fourth Ma or perfect fifth Pa as part of the gamut; it always has a sonant or vadisvara; the pitch distance between the vadi and the samvadi must be either 9 srutis or 13 srutis.

The sonant note is the single fixed point of reference from which all other tonal affinities are reckoned. When actualized in a particular scale formation or melodic construction, it becomes the 'denominator' (as in a fraction), and by extension, it is this pitch that determines the individual roles and limitations of the remaining pitches (there is a hint of mathematical relationship of tonalities here).

In performance, a raga is introduced by way of a handle/lever or 'pakad' constituted of the leading svaras, especially vadi, samvadi and anuvadi. The 'pakad' may be thought of as the core musico-acoustic signature of the raga. The nearest analogy of the 'pakad' to any entity within the western system of music is the dominant chord played out in arpeggio.

There are three possible ways in which one can visualize the relationship between the octave and the srutis:

a. Divide octave equally into 22 identical or roughly equal srutis;
b. Allow the srutis to remain unequal and unspecified in size, allowing the intervals to gravitate towards 'natural' intervals - those expressible in simple integral ratios: 3/2, 4/3, 9/8, 10/9 and so on, approximating the system of just intonation;
c. Come to terms with the prospect that the sruti intervals never achieve an exact size; that the srutis will be determined on the basis of oral instruction within a particular teacher-student line and will always remain rough approximations of the intervals mentioned.

The Indian system is closer, at least in concept, to just intonation than to the celebrated Pythagorean tuning, in which all the whole tones are of equal size: 9:8 = 204 cents.

Bharata's experiment around 2000 years ago with two vinas in which he demonstrated the existence of the twenty-two srutis is interpreted as argument for the system as conforming to just intonation. This is because the measure Bharata mentions - the so-called pramana-sruti, the standard measure of tuning - turns out to be the well-known syntonic comma (81:80 = 22 cents), the difference between the two whole tones of just intonation.

Musical experience in the Indian tradition has been summed up in the following passage: "*To put the entire question of tuning into perspective, it is useful to recall that musical experience in ancient India - music as heard, as taught in face to-face instruction, as learned, and as remembered - had accustomed practitioners to live in a world of subtle microtonal shadings which to them were more familiar, more tangible, and more real than anything they could hope to set down in the form of written knowledge. Modern Indian practice has evolved into a close approximation of twelve-tone equal temperament, from which seven svaras (or in certain cases, more or fewer) are ordained for each raga, but in which individual notes may be shaded up or down from this standard. Once learned and 'locked into place' at the beginning of a performance, each set of svaras becomes a collection of unique possibilities, and the difference between one scale step and another is understood, not as a calculated interval with an identity of its own, but as a matter of jumping or sliding from one fixed location to another*" (Rowell, 1990, 149).

Recent history of Indian classical music has witnessed a curious experience with attempt to introduce the idea of equal temperament through the harmonium. The harmonium features a hybrid between pure and tempered tuning. Working with Indian musicians in performance, Mark Levy has shown empirically that the intonation of the human voice is handled along similar lines: for each step in the scale there is a certain bandwidth within which the intonation of the voice still sounds correct. According to Levy's data the upper and lower limits of this bandwidth can be as much as 50 cent apart within one performance (see Abels, 2010, 124).

The combination of several notes woven into a composition in a way which is pleasing to the ear is called a raga. Each raga creates an atmosphere, which is associated with feelings and sentiments. A performer with sufficient training and knowledge alone can create the desired emotions. A raga is a subtle and aesthetic melodic form with its own peculiar ascending and descending movement consisting of either a full seven note octave, or a series of six or five notes (or a combination of any of these) in a rising or falling structure. It is the subtle difference in the order of notes, an omission of a dissonant note, an emphasis on a particular note, the slide from one note to another, and the use of microtones together with other subtleties that demarcate one raga from the other.

It may be useful at this stage to take a brief look at the thinking on music in other traditions. A certain degree of previous exposure has been seen as important to aesthetic reception and pleasure even in the western philosophical traditions. Hanslick (1957) saw the influence of aesthetic judgment as preceding feelings invoked by artistic expression. Meyer (1903) believed that repeated listening to a novel musical work tends to increase reported satisfaction. Lundin (1947) saw consonance and dissonance preferences as cultural in origin. In the cognitivist view, music evokes emotion only after passing through a cognitive/interpretative filter (Huron, 2016).

According to Huron, "*sometimes art is successful because it educates us, inspires us, challenges us, disturbs us, or even insults us. But if art never offered any element of pleasure, it would cease to play much role in human affairs*" (Huron, 2006, 367). Pleasure circuits in the brain were discovered accidentally half a century ago and more recently, peak musical experiences have been shown to produce dopamine release in brain regions implicated in many pleasure-related experiences (Salimpoor et al, 2011).

However, in order to achieve the microtonal aesthetics that the Indian musical system encourages, the practitioner would still need to be able to produce the twelve pitches in a manner that allows for the flexibility needed to negotiate the srutis, or inter-interval pitch (frequency) spaces. For a beginning, the schema to be presented here seeks to play the role of facilitator.

Vocal Tract Modeling

The sound emerging from the lips is the result of aerodynamic processes in the vocal tract. It is the result of vibrating column/s- and eddies - of air, and amenable to experience as such. The shaping of the sound in the vocal tract can be viewed from a geometric perspective. In modeling the acoustic modulation of speech, the vocal tract has been viewed as a concatenation of cylinders, of varying radii (Fant, 1970). The pharyngeal part has been viewed as cylinders piled one on top of another vertically. The remaining part of the vocal tract, through which air flows towards the lips, has been viewed as a series of horizontally telescoping cylinders. Air, thus, flows first along a vertical axis and then through a horizontal axis, along the space of the vocal tract with its modulating parts.

Another caveat is called for here. This geometry has to be viewed as virtual and as heuristic imagery to support a particular series of conformations of the vocal tract. In the same way as linguists seek to map the vocal tract through approximations to describe resultant speech sounds, an approximation is being sought to be made here, and to only model one vowel for a beginning. In fact, x ray views of the flows of air show up the complexity of modeling physical reality. So, the model is inspired more by linguistics than it is by physics.

In this assumed geometry that is conceptualized, the 'upper palate' and 'lower palate' introduced earlier share a common plane along their rims. The plane runs horizontally from the midline of the pharyngeal arch, just below the velum, to the midline just below the upper lip, both lines being central in an 'axiomatic' sense. During sustained vowel phonation, air travels from the glottis along this horizontal plane to and beyond the lips. It may be seen as approximating a columnar flow along a virtual straight line drawn between the midpoints of these two midlines. The straight line is at the intersection of the horizontal plane mentioned earlier and a vertical plane perpendicular to it, intersecting the upper and lower palates along a sagittal ridge, touching both virtual poles, upper and lower. Viewed from above, the planes divide the cavity into left and right halves and, viewed laterally, into upper and lower halves. The virtual straight line is chosen as x axis with its positive side oriented towards and beyond the lips.

Let us consider the production of the sustained vowel /ɑ/. In classical phonetics, it is produced by a tongue position that is lowered and pushed to the back. Of course, the same vowel has different actualizations in different languages, and as different speech events in the same speaker at different times. This manner of movement of the tongue causes the pharyngeal space to shorten and the oral cavity to enlarge. The shortened pharyngeal cavity resonates to higher frequency harmonics in the source, generating a relatively high-frequency first formant, whereas the large oral cavity resonates to low-frequency harmonics in the source and thus generates a relatively low frequency second formant. The relationship between the two formants - and how an entire range of vowels may be produced by shifts in their relative positions (using movements of the mandible, lips and larynx) has been worked out in detail by Sundberg (2019). In his work on the acoustics of the vocal tract, Fant (1970) approximates the passage to a series of cylindrical tubes of differing radii. Assuming a particular conformation of muscles for the production of the vowel /a/ sustained over a period, arguably, one can ignore the larger system complexity. Author and associate are currently studying spectrograms of the sustained and sung vowel /a/ to identify, if any, invariant acoustic features of the vowel. The geometry being considered here is intuitive and meant to serve a heuristic, mnemonic device to aid pedagogy of the sung pitches in the Indian musical context.

The next step in building the geometry is the creation of a shared aural framework between singer/speaker and listener. The framework must help singer and listener judge levels of consonance of the produced pitches. A steadily maintained pitch held outside of the vocal tract - as in a drone or played on the harmonium or piano - would be a good reference source. There are three functions for the drone in melodic music such as the Indian system. The drone provides a rich tonal background against which the musician places the notes of the raga being played or sung. It also provides the base for accurate intonation of tones. With the drone as the background of tonality, the musician is able to perceive, psycho-acoustically, the relationship among isolated tones and the relationships of their relationships. Both acoustic and neuromuscular mechanisms are involved in this process. The drone, externally representing the tonic, functions as a rest note. It is the note from which any melody must start and finish. While any departure from it causes neuromuscular unrest, return to the tonic resolves the unrest (Deva, 2011, 49).

The role of the drone may be played by any instrument as such but typically it is the four-stringed tanpura. The tanpura is a fretless musical instrument that is played traditionally for accompaniment in Hindustani music. At one end it has a large gourd resonator and a long voluminous neck with four or five metal strings supported at the lower end by a meticulously curved bridge made of bone or ivory. The strings are plucked one after the other in slow cycles of several seconds generating a buzzing drone sound. In modern day construction, the first string of the tanpura is tuned to the perfect fifth or fourth, depending on which is the dominant pitch in the raga being sung. The second and third are tuned to the mid tonic Sa and the fourth, to Sa an octave lower. The fourth string has thicker gauge to produce a lower pitch. Male vocalists pitch their tonic note to about C#; female singers usually a fifth higher. The male instrument has an open string length of approximately one metre, the female is three-fourths of the male. The standard tuning is 5-8-8-1 (sol do' do' do) or, in Indian sargam, P-S-S-S_* (S_* is the tonic one octave lower). For ragas that omit the fifth, the first string is tuned down to the natural fourth: 4-8-8-1 or M-S-S- S_*. Both in its musical function and how it works, the tanpura is unique in many ways. It does not partake in the melodic part of the music but it supports and sustains the melody by providing a colourful and dynamic harmonic resonance field based on one precise tone, the basic note or key-note, often referred to as mean-tone (Datta et al, 2017, 169).

Vocal Tract as Cavity Resonator

The idea of the schema was inspired by the possibility of imagining the vocal tract as topologically equivalent to a cavity resonator. In a cavity resonator, as in a laser, the light signal is amplified through multiple collisions - stimulated emissions - with the container wall. This insight will do for us. We need not go into details of light signal amplification.

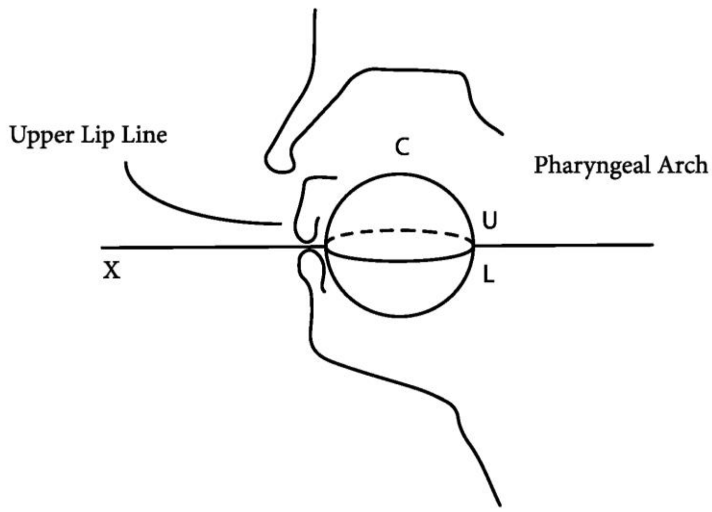

Figure 1. Vocal tract as cavity resonator

The vocal tract may be viewed as a cavity C. C is comprised of two hemispheres, upper hemisphere U and lower hemisphere L, interlocking at their rims and sharing the same plane. A horizontal straight line (x axis) marks two important midpoints on this plane, one at the back of the vocal tract, the other on the lip line of the upper lip.

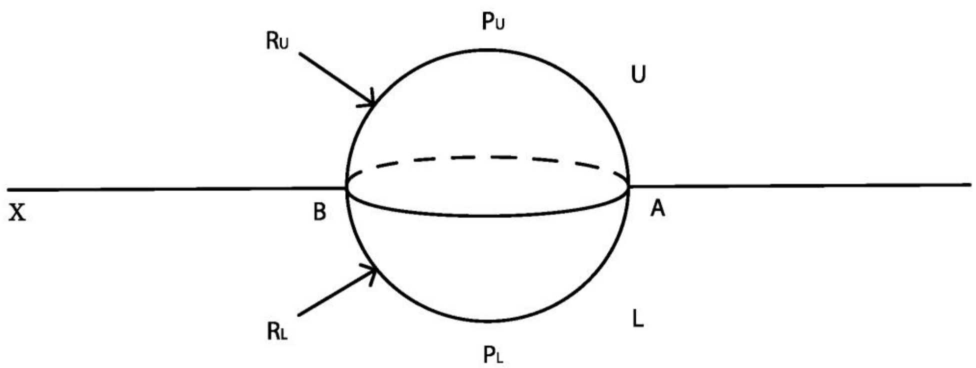

Figure 2. Proposed geometry

The x axis intersects the pharyngeal arch at its midpoint A and the lip line of the upper lip at its midpoint B. Two ridge lines, RU and RL run from A to B along the upper and the lower ridges - longitudes - of the respective hemispheres, as well as through the apices PU and PL.

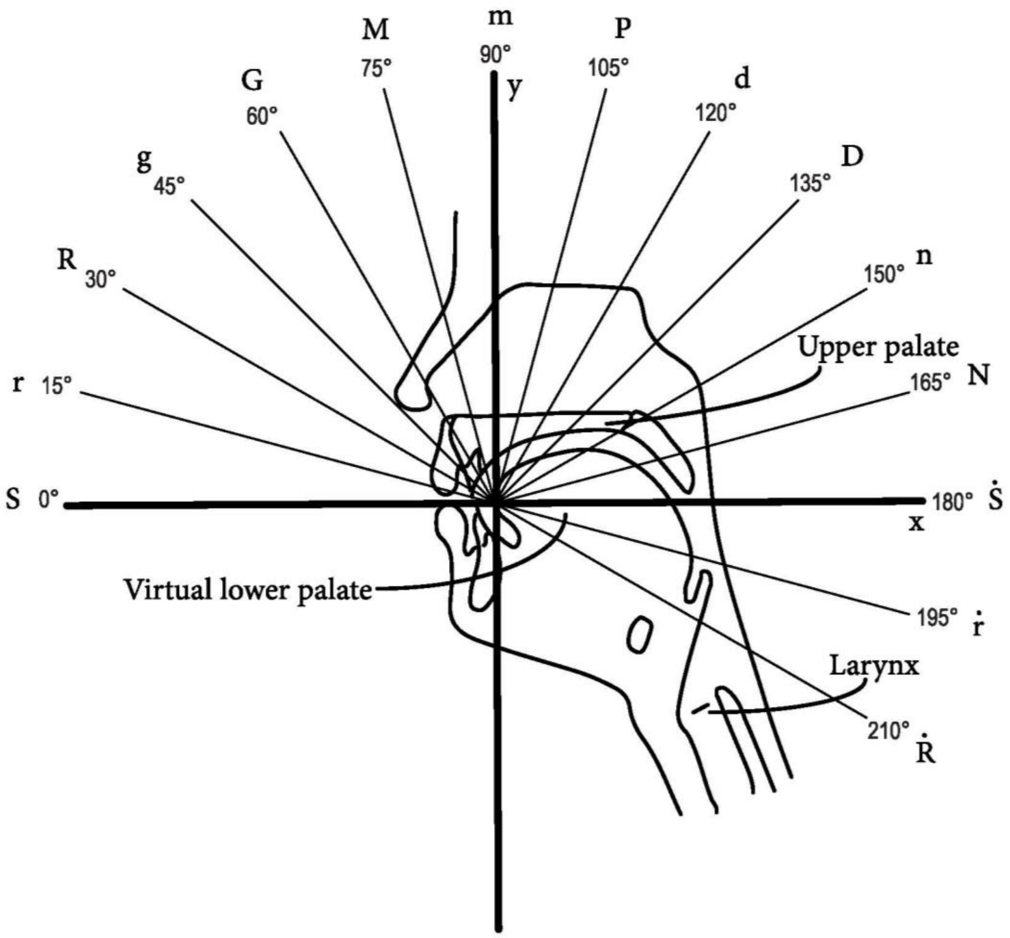

Figure 3: Notes by division

It is possible to produce the 12 musical pitches comprising an octave by aiming the post-pharyngeal, epiglottal air column - e.g. leading to the production of the sustained vowel /a/ - along points on a palatal geodesic that divides the vocal tract into arcs of 15 degrees each, starting with the horizontal axis as the reference note or tonic - S, and moving along the sagittal ridge, from zero degrees to 180 degrees, to attain the tonic in one octave higher - S*. The 12 musical pitches are, in the ascending order: S r R g G M m P d D n N S*.

In the author’s practice, the outside reference pitch was produced by an electronic version of the Indian tanpura. The vowel /a/ is repeated several times to ensure that the tone produced is in unison with the drone. The air flow that is responsible for the production of the tone is experienced as moving 'smoothly' from the glottis to, through and beyond the vocal tract, and out into the open medium where the acoustic signal is carried as pressure waves. An optimal level of air flow is a rate that is held steady for at least ten seconds, the steadiness being more important than the duration. The reference pitch is assumed to match the uttered vowel if the humming that is produced by stopping the air flow by a closure of the lips matches the drone pitch; if they are aligned, there is both an aural experience of the sonance/resonance and a subtle vibration palpated at the skin level. In the attainment of the open vowel /a/, the lips are parted and closed alternately to produce a humming tone and an /a/.

An electronic tanpura has a light emitting diode window displaying pitches according to the nomenclature used in the western system: A A# B C C# and so on. It has now become an ineluctable appendage to Indian classical musicians of both the Hindustani and the Karnataka styles of singing. The author chose C sharp as the reference pitch, the tonic. Teachers often use the alternating hum and the /a/ as part of warming up exercises. In the Indian musical context, singers do not use a score (except in the context of modern day film music where the orchestration is elaborate and one can clearly see the influence of western harmonic instrumentation).

In order to experience the production of the /Sa/, the canonical tonic with its fricative onset, author sought to keep the utterance of the fricative part /s/ as brief as was possible and quickly moved on to co-articulate the /ɑ/ vowel, the energy carrier. Previous to the use of the geometry of the schema, the singing of /Sa/ was gone through unreflectively. Awareness of the geometry allowed bringing the experience into conscious proprioception. This awareness is a critical transitional moment for pedagogy, from unreflective singing to reflective conscious singing. Of course, there are many questions that would need to be answered before one could definitely make a claim on behalf of the new vocal geometry-based singing. But the experience so far has been encouraging.

Once the /Sa/ is experienced, pitches that rise with respect to it - higher frequency pitches - are tested gradually, one step at a time, by bringing the vocal tract muscles to produce air flow at different angles. /Sa/ has been determined by us - experientially - as the result of air column striking the walls of the virtual cavity along the x axis, from the pharyngeal arch to the lips. Now by aiming the air column towards the pharyngeal end of the x axis we experience the production of pitch /Sa/ one octave higher (represented in Fig. 3 as S*). This is the tonic one octave higher; the air column now subtends an angle of 180 degrees with the x axis (see Fig. 3). This process seems abstract at the level of verbal description but empirical effort proves this to be accessible and much easier to comprehend. The term calibration may be used to ascribe and map tone values to angles at which air columns strike the oral cavity walls along the sagittal ridge, between angle zero - /Sa/-1 to angle 180 degrees - /Sa/-2 - of our geometry (S* of Fig. 3). This provides the framework for the production of pitches that lie in between. Although experienced subjectively, the production of pitches follows a set of images that is objectively given. A set of primitives or postulates will help us systematize the geometry to be meaningful. It is easily seen that the postulates are self-evident and collectively shared.

A Set of Postulates

1. The vocal tract - or mouth - may be viewed as enclosed space, a cavity (designated **C** in Figure 1). It is a cavity that resonates acoustically to vibrating air columns.

2. The space of the cavity can be viewed as being limited by boundaries. It is made up of an upper hemisphere (**U**) and a lower hemisphere (**L**), inverted over each other, interlocking at the rims and sharing a common plane (see Figure 1).
3. A horizontal straight line - x axis - may be drawn between a midpoint at the back of the vocal tract and a midpoint of the lip line of the upper lip. The midpoints comprise, one, the point at which the rim of the 'upper palate' meets the wall of the pharyngeal arch at its intersection with the x axis (**A**) and, two, the point constituting the midpoint of the lip line of the upper lip at the intersection with the x axis (**B**) (see Figure 2).
4. A ridge line, $\mathbf{R_U}$, runs from **A** to **B** forming one half of a sagittal longitude and passing through the pole, $\mathbf{P_U}$. A similar ridge line $\mathbf{R_L}$, runs from **A** to **B** in the lower half, complementing the longitude in the upper half, and passing through the pole, $\mathbf{P_L}$ (see Figure 2).
5. With the help of the appropriate muscular conformation and shaping of the vocal tract, the air column may be made to vibrate between diametrically opposite points in the two hemispheres along the sagittal ridges $\mathbf{R_U}$ and $\mathbf{R_L}$ striking the walls of the cavity **C** at diametrically opposite points. In the process, it is possible to produce S and S*, r and r*, R and R* (where the * refers to a note pitched one octave higher) in the corresponding hemisphere. This relationship is also seen to hold with respect to attainment of notes an octave lower; this involves minimally three subjective processes: perception of the pitch produced in comparison with the reference pitch, cognition of pitch identity between the two, and sensorimotor muscle articulation (see Figure 3).

These postulates will be the framework of ideas and images to build a schema for articulation of the musical notes in singing. It is a schema that has worked in practice and its pedagogical value has been established through trials with students.

The tongue plays a central role in navigating the air column. For purposes of the geometry building it is assumed to form the centre point of the air column. It is a virtual centre which the air column engages as fulcrum, as it shifts angle and negotiates points along the sagittal ridge of the 'palates' to produce different pitches.

Pitches are produced when air column is made to strike points on the walls of the oral cavity. In the geometry here conceptualized they strike points along the ridge - sagittal ridge - formed when a plane passing through the centre of the upper palate and the lower palate also intersects the straight line aligned with the x axis (see Figures 1 and 2). From previous discussion, we have the two /Sa/ pitches one octave apart, subtending zero and 180 degrees with the x axis. These have been designated as S and S*. In this geometry, the 12 pitches between these two ends of the octave will span the 180 degrees. Since we are dealing with a system that is tolerant of microtonal deviations, we are looking for a general region in which pitches occur so we are justified in dividing 180 by 12, to get a 15 degree difference in angle from pitch to pitch: as the air column rises, therefore, a change in pitch occurs every 15 degrees. In this manner are produced the 12 pitches, namely: /S/, /r/, /R/, /g/, /G/, /M/, /m/, /P/, /dh/, /Dh/, /n/, and /N/.

A note on pronunciation: the symbols represent pitches, each carrying a consonant head and vowel coda. So, /S/, /G/, /g/, /m/, /M/ and /P/ are to be pronounced with the vowel /a/ suffixed to them; /dh/ and/Dh/ are to be pronounced by stretching the vowel

/a/ just a little; /r/ and/R/ are pronounced with the vowel /e/ in suffix and /n/ and /N/ with the vowel /i/.The /S/ is led by the fricative /s/; the /R/ and /r/ are defined by the liquid /r/; /g/and /G/ are velar; /m/ and /M/ are nasal-bilabials; /P/ is a bilabial; /dh/ and /Dh/ are soft, voiced alveolar; and /n/and /N/ are nasal velar.

A note on symbolism: Where a letter carries a * or dot above its body, it is to be read as a pitch one octave higher; if it occurs below the body of the letter, it is to be read as a pitch in the lower octave. S, Sa and Sa-1 both refer to the same note; S* and Sa-2 refer to the tonic in the upper octave.

Although the octave of the Indian system is not laid out in the same way as the equal tempered octave of the western classical system, a rough and ready approximation, a one-to-one correspondence, might be set up between: C C# D D# E F F# G G# A A# and B, on the one hand, and /S/ /r/ /R/ /g/ /G/ /M/ /m/ /P /dh//Dh/ /n/ and /N/, on the other. There are significant differences between the two systems because of considerations historical and aesthetic-philosophical (see Huron, 1994).

It is possible to produce the 12 musical pitches comprising an octave by aiming the post-pharyngeal, epiglottal air column along points on a palatal map that divides the vocal tract into arcs of 15 degrees each, starting with the horizontal axis as the reference note or tonic and moving along the sagittal ridge, from zero degrees to 180 degrees (see Fig. 3).

Discussion and Conclusions

Is there a way to reconcile the phonetic chart and the geometry of the schema with respect to the 12 pitches?

Some phonologists see the production of consonant sounds as being initiated in one of three distinct zones of the roof of the vocal tract. The principal zones or places of origin are labial (lateral), coronal (or ventral) and dorsal (back of velar). A consonant is co-articulated with a vowel so that it derives from and moves toward another region. In the case of [g, k], for instance, the consonant is initiated in the dorsal (phonetically velar) zone, spreading thereafter to coronal or labial places (Wyn & Reimers, 2010).

Using this logic, and focusing on the second part of the sound - the vector of the vowel - we now seek to map the sounds of the 12 pitches. The sound value of */P/,* a bilabial initiated by the parting of the lips, spreads on the vowel /ɑ/in a vertical direction. The same bilabial, nasalized and articulated more softly, yields */M/*and */m/*. These sounds are articulated upwards into the roof or coronal zone. */P/* starts at the inner line of the lips and */M/* from a point just a little inwards and closer to the upper dental ridge. Both pitches are sustained in the mid region of the vocal tract, spanning - as experienced in practice - an angle of 45 degrees in the upper dental-alveolar-palatal swath. */M/* is sounded at 75 degrees and */m/* is sounded at 90 degrees; */P/* is sounded at 105 degrees.

The positioning of */g/* and */G/* is not very straightforward. Phonetically born in the back of the velar, in a swath between 225 and 240 degrees, they are co-articulated with /ɑ/ to sound in the labial-dental region closer to the upper lip. The two pitches span, respectively, an arc between 45 and 60 degrees, making allowance for the lowering of the mandible which causes the angles to rise along the arc.

The nasalized consonants /N/ and /n/ are sounded in the velum and nasal resonating region. But as they are co-articulated with vowel /i/, which is produced along an arc roughly between 15 and 30 degrees, below the lower lip, respectively, they produce two pitches that are experienced as being lower than the tonic /Sa/.If these two columns are projected backwards into the velar region, we account for the production of upper pitches /n/ and /N/ at 150 and 165 degrees, respectively.

Radiographs of the vocal tract of a speaker saying [i] have shown that the tongue mass, fills most of the oral cavity, leaving a small volume of air to vibrate in the space anterior to the constriction formed by the tongue. The pharynx, in contrast, enlarges because the posterior part of the tongue has been raised, that is, lifted out of the pharyngeal space. The oral cavity is larger and the pharyngeal cavity smaller for the vowel /ɑ/ than for /i/ (Raphael et al, 2011).

The size of the oral cavity for /ɑ/ may be increased in two ways: lowering the tongue passively by lowering the jaw or actively by depressing the tongue. It is also possible to combine these two strategies. Active or passive lowering of the tongue for /ɑ/ provides the large oral cavity and small pharyngeal cavity volumes that characterize this vowel (Raphael et al, 2011).

The pitches /dh/ and/Dh/ need a somewhat convoluted justification. In the classical chart, they are born in the alveolar region. In practice, they resonate to coronal harmony. They are produced when air column strikes the upper palate at angles 120 and 135 degrees, respectively.

By extending the arc of the angle to beyond 180 degrees, it is possible to produce members of the pitch sets /r/, /R/ and so on in the higher octave. In the other direction, in the arc of angle moving down from the lips into the lower palate, pitch sets /N/, /n/ and so on are obtained in the register one octave lower.

In the vocal tract it is air that moves to produce the sound output. Air would need to move steadily to justify use of the cavity resonator analogy. We assume that it would remain steady if the muscles that shape the articulatory part of the vocal tract are held steady. Imagining the movement of air column with a fixed point in the middle, invoked the image of a see-saw plank that was free to move in a circle, carving out an upper half and a lower half. The idea of the upper palate and the lower palate being navigated by such a see-saw like air column was tempting. It is also fitted in with the image of an old-fashioned wall clock with its second and minute arrows fused along a diameter.

As a pedagogical tool to learn the pitches, the schema provided an imagery that was easy to relate to. Students recognized its usefulness in attaining the desired pitches, both canonically and in the context of particular ragas. The canonical form represents something close to the twelve notes of the equal tempered scale. And with respect to the srutis, the schema facilitates imagination of a continuum of pitches which can be executed through mappable conformations of muscles, repeatedly and consistently.

The schema is testable. It extends the classical phonetic chart. It also extends the observation with regard to the placement within the articulatory space of the vowels /i/ and /e/ (Zavadska & Davidova, 2017), as well as the observation that the lower part of the pharynx is important for classical singing (Mainka et al., 2015). It illustrates the idea of a relationship between imagery and actualization in utterance. It is likely to help create long term muscle memory of the conformations needed for singing (see Pfordresher et al, 2015). It is available to the singer's proprioception within the vocal tract. It becomes part of the sensorimotor system such that imagery and translation are separated by the least temporal distance. It has implications for the shape and timbre of musical pitches. Familiarity with it offers a degree of control over the production of notes and the confidence that one will not go wrong. Its geometry envisages 12 seesaw-like diagonals crisscrossing the vocal tract. It is pitch invariant.

The aim of the training is to acquire internalized representations of the pitches to elaborate a raga or a set of ragas. A raga is like a language. And like language, one learns the basic vocabulary and the idioms before venturing narratives in it. This usually takes years of practice, partly because of the problems of enunciation, partly problems of cognition and partly imagination. The ability to perform raga elaboration is considered an attainment in melodic singing. The author sees value for the schema presented here as an aid to raga melodic practice and performance.

In summary, two quotations would serve as a suitable valediction:

- *"The marksman ends by thinking only of the exact position of the goal, the singer only of the perfect sound..."* (James, 1880, 774);
- The play of imagery and its actualization in performance is like *"... a duet between the music in your head and the music you are performing"* (Green & Gallwey, 1986, 75).

Acknowledgements

The author is pleased to acknowledge the following individuals: Dr T.S Channesh, Director, Centre for Public Understanding of Science, Bangalore, for critical inputs throughout the preparation of the manuscript; Nidhi Nambiar, Bangalore, for the diagrams and; Dr T.V Ananthapadmanabha, Bangalore, for opening up a new front in the investigation of vocal tract dynamics with respect to singing.

References

Abels, B. (2010). *The Harmonium in North Indian Music.* New Delhi: New Age Books.

Cox, A. (2011). Embodying music: Principles of the mimetic hypothesis. *Music Theory Online, 17(2),* 4-24.

Datta, A.K., Solanki, S.S., Sengupta, R., Chakraborty, S., Mahto, K. & Patranabis, A. (2017). *Signal Analysis of Hindustani Classical Music.* Singapore: Springer.

Deva, B.C. (2011). *The Music of India.* New Delhi: Munshiram Manoharlal Publishers Pvt. Ltd.

Fant, G. (1970). *The Acoustic Theory of Speech Production*. Mouton: The Hague.

Green, B. & Gallwey, W.T. (1986). *The Inner Game of Music.* New York: Doubleday.

Hanslick, E. (1957). *The Beautiful in Music.* Translated by Gustav Cohen. New York: The Liberal Arts Press.

Huron, D. (2006). *Sweet Anticipation: Music and the psychology of expectation.* Cambridge MA: The MIT Press.

Huron, D. (2015). Interval-class content in equally tempered pitch-class sets: Common scales exhibit optimum tonal consonance. *Music Perception, 11(3),* 289-305.

Huron, D. (2016). Aesthetics, In S.Hallam, I.Cross, & M.Thaut (Eds.), *The Oxford Handbook of Music Psychology* (2nd ed.) (pp. 233-245). Oxford University Press.

Iacoboni, M. (2009). Neurobiology of imitation. *Current Opinion in Neurobiology, 19(6),* 661-665.

James, W. (1880). *Principles of Psychology.* New York: Holt.

Khosla, S., Murgugappan, S., Paniello, R. & Ying, J. (2009). Role of vortices in voice production: Normal versus asymmetric tension. *Laryngoscope, 119(1),* 216-221.

Lundin, R.W. (1947). Toward a cultural theory of consonance. *Journal of Psychology, 23,* 45-49.

Mainka, A., Poznyakovsky, A., Platzek, I., Fleischer, M., Sundberg, J. & Mürbe, D. (2015). Lower vocal tract morphologic adjustments are relevant for voice timbre in singing. *PLOS ONE, 10(7)*, July 17. Retrieved from https://doi.org/10.1371/journal.pone.0132241

McDermott, J., H.Schultz, A., F.Undurraga, E.A. & Godoy, R.A. (2016). Indifference to dissonance in native Amazonians reveals cultural variation in music perception. *Nature*, *535(7613),* 547-550.

Merker, B. (2002). Music: The missing Humboldt system. *Musicae Scientiae, 6,* 3-21.

Meyer, M. (1903). Experimental studies in the psychology of music. *American Journal of Psychology, 14,* 456-475.

Pfordresher, P.Q. (2010). Imprecise singing is widespread. *The Journal of the Acoustical Society of America*, *128(4).* Retrieved from https://doi.org/10.1121/1.3478782

Pfordresher, P.Q., Halpern, A.H. & Greenspon, E.B. (2015). A mechanism for sensorimotor translation in singing: The Multi-Modal Imagery Association (MMIA) Model. *Music Perception*, *32(3),* 242-253.

Raphael, L., Borden, G.J. & Harris, K.S. (Eds.) (2011). *Speech Science Primer*. Lippincott Williams & Wilkins.

Rowell, L. (1990). *Music and Musical Thought in Early India.* The University of Chicago Press.

Salimpoor, V.N., Benovoy, M., Larcher, K., Dagher, A. & Zatorre, R.J. (2011). Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. *Nature Neuroscience, 14(2),* 257-262.

Sundberg, J. (2019). The singing voice. In S. Fruhholz, & P. Belin (Eds.), *The Oxford Handbook of Voice Perception* (pp. 107-142). Oxford University Press.

Tatham, M. & Morton, K. (2011). *A Guide to Speech Production and Perception*. Edinburgh University Press.

Wyn, J. & Reimers, P. (2010). *Patterns in Child Phonology*. Edinburgh University Press.

Zatorre, R.J. (2015). Musical pleasure and reward: Mechanisms and dysfunction. *Annals of the New York Academy of Science, 1337(1)*. Retrieved from https://doi.org/10.1111/nyas.12677

Zavadska, G. & Davidova, J. (2017). Physiological features of developing 6-8-year-old children's vocal apparatus. *Problems in Music Pedagogy*, *16(1*), 95-105.

Received 07.07.2019
Accepted 21.12.2019