Surface Electromyography for Sensing Performance Intention and Musical Imagery in Vocalists

Courtney N. Reed
School of Electronic Engineering & Computer Science
Queen Mary University of London
London, United Kingdom
c.n.reed@qmul.ac.uk

Andrew P. McPherson
School of Electronic Engineering & Computer Science
Queen Mary University of London
London, United Kingdom
a.mcpherson@qmul.ac.uk

ABSTRACT

Through experience, the techniques used by professional vocalists become highly ingrained and much of the fine muscular control needed for healthy singing is executed using well-refined mental imagery. In this paper, we provide a method for observing intention and embodied practice using surface electromyography (sEMG) to detect muscular activation, in particular with the laryngeal muscles. Through sensing the electrical neural impulses causing muscular contraction, sEMG provides a unique measurement of user intention, where other sensors reflect the results of movement. In this way, we are able to measure movement in preparation, vocalised singing, and in the use of imagery during mental rehearsal where no sound is produced. We present a circuit developed for use with the low voltage activations of the laryngeal muscles; in sonification of these activations, we further provide feedback for vocalists to investigate and experiment with their own intuitive movements and intentions for creative vocal practice.

CCS CONCEPTS

• **Applied computing** → Performing arts; Sound and music computing; • **Human-centered computing** → *HCI theory, concepts and models*; **Gestural input**.

KEYWORDS

Electromyography; Mental Imagery; First-Person Perspectives; Performer Intention; Biofeedback

ACM Reference Format:
Courtney N. Reed and Andrew P. McPherson. 2021. Surface Electromyography for Sensing Performance Intention and Musical Imagery in Vocalists. In *Fifteenth International Conference on Tangible, Embedded, and Embodied Interaction (TEI '21), February 14–17, 2021, Salzburg, Austria.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3430524.3440641

TEI '21, February 14–17, 2021, Salzburg, Austria

ACM ISBN 978-1-4503-8213-7/21/02…$15.00
https://doi.org/10.1145/3430524.3440641

1 INTRODUCTION

The singing voice presents a particular challenge for interaction design as it relies on well-defined control over muscles which are mostly obscured from view and are not easily accessible; mainly, we get information from the voice based on its sound. Vocal pedagogy itself revolves around this type of interaction, where educators address singing technique largely based on how a student sounds. Teaching occurs through the use of abstract metaphor and any progress the student has made to improve their technique is determined on the change in the sound [15, 45]. Although educators can adjust some aspects such as posture, there is little way to physically move structures existing within the body, namely the larynx. Therefore, much of voice teaching exists in metaphor or through sharing of personal experience, rather than being rooted in physiology [9, 64]. Singers build a relationship between their sound and bodily feeling during sound production which becomes rooted in mental imagery [33, 72]. The technique remains highly abstract in this mental representation and is tied closely to bodily sensations.

Due to the covert mechanics of singing and the mostly auditory ways in which we work with the voice, audio analysis also dominates most computer-based interactions. Although audio analysis of speech and singing has become much more accurate in recent years, audio result-based interaction is not currently able to relay information about body movements, such as posture and the positioning of the head or limbs, which are an integral part of sound production. Different actions may yield very similar sounds and misinterpretations of user intention can arise when there is a lack of insight on how the sound is being produced. Performer intention and aspects of emotional expression conveyed through the body can be lost if only audio information is collected from a performance. This is especially so for the voice, an instrument which is highly personal; vocalists often feel that their singing is a part of their identity and vocal health is a critical daily practice [1, 68, 72].

We describe in this paper a method for body-based interaction with the voice through surface electromyography (sEMG). Electromyography (EMG) involves measuring the electrical impulses produced by the nervous system which cause muscular contraction [10]. Autobiographical design utilising lived experience [18, 41, 65] informed the development of a system to sense sEMG with the laryngeal muscles. We present, through a first-person use case, how other potential users can learn through its application in a creative improvisation context and how other designers might work with sEMG in similar contexts. We further assess the benefits of sEMG in working with embodied relationships and observing how intention and movement are used in creative interaction. Our contributions are therefore:

- Summarising the relationship between mental imagery and embodied interaction, uniting these parallel ideas from cognitive science and design practice in a musical context,
- Describing how sEMG can be used to sense performance intention and provide information about conscious and unconscious movement, as well as executed and imagined action,
- Providing an autobiographical design perspective for sEMG sensing to be used with the physiology of the singing voice to sense intention and imagery use, which can be reproduced further for the study of other embodied interactions, and
- Applying sEMG in a creative context and assessing first-person experience with the system to show how vocalists can use this interaction to explore their technique.

2 RELATED WORK

2.1 Musical Mental Imagery

The ability to imagine something happening without needing to actually observe it relies on mental imagery. From a cognitive science perspective, a mental image is composed of internalised representations of an action, which are accumulated through experiences [26, 49, 81]. In a musical context, a soloist preparing for a performance would be able to use mental imagery to imagine what the stage lighting might look like, the sound of the orchestra playing the beginning of the piece, and how it should feel to breathe before singing their first note. Musical imagery is therefore the multimodal mental representation of a musical experience and the imagining of all of the necessary components to perform, without needing to experience or before experiencing it. [12, 26].

Musical imagery is critical for musicians. While practicing, performers build images of what their performance should look, feel, and sound like, and these images are later recalled so that the musician can recreate their desired performance [27, 58]. Strong mental images are useful for adaptation in performance [8, 37, 91] and can be applied to incorporate emotional expression [5, 12, 39, 57] and to execute difficult tasks through sensorimotor recall [20, 49, 54, 70]. Musicians often rehearse in ways which solidify these mental images, such as rehearsing in an environment which mimics the performance space [90] or rehearsing only within the mind. Mental musical rehearsal through "audiation" occurs as musicians imagine playing (or singing) only within the mind, and can be thought of as "hearing the music in your head" or "listening with the mind's ear [5, 31]." Audiation allows musicians to consider technical and expressive demands outside of performance [8, 81] and builds an expectation of the sound before any attempt to create it.

2.2 Neural Correlates to Imagery Ability

When employing mental imagery through the imagination of an action, neural activation occurs in a similar way to activation during execution [12, 31, 51, 53, 92]. The imagination of an action excites the areas of the brain needed to perform that action; audiation can therefore help train and internalise action patterns to be recalled more easily and consistently. There is a growing body of research around auditory musical imagery and its neural correlates; self-reports of auditory imagery ability as assessed using the Bucknell Auditory Imagery Scale (BAIS) [30] are found to strongly correlate with the strength of activation in areas of the brain tied to auditory processing [36, 93], working memory [36], and motor control [56]. Individuals self-assessed to have higher imagery ability are also found to be better with melodic pattern recall [29], pitch imitation and control [71, 91], and prediction of melodic movement [25] and tempo changes [30] in a musical piece. We see here that, using well-defined musical imagery, musicians are able to draw on past experiences to anticipate and react during performance.

2.3 Mental Imagery Forms Embodiment

Embodied musical technique arises then from use of well-refined and internalised mental imagery. This allows a musician to execute complex actions through focus on their big-picture intentions of a performance, with less attention on finer actions such as motor control over their instrument. With experience, musicians learn to match their gestures with desired sounds and, over time, the fundamentals of sound creation can be done without constant or complete attention. Embodiment, in opposition to more Cartesian mind-body dimensions, describes human interaction as depending on these types of internalised working routines which do not need to be actively monitored, yet are still highly nuanced and driven by experience [35, 83]. In highly skilled musicians, the body and the instrument here become one functioning entity [66]; in the case of the voice, already existing within the body, this connection to the instrument is highly personal from the singer's point of view [68].

The principles of mental imagery, embodiment, and lived experience are therefore very similar, although these ideas have formed somewhat separately between cognitive science, HCI, and design practice, respectively. The key similarity is that the mind and body are indeed inseparable [3, 84–86] and that, rather than actively moving through each part of a refined action, the execution of a task relies on a broader mental image and intention. Recreation of an action and behaviour depends on personal lived experience and environment and is constantly changing with further knowledge. A vocalist's technique and application of their experience, rooted in mental imagery, become more ingrained through information gathered from the bodily sensations while achieving a desired outcome. This creates a feeling of "one-ness" with the voice.

Mental imagery is therefore a key component in enactivism, post-phenomenology, and entanglement-driven design in the sense that it is formed through and depends on the user and system or environment interacting with each other [3, 19, 21]. Through the subjective experience of the user and their application of past experience, mental imagery provides the basis to examine the co-dependency and collaboration of a user with their environment [40, 83, 87]. Mental musical imagery forms as musicians participate in rehearsal, performance, instruction, collaboration, and so on, and then use the experience to relate back to their world [35, 43]. There is a cyclical relationship between performer and performance, where they continually shape each other and act, in a way, as a feedback mechanism - the performer learns from the space they are in or from whomever they perform with and the mental image is updated and refined as this interaction goes on. There is no permanency to the musical imagery and it evolves through the embodied relationship of the musician to their world [3, 43]. Imagery and embodiment are therefore inherently linked.

2.4 Designing for Interaction with the Voice

One major difficulty in developing technology which works with this embodied interaction comes from the fact that systems require some way of measuring gesture and action for input. When the machine's interpretation relies only on the results of a movement or interaction, rather than the lived experience and mental imagery driving that gesture, we lose its meaning and the intention in that action [83]. Considering the covert ways in which most interactions with the voice occur, this is especially the case.

The body's role in vocal sound production is often not considered when there is a preference on audio feature extraction alone. Many different vocal gestures may produce the same or very similar sounds, with nuances either going unnoticed or undetectable by current audio processing algorithms. We have little way of seeing or sensing how the vocalist moves and what technical knowledge and experience is driving their actions. In an expressive sense, much of the emotional intent in performance may also go unnoticed, including vocal embellishments such as vibrato, as they are currently not able to be detected with great accuracy through computer-based audio analysis. In a more serious sense, this interpretation or mapping of audio in an arbitrary way can lead to unnatural movement or miscommunication of gesture [83]. With the voice, this could also include ignoring or potentially encouraging strenuous or damaging activity. It can be difficult to tell if someone is using unhealthy technique, as the resulting sound may still be deemed "desirable," to an interaction system, but the action may impact the health of the voice or cause long-lasting damage [88]. In these cases, audio-based interactions without insight into the body may encourage singers with less-developed imagery and technical control to engage in unsustainable vocal practice. Therefore, it is important to keep bodily sensations and physiology in focus with such highly embodied practices, not only so that important information is not lost, but also so that the user can work with their body instead of against it.

Musical imagery and the performer's intent as a focus for design provide a more body-based representation of how a vocalist (or other musician) might perceive and interact with their instrument and communicate through it. In order to achieve this, we must provide an alternative to using only audio-based vocal interaction and return to the bodily origin of the sound. Most of the current methods of looking at physiology related to singing, while effective and highly informative, would be largely unsuitable for performance or creative contexts. Due to the internal nature of the vocal mechanisms, medical equipment is often needed to observe voice physiology, making this potentially an invasive or expensive task. For this reason, we propose surface electromyography for sensing vocalists' performance intention.

2.5 sEMG and the "Negative Latency"

Using minimally invasive sEMG, captured with electrodes adhered to the skin at the muscle site, we are able to sense the electrical neural impulses which are sent to the muscles to trigger activation. These signals convey information such as the intensity of contraction and work done by a muscle [6, 10, 78, 82]. The most notable difference in using sEMG, compared to other biosignals, is that the impulses being measured are present as a precursor to movement. sEMG therefore provides information about the intention of a wearer just before the gesture is made [77]. As described by Tanaka, this results in a sort of "negative latency," where sEMG signals can indicate the occurrence of a motion a few milliseconds before it occurs [79]. This is because of where sEMG lies in the action path: "A classical sensor, then, is at the 'output' of a gesture while the EMG is a signal that is the 'input' to a gesture [77, 79]."

2.6 Sensing Subvocalisation

In design for human interaction, this provides not only an opportunity to make up for system latency introduced during signal processing or analysis, but also in examining gesture separated from visible movement or resulting sound. In vocal interaction, we are able to sense activation not only in vocalised singing, but also during imagined mental rehearsal, where there is still muscular activity but no sound or visible motion produced. The activation of the vocal and articulatory muscles, involved in either singing or in speech, during imagined activity is called "subvocalisation," and occurs when we read or talk internally or audiate through a musical passage [4, 47]. As previously discussed, mental action and executed action share the same neural pathways; because of this, the muscles are still excited by neural impulses when imagination occurs. It has been found previously that the subvocalisation which occurs during use of this inner voice is linked to both auditory and kinetic imagery [2]. sEMG taken off the facial and jaw muscles during subvocalisation has been successfully used in the recognition and classification of internally spoken words [22, 47, 63]. In subvocalised singing, vocal exercises have produced similar patterns of sEMG activation as when sung aloud, further highlighting how the same imagery is used for both activities [73]. We can thus utilise EMG not only for intention in executed action, but also in imagined practice where imagery is extensively used.

2.7 Ambiguity and Exploration

Due to the fact that sEMG signals are notoriously noisy [10] and we are not always consciously moving all of our muscles, there are inherent aspects of controllability and disorganisation when using sEMG-based control [17]. This, however, can be beneficial to studying embodied relationships in design. This ambiguity allows for users to develop their own relationship with a system, learn through play and experience, and act intuitively towards design without any pre-existing ideas of how the system should operate [24, 75]. sEMG is highly responsive to movement, which would allow musical systems to take advantage of the very refined and subtle movements of experienced performers. In the opposite sense, these systems can also incorporate unconscious movement into interaction, allowing for users to explore their existing relationship with their bodies and aspects of technique which may have been adapted into the subconscious [17]. In this sense, sonic exploration with sEMG allows for intention and also effort and restraint to be used for control [77]. Performers can adapt the force and exertion of their movements in addition to the overall shape of the movement in interaction. Restraint, conscious loosening of muscular tension, or explicit "letting-go" of control over muscles is a critical part of the embodied musical experience [77], making sEMG-based sensing well-suited to the existing highly-defined control musicians have over their instruments and bodies in performance.

2.8 sEMG in Musical Interfaces

sEMG has already been employed in a number of human interaction studies and tangible wearable systems within TEI and the broader CHI community to extend natural body language. sEMG has been used previously to convey information to the user or another system about the force and direction of a gesture [6, 52, 55, 80], or as an alternative to other sensors for emotional communication through movement [11, 34, 52, 55, 89]. Converted to an auditory signal, sEMG is capable of providing discernible feedback about complex muscular activity through frequency and rhythmic conversion [82]. Within a musical performance context, sEMG has similarly been used to detect different gestures, particularly movements of the head and limbs [13, 79], for control over digital audio synthesis [78, 79]. The MYO armbands (Thalmic Labs) in particular have been used for wearable controllers which incorporate both rotational and sEMG sensing of the arms and fingers in performance. In studies involving the use of EMG for detection of performance gesture, it has been found that, although it is difficult to classify some gestures, amateur users were able to quickly adapt their movements according to audio feedback and create new gestures to achieve desired sound [44, 48, 59, 60, 67]. This interaction allows for users to explore how their movements impact audio synthesis and can be useful for spontaneous composition using the body. As imagery is built through these associations of movement and sound, sEMG can be used to reflect and represent changes in a user's intention informed by learning with a system. sEMG in this sort of creative context does not capture the user's existing imagery, but rather an image of the users adaptation and interaction with it.

3 SYSTEM DESIGN AND IMPLEMENTATION

We here describe a method for acquisition, filtering, and application of vocal sEMG. We begin with a basic overview of voice physiology, followed by an outline of the sensor development and how the setup works and can be recreated for other applications.

3.1 Vocal Physiology

Vocalisation depends on the larynx, the vocal tract (which includes the soft tissues in the throat and mouth), and the respiratory muscles. The respiratory system provides airflow through the larynx, causing the vocal folds (also known as vocal cords) to vibrate. This base vocal sound is further shaped by the vocal tract [32, 76] and provides each individual with their unique voice characteristics [76]. We focus on sEMG signals from the extrinsic muscles which control the position of the larynx (Figure 1) and which can be measured with surface electrodes placed around the throat.

3.2 Long-Term Autobiographical Design

We designed a custom setup for vocal sEMG acquisition following similar methods previously used in speech applications [47]. Development of a system to collect sEMG from the laryngeal muscles was created from scratch in order to tailor the setup around vocal practice. As well, currently, only medical-grade sEMG systems provide enough precision to examine such small muscles; other systems such as the MYO bands previously discussed, are only designed to be worn on the limbs and further are no longer commercially available. The design involved iterative testing in an autobiographical context by the first author of this paper, a classically trained mezzo-soprano who regularly sings in professional choral and solo settings. At the time of this writing, the author had been developing and living with the system for approximately 9 months and so had a long-term view of its use in a daily context. Although autobiographical design is only recently becoming more reported on in HCI research [18, 65], we believe that this form of design is beneficial to focusing on incorporating lived-experience and embodiment in interaction [28, 41, 65] and is a well-suited method for this exploration of intention with the voice. As well, the autobiographical design also allowed for development in its intended environment and immediate implementation by a user who had well-established vocal practices and imagery which could be called upon in active reflection during use to meet any challenges [74].

It is important to note then that this system and experience is not intended to provide a universal view of how such a sEMG system can be implemented. Living with the system through its development and designing for use by the designer, rather than a hypothetical imagined user, the design became highly specific to individual use. The use of the design grew up around this long term relationship and evolved as it did in a way that is potentially not usable to others [18], who might approach the system from a completely different perspective. However, this type of long-term personal use allows for evolution through customisation [18], and means that the sEMG system in its current state actually describes the user's perspectives and intentions in using it. The act of tinkering with the system to meet a more tailored use as time went on created a relationship with the system, which has been previously observed in similar long-term interactions [18, 23, 87]. Similarly, it would be useful to observe another singer (or even a non-musical user) adapting this system and growing through it [94]; this could either lead to generic features which are useful to many or implementation of features which allow any potential user to customise the setup around their own practice, which inherently would reflect the user in the system.

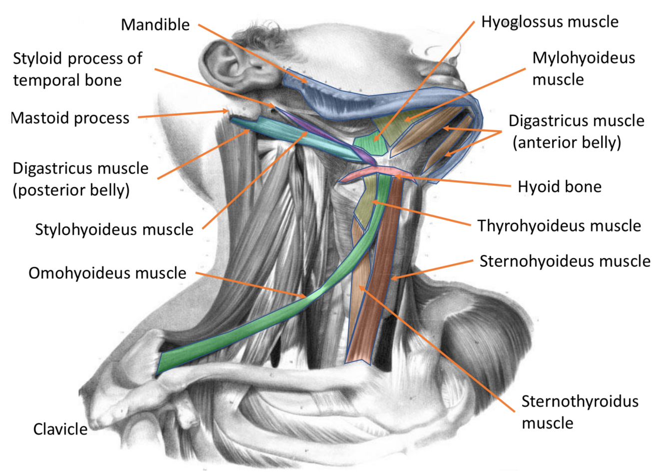

Figure 1: Extrinsic muscles of the larynx (adapted from image available in the public domain, retrieved from Flickr: https://flic.kr/p/u6pAzL).

3.3 Sensing and Setup

The currently used version of the EMG circuit can measure up to two muscles simultaneously. For a single muscle, three reusable 10mm gold-plated silver cup electrodes (Medimaxtech, New Malden, UK) are needed. One electrode is used as a reference to the body and should be placed on a non-muscular area of the body (we use the earlobe, for instance); the other two electrodes are placed on the desired muscle, with one in the middle of the muscle body and one at the end of the muscle, close to where it attaches to the bone. If a second muscle is being measured, the singular reference electrode can still be used, so only an additional 2 electrodes are needed[1]. A conductive adhesive paste (Ten20 Conductive Paste, Weaver and Company, Aurora, CO, USA) is used to secure the cup electrodes and reduce skin impedance; the electrodes are further secured using an adhesive non-woven fabric tape (Hypafix, BSN Medical GmbH, Hamburg, Germany) which is placed over the electrodes to ensure they are in close contact with the skin and do not move around.

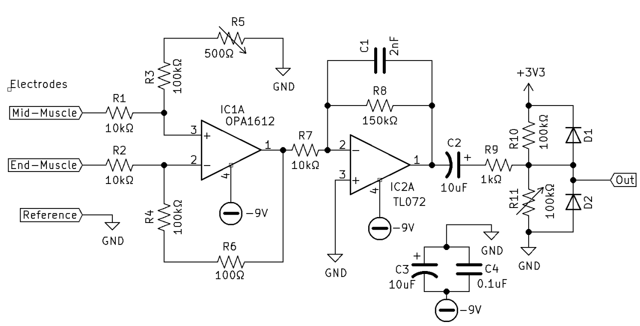

Figure 2: The current version of the circuit for sEMG signal acquisition and pre-amplification for a single muscle using three electrodes, first presented in [73]. It is important in reproduction of this circuit that high-precision (ideally 0.1%) metal film resistors be used for R1-R4 to ensure the resistance in series with both electrodes is the exactly the same. The OPA1612 (ICA1) was chosen for its extremely low voltage noise density ($1.1nV/\sqrt{(Hz)}$).

The informative components of the raw sEMG signals lie mainly below 500 Hz and have an amplitude of less than 10 mVpp [10]. The small sEMG signal is therefore acquired using an analogue pre-amplification circuit (Figure 2). The circuit is powered by two 9V batteries, which allow for integration into a wearable setup detached from a computer, as well as reduction of grid power noise sources. The circuit amplifies the difference in electrical potential between the mid- and end-muscle electrodes; during activation, this potential will become uneven as the muscle contracts. Taking the difference between the electrodes also allows for common mode rejection, removing some noise present at both points, including artifacts from other bodily processes. The signal is filtered into an appropriate frequency range; the Common Mode Rejection Ratio (CMRR) for this circuit averages approximately −37.96 dB across the frequency band for sEMG signals. The amplified and filtered signal is then passed to a Bela board [62], an open-source embedded computing platform. We use Bela in this context for digital signal processing due to its ultra-low latency and small size, which further adds to the portability of the system. Within the digital Bela context, the sEMG signal can be further processed to remove any remaining noise and used for digital interaction. For example, we use the sEMG signal with SuperCollider, a computer language and platform for audio synthesis and algorithmic control, where we can use the laryngeal muscles to control aspects of digital vocal processing.

4 AUTOBIOGRAPHICAL STUDY

In the next section, we describe a use case for sEMG in vocal improvisation performed by one of the authors (from here, I will therefore use first-person). The system was designed through iterative testing by myself, with a focus on my own internalised metaphors, imagery, and personal experience as a vocalist. Keeping this in mind, while my own interaction with the system is unique and the specifics of my interaction would not necessarily be translatable to other vocalists, there are aspects of this interaction which we feel can be ubiquitous. For instance, the ways in which I learned to interact with my body and adapt some of my technique, as well as some perhaps unconventional or unnatural behaviours I attempted, could be distinguishable interaction patterns which are common amongst others using the system. Even in non-musical contexts, the methods of exploration and balance between user action and system reaction may be experienced similarly by individual users.

The interaction involved sEMG taken from my suprahyoid region, namely on the digastric, mylohyoid, and geniohyoid muscles (Figure 3); these muscles, located under the chin, are active during laryngeal elevation and move with the tongue. The tongue is important in singing firstly for lyric articulation, but also can be moved in a way that creates different kinds of resonant spaces in the mouth. Changing the position of the tongue helps to get different vowels and place the sound in different portions of the face, with either a rounded open sound or a pinched nasally sound in either extreme. I used these sEMG signals in a SuperCollider script to control the frequency of a sine wave carrier used in a ring modulator. With increases in suprahyoid activity, I was able to increase the frequency of the sine wave which modulated my direct voice input. As well, I included some other synthesis effects to my direct voice input such as a pitch follower and chorusing to create an environment in which I could play around with the sEMG signals.[2]

4.1 Interaction Perspectives

Before I begin to describe my experience with the sEMG sensor in this application, it is important that I reconcile the differences between first-person and third-person perspectives with this system and its usage. sEMG can be used in both ways and it depends on who is the spectator of the information it provides–it can most certainly be used in a third-person context where information about the user's muscular activations are analysed and used to convey objective information such as force and grasp, as utilised in previous research for controllers using sEMG [11, 47, 55, 63, 80]. Here,

[1] Note that this schematic depicts the acquisition for a single-muscle setup; in order to capture two muscles at once, the same signal path can simply be duplicated using the other side of the two op-amps used here (input pins 5-8 on each IC).

[2] An excerpt of an improvisation using this setup can be found at: http://instrumentslab.org/data/courtney/TEI2021_demo.mp3

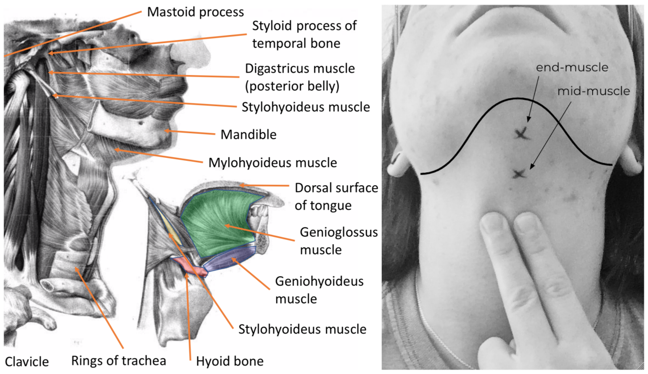

Figure 3: The suprahyoid muscles (adapted from image available in the public domain, retrieved from Flickr: https://flic.kr/p/u6pB3Q), and electrode placement on the suprahyoid region (right, finger placement indicates the position of the hyoid bone on the wearer).

predictions are made about body movement or the muscular contractions are used to interact with digital object, but the gesture is not necessarily adapted based on the machine's interaction. A movement classifier to control a robotic hand, for instance, determines a gesture and mimics the user's action and the user acts through the machine [50]. sEMG also allows for useful first-person interaction, where this reciprocal interaction between human and machine is at the forefront. This is more common in musical [77–79] and emotional interaction [34], where the user is the spectator of their own gestures and learns by observing how the system responds and learning how it acts as an individual agent.

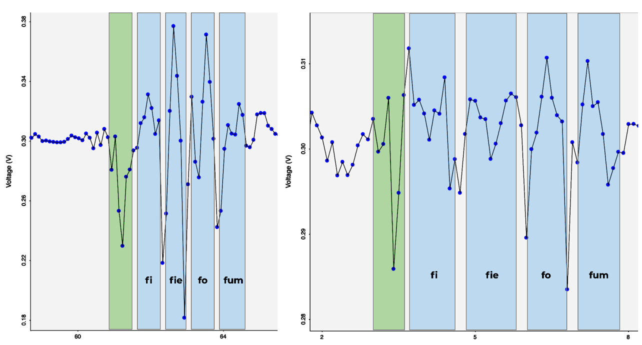

Figure 4: Activation of the omohyoid muscle while singing (left) and imagining (right) a major arpeggio, as displayed in the Bela GUI. The activation during the breath is highlighted in green, with the sung syllables in blue.

The same can be said for this particular vocal sEMG system. I can use it to simply measure gesture in an objective way, and this is beneficial if I want to see how my imagined action mirrors anything I actually do. For example, if I look at activations in the omohyoid moving while I sing a major arpeggio, and compare it with my imagining of that same exercise, I can see that this activation occurs in a similar pattern (Figure 4). In this exercise, I sing a simple major arpeggio (C4-E4-G4-C5) on the syllables 'fi, fie, fo, fum.' The muscle activates as I articulate each note (or imagine articulating each note). Note that the subvocalised exercise (right) lies within a much smaller voltage range (approximately 0.285-0.31 V, 0.025 V peak-to-peak) than the vocalised exercise (left, 0.18-0.38 V, 0.2 V peak-to-peak) but similar activation of the muscle occurs in subvocalisation as in vocalisation. This is by no means a quantitative analysis of comparison between vocalised and subvocalised activity, but rather a means to see that activity does indeed occur in the same exercise both sung and imagined, as utilised in previous studies using the muscles of the face, neck, and jaw [47]. This type of information could be further used for gesture classification or force measurement or a multitude of other analyses; in this sense, it is information about my performance but not about any sort of interaction with the system. The system examines me, and in this sense, it views my performance from a third-person perspective to measure aspects about my movement or imagery.

On the other hand, in my exploration of the system, I was not necessarily looking to determine any gestures or to extract any sort of data. However, I do get to observe how the machine responds to my input and how I can adjust accordingly to change the result if I wish. We are co-dependent in a very entangled way and the interaction is based within a first-person perspective, where I am in control of my responses to the system as my collaborator or accompanist. It is in this type of first-person perspective which we see the benefits of sEMG and musical imagery as tools for embodied interactions; the output is not always predictable or expected, so manipulation of existing knowledge, surprise, and reaction are used to relate to the machine in an improvised performance [19, 43]. I am able to monitor and change my expectations and work with the system to create something meaningful to me and rooted in both my existing experience and the learning I did while using the system, which I will now elaborate on.

4.2 Re-Learning and Reacting with the Body

Firstly, the decision to use the suprahyoid region was made because of the flexibility of interaction I learned through testing and playing with the sEMG sensor. In testing the system, I found as well that the larger infrahyoid muscles (those underneath the hyoid bone, Figure 1), particularly the omohyoid muscle, provided sources of reliable activations, although these are laryngeal depressors and so are most active at the bottom of the vocal range. On the other hand, because the suprahyoid muscles move not only to elevate the larynx but also with the chin, jaw, and tongue movement, I felt I would not have to rely on pitch production or singing within a certain range, and so might be more flexible in this case; however, restricting movement by focusing on muscles which perform a more defined action would also be worth further consideration when it comes to observing how learning occurs through interaction with a system.

With the suprahyoid, I know I can choose between a number of techniques which actively utilise these muscles, ranging from vowel formant shaping and text articulation to some unorthodox practices in singing, namely stretching my neck all the way out or squishing it towards my throat (ie. the "double-chin"). These sorts of movements are completely undesirable in singing, as they create lots of tension and make the voice sound constricted; in students who are soprano voices like myself, it is a common tendency in an untrained voice to "reach with the chin" when trying to sing higher notes, and several voice teachers worked very hard to hammer this out of me. However, when it came to actually playing with the system, I realised I could get a lot of activity out of these activities and so I started to work them into my practice, leaning into some of the more constricted sounds for a tense effect. For me, this was a very useful practice tool–I had to recall motions which had been trained out of me and try to work with them while still keeping my "well-trained" open and supported sound. An untrained singer might have less awareness of these tensions in the chin, and providing a source of feedback to it would likely call attention to it.

In an improvisation sense, I found I behaved normally when it came to the pitches I was singing. Being very comfortable with the sorts of fast melisma-style phrases you'd find in opera (where a singer strings many notes alone to one lyric), I would have previously focused my vocal contributions in improvisation on moving the melodic line, toying with dissonances, and mimicking any other instruments I was performing with. In the process of coding the synthesis in SuperCollider, I added just a small sine wave chorus in the background so I had some context to sing with. I play a game with myself where I try to sing complex melodies over the drone of the microwave while heating up my lunch, so I imagined doing something similar in this improvisation; however, I found it was more interesting to focus on changing my vowel and resonant space over a singular pitch instead. I became a part of this background drone, trying to harmonise with it, but not doing very much over top of it, in favor of playing with formant placement instead.

4.3 From Unconscious to Conscious

As well, some movements of the suprahyoid I had not noticed until I started getting feedback from the activation during testing of the sEMG system. sEMG in this way can provide a basis for *re*action to a system and adaptation to the unexpected. My tongue moves down and out of the way when I breathe so I can take in more air and then moves again as I start to articulate any text. I had observed this while refining and testing the circuit, but I did not realise how prevalent it was until I began to sing with it. As someone who has been singing for quite some time, the act of supported breathing is very natural and I do not have to think how to get enough air and sustain a note for a long period of time. The beginning of phrases were modulated almost always because of this movement of my tongue after breathing. Certain consonants require more movement and so have stronger activation; for instance, plosives made with the lips, such as 'p' and 'b,' minimally involve the tongue compared to sharp consonants such as 't' and 'k.' This awareness of my breathing made its way into the improvisation as well, resulting in sharp or exaggerated breaths, exhaling without producing pitch, and using different consonants.

Being aware of both conscious and unconscious movements, especially those which do not produce any sound (such as the extreme chin movements I describe previously), are extremely useful when considering vocal technique and voice pedagogy. Many younger or inexperienced singers move in all sorts of ways which are incompatible with singing in a healthy and sustained way, but without the help of a voice teacher to physically point them out, they can go unnoticed. As well, the act of a teacher correcting posture or head positioning or some other type of movement during an exercise does not allow a student to examine their behaviour in the moment, but rather by comparing and contrasting as a before and after correction. Sometimes tension can be introduced through, for example, a jutting chin, and be heard in the sound; however, these types of unsustainable movement and overexertion are more likely to be felt in the long-term [88]. Myoelectric potential captured through sEMG in beginner artisans is found to be unstable and power used unnecessarily, compared to experts, making this a valuable indicator of control or lack thereof [38]. If this type of information were relayed through digital modification of a sound, as done here with my chin movements and resulting muscular tension, this type of soundless movement could be relayed back to the student, who would potentially then notice their straying posture and correct themselves, as done successfully with other auditory biofeedback [14, 69], leading to in-the-moment learning of action-sound relationships. In the same way, this reinforcement could provide additional information to educators.

In these couple examples of my interaction with the sEMG sensing, what was most striking was that, by drawing attention to one aspect of my technique, I actually found myself singing less, at least in a conventionally attractive way, and just paying attention to my movement. The tongue is constantly moving during most vocal exercises and diction and text clarity are common focuses in rehearsal; however, the movements of the tongue and chin need to also remain relaxed to keep resonant spaces and projection. I know how to open this space in my mouth and throat for a proper operatic or choral sound, and I don't usually think about it. When I gave presence to this aspect of my voice, I found myself putting it at the forefront of the interaction, often at the expense of any clarity or support in my pitch. Even though I designed this system to provide an alternative to audio-only voice interaction, I was surprised to find that I quickly abandoned some practices which I have been refining for years; in fact, I felt it was because I know so well what not to do while singing that I could actually supplement a number of new ways to appropriate these "poor" techniques.

5 DISCUSSION

The system was developed to collect information about the singing voice, yet the experience of singing with it was completely changed to work with the system in this musical exercise. In this context, it is a poor tool for collecting data on musical practice; however, it is an excellent tool in allowing the user to practice manipulating their body or how they move in different contexts. This resonates with ideas of postphenomenology or entanglement and using design practices in HCI [43, 94], where we as users are inseparable from our tools and, through their use, they have the potential to change our goals [21, 87]. Rather than demonstrating any particular technical skill or performing in a way shaped by many years of experience or instruction, the intentions through performance became focused on this bodily relationship in the moment and on how the ingrained musical imagery and internalised movements could be manipulated. When applying the system in this way, the user and their embodied techniques and experience are inseparable from the system itself and they cannot exist alone; the sEMG system was developed around the vocalist's body, and the vocalist's practice then revolves around how the system reacts. The system becomes a collaborator in this performance, assisting the improvisation by providing feedback about the body and sonifying performer intentions which have become background processes over time.

5.1 Insights for Further Application

Through this account of a first-person interaction, there are several takeaways which can be applied to both the use of sEMG systems specifically and other feedback systems with which the user collaborates and adapts. As previously discussed, this system and other biofeedback devices allow for both first- and third-person interaction; the user's movements can be observed, measured, and utilised in a quantitative manner, or the user is free to explore their movement through qualitative sensation and reaction to the system. Although we have only begun to explore quantitative aspects of vocalisation through sEMG and focus on mainly qualitative aspects through this paper, this method of interfacing with the body has been able to be measured and used successfully in classifiers [47, 55]. In this sense, such a system provides potential to not only observe the process of adaptation and learning through biofeedback, but also a way to measure it and begin to describe those processes.

Because it does not rely on visual movement, sound, or any perceptible action, sEMG provides a concrete way to examine action apart from complex image and audio analysis. Further, because muscles are active during mental rehearsal and imagined action, sEMG can be used to examine more internalised aspects such as preparation and adaptation through imagery application. sEMG also provides a way to experience bodily sensations externally, either visually or aurally as presented here, and potentially with other senses, which is known to reinforce learning [7] and provide understandable information about the amount of and change in muscular activity [82]. This kind of interaction is beneficial in that it provides a way for users to not only explore their conscious movements, but to also become aware of the unconscious action paths which have become internalised through prolonged experience. Where action has become embedded in the imagery of a desired result, biofeedback from sEMG can force users to question their actions and recall learned behaviours; in this case, through the sonification of movements which were being done subconsciously or had been intentionally trained out of the user. This is useful not only in performance contexts, but also in teaching and training usage; as described in vocal pedagogy, the ability to bring attention to particular actions and challenge personal technique. The visibility of sEMG through biofeedback is found to be useful in training and rehabilitation as it also enforces understanding between practitioners and patients [42]. This is critical in many teacher-student relationships, such as those in voice teaching and other arts practices.

Finally, the coordination and decisions on the part of the user, some of which were observed through the first-person interaction described here, can potentially inform about action paths in interaction in more universal senses. sEMG allows the user to tease apart ambiguous responses from the system and act as a detective to determine which aspects they have control over and how to work with the things they cannot; this is critical in performance art generally, and a worthwhile exploration for other fields besides vocal performance. There is potential that other users may be inclined to act through play, doing somewhat silly, creative, or unorthodox things just to see what happens when presented with this interaction, and then refine them further through practice. Creativity can be similarly supported in different adaptations by the user to match the feedback they are given, whether that be by changing behaviour or by changing the focus of what constitutes "interesting" or "desirable" performance, as described through this personal use case. By identifying similar patterns of exploration or determining multiple ways of interaction, we can better engage with personal interpretation and observe emotional connection to experience in design and the development of this and similar feedback systems.

5.2 Future Studies

We can imagine a variety of additions to the system or potential applications to explore in both a design and research context. Two key areas to explore further with performer intention while using the sEMG system would be through user studies playing on the flexibility and bodily learning with the system, as described previously in the author's use case, and in learning contexts where the system can be used to highlight and exaggerate behaviours and build imagery around the body. Because sEMG interacts with user intention and we can use it build systems centered on the body, it provides the basis for further uniting HCI research and interaction design practices [28]. We plan for further first-person study of the setup and including users documenting their prolonged relationship with the system and giving them the ability to control aspects of their interaction in the same way it was developed; this might include further flexibility with sensor placement to let users focus on their desired movement aspects and studies involving play and learning through interaction to observe how the relationship with existing technique and musical imagery can be further used or subverted.

As well, use of this system in learning environments would be a beneficial area of exploration, as previously described in Section 4.3. For inexperienced vocalists who are still building musical imagery and relationships with their bodies, there are benefits in calling attention to gestures which may be done subconsciously. Potentially harmful singing practices can be indicated back to the user and additional feedback, as in the case of the musical system described in this paper, can allow students to play with their bodies and movements. As well, multimodal feedback from visualisation and sonification of gesture can strengthen imagery formed during rehearsal [27, 58]. Sonification can lead to better reproduction accuracy in movement patterns [16] and allow for voice teachers to better relate to students [42] through this sensory feedback, in addition to relaying experiential knowledge. Sonification of biofeedback is found to be engaging and rewarding in rehabilitation through the emotional cues provided within music [46, 61], which may further strengthen the learning process in terms of skill practice.

It is of course critical to again state that, while the interaction observed provides a detailed account of a prolonged interaction with biofeedback through sEMG, this interaction is highly specific to the user. As well, there is a strong potential bias for the interaction capabilities of the system due to the fact that the vocalist observed is also one of the authors of this paper. As mentioned previously, further studies to conduct similar trials and autobiographical use cases with the system will be necessary to validate the universality or differences in the experiences. As well, there is little to be said quantitatively regarding this system while currently N = 1; these further studies will also need to work to detect clearer similarities of activation between vocalised and subvocalised singing, as well as in patterns of activity and work done by the muscles for different vocal techniques. This is especially critical if classifier systems and further pedagogical applications are to be created; this paper is limited in terms of such future developments due the lack of a qualitative study. However, in our current world of self-isolation and limited capacity for singing, we hope similar case studies and documented long-lived interactions will be come a larger part of design and development practice. The personal experiences this study provides form the basis for future imagery and sensory-based exploration in biofeedback systems, and the beginnings of voice related studies which bring focus back to the artist's intention and movement and in addition to interpretation of their sound.

6 CONCLUSION

In this paper, we demonstrate how sEMG can be used to detect musical mental imagery use and sense vocalists' intention in performance. We discuss how the development of musical imagery through experience allows performers to focus on their intention and high-level action, rather than on technique. With sEMG, musical systems can be based around this performer intention and use both conscious and unconscious body movement to form interaction. We find sEMG to be useful especially in bringing attention to embodied techniques which have become highly internalised and allow performers to play with their knowledge and experience in performance, and we demonstrate the potential of such an interaction through a first-case use by a professional vocalist. We discuss how highlighting a particular body-based technique for interaction in digital synthesis can shape the content of a performance and strengthen the performer's relationship with the body and use it for creative control. We thus identify sEMG as an effective tool for uniting aspects of cognitive science, HCI research, and design practices through this focus on user intention and experience.

ACKNOWLEDGMENTS

CNR is funded by a Principal Electronic Engineering and Computer Science Studentship from Queen Mary University of London. APM is funded through EPSRC grant EP/N005112/1 (Design for Virtuosity). The authors wish to thank Morghan Hartmann, who has provided essential remote electronics support during our COVID-19 university closure.

REFERENCES

[1] M. A. Achey, M. Z. He, and L. M. Akst. 2016. Vocal Hygiene Habits and Vocal Handicap Among Conservatory Students of Classical Singing. *Journal of Voice* 30, 2 (2016), 192–197. https://doi.org/10.1016/j.jvoice.2015.02.003

[2] A. Aleman and W. van't Wout. 2004. Subvocalization in auditory-verbal imagery: just a form of motor imagery? *Cognitive Processing* 5, 4 (2004), 228–231. https://doi.org/10.1007/s10339-004-0034-y

[3] N. B. Armstrong. 2007. *An enactive approach to digital musical instrument design–Theory, Models, Techniques.* VDM Verlag.

[4] A. Baddeley, M. Eldridge, and V. Lewis. 1981. The role of subvocalisation in reading. *The Quarterly Journal of Experimental Psychology Section A* 33, 4 (1981), 439–454. https://doi.org/10.1080/14640748108400802

[5] F. Bailes. 2006. The use of experience-sampling methods to monitor musical imagery in everyday life. *Musicae Scientiae* 10, 2 (2006), 173–190. https://doi.org/10.1177/102986490601000202

[6] V. Becker, P. Oldrat, L. Barrios, and G. Sör. 2018. TouchSense: Classifying and Measuring the Force of Finger Touches with an Electromyography Armband. In *Proc. Augmented Human International Conference (AH '18).* ACM, New York, NY, USA. https://doi.org/10.1145/3174910.3174947

[7] G. Bernal, D. Ahuja, and F. Casalegno. 2010. EMG-based biofeedback tool for augmenting manual fabrication and improved exchange of empirical knowledge. In *Proc. ACM XVI International Conference on Human Computer Interaction (Interacción '15).* 1–8. https://doi.org/10.1145/2829875.2829932

[8] R. Brown and C. Palmer. 2012. Auditory–motor learning influences auditory memory for music. *Memory & Cognition* 40, 4 (2012), 567–578. https://doi.org/10.3758/s13421-011-0177-x

[9] J. Callaghan. 1998. Singing Teachers and Voice Science - An Evaluation of Voice Teaching in Australian Tertiary Institutions. *Research Studies in Music Education* 10, 1 (1998), 25–41. https://doi.org/10.1177/1321103X9801000103

[10] R. Chowdhury, M. Reaz, M. Ali, A. Bakar, K. Chellappan, and T. Chang. 2013. Surface Electromyography Signal Processing and Classification Techniques. *Sensors* 13, 9 (2013). https://doi.org/10.3390/s130912431

[11] E. Costanza, S. A. Inverso, and R. Allen. 2005. Toward subtle intimate interfaces for mobile devices using an EMG controller. In *Proc. CHI'05: SIGCHI Conference on Human Factors in Computing Systems.* ACM, New York, NY, USA, 481–489. https://doi.org/10.1145/1054972.1055039

[12] J. Cumming and S. E. Williams. 2012. *The Role of Imagery in Performance.* Oxford University Press, Oxford. https://doi.org/10.1093/oxfordhb/9780199731763.013.0011

[13] M. Donnarumma, B. Caramiaux, and A. Tanaka. 2013. Combining EMG and MMG sensing for musical practice. In *Proc. New Interfaces for Musical Expression (NIME), Daejeon, Korea, May 27-30, 2013.*

[14] M. Dozza, F. B. Horak, and L. Chiari. 2007. Auditory biofeedback substitutes for loss of sensory information in maintaining stance. *Experimental Brain Research* 178, 1 (2007), 37–48. https://doi.org/10.1007/s00221-006-0709-y

[15] R. Dunbar-Wells. 1999. The Relevance of Metaphor to Effective Voice Teaching Strategies. *Australian Voice* 5 (1999), 50–59.

[16] A. O. Effenberg. 2004. Using Sonification to Enhance Perception and Reproduction Accuracy of Human Movement Patterns. In *Proc. International Workshop on Interactive Sonification, Bielefeld, Germany.* 1–5. https://doi.org/10.1109/MMUL.2005.31

[17] C. Erdem and A. R. Jensenius. 2020. RAW: Exploring Control Structures for Muscle-based Interaction in Collective Improvisation. In *Proc. New Interfaces for Musical Expression (NIME), Royal Birmingham Conservatoire, Birmingham, UK.*

[18] T. Erickson. 2010. The design and long-term use of a personal electronic notebook. In *Proc. CHI '96, ACM Press.* 11–18. https://doi.org/10.1145/238386.238392

[19] G. Essl and S. O'Modhrain. 2006. An enactive approach to the design of new tangible musical instruments; The design of new tangible musical instruments. *Organised Sound* 11, 3 (2006), 285–296. https://doi.org/10.1017/S135577180600152X

[20] F. Fontana, H. Järveläinen, S. Papetti, F. Avanzini, G. Klauer, and G. Malavolta. July 30-31 & Aug 1, 2015. Rendering and Subjective Evaluation of Real Vs. Synthetic Vibrotactile Cues on a Digital Piano Keyboard. In *Proc. Sound and Music Computing (SMC-15), Maynooth, Ireland.* 161–167.

[21] C. Frauenberger. 2019. Entanglement HCI The Next Wave? *ACM Transactions in Computer-Human Interaction* 27, 1 (2019), 2:27. https://doi.org/10.1145/3364998

[22] L. I. Garrity. 1977. Electromyography: A review of the current status of subvocal speech research. *Memory & Cognition* 5, 6 (1977), 615–622. https://doi.org/10.3758/BF03197407

[23] W. Gaver. 2006. The video window: my life with a ludic system. *Personal and Ubiquitous Computing* 10, 2-3 (2006), 60–65. https://doi.org/10.1007/s00779-005-0002-2

[24] W. W. Gaver, J. Beaver, and S. Beford. 2003. Ambiguity as a Resource for Design. In *Proc. CHI. 2003.* 233–240. https://doi.org/10.1145/642611.642653

[25] R. W. Gelding, W. F. Thompson, and B. W. Johnson. 2015. The pitch imagery arrow task: Effects of musical training, vividness, and mental control. *PloS One* 10 (2015), e0121809. https://doi.org/10.1371/journal.pone.0121809

[26] R. I. Godøy and H. Jørgensen. 2001. *Musical Imagery.* Number 5 in Studies on New Music Research. Swets & Zeitlinger, Lisse, Netherlands.

[27] W. Goebl and C. Palmer. 2008. Tactile feedback and timing accuracy in piano performance. *Experimental Brain Research* 186, 3 (2008), 471–479. https://doi.org/10.1007/s00221-007-1252-1

[28] E. Goodman, E. Stolterman, and R. Wakkary. 2011. Understanding Interaction Design Practices. In *Proc. CHI 2011, May 7–12, 2011, Vancouver, BC, Canada.* 1061–1070. https://doi.org/10.1145/1978942.1979100

[29] E. B. Greenspon, P. Q. Pfordresher, and A. R. Halpern. 2018. Pitch Imitation Ability in Mental Transformations of Melodies. *Music Perception: An Interdisciplinary Journal* 34, 5 (2018), 585–604. https://doi.org/10.1525/mp.2017.34.5.585

[30] A. R. Halpern. 2015. Differences in auditory imagery self-report predict neural and behavioral outcomes. *Psychomusicology: Music, Mind, and Brain* 25, 1 (2015), 37–47. https://doi.org/10.1037/pmu0000081
[31] A. R. Halpern and R. J. Zatorre. 1999. When that tune runs through your head: A PET investigation of auditory imagery for familiar melodies. *Cerebral Cortex* 9 (1999), 697–704. https://doi.org/10.1093/cercor/9.7.697
[32] W. J. Hardcastle. 1976. *Physiology of Speech Production: An Introduction for Speech Scientists.* Academic Press Inc., London.
[33] D. Hargreaves, D. Miell, and R. MacDonald. 2011. *Musical Imaginations: Multidisciplinary perspectives on creativity, performance and perception.* Oxford University Press, Oxford.
[34] K. Hartman, B. Kourtoukov, and E. Lewis. 2018. Kinetic Body Extensions for Social Interactions. In *Proc. Tangible, Embedded, and Embodied Interaction (TEI '18).* ACM, New York, NY, USA, 736–739. https://doi.org/10.1145/3173225.3173333
[35] M. Heidegger. 1967. *Being and Time.* Blackwell, Oxford.
[36] S. C. Herholz, A. R. Halpern, and R. J. Zatorre. 2012. Neuronal correlates of perception, imagery, and memory for familiar tunes. *Journal of Cognitive Neuroscience* 24 (2012), 1382–1397. https://doi.org/10.1162/jocn_a_00216
[37] Z. Highben and C. Palmer. 2004. Effects of Auditory and Motor Mental Practice in Memorized Piano Performance. *Bulletin of the Council for Research in Music Education* (2004), 1–8.
[38] A. Hiyama, Y. Doyama, K. Kakurai, H. Namiki, M. Miyasako, and M. Hirose. 2010. Archiving and transferring of traditional artisanship focused on interaction between artisan and tools. In *Proc. 16th International Conference on Virtual Systems and Multimedia, Seoul, South Korea.* 171–176. https://doi.org/10.1109/VSMM.2010.5665987
[39] P. Holmes. 2005. Imagination in practice: a study of the integrated roles of interpretation, imagery and technique in the learning and memorisation processes of two experienced solo performers. *British Journal of Music Education* 32, 3 (2005), 217–235. https://doi.org/10.1017/S0265051705006613
[40] K. Höök. 2010. Transferring Qualities from Horseback Riding to Design. In *Proc. NordiCHI 2010, October 16–20, 2010, Reykjavik, Iceland.* 226–235. https://doi.org/10.3390/informatics5010008
[41] K. Höök, B. Caramiaux, C. Erkut, J. Forlizzi, N. Hajinejad, M. Haller, C. C. M. Hummels, K. Isbister, M. Jonsson, G. Khut, L. Loke, D. Lottridge, P. Marti, E. Melcer, F. F. Müller, M. G. Petersen, T. Schiphorst, E. M. Segura, A. Ståhl, D. Svanæs, J. Tholander, and H. Tobiasson. 2015. Embracing First-Person Perspectives in Soma-Based Design. *Informatics* 5, 8 (2015), 1–26. https://doi.org/10.3390/informatics5010008
[42] N. Igarashi, K. Suzuki, H. Kawamoto, and Y. Sankai. 2010. bioLights: Light emitting wear for visualizing lower-limb muscle activity. In *Proc. International Conference of the IEEE Engineering in Medicine and Biology, Buenos Aires, Argentina.* 6393–6396. https://doi.org/10.1109/IEMBS.2010.5627306
[43] D. Ihde. 1975. The Experience of Technology: Human-Machine Relations. *Cultural Hermeneutics* 2, 3 (1975), 267–279. https://doi.org/10.1177/019145377500200304
[44] A. R. Jensenius, V. G. Sanchez, A. Zelechowska, and K. A. V. Bjerkestrand. 2017. Exploring the Myo Controller for Sonic Microinteraction. In *Proc. New Interfaces for Musical Expression (NIME).*
[45] J. A. Jestley. 2011. *Metaphorical and Non-Metaphorical Imagery Use in Vocal Pedagogy: An Investigation of Underlying Cognitive Organisational Constructs.* Ph.D. Dissertation. University of British Columbia.
[46] P. Kantan and S. Dahl. 2019. Communicating Gait Performance Through Musical Energy: Towards an Intuitive Biodfeedback System for Neurorehabilitation. In *Proc. International Workshop on Interactive Sonification, Stockholm, Sweden.* 108–115. https://doi.org/10.5281/zenodo.3756783
[47] A. Kapur, S. Kapur, and P. Maes. Mar 7-11, 2018. AlterEgo: A Personalized Wearable Silent Speech Interface. In *Proc. Human Information Interaction & Retrieval (IUI '18), Tokyo, Japan.* 43–53. https://doi.org/10.1145/3172944.3172977
[48] J. Karolus, A. Kilian, T. Kosch, A. Schmidt, and P. W. Woźniak. 2020. Hit the Thumb Jack! Using Electromyography to Augment the Piano Keyboard. In *Proc. DIS 2020.* https://doi.org/10.1145/3357236.3395500
[49] P. E. Keller. 2012. Mental imagery in music performance: underlying mechanisms and potential benefits. *Annals of the New York Academy of Sciences* 1252, 1 (2012), 206–213. https://doi.org/10.1111/j.1749-6632.2011.06439.x
[50] Z. O. Khokhar, Z. G. Xiao, and C. Menon. 2010. Surface EMG pattern recognition for real-time control of a wrist exoskeleton. *BioMedical Engineering OnLine* 9, 41 (2010), 1–17. https://doi.org/10.1186/1475-925X-9-41
[51] B. Kleber, N. Birbaumer, R. Veit, T. Trevorrow, and M. Lotze. 2007. Overt and imagined singing of an Italian aria. *NeuroImage* 36, 3 (2007), 889–900. https://doi.org/10.1016/j.neuroimage.2007.02.053
[52] Y. Koike, K. Nakakoji, and Y. Yamamoto. 2006. Tele-kinesthetic interaction: using hand muscles to interact with a tangible 3D object. In *Proc. SIGGRAPH '06: ACM SIGGRAPH 2006 Emerging Technologies.* ACM, New York, NY, USA, 33–es. https://doi.org/10.1145/1179133.1179167
[53] S. M. Kosslyn, G. Ganis, and W. L. Thompson. 2001. Neural foundations of imagery. *Nature Reviews Neuroscience* 2, 9 (2001), 635–642. https://doi.org/10.1038/35090055
[54] M. Leman and P. Maes. 2015. The Role of Embodiment in the Perception of Music. *Empirical Musicology Review* 9, 3-4 (2015), 236–246. https://doi.org/10.18061/emr.v9i3-4.4498
[55] C. G. Lim, C-Y. Tsai, and M. Y. Chen. 2020. MuscleSense: Exploring Weight Sensing using Wearable Surface Electromyography (sEMG). In *Proc. Tangible, Embedded, and Embodied Interaction (TEI '20), February 9– 12, 2020, Sydney, NSW, Australia.* ACM, New York, NY, USA. https://doi.org/10.1145/3374920.3374943
[56] C. F. Lima, N. Lavan, S. Evans, Z. Agnew, A. R. Halpern, S. Meekings, D. Boebinger, M. Ostarek, C. McGettigan, J. E. Warren, and S. K. Scott. 2015. Feel the Noise: Relating Individual Differences in Auditory Imagery to the Structure and Function of Sensorimotor Systems. *Cerebral Cortex* 25, 11 (2015), 4638–4650. https://doi.org/10.1093/cercor/bhv134
[57] N. Loimusalo and E. Huovinen. July 5-9, 2016. Silent Reading and Aural Models in Pianists' Mental Practice. In *Proc. International Conference on Music Perception and Cognition (ICMPC), San Francisco, California, USA.* 609–614.
[58] J. MacRitchie and A. J. Milne. 2017. Exploring the Effects of Pitch Layout on Learning a New Musical Instrument. *Applied Sciences* 7, 1218 (2017), 1–19. https://doi.org/10.3390/app7121218
[59] C. P. Martin, A. R. Jensenius, K. A. V. Bjerkestrand, and V. Johnson. 2017. Stillness Under Tension: Performance for Myo armbands and Bela embedded computers. In *MusicLab Vol.1: Biophysical Music.*
[60] C. P. Martin, A. R. Jensenius, and J. Torresen. 2018. Composing an Ensemble Standstill Work for Myo and Bela. In *Proc. New Interfaces for Musical Expression (NIME).*
[61] M. Matsubara, H. Kadone, M. Iguchi, H. Terasawa, and K. Suzuki. 2013. The Effectiveness of Auditory Biofeedback on a Tracking Task for Ankle Joint Movements in Rehabilitation. In *Proc. International Workshop on Interactive Sonification, Erlangen, Germany.* 81–86.
[62] A. P. McPherson. 2017. Bela: An embedded platform for low-latency feedback control of sound. *Journal of the Acoustical Society of America* 141, 3618 (2017). https://doi.org/10.1121/1.4987761
[63] G. S. Meltzner, J. Sroka1, J. T. Heaton, L. D. Gilmore, G. Colby, S. Roy, N. Chen, and C. J. De Luca. 2008. Speech Recognition for Vocalized and Subvocal Modes of Production using Surface EMG Signals from the Neck and Face. In *INTERSPEECH 2008, 9th Annual Conference of the International Speech Communication Association, Brisbane, Australia, September 22-26, 2008.* 2667–2670.
[64] R. Miller. 1996. Imagery and the Teaching of Singing. In *On the Art of Singing.* Oxford University Press, Oxford, Chapter 1, 3–5. https://doi.org/10.1093/acprof:osobl/9780195098259.001.0001
[65] C. Neustaedter and P. Sengers. 2006. Autobiographical design in HCI research: designing and learning through use-it-yourself. In *Proc. DIS 2012, June 11-15, 2012, Newcastle, UK.* 514–523. https://doi.org/10.1145/2317956.2318034
[66] L. Nijs, M. Lesaffre, and M. Leman. 2013. The Musical Instrument as a Natural Extension of the Musician. In *Music and Its Instruments*, H. Castellengo, M. Genevois and J.-M. Bardez (Eds.). Editions Delatour France, 467–484.
[67] K. Nymoen, M. R. Haugen, and A. R. Jensenius. 2015. MuMYO — Evaluating and Exploring the MYO Armband for Musical Interaction. In *Proc. New Interfaces for Musical Expression (NIME).*
[68] J. O'Bryan. 2015. "We ARE our instrument!": Forming a singer identity. *Research Studies in Music Education* 37, 1 (2015), 123–137. https://doi.org/10.1177/1321103X15592831
[69] K. Okada and S. Hirai. 2019. Interactive Sonification for Correction of Poor Sitting Posture While Working. In *Proc. International Workshop on Interactive Sonification, Stockholm, Sweden.* 101–107. https://doi.org/10.5281/zenodo.3743339
[70] P. Q. Pfordresher. 2019. *Sound and Action in Music Performance.* Academic Press, London.
[71] P. Q. Pfordresher and A. R. Halpern. 2013. Auditory imagery and the poor-pitch singer. *Psychonomic Bulletin & Review* 20, 4 (2013), 747–753. https://doi.org/10.3758/s13423-013-0401-8
[72] D. Prem and R. Parncutt. 2008. Corporality in the timbre vocabulary of professional female jazz vocalists. In *M. M. Marin, M. Knoche, & R. Parncutt (Eds.), Proc. Students of Systematic Musicology (SysMus08), Graz, Austria, November 14-15, 2008.* 69–71.
[73] C. N. Reed and A. P. McPherson. 2020. Surface Electromyography for Direct Vocal Control. In *Proc. New Interfaces for Musical Expression (NIME 2020) Royal Birmingham Conservatoire, Birmingham, UK.* 1–6.
[74] D. Schön. 1984. *The Reflective Practitioner: How Professionals Think In Action.* Basic Books.
[75] P. Sengers and W. W. Gaver. 2006. Staying Open to Interpretation: Engaging Multiple Meanings in Design and Evaluation. In *Proc. DIS 2006, June 26–28, 2006, University Park, Pennsylvania, USA.* 99–108. https://doi.org/10.1145/1142405.1142422
[76] J. Sundberg. 1994. Perceptual Aspects of Singing. *Journal of Voice* 8, 2 (1994), 106–122. https://doi.org/10.1016/S0892-1997(05)80303-0
[77] A. Tanaka. 2015. Intention, Effort, and Restraint: The EMG in Musical Performance. *Leonardo: Transactions in Live Interfaces* 43, 8 (2015), 298–299. https://doi.org/10.1162/LEON_a_01018

[78] A. Tanaka and R. B. Knapp. 2017. Multimodal Interaction in Music Using the Electromyogram and Relative Position Sensing. In *A NIME Reader: Fifteen Years of New Interfaces for Musical Expression*. Springer.

[79] A. Tanaka and M. Ortiz. 2017. Gestural Musical Performance with Physiological Sensors, Focusing on the Electromyogram. In *The Routledge Companion to Embodied Music Interaction*. Routledge.

[80] M. Theiss, P. M. Scholl, and K. Van Laerhoven. 2016. Predicting Grasps with a Wearable Inertial and EMG Sensing Unit for Low-Power Detection of In-Hand Objects. In *Proc. Augmented Human International (AH '16)*. ACM, New York, NY, USA, 1–8. https://doi.org/10.1145/2875194.2875207

[81] W. H. Trusheim. 1991. Audiation and Mental Imagery: Implications for Artistic Performance. *The Quarterly* 2, 1-2 (1991), 138–147.

[82] Y. Tsubouchi and K. Suzuki. 2010. BioTones: A wearable device for EMG auditory biofeedback. In *Proc. International Conference of the IEEE Engineering in Medicine and Biology, Buenos Aires, Argentina*. 6543–6546. https://doi.org/10.1109/IEMBS.2010.5627097

[83] K. Tuuri, J. Parvainen, and A. Pirhonen. 2017. Who Controls Who? Embodied Control Within Human–Technology Choreographies. *Interacting with Computers* (2017), 1–18. https://doi.org/10.1093/iwc/iww040

[84] F. Varela. 1979. *Principles of biological autonomy.* Elsevier (North Holland), Amsterdam.

[85] F. Varela and E. Thompson. 1991. *The embodied mind: Cognitive science and human experience.* MIT Press, Cambridge, MA.

[86] F. Varela and E. Thompson. 2001. Radical embodiment: Neural dynamics and consciousness. *Trends in Cognitive Sciences* 5, 10 (2001), 418–425.

[87] R. Wakkary, D. Oogjes1, H. W. J. Lin, and S. Hauser. 2018. Philosophers Living with the Tilting Bowl. In *Proc. CHI 2018, April 21–26, 2018, Montréal, QC, Canada*. 1-12. https://doi.org/10.1145/3173574.3173668

[88] A. H. D. Watson. 2009. *The Biology of Musical Performance and Performance-related Injury.* Scarecrow Press, Lanham, Maryland, USA, Chapter The Voice: Management and Problems, 139–192.

[89] K. Woodward, E. Kanjo, S. Burton, and A. Oikonomou. 2018. EmoEcho: A Tangible Interface to Convey and Communicate Emotions. In *Proc. ACM International Joint Conference, International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*. ACM, New York, NY, USA, 746–749. https://doi.org/10.1145/3267305.3267705

[90] D. J. Wright, C. J. Wakefield, and D. Smith. 2014. Using PETTLEP imagery to improve music performance: A review. *Musicae Scientiae* 18, 4 (2014), 448–463. https://doi.org/10.1177/1029864914537668

[91] J. M. Zarate and R. J. Zatorre. 2008. Experience-dependent neural substrates involved in vocal pitch regulation during singing. *NeuroImage* 40, 4 (2008), 1871–1887. https://doi.org/10.1016/j.neuroimage.2008.01.026

[92] R. J. Zatorre, J. L. Chen, and V. B. Penhune. 2007. When the brain plays music: auditory–motor interactions in music perception and production. *Nature Reviews Neuroscience* 8, 7 (2007), 547–558. https://doi.org/10.1038/nrn2152

[93] R. J. Zatorre, A. R. Halpern, and M. Bouffard. 2010. Mental reversal of imagined melodies: A role for the posterior parietal cortex. *Journal of Cognitive Neuroscience* 22 (2010), 775–789. https://doi.org/10.1162/jocn.2009.21239

[94] J. Zimmerman, J. Forlizz, and S. Evenson. 2007. Research through Design as a Method for Interaction Design Research in HCI. In *Proc. CHI, ACM Press*. 493–502. https://doi.org/10.1145/1240624.1240704