

# Beyond AUM:

## The Acoustics of the Fourth Element

Iannis Psallidakos

*Overtone Voices: Overtone Singing, Research and Spectral Listening*

overtonevoices.com | ORCID: 0009-0004-2234-0476

### Abstract

The Vedic syllable AUM is traditionally analysed as three phonemes (A, U, M), yet Indian philosophy has long described a fourth element beyond these audible segments. This paper reports a spectrographic observation of what occurs during the sustained nasal /m/ when chanted with deliberate articulatory awareness.

Using VoceVista Video and Praat analysis of sustained vocalisations at four fundamental frequencies ( $f_0 \approx 54$  Hz creaky voice, 105 Hz inert control, 108 Hz modal phonation, and 136 Hz modal phonation), I document a continuous reorganisation of the upper formant structure during the nasal /m/: F2 rises from  $\sim 700$  Hz to  $\sim 2500$  Hz, F3 rises from  $\sim 2500$  Hz to  $\sim 3700$  Hz, and the nasal anti-resonance (a travelling band of spectral silence) migrates from  $\sim 500$  Hz to  $\sim 3000$  Hz. The rising F2 and the rising anti-resonance converge at  $\sim 1900$  Hz, where they encounter a steady peak (N3, the third resonance of the pharyngo-nasal tube), producing brief deep attenuation before F2 recovers. The trajectory is reproduced across all three living fundamentals (54, 108, 136 Hz); the N3 anchor remains within a narrow frequency band regardless of  $f_0$ . A pitch-matched comparison between a living /m/ at 108 Hz and an inert /m/ at 105 Hz, from the same performer in the same recording session, demonstrates that the acoustic contrast between the two variants is attributable to articulatory behavior, specifically deliberate tongue advancement, and not to pitch-related or register-related differences.

I distinguish this 'living' /m/ from an 'inert' variant with stationary spectral features, and propose the living trajectory as the acoustic correlate of the graded sonic dissolution described in the nāda-yoga strand of the Indian philosophical tradition: the same continuous sonic progression named by the tradition from the inside, documented here from the outside. The acoustic contrast between living and inert /m/ aligns with the traditional distinction between 'struck'

(*āhata*) and 'unstruck' (*anāhata*) sound, and with the fourth element long described in the Upaniṣadic and Yoga traditions.

**Keywords:** AUM, OM, formant trajectory, nasal anti-resonance, pole-zero crossing, N3 nasal formant, vocal tract acoustics, anāhata nāda, turīya, bindu, spectrographic analysis

## 1. Introduction

In a recent companion paper (Psallidakos 2026c), I documented a continuous F2 rise during the sustained nasal murmur /m/ of the Vedic syllable AUM, as part of a comparative spectrographic study of three sacred vowel sequences. That paper proposed the term *pneumatic sweep*, defined as the complete traversal of the formant space (from open-back through close-front vowel regions) on a single sustained phonation, to describe the shared acoustic principle observed across the Greek seven-vowel sequence of the Papyri Graecae Magicae, the Japanese kototama Futonorito, and the Vedic AUM: a complete traversal of the cardinal vowel triangle on a single sustained phonation. In the PGM and kototama sequences, this traversal proceeds through discrete vowel stations. AUM, by its phonemic notation, appears to traverse only two vowel positions and then to close into a nasal consonant that, on the phonemic level, contributes no further vowel movement.

The present paper reports what happens during that nasal phase when it is performed with deliberate articulatory awareness. The /m/ of AUM is not acoustically inert. When the tongue advances slowly inside the closed mouth, the spectrogram shows a continuous reorganisation of the upper formant structure.

Previous acoustic studies of OM chanting (Gurjar & Ladhake 2008, 2009; Schwonek et al. 2018; Wani et al. 2024; Pattanayak et al. 2025) have treated the /m/ phase as a steady-state feature and focused on F0, formant stability, or perceptual effects; none have reported time-resolved trajectories inside the sustained nasal phase.

This paper makes one empirical contribution and offers one interpretive proposal. The empirical contribution is a high-resolution spectrographic description of the trajectory inside the sustained /m/, carried jointly by F2, F3, and the nasal anti-resonance, together with a fourth steady feature at ≈1900 Hz that I identify as the third resonance of the pharyngo-nasal tube (N3). The interpretive proposal is that this continuous four-feature reorganisation inside the sustained nasal phase may be read as the acoustic correlate of the graded sonic dissolution described in

the nāda-yoga strand of the Indian philosophical tradition. The author is an overtone singing practitioner and vocal acoustics researcher, not an Indologist, and defers to specialists in Indian philosophy for any judgement on the parallel.

### **1.1 What This Paper Does Not Claim**

Before describing the findings, I wish to be explicit about what this paper does not claim. It makes no claim that the articulation described here represents how AUM was chanted in antiquity, how it should be chanted today, or how traditional practitioners do chant it. It does not claim that the acoustic trajectory **is identical with, or produces, *anāhata nāda* or *turīya*** as inner experiential categories; the measurable acoustic domain and the phenomenological domain of contemplative practice remain distinct, and any claim to have 'found' an inner state on a spectrogram would be philosophically naive. The deeper parameters of AUM as a contemplative practice (intention, presence, the listening awareness) exceed any acoustic measurement.

What this paper does propose, in §4.2, is a more specific claim. The Indian tradition contains two distinct framings of the fourth element. In the Advaitic strand of the Māṇḍūkya Upaniṣad, the fourth is *amātra* (without measure), soundless, and lies beyond phonation; the paper makes no claim about this strand. In the nāda-yoga strand (Nāda-bindu, Dhyāna-bindu, Haṭha Yoga Pradīpikā, Śārngadeva), the path toward the fourth is explicitly described as a continuous sonic progression from gross to progressively subtler inner sound. It is specifically within this second strand that the present paper proposes the living trajectory as the acoustic correlate: the same graded sonic dissolution described by the tradition from the inside, documented here from the outside in the measurable acoustic domain of the same sustained practice.

### **1.2 The Indian Philosophical Framework**

The Māṇḍūkya Upaniṣad maps AUM's phonemic structure onto states of consciousness (A to waking, U to dream, M to deep sleep), and identifies a fourth element, *turīya*, that lies beyond the three audible phonemes. Verse 12 calls this fourth element *amātra* (without measure) and *advaita* (non-dual). The visual symbol ॐ marks it with a dot (*bindu*) placed above the three curves of A, U, and M. The Yoga Upaniṣads (Nāda-bindu, Dhyāna-bindu, Haṃsa) describe the approach to the fourth element as a **graded dissolution**, a progression through increasingly subtle stages (including *bindu*, *nāda*, *kalā*) from articulate sound toward the silence of *anāhata nāda*, the 'unstruck sound.' Classical musicological treatises, especially Śārngadeva's Saṅgīta-

Ratnākara (13th c.), draw the technical distinction between *āhata* ('struck' sound, produced by ordinary physical mechanisms and perceptible to the senses) and *anāhata* ('unstruck', belonging to a different order of sonic experience). While the Māṇḍūkya Upaniṣad describes the fourth as soundless (*amātra*), tantric and nāda-yoga sources associate the *anusvāra* with a graded dissolution from articulate sound through progressively subtler *nāda* toward *anāhata* (Beck 1993; Padoux 1990; Rowell 1992; Wilke and Moebus 2011).

In an interview at the 19th World Sanskrit Conference (Kathmandu, June 2025), Finnian M. Gerety described the Vedic practice of *uccāra* ('ascending utterance') as a continuous vertical rising of sound through the body: 'the sound is vertical but then also the soul's journey is vertical' (Gerety 2025). In the same interview, he contrasts modern wellness readings of mantra with the traditional understanding: in Vedic, tantric, and Buddhist contexts mantras are 'not much about de-stressing. It's a kind of serious thing, it's a matter of life and death. A mantra is a kind of tool of transforming your body and your state of mind' (Gerety 2025). These two framings from the same source (sound as ascending vertical movement, and mantra as a matter of life and death) suggest structural parallels with two features of the acoustic observation reported below: a continuous spectral ascent inside the sustained nasal phase, and the co-presence within it of trajectories of energy and a trajectory of silence.

A closely related phenomenological observation is given by Reznikoff (2004/2005), who describes a simple body-resonance exercise in which the sequence A-O-U-M is felt to rise progressively through the body, from the chest on A, to the throat on O, to the lower face on U, and finally to the upper part of the head on M. Reznikoff presents this ascent as a primitive, pre-linguistic element of sound perception and a foundation for his work in sound therapy and the acoustics of prehistoric caves. The ascent he describes phenomenologically, from chest to head, runs in the same direction as the spectral ascent reported acoustically in Section 3 of the present paper.

It is this nāda-yoga and musicological strand, together with Reznikoff's phenomenological ascent, rather than the Māṇḍūkya's soundless *amātra*, that provides the framework for the acoustic correlate proposed in Section 4.

## 2. Materials and Methods

### 2.1 Performer and Recording Conditions

All recordings were made by a trained overtone singer with over ten years of experience in sustained vowel production and harmonic awareness. Recordings were made in a quiet domestic room using an Audio-Technica AT2035 large-diaphragm cardioid condenser microphone and analysed with VoceVista Video (Sygyt Software) and Praat (Boersma and Weenink 2024). As self-experimentation on the performer's own voice, this study did not require ethics board approval. These are single-speaker exploratory measurements; replication with additional singers and traditional practitioners is a priority for future work.

## **2.2 Recording Protocol**

AUM was performed as a single sustained vocalisation on one exhalation at  $f_0 \approx 108$  Hz (adult male chest voice, modal phonation) as the principal reference condition. To test the pitch-independence of the observed trajectories, additional **living /m/** recordings were made at  $f_0 \approx 136$  Hz (chest voice, modal phonation) and at  $f_0 \approx 54$  Hz (creaky voice with subharmonic register). As a pitch-matched control, an **inert /m/** recording was made at  $f_0 \approx 105$  Hz (chest voice, modal phonation), in the same session and from the same register as the 108 Hz living recording; this close pitch match ( $\Delta f_0 \approx 3$  Hz) allows the living/inert contrast to be attributed to articulatory behaviour rather than to pitch, register, or tract-length differences. The choice of fundamentals spans approximately 1.3 octaves within the performer's comfortable register and is not proposed as historically privileged.

The creaky-voice condition at  $f_0 \approx 54$  Hz produces densely packed harmonics that reveal formant envelopes and anti-resonant features as clear spectral contours. It serves an analytical purpose only and is not proposed as a normative chanting register. Two performance variants of AUM were recorded as the central contrast of this paper:

**Inert /m/ (control condition).** /a/ → /u/ → /m/, with the nasal murmur sustained but the tongue allowed to remain near its retracted /u/ position throughout, without deliberate advancement.

**Living /m/ (experimental condition).** As above, but with the anterior and middle tongue body advancing slowly and continuously during the sustained /m/, from its retracted /u/ position toward broad contact with the hard palate (the position of /n/, though no /n/ is heard because the lips remain closed throughout). The lips close before any tongue-palate contact is made: the /m/ is formed by lip closure, and the living trajectory inside it is produced by the tongue's subsequent slow advance. The endpoint approximates the rest tongue position described in myofunctional orofacial literature, in which the entire tongue rests gently against the palate with

the tip at the alveolar ridge (the small bony ridge just behind the upper front teeth). The advancement is performed with deliberate somatic awareness, over the full duration of the available exhalation (approximately 8 to 12 seconds), and the chanter listens actively to the resulting spectral reorganisation throughout.

## 2.3 Spectrographic Analysis

F1, F2, F3, and the location of the nasal anti-resonance were estimated visually from spectrogram peaks and troughs in VoceVista Video, and cross-checked against Praat's Burg LPC formant tracker output (6 formants, 5000 Hz ceiling, 25 ms window) wherever the tracker produced a stable label. Frequency values reported in this paper should therefore be read as visual estimates consistent with both instruments: ranges and qualitative trajectories are taken from VoceVista, while specific landmark values (N3, N2, F2 endpoints) are corroborated by Praat formant listing output. The trajectories described in Section 3 were observed consistently across multiple exploratory takes at each fundamental frequency; the values reported are representative of these takes. Because all-pole LPC algorithms cannot distinguish nasal poles from oral-tract poles (Pruthi, Espy-Wilson, and Story 2007), the tracker's sequential formant labelling (F1, F2, F3 from bottom to top) does not track a single physical resonator through the pole-zero crossing region. Manual F2 reconstruction was therefore performed from Praat's formant listing output for the three LPC-tracked recordings (108 Hz living, 136 Hz living, 105 Hz inert), selecting at each time step the spectral peak continuous with the preceding F2 estimate.

## 3. Results

### 3.1 The Inert /m/ vs. the Living /m/: a pitch-matched contrast

The principal empirical finding of this paper is presented in Figure 1: an **inert /m/** recorded at  $f_0 \approx 105$  Hz and a **living /m/** recorded at  $f_0 \approx 108$  Hz, both from the same performer in the same recording session using modal chest phonation. The close pitch match ( $\Delta f_0 \approx 3$  Hz) allows the acoustic difference between the two to be attributed to deliberate tongue advancement rather than to pitch, register, or tract-length differences.

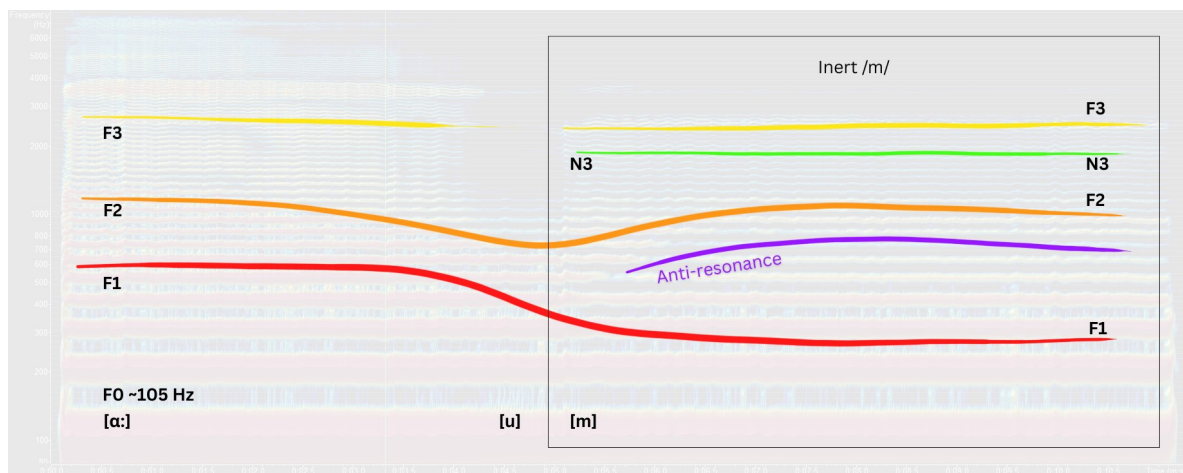
In the **inert /m/** (Figure 1, top pair), with the tongue remaining near its retracted /u/ position throughout the sustained murmur, the spectrogram is essentially static: F2 descends from  $\sim 1100$  Hz during /a/ to  $\sim 700$  Hz at the onset of /m/ and remains there; the anti-resonance sits

stationary in the region 700 to 800 Hz between F1 and F2; a steady peak is visible at  $\approx 1900$  Hz throughout the nasal phase. This configuration is consistent with published descriptions of the static acoustic structure of /m/ (Fujimura 1962; Stevens 1998, Chapter 9) and with static measurements of the /m/ portion of AUM reported in earlier studies (Schwonek et al. 2018; Pattanayak et al. 2025).

In the **living /m/** (Figure 1, bottom pair), with the tongue advancing deliberately during the sustained murmur, the spectrogram shows a continuous reorganisation. Three upper-spectrum features move in parallel: F2 rises from  $\sim 700$  Hz to  $\sim 2500$  Hz, F3 rises from  $\sim 2500$  Hz to  $\sim 3700$  Hz, and the anti-resonance migrates upward from  $\sim 500$  Hz to  $\sim 3000$  Hz. A fourth feature, the  $\approx 1900$  Hz steady peak already visible in the inert recording, remains fixed throughout. The rising F2 and rising anti-resonance converge at  $\sim 1900$  Hz, producing a brief deep attenuation of F2 before it recovers and continues upward.

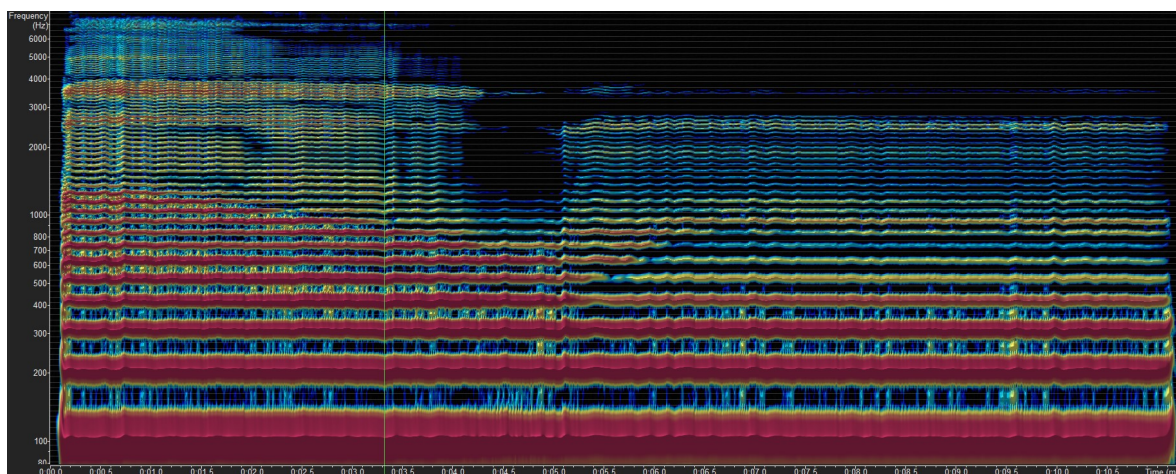
The contrast is visible side-by-side in Figure 1 at matched pitch. The detailed structure of the living trajectory, namely the four spectral features and the crossing event, is described in §3.2.

**Figure 1.**

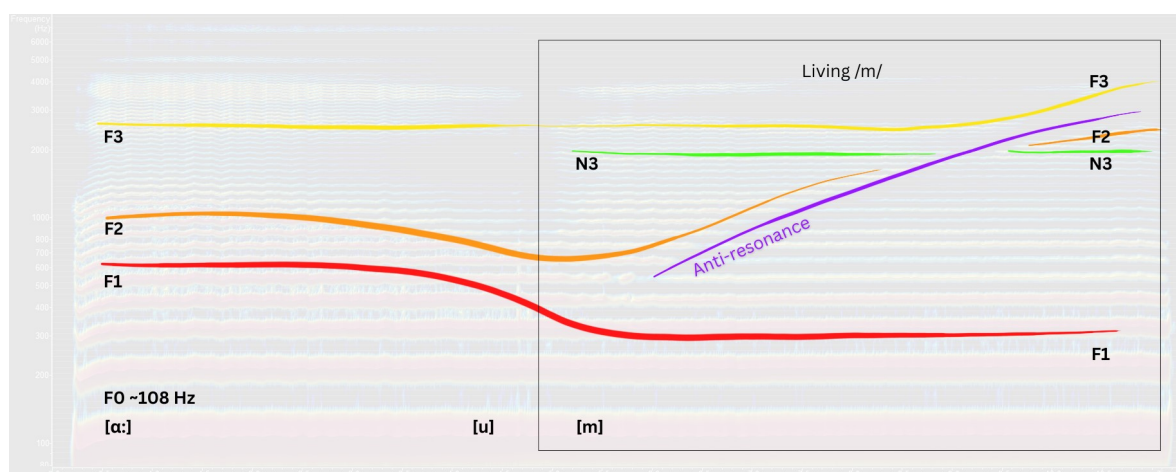


*Inert /m/,  $f_0 \approx 105$  Hz, annotated trajectory*

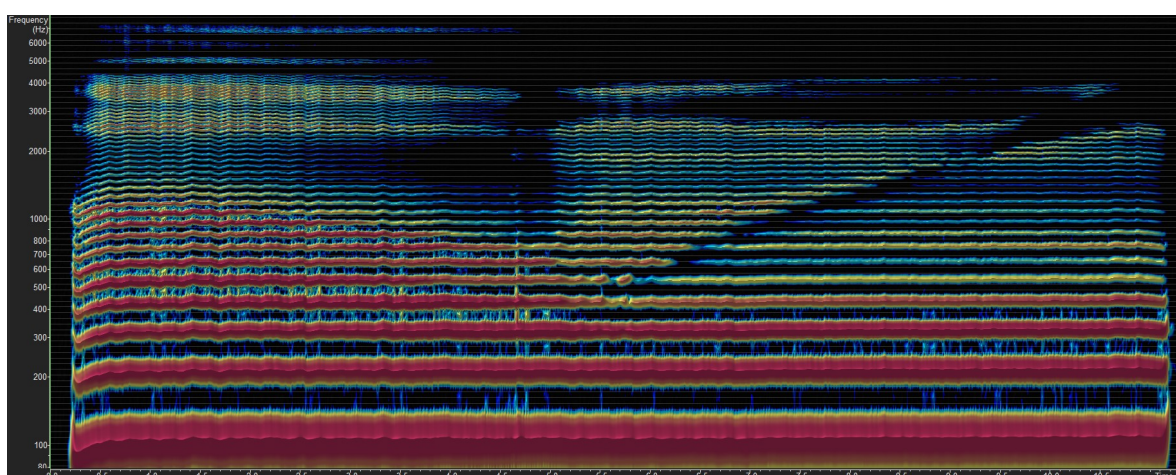




*Inert /m/,  $f_0 \approx 105$  Hz, raw spectrogram*



*Living /m/,  $f_0 \approx 108$  Hz, annotated trajectory*



*Living /m/,  $f_0 \approx 108$  Hz, raw spectrogram*



*Figure 1. The inert/living contrast at matched pitch (chest voice, modal phonation). Top pair: the inert /m/ at  $f_0 \approx 105$  Hz, showing the stationary anti-resonance, stationary F2, and static nasal envelope. Bottom pair: the living /m/ at  $f_0 \approx 108$  Hz, same performer, same recording session, showing the continuous upward reorganisation of F2 (orange) and the anti-resonance (purple) rising through the stable N3 anchor (green) at approximately 1900 Hz. The pitch difference between the two recordings is  $\approx 3$  Hz; the acoustic difference is not.*

### **3.2 Structure of the living trajectory: three moving features and a fixed anchor**

The reorganisation inside the living /m/ is carried by three co-evolving spectral trajectories together with one stable anatomical anchor. Their structure is summarised here; the identification of the  $\approx 1900$  Hz anchor as the third resonance of the pharyngo-nasal tube (N3) is discussed in §4.1, together with the canonical literature values against which it is interpreted.

**Rising F2.** F2 rises continuously from  $\sim 700$  Hz (carried over from the preceding /u/) toward  $\sim 2500$  Hz as the anterior and middle tongue body advances toward broad contact with the hard palate.

**Rising F3.** F3 rises from  $\sim 2500$  Hz toward  $\sim 3700$  Hz, displaced upward as the rising anti-resonance approaches it from below through the well-known repulsion between poles and zeros in the coupled nasal-oral system (Fant 1960; Fujimura 1962).

**Rising anti-resonance.** The nasal anti-resonance (the frequency band in which the coupled nasal-oral system absorbs rather than amplifies sound) rises continuously from  $\sim 500$  Hz toward  $\sim 3000$  Hz. The anti-resonance frequency of /m/ is determined by the length of the oral side-branch from the velopharyngeal port to the lip closure (Fujimura 1962; Stevens 1998, Chapter 9); as the anterior and middle tongue body advances toward broad contact with the hard palate, this side-branch shortens progressively and the anti-resonance rises accordingly. The anti-resonance is not a formant but a **trajectory of silence**: a travelling band of spectral absence migrating upward through the harmonic spectrum in parallel with the rising formants, particularly distinct in the creaky-voice condition (Figure 2) where densely packed harmonics make the trough sharply defined.

**Fixed anchor at  $\approx 1900$  Hz.** From the moment of lip closure onward, a steady peak at  $\approx 1900$  Hz is visible throughout the sustained /m/, independent of tongue motion. A secondary steady peak at  $\approx 900$  Hz is also present. These two peaks are consistent with N3 and N2 of the pharyngo-

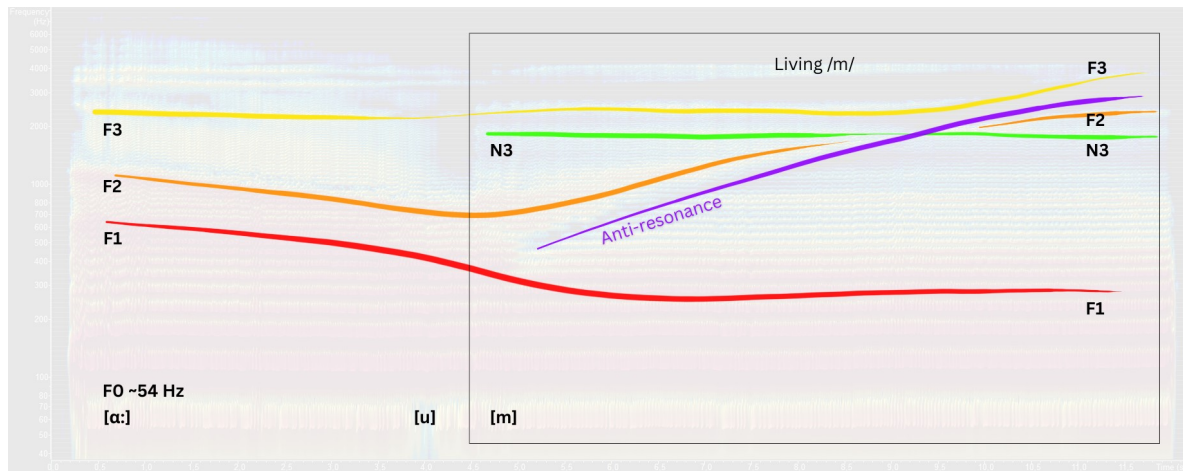
nasal tube respectively (see §4.1 for the identification); because the nasal cavity is bounded almost entirely by bone and cannot be reshaped by any articulatory gesture (Maeda 1993; Pruthi, Espy-Wilson, and Story 2007), the  $\approx 1900$  Hz peak is a fingerprint of the speaker's own anatomy.

**The crossing.** The rising F2 and the rising anti-resonance converge at  $\approx 1900$  Hz, precisely at the frequency of the fixed anchor. At this moment F2 is strongly attenuated by the coincident anti-resonance, the formant track narrows to a deep spectral dip, and harmonic energy is suppressed nearly to background level. In a real damped vocal tract both pole and zero have finite bandwidth, so energy is attenuated rather than mathematically erased; the dip is brief. F2 recovers full amplitude by  $\sim 2100$  Hz and continues upward toward its endpoint at  $\sim 2500$  Hz. This crossing is predicted by classical nasal acoustics: the 1500 to 2000 Hz region is what Fant, Fujimura, and Stevens describe as a "battleground" between nasal poles and oral-cavity zeros, and when a moving oral-origin zero approaches a nearby nasal pole, pole-zero interaction produces exactly these characteristic spectral effects (Fant 1960; Fujimura 1962).

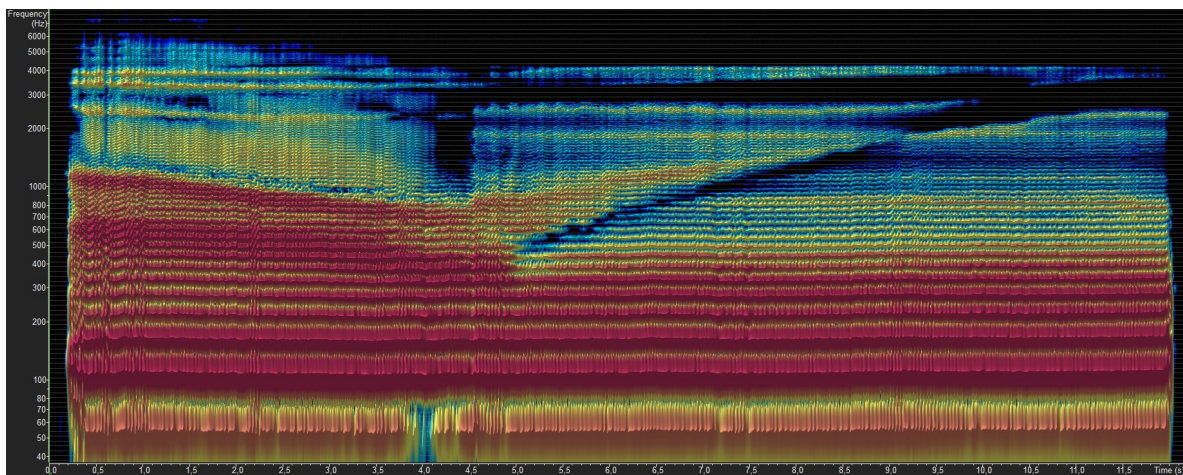
**Pitch-independence of the trajectory.** The four-feature reorganisation (three rising trajectories plus one fixed anchor) was observed at all three living fundamentals:  $f_0 \approx 54$  Hz (creaky voice, Figure 2),  $f_0 \approx 108$  Hz (chest voice, Figure 1, bottom pair), and  $f_0 \approx 136$  Hz (chest voice, Figure 3). The N3 anchor remained within a narrow frequency band ( $\approx 1850$  to  $1950$  Hz) across all three fundamentals, while the harmonic positions on which F2 landed differed with the harmonic spacing of each  $f_0$ . This constancy of the N3 frequency across fundamentals supports its identification as an anatomically determined nasal resonance rather than a harmonic artefact.

**Qualitative features of the trajectory.** Three features of the recordings distinguish the living trajectory from incidental tongue advancement: the ascent is *progressive and continuous* (not stepped), the progression unfolds on a *slow timescale* comparable to *uccāra* in sustained Vedic chanting (several seconds, not the tens of milliseconds of coarticulatory transitions in ordinary speech), and the crossing zone at  $\approx 1900$  Hz exhibits *dense spectral resonance* rather than a faint or noise-like region. These features matter for the contemplative dimension discussed in §4: a tongue that advances too quickly or without sustained pharyngeal resonance produces some version of a rising trajectory but not the coherent, graded, resonant reorganisation documented here.

**Figure 2.**



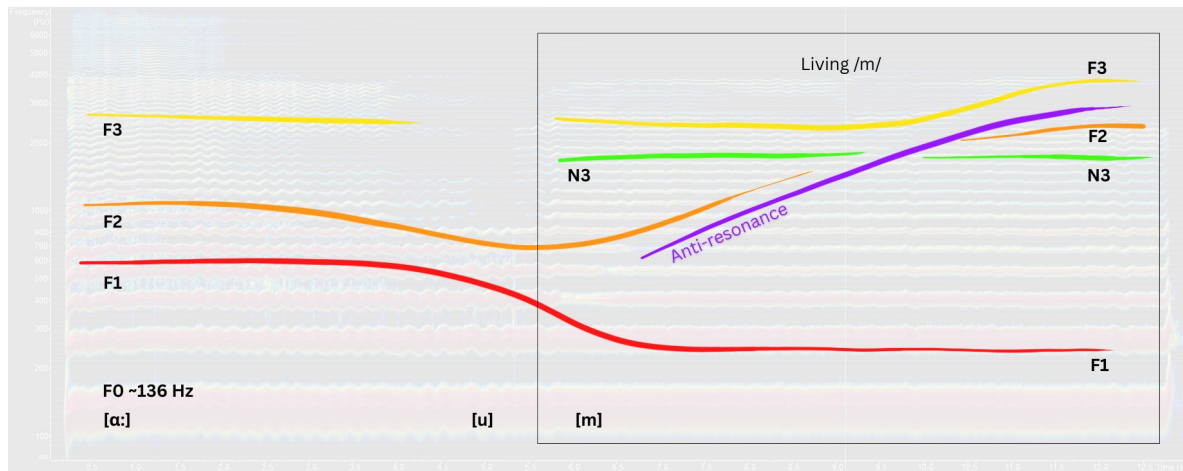
***Living /m/,  $f_0 \approx 54$  Hz (creaky voice), annotated trajectory***



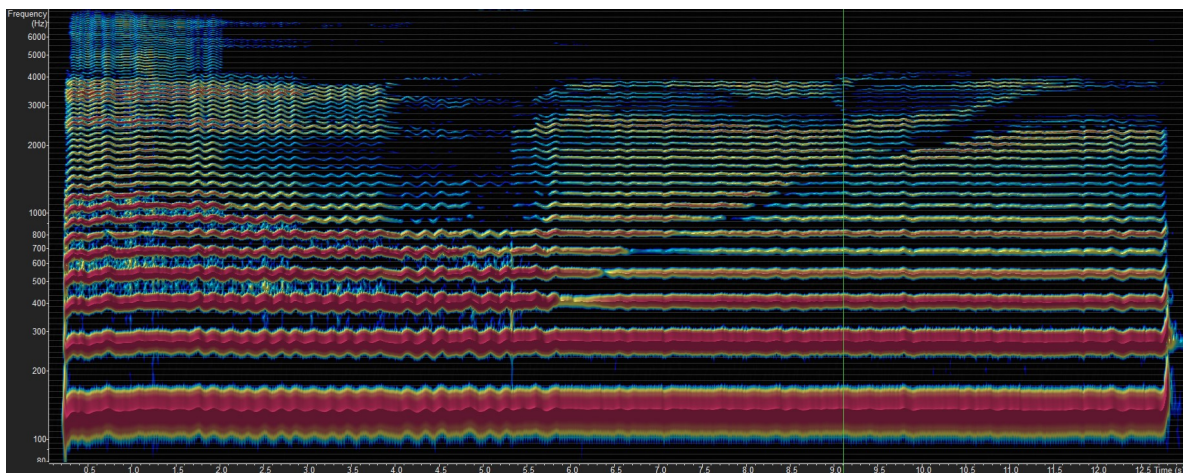
***Living /m/,  $f_0 \approx 54$  Hz, raw spectrogram***

Figure 2. Living /m/ at  $f_0 \approx 54$  Hz (creaky voice with subharmonic register). The dense harmonic packing characteristic of creaky voice makes the anti-resonance trajectory (purple, top panel) particularly distinct as a dark migrating band in the raw spectrogram (bottom panel). The rising F2 (orange), rising anti-resonance, and steady N3 anchor (green) at  $\approx 1900$  Hz are all reproduced at this low fundamental.

**Figure 3.**



***Living /m/,  $f_0 \approx 136$  Hz (chest voice), annotated trajectory***



***Living /m/,  $f_0 \approx 136$  Hz, raw spectrogram***

Figure 3. Living /m/ at  $f_0 \approx 136$  Hz (chest voice, modal phonation). The three-trajectory reorganisation is reproduced at this higher fundamental, with the N3 anchor remaining in the same frequency region ( $\approx 1900$  Hz) as at  $f_0 \approx 108$  Hz and  $f_0 \approx 54$  Hz, and with F2 rising through N3 toward  $\sim 2400$  Hz. The wider harmonic spacing at 136 Hz makes the anti-resonance trajectory less visually dense than in the 54 Hz creaky-voice condition of Figure 2, but the same structural pattern is present.

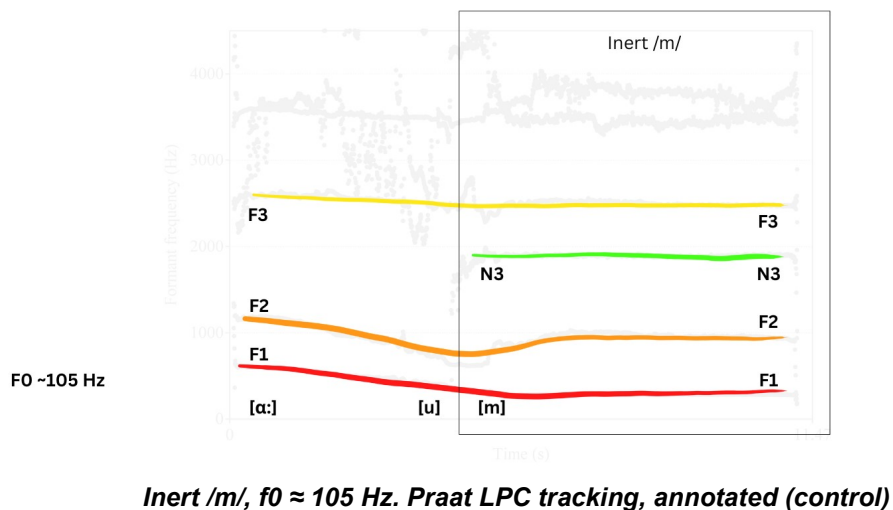
### 3.3 Automatic formant tracking validation

To validate the visual trajectories, Praat's Burg LPC formant tracker was applied to three recordings: the 108 Hz living /m/, the 136 Hz living /m/, and the 105 Hz inert /m/ (Figure 4). For the two living recordings, the automatically tracked F2 contour confirms the visually estimated

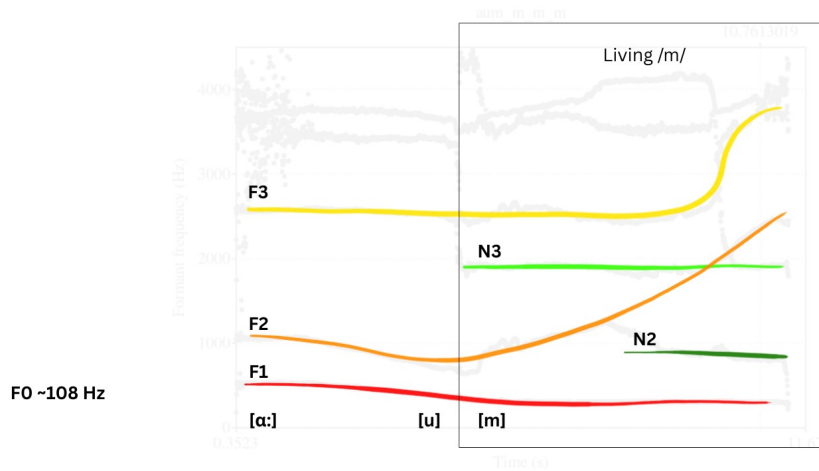
trajectory, rising from the /u/ region through the  $\approx 1900$  Hz crossing zone toward  $\sim 2500$  Hz (at 108 Hz) and  $\sim 2400$  Hz (at 136 Hz). A discontinuity appears at lip closure as the vocal tract coupling shifts abruptly from oral to nasal radiation. Inside the sustained /m/, a region of tracker instability appears in the  $\approx 1900$  Hz zone, where the algorithm's sequential formant labelling relabels the steady N3 peak, first assigning it to its 'F3' track (when the rising oral F2 is still below N3), then to its 'F2' track (after oral F2 has risen above N3). This tracker confusion is itself consistent with the pole-zero interaction described in §3.2: the relabelling occurs precisely where a nasal pole and a rising oral formant occupy the same spectral region.

For the 105 Hz inert control, the Praat tracker shows a fundamentally different picture: F2 descends smoothly from  $\sim 1200$  Hz during /a/ to  $\sim 900$  Hz during /u/ and /m/ and remains stationary; no rising oral F2 crosses the N3 frequency; the N3 peak at  $\approx 1900$  Hz appears at the onset of /m/ and remains stable throughout, without relabelling instability. The presence of N3 in both the living and inert recordings confirms that this anchor is anatomical and produced by the act of lip closure itself, while the rising F2 that crosses N3 is present only in the living variant. A steady peak at  $\approx 900$  Hz, consistent with N2, is present in all three recordings.

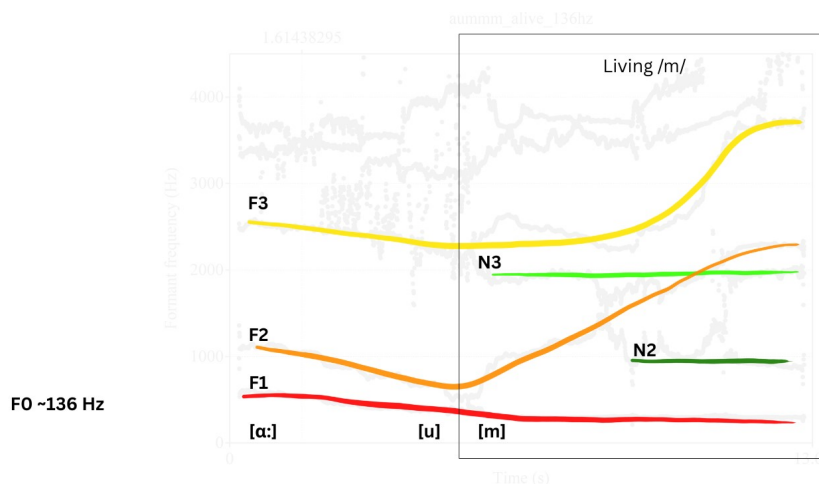
**Figure 4.**







***Living /m/,  $f_0 \approx 108$  Hz. Praat LPC tracking, annotated***



***Living /m/,  $f_0 \approx 136$  Hz. Praat LPC tracking, annotated***

*Figure 4. Automatic formant tracking output (Praat, Burg LPC, 6 formants, 5000 Hz ceiling, 25 ms window). Top panel: inert /m/ at 105 Hz, showing the N3 anchor at  $\approx 1900$  Hz present throughout /m/ but no rising oral F2 and no crossing event. Middle and bottom panels: living /m/ at 108 Hz and 136 Hz. In both, a visible discontinuity appears at lip closure, and a region of instability appears around 1900 Hz where the steady N3 peak is relabelled between formant tracks as the rising oral F2 crosses through the N3 frequency. A steady peak at  $\approx 900$  Hz, consistent with N2, is present in all three recordings.*

## 4. Discussion

### 4.1 *Acoustic novelty and theoretical context*

The individual phenomena combined in the living /m/ are each well established in classical acoustic phonetic theory: rising F2 and F3 during tongue advancement, nasal anti-resonances shifting with oral side-branch geometry, pole-zero interactions in the coupled nasal-oral system, and stable nasal tract resonances from the rigid bony anatomy (Fant 1960; Fujimura 1962; Maeda 1993; Stevens 1998; Pruthi, Espy-Wilson, and Story 2007). What the present paper adds is the time-resolved spectrographic documentation of these phenomena unfolding together within a single sustained nasal phonation, producing a coordinated reorganisation of the upper formant structure over several seconds of continuous articulatory motion. This phenomenon does not arise in ordinary speech, where nasal consonants last 50 to 120 ms and speakers actively compensate against pole-zero collisions that would compromise consonant identification (Engwall et al. 2006). The pitch-matched living/inert comparison reported here (Figure 1) further shows that the difference between the two variants is not reducible to pitch, register, or tract-length differences: it is produced by deliberate tongue advancement within the sustained nasal phase.

**Identification of the  $\approx 1900$  Hz anchor as N3.** The steady  $\approx 1900$  Hz peak observed in both living and inert recordings is most parsimoniously interpreted as N3, the third resonance of the pharyngo-nasal tube. Fujimura (1962) reported poles in /m/ at 1450 and 2000 Hz; Stevens (1998, Chapter 9) gives the nasal murmur formant series as approximately 250–300 Hz (N1),  $\approx 1000$  Hz (N2),  $\approx 2000$  Hz (N3), and  $\approx 3000$  Hz (N4); a simple quarter-wave model of a pharyngo-nasal tube slightly longer than the standard  $\approx 20$  cm predicts N3 at the observed  $\approx 1900$  Hz. Direct confirmation of the N3 identification would require MRI imaging of the performer's pharyngo-nasal tract or transmission measurements (see §7d); alternative explanations, including sphenoidal sinus coupling in the 1700–1900 Hz range (Dang and Honda 1996), cannot be excluded at this resolution.

### 4.2 *The acoustic correlate of the nāda-yoga dissolution*

The Indian tradition contains two distinct framings of the fourth element of AUM. The Advaitic strand of the Māṇḍūkya Upaniṣad describes it as *amātra* (without measure), soundless, and beyond phonation; the paper's §1.1 already sets that strand aside. The nāda-yoga strand



frames the fourth differently. The Nāda-bindu, Dhyāna-bindu, and Haṃsa Upaniṣads, the Haṭha Yoga Pradīpikā (*nāḍānusandhāna*), and Śārṅgadeva's Saṅgīta-Ratnākara describe the path toward *anāhata* as a **continuous sonic progression**: the practitioner hears a graded dissolution from gross, external sound into progressively subtler inner sounds (ocean, bell, flute, vīṇā, bee-hum, and so on), arriving at *anāhata* not as silence but as a qualitatively different order of sonic experience. Śārṅgadeva's distinction between *āhata* (struck, externally perceptible sound) and *anāhata* (unstruck, inner sonic experience) makes this explicit. Within this strand, the path toward the fourth is a sonic journey throughout.

The continuous four-feature reorganisation documented in Section 3 is, within the measurable acoustic domain, precisely such a graded sonic progression inside a single sustained phonation. Three moving trajectories (F2, F3, and the anti-resonance) rise continuously over several seconds from their /u/-region starting values into progressively higher spectral regions. A fixed anatomical anchor at ≈1900 Hz (N3) marks the spectral region where the rising F2 and the rising anti-resonance converge. The convergence produces a brief deep attenuation of F2, after which F2 recovers and continues upward into an upper-register configuration structurally different from the starting one. The trajectory is absent in the inert /m/ (Figure 1, top pair) and present in the living /m/ (Figure 1, bottom pair) at matched pitch.

I propose the living trajectory as the **acoustic correlate** of the nāda-yoga dissolution: not a resemblance from outside, but a parallel process realised in the measurable acoustic domain of the same practice that the tradition describes from the inside. Within a tradition that frames the approach to *anāhata* as a sonic journey from gross to subtle, the spectrogram of the living /m/ documents what that sonic journey looks like inside the vocal tract. Two features of the observation resonate with the tradition's own framings. The anti-resonance trajectory, the trajectory of silence, is a moving band of spectral absence interacting with trajectories of energy; within nāda-yoga, silence and sound are not opposed but interwoven in the graded dissolution. And the rising trajectories meet, at the crossing event, at the precise frequency of the fixed N3 anchor given by the chanter's own anatomy: the moving features are what the chanter produces, the anchor is what the body contributes.

Falsifiability operates at two distinct levels. At the **acoustic level**, the trajectory with its qualitative features (progressive ascent, slow uccāra-paced timescale, dense resonance at the crossing) should be reproducible in other singers. The paper does not claim that the trajectory is rare or unique to contemplative practice: any performer advancing the anterior and middle

tongue body during sustained /m/ will produce some version of the reorganisation. Multi-speaker replication would establish robustness of the qualitative features, not exclusivity. At the **neurophysiological level**, whether a given production constitutes a contemplative event depends on the producer's listening state: the acoustic trajectory is *necessary but not sufficient*. Recent MEG evidence shows that overtone-rich stimuli engage right-hemispheric theta-dominant processing in listeners (Saus, Seither-Preisler, and Schneider 2025); a companion hypothesis paper (Psallidakos 2026b) proposes that active production of such stimuli, including the living AUM described here, generates endogenous theta-gamma cross-frequency coupling. A within-subject protocol recording the same participants across three conditions (baseline AUM, inert /m/ training, living /m/ training), with simultaneous spectrographic and EEG or MEG measurement, would isolate the effect of the acoustic configuration on brain state and provide the test of the neurophysiological level of the claim.

### ***4.3 Acoustic possibility and contemplative practice***

The acoustic findings describe what the vocal tract is capable of doing during sustained /m/ when the tongue advances with deliberate awareness. They describe nothing about what the chanter perceives while doing it. Whether the chanter inhabits the inert or living variant in actual practice depends on whether the chanter is listening, that is, whether the spectral reorganisation becomes an object of awareness. One observational point follows, offered without prescription. In repetitive chanting of AUM or OM, the tongue's articulatory state during the short interval between repetitions matters for what the next /m/ can become: if the silence is traversed hastily with the tongue already anticipating the next /a/, the subsequent /m/ retains its retracted /u/-position configuration; if the silence is given its own duration and the tongue rests in broad palatal contact during that interval, the subsequent /m/ begins from a place where the living trajectory becomes available. The acoustic living /m/ is therefore conditioned not only on what happens during sustained phonation but on what happens in the silence preceding it.

This paper does not claim that the living variant is more authentic, more correct, or more contemplatively powerful than the inert variant; the parameters of contemplative practice exceed any acoustic measurement. What it offers is the acoustic description and the proposal made in §4.2. The listening, and everything that follows from it, belongs to the practitioner.

## **5. Limitations**

All measurements are single-speaker exploratory data from a trained overtone singer. Frequency values are visual estimates from VoceVista Video spectrograms cross-checked against Praat's Burg LPC formant tracker output for the three principal recordings (105 Hz inert, 108 Hz living, 136 Hz living), with manual F2 reconstruction through the pole-zero crossing region. All-pole LPC trackers cannot distinguish nasal poles from oral-tract poles (Pruthi, Espy-Wilson, and Story 2007); replication studies would benefit from pole-zero modelling methods explicitly designed for nasalised phonation. Multi-speaker replication is needed to establish the robustness of the trajectory and of its qualitative features (progressive ascent, slow timescale, dense resonance at the crossing) across performers. Recordings were made in a domestic room under analytical conditions, not in a ritual or meditative context. The identification of the steady  $\approx 1900$  Hz peak as N3 is consistent with classical nasal acoustics but definitive confirmation would require MRI imaging of the performer's pharyngo-nasal tract or transmission measurements. The author is not a specialist in Indian philosophy; the correlate proposed in §4.2 is offered to scholars of Indian philosophy for judgement.

## 6. Conclusion

When AUM is chanted with deliberate articulatory awareness, the sustained nasal murmur /m/ is not an inert closing consonant. It is a continuous reorganisation of the upper formant structure in which F2 rises from  $\approx 700$  Hz toward  $\approx 2500$  Hz, F3 rises from  $\approx 2500$  Hz toward  $\approx 3700$  Hz, and the nasal anti-resonance (a travelling band of spectral silence) migrates from  $\approx 500$  Hz toward  $\approx 3000$  Hz.

A fourth feature appears at the moment of lip closure and persists throughout: a steady peak at  $\approx 1900$  Hz consistent with N3, the third resonance of the pharyngo-nasal tube, the fingerprint of the speaker's own anatomy, independent of tongue motion. The rising F2 and the rising anti-resonance meet at the frequency of this fixed anchor, producing a brief deep attenuation of F2 before its upward recovery. The pitch-matched comparison between a living /m/ at 108 Hz and an inert /m/ at 105 Hz, from the same performer in the same recording session, demonstrates that the acoustic contrast is produced by articulatory behaviour rather than by pitch or register.

The contrast between the inert and living variants is not a contrast of degree but of kind: the same phonetic label covering two qualitatively different acoustic events. This difference can be read as an acoustic analogue of the musicological distinction between *āhata* and *anāhata*.

I propose the living trajectory as the acoustic correlate of the nāda-yoga dissolution: the graded sonic progression from gross to subtle that the tradition describes from the inside, documented here from the outside. Confirmation operates at two levels. Multi-speaker replication tests the acoustic level: whether the trajectory and its qualitative features are robust across performers. Simultaneous spectrographic and neuroimaging measurement tests the neurophysiological level: whether the listening brain state during living /m/ production differs from that during inert /m/ or untrained chanting. The acoustic trajectory is necessary but not sufficient for the contemplative correlate; the full claim is testable only through the combined acoustic and neurophysiological measurement (§7).

## 7. Future Directions

(a) Simultaneous spectrographic and EEG or MEG measurement during the within-subject three-condition protocol proposed in §4.2 (baseline chanting, inert /m/ training, living /m/ training), with a parallel group of experienced contemplative practitioners through the same three conditions. This is the primary test of the neurophysiological level of the correlate claim, in connection with the theta-gamma cross-frequency coupling hypothesis proposed independently (Psallidakos 2026b) and the listener-side evidence of Saus, Seither-Preisler, and Schneider (2025).

(b) Multi-speaker acoustic replication with singers and traditional practitioners from Vedic, tantric, and nāda-yoga lineages, to establish the robustness of the trajectory and its qualitative features across performers.

(c) Pole-zero modelling methods (rather than all-pole LPC) to quantify the three trajectories and the anti-resonance with greater precision than Burg formant tracking permits in the presence of nasal poles.

(d) MRI or real-time articulatory imaging to confirm the N3 identification of the steady  $\approx 1900$  Hz peak through direct measurement of the pharyngo-nasal tract.

(e) Interdisciplinary dialogue with scholars of Indian philosophy on the acoustic correlate proposed in §4.2, and with scholars of other contemplative and chanting traditions in which sustained vocal practice plays a central role.

## References

- Beck, G.L. (1993). *Sonic Theology: Hinduism and Sacred Sound*. Columbia, SC: University of South Carolina Press.
- Boersma, P., and Weenink, D. (2024). Praat: doing phonetics by computer (version 6.4). Available at: <http://www.praat.org/>.
- Dang, J. and Honda, K. (1996). Acoustic characteristics of the human paranasal sinuses derived from transmission characteristic measurement and morphological observation. *Journal of the Acoustical Society of America* 100(5): 3374-3383. DOI: 10.1121/1.416978
- Engwall, O., Delvaux, V., and Metens, T. (2006). Interspeaker variation in the articulation of nasal vowels. *Proceedings of the 7th International Seminar on Speech Production*, 3-10.
- Fant, G. (1960). *Acoustic Theory of Speech Production*. The Hague: Mouton.
- Fujimura, O. (1962). Analysis of nasal consonants. *Journal of the Acoustical Society of America* 34(12): 1865-1875. DOI: 10.1121/1.1909142.
- Gerety, F.M.M. (2015). *This Whole World is OM: Song, Soteriology, and the Emergence of the Sacred Syllable*. PhD dissertation, Harvard University. URI: <http://nrs.harvard.edu/urn-3:HUL.InstRepos:17467527>
- Gerety, F.M.M. (2021). 'Between Sound and Silence in Early Yoga: Meditation on Om at Death.' *History of Religions* 60(3): 209-244. DOI: 10.1086/711944.
- Gerety, F.M.M. (2025). Interview at the 19th World Sanskrit Conference, Kathmandu, Nepal, 26-30 June 2025. International Association of Sanskrit Studies. YouTube: <https://m.youtube.com/watch?v=KGDpjnHbN4w>.
- Gurjar, A.A. and Ladhake, S.A. (2008). 'Time-Frequency Analysis of Chanting Sanskrit Divine Sound OM Mantra.' *International Journal of Computer Science & Network Security* 8(8): 170-175. [http://paper.ijcsns.org/07\\_book/200808/20080825.pdf](http://paper.ijcsns.org/07_book/200808/20080825.pdf)
- Gurjar, A.A. and Ladhake, S.A. (2009). Spectral Analysis of Sanskrit Divine Sound OM. *Information Technology Journal* 8(6): 781-785. DOI: 10.3923/itj.2009.781.785
- Maeda, S. (1993). Acoustics of vowel nasalization and articulatory shifts in French nasal vowels. In M.K. Huffman and R.A. Krakow (eds.), *Nasals, Nasalization, and the Velum*, 147-167. San Diego: Academic Press. DOI: 10.1016/B978-0-12-360380-7.50010-7
- Padoux, A. (1990). *Vāc: The Concept of the Word in Selected Hindu Tantras*. Albany: State University of New York Press. DOI: 10.1017/S0041977X00001981
- Pattanayak, J., Belur, Y., and Ilavarasu, J. (2025). Acoustic and Psychological Effects of Omkar Chanting in Novice Indian Adults: A Quasi-experimental Study. *Journal of Clinical and Diagnostic Research* 19(6): VC07-VC12. DOI: 10.7860/JCDR/2025/77365.21111.
- Pruthi, T., Espy-Wilson, C.Y., and Story, B.H. (2007). Simulation and analysis of nasalized vowels based on magnetic resonance imaging data. *Journal of the Acoustical Society of America* 121(6): 3858-3873. DOI: 10.1121/1.2722220

Psallidakos, I. (2026a). The Formant Trajectory of the Greek Seven-Vowel Sequence. Preprint, Zenodo. DOI: 10.5281/zenodo.19115650.

Psallidakos, I. (2026b). Overtone Singing as Natural Theta-Gamma Cross-Frequency Neuromodulation. Preprint, Zenodo. DOI: 10.5281/zenodo.18861836.

Psallidakos, I. (2026c). The Pneumatic Sweep: Formant Trajectories in Three Sacred Vowel Sequences. Preprint, Zenodo. DOI: 10.5281/zenodo.19438733.

Reznikoff, I. (2004/2005). 'On Primitive Elements of Musical Meaning.' JMM: The Journal of Music and Meaning 3, Fall 2004/Winter 2005. Available at: <http://www.musicandmeaning.net/issues/showArticle.php?artID=3.2>.

Rowell, L. (1992). Music and Musical Thought in Early India. Chicago: University of Chicago Press.

Saus, W., Seither-Preisler, A., and Schneider, P. (2025). Harmonic vowels and neural dynamics: MEG evidence for auditory resonance integration in singing. *Frontiers in Neuroscience* 19: 1625403. DOI: 10.3389/fnins.2025.1625403.

Schwonek, J., Singh, R., Okanime, O., and McCraw, N. (2018). Fourier Analyses of Certain Mystical Vowel Sound Intonations Using PRAAT Software. *Rose+Croix Journal* 12: 69-89.

Shringy, R.K. and Sharma, P.L. (1978-1989). *Saṅgīta-Ratnākara of Śārṅgadeva: Text and English Translation*, 2 vols. Delhi: Motilal Banarsidass.

Stevens, K.N. (1998). *Acoustic Phonetics*. Cambridge, MA: MIT Press.

Svātmārāma. (15th c.). *Haṭha Yoga Pradīpikā*. Translated by Pancham Sinh (1914). Allahabad: Panini Office.

Wani, L.K., Upasani, D.E., Dhumal, P.S., Deshpande, A., and Jagadish Prasad, S. (2024). Analysis of Om (AUM) chanting using machine learning. *Educational Administration: Theory and Practice* 30(4): 220-232. DOI:10.53555/kuey.v30i4.1441.

Wilke, A. and Moebus, O. (2011). *Sound and Communication: An Aesthetic Cultural History of Sanskrit Hinduism*. Religion and Society 41. Berlin: De Gruyter.

## **Acknowledgements**

The author thanks colleagues in the voice acoustics for stimulating conversations that informed the development of this work.

## **Disclosure of interest**

The author is a professional overtone singing instructor and vocal acoustics researcher. The spectrographic measurements reported here were performed by the author on his own voice.

The research was conducted independently and without commercial influence. No other competing interests are declared.

### **Funding**

This work received no external funding.

### **Data Availability**

Spectrographic figures (VoceVista Video screenshots with annotated formant and anti-formant trajectories) are included in the manuscript. Audio recordings are available from the corresponding author upon reasonable request.

### **Statement on AI Assistance**

Portions of the editorial revision, literature cross-referencing, and English prose refinement of this manuscript were assisted by the large language model tool Anthropic Claude. All scientific claims, research hypotheses, interpretations of the spectrographic analyses, and all decisions about the selection and interpretation of the cited literature and the final text remain entirely the responsibility of the author.