A Multimodal Wearable Sensing System for Vocal Muscle Biofeedback in Singing Pitch Training

Kanyu Chen
KMD, Keio University
Yokohama, Japan
cady.cky@kmd.keio.ac.jp

Akira Kato
KMD, Keio University
Yokohama, Japan
kato@kmd.keio.ac.jp

Kai Kunze
KMD, Keio University
Tokyo, Japan
kai@kmd.keio.ac.jp

Abstract

We introduce a wearable system that uses EMG, ultrasound, and audio features to support vocal muscle biofeedback in singing pitch training. Through multi-phase evaluations with novice and expert singers, we demonstrate how physiological feedback enhances vocal control, reveals skill differences, and informs predictive models for real-time guidance. Our findings highlight the potential of muscle-acoustic sensing in embodied vocal learning.

CCS Concepts

• **Human-centered computing → Sound-based input / output**; **Visualization design and evaluation methods**; **Field studies**.

Keywords

physiology, emg, ultrasonography, microphone, singing training, biofeedback, user-centered design

ACM Reference Format:
Kanyu Chen, Akira Kato, and Kai Kunze. 2025. A Multimodal Wearable Sensing System for Vocal Muscle Biofeedback in Singing Pitch Training. In *Companion of the 2025 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp Companion '25), October 12–16, 2025, Espoo, Finland.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3714394.3750550

1 Problem Statement

Singing is a highly embodied skill that relies on complex coordination of internal physiological subsystems, including vocal fold muscles, breath control, and resonance shaping [48, 54]. Unlike many motor skills involving visible movement, the vocal muscles are hidden from conscious perception, making it difficult for singers—especially novices—to develop accurate muscle control during pitch production. Traditional singing instruction primarily relies on auditory feedback and verbal coaching [28, 63], which often fails to provide actionable cues on internal laryngeal adjustments and can overwhelm learners' cognitive resources [52].

While recent advances in biofeedback have demonstrated the value of physiological sensing for supporting embodied learning [18, 46], most existing systems focus on external motor tasks or affective regulation [6, 7, 24]. Real-time biofeedback for singing remains underexplored, particularly for capturing the fine-grained neuromuscular coordination involved in pitch control [14, 39, 60]. Metrics such as Electromyography (EMG) [4, 40], Ultrasonography (USG) [3, 33], and Singing Power Ratio (SPR) [58, 62] show promise for revealing vocal muscle activity and breath support. However, integrated multimodal sensing systems that combine these measures into wearable, real-time feedback interfaces for vocal skill acquisition remain scarce [38].

This research aims to address these gaps by developing multimodal wearable biofeedback systems that visualize internal vocal muscle activity and breath dynamics. Through real-time physiological feedback, the system seeks to make the "invisible" embodied processes of singing accessible to learners, enabling more effective training and refined control over vocal pitch. Specifically, we investigate whether EMG and USG can be used to capture different levels of proficiency in singing and support singers of various skill levels (RQ1). We further examine how EMG and USG perform in training novice singers (RQ2), and how a portable EMG setup may support professional singers in comparison to traditional feedback methods (RQ3). Finally, we explore how an optimized USG interface (RQ4), together with additional factors such as breath control, can enhance vocal training for amateur singers.

2 Related Works

This work represents the first EMG and USG measurement techniques on singing pitch. In this section, we review prior work in vocal skill learning, muscle sensing technologies, and biofeedback interface design, which together inform the development of our multimodal feedback system for singing training.

2.1 Vocal Physiology and Pitch Control

Voice is a natural instrument in human singing. Widely adopted pedagogical frameworks — such as Estill Voice Training (EVT) [48], Speech Level Singing (SLS) [34], and Complete Vocal Technique (CVT) [51], emphasize elements like breath control, pitch accuracy, volume modulation, and rhythmic precision [63]. Among these, Vocal pitch control plays a particularly central role, often linked to the tension regulation in the vocal folds muscles [2, 30, 48, 54]. The thyroarytenoid and cricothyroid muscles play key roles in pitch production by adjusting vocal fold tension through movement of the cricoid and arytenoid cartilages [4, 15, 23, 54], which changes the fundamental frequency (F0) of the voice. Each note corresponds to a specific F0, but what gives it character is not just pitch - it's the way sound is shaped by the vocal tract. As air travels from the vocal folds to the lips, it resonates in the vocal tract, creating *Formants*

UbiComp Companion '25, Espoo, Finland

ACM ISBN 979-8-4007-1477-1/2025/10
https://doi.org/10.1145/3714394.3750550

- peaks in the sound spectrum shaped by the position of the lips, tongue, and other articulator [30, 49]. These *Formants* determine the timbre and perceived quality of the sound.

This is especially crucial in musical singing, where techniques such as belting [12] demand precise and sustained high-pitch production, supported by coordinated muscle engagement and subglottal pressure control involving dominant thyroarytenoid activity with limited cricothyroid elongation. These mechanisms form the theoretical basis for our focus on pitch muscle sensing.

Despite clear physiological mechanisms, a major challenge in vocal pedagogy remains: the **lack of accessible, real-time feedback mechanisms** that reflect muscle activity [63]. In the absence of continuous expert coaching, most learners relies on auditory or spectrographic evaluation [28, 32, 37], which may overlook subtle muscle misalignments. Studies such as Yiu et al. [64] show that muscle fatigue significantly affects EMG patterns during singing, reinforcing the importance of direct muscular feedback in pedagogy. Therefore, emerging trends in vocal training have begun to integrate technological innovations to offer more precise and informative feedback to learners research [63]. This leads to the next consideration: how vocal metrics can be used as effective input signals for vocal training systems.

2.2 Acoustic and Muscle Sensing in Vocal

Scientific interest in the physiological basis of professional singing dates back to Bartholomew's pioneering work in 1934 [1], which analyzed vibrato, pitch intensity, and the presence of high and low formants in professional vocal production. Building on these early insights, recent advances in acoustic analysis have enabled more detailed processing of complex vocal performances using microphone input [19]. Among the established acoustic metrics, the **Singing Power Ratio (SPR)** has proven effective for assessing vocal resonance and clarity—key qualities in classical and musical singing [29, 50]. SPR captures the ratio of high-frequency to low-frequency energy, typically between 2000–4000 Hz, and is strongly correlated with professional-level vocal projection [58, 62], creating Singer's Formant. Despite these advances, microphone-only methods are limited in their ability to reveal the internal muscular mechanisms that shape sound production.

Ultrasound stands out as a traditional yet potent modality for visualizing muscle movements, including those involved in vocal production [33, 36, 45]. Unlike laryngoscopy, which requires inserting equipment into the oral or nasal cavity, ultrasound offers a more accessible and repeatable option for examining laryngeal structures [3]. Clinically, laryngeal ultrasound has been applied to assess vocal fold mobility, diagnose vocal fold paralysis, and monitor recovery after surgery [33, 53]. Its ability to capture real-time vocal fold oscillation during phonation makes it particularly valuable for studying voice production [3].

In this research, we leverage ultrasonography (USG) not for diagnostic purposes, but as an interactive training tool that enables singers to visualize their own vocal fold movements in real time. However, applying ultrasound in training settings remains challenging, as real-time interpretation of laryngeal images requires clear landmark extraction, stable probe positioning, and robust signal processing [38]. Quantifying meaningful metrics such as vocal fold length, symmetry, and oscillation patterns often involves complex image tracking and segmentation algorithms. These technical challenges motivate our ongoing work to optimize real-time USG-based feedback interfaces for vocal training.

Recent advancements in wearable muscle sensing have enabled real-time, non-invasive monitoring of vocal muscle activity, introducing a new class of input for vocal training and assessment. Among these, **surface electromyography (sEMG)** has emerged as a prominent technique for measuring muscle engagement and coordination. EMG has been widely used in both music and motor skill learning to distinguish expertise levels and capture fine motor control [9, 14, 39]. Studies such as Pettersen et al. [39, 40] demonstrated that professional singers activate neck and respiratory muscles differently compared to students, particularly under high-pitch conditions. Such findings align with work by Visentin and Shan [60], who reviewed EMG's value for skill assessment, biofeedback-based learning, and injury prevention in music contexts.

Emerging technologies such as Electromyography (EMG) and ultrasonography have enabled researchers to examine vocal physiology in greater detail. Existing research has explored various facets of speech and vocal dynamics, including the correlation between uttering vowels or sentences [65] and the engagement of speech muscle, the use of pitch and EMG for omohyoid detection [61], analysis of face and neck muscle movements during sentence speech [42], ultrasonic techniques for capturing speech dynamics from tongue movement to sound production [25]. Certain wearable designs worn on the throat or chest offer a potential avenue for enhancing vocal performance, the development of a custom wearable collar [43] for enhancing vocal performance, and the design of breath-related soma for authentic vocal expression [10]. Despite significant progress, a fully integrated sensing system specifically tailored for detecting vocal pitch muscles remains absent. Nevertheless, the integration of EMG and ultrasound into practical, wearable systems for real-time vocal analysis is still an area ripe for exploration, with promising applications in both clinical settings and vocal pedagogy.

Therefore, integrating electromyography (EMG) and ultrasonography (USG) with acoustic metrics such as the Singing Power Ratio (SPR) through wearable sensing systems offers a promising approach to vocal training, especially in mobile, real-world settings. This dual-sensing method is particularly effective in collective learning group environments, where EMG visualizations remain reliable despite ambient sound, and can be further enhanced with advanced microphone-based audio processing to support on-the-go vocal learning.

2.3 Design Biofeedback in Dynamic Contexts

Effective biofeedback systems for vocal training need to convert EMG, USG and audio signals into **interpretable, actionable insights** — especially in dynamic contexts. Gamified voice interfaces, such as those supporting vocal therapy for Parkinson's patients [26], show how feedback and repetition can increase vocal loudness and engagement. Traditional microphone-based tools, such as *Celestia* [47] for amateurs and commercial platforms like *SingStar* [44]

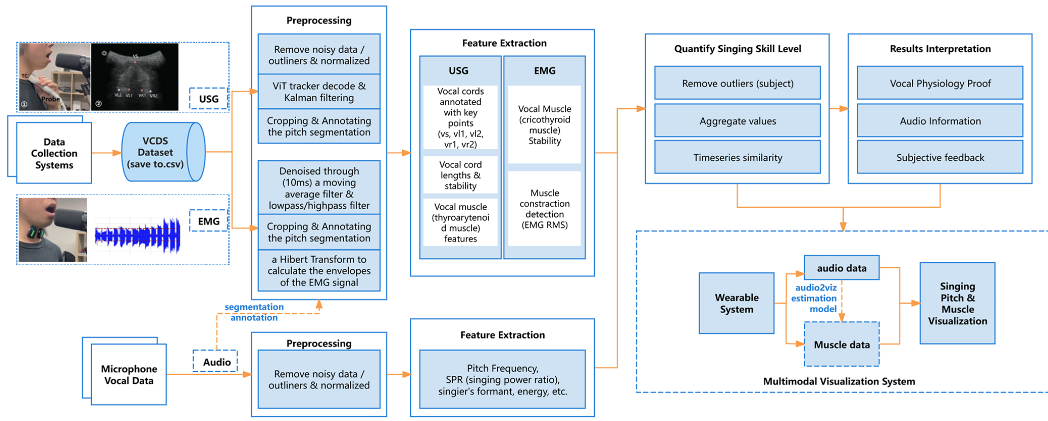

Figure 1: Proposed physiological data process method to quantify the singing skill using a wearable multimodal system.

for musician, offer real-time pitch-based scoring and visual feedback. While these systems improve user engagement and pitch awareness, they focus solely on acoustic output and offer **little visibility into the underlying muscular control** required for vocal technique.

In contrast, embodied and interactive music systems have shown strong potential for supporting physical awareness and expressive control. *BrainiBeats* [5] translates biosignals like EMG and EEG into generative musical output, linking emotion and performance. Installations like *The Music Room* [35], *The Throat III* [55], and *The Vocal Corder* [56, 57] enable users to shape vocal output through bodily gestures, highlighting how feedback can shape expressive control. Diaz [11] similarly argues that pitch-processing tools like Auto-Tune have redefined perceptions of vocal clarity and expression, suggesting that feedback technologies do more than measure — they influence what is considered correct or desirable in vocal performance.

Recent research has explored how electromyography (EMG) can support such feedback systems by providing real-time data on muscle activation. Systems like *FitBack* [20] and *MappEMG* [41, 59] deliver visual or haptic muscle feedback, enhancing posture and technique learning in sports and music. Devices such as BITalino allow these feedback methods to extend beyond laboratory settings into classrooms, home practice, or stage rehearsals. Karolus et al. [20] showed that EMG feedback improved posture awareness in non-expert users, while Gagnon et al. [14] found that expert and novice muscle coordination patterns could be distinguished via EMG sensing. Beyond voice, EMG has also enabled expressive control in instrument learning—such as pitch modulation in piano [21] and chord recognition in guitar [22] — demonstrating its versatility in motor-based musical interaction.

Overall, these systems demonstrate how muscle sensing and feedback can support **on-the-go, embodied learning** in real-world musical contexts—laying the groundwork for vocal-specific applications that integrate both physiological and acoustic feedback.

3 Methodology

This research adopts a Design Science Research (DSR) approach [16, 17] to guide the iterative development of a multimodal wearable system for vocal pitch training. The DSR framework integrates theoretical insights from vocal physiology, sensing technologies, and skill learning with the practical challenges of system design.

3.1 Design Science Framework

This research follows the Design Science Research (DSR) three-cycle model. The relevance cycle addresses the lack of internal muscle feedback in singing practice. The rigor cycle builds on studies in vocal physiology, sensing (EMG, USG, SPR), and embodied learning. The design cycle guides the system's development and evaluation to support vocal training.

3.2 Hardware Sensing Setup

We designed a multimodal sensing system targeting laryngeal muscle control. Surface electromyography (EMG) was captured using Delsys Trigno Wireless sensors positioned between the thyroid cartilage and the cricoid cartilage, with signals sampled at 2000 Hz and downsampled to 30 Hz for processing. Vocal fold dynamics were visualized using a CONTEC CMS600P2 B-mode ultrasound system, which recorded ultrasound images at 3.5 MHz and 30 frames per second. Synchronized high-quality audio was recorded using Shure SM7B or wireless microphones, depending on the experimental context. In addition, we plan to integrate wearable respiratory belts into the system to capture breath dynamics as an additional physiological parameter.

3.3 Data Processing and Feature Extraction

EMG Processing. Signals are denoised with a moving average filter (10 ms window), followed by Hilbert transform to extract the amplitude envelope. Muscle stability is quantified using shimmer-inspired metrics [13]:

$$s = \frac{1}{N-1}\sum_{t=1}^{N-1}\left|20\log\frac{A_{t+1}}{A_t}\right| \tag{1}$$

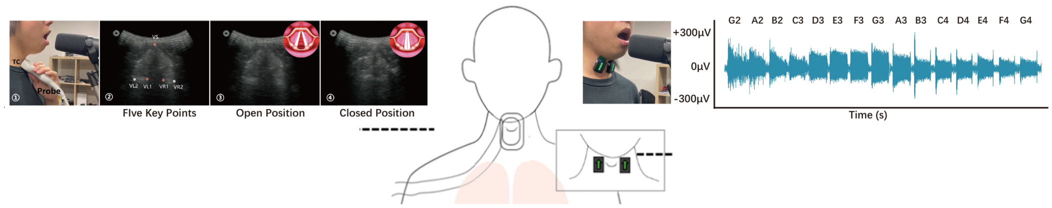

Figure 2: Left: EMG sensor placement and visualization of a sample of raw EMG data accompanied by muscle/cartilage position annotations. Right: Positioning of the ultrasonography probe and sample of raw ultrasound imaging data accompanied by muscle/cartilage position annotations.

where A_t is the EMG envelope at time t.

Ultrasound Processing. USG frames are manually annotated with five vocal fold landmarks following prior work [27]. Landmarks are automatically tracked using Kanade-Lucas-Tomasi (KLT) optical flow to compute vocal fold length dynamics.

Audio Feature Extraction. Singing Power Ratio (SPR), pitch (F0), MFCCs (Mel-frequency cepstral coefficients), and other spectral features are extracted for acoustic analysis and estimated modeling.

3.4 Research in Practice

3.4.1 Dataset Collection and Results (RQ1). We collected the **Vocal Cord Sensing Dataset (VCSD)** from 16 singers with varying skill levels to evaluate whether EMG and USG can distinguish novice and expert muscle activity. The dataset includes around 1.5 hours of synchronized EMG and USG recording across controlled pitch exercises. Repeated-measures ANOVA resulted that both EMG stability and USG vocal fold length differentiate singing proficiency levels [9].

3.4.2 Phase 1: Biofeedback for Novices (RQ2). We implemented real-time visual biofeedback interfaces based on expert reference data to guide novice singers (N=12) through vocal training sessions. Three conditions were compared: (1) Baseline (audio feedback only); (2) EMG feedback (muscle stability visualization); (3) USG feedback (vocal fold motion visualization).

The study revealed that EMG enhanced perceived controllability but induced cognitive load during real-time control. USG provided clearer visualization of vocal fold dynamics, facilitating vocal length control improvements, though users reported probe usability challenges.

3.4.3 Phase 2: Portable EMG-Microphone Setup for Experts (RQ3). To enable mobile training scenarios, we developed a portable EMG-microphone system evaluated with 16 professionally trained singers during live stage rehearsal. EMG muscle engagement and SPR vocal power showed strong correlations among expert performers, confirming that EMG-SPR sensing can capture nuanced vocal control in performance contexts [8].

3.4.4 Phase 3: Estimated Ultrasound Visualization and Breath Integration (RQ4 - Current Stage). We are developing predictive models to estimate ultrasound-based vocal fold motion from audio features (SPR, MFCC, F0) using MLP and Transformer-based neural networks. Additional data were collected from 9 expert singers. Building upon prior work [31], we plan to integrate breath sensing into the system pipeline to provide real-time multimodal feedback on breath-muscle coordination during singing.

4 Evaluation

We conducted multi-phase evaluations to test our system design. First, we used the VCSD dataset from 16 singers to confirm that EMG stability and vocal fold length (USG) could distinguish novices from experts [9]. In Phase 1, 12 novice singers tried real-time EMG and USG feedback. EMG helped with control but increased mental load, while USG improved vocal fold length control but required more effort to use. NASA-TLX showed both methods felt more demanding than audio-only training, especially USG. In Phase 2, 15 professional singers used a portable EMG-microphone system during rehearsals. Results showed experts had stronger EMG-SPR correlations, suggesting the system captures detailed muscle-acoustic coordination [8]. Currently, in Phase 3, we are training models to predict vocal fold movement from audio features (SPR, pitch, MFCC) using data from 9 experts, and testing the predictive interface with 20 amateur singers, including breath feedback integration.

5 Expected Contribution

This research contributes to both wearable vocal sensing and embodied skill learning. First, we propose and validate a novel multimodal wearable system that integrates EMG, ultrasound, and audio-based metrics for real-time vocal pitch training. Second, we also release a large dataset (VCSD) of synchronized muscle, vocal fold, and audio data from singers of different skill levels. Third, our iterative studies provide empirical evidence on how different sensing modalities affect vocal training, highlighting trade-offs between interface intuitiveness, controllability, and cognitive load. Fourth, we develop predictive ultrasound models using audio features to simplify sensing hardware while preserving physiological feedback. Finally, this work offers design insights into how biofeedback systems can support embodied skill learning, with potential implications for voice pedagogy, rehabilitation, and broader HCI applications involving internal motor processes.

Acknowledgments

This work is conducted under the IoT Accessibility Toolkit Project supported by JST Presto Grant Number JPMJPR2132 and partly funded by JST SPRING H09GQ24152.

References

[1] Wilmer T Bartholomew. 1934. A physical definition of "good voice-quality" in the male voice. *the Journal of the Acoustical Society of America* 5, 3_Supplement (1934), 224–224.

[2] Michel Belyk, Yune S Lee, and Steven Brown. 2018. How does human motor cortex regulate vocal pitch in singers? *Royal Society Open Science* 5, 8 (2018), 172208.

[3] Michael S Benninger. 2016. *Sataloff's Comprehensive Textbook of Otolaryngology Head and Neck Surgery: Laryngology.* Jaypee Brothers Medical Publishers.

[4] Fritz Buchthal. 1959. Electromyography of intrinsic laryngeal muscles. *Quarterly Journal of Experimental Physiology and Cognate Medical Sciences: Translation and Integration* 44, 2 (1959), 137–148.

[5] Caterina Ceccato, Ethel Pruss, Anita Vrins, Jos Prinsen, and Maryam Alimardani. 2023. BrainiBeats: A dual brain-computer interface for musical composition using inter-brain synchrony and emotional valence. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI EA '23)*. Association for Computing Machinery, New York, NY, USA, Article 56, 7 pages. doi:10.1145/3544549.3585910

[6] Kanyu Chen, Jiawen Han, Holger Baldauf, Ziyue Wang, Dunya Chen, Akira Kato, Jamie A Ward, and Kai Kunze. 2023. Affective Umbrella–A Wearable System to Visualize Heart and Electrodermal Activity, towards Emotion Regulation through Somaesthetic Appreciation. In *Proceedings of the Augmented Humans International Conference 2023*. 231–242.

[7] Kanyu Chen, Jiawen Han, George Chernyshov, Christopher Kim, Ismael Rasa, and Kai Kunze. 2021. Affective Umbrella–Towards a Novel Sensor Integrated Multimedia Platform Using Electrodermal and Heart Activity in an Umbrella Handle. In *Proceedings of the 20th International Conference on Mobile and Ubiquitous Multimedia*. 208–210.

[8] Kanyu Chen, Emiko Kamiyama, Ruiteng Li, Yichen Peng, Daichi Saito, Erwin Wu, Hideki Koike, and Akira Kato. 2024. Phantom Audition: Using the Visualization of Electromyography and Vocal Metrics as Tools in Singing Training. In *SIGGRAPH Asia 2024 Posters (SA '24)*. Association for Computing Machinery, New York, NY, USA, Article 113, 2 pages. doi:10.1145/3681756.3697908

[9] Kanyu Chen, Erwin Wu, Daichi Saito, Yichen Peng, Chen-Chieh Liao, Akira Kato, Hideki Koike, and Kai Kunze. 2024. Novel Sensing Methods for Vocal Technique Analysis: Evaluation on Electromyography and Ultrasonography. In *2024 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. IEEE, 121–125.

[10] Kelsey Cotton, Pedro Sanches, Vasiliki Tsaknaki, and Pavel Karpashevich. 2021. The Body Electric: A NIME designed through and with the somatic experience of singing. In *NIME 2021*. https://nime.pubpub.org/pub/ntm5kbux.

[11] Joe Diaz. 2009. The Fate of Auto-Tune. *Music and Technology* (2009).

[12] Jo Estill. 1988. Belting and classic voice quality: some physiological differences. *Medical problems of performing artists* 3, 1 (1988), 37–43.

[13] Mireia Farrús, Javier Hernando, and Pascual Ejarque. 2007. Jitter and shimmer measurements for speaker recognition. In *8th Annual Conference of the International Speech Communication Association; 2007 Aug. 27-31; Antwerp (Belgium).[place unknown]: ISCA; 2007. p. 778-81.* International Speech Communication Association (ISCA).

[14] Denis Gagnon, André Plamondon, and Christian Larivière. 2016. A biomechanical comparison between expert and novice manual materials handlers using a multi-joint EMG-assisted optimization musculoskeletal model of the lumbar spine. 49, 13 (2016), 2938–2945. doi:10.1016/j.jbiomech.2016.07.009

[15] Thomas Gay, Marshall Strome, Hajime Hirose, and Masayuki Sawashima. 1972. Electromyography of the intrinsic laryngeal muscles during phonation. *Annals of Otology, Rhinology & Laryngology* 81, 3 (1972), 401–409.

[16] Alan R Hevner. 2007. A three cycle view of design science research. *Scandinavian journal of information systems* 19, 2 (2007), 4.

[17] Alan R Hevner, Salvatore T March, Jinsoo Park, and Sudha Ram. 2004. Design science in information systems research. *MIS quarterly* (2004), 75–105.

[18] Kristina Hook. 2018. *Designing with the body: Somaesthetic interaction design.* MIt Press.

[19] Peng Huang, Yuan Li, and Zheng Wang. 2021. Enhanced waveform analysis for detecting vibrato and pitch modulation in vocal performances. *Journal of the Acoustical Society of America* 150, 3 (2021), 1753–1762.

[20] Jakob Karolus, Felix Bachmann, Thomas Kosch, Albrecht Schmidt, and Paweł W. Woźniak. 2021. Facilitating Bodily Insights Using Electromyography-Based Biofeedback during Physical Activity. In *Proceedings of the 23rd International Conference on Mobile Human-Computer Interaction* (New York, NY, USA, 2021-09-27) *(MobileHCI '21)*. Association for Computing Machinery, 1–15. doi:10.1145/3447526.3472027

[21] Jakob Karolus, Annika Kilian, Thomas Kosch, Albrecht Schmidt, and Paweł W. Wozniak. 2020. Hit the Thumb Jack! Using Electromyography to Augment the Piano Keyboard. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference* (Eindhoven, Netherlands) *(DIS '20)*. Association for Computing Machinery, New York, NY, USA, 429–440. doi:10.1145/3357236.3395500

[22] Jakob Karolus, Hendrik Schuff, Thomas Kosch, Paweł W. Wozniak, and Albrecht Schmidt. 2018. EMGuitar: Assisting Guitar Playing with Electromyography. In *Proceedings of the 2018 Designing Interactive Systems Conference* (Hong Kong, China) *(DIS '18)*. Association for Computing Machinery, New York, NY, USA, 651–655. doi:10.1145/3196709.3196803

[23] Gail B Kempster, Charles R Larson, and Michael K Kistler. 1988. Effects of electrical stimulation of cricothyroid and thyroarytenoid muscles on voice fundamental frequency. *Journal of Voice* 2, 3 (1988), 221–229.

[24] Misaki Kikuchi, Kanyu Chen, Rebecca Panskus, Dunya Chen, Keiko Okawa, and Kai Kunze. 2025. A Wearable System to Visualize Singing Breath towards Musical Perception in MR Context. In *Proceedings of Mensch und Computer 2025 (MuC '25)*. ACM, Chemnitz, Germany. doi:10.1145/3743049.3743054 ACM Conference Publication.

[25] Naoki Kimura, Michinari Kono, and Jun Rekimoto. 2019. SottoVoce: An ultrasound imaging-based silent speech interaction using deep neural networks. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–11.

[26] Markus Krause, Jan Smeddinck, and Ronald Meyer. 2013. A digital game to support voice treatment for parkinson's disease. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems* (Paris, France) *(CHI EA '13)*. Association for Computing Machinery, New York, NY, USA, 445–450. doi:10.1145/2468356.2468435

[27] Amarjeet Kumar, Chandni Sinha, Akhilesh Kumar Singh, and Umesh Kumar Bhadani. 2017. Vocal cord dysfunction: Ultrasonography-aided diagnosis during routine airway examination. *Saudi Journal of Anaesthesia* 11, 3 (2017), 370–371.

[28] Pauline Larrouy-Maestri, Yohana Lévêque, Daniele Schön, Antoine Giovanni, and Dominique Morsomme. 2013. The evaluation of singing voice accuracy: A comparison between subjective and objective methods. *Journal of Voice* 27, 2 (2013), 259–e1.

[29] Jihun Lee, Hyun Kim, and Soyeon Park. 2022. Real-time feedback system for classical singing training using singing power ratio. *Journal of Voice* 36, 1 (2022), 129–136.

[30] Ilaria Leocata. 2018. *Singing voice quality assessment in professional singers through acoustic parameters obtained with different microphones.* Ph. D. Dissertation. Politecnico di Torino.

[31] Yinmiao Li, Ziyue Piao, and Gus Xia. 2021. A Wearable Haptic Interface for Breath Guidance in Vocal Training. In *NIME 2021*. https://nime.pubpub.org/pub/cgi7t0ta.

[32] Kally M Luck, Dorothea C Lerman, Wai-Ling Wu, Danielle L Dupuis, and Louisa A Hussein. 2018. A comparison of written, vocal, and video feedback when training teachers. *Journal of Behavioral Education* 27 (2018), 124–144.

[33] Inita R Matta, Kanupriya B Halan, Ramesh H Agrawal, and Mandar S Kalwari. 2014. Laryngeal ultrasound in diagnosis of vocal cord palsy: An underutilized tool? *Journal of Laryngology and Voice* 4, 1 (2014), 2–5.

[34] Josef William McClellan. 2011. A comparative analysis of speech level singing and traditional vocal training in the United States. (2011).

[35] Fabio Morreale, Raul Masu, Antonella De Angeli, and Paolo Rota. 2013. The music room. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems* (Paris, France) *(CHI EA '13)*. Association for Computing Machinery, New York, NY, USA, 3099–3102. doi:10.1145/2468356.2479620

[36] Masaaki Nagae, Hiroyuki Umegaki, Akito Yoshiko, and Kosuke Fujita. 2023. Muscle ultrasound and its application to point-of-care ultrasonography: a narrative review. *Ann. Med.* 55, 1 (Dec. 2023), 190–197.

[37] Andrew Naseth. 2012. Constructing the voice: Present and future considerations of vocal pedagogy. *The Choral Journal* 53, 2 (2012), 39.

[38] Michael T Paris and Marina Mourtzakis. 2021. Muscle composition analysis of ultrasound images: A narrative review of texture analysis. *Ultrasound Med. Biol.* 47, 4 (April 2021), 880–895.

[39] V. Pettersen, , and R. H. Westgaard. 2004. Muscle activity in professional classical singing: a study on muscles in the shoulder, neck and trunk. 29, 2 (2004), 56–65. doi:10.1080/14015430410031661 Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/14015430410031661.

[40] Viggo Pettersen, Kåre Bjørkøy, Hans Torp, and Rolf Harald Westgaard. 2005. Neck and Shoulder Muscle Activity and Thorax Movement in Singing and Speaking Tasks with Variation in Vocal Loudness and Pitch. 19, 4 (2005), 623–634. doi:10.1016/j.jvoice.2004.08.007

[41] Ziyue Piao, Marcelo M. Wanderley, and Felipe Verdugo. 2024. MappEMG: Enhancing Music Pedagogy by Mapping Electromyography to Multimodal Feedback. In *ArtsIT, Interactivity and Game Creation* (Cham, 2024), Anthony L. Brooks (Ed.). Springer Nature Switzerland, 325–341. doi:10.1007/978-3-031-55312-7_24

[42] Courtney Reed, Andrew McPherson, et al. 2020. Surface electromyography for direct vocal control. International Conference on New Interfaces for Musical Expression (NIME).

[43] Courtney N Reed, Sophie Skach, Paul Strohmeier, and Andrew P McPherson. 2022. Singing knit: soft knit biosensing for augmenting vocal performances. In *Proceedings of the Augmented Humans International Conference 2022*. 170–183.

[44] Rui Rolo. 2011. Singstar-applying to music education. In *EdMedia+ Innovate Learning*. Association for the Advancement of Computing in Education (AACE), 3192–3201.

[45] Fabio Sarto, Jörg Spörri, Daniel P Fitze, Jonathan I Quinlan, Marco V Narici, and Martino V Franchi. 2021. Implementing ultrasound imaging for the assessment of muscle and tendon properties in elite sports: Practical aspects, methodological considerations and future directions. *Sports Med.* 51, 6 (June 2021), 1151–1170.

[46] Lawrence Shapiro. 2019. *Embodied cognition.* Routledge.

[47] Yang Shi and Cheng Yang. 2013. Celestia: a vocal interaction music game. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems* (Paris, France) *(CHI EA '13)*. Association for Computing Machinery, New York, NY, USA, 2647–2650. doi:10.1145/2468356.2479485

[48] Kimberly M Steinhauer and Mary McDonald Klimek. 2019. Vocal Traditions: Estill Voice Training®. *Voice and Speech Review* 13 (2019), 354–359. Issue 3. doi:10.1080/23268263.2019.1605707

[49] Johan Sundberg. 1974. Articulatory interpretation of the "singing formant". *The Journal of the Acoustical Society of America* 55, 4 (1974), 838–844.

[50] Johan Sundberg. 2020. Singing power ratio and vocal health in professional classical singers. *Journal of Voice* 34, 5 (2020), 690–696.

[51] Johan Sundberg, Maddalena Bitelli, Annika Holmberg, and Ville Laaksonen. 2017. The "Overdrive" Mode in the "Complete Vocal Technique": A Preliminary Study. *Journal of Voice* 31 (2017), 528–535. Issue 5. doi:10.1016/j.jvoice.2017.02.009

[52] John Sweller. 2011. Cognitive load theory. In *Psychology of learning and motivation.* Vol. 55. Elsevier, 37–76.

[53] Ingo R Titze, Erich S Luschei, and Minoru Hirano. 1989. Role of the thyroarytenoid muscle in regulation of fundamental frequency. *Journal of Voice* 3, 3 (1989), 213–224.

[54] Ingo R Titze and Brad H Story. 2002. Rules for controlling low-dimensional vocal fold models with muscle activation. *The Journal of the Acoustical Society of America* 112, 3 (2002), 1064–1076.

[55] Carl Unander-Scharin, Kristina Höök, and Ludvig Elblaus. 2013. The throat III: disforming operatic voices through a novel interactive instrument. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems* (Paris, France) *(CHI EA '13)*. Association for Computing Machinery, New York, NY, USA, 3007–3010. doi:10.1145/2468356.2479596

[56] Carl Unander-Scharin, Åsa Unander-Scharin, and Kristina Höök. 2014. The vocal chorder: empowering opera singers with a large interactive instrument. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 1001–1010. doi:10.1145/2556288.2557050

[57] Carl Unander-Scharin, Åsa Unander-Scharin, Kristina Höök, and Ludvig Elblaus. 2014. Interacting with the vocal chorder: re-empowering the opera diva. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI EA '14)*. Association for Computing Machinery, New York, NY, USA, 603–606. doi:10.1145/2559206.2574798

[58] M Usha, YV Geetha, and YS Darshan. 2017. Objective identification of prepubertal female singers and non-singers by singing power ratio using matlab. *Journal of Voice* 31, 2 (2017), 157–160.

[59] Felipe Verdugo, Amedeo Ceglia, Christian Frisson, Alexandre Burton, Mickael Begon, Sylvie Gibet, and Marcelo M. Wanderley. 2022. Feeling the Effort of Classical Musicians - A Pipeline from Electromyography to Smartphone Vibration for Live Music Performance. In *International Conference on New Interfaces for Musical Expression* (2022-06-28). doi:10.21428/92fbeb44.3ce22588

[60] Peter Visentin and Gongbing Shan. 2011. Applications of Emg Pertaining to Music Performance - a Review. 1, 1 (2011), 15–32. https://www.proquest.com/docview/1705960673/abstract/809476B3202E4591PQ/1 Num Pages: 18 Place: Hauppauge, United States Publisher: Nova Science Publishers, Inc..

[61] Jennifer M Vojtech, Michael D Chan, Bhawna Shiwani, Serge H Roy, James T Heaton, Geoffrey S Meltzner, Paola Contessa, Gianluca De Luca, Rupal Patel, and Joshua C Kline. 2021. Surface electromyography–based recognition, synthesis, and perception of prosodic subvocal speech. *Journal of Speech, Language, and Hearing Research* 64, 6S (2021), 2134–2153.

[62] Christopher Watts, Kathryn Barnes-Burroughs, Julie Estis, and Debra Blanton. 2006. The singing power ratio as an objective measure of singing voice quality in untrained talented and nontalented singers. *Journal of voice* 20, 1 (2006), 82–88.

[63] Graham F Welch, David M Howard, and John Nix. 2019. *The Oxford handbook of singing.* Oxford University Press.

[64] Edwin M-L Yiu, Gary WH Lau, and Feifan Wang. 2023. Fatigue-related change in surface electromyographic activities of the perilaryngeal muscles. *Journal of Speech, Language, and Hearing Research* 66, 1 (2023), 98–109.

[65] Mingxing Zhu, Xin Wang, Hanjie Deng, Yuchao He, Haoshi Zhang, Zhenzhen Liu, Shixiong Chen, Mingjiang Wang, and Guanglin Li. 2022. Towards evaluating pitch-related phonation function in speech communication using high-density surface electromyography. *Frontiers in Neuroscience* 16 (2022), 941594.