Art and Performance Letters (2026)
Clausius Scientific Press, Canada

DOI: 10.23977/artpl.2026.070116
ISSN 2523-5877 Vol. 7 Num. 1

Voicing the Five Notes: Reframing Five-Element Music Therapy as a Vocal-Articulatory Tradition

Jianhua Zhang

Zhejiang Conservatory of Music, Hangzhou, Zhejiang, China

Keywords: Five-element music therapy, Vocal articulation, embodied singing, Appoggio, Music therapy, Yizong Jinjian

Abstract: Five-Element Music Therapy (FEMT) has built an international evidence base over the past decade documenting effects across post-stroke aphasia, cancer-related distress, and late-life depression. The subfield treats the five notes — gong, shang, jiao, zhi, yu — as listening stimuli administered through recordings, with the mode label as the operative therapeutic variable. The classical sources operate at a different register. The Mengzi, the Tongzhi: Qi Yin Lüe, and the Yizong Jinjian describe the five notes as vocal-articulatory configurations, locating each at a specific site of tongue, jaw, lip, and breath. Song Boyuan's philological critique reaches the modal reading and dissolves it; the vocal-articulatory reading is left untouched and was the older one. FEMT has been resting on the weakest available reading of its own sources. An international research program on singing as therapeutic practice — Sing Well, Melodic Intonation Therapy, the choral physiology trials of Vickhoff, Fancourt, and others — has independently developed methods the FEMT subfield could absorb but has not. Italian appoggio, the Chinese qi chen dan tian tradition, and the recent theory of embodied singing converge on a single anatomical zone of breath support, providing a cross-cultural conceptual bridge with consequences for protocol specification, outcome measurement, and clinical collaboration.

1. Introduction

Five-Element Music Therapy (FEMT) — the contemporary clinical formalization of a tradition that aligns the five notes of the Chinese pentatonic scale (gong, shang, jiao, zhi, yu) with the five visceral systems and emotions of the Huangdi Neijing — has crossed in the past decade from Chinese-language journals into international indexed publications, with meta-analyses now covering post-stroke aphasia, cancer-related distress, and late-life depression at effect sizes from moderate to large [1][2][3]. Across this growing body of evidence, the five notes are administered as recorded listening stimuli and the mode label is treated as the operative therapeutic variable. The classical sources from which FEMT draws — particularly the Mengzi, the Tongzhi: Qi Yin Lüe, and the Qing-dynasty Yizong Jinjian — describe the five notes in vocal-articulatory terms, locating each note at a specific site of tongue, jaw, lip, and breath rather than at a position within a scalar mode [4][5]. The contemporary subfield has therefore been resting on the weakest reading available of its own classical foundations: the philological critique advanced by Song Boyuan in 2016, which targets the modal reading and largely succeeds against it, leaves the vocal-articulatory reading not

only untouched but, on the textual evidence Song herself adduces, more textually defensible [6]. This article reframes FEMT as a vocal-articulatory tradition. The argument proceeds by reviewing the FEMT clinical literature against the philological record, by setting that record alongside an international research program on singing as therapeutic practice that the FEMT subfield has not yet engaged with [7][8][9], and by identifying a cross-cultural conceptual bridge — between Italian appoggio pedagogy, the Chinese qi chen dan tian tradition, and the recent theory of embodied singing — that the reframing makes available to both fields.

2. Five-Element Music Therapy as a Listening Paradigm: A Brief Review

2.1 Historical Formation and Core Formulation

The classical textual basis of FEMT lies in the Huangdi Neijing, particularly the Suwen chapter Yinyang Yingxiang Dalun, which establishes the canonical correspondences between the five notes, the five visceral systems, and the five emotional states: the liver corresponds in sound to jiao and in emotion to anger; the heart to zhi and to joy; the spleen to gong and to thought; the lung to shang and to grief; the kidney to yu and to fear [10][11]. A parallel passage in the Lingshu, the Xieke chapter, frames the relationship cosmologically: "heaven has the five notes, the human has the five viscera; heaven has the six pitches, the human has the six bowels" [12]. Contemporary FEMT practitioners derived from these passages a diagnostic-prescriptive scheme in which music in a given modal classification is matched to the affected visceral or emotional system — gong-mode pieces for the spleen-deficient patient, zhi-mode pieces for the patient with insufficient heart-yang, and so on [13][14]. The prescriptive logic uses two principles drawn from broader Chinese medical theory: xiang ying (resonance with), which selects modally similar music to validate the patient's current state, and xiang sheng or xiang ke (generation across or constraint upon), which selects modally complementary music to redirect it [15]. The two principles, taken seriously, already point past the listening paradigm. Validating a patient's current state and redirecting it are acts requiring the patient to do something with the music — to enter it, to be moved by it, to find a way out of it — and that act is more naturally specified at the level of bodily participation than at the level of acoustic exposure. The classical scheme was built on assumptions about the patient's relation to music that the contemporary recorded-listening protocol has quietly discarded.

2.2 The Contemporary Clinical Evidence Base

Over the past fifteen years, FEMT has been tested in an expanding range of clinical populations. Yang and colleagues' systematic review and meta-analysis on post-stroke aphasia documented significant improvements in language recovery measures with FEMT as an adjunct to standard rehabilitation [1]. The same group's later meta-analysis on cancer patients reported moderate effects on cancer-related distress, anxiety, and quality-of-life indicators [2]. More recent meta-analyses have extended the evidence base into post-stroke depression and into late-life depression and anxiety in older adults [16][17]. Two multicenter trials are registered on ClinicalTrials.gov: NCT00964314, which evaluated FEMT for advanced cancer quality of life, and NCT07048938, currently recruiting at Shanghai Mental Health Center, which is examining FEMT for depression with neuroimaging endpoints [18]. The domestic Chinese literature is correspondingly large and diverse, with applications across insomnia, perimenopausal syndromes, post-stroke depression, and oncology supportive care [19][20][21]. FEMT is, by any reasonable accounting, no longer a parochial topic. What remains striking, given the size of the evidence base, is the degree to which its mechanistic vocabulary has lagged behind its clinical traction. The meta-analyses cited above credit the effects to "five-element music," but the construct doing the explanatory work is

essentially undifferentiated — a recorded stimulus assigned to a modal category, exposure measured by minutes per session and weeks of intervention. Effect sizes accumulate; the variable producing them remains conceptually unresolved. Clinical effects without specified mechanism do not refute a research program, but they leave it vulnerable to the next generation of well-controlled trials and limit the protocol refinement that would translate moderate effects into clinically decisive ones.

2.3 The Descriptive Lacuna

Read from a music-educator's vantage, this body of evidence has a curious blind spot. The music itself is described almost exclusively in second-order vocabulary — by mode name, by classical-piece title, by approximate duration of exposure — and almost never by what a musician would call its sonic specifics: instrumentation, tempo, dynamic profile, vocal versus instrumental rendering, recording properties. The consequence is one the meta-analyses themselves cannot register: from the published reports alone, it is genuinely difficult to know whether two trials of "gong-mode therapy" administered the same musical stimulus or two musically dissimilar stimuli that happened to share a modal label. Two recent contributions have begun to register this limitation from within the field. Yao and colleagues' 2025 trial administered the same five-element modal material to medical students on three different traditional Chinese instruments — guzheng, erhu, and bamboo flute — and documented significantly different antidepressant effects across instruments, with the guzheng outperforming the other two [22]. Kwon and colleagues, writing from the vantage of Korean medicine and biomedical informatics, propose an artificial intelligence-driven acoustic decomposition of the five-element stimulus as the remedy, pushing descriptive precision into finer-grained modal-acoustic features [23]. Domestic Chinese reviewers have noted that the absence of a unified prescriptive system, the variability of musical selections across studies, and the lack of standardized acoustic descriptors are now widely acknowledged limitations of the field [24][25]. The two strands of internal critique converge on the same conclusion from opposite directions. Yao's instrument-comparison result implies that the mode label is not doing all the therapeutic work; Kwon's AI proposal implies that the recorded stimulus needs finer description. Both, however, remain inside the listening paradigm: the patient is still the recipient, the music is still the stimulus, and refinement is sought in the upstream specification of what the patient hears. The possibility that the operative variable might lie at a register the listening paradigm cannot describe — at the level of what a body does when it produces the note — is not foreclosed by either Yao or Kwon. It is simply not in their analytical vocabulary. Recovering that vocabulary is what the next section undertakes.

3. Returning to the Sources: The Vocal-Articulatory Reading of the Five Notes

3.1 Wu Yin and Wu Sheng in Pre-Qin and Han Texts

A philological complication that bears directly on the foregoing arises from a careful reading of the early sources. The classical Chinese tradition distinguishes between wu yin and wu sheng in ways the modern conflation "five notes" tends to obscure. Tai and colleagues, in a 2023 analysis published in the China Journal of Traditional Chinese Medicine and Pharmacy, document that the Mengzi already provides a vocal-articulatory account of the five sounds: "issuing from the spleen, the lips closed and resonant, this is called gong; issuing from the lung, the mouth opened and expelled, this is called shang; issuing from the liver, the mouth opened and the lips swelling, this is called jiao; issuing from the heart, the teeth closed and the lips opened, this is called zhi; issuing from the kidney, the teeth opened and the lips pursed, this is called yu" [26]. This is unambiguously a third description — neither pitch frequency nor modal scale, but vocal configuration: where in the

mouth the sound is formed, what the tongue and lips do, what posture the jaw assumes. The same passage is preserved with slight variation in the Liji Yueji and in later medical compilations, and the consistency across these transmissions is itself evidence that the articulation-based reading was a stable strand of the tradition rather than a marginal aberration [27][28]. Jiang and Liu, in a 2024 paper in the same journal, push the analysis further: they argue that the Huangdi Neijing's own use of wu sheng to encompass the five vocal expressions hu (call), xiao (laugh), ge (sing), ku (cry), and shen (groan) places voice, not mode, at the core of the text's auditory schema [29]. Two independent lines of evidence therefore converge. The Mengzi specifies five vocal articulations and assigns each to a viscera. The Neijing's own definition of wu sheng centers on five vocal acts, not five musical pitches. The articulation-based reading is not a modern reconstruction imposed on the sources; it is the reading the sources themselves, read together, most readily support. The conflation of wu yin with five scalar modes is the late development; the vocal reading is the older one.

3.2 The Yizong Jinjian's Articulation-Site Scheme

The most systematic vocal-articulatory account of the five notes in the Chinese medical tradition is preserved in the Qing-dynasty Yizong Jinjian (Golden Mirror of the Medical Lineage), as later reconstructed and disseminated by Hu Jiexu in his Yinyue yu Baojian Yiliao [30]. Hu's reconstruction, which Zhang Yong's 2022 study in the Journal of Art Communication has carefully documented, gives each of the five notes a distinct site of articulation, characteristic acoustic quality, and ascribed functional effect [31]. Gong is described as "rising from beneath the tongue," with an open, grounded resonance; its acoustic quality is "extremely long, extremely low, extremely thick, with a sunken, expansive, robust quality," and it is associated with the spleen and with nourishment of qi and blood. Shang is described as "issuing with an opened jaw, the sound coming from the mouth"; its quality is "next-long, low and thick, with a clear, ringing, restrained quality," and it is associated with the lung and the strengthening of respiratory function. Jiao is described as "formed with the tongue drawn back at the midline"; its quality is "of moderate length and pitch, clear and balanced," and it is associated with the liver and with the smooth movement of qi. Zhi is described as "formed with the tongue touching the teeth"; its quality is "next-short, next-high, next-clear, with an inflected, lingering quality," and it is associated with the heart and with the elevation of neural function. Yu is described as "formed with the lips pursed, the sound issuing from the lips"; its quality is "extremely short, extremely high, extremely clear, with a delicate, fine quality," and it is associated with the kidney and with mental vitality. Hu drew further on Zheng Qiao's Tongzhi: Qi Yin Lüe for the supporting classification of the five notes as labial, lingual, dental, guttural, and velar — a phonological classification that confirms the vocal-articulatory rather than modal nature of the source description [32]. Three features of this scheme deserve to be marked. The description proceeds from articulation site outward to acoustic quality and then to physiological effect — the order in which a singing teacher actually thinks. The five articulation sites describe a progression from posterior-low to anterior-high in the vocal tract, an ordering that corresponds to a real anatomical gradient and not to an arbitrary alignment with the five elements. The ascribed physiological effects are, in each case, plausible consequences of the articulatory specification: the open low resonance of gong does engage the abdominal musculature in ways that affect digestive support; the lip-pursed high resonance of yu does demand fine breath control of the kind associated with mental concentration. The Yizong Jinjian's account is not a forced cosmological scheme tacked onto vocal observation. It is a description of vocal practice an attentive teacher could have written, with a precision the modal reading has never matched.

3.3 Engaging the Philological Critique

The most significant recent critique of FEMT, advanced by Song Boyuan in a 2016 paper in Chinese Musicology, addresses the modal reading of the five notes and finds it textually untenable [6]. Song's argument has four main components: that the Neijing lacks the concept of "modal tonic" that contemporary FEMT practitioners presume; that wu yin and wu sheng in pre-Qin sources denote pitch-names or pitch generally rather than pentatonic modes; that the Neijing's use of wu sheng to encompass the five vocal sounds (hu, xiao, ge, ku, shen) places voice at the core of the text's auditory schema; and that the mathematical order of modal generation by the sanfen sunyi (three-fold subtraction-addition) method does not align with the five-element generative cycle, so that no coherent element-mode therapeutics can be read off the text. Song's critique is sharp, and on the terms in which it is pitched — the interpretation of the five notes as five scalar modes — it is largely correct. The text-historical literature appearing since 2016 has, on balance, accepted the central thrust of Song's argument [29][33]. The reading proposed in this article begins, however, from an asymmetry in Song's own evidence that the FEMT response to her critique has not adequately registered. Three of Song's four arguments — the absence of modal tonic, the non-modal meaning of wu yin and wu sheng in pre-Qin sources, and the mismatch between sanfen sunyi and five-element ordering — are arguments against reading the five notes as scalar modes. The fourth — that wu sheng centrally encompasses the vocal expressions — is an argument for reading the five notes as vocal acts. Song's critique therefore does two things at once that are too easily collapsed into one: it dissolves the modal reading and, at the same time, points toward the vocal reading that the modal reading had displaced. The Mengzi passage and the Yizong Jinjian's articulation-site scheme described in sections 3.1 and 3.2 are not vulnerable to Song's argument. They do not rest on the concept of modal tonic, do not depend on identifying wu yin with five scalar modes, and do not require the order of modal generation to match the five-element cycle. They identify each of the five notes with a specific configuration of tongue, jaw, lip, and breath. The therapeutic specificity these descriptions name resides in the production of sound by a body, not in its reception by an ear. Song's critique, on this reading, does not undermine FEMT's classical foundations so much as relocate them: it tells us where the foundations cannot be — in the modal scheme — and the textual evidence Song herself adduces points to where they can be. The contemporary FEMT subfield has been defending the classical tradition on the philologically weakest available terms while a stronger reading lay in the same archive untouched.

4. The Parallel Literature: International Research on Singing as a Therapeutic Practice

4.1 The Sing Well Program and Its Companion Initiatives

While the FEMT subfield has continued, by and large, within the listening paradigm, an international research program on singing as therapeutic practice has taken shape over the past half-decade with little institutional overlap. The Sing Well initiative, led by Russo and Good at Toronto Metropolitan University and launched in 2021, is a seven-year cross-institutional program spanning aphasia, Parkinson's disease, hearing loss, stuttering, respiratory disorders, and second-language learning [7]. Its institutional model — a network of conservatories, clinical research centers, and community singing organizations — was deliberately constructed to support clinically rigorous voicing interventions on a scale no single institution could mount. Adjacent work by Loutrari and Georgiadou on adapted Melodic Intonation Therapy for voice feminization in transgender women extends the same sensibility into further populations [34]. The bibliometric and review literature on music therapy as a whole documents a steady growth in singing-based interventions over the past two decades and an increasing methodological sophistication in their evaluation [35]. What unites

these initiatives is a shared assumption the FEMT subfield has not made: that the therapeutic effect resides in what the patient does with the voice rather than in what the patient hears. The two literatures have been working on the same physiological substrate from opposite framings. Sing Well asks what happens when patients sing; FEMT asks what happens when patients listen to recordings derived from a singing tradition. The reading of the classical sources advanced in section 3 suggests these two framings are not equally faithful to the textual record FEMT claims to inherit; the framing more faithful to the record is the one Sing Well has independently arrived at on grounds owing nothing to the Yizong Jinjian.

4.2 Melodic Intonation Therapy and the Neuroplasticity of Sung Phonation

The strongest existing demonstration that produced singing — not heard music — remodels the nervous system comes from the literature on Melodic Intonation Therapy (MIT) for non-fluent aphasia. MIT asks patients with Broca's-type aphasia, who have lost fluent speech, to intone short phrases on simple melodic contours, gradually fading the melody until ordinary speech returns. The much-cited study by Schlaug and colleagues found that an intensive course of approximately seventy-five to eighty MIT sessions produced measurable increases in the fiber number and volume of the right-hemisphere arcuate fasciculus in chronic aphasia patients, suggesting that the act of sung phonation had driven structural change in a white-matter tract known to subserve speech production [8]. More recent work has refined and partially revised this picture: Sihvonen and colleagues' 2024 study in eNeuro on structural neuroplasticity effects of singing in chronic aphasia documents broader white-matter changes than the original Schlaug findings suggested, and Pitkäniemi's 2024 doctoral work proposes that preserved singing in Broca-type cases may rely on left-hemisphere ventral-stream pathways rather than on the right-hemisphere compensation Schlaug's work had suggested [36][37]. The specific neuroanatomy continues to be revised. The robust and older finding — that repeated sustained phonation measurably reorganizes the nervous system of the person producing it — has only become more secure with each successive study. The interpretive habit of the field has been to frame these results as evidence that MIT works because it remodels specific white-matter tracts. The order of explanation, on closer inspection, runs the other way. MIT works because it forces a specific kind of sustained vocal production; the white-matter remodeling is the neural sediment of that production. Read in this direction, the implication for music therapy research is substantial: outcome measures capturing the bodily quality of phonation — breath length, pitch stability, resonance placement — should become first-class endpoints alongside the imaging that has so far carried most of the explanatory burden. Every measured brain change in this literature is the trace, accumulated across dozens of sessions, of a patient sustaining phonation against a held breath, shaping a vowel, placing a consonant — that is, of a patient doing, in clinically simplified form, the work a singer does in a voice lesson. The MIT literature provides the proof of principle for the therapeutic specificity of voicing as distinct from hearing. The FEMT subfield has not yet drawn on it.

4.3 Choral Singing, Breath, and Autonomic Synchronization

A third strand of international research has examined the physiological and affective consequences of singing in groups. Vickhoff and colleagues' 2013 study, now widely cited in the field, demonstrated that when singers share a song the length of the musical phrase entrains their respiration, and through respiration their cardiac rhythms — a phenomenon the authors explicitly linked to Porges's polyvagal theory of social engagement [9][38]. Subsequent work by Tatschl and Schwerdtfeger has shown that paced singing at approximately 0.1 Hz — corresponding to phrase lengths of roughly six seconds — produces cardiovascular signatures including elevated vagally-

mediated heart rate variability alongside elevated heart rate and blood pressure, indicating a coordinated autonomic response with both parasympathetic and sympathetic components [39]. Fancourt and colleagues, in a multicenter study of cancer patients, bereaved carers, and caregivers participating in five Welsh community choirs, documented that one hour of group singing was followed by acute reductions in salivary cortisol alongside increases in mood and reductions in measured stress [40]. Schladt and colleagues directly compared choral singing with solo singing and found that twenty minutes of either was sufficient to increase happiness and decrease sadness and worry, with the two conditions producing distinguishable hormonal profiles in oxytocin and cortisol [41]. Earlier work by Grape and colleagues had observed that even a single one-on-one singing lesson with a teacher produced measurable shifts in plasma oxytocin and cortisol alongside subjective reports of greater joy and energy [42]. The most rigorous recent test of choral singing as an intervention for affective health is Ng and colleagues' two-year randomized controlled trial in older Singaporean adults, published in 2025, which compared weekly choral singing against a deliberately matched health-education active comparator and found no significant difference between the two arms in depression and anxiety screening positivity at one and two years [43]. The Ng result has been read in the field as a setback for choral interventions; the present article reads it differently. The trial defined "choral singing" at the level of group musical activity and matched the active comparator on the non-specific components long suspected of carrying part of the apparent effect in earlier weakly-controlled trials. What the design did not isolate was any third element of the choral activity: the breath patterns sustained, the resonance configurations engaged, the articulatory work undertaken before any phrase was sung. A choir whose director devotes fifteen minutes per session to coordinated breath training is doing something somatically distinct from a choir whose director leaps directly into repertoire; nothing in Ng's published methodology rules out that the participating choirs varied substantially on this dimension and that the effects were averaged out as a result. The null finding is not evidence that choral singing does not work. It is evidence that "choral singing" as currently operationalized is too coarse a category to isolate the variable doing the work. Ng and colleagues have, paradoxically, advanced the field more than a positive finding would have: they have set the methodological bar at which a more interesting question becomes available.

5. A Conceptual Bridge: Appoggio, Qi Chen Dan Tian, and Embodied Singing

5.1 The Italian Appoggio Tradition

The Italian bel canto tradition of vocal pedagogy developed, over roughly three centuries, a coordinated breathing technique it calls appoggio. The word means "to lean" or "to rest upon," and the technique refers to the active maintenance of the inspiratory posture of the ribcage during sung exhalation, so breath is released slowly against continued engagement of the inspiratory musculature, rather than collapsing under the elastic recoil of the lungs. The canonical modern exposition by Richard Miller in The Structure of Singing systematizes a tradition reaching back to the Italian voice schools of the seventeenth and eighteenth centuries; the more recent reflective account by Janice Chapman has carried the tradition into a contemporary holistic framework [44][45]. Trained singers under appoggio sustain phrase exhalation between four and twelve seconds, sometimes longer; their respiration during sustained vocal work drops from a resting rate of twelve to sixteen breaths per minute to four or six. This respiratory geometry is the same geometry that Tatschl and Schwerdtfeger have shown produces the autonomic signature of paced singing at 0.1 Hz, and the same geometry Vickhoff has shown entrains cardiac rhythms across choristers [38][39]. The convergence is not coincidental. Bel canto pedagogy was developed by working voice teachers, over generations, by trial and error against the constraint of having to

sustain audible musical lines without strain. The constraint they were optimizing against turns out to be the same constraint producing the physiological signatures the autonomic literature now measures. Appoggio is, in this sense, an empirically discoverable bridge between voice pedagogy and autonomic physiology — a bridge constructed centuries before the physiology was characterized, visible as a bridge only in retrospect once both sides became measurable.

5.2 The Chinese Qi Chen Dan Tian Tradition and Embodied Singing Theory

The comparable Chinese description — qi chen dan tian, "breath sunk to the cinnabar field" — points to the same anatomical coordination, with the rhetorical weight placed on the destination of the inhalation rather than on the mechanics of the exhalation. The qi chen dan tian tradition has been embedded for centuries in opera training, in qigong practice, and in temple chant. It shares with appoggio both an attention to lower abdominal engagement and a concern for sustaining breath against a resistance maintained by inspiratory musculature. In 2025, Wang Haitao and Liu Hong proposed a framework they call shenti gechang ("embodied singing"), according to which singing is to be understood as the coordinated interaction (which they call gongge, "conjugation") of approximately twelve bodily subsystems: phonatory apparatus, resonating cavities, singing breath, singing proprioception, singing auditory perception, singing visual perception, posture and expression, singing psychology, singing synesthesia, external environment, central nervous system oversight, and the inherited physiological substrate of the individual singer. Wang and Liu's framework is philosophical in orientation rather than clinical, and explicitly aims to provide a theoretical grounding for music therapy practice in China. The framework's central claim — that singing must be studied as a full-body coordinated event rather than a vocal-tract event incidentally accompanied by respiration — does in contemporary medical-philosophical idiom what qi chen dan tian had encoded in pedagogical practice and what appoggio had encoded in the parallel European register. Three traditions, three vocabularies, one underlying account: the singing body as an integrated coordinated system in which breath support is anatomically prior to articulation, and in which the work the system does on itself is more than the sum of its parts.

5.3 The Convergent Anatomical Zone and What It Suggests

The Italian appoggio tradition, the Chinese qi chen dan tian tradition, and the recent theory of embodied singing converge on the same zone of the lower ribcage and upper abdomen as the operative site of sustained singing. The convergence is striking precisely because the three traditions developed independently, in different cultural and intellectual contexts, and arrived at intersecting practical knowledge none of them framed in the conceptual vocabulary of the others. Appoggio speaks the language of vocal craft. Qi chen dan tian speaks the language of qi cultivation. Embodied singing theory speaks the language of contemporary medical philosophy. Each, in its own idiom, identifies the same anatomical region as the support point on which sustained singing must rest. For the FEMT reframing this article advances, the convergence has a specific consequence. The Yizong Jinjian's articulation-site scheme described in section 3.2 specifies the upper end of the vocal tract — tongue, jaw, lips — for each of the five notes; the appoggio and qi chen dan tian traditions specify the lower end — the breath support that makes sustained articulation possible. Together they describe a coordinated body engaged in the production of sound, and the body they describe is exactly the body the listening paradigm of contemporary FEMT does not engage. A music therapist trained in either the appoggio tradition or the qi chen dan tian tradition possesses, in working knowledge, the technical resources the FEMT subfield has not yet learned to draw upon. A FEMT researcher trained on the Neijing's emotional-visceral correspondences possesses the conceptual framework the international singing-as-therapy literature

has not yet had occasion to engage with in detail. The two are complementary halves of a single account no single tradition has yet articulated as a whole. This is the cross-cultural conceptual bridge the present article offers as its distinct contribution. The bridge requires only that the upper-vocal-tract specifications of the Yizong Jinjian and the lower-thoracic specifications of appoggio and qi chen dan tian be read as describing the same body, and that the international research program on singing as therapeutic practice be read as the empirical complement to a classical Chinese textual archive that has been waiting, undisturbed, for it.

6. Discussion: Implications for Music Scholarship and Therapy Practice

6.1 For Voice Pedagogy and Music Education

The implications for voice pedagogy and music education are conservative in substance and ambitious in framing. Voice teachers across cultures already do most of what a vocal-articulatory reading of the five notes would ask for: they shape breath, place resonance, refine articulation in a daily practice whose accumulated craft knowledge has rarely been translated into the vocabulary of clinical research. What this article offers is a reason for that translation to be attempted. A music conservatory taking the cross-cultural conceptual bridge seriously can offer students an articulation of vocal craft engaging both the Italian and the Chinese traditions in their full technical specificity, rather than treating one as foreign reference and the other as cultural inheritance. Voice teachers gain a body of clinical and physiological evidence establishing their work as therapeutically consequential beyond the artistic context — Grape and colleagues' single-lesson study produced measurable hormonal changes in one session [42], a finding that has not yet been adequately registered in conservatory pedagogy as an argument for what voice teachers do.

6.2 For Music Therapy Practice

The implications for music therapy practice are more substantial. A music therapist who decides to engage the body that produces sound — rather than only the body that hears it — is choosing to draw on a body of craft knowledge her training likely did not include. The Italian appoggio tradition's three centuries of pedagogical refinement, and the much older Chinese qi chen dan tian tradition embedded in opera and qigong practice, together constitute a working anatomy of sustained vocal production no contemporary clinical training program has fully absorbed. The Sing Well program offers one model of cross-disciplinary collaboration in the international context [7]; what is currently lacking is a parallel structure connecting Chinese voice pedagogy traditions and Chinese-medical music therapy practice, which the conceptual bridge described in section 5 could in principle support. In the FEMT-specific context, the most concrete implication is that protocol descriptions in clinical trials need to specify the vocal-articulatory dimension of the intervention — what the patient is doing with the voice, not only what the patient is hearing — even when the intervention itself remains primarily receptive. Without such specification, the field cannot distinguish the variable doing the work from the recording delivering it.

6.3 Future Research Directions

The reframing advanced here makes available three classes of empirical question the listening paradigm by itself does not generate. The first concerns the differential physiological correlates of producing the five notes under the Yizong Jinjian's articulatory specifications, which can be examined with breath-analysis, electromyography, and heart rate variability instrumentation already standard in the broader singing-and-physiology literature [38][39]. The second concerns head-to-

head comparisons of vocally-produced versus listened-to FEMT interventions in the same clinical population, which would establish more precisely the differential contribution of the production component beyond the shared musical content. The third concerns the third-stratum specification of choral singing interventions in the lineage of the Ng et al. trial: future trials should compare choral arms differing on the presence or absence of structured breath-articulation training while otherwise matched on conductor experience, repertoire selection, session length, and group composition [43]. None of these directions is original to this article in its specific empirical components. What the article makes available is their joint formulation as crisp empirical hypotheses, possible once FEMT and the international singing-as-therapy literature are read as a single conceptual landscape rather than two separate research programs.

7. Conclusion

This article has argued that contemporary Five-Element Music Therapy has been resting on the philologically weakest reading available of its own classical foundations. The Mengzi, the Tongzhi: Qi Yin Lüe, and the Yizong Jinjian describe the five notes as vocal-articulatory configurations, not scalar modes. Song Boyuan's 2016 critique dissolves the modal reading and, on the textual evidence Song herself adduces, points toward the vocal-articulatory reading the modal reading had displaced. The international research program on singing as therapeutic practice — Sing Well, the MIT neuroplasticity literature, the choral physiology trials of Vickhoff, Fancourt, and others — has independently arrived at the framing more faithful to FEMT's own classical sources without engaging with those sources at all. The Italian appoggio tradition, the Chinese qi chen dan tian tradition, and the recent theory of embodied singing converge on a common anatomical zone of breath support, providing a cross-cultural conceptual bridge neither tradition alone supplies. The contribution advanced here is the philologically grounded reframing of the five notes as vocal-articulatory configurations, the parallel placement of two literatures developed in isolation, and the cross-cultural bridge connecting three pedagogical-philosophical traditions as describing the same singing body. Three limitations should be marked. The article presents no original empirical data. The vocal-articulatory specifications recoverable from the Yizong Jinjian were developed for trained singers and may require adaptation for clinical populations whose vocal capacity is reduced by their condition. The cross-cultural bridge as described privileges the European and Chinese traditions and does not engage other vocal pedagogical traditions whose engagement would enrich the picture. Each marks a place where further work is needed. The reframing is on the table.

References

[1] Yang T, Wang J, Wang R, Yang Y, Liu Y, Zhang C. Effects of Five-Element Music on Language Recovery in Patients with Poststroke Aphasia: A Systematic Review and Meta-Analysis. Journal of Alternative and Complementary Medicine. 2019; 25(10): 983-991.

[2] Yang T, et al. Effectiveness of Five-Element Music Therapy in Cancer Patients: A Systematic Review and Meta-Analysis. Complementary Therapies in Clinical Practice. 2021; 44: 101416.

[3] Du QQ, Bai L. Research Progress of Traditional Chinese Music Therapy in China. Modernization of Traditional Chinese Medicine and Materia Medica. 2020; 22(8): 2876-2881.

[4] Ma Y, Liu MM, Gao SH. Discussion of TCM Music Therapy Based on the Five-Note Theory of the Huangdi Neijing. China Journal of Traditional Chinese Medicine and Pharmacy. 2014; 29(5): 1294-1297.

[5] Hu JX, Music and Health Medicine. Beijing: China Federation of Literary and Art Circles Publishing House; 2004.

[6] Song BY. A Critique of Five-Element Music Therapy. Chinese Musicology. 2016; (2): 40-43.

[7] Russo FA, Good AJ. Sing Well: An International Research Program on Singing-Based Interventions. Toronto Metropolitan University; launched 2021. Available from: https://singwell.ca [Accessed 15th April 2026].

[8] Schlaug G, Marchina S, Norton A. Evidence for Plasticity in White-Matter Tracts of Patients with Chronic Broca's Aphasia Undergoing Intense Intonation-Based Speech Therapy. Annals of the New York Academy of Sciences. 2009;

1169: 385-394.
[9] Vickhoff B, Malmgren H, Astrom R, Nyberg G, Ekstrom SR, Engwall M, Snygg J, Nilsson M, Jornsten R. Music Structure Determines Heart Rate Variability of Singers. Frontiers in Psychology. 2013; 4: 334.
[10] Lin FC, He NN, Huang DH. Primary Exploration on Music Therapy Based on Five Elements in Inner Canon of Huangdi. China Journal of Traditional Chinese Medicine and Pharmacy. 2015; 30(11): 4161-4162.
[11] Tai XL, Liu XY, Pei F, Cui YD, Zhang YX, Huang XJ, Zhang WJ, Liang YX. Brief Analysis of the Five Yin and Five Sheng in Huangdi Neijing. China Journal of Traditional Chinese Medicine and Pharmacy. 2023; 38(12): 5695-5698.
[12] Lin FC. Mechanism of Five-Notes Therapy Based on Meridian Theory. Beijing University of Chinese Medicine Journal. 2014; 37(7): 466-469.
[13] Zhang XF, Liao LH. Application Research of Music Therapy Based on Five-Zang Sound Correspondence in TCM Diagnosis and Treatment. China Journal of Traditional Chinese Medicine and Pharmacy. 2020; 35(7): 3726-3729.
[14] Zhang Y. An Exploration of the Thought-System of China's Traditional Five-Note Therapy. Journal of Art Communication. 2022; (2): 46-58.
[15] Cai ZX. The History and Current Status of Music Therapy for Emotional Diseases. Chinese Journal of Information on Traditional Chinese Medicine. 2014; 21(7): 124-127.
[16] Fan XX, Zhao ST, Yong LL, Zhai ZW, Li Y, Zhong YM, Wen CB. Clinical Research Progress of Insomnia Based on the Theory of Five Tones in Traditional Chinese Medicine and Analysis of Problems and Strategies. Modernization of Traditional Chinese Medicine and Materia Medica. 2022; 24(11): 4304-4309.
[17] Li XH, et al. Research Progress of Traditional Chinese Medicine External Treatment for Depression in Recent Ten Years. Chinese Journal of Integrated Traditional and Western Medicine. 2024; 44(5): 612-617.
[18] ClinicalTrials.gov. NCT00964314, Prospective Study with RCT Designed for Evaluation of TCM Five Elements Music Therapy Improving Quality of Life in Patients with Advanced Cancer; and NCT07048938, Clinical Effect and Mechanism Study of Five Elements Music Therapy on Depression Disorder. Available from: https://clinicaltrials.gov [Accessed 15th April 2026].
[19] Zhao LZ, et al. Research Progress of Traditional Chinese Medicine Five-Tone Therapy. Chinese Journal of Traditional Chinese Medicine and Pharmacy. 2016; 31(11): 4666-4669.
[20] Zheng ZW. Origin Exploration of Five-Tone Establishing Movement. Liaoning Journal of Traditional Chinese Medicine. 2024; 26(1): 1-4.
[21] Wang F. Research on Combining Modern Electronic Music Technology with TCM Five-Tone Therapy in Clinical Application. Chinese Folk Therapy. 2024; 32(15): 88-92.
[22] Yao XY, Xu C, Yang LL, Wu AL, Xiong L. An Empirical Study on the Intervention of Traditional Chinese Medicine Five-Element Music in Alleviating Anxiety and Depression Among Medical Students: Differences in Effectiveness Among Three Different Musical Instruments. Frontiers in Psychology. 2025; 16: 1625325.
[23] Kwon N, Leem J, Park S, et al. The Modernization of Oriental Music Therapy: Five-Element Music Therapy Combined with Artificial Intelligence. Healthcare. 2024; 12(3): 411.
[24] Xu JZ, Qiao XC, Shi YJ, Li YM. To Define the Chinese Ancient Standard Pitch by Resonate Viscera and Collateral Channels. (Journal of Wuhan Conservatory of Music). 2010; (4): 145-150.
[25] Hou Y. Basic Methods and Operations of Music Therapy. (Journal of Wuhan Conservatory of Music). 2013; (2): 167-172.
[26] Mengzi., Book of Rites: Record of Music. In: Tai XL, Liu XY, Pei F, et al. Brief Analysis of the Five Yin and Five Sheng in Huangdi Neijing. China Journal of Traditional Chinese Medicine and Pharmacy. 2023; 38(12): 5695-5698.
[27] Wang ST. Research on Traditional Chinese Music Therapy in the Context of Cultural Anthropology. Ph.D. Thesis, Nanjing University of Chinese Medicine, Nanjing; 2020.
[28] Zheng Q. Tongzhi: Seven Notes Summary (Comprehensive Treatise: Brief on the Seven Sounds). As cited and reconstructed in Hu JX, Beijing: China Federation of Literary and Art Circles Publishing House; 2004.
[29] Jiang X, Liu Y. Analysis of Five-Tones and Treatment of Disease with Five-Tones. China Journal of Traditional Chinese Medicine and Pharmacy. 2024; 39(7): 3779-3782.
[30] Wu Q, et al. Golden Mirror of the Medical Lineage. Qing Dynasty compilation; modern edition Beijing: People's Medical Publishing House; 2006.
[31] Zhang Y. An Exploration of the Thought-System of China's Traditional Five-Note Therapy. Journal of Art Communication. 2022; (2): 46-58.
[32] Zheng ZW. Origin Exploration of Five-Tone Establishing Movement. Liaoning Journal of Traditional Chinese Medicine. 2024; 26(1): 1-4.
[33] Loutrari K, Georgiadou E. Adapted Melodic Intonation Therapy for Voice Feminization in Transgender Women: A Preliminary Investigation. Journal of Voice. 2024; advance online publication.
[34] Li K, Weng LM, Wang XQ. The State of Music Therapy Studies in the Past 20 Years: A Bibliometric Analysis. Frontiers in Psychology. 2021; 12: 697726.
[35] Sihvonen AJ, et al. Structural Neuroplasticity Effects of Singing in Chronic Aphasia. eNeuro. 2024; 11(5):

ENEURO.0408-23.2024.
[36] Pitkaniemi A. Neural Correlates of Preserved Singing Ability in Broca-Type Aphasia: Evidence for Left-Hemisphere Ventral-Stream Involvement. Ph.D. Thesis, University of Helsinki, Helsinki; 2024.
[37] Porges SW. The Polyvagal Theory: Neurophysiological Foundations of Emotions, Attachment, Communication, and Self-Regulation. New York: Norton; 2011.
[38] Tatschl JM, Schwerdtfeger AR. Singing at 0.1 Hz as a Resonance Frequency Intervention to Reduce Cardiovascular Stress Reactivity? Frontiers in Psychiatry. 2022; 13: 876344.
[39] Fancourt D, Williamon A, Carvalho LA, Steptoe A, Dow R, Lewis I. Singing Modulates Mood, Stress, Cortisol, Cytokine and Neuropeptide Activity in Cancer Patients and Carers. E-cancer medical science. 2016; 10: 631.
[40] Schladt C, Nguyen GC, Tiliopoulos C, et al. Choir versus Solo Singing: Effects on Mood, and Salivary Oxytocin and Cortisol Concentrations. Frontiers in Human Neuroscience. 2017; 11: 430.
[41] Grape C, Sandgren M, Hansson LO, Ericson M, Theorell T. Does Singing Promote Well-Being? An Empirical Study of Professional and Amateur Singers during a Singing Lesson. Integrative Physiological and Behavioral Science. 2003; 38(1): 65-74.
[42] Ng TKS, Lim ZH, Todd M, Sun F, Ray K, Qi X, Guo J, Ye KX, Maier AB, Mahendran R, Goh LG, Tsakok M, Kua EH, Feng L. Effects of Choral Singing on Depression and Anxiety in Older Adults: A Randomized Controlled Trial. Journal of Psychiatric Research. 2025; 188: 162-168.
[43] Miller R. The Structure of Singing: System and Art in Vocal Technique. New York: Schirmer Books; 1986.
[44] Chapman J. Singing and Teaching Singing: A Holistic Approach to Classical Voice. 3rd ed. San Diego: Plural Publishing; 2017.
[45] Wang HT, Liu H. The Practical Application of Medical Body Philosophy: A Study of the Theory of Embodied Singing. Medicine and Philosophy. 2025; 46(5): 74-75.