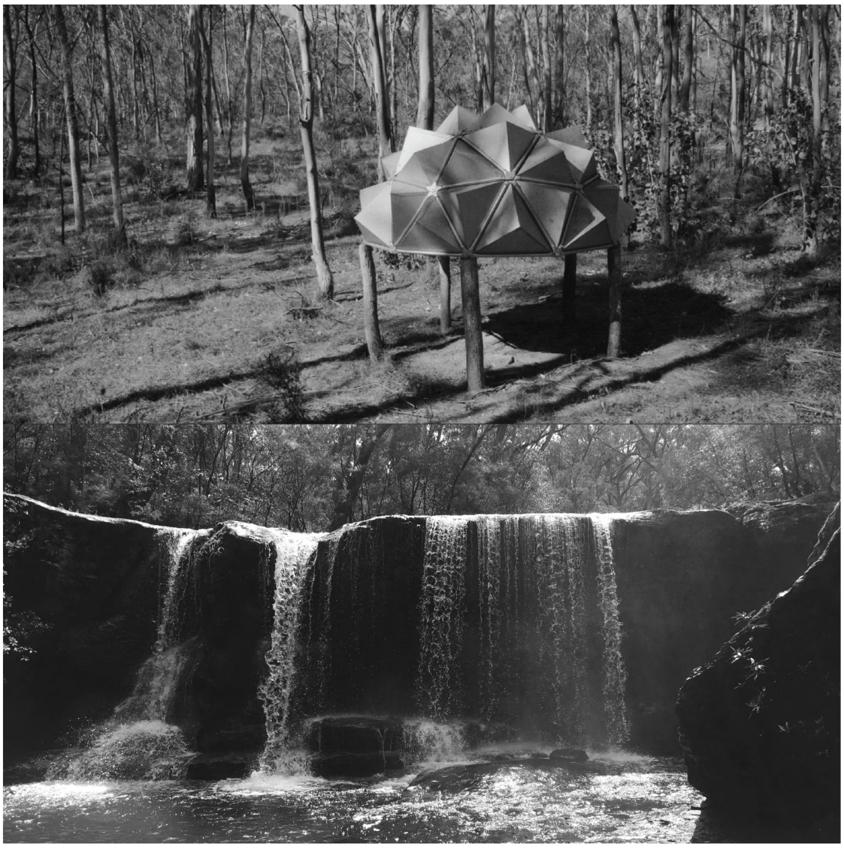
Retroreflection and waterfalls: speaking and singing in extreme acoustic environments.



Mary Rapp

The University of Sydney

School of Architecture, Design, and Planning

A thesis submitted in partial fulfilment of requirements for the degree of Doctor of Philosophy

2023

Declaration

This is to certify that to the best of my knowledge; the content of this thesis is my own work. This thesis has not been submitted for any degree or other purposes.

I certify that the intellectual content of this thesis is the product of my own work and that all the assistance received in preparing this thesis and sources have been acknowledged.

The experiment described in Chapters 2 was conducted as per the University of Sydney Human Research Ethics Committee (HREC) guidelines and was approved by the HREC under the protocol number 2019/101.

Mary Rapp

08/10/2022

Acknowledgements

To Dr Densil Cabrera, thank you for your wisdom, guidance, open mind, endless creativity, and for consistently repeating yourself until I finally grasped what you were trying to say.

To Dr Simon Barker, thank you for your friendship, support, and guidance throughout this project, and well before this project began.

To Bae Il Dong, thank you for being the most generous teacher and friend possible.

To Alison and Boyd, thank you for your generosity, passion, technical support, and your wonderful Burragorang Valley. This project really wouldn't have been possible without your help.

To Chris Abrahams and Peter Farrar, many thanks for your amazing creativity and musicianship.

To Lucy and the Crags community, thank you so much for lending me your waterfall.

To my love, Carl Dewhurst, thank you for your support, beautiful guitar playing, and sorry for all the tears.

To Hannah Czaban, thank you for being my friend: my voice of sanity and insanity in your role as both study and pandemic buddy.

To my family, I feel so lucky to have you.

To my friend family, you complete me.

Thank you to the Australian Government Department of Education for the support of the Skills and Employment Research Training Program scholarship.

This project is dedicated to my father who is the most resilient and loving man I know. Dad: despite what you think, this really does mean I'm a doctor and you must refer to me by that title at all times.

This thesis was proofread by Dave Sampson and Sam Gill. Sam is a researcher in a similar field and is a PhD candidate at the Sydney Conservatorium of Music. Dave and Sam provided advice following the University of Sydney's thesis and examination of higher degrees by research procedures 2020. The advice included (i) matters of structure (the need to structure and reword, deletions, additions); (ii) conventions of grammar and syntax; (iii) the use of clear language; (iv) logical connections between phrases, clauses, sentences, paragraphs, and sections; (v) voice and tone; and (vi) avoiding ambiguity, repetition, and verbosity.

Publications arising from this thesis

Some content of chapters 1 and 2 has been published as:

Rapp, M., Cabrera, D., Yadav, M., 2021, Effect of voice support level and spectrum on conversational speech, *The Journal of the Acoustical Society of America*, 150(4), 2635–2646.

Some content of Chapters 1 and 3 has been published as:

Rapp, M., Cabrera, D., Lu, S., 2022, A polytrihedral dome for acoustic retroreflection, and its application to creative-arts practice-led research, *Applied Acoustics*, 195(108860).

Some content of chapters 1, 2, and 4 has been published as:

Cabrera, D., Holmes, J., Lu, S., Rapp, M., Yadav, M., Hutchison, O., 2021, Voice support from acoustically retroreflective surfaces, in *Proceedings of the Euronoise 2021*, 25–27, October, Madeira, Portugal.

Some content of Chapter 5 has been published as:

Rapp, M., 2021, *Retroreflection and waterfalls* [sound recording], Bandcamp, https://maryrapp.bandcamp.com/album/retroreflection-waterfalls.

Abstract

This is a cross-disciplinary thesis consisting of both acoustics science research and practice-led research focused on improvised music performance. In very different ways, both types of research focus on how a person's surrounding environment affects voice production. The thesis describes two experiments: a behavioural laboratory talking experiment (one chapter) and a fifteen-month singing practice experiment (three chapters) with research outcomes including an original framework for singing practice and a recorded work of ensemble improvised music. Key aspects of both experiments are (i) vocal production response to the acoustic environment, (ii) acoustics that optimise the environment for a singer or talker (as opposed to listeners), and (iii) the acoustic environment created by a retroreflective array whereby sound is reflected back to the person. Two chapters of this thesis have been published in acoustics journals. The experiments can be summarised as follows:

Experiment 1. A behavioural laboratory experiment

A talking experiment that investigates how the sound of one's own voice (autophony), with alterations in gain and spectral balance, influences conversational speech production. Pairs of participants solved puzzles that elicit conversational speech in an anechoic room. The participants' own voice was fed back to their ears in real-time, with changes in voice level, and with one of three spectral filters applied: an All-pass, High-pass, or a Low-pass filter. Resulting voice levels and spectrum of the participants' speech were measured and analysed. The experiment addressed the question: can acoustic treatments with a frequency bias, like a retroreflective array, influence voice production, optimising a room for comfortable conversing?

Experiment 2. A practice-led singing experiment

A singing experiment was completed in three stages. Stage one (i) was the design of a practice framework that facilitated the discovery and creative development of singing techniques and musical materials. The design incorporated empirical research about autophony and voice production response to room reflections and environment sound. The foundation of the experiment design was also an approach to skill acquisition based on the ecological dynamics theory called the constraints-led approach. This stage also included the design, building, measurement, and analysis of a novel practice setting incorporating a retroreflective array. Stage two (ii) was fifteen-months of recorded self-regulated singing practice within the practice framework. Stage three (iii) applied the vocal skills and musical materials discovered and developed in stage two to recorded ensemble improvisation and further vocal training. Stage three also included analysis of the vocal production responses to the practice framework.

Contents

List of embedded of Figures

List of embedded video

List of embedded audio

List of embedded tables

Introduction

The acoustic environment alters all aspects of talking and singing. It alters the transmission, the perception, and, ultimately, how it is produced by a person. The acoustic environment can inspire people to make sound, like singing in the shower or creating an echo with yelling. It can also transform an experience like eating in a restaurant into something intolerably uncomfortable, battling the sound of others talking to converse with your dinner companion. Most architectural acoustics research considers the experience of the listener, including students in a classroom or the audience in a concert hall, examining if speech or singing is intelligible and of a sufficient volume. Comparatively little research considers how built and natural environments alter the sound of one's own voice (autophony), and how it can change voice production in conscious and unconscious ways. Research into this aspect of acoustics can be applied to optimise acoustic environments, contributing to the wellbeing and comfort of talkers, singers, and listeners. It also has applications outside the norms of acoustic science, for example, as in this thesis, to the discovery and development of new ways to make music.

This thesis investigates how the sound of one's own voice, altered by the acoustic environment, influences vocal production in both talking and singing, and how this mechanism can expand opportunities for creative development. The acoustic environment of primary interest is that which is created by a retroreflective array that reflects sound back to the source of the sound. This thesis outlines two experiments: (i) a behavioural talking experiment; and (ii) a practice-led research singing experiment which includes novel acoustic design and creative research outcomes.

This is a cross-disciplinary thesis that uses empirical research to answer questions applicable to architectural acoustics. The same empirical research is used in conjunction with theoretical frameworks for skill acquisition and creative-arts research methods to develop a unique framework for singing practice. The framework is applied to singing practice over fifteen months to instigate changes in voice quality and creative development of musical skills and materials for improvised music performance. Due to the cross-disciplinary nature of the thesis, the purpose of this introduction is to orientate the reader with a broad overview of how these disciplines are applied.

The prior empirical research into voice regulation and retroreflection that informs both the singing and talking experiments is outlined in Chapter 1. The behavioural talking experiment is described in Chapter 2 with an additional review of some prior research, experiment methods, results, and discussion. Chapters 3, 4, and 5 describe different aspects of the singing experiment. Chapter 3 begins with outlining the theoretical and empirical research (found in Chapter 1) that are both applied to the singing experiment design. Chapter 3 continues with the singing experiment design and methods, outlining the details of practice settings and procedures. Chapter 4 is a scientific analysis of aspects of the acoustics of the singing experiment practice settings, and the changes in the voice that occur in practice. Chapter 5 presents the singing experiment's creative outcomes including a taxonomy of new musical materials and recordings of ensemble improvised performance, demonstrating the application of the materials.

Thesis overview

Summary of the acoustic science applied in both experiments

The acoustic environment influences how people talk and sing, this influence is referred to here as voice regulation. The Lombard effect is a widely known voice regulation phenomenon, whereby talkers instinctively increase their vocal output to compensate for noise in the environment, maintaining a speech to noise ratio suitable for communication (Lane, Catania & Stevens, 1961). The Lombard effect is not only linked to communicative factors like background noise or distance from a listener, but also linked to the way a talker hears their own voice, described in this study as autophony. If a person talking cannot hear their own voice well, i.e., when autophony is disrupted or masked, they will increase their vocal output. Conversely, when autophony is high, vocal output decreases. Depending on the context, this voice regulation phenomenon is referred to as sidetone compensation (Lane, Tranel & Sisson, 1970) or the room effect (Pelegrín-García & Bruskog, 2012). The Lombard effect and sidetone compensation or the room effect are inversely related, described by Lane and Tranel (1971) as being two sides of the same coin. These voice regulation phenomena have been demonstrated to be systematically present, difficult to inhibit, and a talker/singer may not always be aware of it (Pick et al., 1989).

The voice that is returned to the speaker's or singer's ears via the sound reflections at room boundaries, indirect autophonic sound, varies from room to room, mostly determined by the characteristics of the reflective surfaces. Indirect autophonic sound can vary by the number of reflections, arriving at different delays, with different amplitudes, and spectral weightings.

The amount of sound reflected is referred to as voice support and can have a significant effect on the vocal output of a talker/singer. High voice support increases indirect autophonic sound and thus reduces vocal output (Lane, Tranel & Sisson, 1970). Conversely, low voice support increases vocal output. High levels of voice support can be provided in many ways, for example through a room's diffuse field (for small reverberant rooms), directed reflections, focussing, or with electro-acoustics. This thesis examines support that might be achieved with acoustically retroreflective arrays.

Retroreflection occurs when a surface reflects incident rays back to the direction from whence they came. An array of retroreflectors can provide high levels of voice support in a unique way where reflections of a talker's voice is returned only to the talker themself (Cabrera et al., 2018). This is a unique form of voice support that could reduce the vocal output of a talker or singer through high levels of indirect autophonic sound.

Retroreflective arrays do not reflect the whole spectrum of the voice due to size limitations of retroreflectors in a room and the size of sound waves of different frequencies. Research has demonstrated that for retroreflectors of a size suitable to applications in architectural acoustics, reflections are limited to a high frequency spectrum: generally 2 kHz and above (Cabrera et al. 2018; 2020; 2021; 2022).

Talking experiment

The first experiment in this thesis arises from the question; could an acoustic treatment with a frequency bias like a retroreflective array optimise the acoustic environment for a talker? Where as the room effect and side tone compensation have been demonstrated in previous research, this experiment examines these effects according to variations in the frequency ranges of room reflections. A behavioural experiment in a laboratory recorded participants conversing with the sound of their own voice fed back to their ears. The returned voice has a range of spectrum limitations and gains, measuring how the talker's vocal output changes.

Singing experiment

The second experiment in this thesis uses practice-led research alongside acoustics and voice regulation empirical research, not for influencing talking, but for (i) the discovery and development of new singing techniques and musical materials for use in improvised music, and (ii) changing the quality of my own singing voice (characterised here as spectral centroid). Acoustics science was used in conjunction with learning methodologies to investigate whether the environment in which one practices can positively influence the development of an original improvisational style.

A practice framework was designed and imposed on self-regulated singing practice. I am a professional musician and was the subject of the experiment thus the experiment can be described as autoethnography as I was both the experimenter and the subject (Adams, 2015). I completed three hundred and fifty-four hours of self-regulated singing practice over 15 months in two atypical settings with extreme and opposing acoustics: a waterfall with constant broadband noise, and a geodesic dome with a retroreflective array (here called a polytrihedral dome) that provided extreme levels of voice support. These atypical practice settings were chosen as environmental constraints for practice in line with the constraints-led approach, a theory for learning that was applied in this experiment.

The constraints-led approach is a pedagogical methodology for learning and skill acquisition, underpinned by ideas in the theoretical framework of ecological dynamics. It is a methodology for skill acquisition, development and performance based upon the understanding of the learner as adaptive and self-organising under interacting constraints (Chow et al., 2016; Davids, Button & Bennett, 2008). Constraints can be conceived as boundaries that shape self-organisation and can be separated into three categories: individual, environmental, and task constraints (Newell, 1986). Through the interaction of different constraints, a learner will self-organise in attempts to generate effective movement solutions (Renshaw et al., 2011).

The constraints-led approach has recently been applied to practice-led research for musical skill development (McMahon, 2022; Slater, 2020). This research focuses on constructing practice programs that consider and integrate all areas that affect performance, such as the performer's intention (i.e., motivations and goals), attention (i.e., being sensitive to and exploring environmental information and variables), and calibration (i.e., the attunement of performance actions to perceptual judgements). The objective of this research is to develop expert performance abilities through educating a performer's continual attunement and calibration to environmental information. This is achieved by creating information-rich learning environments for practice, guiding the learner towards individual, effective, and self-discovered performance solutions. The result is the development of flexible interactions with a dynamically changing environment through stable yet adaptable performance actions.

The singing experiment outlined in this thesis applies the constraints-led approach to improvised singing skill acquisition and development through the use of atypical performance contexts as an environmental constraint on practice activities. The polytrihedral dome and the waterfall were chosen as atypical practice settings in the constraints-led singing experiment for i) their extreme and opposing acoustics and ii) being information rich environments.

i) The voice regulation that occurs in extreme acoustics provided a unique way to impose changes in singing and thus the education of performance activities. The waterfall was used to induce a Lombard effect in singing and, inversely, the polytrihedral dome produced the room effect. The extreme acoustics and the application of voice regulation as an educational tool is a novel extension of a body of research that applies the constraints-led approach to music skill development and acquisition (McMahon, 2022; Slater, 2020).

ii) Outdoor practice settings are information-rich and dynamic, with information sources including auditory, visual, and haptic forms that continuously change. The experiment used outdoor practice settings for skill and material discovery and development via the continual attunement of performance actions to this rich and dynamic information.

The polytrihedral dome was designed and built for this experiment as a unique practice setting, the acoustics of which were contrasted to that of the waterfall. It is an unenclosed pavilion that maintains the assets of an outdoors practicing setting whilst providing extreme voice support that far surpasses everyday rooms.

The experiment has several research outcomes:

i) The framework for practice as detailed in Chapter 3. The framework is an experimental intervention on practice to expand opportunities for change in the voice, skill acquisition, and creative development of musical materials.

ii) The design and acoustic analysis of the polytrihedral dome and the analysis of both the voice regulation at practice settings as well as long-term changes in the spectrum of the singing voice (Chapter 4).

iii) The materials discovered and developed throughout the experiment, categorised as organisational devices or singing idioms. They are described in Chapter 5, with the idioms demonstrated by field recordings. At the completion of the experiment, new skills and materials were applied and explored in recorded works of improvised music duets. It is important to note that the singing experiment research contribution being made is oriented towards a conceptual framework for practice and performance, not musical analysis. As such, these recordings are presented as creative research outcomes, the recordings are not analysed.

Conclusion

In summary, this thesis outlines two experiments that have a multitude of research outcomes including empirical research, and creative contributions. Both experiments investigate acoustic retroreflection, an uncommon acoustic surface treatment in room acoustics which has significant potential for voice support applications. One aspect of this thesis is considering whether this treatment, and others with spectrum-limited voice support, could optimise the acoustics in commonplace rooms in which people converse. Another aspect is the use of a polytrihedral dome for a high voice support practice setting in a singing experiment. The research outcomes of this experiment include a new framework for singing practice that uses extreme acoustics and the constraints-led approach to learning for creative development of an improvised singing style and changes in voice quality. Acoustic analysis of practice settings is presented as a part of this framework. The framework was applied to singing practice over fifteen months. Changes to the singing voice throughout the experiment are presented as both creative outcomes and as empirical research on voice regulation.

Important terms

Autophony: the sound of one's own voice.

Voice regulation: the changes to vocal output due to the acoustic environment.

Voice support: the amount of the sound of one's own voice or autophonic support provided by a room via room reflections.

The Lombard effect: a voice regulation phenomenon where vocal output increases as environmental noise increases.

Room effect: voice regulation where vocal output decreases as voice support increases.

Retroreflection: where an incident ray is reflected by a surface back to the source.

Retroreflective array: a series of geometric shapes that reflect sound back to the source and can be used as a room treatment to increase voice support.

Chapter outline

Chapter 1: Review of the various aspects of prior empirical research on voice regulation, voice support, and retroreflection that provide the basis for both the talking and singing experiments.

Chapter 2: Aims, hypothesis, methods, results, and discussion of the behavioural talking experiment. Most of this chapter has been published in the Journal of the Acoustical Society of America (Rapp et al., 2021).

Chapter 3: Theoretical frameworks that are used in the singing experiment and how they related to the empirical research of Chapter 1. The starting point/background of the experiment. The methods of the singing experiment including practice timeline, settings, equipment, and procedures.

Chapter 4: A study of the acoustic analysis of the polytrihedral dome, built and used in the singing experiment, and an analysis of my singing voice, voice regulation in the singing practice, and long-term changes in my singing voice. Most of this chapter has been published in the journal Applied Acoustics (Rapp et al., 2022).

Chapter 5: A taxonomy of singing materials discovered and developed throughout the singing experiment, recordings of applied materials as creative research outcomes, and applications of acquired skill to further singing training.

Chapter 6: Consolidates and concludes the findings of the previous chapters.

Navigating the thesis

This project is presented as a multimedia eBook with audio and video illustrations appearing alongside narrative text. The audio and video are designed to support the communication of the project outcomes. Chapter 2 and 4 have been published as journal articles with additional authors (Rapp et al., 2021; 2022). The introductions to these chapters include citations of the publications and outline the other authors' contributions (including to particular sections of the text).

Research outcomes

The empirical acoustics research outcomes within this thesis are the following:

i) An investigation of how the spectral profile and gain of simulated indirect airborne sound, quantified as voice support (ST_V), affect the speaking voice of talkers (Chapter 2).

ii) An examination of acoustic retroreflection from a polytrihedral dome, a unique high voice support singing practice space (Chapter 4).

iii) An analysis of voice regulation in singing practice within environments with extreme and opposing acoustics and the long-term spectral changes in the singing voice whilst practicing in these settings (Chapter 4).

The practice-led research outcomes within this thesis are the following:

i) An addition to the body of practice-led research that applies a constraints-led learning approach to music skill acquisition and creative development (Chapter 3).

ii) A framework for practice that applies novel retroreflective acoustic design and voice regulation research as an original extension of constraints-led creative-arts research (Chapter 3).

iii) An original creative contribution presented as a recording of new works of improvised vocal performance in an ensemble context (Chapter 5).

Chapter 1: Autophony, voice regulation, voice support, and retroreflection empirical research

This chapter outlines the empirical research that provides the basis for the two experiments are detailed in this thesis: the talking experiment (Chapter 2) and the singing experiment (Chapters 3, 4, 5). It includes the prior research on voice regulation, acoustic parameters for autophony, and research on and examples of acoustic retroreflection. Some content of this chapter has been published in the following journal articles:

i) Rapp, M., Cabrera, D., Yadav., M., 2021, Effect of voice support level and spectrum on conversational speech, *The Journal of the Acoustical Society of America*, 150(4), 2635–2646.

ii) Rapp, M., Cabrera, D., Lu, S., 2022, A polytrihedral dome for acoustic retroreflection, and its application to creative-arts practice-led research, *Applied Acoustics*, 195(108860).

iii) Cabrera, D., Holmes, J., Lu, S., Rapp, M., Yadav, M., Hutchison, O., 2021, Voice support from acoustically retroreflective surfaces, in *Proceedings of the Euronoise 2021*, 25–27, October, Madeira, Portugal.

The content of this chapter will be referred to throughout this thesis and identified with section numbers.

1.1 Autophony, voice regulation and voice support

Vocal communication involves a dynamic interaction between the vocal signal of the talker, transmission system through which a message travels (such as a room) and the listener who interprets the message. Any changes to these components can lead to compensatory responses from a talker in the form of voice regulation. Voice regulation studies largely investigate different talking types and scenarios and it has been studied over the entire set of utterances from a talker such as insertion of pauses, changes in long-term spectra and mean fundamental frequency, and type of speech (Klatt, 1976; Krause & Braida, 2002; Lane et al., 1997; Traunmüller & Eriksson, 2000). Voice regulation has been demonstrated to be influenced by the speaking task, interlocutor status, and listener type (Anderson, 1991; Cheyne, 2009; Cooke et al., 2014a; Lane, Tranel & Sisson, 1970; Michael, Siegel & Pick, 1995; Pelegrín-García et al., 2011b; Summers et al., 1988; Williams & Stevens, 1972). It can be difficult to inhibit, systematically present, and a talking-listener is not always aware of it (Pick, 1989).

This dissertation investigates voice regulation as influenced by the sound of one's own voice (autophony). In particular, how changes in autophony via acoustics or background noise can lead to voice regulation in both speaking and singing.

1.1.1 Autophony

Autophony is propagated through three pathways (Figure 1.1):

i) Airborne direct sound from a talker's mouth to their ears (L_D).

ii) Airborne indirect sound via reflections at room boundaries (L_R).

iii) Bone conducted sound via the internal structures of the head to the cochlea.

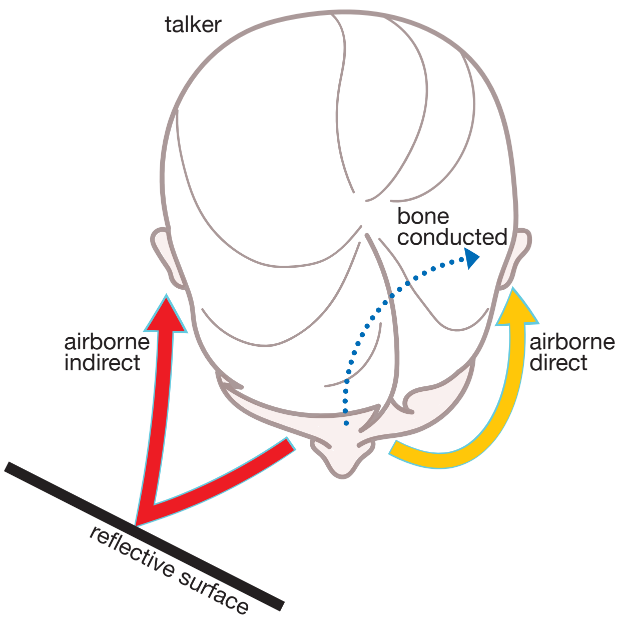


Figure 1.1. An illustration of the three pathways of autophonic sound.

The total airborne autophonic sound (L_E) is the sum of the direct and indirect sound ($L_E = L_D + L_R$, summed logarithmically because values are in decibels). While bone conduction and the long-term direct sound transfer functions from mouth to ears are approximately constant for a given individual, indirect autophonic sound can be altered in substantially different ways according to the acoustic characteristics of a room and/or surfaces.

1.1.2 The Lombard effect

The Lombard effect is a widely known voice regulation phenomenon, whereby talkers instinctively increase their vocal output to compensate for the noise in their environment (Lane & Tranel, 1971). Someone speaking louder than required while listening to music over headphones is an example of the Lombard effect. The Lombard effect can lead to increased voice intensities ranging from 0.1 to 1 dB per 1 dB increase in the background noise (Rapp, Cabrera & Yadav, 2021). The variation is due to the type of background noise, the speaking task and settings (distance, etc.), and the premium placed on intelligible communication for talkers addressing either a real or imagined audience and talkers in actual conversational settings (Cooke et al., 2014b;

Hodgson, Steininger & Razavi, 2007; Lane & Tranel, 1971; Lazarus, 1986), with values within the range of 0.5 – 0.7 dB/dB being used commonly as the Lombard slope (Rindel, 2010; Bottalico, 2018).

Raised voice in Lombard speech is associated with increased fundamental frequency (mean and range), first two formant frequencies, length and duration of utterances, number of pauses, and changes in speaking style from “conversational” to “clear” speech (Smiljanic, 2021). Lombard speech has been shown to change the spectral balance (SB) of speech (Smiljanic, 2021), defined as the ratio of higher-partial (2 – 4 kHz) to lower-partial (<2 kHz) speech energy (Collyer et al., 2009), with greater increases in higher-partial energy than in low-frequency components (Cooke et al., 2014b; Garnier, Henrich & Dubois, 2010; Sjölander & Sundberg, 2004). SB correlates with linguistic stress in speech (Sluijter & Van Heuven, 1996) and enhanced intelligibility in Lombard speech (Cooke & Lu, 2010; Garnier, Henrich & Dubois, 2010). It is also an indicator of voice projection or quality in classical singing (Collyer et al., 2007).

1.1.3 Side tone compensation

In the early days of telecommunication, researchers investigated a voice regulation phenomenon related to the Lombard effect, known as sidetone compensation (Black, 1954). Lane and Tranel (1971) identified that the Lombard effect and sidetone compensation are two sides of the same coin. A sidetone refers to presentation of airborne autophonic sound by electronic means, while the L_D is occluded in one or both ears. In this scenario, studies have shown that talking-listeners automatically compensate for a reduction in sidetone level by increasing their voice level, and vice-versa. Lane et al. (1961) showed that talkers increased voice level due to variations of ambient noise at an equivalent rate as sidetone compensation; ±0.5 dB in voice level for each additional dB of sidetone or ambient noise.

1.1.4 Voice regulation due to room acoustics

Indirect autophonic sound due to reflections (L_R) can be altered in substantially different ways according to the acoustic characteristics of a room and/or surfaces. Studies have shown that such changes to indirect autophonic sound can lead to a range of context-dependent compensatory responses from a talker in the form of voice regulation (Astolfi et al., 2015; Astolfi et al., 2019; Beechey, Buchholz & Keidser, 2018; Black, 1950; Bottalico, Astolfi & Hunter, 2017; Bottalico, Graetzer & Hunter, 2016; Brunskog et al., 2009; Cipriano, Astolfi & Pelegrín-García, 2017; Lane, Tranel & Sisson, 1970; Pelegrín-García & Brunskog, 2012; Pelegrín-García et al., 2011a). In one of the earliest studies of voice regulation due to room acoustics (Black, 1950) participants spoke several phrases in eight rooms to the experimenter 2.4 m away; participants spoke with reduced voice levels in rooms with longer reverberation times. This is consistent with more recent findings including Astolfi et al. (2019) wherein talkers increased their voice levels by 2 dB and around 3 dB in a semi-anechoic room and anechoic room, respectively, compared to a reverberant room in a task requiring effective communication to a listener 6 m away.

1.1.5 Stage support, voice support, and room gain

In 1989 Gade proposed the metric stage support (ST) to characterise the support a room (in this study a concert hall) provides to a performer of an instrument via room reflections (Gade, 1989). Early stage support (ST1, later denoted as ST_{Early} in ISO3382-1:2009) is a measure that assumes the sound returning to the performer delayed between 20 and 100 ms is beneficial (Jeon & Barron, 2004). ST_{Early} is expressed in decibels as the ratio of the early energy (20 – 100 ms) to the direct sound energy (0 – 10 ms), as shown in Equation 1.1 within which $p^2(t)$ is squared pressure as a function of time (i.e., the impulse response wave), which is integrated over time. Very early reflections (e.g., from stage furniture) in the 10 – 20 ms time period are excluded, and the floor reflection (within 10 ms) is included with the direct sound energy. To measure ST_{Early} an omnidirectional source and a microphone are placed on a stage, 1 m apart, and at performer height (1 m or 1.5 m for ST_{Early}, or 1.2 m for ST1). Higher values of support are generally considered preferable in large auditoria (Jeon & Barron, 2004).

$$\mathrm{ST1} = \mathrm{ST_{Early}} = 10\log\frac{\int_{.02}^{0.1} p^2(t)\,\mathrm{dt}}{\int_{0}^{.01} p^2(t)\mathrm{dt}} = L_{0.02-0.1\mathrm{s}} - L_{0-0.01\mathrm{s}} \quad (\mathrm{dB}) \qquad (1.1)$$

Two related metrics were proposed by Bruskog et al. (2009) to characterise the support a room provides not for performance of an instrument, but for talking and singing: room gain (G_{RG}) and voice support (ST_V). Instead of a loudspeaker and microphone separated by 1 m, a physical model of a person is used, with a mouth loudspeaker and ear microphones. G_{RG} (Equation 1.2) is the 'amplification' that a room provides to the sound of one's own voice, expressed as the difference between the total airborne autophonic sound and the direct autophonic sound. It has been shown to have a negatively sloped linear relationship with voice level (Brunskog et al., 2009; Cipriano, Astolfi & Pelegrín-García, 2017; Pelegrín-García et al., 2011a; Pelegrín-García et al., 2011b; Pelegrín-García, 2009; Rollins et al., 2019).

$$G_{\mathrm{RG}} = L_E - L_D \ (\mathrm{dB}) \qquad (1.2)$$

ST_V (Equation 1.3) is an alternative way of expressing G_{RG}, rescaling it over a larger range. It is also negatively correlated with voice levels (Cipriano, Astolfi & Pelegrín-García, 2017; Pelegrín-García & Brunskog, 2012; Pelegrín-García et al., 2011a; Pelegrín-García, Brunskog & Rasmussen, 2014).

$$\mathrm{ST}_V = L_R - L_D \ (\mathrm{dB}) \qquad (1.3)$$

The metrics ST_V and G_{RG} are linked through Equation 1.4.

$$G_{\mathrm{RG}} \approx 10 \log \left(10^{\frac{ST_V}{10}} + 1\right) \text{ (dB)} \qquad (1.4)$$

Early papers reporting ST_V values averaged the values over the six octave bands spanning 125-4000 Hz (Beechey, Buchholz & Keidser, 2018; Collyer et al., 2007; Cooke & Lu, 2010; Pelegrín-García et al., 2011a; Sluijter & Van Heuven, 1996). ST_V parameters (compared to G_{RG}) are used throughout this thesis because of their larger range.

1.1.6 Speech-weighted voice support

In order to obtain a single value descriptor of ST_V based on autophonic sound, Pelegrín-García et al. (2012) applied a frequency weighting to ST_V values in the octave band. The reference spectrum is the typical L_D of speech. This reference relates to a typical speech spectrum at 1 m in front of a talker and the SPL at their eardrums in free-field.

1.1.7 Oral-binaural measurements

G_{RG} and ST_V are derived from a room impulse response that measures the sound transmission from the mouth to the two ears of the same head: an oral-binaural room impulse response (OBRIR) (Cabrera et al., 2009). OBRIR measurements in a room involve a head and torso simulator (HATS) that radiates sound from a loudspeaker at the mouth that is recorded by microphones at the canal entrance of the ears.

1.1.8 Room effect

The slope of the linear relationship between voice level and G_{RG} or ST_V is referred to as the room effect (Equation 1.5).

$$room\ effect = \frac{\Delta\ voice\ levels\ (dB)}{\Delta\ G_{\mathrm{RG}} \text{ or } \mathrm{ST}_V} \qquad (1.5)$$

Pelegrín-García et al. (2011b) analysed the room effect of four acoustically different rooms on the speech produced by 13 male talkers who were describing a map to a listener at four distances in the absence of background noise. They found that talkers adjusted their voice level with each additional decibel of G_{RG} at a rate of -3.6 dB/dB (equivalently, -1.6 dB/dB of ST_V). Pelegrín-García et al. (2012) investigated the relationship between ST_V and voice levels in simulated rooms with three talking tasks. Talkers adjusted voice level linearly with ST_V from a rate of 0.1 to 0.93 dB/dB depending on the task (see below, Section 1.1.9).

1.1.9 Voice regulation and task

As seen in Pelegrín-García et al. (2012, 2011b) and Astolfi et al. (2019) voice regulation is not only influenced by ST_V or room acoustics in general but is also highly dependent on the instruction used on the individuals and the talking scenario. Similar to the variation in the Lombard effect slope (Lane & Tranel, 1971), many studies (Astolfi et al., 2015; Astolfi et al., 2019; Brumm & Zollinger, 2011; Garnier, Henrich & Dubois, 2010; Hazan & Baker, 2011; Junqua, Fincke & Field, 1999; Lane & Tranel, 1971) have argued that voice level regulation has a communicative basis, and thus the results of an experiment are affected by the communication scenario. For example, reading tasks not addressed to anyone in particular do not reflect many aspects of realistic communicative interactions and yield flatter slopes of voice regulation than tasks that require genuine communication (Hazan & Baker, 2011; Lane, Tranel & Sisson, 1970), such as puzzles that involve feedback between interlocuters or talkers addressing an audience. Voice regulation is governed by the interplay between the inherently interactive behaviour of communication, such as distance to the listener (Pelegrín-García et al., 2011b) or dynamic feedback (Schober & Clark, 1989).

1.1.10 Voice regulation and singing

Voice regulation in singing is much less reported than speaking. In studies on the Lombard effect in trained singers, voice level increased with the level of accompaniment of up to 0.17 dB/dB (Bottalico, Graetzer & Hunter, 2015; Coleman & Hicks, 1978; Tonkinson, 1994). Bottalico, Graetzer, and Hunter (2015) demonstrated a room effect in singers by measuring voice level in a soundproof room (ST_V of -18.1 dB), then with increased L_R provided by reflective panels (ST_V of -10.4 dB), and finally decreasing L_E with headphones. The room effect was evident with reductions in voice level of up to 1.17 dB with increased L_R and less evident in trained singers.

1.1.11 Voice regulation, effort, and comfort

The sensation of needing to increase voice level, or behavioural vocal effort (ISO, 2003), has been shown to reduce vocal comfort (psycho-physical response). Dead rooms with excessive absorption overdamp the voice of the speaker, increasing speech power and reducing vocal comfort (Astolfi et al., 2019; Nijs & Rychtáriková, 2011). Rooms with enhanced ST_V can lead to decreased vocal loading and thus increased vocal comfort (Pelegrín-García & Brunskog, 2012).

Voice regulation in response to room acoustical characteristics is of particular interest for professions with higher voice demands, for example teaching. Teachers have been shown to be at elevated risk of vocal injury, most likely due to the prolonged high vocal loading teachers experience at work (Åhlander, Rydell & Löfqvist, 2011; Bottalico, Graetzer & Hunter, 2016; Roy et al., 2004) and poor classroom acoustics often reported as a risk factor (Bottalico, Graetzer & Hunter, 2016; Hodgson, Rempel & Kennedy, 1999; Kob at al., 2006; Pelegrín-García & Brunskog, 2012; Pelegrín-García, Brunskog & Rasmussen, 2014). Pelegrín-García and Brunskog (2012) demonstrated that teachers' vocal comfort scores decreased with the perceived exhaustiveness of speaking in a classroom during a lesson, increased with perceived acoustic support, and decreased with the sensation of having to increase their voice level.

1.1.12 Optimising voice regulation with reflections

Acoustically reflective surfaces can be used to increase indirect autophonic sound, thereby increasing ST_V and potentially inducing the room effect. Bottalico, Graetzer, and Hunter (2016) demonstrated that acoustics can be improved for talkers through the placement of reflective surfaces that increase indirect autophonic sound, reducing voice levels and self-reported vocal effort. Certain acoustic treatments that bolster ST_V may be effective over a limited bandwidth. In particular, if relatively small reflectors are used, diffraction loss at low frequencies will limit their effectiveness to the high frequency range (Rindel, 1986). For example, studies of acoustically retroreflective surfaces (see Section 1.2.5) in architecture have speculated that they could help in reducing speaking levels, even though the unusually strong ST_V from such surfaces in the cases studied is mostly at and above the 2 kHz octave bands (Cabrera et al., 2018; Cabrera et al., 2020).

1.2 Retroreflection

Retroreflectivity is the concept that a surface reflects incident rays back to the source from which they came (Nilsen & Lu, 2004) (Figure 1.2). Architectural acoustics often considers surfaces for their specular reflectivity where the sound is reflected around the surface's normal (a vector perpendicular to the suface), scattering where the sound is scattered in all directions, and absorption when a material, structure, or object absorbs sound energy. Retroreflectivity is much less considered in architectural acoustics, with most prior research investigating incidental retroreflector arrays found on building façades or step wells (Crawford, 1991; Cabrera et al., 2018; Cabrera, et al., 2020; Cabrera, et al., 2022).

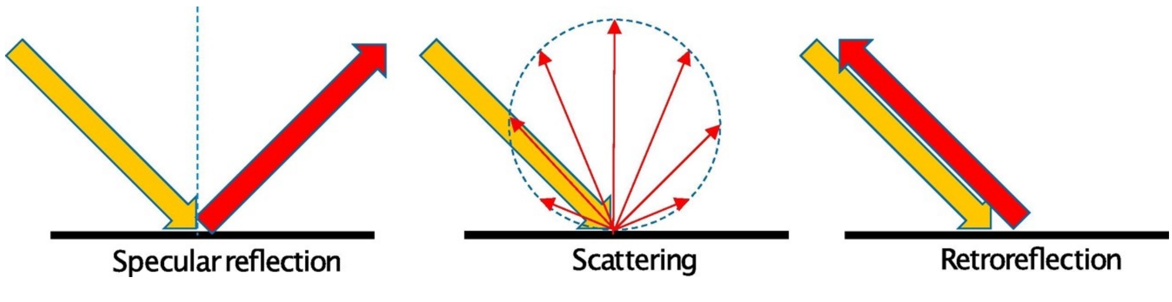


Figure 1.2. Illustration of three reflection types from Cabrera et al. (2018). The sound is reflected around the surface's normal for specular reflection. The sound is scattered in all directions for scattered reflection and is reflected back to the source for retroreflection.

Whilst intentional design of acoustic retroreflection is rare, it has sometimes been used in auditoria to provide acoustic support. A two-dimensional right-angled concave corner is retroreflective. Cox and D'Antonio (2009) noted that right-angled sawtooth surfaces, which can be considered to be 2-dimensional retroreflector arrays, have been used to provide acoustic support to the stage in auditoria. Tuominen, Rämö, and Välimäki (2013) demonstrated a single corner-cube retroreflector (Section 1.2.1) can be used as a replacement of electroacoustic monitoring in music performance when sufficiently large panels are placed at an appropriate distance from performers.

In contrast to architectural acoustics, retroreflective surfaces are of significant interest in optics and radar. They have practical applications in surveying, photoelectric switches, and radar targets (Shan et al., 2013; Somerstein, Fuller & Kennedy, 1989). Retroreflective high-visibility surfaces are experienced in everyday life in safety clothing and treatment of road markings and sign surfaces for driving at night.

1.2.1 Trihedral retroreflector

The simplest retroreflective surface is created by three orthogonal mirrors intersecting to create a concave cube corner. Depending on the source position, an incoming incident ray can be reflected back to the source by reflecting off one (first-order), two (second-order), or three (third-order) surfaces resulting in a reversal of direction (Newman, 2012) (Figure 1.3). Unlike specular reflection, the reflection back to the source occurs over a wide range of source positions and incident ray angles (Cabrera et al., 2018).

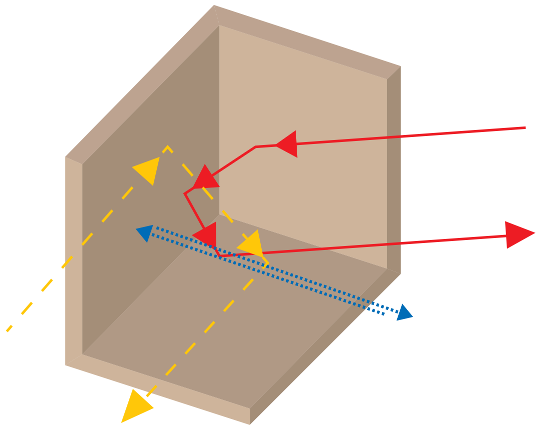


Figure 1.3. Drawing of a corner-cube retroreflector with first (blue), second (yellow), and third (red) order incident rays reflected back to the source.

In optics, the trihedral retroreflector is often referred to as a 'corner-cube reflector' or just a 'corner cube' (Cabrera et al., 2022). Arrays of corner-cube reflectors are used for optical treatments so that the reflected light around the source is intensified as each trihedron in the array reflects back to the source (Barrett & Jacobs, 1979; Wang et al., 2013). Applications of retroreflective arrays in optics include cat's eye road safety reflectors, vehicle and bicycle taillights, photo-electric sensor reflectors, and lunar ranging reflectors that measure the distance between the Earth and the moon (Currie, Dell'Agnello & Delle Monache, 2011). It is not unusual for retroreflective optical arrays to consist of hundreds or thousands of small corner cubes.

1.2.2 Retroreflective arrays for architectural acoustics

In the context of architectural acoustics, a single retroreflector may not be of interest as it returns just one of a vast multitude of reflections that will typically be received in a room (Cabrera et al., 2018). However, an array of retroreflectors, for example covering a wall, ceiling, or multiple surfaces, has the potential to create a strong 'focal region' of sound energy at the source.

This type of focussing is different to the coherent focussing provided by a spherical, hemispherical, or circular surface due to the variety of path lengths from the source to the retroreflectors (Vercammen, 2013). This means the sound returned by a retroreflector array will be spread out in time creating a temporally diffuse focal region (Cabrera et al., 2018). This kind of focussing can be described as 'energy focussing'. With energy focussing the pressure build-up at the focal region is less than an equivalent coherent pressure focal point of a spherical surface with the sound source at the centre (Cabrera et al., 2018). Another important difference is that with coherent focusing provided by a spherical surface, the sound field changes greatly as the source-receiver moves in the vicinity of the focal point (Cabrera et al., 2018). For example, the autophonic sound around the focal point of a hemispherical dome has strong 'phasing' and variations in 'colouration' with relatively small head movements (Cabrera et al., 2018). By contrast, the focal region of a retroreflective array follows the source as it moves. Whilst the sound does change at the focal point of a retroreflective array as the timing of multiple retroreflections change, these changes in the focal region are likely to be subtler when there is not a spatially-fixed coherent focal point as in the centre of a hemispherical dome (Cabrera et al., 2018).

A limitation to the application of a retroreflective array for audible-range acoustic focussing is that the wavelengths involved can be large, and thus the trihedron size needs to be correspondingly large to perform as a retroreflector (Cabrera et al., 2022). This introduces a trade-off for architectural scenarios in the design of retroreflective treatment. The number of reflectors and their size must be balanced for optimum practical design. Many small reflectors, as opposed to fewer large reflectors, should increase the high frequency reflected energy whith a trade-off of a raised high-pass cut-off frequency from diffraction loss (Cabrera et al., 2022).

1.2.3 Retroreflective room

An example of a room with retroreflective arrays was created by Cabrera et al. (2021) and can be seen in Figure 1.4. Fifty percent of the simplified surface area in a private office (with the dimensions of 5.34 × 3.65

× 2.81 m and a volume of 54.8 m3) was treated with square trihedral corner-cube retroreflector arrays (corner-cube edge lengths of 350 mm). The floor is hard and flat and a large amount of porous sound absorptive material was used on the ceiling (above the corner cube array) and two walls (behind and above the corner cube arrays) to minimise reverberant energy. The treatment yielded a reverberation time of 0.3 s in the 1 kHz octave band, and 0.4 s in the 2-8 kHz octave bands (the 2-8 kHz octave bands are those for which retroreflection is strong for this size of retroreflector). Analysis of measured impulse responses from a compact omnidirectional source to an array of 37976 receiver positions over a horizontal plane showed considerable focus of reflected energy close to the source position in the upper octave bands, especially 2 kHz and highe

Figure 1.4. Photograph of a retroreflective room from Cabrera et al. (2021). The orange steel boxes function as both acoustic retroreflectors and storage furniture. The furniture that is not retroreflective is wire mesh so as to minimally disrupt the sound field.

1.2.4 Retroreflective building façades

Research has examined retroreflection of corner-cube arrays in building façades (Cabrera et al., 2018; Cabrera, et al., 2020; Crawford, 1991) and has demonstrated that these façades reflect substantially more energy back to the source than a specular façade, especially at high frequencies. A distinctive case is the Ports 1961 Flagship Store in Shanghai which has an array of 255 trihedral corner-cube reflectors with edge lengths of 300mm (Figure 1.5). The façade was investigated by field impulse response measurements in its vicinity, laboratory measurements, and by finite-difference time-domain (FDTD) simulation of sound reflections from sources near the building façade (Cabrera et al., 2018). Results show that retroreflection from the façade has a strong effect in the 4 kHz and 8 kHz octave bands (wavelengths of 0.086 and 0.043 m respectively), is still clearly evident in the 2 kHz band, and is weakly present in the 1 kHz band when the source is close to the facade.

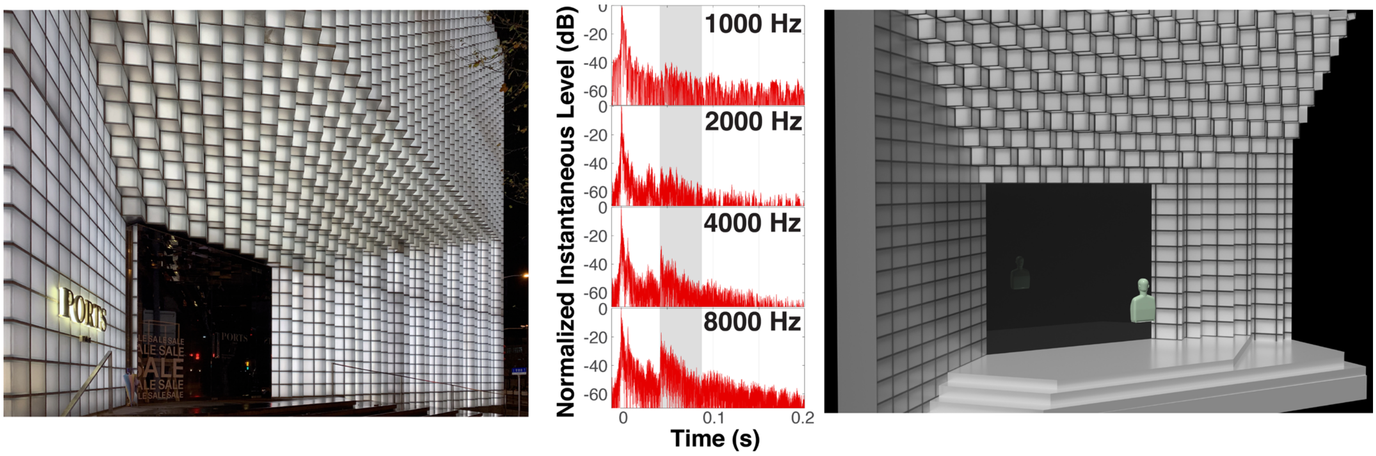


Figure 1.5. Left: a photograph of the Ports 1961 Flagship Store building façade in Shanghai from Cabrera et al. (2021). Middle: a visualisation of octave-band measured impulse response for a collocated source-receiver near the façade, grey shading identifying the retroreflection cluster time-period. Right: a CAD model of the façade and a HATS used in FDTD simulations.

1.2.5 Retroreflective array focussing for voice support and regulation

An array of corner-cube reflectors presents an opportunity for increasing voice support and thus inducing the room effect. The investigations into acoustic retroreflection from building façades (Cabrera et al., 2018; Cabrera, et al., 2020; Crawford, 1991) reported that at positions from which the buildings' corner-cubes were visible, acoustic retroreflection was distinctly audible when listening to one's own voice, especially with high frequency phonemes (\s\ and \t\), and when listening to one's own hand claps. Despite this, ST_V provided by the façades is weak compared to the level of ST_V typical of indoor environments which Brunskog et al. (2009) report to have values between -15 and -10 dB. By contrast, the retroreflective office (Cabrera et al., 2021) has higher than typical ST_V, with a measured speech-weighted ST_V of -4.1 dB when spatially averaged for various positions, orientations, and seated/standing heights (range spanning -5.5 dB to -2.8 dB). The spatially averaged speech-weighted ST_V value was 5.3 dB greater than the theoretical value calculated using diffuse field room acoustics assumptions according to Pelegrín-García and Brunskog (2012). The most important contribution to voice support in the retroreflective office is in the upper octave bands (-9.2 dB at 2 kHz and -6.3 dB at 4 kHz) (Cabrera et al., 2021).

Whilst a high ST_V value may indicate increased vocal comfort for a talker, it may also be associated with difficult conditions for speech transmission, for example due to excessive reverberation, or noise build up like Lombard-influenced babble (Cabrera et al., 2021). The potential benefit of acoustic retroreflection for voice regulation is that it can provide high levels of ST_V without relying on a strong diffuse field throughout a room or long reverberation. Voice support from retroreflection could induce the room effect, increasing vocal comfort, and returning sound to where it is needed in multitalker environments rather than contributing to speech babble build-up or potential speech distraction. The strong voice support provided by retroreflective arrays is typically in the high-frequency range. Studies of the room effect have not previously investigated the role of different parts of the spectrum in voice regulation.

1.3 Conclusion

In conclusion, the prior research outlined in this chapter demonstrates that voice regulation occurs due to autophonic sound as altered by the acoustic environment. A room can be characterised by acoustic parameters that measure how a room is experienced by a listening-talker or a singer, and a retroreflective array can be applied as a room treatment that can modify autophonic sound, thus inducing voice regulation. The research, acoustic parameters, and voice regulation effects described in this chapter are applied in the two experiments outlined in the following chapters. The content of this chapter, identified with section numbers, will be referred to throughout this thesis.

Chapter 2: Talking experiment - effect of voice support level and spectrum on conversational speech

This chapter outlines a talking experiment investigating voice regulation (Section 1.1) and the spectrum of simulated indirect autophonic support (Section 1.1.1). Most of this chapter has been published as:

Rapp, M., Cabrera, D., Yadav, M., 2021, Effect of voice support level and spectrum on conversational speech, *The Journal of the Acoustical Society of America*, 150(4), 2635–2646.

The experiment was designed by both myself and Dr Densil Cabrera. I carried out the experiment. Dr Manuj Yadav did the statistical analysis of the results. All authors contributed to the overall analysis and interpretation and to the editing of the text, and these were further developed in response to the suggestions by anonymous reviewers.

Abstract

One's own voice (autophony) is transmitted to the ears as direct airborne sound, bone conduction, and indirect airborne sound from reflections characterized by overall gain and spectro-temporal features. The experiment outlined in this chapter investigates how the spectral profile and gain of simulated indirect airborne sound, quantified as voice support (ST_V), affect the speaking voice of talkers. Pairs of participants performed a conversation elicitation task in anechoic conditions. The indirect airborne sound was provided in real-time via open headphones that maintain the direct airborne transmission path. Experimental conditions included high-pass, low-pass, and all-pass versions of ST_V, each presented at three overall gains, and a Baseline condition with no electroacoustic contribution to ST_V. The results show an overall speech level reduction of 0.22 dB for every additional dB of speech-weighted ST_V, i.e., a –0.22 dB/dB slope. There was some effect of ST_V spectrum on speech: slope for the high-pass condition was steeper (statistically significant) and significantly different from the all-pass slope; spectral balance (2–4 kHz vs 0–2 kHz) of speech showed an interaction effect between gender and experimental conditions. This experiment's findings may inform acoustic treatments in environments where overall sound reduction is of interest for favourable ergonomics and occupational health for voice professionals.

2.1 Introduction

Section 1.1.1 of this thesis introduces the term indirect autophony: the sound of one's own voice as transmitted to the ears from reflections. Section 1.1.4 outlines how indirect autophony leads to compensatory responses from a talker in the form of voice regulation. This experiment investigates how the spectral profile and gain of simulated indirect airborne autophonic sound, quantified as voice support (ST_V, Section 1.1.5), affects conversational voice characteristics.

Investigations of the effects of room acoustics on speech production began with and continue to examine the acoustic parameter of reverberation time (Black, 1950; Bottalico, Asolfi & Hunter, 2017). In an early study of voice regulation, Black (1950) labelled rooms as "live" or "dead" to characterise the reverberation time of a room, reporting that participants spoke with reduced voice levels in live rooms and increased voice levels in dead rooms. Astolfi et al. (2019) measured participants describing maps and giving a monologue addressing a listener 6 m away in two rooms with very long and very short reverberation times. The speech was measured using a contact-sensor-based device fixed at the base of the neck to disambiguate the build-up of sound provided by the room's reverberation from the voice regulation effect. A 2 dB increase in the sound pressure level (SPL) of speech was observed in a semi-anechoic room compared to a reverberation room when describing maps. A parabolic relationship with teachers' speech SPL and classroom reverberation times in the range $0.4 \text{ s} \leq T30 \leq 1.4 \text{ s}$, with a minimum SPL at $T30 = 0.7\text{s}$, was demonstrated by Puglisi et al. (2017). By contrast, Brunskog et al. (2009), when measuring free speech in six different rooms, found no correlation with speech level and reverberation time, instead finding a correlation with speech level and room gain (Section 1.1.5). In a study by Pelegrín-García et al. (2012), the correlations of acoustic characteristics (voice support, reverberation time, and decay time) with the voice levels of talkers in simulated rooms showed that ST_V had the strongest correlations and was thus the best predictor of voice level.

Pelegrín-García et al. (2012) introduced the term 'room effect' for the linear relationship between voice level and room gain or voice support (Section 1.1.8). It was demonstrated by measuring teachers and students speaking in laboratory experiments under ten different simulated room acoustic conditions, with variations in voice support. The room effect was shown to be highly dependent on the talking task. The tasks included a free-speech three-minute lecture addressing an imaginary group of 30 students, reading text addressing a dummy head 2 m away, and describing paths on maps to an experimenter 3 m away. The average room effect was -0.93 dB/dB of ST_V for free-speech, -0.23 dB/dB for reading text, and -0.10 dB/dB for map reading.

The experiment talking task has been shown to have a significant impact on the magnitude of the room effect, with tasks placing a premium on communication associated with larger changes in voice level. Astolfi et al. (2019) found greater changes to voice level due to room acoustics when talkers were tasked with providing directions from a map, compared to speaking freely for five minutes on a topic of their choosing. The study attributed this difference to the communicative intent of the tasks. By contrast, Pelegrín-García et al. (2012) found map reading to yield smaller changes in voice level than other tasks with free speech or reading a text.

Several studies on voice regulation have considered speech produced during conversations. Cooke and

Lu (2010) and Hazan and Baker (2011) measured acoustic phonetic changes in speech during conversational interactions. Both studies measured conversations between separated talkers: one into different booths and the other with an acoustically transparent screen. Relatively few studies have considered speech modifications that occur within conversations between co-located talkers. Aubanel et al. (2011) measured Lombard speech with pairs of talkers sitting at a table without any visual obstruction. Beechey et al. (2018) measured the speech of co-located talkers, using a novel conversation elicitation framework (applied in the present study) to measure realistic, dynamic conversational speech in the presence of realistic background noise.

Voice regulation has been studied in the context of teaching due to the correlations of vocal loading, comfort, fatigue, and voice injuries necessitated by teaching (Section 1.1.11). School teachers often speak in noisy environments for long periods of time, and they also present with voice problems more often than the average person (Angelillo et al., 2009). Poor classroom acoustics has been identified as a potential hazard for the voice health of teachers (Vilkman, 2000). Brunskog, Astolfi, and Hunter (2017) demonstrated that the vocal load of teachers, which can lead to vocal fatigue, is influenced by classroom reverberation time. Pelegrín-García and Brunskog (2012) measured the room effect for teachers with and without self-reported voice problems under simulated classroom acoustics. The average room effect for teachers with healthy voices was -0.12 dB / dB of ST_V, whereas for teachers with self-reported non-healthy voices it was -0.07 dB / dB. The study demonstrated that teachers with voice problems perceive their environment differently to teachers without voice problems, preferring talking in rooms with longer decay times.

A study by Bottalico, Graetzer, and Hunter (2016) increased self-reported vocal comfort of talkers by increasing autophonic sound at the talker position with a reflective panel that created a strong first reflection. Students' speech was measured in an anechoic, semi-reverberant, and reverberant rooms, at normal and loud speaking styles, and in the presence of classroom babble noise. Speech SPL and self-reported effort increased with the loud style and decreased with the reflective panels. Self-reported comfort and control decreased in the loud style, while self-reported clarity increased when panels were present.

Certain acoustic treatments that can bolster autophonic sound, like the panels used by Bottalico, Graetzer, and Hunter (2016), may be effective over a limited bandwidth. In particular, if relatively small reflectors are used, diffraction loss at low frequencies will limit their effectiveness to the high frequency range (Rindel, 1986). For example, studies of acoustically retroreflective surfaces in architecture (Section 1.2.4) have speculated that a retroreflective array could help in reducing speaking levels, even though the unusually strong ST_V from such surfaces in the cases studies is mostly at and above the 2 kHz octave bands (Cabrera et al., 2018; Cabrera et al., 2020). However, I am not aware of prior studies of voice level regulation from ST_V restricted to the high frequency range. This chapter asks the question of whether ST_V restricted to either low or high frequencies has a similar effect as broadband ST_V.

2.2 Method

An experiment was conducted in an anechoic room (6.5 m × 3.6 m × 3.0 m height), in which pairs of participants conversed in ten conditions (Figure 2.1). Experimental conditions were presented via open headphones to both participants in five-minute blocks during which they jointly solved puzzles designed to elicit conversation. Participants' voices were recorded to determine voice regulation response to ST_V levels and spectral weightings. The protocol was approved by the University of Sydney Human Research Ethics Committee.

Figure 2.1. A photograph of two people posing as participants wearing ear loudspeakers and head-mounted microphones in the anechoic room, as per the experiment setup.

2.2.1 Participants

Twenty-two participants (14 males and 8 females, ages 22 – 45, median age 31) took part. Participants reported no known hearing or voice problems and were fluent in English. They were naïve to the purpose of the study.

2.2.2 Conversation task

The experiment tested pairs of participants engaged in a conversation elicitation task designed and qualified by Beechey, Buchholz, and Keidser (2018) (Figure 2.2). This task elicits fluent and naturalistic conversations while ensuring both participants provide approximately equal speech data, broadly representative of everyday communication (Beechey, Buchholz & Keidser, 2018). For the experiment, five puzzles were constructed on 10 × 10 grids (on A4 paper) with each square containing a tangram image and one of three colours.

For each puzzle, two complementary participant views were created by removing half of the information from each, either the image or the colour. Participants were instructed to find the unique path from the top left-hand square to the diagonally opposite end square by moving horizontally or vertically between squares containing the same colour or image. They could not look at the other's puzzle to achieve this but spoke freely to gain the required information. Tangram images were chosen by Beechey, Buchholz, and Keidser (2018) as they require detailed descriptions. Completion of the task was not a measure of interest.

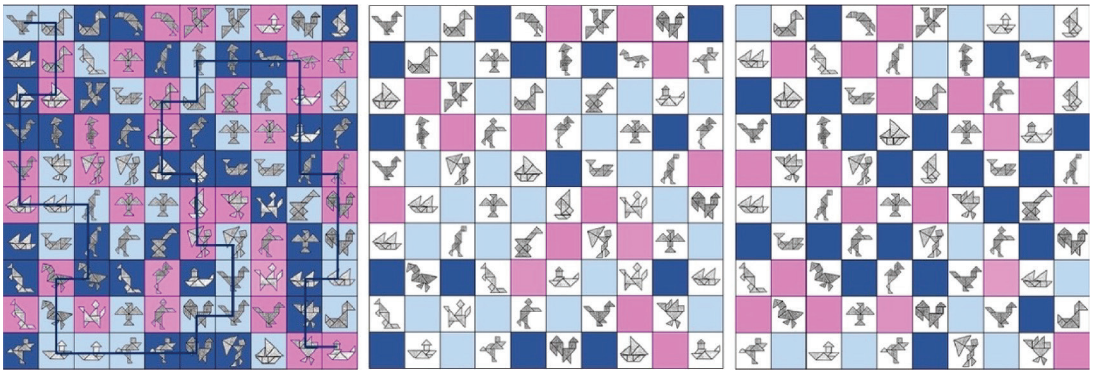


Figure 2.2. Examples of conversation elicitation task designed by Beechey, Buchholz, and Keidser (2018). Left: complete puzzle with solution, unseen by participants. Centre and right: participant view panels (one each and unseen by the other) that have half the information, either the image or the colour.

2.2.3 Experimental setup

The experiment used hardware from an auralisation system that has been used in previous studies of autophonic perception (Yadav, Cabrera & Martens, 2012; Yadav & Cabrera, 2017). The system generated electroacoustic autophonic support with the voice of a talker in real-time. The experiment took place in an anechoic room (Figure 2.1) that has a graded flow-resistivity sound-absorptive lining (Xu, Buchholz & Fricke, 2005), which is anechoic at and above 200 Hz. The floor has the same anechoic termination beneath a metal grill. No reflective floor was added. The room has an $L_{A,eq}$ of 7 dB during normal working hours with building services operating (data from Brüel & Kjær (Nærum, Denmark) type 4179 low-noise microphone system). Pairs of experiment participants were seated facing each other, 4 m apart, holding a clipboard with printed puzzle and a puzzle instruction sheet (Figure 2.1 and 2.2). Based on pilot studies, the 4 m distance between participants was chosen to avoid a too-easy task in the essentially noise-free conditions. As participants spoke, their voice was picked up by an omnidirectional head-mounted microphone (DPA 4066, DPA Microphones, Alleroed, Denmark). If the centre-point of the lips, as viewed from the front of the face, is the origin of a Cartesian plane, the head-mounted microphone was positioned 7 cm from the origin on the abscissa. The audio signal was sampled at 48 kHz by an RME Fireface (RME, Haimhausen, Germany) connected to a computer running the MAX software (Cycling '74, San Francisco, CA). The signal was altered by the software according to a set of ST_V levels, spectral filters, and a reverberation effect, all outlined in Section 2.2.4. The resulting signal was converted into the analogue domain and returned to the talkers' ears via AKG (Los Angeles, CA) K1000 open-ear loudspeakers that maintain direct autophonic sound (< 0.5 dB difference in the 125 Hz – 8 kHz octave bands) (Yadav, Cabrera & Martens, 2012).

Each experimental condition was 5 minutes (10 conditions, 50 min in total), each with an indirect autophonic gain condition (Section 2.2.4) selected randomly per participant, meaning the pair usually did not have the same condition at the one time. Participants were instructed at the onset of the experiment that they could break at any time. However, as reported by Beechey, Buchholz, and Keidser (2018) the task was found to be engaging to the extent that participants typically expressed a desire to stay in the anechoic room to complete the puzzles after the required duration of recordings had been reached.

The only other information that participants received was instructions on how to complete the task. There was a talkback system for communication with the experimenter, who was in a separate room. Audio recordings of each participant's speech signal from the headset microphones and a Brüel & Kjær omnidirectional microphone (1/4–in. type 4135) halfway between the participants at a height of 1 m were stored on the computer in 5–minute increments, aligned with changes in experimental condition. The microphone was on–axis at 2 m allowing calculation of the equivalent 1 m on-axis speaking levels (referred to in studies and standards as an objective measure of voice effort (Nijs & Rychtáriková, 2011)) and spectra. The recordings from this microphone were the main data for analyses.

2.2.4 Voice support conditions

Nine distinct indirect autophonic gain conditions with electroacoustics processing were used. The tenth was the Baseline condition with no indirect autophonic gain provided electroacoustically (Table 2.1).

Table 2.1. The octave-band, average (Av.) and speech-weighted (SW) ST_V values of the ten experiment conditions, categorised in terms of Filter and nominal voice support (*VS*) groups. The reference speech sound pressure level (SPL) at the eardrum, used for speech-weighting, is also shown. Values are in decibels.

Filter	VS	$ST_{V\,(125)}$	$ST_{V\,(250)}$	$ST_{V\,(500)}$	$ST_{V\,(1k)}$	$ST_{V\,(2k)}$	$ST_{V\,(4k)}$	$ST_{V\,(8k)}$	Av. ST_V	SW ST_V
All-pass	Strong	-6.0	-5.0	-4.2	-2.0	3.5	1.8	9.0	-2.0	-1.3
	Moderate	-7.7	-7.8	-7.1	-4.9	0.5	-1.3	6.0	-4.7	-4.2
	Weak	-9.7	-12.5	-12.9	-10.8	-5.5	-7.2	0.0	-9.8	-10.0
High-pass	Strong	-14.3	-24.3	-24.1	-4.4	3.5	1.8	9.0	-8.9	-3.3
	Moderate	-14.3	-24.3	-24.1	-7.4	0.5	-1.3	6.0	-10.4	-6.2
	Weak	-14.3	-24.3	-24.1	-13.3	-5.5	-7.2	0.0	-13.3	-11.9
Low-pass	Strong	-6.0	-5.0	-4.2	-5.6	-20.7	-29.6	-25.9	-11.9	-5.6
	Moderate	-7.7	-7.8	-7.1	-8.6	-21.8	-29.6	-25.9	-13.7	-8.5
	Weak	-9.7	-12.5	-12.9	-14.5	-21.8	-29.6	-25.9	-16.8	-14.0
Baseline		-14.3	-24.3	-26.8	-22.8	-22.2	-29.9	-25.9	-23.4	-23.9
Reference SPL		58.0	69.1	73.5	71.7	69.0	63.0	-	-	-

The electroacoustic indirect autophonic gain was provided by Freeverb~ (Böhm, 2021) (hosted in the software MAX), which is an artificial reverberation processor based on the Schroeder (1962)/ Moorer (1979) reverberation model, providing a simple exponential energy decay. To this artificial reverberation, six additional 'early reflections' were added. The resulting impulse response was used for real-time signal processing in all non-Baseline experimental conditions, which were further distinguished by their gain and filter settings set within MAX. Apart from the Baseline condition, reverberation time was T_{30} = 0.7s in the 125 Hz – 2 kHz octave bands, with shorter reverberation times in the 4 and 8 kHz bands (0.6 and 0.5 s, respectively). The 0.7 s T_{30} was chosen to correspond with previous research (Pelegrín-García, Brunskog & Rasmussen, 2014; Puglisi et al., 2017), which finds that between 0.6 and 0.7 s is the optimal reverberation time in a classroom that meets criteria of vocal comfort and speech intelligibility. See Section 2.2.5 for an example of an oral-binaural impulse response.

ST_V values for the experimental conditions were measured in the anechoic room using a head and torso simulator (HATS; Brüel & Kjær type 4218-C) wearing the AKG K1000 open-ear loudspeakers and DPA 4066 head-mounted microphone as per the above experimental setup. A swept sinusoid was emitted from the HATS mouth simulator, and the L_D and the L_R from the MAX patch fed through the headphones were recorded by the microphones in each ear of the HATS. Once converted to an oral-binaural impulse response (Section 2.2.5), the

L_R was separated from the L_D to calculate ST_V using the window function proposed by Pelegrín-García (2011).

The octave-band measured, averages, and speech-weighted ST_V values (Pelegrín-García et al., 2012) of the experimental conditions are given in Table 2.1. Three filtered electroacoustic indirect autophonic conditions were used: All-pass (AP), Low-pass (LP) and High-pass (HP). Starting with the AP condition, average ST_V values of approximately 2, 5, and 10 dB (based on 125 Hz – 4 kHz octave-band average) were used, nominally referred to as Strong, Moderate, and Weak (Table 2.1) ST_V, respectively, and, along with Baseline, collectively referred to as *VS* groups. The ST_V values of the AP condition were chosen to be representative of real rooms (Pelegrín-García, 2011) without amplification (-10 dB, -5 dB) and a higher value (-2 dB) chosen to explore the possible effects of environments with extreme ST_V. HP and LP filters were then applied, to yield HP and LP variants of the Strong, Moderate, and Weak *VS* (Table 2.1). The HP and LP filters, implemented in the software MAX, had a cut-off frequency of 1 kHz and a roll-off of 18 dB/octave. Hence these variants offered little or no ST_V outside of their passband but the same amount of ST_V as the AP Strong, Moderate, or Weak *VS* groups in their passband (Table 2.1). The use of 1 kHz as the pivot frequency between HP and LP spectral processing had two main reasons: in the authors' experience with real retroreflective arrays, the octave bands 2 kHz and above can be strongly boosted by this type of treatment, and in one sense 1 kHz is the mid-frequency of the 125 Hz – 8 kHz octave-band range that is often used to represent speech. The 18 dB/octave roll-off provides strong frequency selectivity without excessive ringing. Note the octave-band spectrum values of the AP condition are simply the result of gain, in conjunction with the equipment's unmodified transfer functions. Table 2.1 also provides the octave-band reference SPL at the eardrum that was proposed by Pelegrín-García et al. (2012) for speech-weighting of ST_V, which was used to calculate the speech-weighted values in Table 2.1.

2.2.5 Example of an oral-binaural impulse response

Figure 2.3 provides a visualization of a measured oral-binaural impulse response (OBRIR) for one of the experiment conditions (AP, Strong *VS* and unweighted ST_V of -2 dB). Audio 2.1 provides the audio of the OBRIR. The direct sound from mouth to ears is set to 0 ms. Discrete delays are in the 12-35 ms period, merging into reverberant decay. The oral-binaural impulse response was measured using a head and torso simulator (Brüel & Kjær type 4128C) wearing the AKG K1000 open-ear loudspeakers and DPA 4066 head-mounted microphone as per the experiment setup.

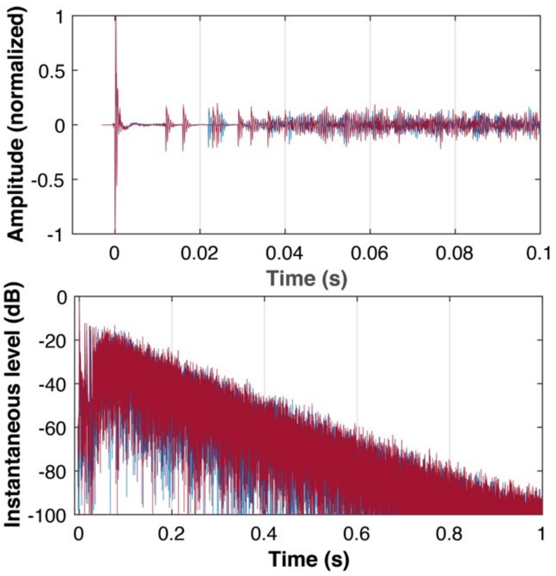


Figure 2.3. Waveform (top) and instantaneous level (bottom) of a measured oral-binaural impulse response that was used in the experiment (All-pass, Strong *VS* and unweighted ST_V of -2 dB).

Audio 2.1 Audio of a measured oral-binaural impulse response that was used in the experiment.

2.2.6 Processing of speech recordings

A MATLAB script identified, selected, and measured windowed speech recordings from the 2 m microphone signal. A 20 ms Hamming window function was applied to all recordings (after relevant filtering, such as octave band, A-weighting, etc.), including the calibration tone, by filtering the squared window function. Automatic comparisons between the A-weighted windowed recordings were used to identify which speech belonged to which participant for the 2 m microphone recording. Windows when one person was speaking without the other speaking were identified using a level threshold with the headset microphone recordings, and windows when neither person was talking were excluded. This threshold was determined by visually inspecting the recorded time-varying speech envelopes. Windows in which both participants were talking were also excluded. The speech L_{eq} of all selected windows was calculated from the 2 m centre microphone, which were converted to 1 m equivalent (labelled $L_{Aeq,1\ m}$) using the distance law for SPLs (no air attenuation). SB, defined as the ratio of energy in the 2 – 4 kHz to the 0 – 2 kHz bands, was used to examine the relative proportion of higher vs lower frequency content in the vocal output. Analysis of SPLs of individual vowels (/a/, /i/, and /u/; corner vowels) was also performed. However, the results did not vary across conditions, which was also statistically verified. Hence, this analysis is not discussed further.

2.2.7 Statistical analysis

Since each participant performed all ten experimental conditions (Table 2.1), this represents a repeated-measures experimental design, wherein results for each participant for the experimental conditions are likely to be intercorrelated to some degree. Furthermore, since each experimental session involved a conversation between two participants, some correlation is likely between results of a pair of participants within the same session. These non-controlled sources of variance or *random-effects*, were explicitly modelled in mixed-effects models that were used to study the relationship between the *fixed-effects*, i.e., the independent variables represented by the experimental conditions. The dependent variable included either the $L_{\text{Aeq,1 m}}$ or the SB.

Two types of statistical analyses were conducted, varying in their representation of ST_V as a fixed-effect: using the *VS* groups or using the actual ST_V values. For the former, the fixed-effects included the three Filter groups (AP, LP, HP), the *VS* (Strong, Moderate, Weak, Baseline), and the Gender (female, male) of the participants. For the latter, besides the three Filter groups and Gender, the other fixed-effect was ST_V; both average and speech-weighted values (Table 2.1) were tested separately. The random-effects in both analyses were modelled by allowing the intercepts of each session and the intercepts of each participant's results clustered within their respective experimental sessions to vary independently.

Statistical analysis was performed within the software R (R Core Team, 2018), using tidyverse (Wickham, 2019) for data management and testing parametric assumptions, lme4 (Douglas et al., 2015) for mixed-effects modelling, and emmeans (Lenth, 2020) for *post hoc* comparisons. Models were created from the ground up, and model complexity was increased in steps, where the significance of an effect was determined using goodness-of-fit comparisons between models without and with this effect. Two goodness-of-fit comparisons were used – the chi-square log-likelihood test and the Akaike information criteria (AIC), where ΔAIC value with and without the effect of < -2 was used as the criterion to determine whether including the effects improves the overall model fit (Burnham & Anderson, 2004). Residuals of the final models met parametric assumptions (i.e., linearity, normality, and homoscedasticity) across all factor levels for all analyses. Orthogonal contrasts between experimental conditions were conducted where appropriate, and 95% confidence intervals (CIs) were derived using parametric bootstrapping (function *confint.merMod* from the lme4 package, 1000 bootstrap simulations). Post hoc comparisons between individual experimental conditions, when required, were performed using *emmeans* function from the emmeans package; the Kenward-Rodger method was used for approximating the degree of freedom of the fixed-effects, and the FDR method was used to adjust p-values. The following presents the modelling procedure for each dependant variable.

A. Using $L_{\text{Aeq,1 m}}$ values as the dependent variable

For the analysis using *VS* (Baseline, Strong, Moderate, Weak), the introduction of random-effects (clustering of participants within experimental sessions) was significant, determined by comparing an intercept-only generalised linear model (GLM) with an intercept-only generalised linear mixed-effects model (GLMM) that included the random-effects ($x^2(2) = 193.29$, $p<10^{-13}$; ΔAIC = -189.3). The fixed-effect of filter ($x^2(3) = 83.79$, $p<10^{-16}$; ΔAIC = -77.8) and the interaction of the fixed-effects of filter and *VS* were significant ($x^2(6) = 37.58$, $p<10^{-6}$; ΔAIC = -25.6). However, introducing the fixed-effect of gender ($x^2(1) = 0.18$, $p<10^{-6}$; ΔAIC = 1.8) did not improve the model fit further, i.e., gender was not a significant effect. The procedure for the analysis with the actual ST_V values instead of *VS* was similar and is not presented here. Hence, the final models predict the $L_{\text{Aeq,1 m}}$ values using the fixed-effects of filter and either ST_V or *VS* depending on the analysis, their interactions, and random-effects of participants' results clustered within their respective experimental sessions.

B. Using SB as the dependent variable

For the analysis using *VS*, the introduction of random-effects was significant ($x^2(2) = 197.2$, $p<10^{-16}$; ΔAIC = -193.2), compared to an intercept-only GLM. In terms of fixed-effects, only the interaction term between the filter and gender was significant ($x^2(7) = 16.5$, $p<0.05$; ΔAIC = -2.5), while the main fixed-effects of filter, gender, and *VS*, each, and other interactions were not significant. Hence, the final model predicts the SB values using the interaction between filter and gender as the fixed-effect. For the analysis using the actual ST_V values (either unweighted or speech-weighted), none of the fixed-effects or their interactions were significant, except for the filter and gender interaction seen above using the *VS* groups. Hence, the results using the ST_V values are not presented further.

2.3 Results

The results are presented starting with the general distribution of values over participants (Section 2.3.1), followed by examining the experimental conditions using *VS* groups (Section 2.3.2), and the effect of ST_V values overall (Section 2.3.3), both (Sections 2.3.2 and 2.3.3 are) in relation to the filter groups.

2.3.1 General results

The speech elicited was similar in general to two studies by Beechey, Buchholz, and Keidser (2018; 2019), where the task was first described and used, and included "disfluencies, interrogatives, and turn-taking behaviors," which are common in typical conversations. Further, the speech elicited in these studies generally included arbitrarily long pauses (>250 ms) between and during utterances, which was true in the current recordings as well. The typical speaking level ($L_{\text{Aeq,1 m}}$) for all experiment participants is 54.9 dB (median of means), with no significant difference between female and male participants (Figure 2.4). The speaking level range is from 3 dB (participant 6m) to 12 dB (participant 10f), with a median range of 6.8 dB (6.3 male; 7.3 dB female). There is no apparent relationship between a participant's mean speaking level and their range. Sixteen of the 23 (70%) participants' Baseline $L_{\text{Aeq,1 m}}$ was their greatest, and only one participant (19f) had a low Baseline $L_{\text{Aeq,1 m}}$. The remaining six non-maximum Baseline participants' values are greater than their respective medians. Octave-band speech levels in this experiment are similar to the "normal" speech values for female and male talkers reported by Pearsons, Bennett, and Fidell (1977), and subsequently published in a journal article by Olsen (1998) (summed from constituent one-third-octave bands to octave bands in Figure 2.5). While many studies of speech level have been conducted in anechoic conditions (Astolfi et al., 2015; Astolfi et al., 2019; Bottalico, Graetzer & Hunter, 2016; Cipriano, Astolfi & Pelegrín-García, 2017; Cheyne et al., 2009 ; Pelegrín-García & Brunskog, 2012; Pelegrín-García et al., 2011), Olsen (1998) reported both anechoic speech and speaking levels in more typical environments. Combined filtered conditions have similar spectral profiles to speech in real rooms despite HP and LP conditions emphasising different parts of the spectrum.

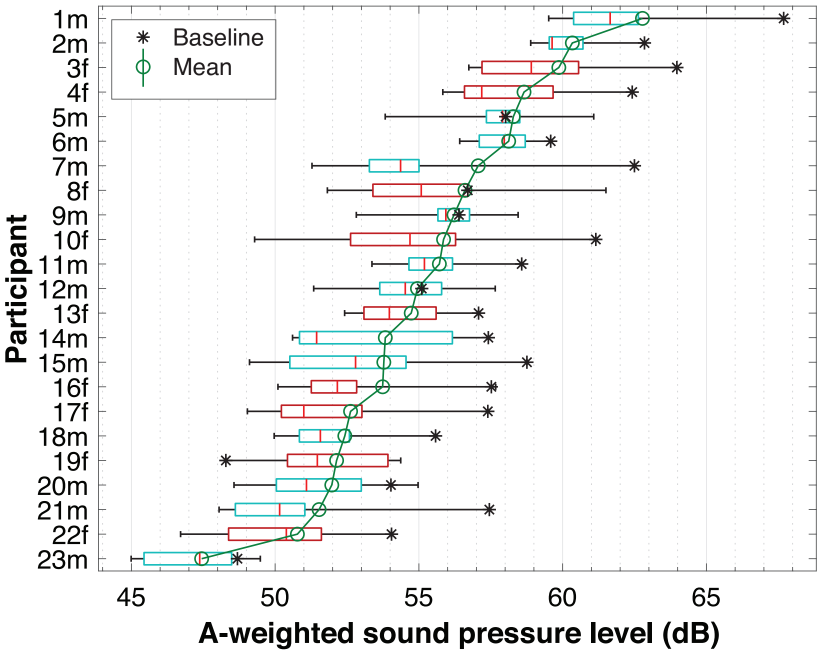


Figure 2.4. Distribution of speech $L_{\text{Aeq,1 m}}$ for each experiment participant ordered by mean value, showing the power mean (circles) and Baseline ('*') values. Female participants are labelled with an 'f'; male participants are labelled with an 'm'.

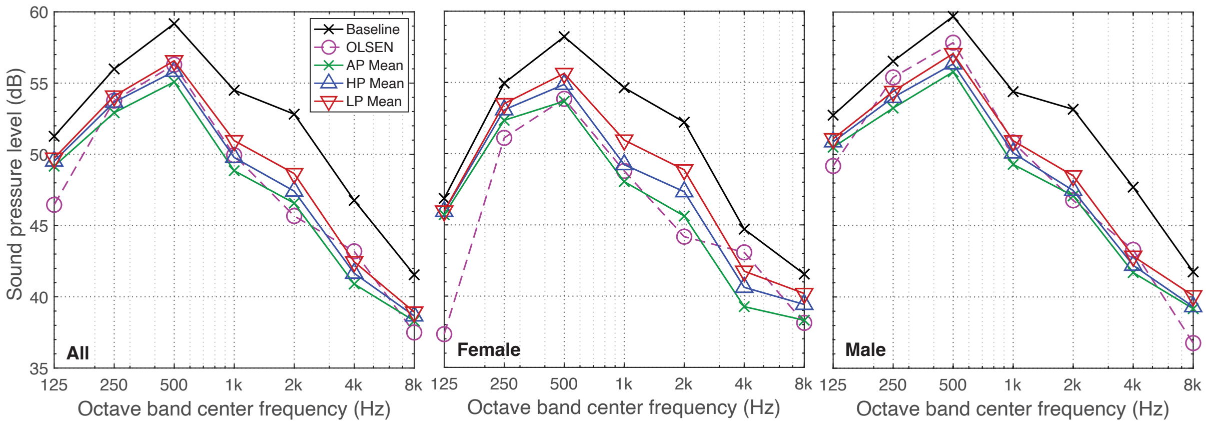


Figure 2.5. Mean speech SPL of all, female and male talkers in octave bands for Baseline and filtered experiment conditions (All-pass, High-pass, Low-pass), along with mean 'normal' speech spectra reported by Olsen (1988).

2.3.2 Results for VS groups

A. Speaking levels

When results from the participants are combined (Figure 2.6), all non-Baseline experimental conditions yield reduced speaking levels ($L_{Aeq,1\ m}$) relative to the Baseline. Strong and Moderate *VS* groups for AP and HP conditions yield similar changes in mean speaking levels of approximately 5 dB from Baseline levels (Figure 2.6 and Table 2.2). This is despite a range of 8.4 dB in the average ST_V for the Strong and Moderate groups of the AP and HP conditions (see Table 2.1). LP conditions demonstrate the most consistent trend of rising speaking levels from Strong to Moderate to Weak *VS*. The participants with the three highest and three lowest average speaking levels respond similarly to the full set of participants. The arithmetic mean values (used in statistical tests in this Section; Table 2.2) are correlated with the power mean values ($r = 0.93$, $p<0.001$), with a mean difference of 6 dB.

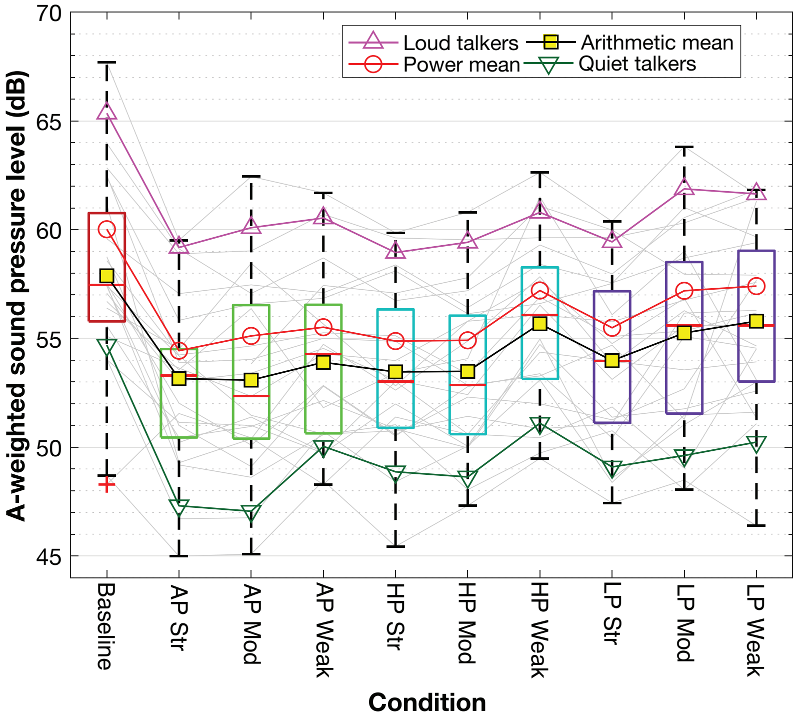


Figure 2.6. Distribution of speech $L_{\mathrm{Aeq,1\ m}}$ for each experimental condition. Gray lines without markers plot individual participant levels. Loud talker markers indicate the mean level of the three participants with the highest overall average speech $L_{\mathrm{Aeq,1\ m}}$. Quiet talker markers indicate the mean level of the three participants with the lowest overall average speech $L_{\mathrm{Aeq,1\ m}}$.

Table 2.2 shows that the Baseline has significant contrasts with all other experimental conditions (also see Figure 2.7). That is expected, given the large ST_V differences between Baseline and the conditions (Table 2.1). AP conditions yielded the largest contrasts with the Baseline condition, with estimated speaking level changes of up to 4.8 dB and the smallest range of 0.82 dB. HP conditions had significant contrasts between the *VS* groups Strong/Weak and Moderate/Weak, with Strong and Moderate conditions yielding similar results. HP and LP conditions have significant contrasts within the Moderate ST_V group, and both have significant contrasts with the AP within the Weak *VS* group. In listening tasks, 0.5 dB is considered the just noticeable level difference (JND) for broadband signals (Burnham & Anderson, 2004). However, the JND for autophonic tasks is less well-established. In this experiment, speech level change of ±0.5 dB is considered a meaningful effect size statistically significant contrasts. Based on the criterion, all significant contrasts in Table 2.2 are considered meaningful.

Table 2.2. Pairwise contrasts between experiment conditions using mixed-effects modelling. *VS* refers to the nominal voice support groups: Baseline, Strong (Str), Moderate (Mod), and Weak. AP, HP, and LP refer to All-pass, High-pass and Low-pass filters, respectively. Contrasts with p<.05 (95% CI not crossing the null value of 0) are in bold.

Category	Filter	Contrast	Estimate (dB)	95% CI
Baseline - others	AP	**Baseline - AP Str**	**4.83**	**3.07, 6.58**
		Baseline - AP Mod	**4.79**	**3.06, 6.53**
		Baseline - AP Weak	**3.98**	**2.24, 5.71**
	HP	**Baseline - HP Str**	**4.42**	**2.68, 6.15**
		Baseline - HP Mod	**4.39**	**2.66, 6.13**
		Baseline - HP Weak	**2.20**	**0.47, 3.94**
	LP	**Baseline - LP Str**	**3.90**	**2.16, 5.63**
		Baseline - LP Mod	**2.63**	**0.89, 4.36**
		Baseline - LP Weak	**2.08**	**0.35, 3.82**
Within each filter group	AP	Str - Mod	-0.03	-1.79, 1.72
		Str - Weak	-0.85	-2.61, 0.91
		Mod - Weak	-0.82	-2.55, 0.92
	HP	Str - Mod	-0.02	-1.76, 1.71
		Str - Weak	**-2.21**	**-3.95, -0.48**
		Mod - Weak	**-2.19**	**-3.92, -0.45**
	LP	Str - Mod	-1.27	-3.01, 0.46
		Str - Weak	**-1.82**	**-3.55, -0.08**
		Mod - Weak	-0.54	-2.28, 1.19
Category	*VS*	Contrast	Estimate	95% CI
Between filter and VS groups	Str	AP - HP	-0.41	-2.17, 1.35
		AP - LP	-0.93	-2.68, 0.83
		HP - LP	-0.52	-2.25, 1.22
	Mod	AP - HP	-0.40	-2.14, 1.33
		AP - LP	**-2.17**	**-3.9, -0.43**
		HP - LP	**-1.77**	**-3.5, -0.03**
	Weak	**AP - HP**	**-1.77**	**-3.51, -0.04**
		AP - LP	**-1.89**	**-3.63, -0.16**
		HP - LP	-0.12	-1.86, 1.61

B. Speech spectrum

Figure 2.7 shows average octave-band speech levels of each experimental condition relative to the Baseline condition. For the gender-combined data, the female data, and most of the male data, two clusters can be seen (also seen in Figure 2.6 based on mean levels), where the LP Moderate and Weak and HP Weak conditions are clustered away from the rest of the conditions, with the largest difference between these clusters being approximately 4.5 – 6 dB at 1 kHz. For these conditions, male talkers have somewhat larger changes in the 2 kHz band than the 1 kHz band, which affects the gender-averaged data, whereas the 1 kHz band consistently has the largest contrast from Baseline for female talkers. In general, the male talkers were more sensitive than female talkers to these conditions, yet the female talkers exhibited larger overall reductions in speaking level for the stronger *VS* conditions, especially in the 1 kHz band.

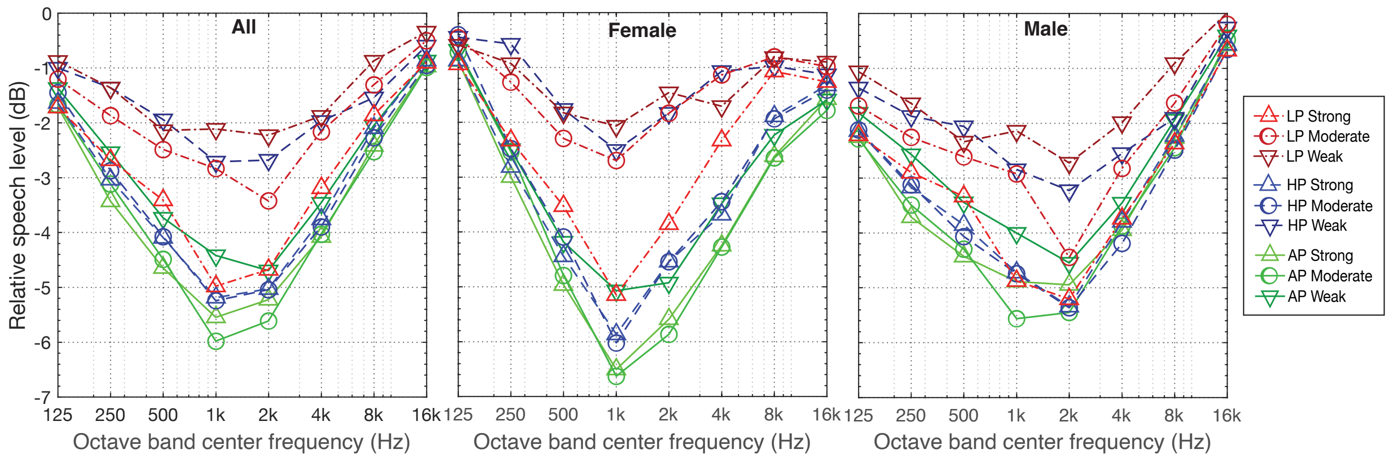


Figure 2.7. Octave-band speech levels relative to Baseline condition of all, female, and male experiment participants.

However, a mixed-effects model considering these two clusters as levels of a fixed-effect used to predict the SB (ratio of energy in the 2 – 4 kHz and 0 – 2 kHz octave bands) was not significant compared to an intercept-only model ($x^2(1) = 2.40$, p=0.12; ΔAIC = -0.41). In other words, there was no significant difference in the SB values between these two clusters. Mixed-effect models with the fixed-effects of gender and the interaction between the gender and the two clusters were also not significant with respect to the intercept-only model. Furthermore, the same results (no statistical significance) applied when speaking levels ($L_{A,eq,1\ m}$) for the 500 Hz (band with greatest on-axis speech SPL), 1 kHz and 2 kHz (often regarded as the most important band for intelligibility (Warren, Bashford & Lenz, 2005)) bands were considered individually as the response variables in respective mixed-effects models. Hence, it is unlikely that participants were speaking to enhance their speech intelligibility as is common in some studies of "clear speech" (Cooke, Mayo & Villegas, 2014; Smiljanic, 2021).

Figure 2.8 shows the SB values for filter groups and Baseline, grouped according to the gender. There is no apparent difference between the SB values for the genders for the Baseline and AP groups. However, there is a marked difference between the genders for the HP and LP groups, with a different trend for each gender in relation to the AP group. Table 2.3 shows the results of the interaction model (Section 2.2.7) between gender and filter groups to predict the SB values. Based on orthogonal contrasts, the AP group was significantly different from the LP and HP groups combined, and when the latter filter groups were considered individually, with the SB values for the female group being higher than the male group for these significant contrasts. Based on pairwise contrasts for each filter group, the female group had significantly higher SB values for only the LP group.

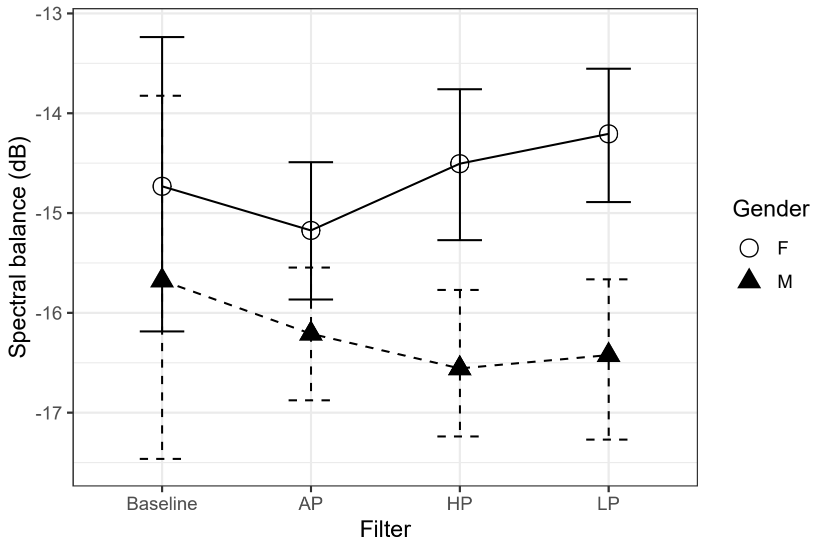


Figure 2.8. Interaction plot for gender and filter groups (AP, HP, and LP, alongwith Baseline), with mean values and bootstrapped 95% CI presented.

Table 2.3. Orthogonal (rows 1–7) and pairwise (rows 8–11) contrasts between SB values for the interaction between the gender (female and male) and filter groups (including Baseline). Statistically significant contrasts (bootstrapped 95% CI not crossing the null value of 0) are highlighted in bold. Each row shows female-male contrast.

Filter	Estimate (dB)
Others vs Baseline	0.21 [0.5,-0.09]
(LP,HP) – AP	**0.34 [0.61,0.09]**
(AP,LP) – HP	-0.13 [0.12,-0.38]
(AP,HP) – LP	-0.21 [0.06,-0.46]
LP – HP	0.08 [0.5,-0.38]
LP – AP	**0.56 [1.03,0.09]**
HP – AP	**0.48 [0.92,0.03]**
Baseline	0.79 [-1.39,2.98]
AP	0.95 [-1.08,2.98]
HP	1.90 [-0.13,3.93]
LP	**2.07 [0.04,4.10]**

2.3.3 Results for actual STV values

Equation (2.1) models participants' speaking levels ($L_{\mathrm{A,eq,1\,m}}$) as a linear function of ST_V ignoring the effect of filter group. The first two rows of Table 2.4 provide the coefficients for Equation (2.2), where ST_V values used are either averages or speech-weighted ones, respectively. Equation (2.2) models participants' speaking levels as a function of both ST_V and filter group and their interaction, after removing the Baseline ST_V, to chart the trend per filter group. RE1 and RE2 refer to the random-effect of the participants grouped within sessions and of the sessions, respectively, followed by the model residual ε, which is also a random-effect.

$$L_{A,eq,1m} = a + b \times ST_V + RE1 + RE2 + \epsilon \quad (2.1)$$

$$L_{A,eq,1m} = a + b \times (ST_V \times Filter) + RE1 + RE2 + \epsilon \quad (2.2)$$

Table 2.4. Mixed-effects regression model coefficients, along with bootstrapped 95% confidence intervals (in brackets) for the fixed-effect of ST_V overall (first two rows), and ST_V and Filter groups (the remaining rows) for the averaged (Av.) and speech-weighted (SW) ST_V values. RE1, RE2, and ε refer to the random-effects (see Equations 2.2 and 2.3) of the models. The R^2_{mm} refers to the coefficient of determination for mixed-effects models (*mm* in the subscript), calculated using the formula in Nakagawa, Johnson, and Schielzeth (2017). Statistically significant intercepts and slopes are highlighted in bold.

ST_v	Filter	Intercept (*a*)	Slope (*b*)	RE1	RE2	ϵ	R^2_{mm}
Av.	-	**51.65 [49.50,53.41]**	**-0.24 [-0.28,-0.20]**	2.39 [1.30,3.28]	2.65 [0.57,4.35]	1.83 [1.64,2.02]	0.81
SW	-	**52.46 [50.63,54.28]**	**-0.22 [-0.26,-0.19]**	2.39 [1.34,3.42]	2.65 [3.02,4.19]	1.81 [1.63,1.98]	0.82
Av.	AP	**52.62 [50.66,54.5]**	-0.11 [-0.24,-3.5·10^{-3}]	2.41 [1.25,3.43]	2.64 [1.4·10^{-4},4.38]	1.64 [1.45,1.82]	0.84
Av.	HP	**48.28 [45.77,50.59]**	**-0.53 [-0.78,-0.29]**	2.41 [1.25,3.43]	2.64 [1.4·10^{-4},4.38]	1.64 [1.45,1.82]	0.84
Av.	LP	49.97 [47.14,52.33]	**-0.34 [-0.57,-0.11]**	2.41 [1.25,3.43]	2.64 [1.4·10^{-4},4.38]	1.64 [1.45,1.82]	0.84
SW	AP	**52.72 [50.91,54.50]**	-0.10 [0.21,3.6·10^{-3}]	2.40 [1.37,3.47]	2.63 [1.12,5.29]	1.64 [1.47,1.81]	0.84
SW	HP	**52.15 [51.06,53.25]**	**-0.27 [-0.43,-0.12]**	2.40 [1.37,3.47]	2.63 [1.12,5.29]	1.64 [1.47,1.81]	0.84
SW	LP	53.03 [52.35,54.36]	-0.20 [-0.35, 0.04]	2.40 [1.37,3.47]	2.63 [1.12,5.29]	1.64 [1.47,1.81]	0.84

Table 2.4 shows steepest reductions in speaking levels for the HP group (consistent with Figure 2.6 and Table 2.2) followed by the LP group (not significant for speech-weighted ST_V values) and finally the AP group (not significant). While Table 2.4 provides model coefficients including the slopes for individual filter groups, Table 2.5 compares the slopes of relevant filter groupings for averages and speech-weighted ST_V values. The main finding is that the HP slope is significantly different form the AP slope (Table 2.5), meaning steeper reduction in participants' speech level with HP than with AP filter conditions, more so for the averaged than speech-weighted ST_V values. Slopes of none of the other individual filter conditions differed significantly from each other. For averaged ST_V values, slope of LP and HP groups combined significantly different (steeper reduction) compared to the AP group, and slope of AP and LP groups combined was significantly different

from the HP group (steeper reduction); these contrasts were not significant for speech-weighted ST_V values.

Table 2.5. Orthogonal contrasts between ST_V slopes (averaged (Av.) and speech-weighted (SW)) of various filter groups. Statistically significant contrasts (bootstrapped 95% CIs not crossing the null value of 0) are highlighted in bold.

ST_V	(LP,HP) – AP	(AP,LP) – HP	(AP,HP) – LP	HP – AP	LP – HP	LP – AP
Av.	**-0.11**	**0.10**	0.01	**-0.21**	0.09	-0.12
SW	-0.04	0.04	0.01	**-0.09**	0.04	-0.05

2.4 Discussion

2.4.1 Voice level regulation: The room effect

Similar to previous studies in classrooms (Astolfi et al., 2015; Bottalico, Astolfi & Hunter, 2017), laboratories (Astolfi et al., 2015, Astolfi et al., 2019; Bottalico, Graetzer & Hunter, 2016), and simulated sound fields (Cipriano, Astolfi & Pelegrín-García, 2017; Pelegrín-García & Brunskog, 2012; Pelegrín-García et al., 2011), this experiment shows that increased indirect autophonic gain reduces speaking level. However, in this experiment, a realistic communication scenario was used, and overall, participants reduced their speaking levels linearly with increasing ST_V, for both averaged and speech-weighted values. Speech-weighting (Pelegrín-García et al., 2012) emphasises the mid-frequency bands (Table 2.1), and the greatest changes in speaking level were in the 1 kHz band (Figure 2.7). Speech-weighting also greatly de-emphasises the 125 Hz band, which has little relevance for female speech. Hence, speech-weighting is recommended for future studies for a single number quantity representing ST_V, even though it provided almost the same fit as simple averaging of band values (R^2mm of 0.817 compared to 0.814, Table 2.5; same bootstrapped 95% CI of [0.762,0.833]).

Changes in voice level are not only influenced by the indirect autophonic gain, but by other aspects of the communication scenario, such as inter-talker distance and the instructions or tasks given to the talkers (Lane & Tranel, 1971; Pelegrín-García & Brunskog, 2012; Pelegrín-García et al., 2011; Pelegrín-García et al., 2011b). The current study found an overall -0.22 dB/dB voice regulation slope with a scenario of two participants, 4 m apart, concurrently doing a conversation task where communication was of premium importance. In comparison, the most relevant study by Pelegrín-García and Brunskog (2012), which used a somewhat comparable range of ST_V values (a larger range in the current study), wherein the magnitude range of the room effect depended greatly on the instructions given to participants. An average voice regulation slope of -0.93 dB/dB was found for free speech within the pilot experiments, where talkers addressed an imaginary group of students in a classroom scenario. For the experiment closest in design to the current (group B in Pelegrín-García and Brunskog (2012)), a notably flatter slope of -0.1 dB/dB was produced when participants described a route on a map to the (quiet) experimenter, seated at 3 m distance. A higher slope than the current study of -0.35 dB/dB was found when participants addressed a dummy head (visual cue) at 2 m distance, while another experiment's slope (group A in in Pelegrín-García and Brunskog (2012)) was non-significant. In summary, the current results provide a voice regulation slope likely to be more robust in characterising real-world communication between two people, where the slope reported is higher than those in a previous study that had tasks with a lower-premium on communication.

A caveat in interpretating this slope is that the current study, by design, had some experimental conditions with a clear low- or high-frequency bias in the ST_V. Table 2.3 shows that the slope of the AP group is significantly different from the band-limited LP and HP groups combined for the averaged ST_V values, but not for the speech-weighted ST_V. Hence, combined with the reasoning in the previous paragraph, the speech-weighted slope of -0.22 dB/dB of ST_V is recommended as more characteristic of room effect, or voice regulation due to room acoustics.

2.4.2 Spectral changes in voice regulation

A key question in the present experiment is whether the spectrum of indirect autophonic gain, in the form of reverberant ST_V, affects voice regulation. Broadly speaking, the filter and *VS* groups often yielded significantly different voice level regulation effects which is evident in Figures 2.6 and 2.7 and Table 2.2. More specifically, band-limiting the ST_V, i.e., HP and LP filtering compared to AP, tends to increase voice level regulation (Tables 2.4 and 2.5), and HP filters in particular lead to steeper increases in speech levels with decreasing ST_V values. This indicates that the spectrum of ST_V has some effect on A-weighted speech level – more for octave-band averages than speech-weighted ST_V. The naming of the *VS* groups (Strong, Moderate, Weak) is not intended to be representative of real rooms in which ST_V typically falls within -20 to -5 dB (Pelegrín-García et al., 2011a). Hence the "Moderate" condition has the ST_V of strongly supportive real rooms, while the "Strong" condition is better described as "very strong," probably mostly experienced in small rooms, rooms with amplification, or rooms with directed reflectors or focusing. The contrasts between the Strong and Moderate conditions were not significant for all the filter groups, while there was a significant contrast between Moderate and Weak conditions for the HP group and a similar trend of decreasing levels for the LP group. This suggests that there may be a limit to voice level reduction, a saturation effect; results suggest that this is at the ST_V limit of real rooms for the communication scenario investigated.

The spectral profile of speech produced in response to the three filtered indirect autophonic gain conditions yielded voice spectra that approximately followed the spectrum profile of "normal" speech in real rooms as reported by Pearsons, Bennett, and Fidell (1977) (Figure 2.5), with the Baseline condition in the current experiment having higher in octave-band levels than this previous study. Compared to the Baseline condition, there are several differences in the speech spectrum in response to the filtered conditions, with two clusters of conditions clearly visible in Figure 2.7. While the differences in the SB or level of specific octave bands (500 Hz – 2 kHz) between these two clusters were not significant, more data are needed to determine whether LP and HP filtering of ST_V varies spectral profile systematically compared to mostly unfiltered (i.e., AP) conditions and for ST_V that is typical in rooms (i.e., Moderate and Weak values; see Table 2.1).

Hence, while this work provides a positive indication that band-limited acoustically supportive treatments in rooms may act differently from broadband treatments of the same ST_V, this is a hypothesis for future study with a larger sample of conditions and participants. From a practical standpoint, the important finding is that the band-limited acoustic support yielded voice regulation effects similar to those of broadband support. This is useful because certain acoustic surface treatments are band-limited – particularly treatments based on discrete reflectors for which the reflected spectrum is limited to high frequencies due to diffraction loss (e.g., retroreflective array surfaces (Cabrera et al., 2020)). The results of the present experiment provide some support for the use of such treatments for noise control, subject to further studies using greater realism, as the HP nature of such surfaces may provide higher reduction in voice levels than would be expected with broadband treatment. In general, the contribution of this experiment may be to encourage the development of practical acoustic treatments for ST_V control using techniques that are not necessarily broadband. As observed by Titze and Hunter (2015), ST_V treatment may be effective in reducing vocal loading, effort, and fatigue, which is particularly important for professional talkers, who have elevated risk of vocal injury. If ST_V treatments in architectural acoustics are to be effective in reducing the overall speaking level in multi-talker environments, then they must be effective for loud talkers, who make the greatest contribution to the overall sound field. While the results do show variation between individual talkers in their voice regulation effect, there is no systematic difference between the voice regulation effect for loud versus quiet talkers.

2.5.3 Limitations

A limitation in the experiment design was that experimental conditions for each participant were randomly ordered, meaning that the talkers were usually not experiencing the same indirect autophonic gain. This may have affected the contrast in the results compared to a situation where the autophonic conditions had been matched, as each pair of participants might have influenced each other's speaking level. This is somewhat considered in the statistical modeling with the inclusion of random-effects of participants within sessions and of each session. A further consideration here is that the experiment did not provide any acoustic support of the other person's speech – so the direct voice communication between the participants was anechoic. These are important differences between the experimental and potential real architectural environments in which ST_V could be strongly biased to low or high frequencies.

Alongside the objective results, a subjective assessment of one's own voice is likely to be useful, which is recommended for future research.

2.6 Summary and conclusions

In this experiment, pairs of participants performed a task that involved communicating instructions by talking to each other. The indirect autophonic gain of their voices was provided as electroacoustic ST_V, varying across ten conditions in level and spectrum. The main conclusions of this experiment are as follows:

i) Overall, talkers lowered their speech level by 0.22 dB for every additional dB of speech-weighted ST_V. This slope is different from previous studies that had lower premium on communication and is more likely to be representative of real communication at short distances between talkers (4 m in the experiment) in relatively quiet background noise conditions. More studies are needed to investigate the effect of distance, higher background noise, different communication tasks, and the subjective impressions of the talkers.

ii) HP ST_V conditions had the steepest reductions in speaking levels and were significantly different from slopes for the AP conditions. This may have relevance for designing acoustical treatment for spaces used for speech communication, since from a noise control perspective, rooms generally have high sound absorption for high-frequency octave bands. Future studies could investigate whether ST_V with a high-frequency bias enables talkers to instinctively lower their voices during communication in general over longer periods and within various communication scenarios while maintaining sufficient sound absorption.

iii) Compared to male talkers, female talkers produce more high-frequency voice energy in their speech in band-limited ST_V conditions (HP and LP conditions in this experiment) than in conditions with ST_V spectrum that is more typical of everyday rooms (AP conditions in this experiment) and more high-frequency energy than male talkers in LP filtered ST_V conditions than HP conditions. Combined with the previous point, if speech intelligibility is of concern, it seems that acoustic treatment with low-frequency ST_V boost may need to consider the effect on female voices more, while treatment with high-frequency ST_V boost may need to consider the effect on male voices more.

Chapter 3: Singing experiment - theoretical basis, design, and methods

The singing experiment outlined in this chapter and the following two chapters is a novel framework for singing practice that was imposed as an experimental intervention on established practice methods over fifteen months. The practice framework was designed to achieve the following objectives:

i) To discover and development new singing techniques and musical materials for use in ensemble improvised music.

ii) To change the quality of my singing voice.

The experiment design was informed by prior research and theoretical frameworks from multiple disciplines including creative arts practice-led research, acoustics science research, and ecological dynamics frameworks for learning.

Aspects of the singing experiment are divided into three chapters, the content of which can be summarized as the following.

i) Chapter 3: the research objectives, theoretical basis, experiment design, and methods used in the singing experiment.

ii) Chapter 4: measurements and analysis of the acoustics of the experiment atypical practice settings, measurements and analysis of the singing voice in each practice setting and throughout the entire experiment.

iii) Chapter 5: descriptions and demonstrations of the creative outcomes of this experiment, presented as a taxonomy of discovered and developed materials, and recordings of their application in ensemble improvised music.

3.1 Research objectives

The central goal of this experiment is the development of an original improvisational style of singing, comprised of novel technical singing skills and musical materials, for application in ensemble improvised music contexts. In this thesis I use the label "improvised music" as it is used by Mayas (2019); referring to an approach to performing and composing music in real-time that emerged in the 1960s with influences including free jazz, electronic music, noise, and new music.

This experiment targets skills that are necessary to achieve the spontaneous generation of new musical materials during improvised performance. Targeted materials are those that meet the criterion: 'original', 'efficient', and/or 'transferable' (a criterion introduced by Slater, 2020). Original materials are those that meet my artistic aesthetic. These are materials that I like and want to play and that meet my performance goals and intentions. Efficient materials enable coordination of the physical skills required to sing including physical patterns of coordination and control. Finally, transferable materials are those which are adaptable and flexible. Improvised performance contexts are inherently variable and transferable materials can be utilized in response to multiple situations and aren't rigid in their application.

A secondary goal of the experiment is to change the quality of my singing voice. This experiment uses spectral centroid to describe the quality of the singing voice: the distribution of power as a function of frequency. The analysis of long-term changes in the spectral centroid of singing is included in Chapter 4 (Section 4.3.3C).

3.2 Singing background – experiment starting point

My singing background is an important foundational aspect of the experiment in three ways: i) it informs the experiment design; ii) it is the starting point of my singing technique; and iii) it is explicitly related to some music material outcomes described in Chapter 5 (Section 5.1.2D and Section 5.1.2G, see description below). The vocal techniques and musical constructs taught to me throughout my primary vocal training provide an important theoretical and technical basis for the singing experiment. My vocal training has been with Bae Il Dong, a distinctively innovative *p'ansori* artist (see paragraph below for description of *p'ansori*). The vocal techniques I have learned from Bae are the foundation of my singing technical ability and are thus a starting point for experimental developments. The singing experiment was also influenced by traditional *p'ansori* training procedures, mainly singing at waterfalls (for more details see paragraph below). The descriptions of Bae Il Dong's vocal practice included in this thesis are taken from transcriptions of lessons taught to me in English over eight years. He has approved their inclusion in this thesis.

P'ansori is a story-singing tradition from pre-modern Korea that is characterised by a tonal aesthetic preference for a highly resonant and hard-pressed or attacked voice (Park, 2003). Traditionally, to achieve this vocal aesthetic a *p'ansori* student undertakes a *dok-gong* practice; years of solitary training by a waterfall deep in the mountains (Bae, 2016). Bae Il Dong spent seven years in isolation, training his voice by singing at waterfalls at great volume with the goal of cultivating his unique personal tone (Bae, 2016). He has published two seminal books on his artistic practice (Bae, 2016; 2020) and has collaborated with Australian improvising artists in bands Daorum, Chiri, and the Australian Art Orchestra. His collaboration with Australian drummer Simon Barker and his *dok-gong* training is documented by the documentary 'Intangible Asset No. 82' (Franz, 2008).

My initial motivation for beginning lessons with Bae Il Dong was hearing him perform at the Sydney Conservatorium of Music in 2014. The quality of his voice was unlike anything I had ever heard, and I travelled to Seoul that year to do a month of full-time technical singing training.

Over eight years of lessons in both Australia and Seoul, Bae has taught me fundamental singing techniques and methods for instigating changes in my voice. In these lessons I learned *p'ansori* songs as a means of both acquiring new vocal technique and understanding his theoretical constructs for music expression. However, Bae's techniques and theoretical constructs are not *p'ansori* specific, thus he teaches musicians of all kinds, in addition to traditional Korean musicians.

The techniques taught to me by Bae Il Dong that are most explicitly used in the singing experiment are what I refer to as *resonate* and *grain*. Resonate is the process of directing the voice into the air-filled cavities in the head for enhanced intensity. How this technique and subsequent developments of it are applied in the research outcomes is outlined in Section 5.1.2G. Grain refers to a technique of distorting the tone of the voice and is detailed in the context of research outcomes in Section 5.1.2D. Bae does not call this technique grain, instead he refers to this tonal quality as an 'Earth' tone, in contrast to a clear 'sky' tone.

3.3 Artistic aesthetic

My personal artistic aesthetic shapes key aspects of this experiment. A strong attraction towards specific musical characteristics plays a formative and ongoing basis in the process of practical self-organization and in identifying valuable discovered and developed musical material. Musical characteristics that I am attracted to, that meet my subjective artistic aesthetic, are as follows:

i) The foregrounding of tone colours that are a product of idiosyncratic techniques.

ii) Musical indeterminacy within which nothing is resolved and anything can happen.

ii) Obfuscation of musical pulse.

iv) General avoidance of unified musical activity.

These musical preferences inform ongoing decisions as to what discovered materials are developed or abandoned.

3.4 Singing experiment theoretical framework

This section outlines the theoretical foundation for the singing experiment. All the theoretical frameworks have previously been applied in music research to produce creative research outcomes (Barker & Howard, 2015; Gander, 2017; Hale, 2018; McLean, 2018; McMahon, 2022; Slater, 2020). The primary source for the theoretical approach used in this singing experiment is Slater's (2020) research which applies the constraints-led approach from ecological dynamics to the acquisition and development of jazz trumpet skills. Slater provides a comprehensive and in-depth review of the aspects of this approach and how they can be applied to music practice. The facets of this approach that I applied in the singing experiment are outlined in this section. My experiment uses acoustic science research as a novel extension of Slater's research.

3.4.1 Practice-led research

This experiment can be described as practice-led research, a well established research framework used in the creative-arts within which practitioners use practice procedures to produce creative research outcomes. Both the creative outcomes of practice and the experiment design imposed on practice are outcomes in this kind of research (Candy & Edmonds, 2018; Choi, 2016; Smith & Dean, 2009).

3.4.2 Self-regulated learning and practitioner research

In this experiment, practice-led research is undertaken by means of self-regulated learning, which is described here as the process used by musicians in their solitary practice routines. Within this learning framework, the individual is responsible for goal identification, task specification, practice scheduling, task duration, and performance evaluation (McPherson et al., 2013; McPherson & Renwick, 2011; McPherson & Zimmerman, 2002; Slater, 2020; Zimmerman et al., 1996).

My self-regulated practice can be described as an objective practice mode with explicit goals and well-defined tasks designed to maintain skills and overcome specific weaknesses. Examples of methods used are singing long tones, repetition of materials with variations in tempo, articulation, and rhythm, and material sequence, retrograde, and fragmentation. My self-regulated learning processes have been developed over twenty years of music practice and performance.

This research can be described as practitioner research because I am a professional musician. I consider myself to be an effective self-regulated learner and the singing experiment is an inquiry-based process of continuing professional development. The practice design implemented in this project is an experimental intervention upon my well-established practice methods to expand opportunities for skill acquisition and creative development. This thesis describes the aspects of the singing experiment that are imposed on self-regulated practice: the theoretical basis, experiment design, methods, and research outcomes.

3.4.3 Ecological dynamics

This project applies the constraints-led approach of ecological dynamics to singing practice. Ecological dynamics is a theoretical framework developed in the sports and movement sciences to describe the ways that skill learning, acquisition, and development are regulated by information sources found within the performance environment (Davids, 2010; Davids, Button. & Bennett, 2008). Within the framework, learners are conceptualized as complex adaptive systems who are capable of self-organization under the interactions of one's body, the learning task, and environment within which the learning occurs (Araújo & Davids, 2009; Chow et al., 2016; Davids, 2010; Orth, Kamp & Button, 2019; Renshaw et al., 2019; Uehara et al., 2014). The adaptation a performer makes to the interactions of internal and external information is an important learning tool, effective over multiple timescales: from moment-to-moment instantaneous adjustments to performance decisions and actions, through to developmental change over days, months, and years (Clark, 1995). This learning framework has been empirically tested in both experimental settings and real-world coaching and is supported by athletes and coaches as being a valuable educational tool for understanding and developing movement expertise (Renshaw et al., 2019).

Within the ecological dynamics framework, skill acquisition and development involves the education of a performer's intention, attention, and calibration (Slater, 2020). Intention is the thought processes that guide a performer's actions; these processes include a performer's motivations, directives, and goals. Attention refers to the search and discovery of available information within the performance environment, such as auditory and visual information, which may be of use to the performer. Calibration is the adjustments made to performance actions to enhance attention and the exploitation of environmental information in performance. Effective calibration leads to the discovery of stable yet adaptable movement patterns in dynamic performance contexts.

3.4.4 Constraints-led approach

The constraints-led approach is a learning methodology, supported by ecological dynamics, within which skills are developed through manipulation of performer, environment, and task constraints to guide the discovery of new patterns of coordination (Chow et al., 2009; Davids, Button & Bennett, 2008; Gibson, 1979). Task constraints relate to the intent of the activity; performer constraints pertain to the individual, including their physiological qualities; and environmental constraints include features external to the performer, for example the weather (Teune et al., 2021). In the constraints-led approach, constraints are viewed as boundaries that shape emergent behaviour of individuals and groups (Newell, Liu & Mayer-Kress, 2001). A coach, teacher, or physical therapist manipulates environment and task constraints, as opposed to giving prescriptive instruction, so a student discovers functional movement solutions in the moment. Skilful behaviours are described in terms of the student being able to respond to constraints in increasingly effective ways (Seifert, Button, & Davids, 2013).

3.4.5 Metastabilty and meta-stable performance regions

Metastability is a movement system property that helps performers to remain in a state of relative coordination with their performance environments (Ross, Davids & Renshaw, 2012). In many contexts, environments are information-rich, dynamic, unpredictable, and uncertain. Metastability supports transitions between different states of movement organization, functionally adapting stable movement patterns to remain in relative coordination with a dynamically changing performance environment (Kauffman, 1995). Metastable attuned performers are ready to engage with a multiplicity of affordances, or opportunities for action, allowing for a balance between stable and flexible configurations of movement, adapting their behaviour to what the particular situation requires (Bruineberg, 2021). Designing research and practice tasks with a focus on candidate metastability results in rich, flexible, and varied patterns of movement organization and performance responses (Ross, Davids & Renshaw, 2012), allowing for acquisition of skills and development of metastable attunement (Bruineberg, 2021).

The meta-stable performance region is a physical state of being within which the meta-stable tendency persists. An information-rich performance environment that is capable of a variety of actions forces performers into a meta-stable performance region where rich and varied patterns of functional movement behaviour emerge (Guerin & Kunkle, 2004; Juarrero, 1999; Pinder, Davids & Renshaw, 2012). In the meta-stable performance region, the performer develops the ability to attune to information sources to guide performance actions. The performer is able to choose/move between multiple equally effective performance actions to meet task goals. Ecological dynamics supports a constraints-led approach to creating these useful meta-stable performance regions by manipulating the available resources used by the learner to self-organise responses (Slater, 2020).

3.4.6 An example of the constraints-led approach to music skill acquisition

These theories of learning and their application to music practice-led research are investigated and demonstrated by Slater (2020). Slater advocates the constraints-led approach as an effective and efficient way to encourage emergent performance activities. In this research, Slater performed solo improvisations in 179 practice sessions over twelve months in atypical settings for creative development of musical skills and materials that were subsequently applied to compositions for a jazz quintet. Slater used the environmental constraint of performing in a range of outdoor settings within a 20-minute drive from his house and the task constraint of

20-minutes of recorded improvised performance. The outcome of this research was a new trumpet vocabulary and a jazz quintet recording, 'The Dark Pattern', which Slater describes as representing an important change in his technical and aesthetic preferences and abilities.

3.5 Experiment design and methods

This singing experiment implements environmental and task constraints upon my singing practice to encourage discovery and development of improvisational skills and materials. The specific skills being targeted are those that enable the spontaneous improvisation of original, efficient, and transferable music materials (Section 3.1). This is achieved through attuning my performance intentions, attentions, and calibrations (Section 3.4.3) through the implementation of constraints. The experiment constraints and procedures are outlined in this section. This experiment design is considered as a research outcome.

3.5.1 Environmental constraints

In order to initiate a constraints-effect upon my practice, my experiment design includes atypical practice settings. The chosen practice settings provide rich and variable environmental information to guide my performance attentions and calibrations. Two atypical practice settings were selected: a high flowing waterfall and a polytrihedral dome. These settings were chosen for two reasons: being outdoors (see Section 3.5.1D) and due to their extreme and opposing acoustics (see Section 3.5.1A). These traits introduced perturbations that required me to find new performance solutions. The extreme acoustics of practice settings in this experiment, and the resulting changes to voice production, are considered as a novel extension of the constraints-led approach to music skill acquisition introduced by Slater (2020).

A. Autophony and voice regulation in extreme acoustics

Voice regulation refers to the automatic adjustments a person makes to their speaking or singing voice as a product of the sound of one's own voice (autophonic sound) as altered by the acoustic environment. Autophony and voice regulation are described in detail in Section 1.1. At the onset of the project, it was hypothesized that extreme acoustic environments would result in extreme voice regulation. Often extreme acoustic environments constrain a performer in perceived negative ways. Musicians have reported that whilst they preferred 'live' practice spaces to 'dead' ones, the majority felt that a bathroom (a very live room) would be impossible to practice in (Lamberty, 1980).

By contrast, this experiment prioritizes the extreme regulation of my singing as a result of practicing in extreme acoustics. Voice regulation includes adjustments in both the spectrum and the sound pressure levels of singing (Section 1.1). These changes can educate a singer as to what influence these adjustments have upon sound production and performance choices. Within the ecological dynamics framework, the combination of unintentional and intentional adaptations to these environmental constraints can lead to powerful learning mechanisms, instigating change and discovery of vocal techniques. An analysis of the acoustics of the practice settings and the measured voice regulation of singing at the practice settings is included in Chapter 4.

B. A waterfall as a practice setting

The starting point for choices of practice settings and an example of music practice in extreme acoustics is the traditional practice setting for Korean *p'ansori* singing: a waterfall. High-flowing waterfalls provide constant broadband sound (Galbrun, 2013). A high-flowing waterfall disrupts the direct and indirect autohponic sound through masking. The Lombard Effect (Section 1.1.2) is where talkers and singers instinctively increase their vocal output to compensate for the noise in their environment which disrupts the direct and indirect autophonic sound. The Lombard effect may play a part in the usefulness of waterfalls in practicing singing with high sound pressure levels like *p'ansori*. I practiced at high-flowing waterfalls for the extreme acoustic environment of voice masking and the subsequent voice regulation effects (Figure 3.1).

Figure 3.1. Photos of waterfall practice settings. Left: A photo of Belmore falls 3, Morton National Park, NSW. Right top: A photo of upper Belmore falls, Morton National Park, NSW. Right bottom: A photo of me singing besides the Crags falls located on private property in Mittagong, NSW.

C. A polytrihedral dome as a practice setting

An opposing and alternative extreme acoustic environment to a waterfall is that which provides a high amount of autophonic sound through reflections. Consequently, I designed and built a polytrihedral dome (1m radius, see Figure 3.2) with an array of 40 trihedral retroreflectors (Section 1.2). The polytrihedral dome created extreme levels of autophonic support (voice support, Section 1.1.5) and is a unique acoustic environment in that it reflects sound back to the source. The design and acoustics of the polytrihedral dome are detailed in Chapter 4. Autophonic support from room reflections has been shown to influence talkers to reduce their vocal output (see side tone compensation, Section 1.1.3, and the room effect, Section 1.1.8). The polytrihedal dome provided levels of autophonic sound well beyond that of everyday rooms, particularly at high frequencies. The talking experiment in the previous chapter demonstrated that autophonic support with a high-passed spectrum (frequencies above 1 kHz) reduces voice levels and alters the spectral balance of speech. It was hypothesized at the onset of the project that whilst practicing singing in the polytrihedral dome, the levels and spectral balance of singing would change in systematic and unconscious ways. This voice regulation effect is considered to be a perturbation on performance activities, and thus, a learning tool in vocal skill acquisition.

Figure 3.2. Left: A photo of the polytrihedral dome in the Burragorang Valley, New South Wales. Right top: A photo of me singing in the polytrihedral dome. Right bottom: A photo of the polytrihedral from a distance.

D. Outdoor practice settings

This project uses outdoor locations as practice settings to initiate skill acquisition and creative development. The outdoor practice settings were chosen as they were richly varied, providing a wide variety of visual, auditory, and haptic information to guide my attention and, thus, enable calibration of performance activities and encouraging the discovery of novel performance solutions. Environmental information that both inspired and distracted practice activities included bird song, insects flying and biting, wind moving trees, weather changing, goannas running up trees, frogs croaking, temperature changes, and relentless cicada sounds. The polytrihedral dome is a unique outdoor setting as it was designed to maintain the assets of the outdoor practice environment, whilst providing extreme levels of voice support not usually available in the outdoors.

3.5.2 Timeline

225 in-situ practice sessions were completed over a 15-month timeframe. 12-months was initially planned in line with Slater's (2020) constraints-led practice experiment however, the timeframe was extended due to Covid-19 related interruptions. Practice sessions ranged from 1 to 3 hours, with a total amount of 354 hours and an approximate average of 24 hours a month (range of 11 hours in July 2021 to 54.5 hours in November 2021). A full timetable of practice sessions is attached as Appendix A. The regularity and length of each session was determined by established practice methods (practitioner research, Section 3.3.2) and by limitations including Covid-19 restrictions, weather, and access to practice sites.

3.5.3 Use of and access to practice settings

This section details the practice settings: the locations of, access to, and the amount of practice done at both settings.

A. Polytrihedral dome

A total of 183.5 hours of practice over 67 sessions were completed in the polytrihedral dome. The polytrihedral dome was built on private property in the remote Burragorang valley in New South Wales, Australia. It is the only human-made structure in the immediate vicinity and is accessed via 10 km trail (4-wheel drive access only), down a steep escarpment and then a further 20-minute walk. The owners of the property granted and facilitated ongoing access. Video 3.1 shows the polytrihedral dome in the Burragorang Valley.

Video 3.1. A video of the polytrihedral dome installed on private property in the Burragorang Valley.

B. Waterfall

A total of 170.5 hours of practice over 51 sessions were completed at a waterfall. A range of waterfalls were used for practice that were chosen for access ease and low visitor traffic. They were alternated depending on weather and water levels as high-flowing waterfalls were prioritised. The waterfalls are located in national parks or on private property where owners granted ongoing access. Regularly visited waterfalls include:

i) Nattai creek falls, The Crags (private property), Mittagong, NSW

ii) Dog trap creek falls, Burragorang valley (private property), High Range, NSW

iii) Belmore falls, Morton National Park, NSW

iv) Curracurrang gully falls, Royal National Park, NSW

v) Edendarry falls, Blue Mountains National Park, NSW

Video 3.2 provides an example of waterfall practice environment with the environmental sound as recorded by in-ear microphones, the DPA 4560 Binaural Headset Microphone (DPA Microphones, Alleroed, Denmark) (see Section 3.3.5 for a list of field recording equipment). Trails to waterfalls required a range of hiking times, from 20-minutes to 3-hours, some of these trails being accessible only by 4-wheel drive roads.

Video 3.2. A video of upper Belmore falls (for scale: approximately 6 m high) with the environmental sound as recorded by in-ear microphones.

3.5.4 Remote Fieldwork

This experiment required remote fieldwork that included hiking, camping, and driving on roads that require 4-wheel drive vehicles. It was carried out under University of Sydney Occupational Health and Safety guidelines. I have extensive experience in hiking and navigating and, prior to the experiment, I received advice from a professional outdoor safety guide and 4WD training. A vehicle with dual range four-wheel drive capability, repair items, and recovery equipment was used. A first-aid kit, maps, compass, and a personal location beacon were always carried.

3.5.5 Data collection

Audio and visual recording was used for primary data collection in practice sessions. Equipment for data collection was chosen to meet the requirements for environment and voice sound analysis (Chapter 4) and assist in ongoing development of musical materials. The equipment also had to facilitate remote fieldwork procedures, mainly being compact and lightweight enough for hiking. Figure 3.3 shows all data collection equipment.

Figure 3.3. A photo of the equipment used for data collection.

For sound analysis, practice sessions were recorded with the following equipment:

i) DPA 4560 Binaural Headset Microphone (DPA Microphones, Alleroed, Denmark).

ii) F6 Zoom field recorder (Zoom Corporation, Tokyo, Japan).

ii) Jez Riley French C-series contact microphone (United Kingdom).

The procedure and results of these recordings are discussed in Chapter 4.

For musical material analysis, practice sessions were recorded with a tripod mounted Zoom Q8 handy video recorder (Zoom Corporation, Tokyo, Japan). It was not calibrated, and the gain and position were changed each session to ensure the singing was recorded clearly despite variable background noise. At the polytrihedral dome the recorder was placed outside the perimeter of the dome to record singing and limit the recording of reflections from the retroreflective array. The recorded audio material was later transferred from the Q8 Zoom soundcard to a computer and checked for technical errors. Recordings ranged from 10 to 30 minutes at the end of a practice session and intermittently to document recognised moments of discovery and change. Some practice sessions were not recorded due to weather or restricted access to battery charging. Recordings were reviewed after every practice session as a part of the singing practice procedure, which is outlined in Section 3.5.6D.

3.5.6 Practice procedures

The experimental practice procedures are described below. The described procedures are designed as an experimental intervention on my pre-project, established self-regulated practice methods.

A. Practice journal

In the singing experiment, journaling was used as a tool to track the progress of the singing practice. Practice journaling has been demonstrated to be an effective way to enhance learning activities within self-regulated processes and is a regular feature of practice-led research projects (Barrett & Bolt, 2007; McPherson & Renwick, 2011; McPherson & Zimmerman, 2002; Pedgley, 2007; Slater, 2020; Zimmerman, 2000). Throughout the experiment, practice notes were written in the same journal at every session so that previous notes could be reviewed whilst practicing. The notes throughout the project primarily included the following:

i) What, when and where of what I was practicing.

ii) Environment conditions and sounds.

iii) Technical and organisational discoveries.

iv) Questions to myself about possibilities of developments and concepts.

v) Direction for subsequent practice sessions.

vi) Practice recording details.

Throughout the project, journaling played an essential role in documenting discoveries of musical materials and directing subsequent developments. It facilitated the process of self-teaching through critical self-reflection and provided an overview of the trajectory of the project. A transcription of a selection of practice notes is included in Figure 3.4.

15/9/20 Practice session 6 dome
9:30 am start
- beginning with long notes
- same note as a fly
- the fly has great tone changes and texture pulsing
- playing with my tone, it is easy to experiment with tone in the dome
- playing with long notes and palate pressure/ air direction
- changing with rhythmic pulse
- much harder at a higher register to beat between textures
- reveals idea of absolute control over every part of the voice
- maybe the waterfall allows discovery, and the dome allows detailed control?
- recording of 12 minutes
11:30 finish – 2 hours

17/12/20 Practice session 59 dome
5:30pm start
- very wet so no recording
- why does mouth shape help pitch change?
- weird screaming overtone thing, sounds so funny, making me laugh a lot
- different textures, increasing pressure and taking away pitch
- rhythmic pulsing with air
- sound vs integration or phrasing
- how can I organise?
6:30 finish

2/6/21 Practice session 139 crags falls
2pm start
- word set "in a white museum room full of sunlit pink nudes sixty feet high"
- idea to collect pitch sets I like in the same way I collect words
- birds are a big pitch set inspiration
- pitch and word sets don't matter like exploration of sounds
- not interested in creating long song-like melodies
- Start sound (usually consonant) at pitch
    like plucking a string
    dynamic – often dictated by technique, register, vowel in word
    texture – variables, front and back palate
    oscillate ***try oscillate whilst sliding, can they be separated?
    move pitch – slide, step
    stop pitch by stopping air ***is that the end of sound?

stop note – usually on consonant
- work on big oscillate, it's not happening
- no recording, no battery
4pm finish – 2 hours

22/10/21 Practice session 189 crags falls
1:30 start
- word set "I realised people are decent and I feel sad slash fat"
- practicing oscillate with 1 and 2 semitones
- trying oscillate with alternating 2 3 2 semitones
- keep working on specific ricochet numbers
- what if I practiced and really knew every oscillation interval?
- could I really transfer through oscillate pitches?
- maybe pitch should be seen as a spectrum ranging in complexity
- 20 mins recording
3:30 finish

Figure 3.4. Transcribed excerpts from handwritten practice notes as recorded in the practice journal.

B. Practice as solo improvised performance

Large parts of practice sessions were treated as solo improvisation performance. This kind of practice is described by Johansen (2017) as explorational practice as opposed to deliberate practice, which is characterised as a targeted and task centered process (Huoven, Tenkanen & Kuusinen, 2011). Research has challenged that deliberate practice is sufficient to develop expertise in music performance (Bonneville-Rousy & Bouffard, 2015; Macnamara, Hambrick & Oswald; 2014) and studies have suggested that pedagogues should give more attention to the development of creative, spontaneous, and interactive dimensions of music through explorative practice (Hickey, 2015; Huoven, Tenkanen & Kuusinen, 2011; Johansen, 2017; Johnston, 2013; Ward-Steinmen, 2014). Ecological dynamics theory promotes the importance of this kind of practice in the development of performance expertise.

In this singing experiment, the discovery and development of musical materials was achieved through exploration of both internal and external information with improvised singing. New musical materials often occurred with the imitation of or interaction/entrainment with environmental sounds and the exploration of technical skills with adjustments to body organisation. Whilst improvising, off-task spontaneous thought, random association, and mind-wandering produced musical materials and new vocal techniques. These implicit and unplanned occurrences during singing improvisation in the practice settings facilitated discovery and development of musical materials and contributed to my intention, attention, and calibration capabilities.

C. Development procedures

Development of musical materials was achieved through improvised singing practice and the deliberate task-related practice of my self-regulated learning processes (Section 3.4.2). Of particular interest in this experiment was voice-regulation as a development procedure. For example, a vocal technique discovered at a waterfall can change when sung in the polytrihedral dome due to the automatic calibrations in air pressure from voice regulation. This development procedure emphasises the defining aspect of this singing experiment: focusing upon the 'where' of practice just as much as the 'what' of practice for development of music materials and skill acquisition.

D. Evaluation and selection

Evaluation of musical materials was applied throughout the singing project in both immediate and reflective ways. Throughout practice sessions, immediate evaluation of musical materials took place in ways that were both conscious and intrinsic, pre-reflexive, and unconscious. Additionally, practice recordings were listened to after sessions for evaluation of musical materials and practice choices through critical reflection. This process of reflective self-evaluation has been demonstrated to be useful in self-directed learning ability and academic self-efficacy research in many fields including medicine (Song & Kim, 2015; Yoo, M.S., Yoo, I.Y., & Lee, 2010), teaching (Baecher et al., 2013; Colwell, 1995), sports pedagogy (Downs, Miltenberger, & Biedronski, 2015; Kitsantas & Zimmerman, 1998), and music performance (Hewitt, 2001). This post-practice self-evaluation process allowed for reflective evaluation of materials discovered and developed. Sometimes the recordings revealed the discovery of materials that were not recognized in the moment of practice. Discovered materials and techniques are selected to be developed or abandoned according to the criteria of originality, efficiency, and transfer (see research objectives, Section 3.1).

3.6 Conclusion

This chapter outlined the theoretical basis, experiment design, and methods of the singing experiment. All these aspects were combined to create an experimental intervention on singing practice to expand opportunities for changes in voice quality, and discovery and development of musical materials. Two opposing atypical practice settings were selected to explore with improvised singing, one of these settings was designed and built for the experiment. This practice framework is considered to be a research outcome and a contribution to the body of practice-led research. The experiment included novel architectural design, an analysis of which is provided in Chapter 4. The implementation of the experiment successfully led to changes in the quality of the voice, also outlined in Chapter 4, and new musical techniques and materials, outlined in Chapter 5. The following chapter includes an analysis of practice settings acoustics, my voice whilst singing in practice settings, and my singing voice quality (as described by spectral centroids) over the length of the experiment. The creative outcomes that were a product of the singing experiment are outlined and included as audio recordings in Chapter 5.

Chapter 4: A polytrihedral dome for acoustic retroreflection, voice regulation in extreme acoustics, and changes to the singing voice

This chapter examines the acoustics, voice regulation, and changes to voice quality in the singing experiment. The goals, theoretical basis, and methods of the singing experiment are outlined in the previous chapter. The prior empirical research on which this experiment is based is outlined in Chapter 1. Most of this chapter has been published as:

Rapp, M., Cabrera, D., Lu, S., 2022, A polytrihedral dome for acoustic retroreflection, and its application to creative-arts practice-led research, *Applied Acoustics*, 195(108860).

The polytrihedral dome was designed by Dr Densil Cabrera and myself. I built the dome, performed the measurements, and was the sole-participant in the singing experiment. The computational acoustic simulations were done by Dr Shuai Lu. All authors contributed to the interpretation of the results, and to editing the final article text (also following the suggestions by anonymous reviewers).

Abstract

This chapter examines acoustic retroreflection from a polytrihedral dome that was used in the singing experiment as a high voice support singing practice space. The polytrihedral dome was investigated with laboratory measurements, field recordings of singing practice, and by finite-difference time-domain (FDTD) simulation of sound reflections within the polytrihedral dome and other kinds of domes for comparison. Measurements and FDTD simulations show that the polytrihedral dome is retroreflective, especially in the 2 kHz octave band and higher, and provides high levels of voice support. Oral-binaural measurements of voice support exhibit a 4 dB/octave spectral slope from the 500 Hz to the 4 kHz octave bands. The power spectrum of singing is measured and used to weight voice support, yielding measured mean values of +3.5 dB (maximum +6 dB). This value is substantial partly because the on-axis singing energy is greatest in the 2 kHz octave band where retroreflection is particularly evident. The polytrihedral dome was installed at a remote acoustically pristine location for sustained use in singing practice. In-situ field measurements of singing practice show reduced voice levels (room effect, Section 1.1.8) from the dome's voice support, contrasted with increased voice levels (Lombard effect, Section 1.1.2) at a second practice setting, a waterfall. Average voice output between the two environments as measured by a contact microphone differ by 9 dB while the level at the ears remains approximately the same. Using these approaches to investigate reflections in the polytrihedral dome, this chapter documents retroreflective acoustic design that, on a larger scale, might ameliorate sound focussing problems in domed room designs.

4.1 Introduction

Domes and curved surfaces in architecture have been noted for the way they capture and concentrate sound through whispering wall and focussing phenomena. In many domed buildings the geometric focus is not accessible, but whispering wall phenomena are well-known at St Paul's Cathedral (London, UK) and Gol Gumbaz (Bijapur, India) (Cox, 2014). The spherical Mapparium (Boston, USA) provides access to the sphere's focus, entertaining tourists and especially acousticians with the sound of their own or each other's voice (Foulkes, 2013; Hartmann, Colburn & Kidd, 2006). At a spherical dome's focal point one can hear one's own voice concentrated to an extraordinary extent, but the focus shifts away from oneself at other locations.

While architecturally spectacular domes have been objects of acoustic intrigue at tourist sites, acoustic designers often aim to avoid the uneven intense sound concentration from them, e.g., in planetariums (Cremer & Müller, 1982). Acoustically retroreflective architectural surfaces—formed by arrays of concave right-angled trihedra, e.g., in building facades or steps—also focus sound (Crawford, 1991; Cabrera et al., 2018; Cabrera, et al., 2020; Cabrera, et al., 2022). However, retroreflective focusing is different to that of domes in that the focus may be better characterised as an energy focus (rather than a coherent pressure focus) and the focus is at the source regardless of the source's position (within limits). This chapter introduces a hybrid of these architectural forms: an acoustically retroreflective dome. It explores its characteristics in terms of sound concentration and support of one's own voice (autophony) and applies it to a particular type of singing practice (outlined in Chapter 3).

Acoustic environments have been explored by many sound artists and musicians. Acoustics have been used to colour sound (Bandt, 1980) and, in some works, the acoustic environment is the central mechanism of the artwork (Lucier, 1990). Sound sculptures have been designed to have specialized acoustic systems that are passive until activated by something. An example of this is the 'Singing Ringing Tree', a three-meter-tall sound sculpture made of galvanized steel open-ended pipes, the acoustic resonances of which are excited by the wind (Liu & Tonkin, 2022). The artists designed the sculpture to produce dissonant sound fields that vary with localized environmental conditions (Lacey, 2016). The sculpture 'Tvísöngur' by Lukas Kühne (2012) is designed as an acoustic system activated by a person singing or playing an instrument. Kühne (2012) describes the sculpture as five concrete, interconnected domes, each of which has a resonance that corresponds to a tone in the Icelandic musical tradition of five-tone harmony, and works as a natural amplifier to that tone.

By contrast to 'Tvísöngur', acoustic treatments have been introduced to ameliorate the coherent focussing in domed architectural forms which is described as an adverse acoustic effect. Ismail and Elday (2018) presented acoustic design strategies of residential monolithic dome structures and outlined mitigation measures to the concentration of sound energy caused by complex internal concave surfaces. Simulated domes of common sizes were analyzed with ray tracing. Absorptive materials to the internal upper cap of the structure were introduced and a geometric relationship of the acoustic treatment and dome radius was suggested. Reinhardt, Martens, and Miranda (2013) proposed a parametric design process of adapting the height, dimension, and centre point of a dome structure to improve the acoustic behaviour of a domed performance space. Adapted simulated domes were analysed for sound centration patterns to optimize the audience experience in spaces of temporal arts performance.

In this experiment, a polytrihedral dome (Figure 4.1) was designed to optimize sound for a singer, focusing sound back to a source within it through retroreflection (Section 1.2). Forty triangular trihedra were used as a hemispherical retroreflector array (Section 1.2.5). Similar to other studies with acoustically retroreflective arrays (Cabrera et al., 2018; Cabrera, et al., 2020; Cabrera, et al., 2022), sound concentration on the source from retroreflection was expected to be most effective at high frequencies, for which the wavelengths are sufficiently short relative to the size of individual retroreflectors.

The polytrihedral dome was constructed and tested acoustically in a laboratory, and then installed in a remote location for the singing experiment lasting fifteen months (the experiment is outlined in Chapter 3). It was designed as an opposing and alternative extreme acoustic practice space to that of a waterfall, using retroreflection to provide a high amount of voice support (Section 1.1.5).

Figure 4.1. Photograph of the polytrihedral dome situated in the Burragorang Valley (NSW Australia).

This chapter will examine the acoustic characteristics of the polytrihedral dome, for a simple sound source and in relation to the voice of a person in it, the voice regulation that was measured in the polytrihedral dome and at a waterfall in singing practice, and the long-term changes to voice quality over the singing experiment. Key concepts in the examination of the polytrihedral dome, for example voice regulation and voice support, are outlined in Chapter 1.

4.2 Materials and Method

4.2.1 Design and description of the polytrihedral dome

The polytrihedral dome was made from forty triangular trihedral retroreflectors (Figure 4.2) of 2 mm galvanized steel mounted on a nominally 1 m radius 2V steel geodesic dome frame (Szmit, 2017). The geodesic frame was commercially available and two sizes of trihedra (Table 1) were designed and made to fit to the frame. The geodesic area is the triangular aperture formed by the outer edges of the trihedron (the edges that do not intersect the trihedron's apex), the sum of which is 5.82 m2, a little smaller than that of a 1 m radius hemisphere (2π m^2).

Table 4.1. Geometric characteristics of the trihedra used in the dome.

Type	Number	Inner edge lengths	Outer edge lengths	Geodesic area
A	10	437 mm, 437 mm, 437 mm	618 mm, 618 mm, 618 mm	0.165 m^2
B	30	437 mm, 437 mm, 328.2 mm	618 mm, 546.5 mm, 546.5 mm	0.139 m^2

Triangular trihedra were used as retroreflectors because they fit a geodesic dome frame. Although various geometries of trihedral retroreflectors have been extensively studied in optics and radar, such studies are concerned with far-field scenarios. However, the polytrihedral dome has trihedra within a short distance of any envisaged sound source. Simply considering this from a geometric perspective, the rays diverge considerably from source to trihedron, and continue to diverge within the trihedron. Retroreflection by a trihedron involves three successive reflections (first- to third-order; Figure 4.2), and the rays that survive to the third-order reflection are retroreflected. For a source 1 m from the triangular aperture on-axis to a type-A trihedron (Table 4.1), almost all of the trihedron's aperture contributes to the third-order reflection, but those surviving rays fill only 36% of the triangular aperture for the corresponding first-order reflection, with the remaining 64% surface area reflective but not necessarily retroreflective. By contrast, using ray assumptions for far-field on-axis (boresight) incidence, the effective retroreflective area of a regular triangular trihedron is 67% of its triangular aperture, with the tips of triangular trihedra uninvolved in the third-order reflection (Kim & Lee, 2007).

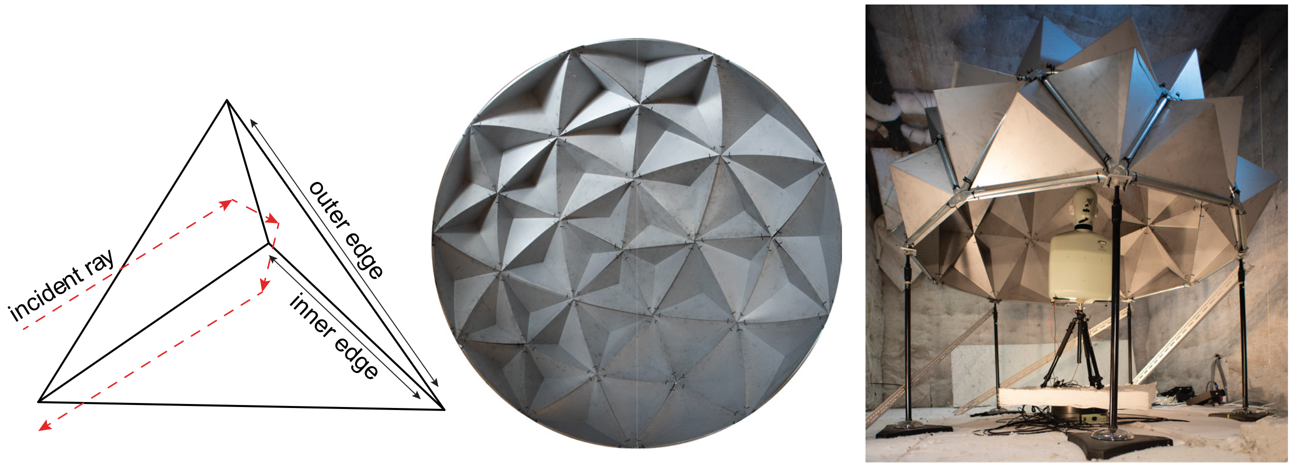


Figure 4.2. Left: diagram of a triangular trihedral retroreflector, with a third-order reflected ray illustrated. Middle: photograph of the dome taken from below (using a 180º fisheye lens). Right: a photograph of the polytrihedral dome in the anechoic room and a head and torso simulator (HATS; Brüel & Kjær type 4128C) in the centre-low measurement position (Section 4.2.2) on a turntable for voice support meaurements at 5º increments.

4.2.2 Acoustic measurements

For the purposes of acoustic measurement, the dome was erected on five posts in an anechoic room (6.5 m × 3.6 m × 3.0 m height) (Figure 4.2). The anechoic room has a graded flow-resistivity sound-absorptive lining (Xu, Buchholz & Fricke, 2005) which is anechoic at and above 200 Hz. The floor has the same anechoic termination beneath a metal grill. Some additional polyester wool sound-absorptive material was placed above the grill. The room has an $L_{\mathrm{A,eq}}$ of 7 dB during normal working hours with building services operating (measured with a Brüel and Kjær (Nærum, Denmark) type 4179 low-noise microphone system).

The polytrihedral dome was measured to quantify ST_V (Section 4.3.1) and to validate FTDT simulations of sound concentration (Section 4.3.2A). Measurements and simulations were done with the the sound source in four positions (Figure 4.3).

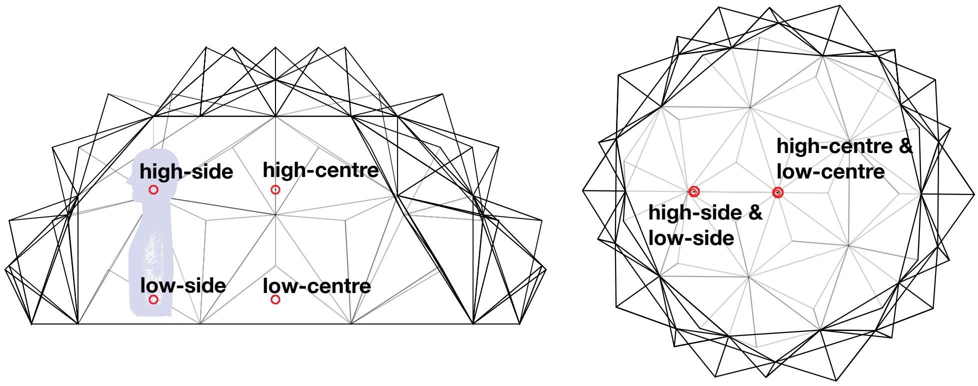


Figure 4.3. Left: four source positions inside the polytrihedral dome used for measurements and simulations. The head and torso simulator positions in relation to the measurement points are shown by the HATS outline in the high-side position. Right: measurement positions from above.

4.2.3 Oral-binaural measurements

Oral-binaural measurements were conducted in 4 positions (Figure 4.3) using a Brüel & Kjær 4128C head and torso simulator (HATS). The HATS was placed on a turntable and measurements were in 5º increments around the full 360º rotation. The test signal was a 10 s exponential sweep (80 Hz – 20 kHz, with an audio sampling rate of 48 kHz), from which impulse responses were derived.

4.2.4 Field and anechoic singing measurements

The dome was installed in a remote location (Figure 4.1) for sustained use in singing practice (see Chapter 3). To measure voice regulation over the fifteen-month singing experiment, field recordings of singing practice were taken in the polytrihedral dome and at a waterfall. Audio recordings were made at the ears, capturing the combined L_D and L_R, and using a contact microphone on the throat (Figure 4.4) to provide audio recordings without any acoustic reflections from the environment, similarly to Astolfi et al. (2019). DPA 4560 Binaural Headset microphones (DPA Microphones, Alleroed, Denmark) were positioned in the pinnae next to the open ear canals. The contact microphone (Jez Riley French C-series, United Kingdom) was fixed at the jugular notch using sports strapping tape. Audio recordings were made by a portable audio recorder (F6 Zoom field recorder) using a sampling rate of 48 kHz. The field recorder's gains were not changed during the 15-month recording process, and the binaural microphones were calibrated so that sound pressure levels could be derived from the recordings.

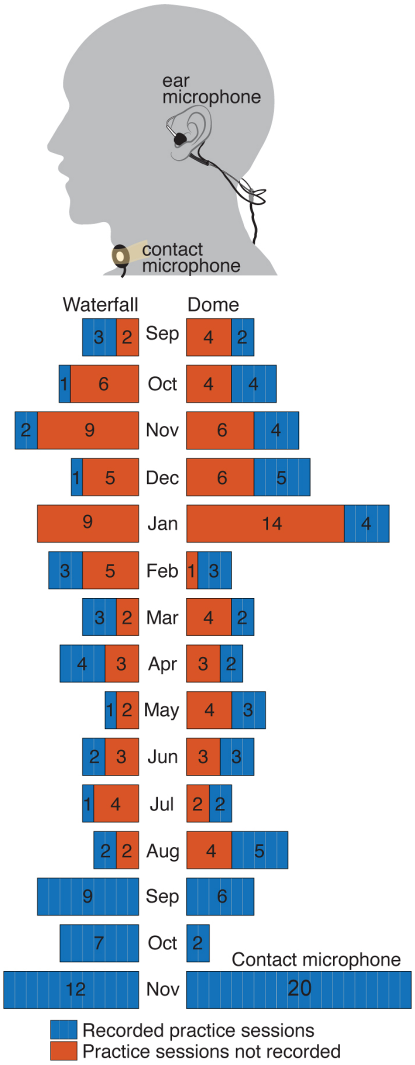


Figure 4.4 Top: Microphone placement for field recordings. Bottom: timeline and numbers of recorded (in-ear microphones), and not-recorded practice sessions at the waterfall (left) and polytrihedral dome (right).

Of 225 practice sessions, 67 in the polytrihedral dome and 51 at waterfalls were recorded with in-ear microphones. Figure 4.4 shows the 15-month timeline, the number of practice sessions in each month, and the amount of those that were recorded. Practice sessions ranged from 1 to 3 hours (average of 1.5 hours, approximate sum of 354 hours) and the recordings were 10 to 30 minutes at the end of the session. If the session was not recorded it was due to the weather or equipment limitations. The contact microphone was added in the final month of the experiment (November 2021) and used in conjunction with the in-ear microphones. Ten practice sessions in the polytrihedral dome and six practice sessions at waterfalls were chosen for voice regulation analysis (Section 4.3.3) due the contact microphone addition, the quality of the recording, and absence of unwanted background sounds. An analysis of all 118 in-ear recordings is included in Section 4.3.3C.

A singing recording session was also conducted in an anechoic room using the same equipment, supplemented with a calibrated omnidirectional microphone at 2 m on-axis (Earthworks M30, Milford NH, USA). The 2-meter on-axis microphone allowed for calculation of the equivalent 1 m on-axis singing levels (frequently used in studies and standards to quantify voice level (e.g., ISO, 2003)) and spectra. Furthermore, an octave-band transfer function between the 1 m reference and the other transducers could be derived—facilitating the interpretation of the field contact microphone measurements.

Another purpose of the anechoic recording session was to find a relevant spectral weighting function for voice support of singing. Using the method outlined by Pelegrín-García et al. (2012) for applying a speech-weighting to voice support (Section 1.1.6), the spectral data from the on-axis microphone was measured to calculate a relevant singing weighting function.

4.2.5 FDTD simulation

Simulation of the sound field in the entire polytrihedral dome was done using finite-difference time-domain (FDTD) method. For comparison two variants of the dome were also simulated, a flat-faced geodesic dome (of the same basic geometry) and a hemispherical dome of the same radius. The purpose of these simulations was to examine the distribution of sound in the dome from omnidirectional sources at each of the four positions, in comparison to the hemispherical and geodesic domes. FDTD is a wave-based simulation method, which inherently accounts for the diffraction and high order propagation effects from architectural forms, and is better-suited than ray-based methods for problems where important architectural elements are similar in size to wavelengths. The codes for FDTD simulation in this research were developed in C+ (for the simulation kernal to improve speed) and Matlab (for other parts including the user interface) based on the framework proposed by Sheaffer et al. (2014). As a feature of our FDTD package, simulation is significantly accelerated by GPU parallel computing implemented based on the CUDA toolkit (NVIDIA, 2022). To perform a simulation, the geometry of the dome is first modelled in Rhinoceros as a solid object, and then exported as a .obj file. The geometry is next voxelized in the simulation volume with a 3D rectilinear grid. The modelled volume was 2.6 m × 2.6 m × 1.4 m = 9.5 m3, excluding the perfectly matched layer, with a grid spacing of 4.3 mm. The simulations were done at a frequency of 144,000 Hz, with the upper cut-off frequency of the simulation results to be 14,400 Hz to limit the dispersion error to below 2% (Bilbao, 2013). The source in the simulation was an omnidirectional soft source, with a Gaussian pulse waveform (Schneider, Wagner & Broschat, 1998), and the receivers were spread over the full enclosed volume, with a cubic grid spacing of 12.8 mm. The dome surfaces were set to be fully reflective, and a 10-voxel perfectly matched layer was added to create anechoic boundaries for the volume (Chern, 2019). The simulations were run for 80 ms (11520 time samples). Longer simulation cut-off times up to 150 ms were also tested to verify that 80 ms sufficed to capture the non-negligible reflected energy.

4.2.6 Validation measurements

For the purpose of validating the FDTD simulations, measurements of the reflected energy inside the polytrihedral dome were made with a 100 mm diameter dodecahedral omnidirectional loudspeaker at the 4 positions in the dome (Figure 4.3). The loudspeaker model (Dr Three 3D-032, Tokyo) meets the BS ISO 26101:2021 omnidirectional criteria for anechoic room qualification traverses up to the 3.15 kHz 1/3-octave band, directionality deviations within ±2 dB. A Brüel & Kjær omnidirectional microphone (1/4-in. type 4135) was positioned initially 10 mm from the source surface (approximately 60 mm from the source centre) and then at 50 mm increments along a horizontal linear traverse to the edge of the dome. Impulse responses were acquired using an exponential sinusoidal sweep (90 Hz – 20 kHz, sampling rate 48 kHz, duration of 20 s per sweep).

4.3 Results

4.3.1 Voice Support

Voice support for each measuring position was calculated following the method of Brunskog et al. (2009) and Pelegrín-García (2011) in the 125 Hz – 4 kHz octave bands. The arithmetic mean of octave band ST_V for the combined measurement positions is 0.03 dB (range -2.7 dB to 4.5 dB). ST_V as a function of frequency is shown in Figure 4.5 and shows a consistent growth of approximately 4 dB/octave from 500 Hz to 4 kHz. The 4 kHz octave band has the highest levels of ST_V with a range from 5.4 dB to 11 dB and an arithmetic mean of 8.6 dB.

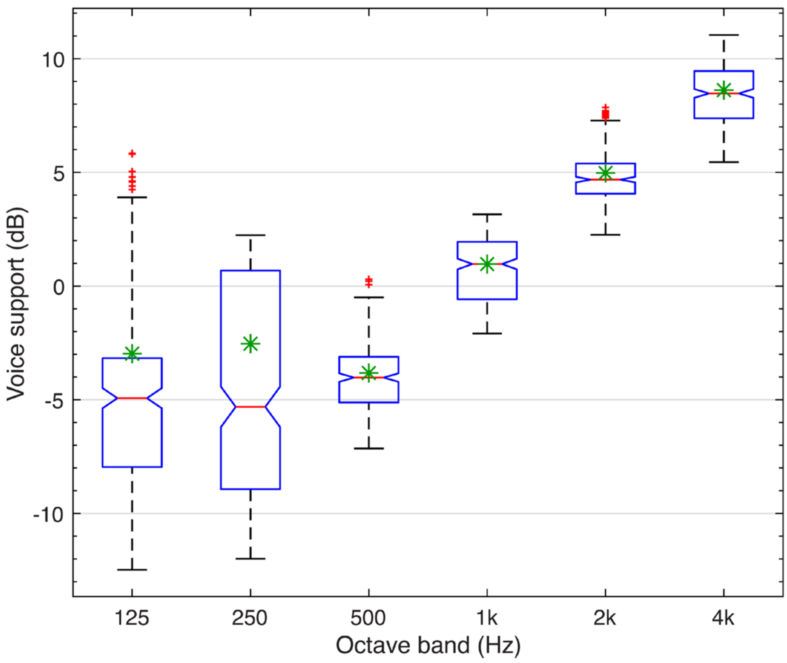


Figure 4.5. Octave band voice support for the polytrihedral dome (72 orientations at four positions). Star (*) markers show means.

A. Measurements for calculating singing-weighted voice support

The on-axis anechoic singing recordings were used to calculate the LD received at the eardrum, following the method that Pelegrín-García et al. (2012) used to derive speech-weighting. Figure 4.6 shows the on-axis spectrum of the singing in comparison with 8 professional opera singers (Cabrera, Davis & Connolly, 2010), 15 people with at least 2 years of post-high school private singing voice training (Monson, Hunter & Story, 2012), and the speech spectrum used for speech-weighted ST_V (Pelegrín-García et al., 2012). The measured singing spectrum is similar to the professional opera singers' spectrum, with a spectral centroid (1.35 kHz) a little higher than the opera singers (1.15 kHz).

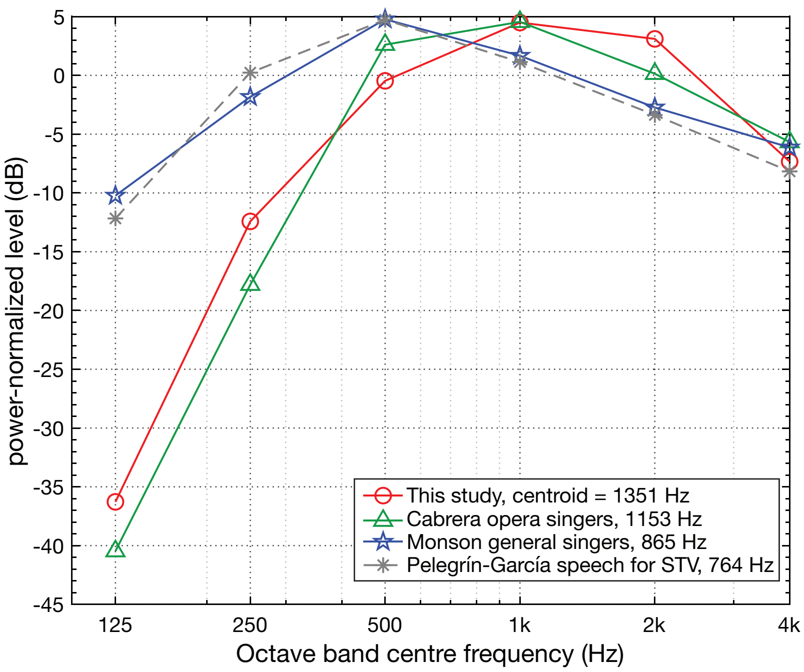


Figure 4.6. Power-normalised octave band spectra of my singing, professional opera singers (Cabrera, Davis & Connolly, 2010), general trained singers (Monson, Hunter & Story, 2012), and the speech spectrum used for speech-weighted ST_V (Pelegrín-García et al., 2012).

B. Speech and singing-weighted voice support

The overall arithmetic mean of the speech-weighted ST_V in the polytrihedral dome is 1.1 dB (range -0.7 dB to 3.6 dB). With its greater emphasis on high frequency octave bands, the singing-weighted ST_V measurements in the polytrihedral dome are greater than the speech-weighted counterparts, with an arithmetic mean of 3.5 dB (range -0.9 dB to 6 dB). Figure 4.7 shows both the speech and singing-weighted ST_V values for each measurement position. Singing-weighted ST_V is consistently greater than speech-weighted, with an average difference of 2.4 dB. The high-side position has the greatest levels of ST_V.

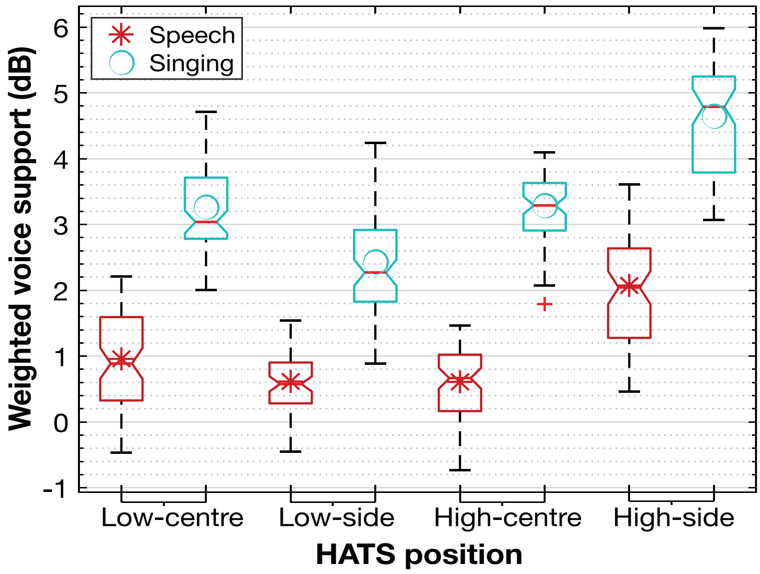


Figure 4.7. Speech-weighted and singing-weighted ST_V levels of the four HATS measurement positions (72 measurements at 5º increments for each position). Markers show means.

4.3.2 Sound concentration

A. FDTD validation

Figure 4.8 compares the sound concentration validation traverse measurements made of the polytrihedral dome in the anechoic room with the results from FDTD simulation measurements. Impulse responses were analysed by filtering into 8 octave bands (125-8000 Hz) and Figure 4.8 shows the 250 Hz and 2 kHz octave bands as examples. The R2 and RMSE of each octave band are shown in Table 4.2. The measurements and simulations are aligned, more closely in the higher octave bands which are the most pertinent when examining retroreflection. These measurements indicate that the FDTD simulation matches measurements closely in the high frequency range, with larger deviations at low frequencies that might be attributed to the metal panels being neither perfectly rigid nor sealed. The smaller deviations at very high frequencies are more likely due to geometric errors and loudspeaker directivity.

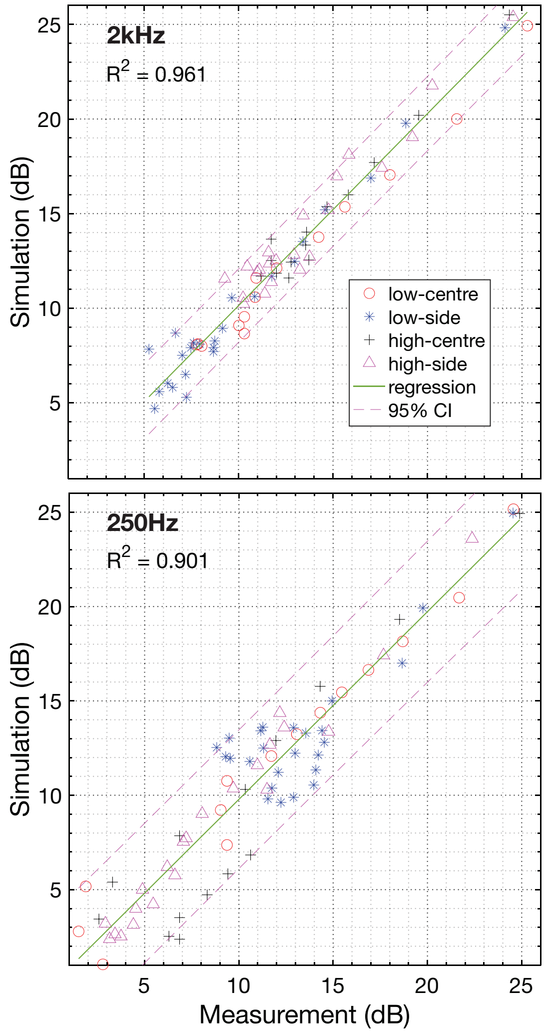


Figure 4.8. Measurement and simulation comparison for traverses in the polytrihedral dome for the 2 kHz and 250 Hz octave bands.

Table 4.2. The coefficient of determination (R2) and root mean square error (RMSE) of octave band measurement and simulation comparisons.

Octave band	125 Hz	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz	8 kHz
R^2 RMSE (dB)	0.924 1.64	0.901 1.82	0.831 1.96	0.924 1.27	0.961 0.96	0.945 1.19	0.921 1.43

In addition, the impulse responses acquired from the measurements and the FDTD simulations are compared. Impulse responses were filtered into octave bands and Figure 4.9 shows the 2 kHz band as an example. Data in the figure is the absolute value of the Hilbert transform showing the envelope of the impulse response for the first 12 ms, and the Y-axis shows how the impulse response varies with distance from the source (starting from 6 cm from source centre to the edge of the dome in 5 cm increments; note that the farthest distance varies between traverses). The amplitude values are normalized between 0 and 1 (relative amplitudes between impulse responses are preserved), with the colormap clipped at 0.1 so that sound reflections can be shown clearly. The measured and simulated impulse responses are in good general agreement, both spatially and temporally, with minor deviations that are potentially caused by sound speed discrepancy, geometric errors, and loudspeaker directivity.

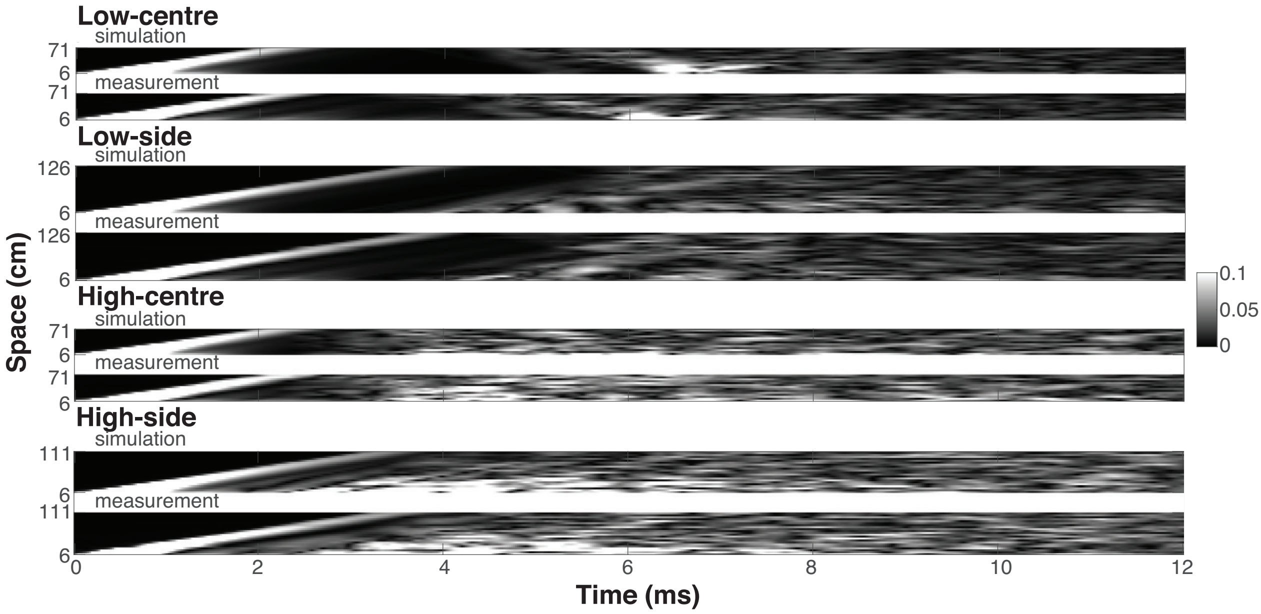


Figure 4.9. Envelopes of the 71 impulse responses (up to 12 ms) from the four traverses, comparing simulations with measurements. For each traverse, the upper subplot shows the FDTD simulation, and the lower subplot shows the corresponding set of measurements.

B. Sound concentration simulations

Figure 4.10 shows the frequency-dependent sound concentration of reflected energy as calculated by FDTD simulation for the polytrihedral dome ('poly') and, for comparison, a flat geodesic dome ('geo') and a hemispherical dome ('hemi') of the same radius. All four source positions are shown up to 12 ms (i.e., the first reflections), and then in the final cumulative state at 80 ms. The figure excludes the direct sound, which was removed by subtracting the corresponding fully anechoic simulations.

The value L in Figure 4.10 is the reflected sound energy level at the source position, relative to the emitted sound energy from the source at 1 m distance. An acceptance sphere 0.1 m radius is used for the reflected energy at the source position. The value ΔL is the difference between the reflected energy level at the source position and the average energy level over the dome's entire volume (using the same 0.1 m source position acceptance sphere). The colour-mapped visualisation is across a horizontal plane at the height of the source, with values relative to the emitted sound energy from the source at 1 m distance.

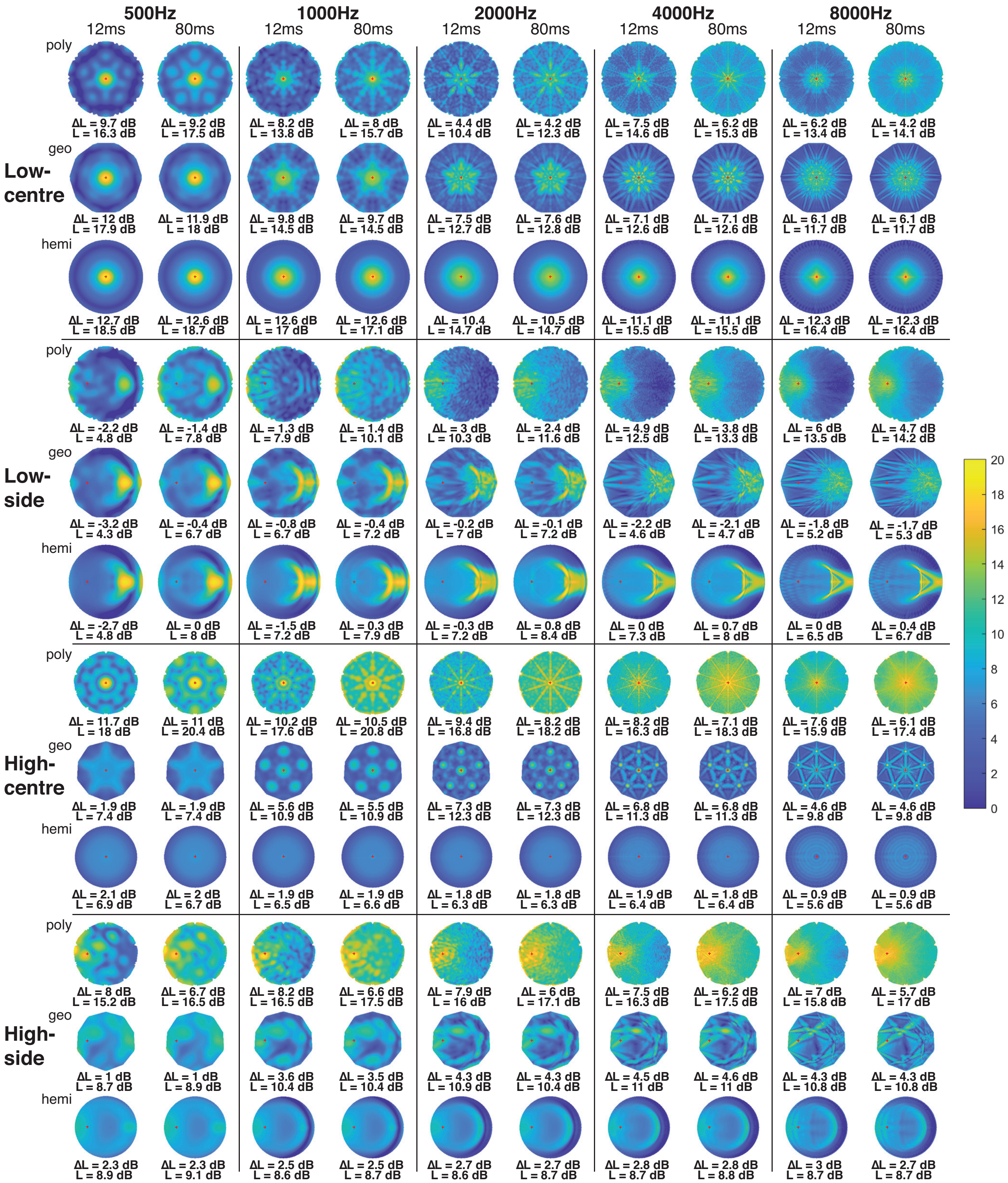


Figure 4.10. Reflected sound concentration derived from FDTD modelling for 1. Polytrihedral dome ('poly'), 2. Geodesic dome with flat surfaces ('geo') and 3. Hemispherical dome ('hemi'), all with a radius of 1 m. The sound source position is shown by the red cross. Values are in decibels and the total energy (L) and the spatial contrast (ΔL) are written on each chart. Energy integrated to 12 ms captures the first-order reflections, energy integrated to 80 ms captures essentially all the reflected sound.

Even though none of the source positions is at the hemispherical dome's geometric focus, it shows considerable focusing behaviour. The low-centre source position is not far from the sphere's centre and shows a large concentration of reflected energy onto the source (with $L \geq 14.7$ dB, $\Delta L \geq 10.4$ dB) This energy is almost entirely from the first reflections. On the other hand, the side positions show a reflective focus away from the source, on the opposite side of the dome. For the side positions, the spatial form of the focus as mapped on the horizontal receiver plane becomes more finely structured with frequency. The high-centre position has no concentration of reflected sound energy at the source (and presumably focuses to a position below the dome's equator). The L and ΔL do not have significant changes from 12 ms to 80 ms (except the low-side position in the 500 Hz band) due to the sound leaving the dome.

The polytrihedral dome exhibits retroreflective focusing at high frequencies, as well as sphere-related sound concentration at low frequencies. For the low-centre, the result is not very different to the spherical dome result. However, for the side positions the transition from spherical behaviour to retroreflective behaviour is evident as the reflected sound concentration shifts from right (at low frequency bands) to left (high frequency bands). Unlike the hemispherical dome, the polytrihedral dome concentrates reflected sound on the high-centre source position, for all the frequency bands and beyond the first reflections. The high-centre has the largest contrast in sound concentration to the other domes, with a difference of up to 13 dB in L and 9 dB in ΔL. Changes in both L and ΔL over time demonstrate that the sound in the polytrihedral dome is more contained than the hemispherical and geodesic domes. This is consistent with the notion that retroreflection reflects back to the source, which is above the dome's equator.

The geodesic dome with flat surfaces has similar results to the hemispherical dome, except for the high-centre source position, which shows a pentagonal pattern of sound concentration (following the pentagonal rotational symmetry of the geodesic dome). Like the hemispherical dome, L and ΔL have almost no change between the first reflections and the final state.

4.3.3 Singing field measurements

Sections 4.3.3A and 4.3.3C describe the results are from final month of the singing experiment when the contact microphone was used in conjunction with in-ear microphones. The long-term analysis of the singing (sound levels and the spectral centroid at the ears) spanning the whole experiment can be found in Section 4.3.3C.

A. Sound at the ears

The in-ear field recordings of practice sessions sound very different at the two practice environments (dome and waterfall). The recordings in the polytrihedral dome include the L_D and L_R of the singing, and the waterfall recordings have the L_D from both the singing and the constant broadband sound of the waterfall (Figure 4.11). The average A-weighted sound pressure level (L_{Aeq}) at the ears when singing in the dome is 98.9 dB and when singing at the waterfall is 101.0 dB. The L_{Aeq} at the ears of the anechoic room singing was 98.8 dB, making the average L_{Aeq} range of the sound at the ears in all singing environments 2.2 dB.

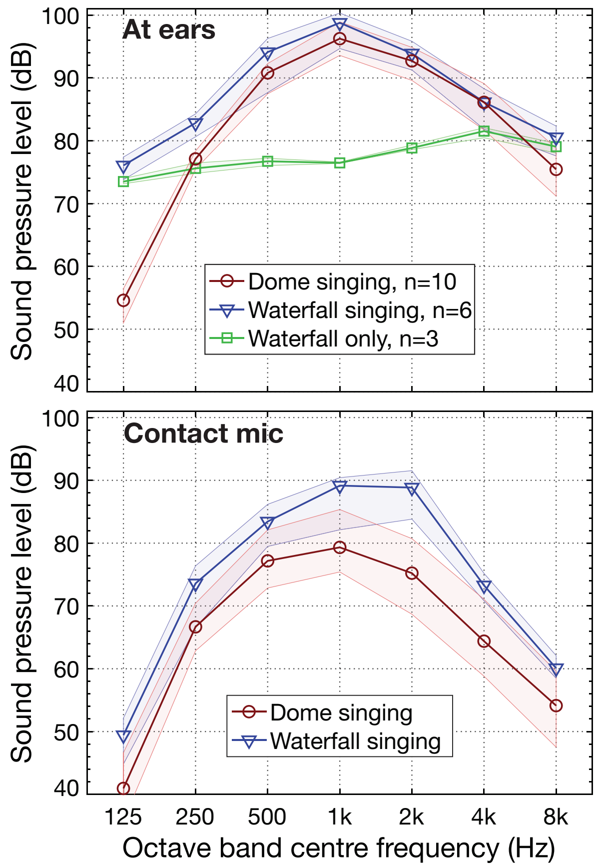


Figure 4.11. Field Leq measurements (power means and ranges) of singing in the polytrihedral dome and at a waterfall measured with in-ear microphones (top) and with a contact microphone taped to the throat. Shading shows range from minimum to maximum. Sound pressure levels derived from the contact microphone are equivalent to 1 m free-field on-axis, derived as described in the text.

Field measurements of practice sessions with in-ear microphones indicate that voice regulation occurred to keep the sound approximately constant at the ears for the two practice environments (Figure 4.11). The average difference between the sound at the ears in the practice environments for the 500 Hz to the 4 kHz octave bands is 3 dB, with the 4 kHz band having a negligible difference. The 125 Hz band is obviously affected by the sound of the waterfall.

B. SPL of voice measured by contact microphone

The contact microphone (adjusted with reference to the anechoic on-axis microphone to 1 m equivalent) measured greater sound pressure level of singing in all octave bands when singing at the waterfall (Figure 4.11).

The greatest difference is 13.6 dB in the 2 kHz octave band. The average difference between all the octave bands is 9.3 dB. The average L_{Aeq} of polyhedral dome singing as measured by the contact microphone (adjusted to 1 m free field equivalent) is 81.1 dB and when singing at the waterfall is 93.0 dB. Hence the range of average L_{Aeq} over the singing environments spans 11.9 dB.

C. Long term sound at ears

Figure 4.12 shows in-ear recordings of 118 practice sessions over 15 months; 51 at a waterfall and 67 in the polytrihedral dome. Sound levels at the ears over time do not show a trend, except for a smaller range of values (no values below 90 dB) in the dome from about 300 days on. The average spectral centroid of singing increases over the length of the project. Here we use power spectral centroid as an indicator of the spectral bias in the voice. It is one of a number of spectral bias parameters reported in the singing science literature, others include singing power ratio and long-term average spectra (Watts et al., 2006).

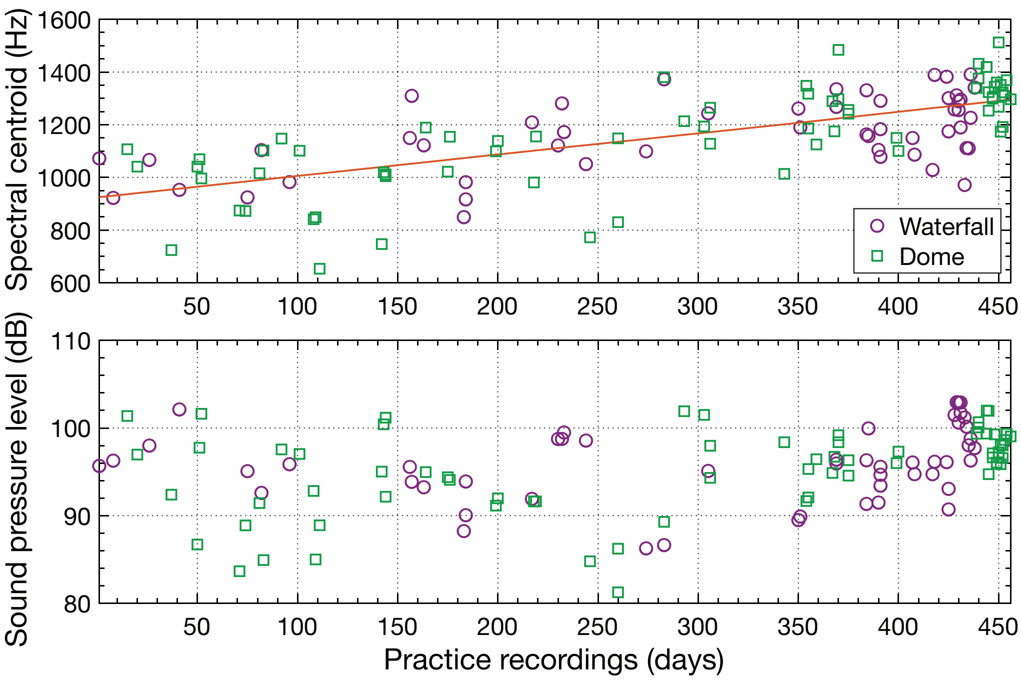


Figure 4.12. The spectral centroid (calculated from octave band power levels) and voice levels of singing practice sessions as measured by in-ear microphones for every practice session over 15 months (or 456 days).

4.4 Discussion

Predominantly retroreflective rooms are rarely seen in architecture. Most research on pre-existing retroreflective architecture is concerned with incidental reflection from building facades (Crawford, 1991; Cabrera et al., 2018; Cabrera et al., 2020). Indian stepped ponds also provide pre-existing cases of acoustic retroreflection, due to the large number of simultaneously visible reasonably proximate concave trihedra in some cases (Cabrera et al., 2022). Research on retroreflection in auditoria for musician acoustic support is concerned with relatively small-scale designs (Toyota et al., 2021; Tuominen, Rämö & Välimäki, 2013), ancillary to the main purpose of such auditoria, which is acoustic transmission to the audience. Some studies consider speculative retroreflective designs, e.g., for open plan office ceilings (Caldwell, 2019; Hannouch et al., 2018) or in a small room (Cabrera et al., 2021). Hence the polytrihedral dome is a distinctive acoustic environment partly due to the sheer intensity of the treatment, considering the number and proximity of the retroreflectors to a person or sound source within it. Furthermore, it is distinctive for its situation, in the remote Burragorang valley (Australia), a largely absorptive environment of sheltered sandstone forest. It is not a closed room, raised 1.3 m in the air by wooden posts, which allowed the feeling of being outside whilst practicing inside it, yet sounds nothing like singing in the open outdoors.

Whether this type of acoustic structure has broader applications in larger scale architectural design remains for future studies or design projects. However, considering that domes are visually appealing but have notorious acoustic focusing problems, it is possible that retoreflective treatment could have a role in both ameliorating problems and contributing to acoustic support in prospective domed room designs.

4.4.1 Retroreflection in the polytrihedral dome

The combination of a hemispherical dome with trihedral array treatment results in hybrid focusing of sound from a source within the dome. At high frequencies the retroreflection overwhelms hemispherical focusing, and the focus follows the source. Retroreflection from the array of trihedra increases up to the 8 kHz octave band and is weakly evident at 500 Hz, at least for the high-side source position. Sound concentration in other domes is prescribed by their larger scale geometry, but in a retroreflective dome the sound concentration is dictated by where the source of the sound is.

While the various propagation phenomena that underly the results are well-understood, a quantitative theory of the polytrihedral dome's reflected sound field would likely be considerably more complicated than the numerical simulation. The hybrid design, near-field reflections, and the wavelength relationships to the sizes of the reflective forms make the wave-based simulation simpler than any potential theory. However, simplistically estimating retroreflected energy from 40 perfect reflectors at 1 m distance (neglecting diffraction, interference, scattering and higher-order reflections) yields an energy sum of 10 dB. Incoherent ideal reflector sums explicitly calculated for the four source positions and 40 trihedron apexes are similar: low-centre 9.4 dB; low-side 9.9 dB; high-centre 10.5 dB; and high-side 10.9 dB. These values are a few to several decibels less than the reflected energy levels (L) found in the simulations at higher frequency octave bands up to 12 ms. Constructive interference (from retroreflectors at similar distances) and scattering (from non-retroreflective parts of the trihedra) would be expected to increase the values, while diffraction would be expected to reduce them (especially at low frequencies).

4.4.2 Voice support of the polytrihedral dome

Voice support measurement results in Section 4.3.1 show that the polytrihedral dome has much greater ST_V than everyday rooms, for example, Brunskog et al. (2009) reported that everyday rooms have ST_V values of between -15 and -10 dB. Higher values may be found in small rooms, but it is rare for reported values to exceed -5 dB. To compare the polytrihedral dome with rooms used for music practice, the Pelegrín-García et al. (2012) prediction model of ST_V was applied to theoretical small music practice rooms with dimensions and reverberation times recommended by BS ISO 23591:2021 (Figure 4.13). Evidently, acoustic conditions as per the recommendation for music practice rooms yield much lower ST_V than the polytrihedral dome.

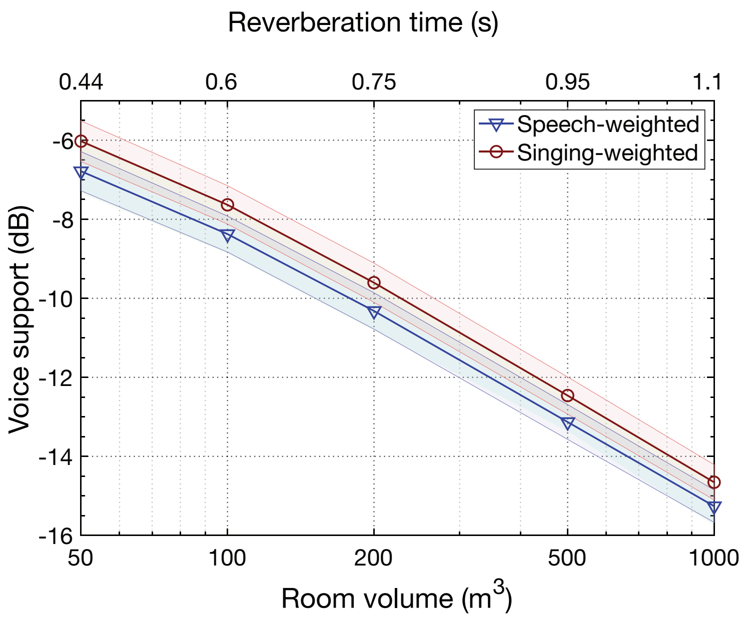


Figure 4.13. Speech-weighted and singing-weighted voice support predictions for theoretical music practice rooms with dimensions and reverberation times recommended by BS ISO 23591:2021. Shading shows the range of voice support values derived from the range of recommended octave-band reverberation times. Average (unweighted) ST_V values are not shown, but are similar to speech-weighted.

Octave band ST_V in typical rooms tends to increase with frequency. This is because the mouth simulator transmits little sound at high frequencies directly to the ears and both mouth and ears are increasingly directional with frequency, so the denominator of the ratio used to calculate ST_V becomes relatively small (Cabrera et al., 2018). Pelegrín-García et al. (2012) measured frequency dependent voice support in 27 small and medium school classrooms (volume < 500 m^3). The average ST_V of the classrooms increased by between 0.7 to 1.9

dB per octave between the 500 Hz and 4 kHz octave bands. The polytrihedral dome has a much steeper ST_V spectral slope of approximately 4 dB per octave. Similarly steep ST_V spectra have been observed in other studies of retroreflective environments and are partly attributable to diffraction loss from the retroreflectors at low frequencies (Cabrera et al., 2021). The combination of singing spectrum (of the profile observed in this experiment, which is similar to that of professional opera singers) and the ST_V spectrum increasing steeply with frequency leads to overall increased ST_V for a singer in the polytrihedral dome.

4.4.3 Voice regulation

Field singing measurements showed that voice regulation occurred to maintain a similar sound level at the ears between the singing environments. This is consistent with findings established in early studies, where sidetone was altered by dampening the direct auditory feedback or amplifying the voice of the talker to the ears. This led to talkers regulating their voice level to maintain a constant autophonic sound at their ears (Lane, Catania & Stevens, 1961; Lane, 1963).

When comparing this experiment with other voice regulation research, it is important to note that research on voice regulation in singing is very limited, with most reasearch using speaking scenarios. Comparison of voice regulation in singing and talking has limitations as the sound of the singing voice differs from the ordinary speaking voice in a variety of ways (Sundberg, 1990; Hutchins & Peretz, 2011; Monson, Hunter & Story, 2012). Even comparisons between speaking studies is challenging, because changes in voice level are not only influenced by the indirect autophonic gain, but also by the communication scenario, such as inter-talker distance and the instructions or tasks given to the talkers (Astolfi et al., 2015; Astolfi et al., 2019; Garnier, Henrich & Dubois, 2010; Cooke & Lu, 2010).

Voice regulation has been observed in studies of singing due to acoustic reflections or room reverberation (Kittimathaveenan & Park, 2021; Bottalico, Graetzer & Hunter, 2015; Yadav & Cabrera, 2017) and due to sound in the environment (Coleman & Hicks, 1978; Tonkinson, 1994; Bottalico, Graetzer & Hunter, 2015). Bottalico et al. (2015) measured a voice regulation range of 1.17 dB for nonprofessional singers and 0.13 dB for professional singers associated with a 7.7 dB change in voice support (from the introduction of sound reflectors near the singer). Bottalico et al. (2015) also observed larger compensatory voice regualtion for the sound pressure level of musical accompaniement: singing SPL increased by 4.3 dB for nonprofessionals and 2.6 dB for professionals for a 20 dB increase in accompaniement, which is arguably a form of the Lombard effect.

I am a professional and trained singer and was not self-consciously regulating my voice throughout the singing experiment, yet the average voice regulation range is 11.9 dB from the acoustically supportive dome to the noisy waterfall. This is probably due to the extreme contrast between the acoustic environments, however, it must be acknowledged that there are limitations in getting reliable results from a contact microphone in measuring voice level as it is sensitive to placement and attachment which cannot be replicated perfectly for each recording.

4.5 Conclusions

This chapter has presented an analysis of an unusual hybrid dome design, the polytrihedral dome, which combines acoustic retroreflection with the form of a hemisphere. Numerical acoustic simulations indicate that retroreflection has a strong influence on the reflected sound field in the high frequency range, evidenced by a concentration of reflected energy at the source. This contrasts with the sound focusing behaviour of a hemispherical dome. Consequently, the dome is shown to provide unusually strong acoustic support for one's own voice in the high frequency range, based on physical measurements.

Installed in a near-pristine soundscape, the polytrihedral dome was used in the singing practice experiment over many months. The singing voice can have a power spectrum that is weighted more to high frequencies than conversational speech, which is attuned to the polytrihedral dome's high frequency bias. Measured voice levels over the course of the experiment are consistent with voice regulation from the dome's acoustic support (i.e., reduced voice power), contrasted with Lombard effect voice regulation at a waterfall (i.e., increased voice power), such that the sound level received at the ears was similar in each environment.

Chapter 5: Creative outcomes- a taxonomy of singing materials and organisational devices, and applications in ensemble improvisation and further training

This chapter details the creative outcomes of the singing experiment. These outcomes are as follows:

i) A taxonomy of musical materials discovered and developed during the singing experiment. Musical materials that are products of specific vocal techniques are described as idioms. Audio documentation of idiom development throughout practice sessions supports the idiom descriptions and are offered here as primary data outcomes of the experiment. The ways in which singing idioms are applied and structured in a performance are described as organisational devices.

ii) Studio recordings of the application of materials in a solo piece, and three duo improvisations. The materials were recorded at the conclusion of the experiment and are described with only recording details; there is no further exegesis or analysis of music.

iii) Application of new vocal techniques to further vocal training with Bae Il Dong, including a description of a post-singing experiment period of teacher/student training, within which the newly acquired skills were used to enhance student understanding of and ability to acquire the teacher's technical instruction. An improvised music recording is included as a further creative research outcome and represents evidence of applied skills from further training.

5.1 Taxonomy of discovered and developed materials

The musical materials discovered and developed in the singing experiment are best described via two categories: i) singing idioms, and ii) organisational devices. Both singing idioms and organisational devices are described below. Singing idioms are also presented with audio excerpts from the practice recordings.

5.1.1 Nomenclature

The names of both the idioms and organisational devices emerged throughout the practice sessions as a way of identifying the materials to myself and signifying a performance action. Whilst the names have connections to well-known music terminology, these connections are superficial and aren't meant to offer an alternative terminology. The names served the purpose of talking to myself in both practice and performance via a shorthand communication.

5.1.2 Singing idioms

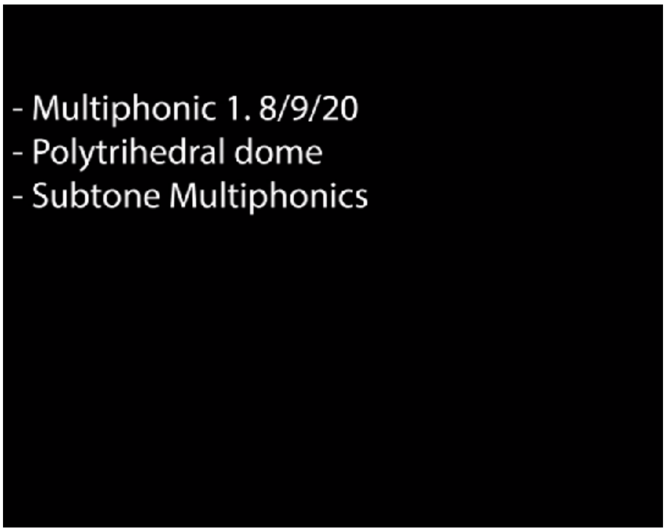
Slater (2020) introduced the term idiomatic types to categorise musical motives that emerge from instrument-specific affordances, including articulation patterns, technical adaptations, and pitch-space organisation. The term idiom is used here to describe a foundational singing technique that was discovered in practice or taught to me by my teacher Bae Il Dong (see singing background, Section 3.2). The idioms all have multiple variations and iterations that are a result of development practice procedures (Section 3.5.6C). Idioms are applied to target pitches (an organisational device, see Section 5.1.3B) and several idioms can be overlapped.

The singing idioms are described below and demonstrated with audio excerpts from practice sessions at both the polytrihedral dome and the waterfall. Most singing examples are from dome practice session Q8 Zoom recordings as they provide the most clarity and demonstrate the singing without the strong frequency dependent reflections from the polytrihedral dome to the in-ear microphones (see Sections 3.5.4, 4.2.4, and 4.3.3A). The contrast of the sound of singing in the polytrihedral dome between the Q8 Zoom and the in-ear microphones is demonstrated by the practice recording Audio 5.1 which includes the same moment of practice as recorded by first the Q8 Zoom and then the in-ear microphones. Excerpts from waterfall practice sessions are selected from in-ear microphone recordings due to waterfall noise dominating Q8 Zoom recordings. Examples of idioms are presented as visual/audio content in chronological order with the date of the recorded example and some additional information included as text.

Audio 5.1. The contrast of a moment in singing practice between i) the Zoom Q8 recorder (0:00 - 00:24) and ii) the DPA Binaural Headset in-ear microphones (00:25 - 00:49).

A. Multiphonic idioms

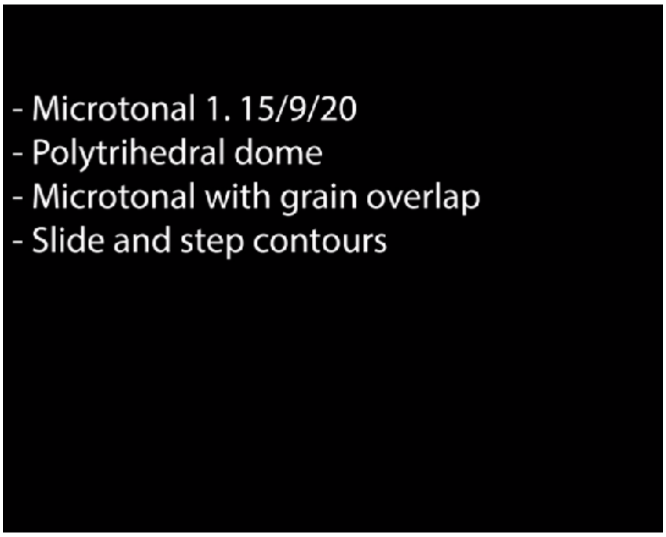
Multiphonic idioms can be categorised into three types of effects: cluster, subtone, and whistle. These are applied in degrees that result in relatively stable multiple intervals through to very unstable distortions of sound. Multiphonic idioms were discovered through random experiments in air control and mouth/palate shapes. A selection of multiphonic idioms is shown in Video 5.1.



Video 5.1. Multiphonic idioms selection.

B. Microtonal idioms

Microtonal idioms are used to expand the tonal palette. They are seen as a pitch degeneracy of the target pitch where the strength of the fundamental is degraded through micro pitch movements. These idioms can be explored and exploited in various ways to create interesting melodic motions. Many of the microtonal idioms were discovered through attempting to integrate with the nontonal sounds of the practice environments. Video 5.2 presents several microtonal idiom iterations.

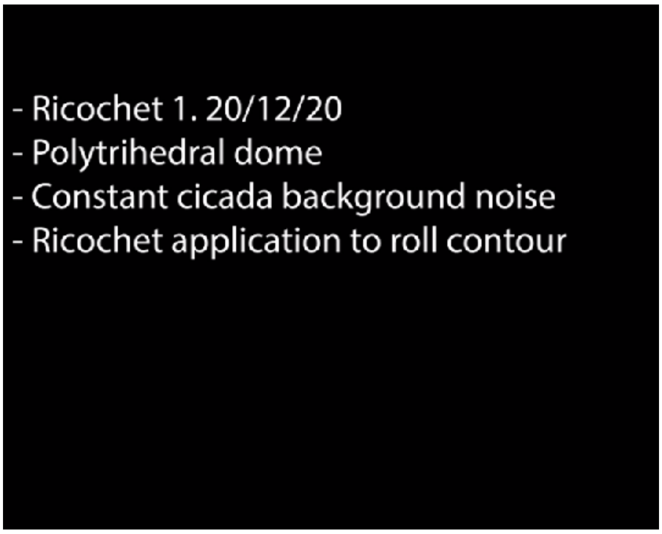


Video 5.2. Microtonal idioms selection.

C. Ricochet idioms

Ricochet idioms are repetitive fragmentations of the target pitch. Manipulations of ricochet idioms can imply a consistent tempo or create momentum changes (slowing down or speeding up). These idioms are the result of birdcall imitation, and the moment of discovery was documented in a practice recording (Audio 5.2). A selection of the discovered and developed ricochet idioms are provided in Video 5.3.

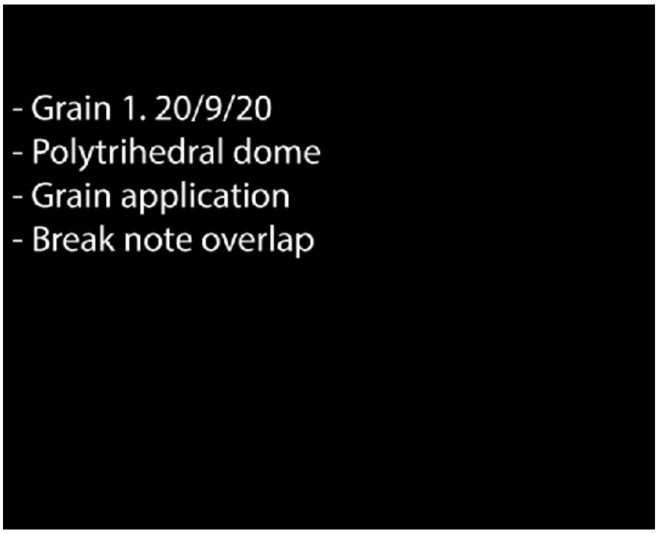
Audio 5.2. Ricochet idiom discovery through bird call imitation from a polytrihedral dome practice session on the 10/12/20.



Video 5.3. Ricochet idioms selection.

D. Grain idioms

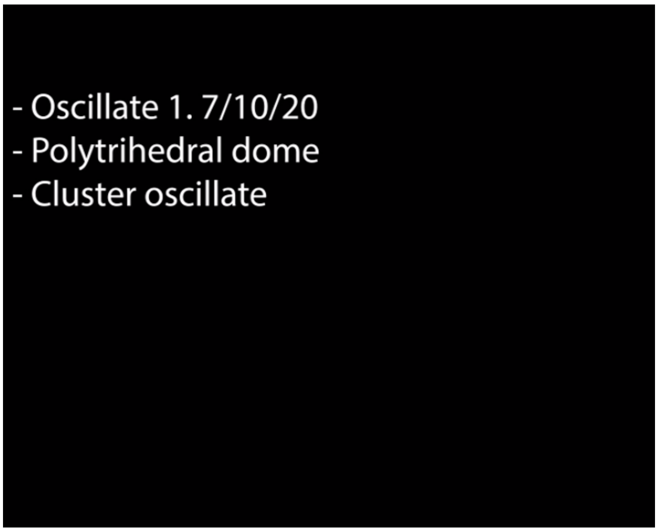
Grain idioms use both pitched and un-pitched sounds simultaneously. Bae Il Dong taught me the vocal technique used to create un-pitched sounds in singing prior to the singing experiment. The subsequent developments of this technique came from experimentation in air pressure and air direction, and imitation of non-tonal environmental sounds throughout practice sessions. The idioms can be manipulated to change the balance between pitched and unpitched sounds. See Video 5.4 for a selection of examples.



Video 5.4. Grain idioms selection.

E. Oscillate idioms

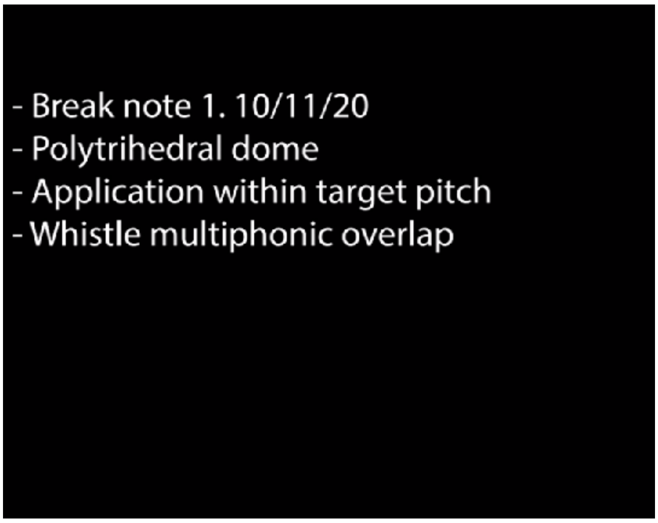
Oscillate idioms are a repeating pitch movement away from the target pitch. The secondary pitch in the oscillation is a register selection, as opposed to a secondary target pitch. Register selection is categorised into cluster (approximate intervallic relationship of 4 semitones or below), cadence (between 5 to 8 semitone intervals), or yodel (anything beyond an 8-semitone intervallic relationship). The categories are not strict but allow for spontaneous selection. The yodel register oscillation developed slowly as it requires precise coordination of air and palate. The consistency and rate of oscillation can provide momentary explicit tempo implication or momentum changes. The development of the oscillation idiom is demonstrated in Video 5.5.



Video 5.5. Oscillate idioms selection.

F. Break note idioms

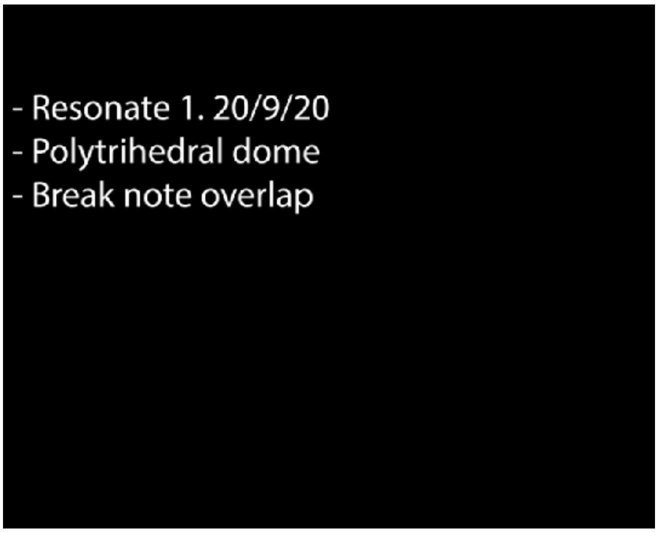
Break note idioms explore the effect of pausing the pitch target without stopping airflow. It is distinct from ricochet idioms as it is an unstable technique by design where the pitch fluctuates unpredictably. The idiom is best afforded by high register targets and is often overlapped with grain idioms or whistle multiphonic idioms. The chronological developments of break note idioms are presented in Video 5.6.



Video 5.6. Break note idioms selection.

G. Resonate idioms

Resonate idioms explore the idea of directing the voice into the resonating cavities in the head to create momentary expansions in sound. Bae Il Dong introduced this singing technique to me prior to the singing experiment and subsequent developments included manipulations of the rate of application and overlapping with other idioms. It is a technique that is more easily afforded by high registers. Video 5.7 demonstrates the resonate idioms as recorded in practice sessions.



Video 5.7. Resonate idioms selection.

H. Contour idioms

Contour idioms are the movement from one target pitch to another. Contour idioms are separated into three categories: step, slide, and roll. Step is the direct movement from one target pitch to another, where the focus is the articulation of both the ending of the first target pitch and the beginning of the next. Both slide and roll contour idioms have intermediary sounds between pitch targets however pitch is disregarded, and the focus is on movement. Slide is the incremental and consistent movement from one target pitch to another. Roll is an irregular and variable slide, where some moments can be paused or dragged through. The slide and roll contour idioms can change direction. A selection of the contour idioms is shown in Video 5.8.

Video 5.8. Contour idioms selection.

5.1.3 Organisational devices

Organisational devices provide a framework for improvised vocal performance. These devices facilitate the construction of an improvised performance that prioritises musical elements that I want to foreground. They were discovered and developed over the practice sessions with the exception of the cells (Section 5.1.3A) and dynamic range (Section 5.1.3E) devices. The organisational devices are described below.

A. Cells device

The cellular organisational approach to improvisation was introduced by Slater (2020). It is described as an approach to melodic invention based upon combining multiple cellular units and was developed in response to respiratory constraints upon performing long melodic lines. This approach revolutionised my improvised vocal performance conception and is the major organisational method in the creative work presented in this research. Within my application of this approach one cell is one vocalisation, the length of which is often what is afforded by natural breathing. It is a modular conception of melody where cells may or may not relate to the previous cell, allowing for regular changes of direction.

B. Target pitch device

Within each cell there are target pitches, either one or multiple. Target pitches are spontaneously chosen because of (i) a preferred intervallic relationship with a previous target, (ii) as a response to external sounds, (iii) and/or are within a register that affords a particular iteration of an idiom. Target pitches are connected by contour idioms.

C. Rhythm device

The rhythm device operates at different structural levels, organising the duration of vocalisations. At a macro level, the rhythmic device is applied to ensure a variety of cell duration and a variety in the duration of the break between cells. Within a cell, the rhythm device is applied for variations in the length of a target pitch and the length of an applied idiom.

D. Word sets device

The word sets device is used as an overarching framework that can span multiple cells. The sets are collections of words that I like that are connected as they would be in a typical sentence. Words sets do not relate to other sets: i.e., the sets do not combine into predetermined paragraphs or stories. The word sets were written, memorised, and used throughout the practice sessions. In performance, the words within a set are not reordered and the sets are chosen spontaneously.

E. Dynamic range device

The dynamic range device is applied as an overarching framework for performance. The singing developed over this experiment is designed to facilitate performance without any electro-acoustic support. Developing a large dynamic range was a primary goal of the singing experiment so the instrumentation of an ensemble performance would not be limited by the average sound levels. Outside of the development phase and the performance context, the dynamic range of singing is a largely unconscious choice, often calibrated in the moment to ranges that facilitate ease in ensemble performance. The resonate idiom (Section 5.1.2G) is connected to dynamic range but operates differently in that it involves more purposeful application through performance. An example of the dynamic range device application is the creative research outcome Retroreflection and waterfalls (Section 5.2) which was recorded in a professional recording studio without any electro-acoustic support, except for the electric guitar amplifier on the track Perfect age. In this recording the singing was performed in duets with a grand piano, electric guitar, and alto saxophone, with both musicians in the same room performing acoustically, that is without any electro-acoustic support.

5.2 Applied research outcomes: Retoreflection and waterfalls

Audio 5.3 – 5.6 presents the major creative output of my singing experiment. The work is a recording of one solo (accompanying my singing with cello) and three duo improvisations, the collection of the recordings is entitled Retroreflection and waterfalls. The work was recorded at Free Energy Device Studio at the conclusion of the 15-month singing experiment and is presented here as a demonstration of the creative performance application of the skills and materials developed throughout the practice period. The constraints on the improvised performances were an approximate five-minute length for each piece, the musicians were in the same room, and no use of electro-acoustic support whilst recording (with the exemption of the guitar amplifier). The recordings are presented here as a primary research output without analysis or exegesis.

The recording features duet performances with Chris Abrahams on piano, Carl Dewhurst on guitar, and Peter Farrar on alto saxophone. On track 4 Drove on I accompany myself on cello. The performance was recorded and mixed by Richard Belkner and mastered by Michael Lynch.

Audio 5.3. An improvised duet with myself and Chris Abrahams on piano entitled Pipe Dreams.

Audio 5.4. An improvised duet with myself and Carl Dewhurst on electric guitar entitled Perfect Age.

Audio 5.5. An improvised duet with myself and Peter Farrar on alto saxophone entitled Chasm.

Audio 5.6. An improvised solo piece with both voice and cello entitled Drove On.

5.3 Reflections on the singing experiment and creative outcomes

Throughout the singing experiment I was practicing at the extremes of my dynamic range. The waterfall, an acoustic environment that functions as a masking tool of the voice, served to develop inner self awareness of vocal processes and control. I could hardly hear my voice as I sung so I focussed on my body coordination. Alternatively, the polytrihedral dome was an environment within which I could monitor the sound of my own voice with extreme clarity. Discovered musical material and vocal techniques felt very different to perform in each environment, and I think this difference allowed for many developments of discovered materials.

In transferring my improvised singing from the atypical practice settings to the recording studio I noticed an enhanced flexibility in performing in different environments and musical situations. Practicing at my dynamic extremes meant that I could easily match the sound of an amplified guitar and grand piano, and easily transfer my skills to improvising at a very low dynamic level with the muted alto saxophone. My improvised singing performance was not compromised by the changing instrumentation and dynamic ranges in the series of duets.

This kind of flexibility is an important part of effectively performing improvised music. Mayas (2019) describes improvised music as a site-specific practice and a profound and ethical engagement with a situation, wherein a range of components (the performance space, the objects or devices for playing and processing, as well as the audience) together constitute a set of constantly changing circumstances and conditions. Practicing at the extremes of my range prepared me for the inherently changing and unpredictable nature of improvised music performance.

5.4 Application of research outcomes to further training with Bae Il Dong

Upon completion of the experiment, I travelled to Seoul, South Korea for 6 months of full-time training with Bae Il Dong, my singing teacher of 8 years (see singing background, Section 3.2). Bae teaches that whilst things can be explained and demonstrated, it is only action in self-regulated practice that can facilitate deep understanding and skill acquisition. The skills acquired and developed through the 15-month experiment were applied to further develop existing techniques and learn new, more advanced, vocal techniques. The process of arriving at an in-depth understanding by moving back and forth between explanation (from a teacher) and self-regulated practice has been described by Rice (1995) as a hermeneutic arc of learning.

The products of this training period are presented as a creative research outcome (Audio 5.7). Whilst there were many singing developments throughout the training period, the material that features in the creative outcome Studies in dimension (Audio 5.7) is outlined below. It can be described as both a vocal technique and a musical construct.

5.4.1 Dimension

The training period deepened my understanding of the resonate idiom (Section 5.1.2G) and introduced the concept of dimension. Bae Il Dong teaches that music is geometric space and describes singing techniques as being a means of manipulating the three-dimensional space of a sound. A musical phrase should have dimension, and this is expressed in the shape of a musical phrase. It starts at a contracted point, then should expand in three-dimensions via resonance. Finally, it contracts again to a point. The contraction and expansion are assisted by the common structure of a word or syllable: consonant, vowel, consonant. A consonant is a short, contracted sound, and a vowel is longer and can facilitate dimension. The amount of expansion and the duration of the dimension often dictates the tone colour of the voice. Large and long dimension enables idioms like break note (Section 5.1.2F), multiphonic (5.1.2A), or grain (5.1.2D) idioms. For these long and large dimension contours the voice needs to resonate in three-dimensions, using all possible resonant cavities including those in the face, mouth, throat, and chest. The singing experiment increased my capacity and access to this resonance, allowing me to start applying the dimension technique as Bae instructs.

5.5 Applied research outcomes: Study in dimension

Audio 5.7 is a creative output of the singing experiment and the subsequent training period with Bae Il Dong. The work is a recording of an improvised duet entitled Study in dimension and was recorded at the conclusion of six-months of singing training with Bae Il Dong. Similarly to Retroreflection and waterfalls, the work was recorded at Free Energy Device Studio with the constraint of the musicians being in the same room without electro-acoustic support whilst recording. The recording is presented here as a primary research output without analysis or exegesis.

The recording features Simon Barker on drums. The performance was recorded and mixed by Richard Belkner and mastered by Michael Lynch.

Audio 5.7. An improvised duet with myself and Simon Barker on drums entitled Study in dimension.

5.6 Conclusion

This chapter has detailed the musical materials that were discovered and developed over the fifteen-month singing experiment. The materials are a product of the experiment design, a framework for practice, being imposed on established self-regulated practice methods. At the end of the experiment, a recording was made titled Retroreflection and waterfalls, that demonstrated applications of the new materials. This recording included a solo piece of improvised singing and cello, and three duets: improvised singing and piano, guitar, and alto-saxophone. The acquired vocal skills were then applied to further training with p'ansori artist Bae Il Dong, as the new skills facilitated a new level of vocal training and access to his technical ability. An improvised music recording Study in dimension was used to document my voice after this training period as the final creative research contribution of this thesis. This chapter has provided the musical evidence of the effectiveness of the singing experiment.

Chapter 6: Conclusion

Acoustic environments can restrict or support speakers and singers. Room acoustics is therefore an important issue when considering the function and performance of a space. A large amount of work has previously been carried out within this field, however, the large body of published articles focuses on the point of view of the listener.

By contrast, this thesis has presented two experiments on the theme of autophony, the sound of one's own voice, considering the experience of a talker or singer. Both experiments investigate how the acoustic environment can alter the sound of one's own voice, supporting the voice, and the unconscious and systematic behavioural responses to this.

Behavioural responses to an environment can be manipulated for multiple purposes. This thesis shows two examples of this: i) how creating an environment with rich and dynamic information can induce a physical state (meta-stable performance region) where expert performance skills can be acquired and developed, and ii) how a built environment can alter the way people converse, optimising the functionality of a space and people's wellbeing.

The connection between the two experiments lies in the changes to vocal production as a result of extreme acoustic environments. Both the experiments examined the effect of alterations of autophonic sound on voice levels and spectrum. The first experiment, a talking experiment, simulated the kind of autophonic support that could be provided by a retroreflective array, comparing the speaking response with other kinds of spectrum limited voice support, and no support at all. The second experiment, the singing experiment, examined the voice support provided by a domed retroreflective array of 40 trihedral reflectors that I designed and built and was contrasted with the autophonic masking of a waterfall.

The secondary connection of these experiments lies in the trajectory of the research. This research began with my interest in retroreflection and the ways a retroreflective array could optimise a multi-talker environment through voice regulation. This investigation and the talking experiment results gave rise to the singing experiment design. Voice regulation from extreme high-passed acoustic support offered an exciting way to develop singing techniques within a learning and skill-acquisition framework as promoted by ecological dynamics and the constraints-led approach. I identified a retroreflective array as an interesting opposing acoustic environment to that of a masking waterfall. I was already thinking about practicing at waterfalls due to my technical singing training background with a *p'ansori* singing expert, and his advocacy of a waterfall as an important singing training environment. Thus, the outcomes of the talking experiment and the previous research into vocal response to autophony were applied in the design of the singing experiment to increase opportunities for creative development.

6.1 Singing experiment summary

The singing experiment can be described as autoethnography, where the experimenter is the subject in the experiment. The autophonic sound in the singing experiment had two 'conditions' as provided by two alternative singing practice settings. A retroreflective array, built for the experiment, provided autophonic sound with a high-frequency bias and levels far exceeding that of every-day rooms. The second practice setting was a high-flowing waterfall with broadband background sound, and thus autophonic masking. The singing experiment included many hours of self-regulated singing practice within the a-typical settings. Practice included improvisatory performance and material and technical development procedures.

This experiment was a success in that it led to changes in my singing voice and a new improvisatory singing style. Changes in voice level were measured at each practice setting, demonstrating voice regulation. Long-term development of the singing can be heard in the field recordings of musical idioms. Another demonstrated change in my singing voice is the increase in spectral centroid over the length of the project as measured at my ears, whilst the overall level of the singing remained broadly the same.

The creative outcomes of the singing experiment include recordings of improvised duets and a solo improvisation, entitled Retroreflection and waterfalls and Study in dimension. The recorded work was intended to demonstrate the application of the technical and creative developments which were discovered during the environmental practice project to ensemble and solo improvisation. The work represents the application of a practice framework and makes an original contribution to vocal improvisation works. It also represents an important change in my technical singing abilities.

Whilst this experiment is a detailed singular case study and is not a repeatable experiment in that specific creative outcomes could not be produced by anyone else, the practice framework offers a new way to encourage improvisatory performance activities. It is an alternative application of existing research and frameworks for learning and can be considered as a novel extension of empirical acoustic science research and the constraints-led approach to skill acquisition.

6.2 Talking experiment summary

The talking experiment measured pairs of participants who were naive to the purpose of the experiment. Autophonic sound was simulated, the participants' voices were recorded, altered by a computer, and fed back to their ears in real-time. Participants talked under ten different conditions of autophonic sound, the gain of which ranged from no sound to levels exceeding that which have been reported in everyday rooms. One of three filters was also applied to the autophonic sound: All-pass, High-pass, and a Low-pass. A conversation elicitation task was used to measure speech that was natural and had a premium on communication.

In the experiment, talkers lowered their speech level at a larger rate for the kind of voice support a retroreflective array might provide (voice support with an emphasis on high frequencies) compared to other kinds of voice support. The results suggest that a retroreflective array could increase voice support in a way that could reduce conversational speech levels in real rooms. This could be useful for acoustic design in multitalker environments, where it is of interest to maintain low voice levels. Future studies could test this directly, especially in environments with notoriously adverse acoustics, for example voice levels and distraction in an open plan office with a retroreflective ceiling.

6.3 Cross-disciplinary research

This thesis provides an example of how cross-disciplinary research can produce a range of outcomes relevant to several research disciplines. The singing experiment uses different approaches to investigate the research question: how can you optimise a context for creative development? Ecological dynamics and the constraints-led approach addresses how solutions to problems, and thus learning, can be created in the moment as you interact with task, environment, and body. Contributing to this is the research in voice regulation, which is a detailed examination into how a person reacts to an acoustic environment and the sound of their own voice as altered by the environment. It provides a prediction model for the attunement of performance with task, environment, and body. I predicted that I would regulate my singing voice in the environments, and that these regulations could be used as a learning and creative tool. The analysis demonstrated the voice regulation in practice settings and, whilst I cannot directly demonstrate an explicit connection of voice regulation and the development of new singing techniques and materials, I did successfully prepare a unique approach to improvised singing within the practice framework.

6.4 Limitations

Both experiments in this thesis had aspects that limited the research outcomes. The talking experiment did not have a viable assessment of the participants comfort in conversing in each acoustic condition. This would have been a valuable research outcome to assess what kind of autophonic sound would optimise a space for conversation. For the singing experiment, it would have been useful to start with an anechoic singing measurement to compare with the later measurement. It would also have been of interest to use a contact microphone that measured voice levels without reflections for the duration of the experiment, here it was only used in the final month.

6.5 Future research

The future trajectory of this research includes the continued development of my singing and descriptions of this process. The development of idiosyncratic vocal techniques involving the ability to texture the voice requires long-term physical changes (as in the grain idioms, Section 5.1.2D). Parallels to this can be identified in descriptions of how *p'ansori* students learn to produce the unique tonal aesthetics that characterise the *p'ansori* voice. Kim (1992) describes this process as singing until the vocal cords have bled and healed, eventually developing callouses, with the singer not reaching their prime well into their fifties. Whilst this kind of extremity of physical change in the voice does not understandably characterise my future research intentions, it does capture the long-term timeline of acquiring idiomatic vocal techniques. Continued investigations into and applications of personal, environmental, and task constraints will provide the contexts to discover uniquely individual and effective vocal techniques as has been demonstrated in this thesis.

The talking experiment indicates that retroreflective treatment in a room used by multiple talkers could optimise the acoustics by reducing speech distraction. Useful future research would involve measuring co-located talkers in real rooms with a retroreflective treatment like that created by Cabrerea et al. (2021, Section 1.2.3). An interesting case study is suggested by Hannouch et al. (2018): a retroreflective ceiling treatment in an open plan office, an acoustic environment described as adverse due to speech distraction causing loss of productivity. In the room created by Cabrera et al. (2021, Section 1.2.3) plastic square trihedral corner cube reflectors with 350 mm edge lengths cover the ceiling, this treatment could also be applied in an open-plan office. The contact microphone taped to throat, as used in the singing experiment, would provide a way to measure the voice production of conversing pairs without the problem of the spectrum dependent voice support provided by the retroreflective array being recorded by a head-mounted microphone in addition to the direct sound of the speaking voice. Subjective effort, comfort, and distraction measurements would indicate whether these kinds of rooms are advantageous from the talking-listener's perspective.

Appendix A: Practice timeline

Table A.1 Full practice timeline with the number, date, and place (‘w’ is waterfall and ‘d’ is polytrihedral dome) of practice sessions. The number of hours practiced (‘time’) and whether part of the sessions was recorded with in-ear microphones (marked with ‘r’) are included.

No.	Date	Place	Time	Rec.	No.	Date	Place	Time	Rec.	No.	Date	Place	Time	Rec.
1	1/9/20	w	1.5	r	39	17/11/20	w	2.5		77	15/1/21	w	1	
2	1/9/20	w	1		40	20/11/20	d	1.5	r	78	16/1/21	w	2.5	
3	8/9/20	w	2	r	41	20/11/20	d	1		79	17/1/21	w	2	
4	9/9/20	d	1		42	21/11/20	w	1	r	80	17/1/21	w	1	
5	9/9/20	d	1		43	21/11/20	w	1		81	20/1/21	d	1	r
6	15/9/20	d	2	r	44	22/11/20	d	3	r	82	20/1/21	d	1	
7	16/9/20	d	2		45	28/11/20	w	2		83	21/1/21	d	2	r
8	20/9/20	d	1.5	r	46	29/11/20	w	1		84	22/1/21	d	1	r
9	21/9/20	d	2		47	29/11/20	w	1		85	22/1/21	d	1.5	r
10	25/9/20	w	2		48	1/12/20	d	1	r	86	24/1/21	w	1	
11	26/9/20	w	2	r	49	1/12/20	d	1		87	24/1/21	w	1	
12	6/10/20	w	2		50	2/12/20	d	2		88	25/1/21	w	2	
13	7/10/20	d	1	r	51	3/12/20	w	2.5		89	25/1/21	w	1	
14	7/10/20	d	1		52	5/12/20	w	1.5	r	90	26/1/21	d	2.5	
15	8/10/20	w	1		53	5/12/20	w	1		91	27/1/21	d	3	
16	8/10/20	w	1.5		54	6/12/20	w	2		92	3/2/21	w	1	r
17	11/10/20	w	1	r	55	10/12/20	d	2	r	93	3/2/21	w	1	
18	11/10/20	w	1		56	11/12/20	w	1.5		94	4/2/21	w	2	r
19	12/10/20	w	2		57	12/12/20	w	2.5		95	10/2/21	w	2.5	r
20	19/10/20	d	2		58	17/12/20	d	1.5	r	96	11/2/21	d	1.5	r
21	20/10/20	d	1.5	r	59	17/12/20	d	1		97	11/2/21	d	1	
22	20/10/20	d	1		60	18/12/20	d	1.5	r	98	12/2/21	w	1	
23	21/10/20	d	1.5	r	61	18/12/20	d	1		99	12/2/21	w	1.5	
24	21/10/20	d	1		62	20/12/20	d	1.5	r	100	22/2/21	d	2	r
25	22/10/20	d	1.5	r	63	20/12/20	d	1		101	23/2/21	d	2	r
26	22/10/20	d	1		64	21/12/20	w	2		102	24/2/21	w	2	
27	2/11/20	d	2		65	6/1/21	d	3		103	24/2/21	w	1	
28	2/11/20	d	1		66	7/1/21	d	1.5		104	2/3/21	w	2	r
29	3/11/20	w	2		67	7/1/21	d	1		105	3/3/21	w	1	r
30	8/11/20	w	2		68	9/1/21	d	1		106	3/3/21	w	1	r
31	10/11/20	d	1	r	69	9/1/21	d	1.5		107	4/3/21	w	1.5	
32	10/11/20	d	1		70	10/1/21	d	2.5		108	4/3/21	w	1	
33	12/11/20	d	2		71	11/1/21	d	1		109	15/3/21	d	2	
34	13/11/20	d	1.5	r	72	11/1/21	d	1		110	16/3/21	d	2.5	
35	13/11/20	d	1		73	12/1/21	d	2		111	19/3/21	d	3	r
36	14/11/20	w	1.5	r	74	13/1/21	d	1		112	18/3/21	d	2	r
37	14/11/20	w	1		75	13/1/21	d	1		113	25/3/21	d	1	
38	16/11/20	w	2		76	15/1/21	w	1.5		114	25/3/21	d	1	

No.	Date	Place	Time	Rec.	No.	Date	Place	Time	Rec.	No.	Date	Place	Time	Rec.
115	5/4/21	w	1	r	152	9/7/21	d	1		189	22/10/21	w	2	r
116	5/4/21	w	1.5		153	20/7/21	w	1		190	23/10/21	w	2	r
117	6/4/21	d	2	r	154	20/7/21	w	1		191	29/10/21	w	2.5	r
118	7/4/21	d	2	r	155	21/7/21	w	1		192	30/10/21	w	1	r
119	11/4/21	w	2		156	21/7/21	w	1.5		193	30/10/21	w	1	r
120	12/4/21	w	2		157	9/8/21	d	1.5	r	194	2/11/21	w	2	r
121	18/4/21	w	2.5	r	158	9/8/21	d	1		195	3/11/21	w	2	r
122	20/4/21	w	2	r	159	16/8/21	w	2	r	196	4/11/21	w	2	r
123	21/4/21	w	2	r	160	17/8/21	w	2	r	197	4/11/21	w	1	r
124	24/4/21	d	1		161	17/8/21	w	1		198	5/11/21	w	1	r
125	24/4/21	d	1		162	20/8/21	d	2.5	r	199	5/11/21	w	1	r
126	25/4/21	d	1		163	21/8/21	d	1	r	200	7/11/21	w	2.5	r
127	2/5/21	w	1.5	r	164	21/8/21	d	1	r	201	8/11/21	w	2	r
128	2/5/21	w	1		165	22/8/21	d	2		202	9/11/21	w	2	r
129	4/5/21	d	2	r	166	22/8/21	d	1		203	10/11/21	w	1	r
130	5/5/21	d	2.5		167	25/8/21	d	1	r	204	10/11/21	w	1	r
131	6/5/21	d	2.5		168	25/8/21	d	1		205	12/11/21	w	2.5	r
132	11/5/21	w	2		169	26/8/21	d	2		206	13/11/21	d	2	r
133	18/5/21	d	1	r	170	2/9/21	d	2	r	207	14/11/21	d	2	r
134	18/5/21	d	1	r	171	3/9/21	d	3	r	208	14/11/21	d	1	r
135	19/5/21	d	1.5		172	4/9/21	w	2	r	209	18/11/21	d	1.5	r
136	19/5/21	d	1		173	4/9/21	w	1	r	210	18/11/21	d	1	r
137	1/6/21	w	1	r	174	5/9/21	d	2	r	211	19/11/21	d	1.5	r
138	1/6/21	w	1		175	5/9/21	d	1.5	r	212	19/11/21	d	1	r
139	2/6/21	w	2		176	10/9/21	d	2	r	213	21/11/21	d	2	r
140	10/6/21	w	1	r	177	10/9/21	d	1	r	214	21/11/21	d	1	r
141	10/6/21	d	1	r	178	19/9/21	w	2	r	215	22/11/21	d	2	r
142	20/6/21	d	1.5	r	179	19/9/21	w	1	r	216	23/11/21	d	3	r
143	20/6/21	d	1		180	20/9/21	w	2.5	r	217	24/11/21	d	2	r
144	21/6/21	d	1		181	25/9/21	w	2.5	r	218	24/11/21	d	1	r
145	21/6/21	d	1.5		182	26/9/21	w	1.5	r	219	25/11/21	d	1.5	r
146	25/6/21	w	2		183	26/9/21	w	1	r	220	25/11/21	d	1	r
147	30/6/21	d	2	r	184	27/9/21	w	1	r	221	26/11/21	d	2	r
148	2/7/21	w	2	r	185	4/10/21	d	2	r	222	26/11/21	d	1	r
149	3/7/21	d	1	r	186	5/10/21	d	2	r	223	27/11/21	d	3	r
150	3/7/21	d	1	r	187	12/10/21	w	2.5	r	224	28/11/21	d	2	r
151	9/7/21	d	1.5		188	13/10/21	w	3	r	225	30/11/21	d	3	r

Glossary of Terms

A-weighting	[Acoustics] Weighting of the frequency spectrum, commonly used in acoustical measurements to account for the sensitivity of the ear to quiet sound. A-weighted levels use the units dB(A) (McMinn, 2013).
Affordances	[Pedagogy] Action possibilities provided to the actor by the environment (Gibson, 1979).
Airborne direct autophony (L_D**)**	[Acoustics] The sound of a talker's voice that is propagated directly from the talker's mouth to their ears (Yadav, 2018).
Airborne indirect autophony (L_R**)**	[Acoustics] The sound of a talker's voice that is propagated from a talker's mouth to their ears via reflections at room boundaries (Yadav, 2018).
Amplitude	[Acoustics] The instantaneous magnitude of an oscillating quantity such as sound pressure. Amplitude can be expressed as a function, as an instantaneous value, as a peak value, or as a root mean square (rms) value. For a given impedance, rms amplitude squared is proportional to power or intensity (Cabrera, 2021; Everest & Pohlmann, 2015).
Anechoic room	[Acoustics] A room designed to suppress internal sound reflections by having internal wall surfaces with extremely high sound absorption coefficients. Anechoic rooms are also highly insulated from external noise. They are used for acoustical measurements (Cabrera, 2021; Everest & Pohlmann, 2015).
Autophony/ autophonic sound	[Acoustics] The sound of one's own voice (Yadav, 2018).
Bandpass filter	[Acoustics] A filter that attenuates signals both below and above the desired passband (Everest & Pohlmann, 2015).
Bandwidth	[Acoustics] The frequency range passed by a given device or structure typically measured in hertz (Everest & Pohlmann, 2015).
Binaural	[Acoustics] Hearing with two ears. Also, a recording and playback configuration emulating hearing with two ears (Cabrera, 2021; Everest & Pohlmann, 2015).
Bone conducted autophony	[Acoustics] The sound of a talker's voice propogated to their ears via the internal structures of the head to the cochlea (Yadav, 2018).
Calibration	[Acoustics] The process of documenting and adjusting the reading of measurement instruments with a traceable reference (Everest & Pohlmann, 2015). [Pedagogy] The process that scales perceptual judgement or action to information (Withagen & Michaels, 2007).
Constraint	[Pedagogy] The boundaries which shape the emergence of functional movement solutions (Browne et al., 2021).

Constraints-led approach	[Pedagogy] The practical application of key theoretical ideas of ecological dynamics, providing guiding principles for the design of learning environments (Roberts, Newcombe & Davids, 2019).
Correlation	[Statistics] The extent to which two sets of numbers are related. Correlation coefficient (r) range between 1 (perfect agreement) through to -1 (perfect disagreement) (Cabrera, 2001).
Cross-correlation	[Acoustics] A measurement that tracks the movements of two or more sets of time series data relative to one another (Meyers, 2002).
Decibel (dB)	[Acoustics] A unit used for a logarithmic representation of a ratio used to measure sound, calculated from $L = 10 \log (x / x_0)$, where L is the value in decibels, x is the measured value, and x_0 is a reference value (Cabrera, 2001, Everest & Pohlmann, 2015).
Dependent variable	[Statistics] A variable that is calculated from other variables (Cabrera, 2001).
Ecological dynamics	[Pedagogy] A theoretical approach explaining how the detection of information regulates actions and actions are coupled to perception of affordances (Gibson, 1979).
Factor	[Statistics] A dimension or construct which is a condensed statements of the relationships between a set of variables (Kline, 2014). One of the dimensions resulting from a factor analysis (Cabrera, 2001).
Factor Analysis	[Statistics] A dimension reduction technique with the aim to simplify complex sets of data, whereby multiple scales are reduced to a smaller number of factors (Cabrera, 2001; Kline 2014).
Fixed-effects	[Statistics] The independent variables represented by the experimental conditions (Borenstein, 2009).
Fragmentation	[Music] The division of a motif into segments to create multiple variations.
Free field	[Acoustics] An acoustic environment in which sound travels only in one direction as a plane wave free from obstructions (Cabrera, 2001).
Frequency	[Acoustics] The measure of the rapidity of alterations of a periodic signal, expressed in hertz (Everest & Pohlmann, 2015).
Hertz	[Acoustics] The unit of frequency, abbreviated Hz. The same as cycles per second (Everest & Pohlmann, 2015).
Improvised music	[Music] Spontaneously generated musical content with real-time negotiation of unfolding musical interactions (Berkowitz, 2010).
Impulse response	[Acoustics] The response of an acoustical system when excited by a short transient, electric or acoustic signal (Cabrera, 2021; Everest & Pohlmann, 2015).
Indirect autophonic sound	[Acoustics] The autophonic sound that is returned to the ears via the sound reflections at room boundaries (Yadav & Cabrera, 2017).

Intensity	[Acoustics] The variation of the energy flux produced by the acoustic perturbation per unit of a sound wave, being the product of pressure and particle velocity. It is sometimes used, less formally, to refer to amplitude or pressure squared (Lefebvre, 1999).
Just noticable difference (JND)	[Acoustics] The magnitude of change in a stimulus necessary for it to be perceived as different from another stimulus (Levine & Shefner, 1981
Level	[Acoustics] A measure of the strength of sound stated in decibels used most commonly to refer to sound pressure level (Cabrera, 2001).
Lombard effect	[Acoustics] A voice regulation phenomenon where vocal output increases in vocal as to environment noise increases (Lane & Tranel, 1971).
Meta-stable performance region	[Pedagogy] The persistence of metastability (Davids et al., 2015).
Metastability	[Pedagogy] The switching between and stabilizing of different functional coordination tendencies to achieve performance goals (Davids et al., 2015).
Oral-binaural room impulse response (OBRIR)	[Acoustics] The acoustic transfer function from mouth to ears measured in a room (Miranda Jofre et al., 2013).
P'ansori	[Music] Korean traditional epic story singing (Howard, 1991).
Phase	[Acoustics] The time relationship between two signals (Everest & Pohlmann, 2015).
Polytrihedral dome	A 1 m radius, 2V geodesic dome with an array of 40 steel trihedral retroreflectors designed and built for the singing experiment.
Practice-led research	[Music] A conceptual framework that allows a researcher to incorporate their creative practice, methods, and output into the research design (Smith & Dean, 2009).
Random-effects	[Statistics] Non-controlled sources of variance (Borenstein, 2009).
Reflection	[Acoustics] For surfaces that are larger than a wavelength of impinging sound, the sound is reflected, with the angle of incidence equalling the angle of reflection (Everest & Pohlmann, 2015).
Resonance	[Acoustics] An acoustic condition where propagating waves reinforce themselves through reflections, thereby forming standing waves (Cabrera, 2001).
Retrograde	[Music] Inversion of a motif.
Retroreflection	[Acoustics] Where an incident ray is reflected by a surface back to the source (Cabrera et al., 2018).
Retroreflective array	[Acoustics] A series of geometric shapes that reflect sound back to the source and can be used as a room treatment to increase voice support (Cabrera et al., 2018).
Reverberation	[Acoustics] The gradual decay of sound due to multiple reflections (Cabrera, 2001; Everest & Pohlmann, 2015).
Reverberation time	[Acoustics] The time required for the diffuse field sound pressure level of a room to decay by 60 dB after a sound source ceases its emission (Cabrera, 2001; Everest & Pohlmann, 2015).

Room effect	[Acoustics] Voice regulation where vocal output decreases as voice support increases (Pelegrín-García et al., 2011b).
Room gain (*G_{RG}*)	[Acoustics] The different between the total airborne autophonic sound and the direct autophonic sound (Brunskog et al., 2009).
Root mean square	[Acoustics] A value calculated by taking the square root of the mean of the square of a waveform (Cabrera, 2001).
Masking	[Acoustics] The amount (or the process) by which the threshold of audibility for one sound is altered by the presence of another sound (Everest & Pohlmann, 2015).
Self-regulated learning	[Pedagogy] A self-direction process and set of behaviours whereby learners transform their mental abilities into skills and habits through a developmental process (Zimmerman, Bonnor & Kovach, 1996).
Sequence	[Music] The transposition of melodic motives to different scale degrees or a new key centre (Slater, 2020).
Sidetone compensation	[Acoustics] The reduction of voice level as a result of increased airborne autophonic sound presented by electronic means (Lane et al., 1961).
Significant	[Statistics] Unlikely to have occurred by chance, usually based on a probability less than 0.05 (Cabrera, 2001).
Sound power level	[Acoustics] A power expressed in decibels above the standard reference level of 1 pW (Everest & Pohlmann, 2015).
Sound pressure	[Acoustics] The pressure component of the sound wave (Cabrera, 2001).
Sound pressure level	[Acoustics] A sound pressure expressed in dB above the standard sound pressure of 2×10^{-5} Pa squared (Everest & Pohlmann, 2015).
Spectral balance	[Acoustics] Ratio of higher partial (2-4 kHz) to lower partial (<2 kHz) (Smiljanic, 2021).
Spectral centroid	[Acoustics] A description of the quality of the signing voice calculated as the distribution of power as a function of frequency (Watts et al., 2006).
Spectrum	[Acoustics] The distribution of the energy of a signal with frequency (Everest & Pohlmann, 2015).
Speech-weighting	[Acoustics] Weighting of the frequency spectrum in reference to the typical direct sound of speech (Pelegrín-García et al., 2012).
Trihedron	[Acoustics] A solid figure formed by three flat faces that intersect in a point (Shaw & Courtine, 1969).
Uncertainty	[Acoustics] The inevitable range of results in repeated measurements due to variability in the source, measurement procedures, and/or sound propagation conditions (Peters, 2020).
Voice regulation	[Acoustics] The changes to vocal output due to the acoustic environment (Rapp et al., 2021).
Voice support (ST_V)	[Acoustics] The amount of autophonic support provided by a room via room reflections (Brunskog et al., 2009).

Wavelength	[Acoustics] The distance a sound wave travels in the time it takes to complete one cycle (Everest & Pohlmann, 2015).

Glossary of Abbreviations

AP	All-pass filter
Av	average
CI	confidence interval
dB	decibel
FDTD	finite-difference time-domain
Geo	geodesic dome
G_{RG}	room gain
HATS	head and torso simulator
Hemi	hemispherical dome
HP	High-pass filter
Hz	hertz
kHz	kilohertz
L	energy level
$L_{Aeq,1m}$	A-weighted equivalent continuous sound levels measured at 2 m, converted to 1 m
L_D	level of airborne direct autophonic sound
L_E	level of total airborne autophonic sound
LP	Low-pass filter
L_R	level of airborne indirect autophonic sound
OBRIR	oral-binaural room impulse response
Poly	polytrihedral dome
SB	spectral balance
SPL	sound pressure level
ST_V	voice support
SW	speech weighted
VS	nominal voice support groups (Chapter 2)

References

Adams, T.E., Holman Jones, S.L., Ellis, C., 2015, *Autoethnography*, Oxford University Press, New York.

Åhlander, V. L., Rydell, R., Löfqvist, A, 2011, Speaker's comfort in teaching environments: voice problems in Swedish teaching staff, *Journal of Voice*, 25(4), 430-440.

Anderson, A. H., Bader, M., Bard, E. G., Boyle, E., Doherty, G., Garrod, S., Weinert, R., 1991, The HCRC map task corpus, *Language and speech*, 34(4), 351-366.

Angelillo, M., Di Maio, G., Costa, G., Angelillo, N., Barillari, U., 2009, Prevalence of occupational voice disorders in teachers, *Journal of Preventive Medicine and Hygiene*, 50(1), 26-32.

Araújo, D., Davids, K., 2009, Ecological approaches to cognition and action in sport and exercise: Ask not only what you do, but where you do it, *International Journal of Sport Psychology*, 40(1), 5-38.

Astolfi, A, Carullo, A, Pavese, L, Puglisi, GE, 2015, Duration of voicing and silence periods of continuous speech in different acoustic environments, *The Journal of the Acoustical Society of America*, 137(2), 565–579.

Astolfi, A., Castellana, A., Carullo, A., Puglisi, G. E., 2018, Uncertainty of speech level parameters measured with a contact-sensor-based device and a headworn microphone, *The Journal of the Acoustical Society of America*, 143(6), EL496–EL502.

Astolfi, A., Castellana, A., Puglisi, G. E., Fugiglando, U., Carullo, A., 2019, Speech level parameters in very low and excessive reverberation measured with a contact-sensor-based device and a headworn microphone, *The Journal of the Acoustical Society of America*, 145(4), 2540-2551.

Aubanel, V., Cooke, M., Villegas, J., Lecumberri, M. L. G., 2011, Conversing in the presence of a competing conversation: effects on speech production. In *Twelfth annual conference of the International Speech Communication Association*, Florence, Italy.

Bae, I.D., 2016, *Dok-gong: hol-lo da-kka gung-geu-ge i-reu-da* [Dok-gong: Reaching the finality by studying alone], Sejong Books, Gwangjin-gu, Seoul.
배일동, 2016, 독공: 홀로 닦아 궁극에 이르다, 초판, 세종서적 (주): 최승구.

Bae I.D., 2020, *Deukeum: soliui ichiwa wonlileul kkaechyeo gung-geug-e ileuda* [Deukeum: Reaching the ultimate by realizing the principles and principles of sound], Kim Seong-sil.

배일동, 2020, 득음: 소리의 이치와 원리를 깨쳐 궁극에 이르다, 초판 1 쇄,김성실.

Baecher, L., Kung, S.C., Jewkes, A., Rosalia, C., 2013, The role of video for self-evaluation in early field experiences, *Teaching and Teacher Education*, 36.

Bandt, R, 1980, *Improvisations in acoustic chambers* [sound recording], Move Records, Carlton South, Australia.

Barker, S., Howard, P. K., 2015, *Korea and the western drumset*, Taylor & Francis Group, London.

Barrett, E., Bolt, B., 2007, *Practice as research: Approaches to creative arts enquiry*, NY: I. B. Taurus & Co, New York.

Barrett, H. H., Jacobs, S. F., 1979., Retroreflective arrays as approximate phase conjugators, *Optics letters*, 4(6), 190-192.

Bates, D., Mächler, M., Bolker, B., Walker, S., 2015., Fitting Linear Mixed-Effects Models Using lme4, *Journal of Statistical Software*, 67(1).

Beechey, T., Buchholz, J. M., Keidser, G., 2018, Measuring communication difficulty through effortful speech production during conversation, *Speech communication*, 100, 18-29.

Beechey, T., Buchholz, J. M., Keidser, G., 2019, Eliciting naturalistic conversations: A method for assessing communication ability, subjective experience, and the impacts of noise and hearing impairment, *Journal of Speech, Language, and Hearing Research*, 62(2), 470–484.

Berkowitz, A. L., 2010, *The improvising mind: cognition and creativity in the musical moment*, Oxford University Press.

Bilbao, S., 2013, Modeling of complex geometries and boundary conditions in finite difference/finite volume time domain room acoustics simulation, *IEEE Transactions on Audio, Speech, and Language Processing*, 21(7), 1524–1533.

Black, J. W., 1950, The effect of room characteristics upon vocal intensity and rate, *The Journal of the Acoustical Society of America*, 22(2), 174-176.

Black, J. W., 1954, The loudness of side-tone, *Communications Monographs*, 21(4), 301-305.

Bonneville-Roussy, A., Bouffard, T., 2015, When quantity is not enough: Disentangling the roles of practice time, self-regulation and deliberate practice in musical achievement, *Psychology of Music*, 43(5), 686–704.

Borenstein, M., 2009, *Introduction to meta-analysis*, U.K: John Wiley & Sons, Chichester.

Bottalico, P., 2018, Lombard effect, ambient noise, and willingness to spend time and money in a restaurant, *The Journal of the Acoustical Society of America*, 144(3), EL209–EL214.

Bottalico, P., Graetzer, S., Hunter, E. J., 2015, Effect of training and level of external auditory feedback on the singing voice: Volume and quality, *Journal of Voice*, 30(4), 434–442.

Bottalico, P., 2018, Lombard effect, ambient noise, and willingness to spend time and money in a restaurant, *The Journal of the Acoustical Society of America*, 144(3), EL209-EL214.

Bottalico, P., Astolfi, A., Hunter, E. J., 2017, Teachers' voicing and silence periods during continuous speech in classrooms with different reverberation times, *The Journal of the Acoustical Society of America*, 141(1), EL26-EL31.

Bottalico, P., Graetzer, S., Hunter, E. J., 2016, Effects of speech style, room acoustics, and vocal fatigue on vocal effort, *The Journal of the Acoustical Society of America*, 139(5), 2870-2879.

Bottalico, P., Graetzer, S., Hunter, E., 2015, Effect of training and level of external auditory feedback on the singing voice: Volume and quality, *Journal of Voic*e, 29.

Browne, P., Sweeting, A. J., Woods, C. T., Robertson, S., 2021, Methodological considerations for furthering the understanding of constraints in applied sports, *Sports Medicine - Open*, 7(1).

Bruineberg, J., Seifert, L., Rietveld, E., Kiverstein, J., 2021, Metastable attunement and real-life skilled behaviour, *Synthese*, 199(5-6), 12819-12842.

Brumm, H., Zollinger, S. A., 2011, The evolution of the Lombard effect: 100 years of psychoacoustic research, *Behaviour*, 148(11-13), 1173-1198.

Brunskog, J., Gade, A.C., Payá Bellester, G., Calbo, L.R., 2009, Increase in voice level and Speaker comfort in lecture rooms, *The Journal of the Acoustical Society of America*, 125(4), 2072–2082.

Brunskog, J., Gade, A. C., Bellester, G. P., Calbo, L. R., 2009, Increase in voice level and speaker comfort in lecture rooms, *The Journal of the Acoustical Society of America*, 125(4), 2072-2082.

Bullock, M., 2010, Self-idiomatic music: An introduction, *Leonardo*, 43(2), 141–144.

Burnham, K. P., Anderson, D. R., 2004, Multimodel inference: understanding AIC and BIC in model selection, *Sociological Methods and Research*, 33(2), 261–304.

Cabrera, D., 2001, *A psychoacoustical study of resonating sound art* (Doctoral dissertation), The University of Sydney.

Cabrera, D., Sato, H., Martens, W. L., Lee, D., 2009, Binaural measurement and simulation of the room acoustical response from a person's mouth to their ears, *Acoustics Australia*, 37(3).

Cabrera, D., Davis, P. J., Connolly, A., 2010, Long-term horizontal vocal directivity of opera singers: Effects of singing projection and acoustic environment, *Journal of Voice*, 25(6), e291–e303.

Cabrera, D., Holmes, J., Caldwell, H., Yadav, M., Gao, K., 2018, An unusual instance of acoustic retroreflection in architecture–Ports 1961 Shanghai flagship store façade, *Applied Acoustics*, 138, 133-146.

Cabrera, D., Yadav, M., Holmes, J., Fong, O., Caldwell, H., 2020, Incidental acoustic retroreflection from building façades: Three instances in Berkeley, Sydney and Hong Kong, *Building and Environment*, 172.

Cabrera, D., Holmes, J., Lu, S., Rapp, M., Yadav, M., Hutchison, O., 2021, Voice support from acoustically retroreflective surfaces, in *Proceedings of the Euronoise 2021*, 25–27, October, Madeira, Portugal.

Cabrera, D., Lu, S., Holmes, J., Yadav, M., 2022, Sound Reflections in Indian Stepwells: Modelling Acoustically Retroreflective Architecture, *Acoustics*, 4(1), 227-247.

Caldwell, H., 2019, *An investigation into ceiling geometries for acoustic control: spatial configurations for absorption and retroreflection* (Masters dissertation), The University of Sydney.

Candy, L., Edmonds, E., 2018, Practice-based research in the creative arts: Foundations and futures from the front line, *Leonardo*, 51(1), 63-69.

Chern, A., 2019, A reflectionless discrete perfectly matched layer, *Journal of Computational Physics*, 381, 91–109.

Cheyne, H. A., Kalgaonkar, K., Clements, M., Zurek, P., 2009, Talker-to-Listener Distance Effects on Speech Production and Perception, *The Journal of the Acoustical Society of America*, 126(4), 2052–2060.

Choi, K. H., 2016, Practice-led origami-inspired fashion design: out of the frame: flight by paper Plane, *International Journal of Fashion Design, Technology and Education*, 9(3), 210-221.

Chow, J, Davids, K, Button, C, Renshaw, I, 2016, *Nonlinear pedagogy in skill acquisition: An introduction*, NY: Routledge, New York.

Chow, J. Y., Davids, K., Button, C., Rein, R., Hristovski, R., Koh, M., 2009, Dynamics of multi-articular coordination in neurobiological systems, *Nonlinear Dynamics, Psychology, and Life Sciences*, 13(1), 27-55.

Chow, J., Davids, K., Button, C., Renshaw, I., 2016, *Nonlinear pedagogy in skill acquisition: An introduction*, NY: Routledge, New York.

Cipriano, M., Astolfi, A., Pelegrín-García, D., 2017, Combined effect of noise and room acoustics on vocal effort in simulated classrooms, *The Journal of the Acoustical Society of America*, 141(1), EL51-EL56.

Clark, J. E., 1995, On becoming skillful: Patterns and constraints, *Journal Research Quarterly for Exercise and Sport*, 66(3), 173-183.

Coleman, R. F., Hicks, D. M., 1978, Singer's compensation for varying loudness levels of musical accompaniment, in Transcripts of the seventh symposium: care of the professional voice, Part I: The scientific papers, *The Voice Foundation*, 80-91.

Collyer, S., Davis, P. J., Thorpe, C. W., Callaghan, J., 2007, Sound pressure level and spectral balance linearity and symmetry in the messa di voce of female classical singers, *The Journal of the Acoustical Society of America*, 121(3), 1728-1736.

Collyer, S., Davis, P. J., Thorpe, C. W., Callaghan, J., 2009, Fundamental frequency influences the relationship between sound pressure level and spectral balance in female classically trained singers, *The Journal of the Acoustical Society of America*, 126(1), 396-406.

Colwell, C. M., 1995, Effect of teaching setting and self-evaluation on teacher intensity behaviors, *Journal of Research in Music Education*, 43(1), 6-21.

Cooke, M., Lu, Y., 2010, Spectral and temporal changes to speech produced in the presence of energetic and informational maskers, *The Journal of the Acoustical Society of America*, 128(4), 2059–2069.

Cooke, M., King, S., Garnier, M., Aubanel, V., 2014a, The listening talker: A review of human and algorithmic context-induced modifications of speech, *Computer Speech and Language*, 28(2), 543-571.

Cooke, M., Mayo, C., Villegas, J., 2014b, The contribution of durational and spectral changes to the Lombard speech intelligibility benefit, *The Journal of the Acoustical Society of America*, 135(2), 874–883.

Cox, T.J., 2014, *The sound book: the science of the sonic wonders of the world*, First American edition, W. W. Norton & Company, New York.

Cox, T. J., d'Antonio, P., 2016, *Acoustic absorbers and diffusers: theory, design and application*, CRC Press, Boca Raton.

Crawford, F. S., 1991, Cube corner retroreflectors for sound waves, *American Journal of Physics*, 59(2), 176–177.

Cremer, L., Müller, H. A., 1982, *Principles and applications of room acoustics*, Applied Science, London.

Currie, D., Dell'Agnello, S., Delle Monache, G., 2011, A lunar laser ranging retroreflector array for the 21st century, *Acta Astronautica*, 68(7-8), 667-680.

Davids, K., Button, C., Bennett, S., 2008, Dynamics of skill acquisition: A constraints-led approach, *Human Kinetics*, USA.

Davids, K., 2010, The constraints-based approach to motor learning: implications for a non linear pedagogy in sport and physical education, in *Motor learning in practice: A constraints-led approach* (eds Renshaw, I., Davids, K., Savelsbergh, G. J. P.), NY: Routledge, New York, 3- 17.

Davids, K., Araujo, D., Seifert, L., Orth, D., 2015, Expert performance in sport: An ecological dynamics perspective. In *Handbook of sport expertise* (eds Baker, J. Farrow, D.), 273-303.

Downs, H. E., Miltenberger, R., Biedronski, J., Witherspoon, L., 2015, The effects of video self-evaluation on skill acquisition with yoga postures, *Journal of Applied Behavior Analysis*, 48(4), 930-935.

Everest, F. A., Pohlmann, K. C., 2015, *Master handbook of acoustics* (6th edition), N.Y: McGraw-Hill Education LLC, New York.

Franz, E., 2008, *Intangible Asset No. 82*. [film], In the Sprocket Productions, Melbourne.

Gade, A. C., 1989, Investigations of musicians' room acoustic conditions in concert halls, Part I. Method and l aboratory experiments, *Acustica*, 65, 193-203.

Galbrun, L., Ali, T. T., 2013, Acoustical and perceptual assessment of water sounds and their use over road traffic noise, *The Journal of the Acoustical Society of America*, 133(1), 227–237.

Gander, A. J., 2017, *Developing a Polyrhythmic Idiolect* (Doctoral dissertation), The University of Sydney.

Garnier, M., Henrich, N., Dubois, D., 2010, Influence of sound immersion and communicative interaction on the Lombard effect, *Journal of Speech, Language, and Hearing Research*, 53(3), 588–608.

Gibson, J. J., 1979, *The ecological approach to visual perception*, MA: Houghton Mifflin, Boston.

Guerin, S., Kunkle, D., 2004, Emergence of constraint in self-organizing systems, *Nonlinear Dynamics, Psychology, and Life Sciences*, 8(2), 131–146.

Hale, C., 2018, *Ritual diamonds and bass hohúp: strategies for cross-domain creative engagement with Korean traditional rhythm* (Doctoral dissertation), The University of Sydney.

Hannouch, A., Caldwell, H., Cabrera, D., Reinhardt, D., 2018, Digital fabrication for retroreflective ceiling treatment: reducing speech distraction in an open work environment, in Learning, Adapting and Prototyping (eds Fukuda, T., Huang, W., Janssen, P., Crolla, K., Alhadidi, S.), Proc. 23rd Int. *Conf. Association for Computer Aided Architectural Design Research in Asia (CAADRIA)*, China, Beijing, 335–344.

Hartmann, W.M., Colburn, H.S., Kidd, G., 2006, Perceptual acoustics in the Mapparium, *The Journal of the Acoustical Society of America*, 119(5), 3209–3209.

Hazan, V., Baker, R., 2011, Acoustic-phonetic characteristics of speech produced with communicative intent to counter adverse listening conditions, *The Journal of the Acoustical Society of America*, 130(4), 2139-2152.

Hewitt, M. P., 2001, The effects of modeling, self-evaluation, and self-listening on junior high instrumentalists' music performance and practice attitude, *Journal of Research in Music Education*, 49(4), 307-322.

Hickey, M, 2015, Learning from the experts: A study of free-improvisation pedagogues in university settings, *Journal of Research in Music Education*, 62(4), 425–445.

Hodgson, M., Rempel, R., Kennedy, S., 1999, Measurement and prediction of typical speech and background-noise levels in university classrooms during lectures, *The Journal of the Acoustical Society of America*, 105(1), 226-233.

Hodgson, M., Steininger, G., Razavi, Z., 2007, Measurement and prediction of speech and noise levels and the Lombard effect in eating establishments, *The Journal of the Acoustical Society of America*, 121(4), 2023-2033.

Howard, K., 1991, Why do it that way? Rhythmic models and motifs in Korean percussion bands, *Asian Music*, 23(1), 1–59.

Huovinen, E., Tenkanen, A., Kuusinen, V.-P., 2011, Dramaturgical and music-theoretical approaches to improvisation pedagogy, *International Journal of Music Education*, 29(1), 82–100.

Hunter, E. J., Bottalico, P., Graetzer, S., Leishman, T. W., Berardi, M. L., Eyring, N. G., …Whiting, J. K., 2015, Teachers and teaching: speech production accommodations due to changes in the acoustic environment, *Energy Procedia*, 78, 3102–3107.

Hutchins, S., Peretz, I., 2011, Perception and action in singing, in Progress in brain research (eds Green, A., Chapman, C., Kalaska, J., Lepore, F.), 191, *Elsevier*, 103–118.

Ismail, M. R., Eldaly, H., 2018, Acoustic of monolithic dome structures, *Frontiers of Architectural Research*, 7(1), 56-66.

ISO 9921:2003, *Ergonomics—Assessment of speech communication*, International Organization of Standardization, Geneva, Switzerland.

ISO 3382-1:2009, *Acoustics — Measurement of room acoustic parameters — Part 1: Performance spaces*, International Organization of Standardization, Geneva, Switzerland.

ISO 23591:2021, *Acoustic quality criteria for music rehearsal rooms and spaces*, International Organization of Standardization, Geneva, Switzerland.

ISO 26101-1:2021, *Acoustics — Test methods for the qualification of the acoustic environment— Part 1: Qualification of free-field environments*, International Organization of Standardization, Geneva, Switzerland.

Jeon, J. Y., Barron, M., 2005, Evaluation of stage acoustics in Seoul Arts Center Concert Hall by measuring stage support, *The Journal of the Acoustical Society of America*, 117(1), 232–239.

Johansen, G. G., 2018. Explorational instrumental practice: An expansive approach to the development of improvisation competence, *Psychology of Music*, 46(1), 49–65.

Johnston, P., 2013, Teaching improvisation and the pedagogical history of the Jimmy Giuffre 3, *International Journal of Music Education*, 31(4), 383–393.

Juarrero, A., 2000., Dynamics in action: Intentional behavior as a complex system, *Emergence*, 2(2), 24-57.

Junqua, J.C., Fincke, S., Field, K., 1999, The Lombard effect: a reflex to better communicate with others in noise. 1999 *IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings*, ICASSP99 (Cat. No.99CH36258), 4, 2083–2086.

Kauffman, S. A., 1995, *At home in the universe: The search for laws of self-organization and complexity*, Oxford University Press, USA.

Kim, H., Lee, B., 2007, Optimal design of retroreflection corner-cube sheets by geometric optics analysis, *Optical Engineering*, 46(9), 094002–0940014.

Kim, K., 1996, *An introduction to classical Korean literature: From Hyangga to P'ansori*, M. E. Sharpe, New York.

Kitsantas, A., Zimmerman, B. J., 1998, Self-regulation of motoric learning: A strategic cycle view, *Journal of Applied Sport Psychology*, 10(2), 220-239.

Kittimathaveenan, K., Park, M., 2021, The influence of stage acoustics on the singers' performance and perception: A pilot study, in *Audio Engineering Society Convention 151*, Audio Engineering Society.

Klatt, D. H., 1976, Linguistic uses of segmental duration in English: Acoustic and perceptual Evidence, *The Journal of the Acoustical Society of America*, 59(5), 1208-1221.

Kline, P., 2014, *An easy guide to factor analysis*, Routledge, London.

Kob, M., Kamprolf, A., Neuschaefer-Rube, C., Goldschmidt, O., Behler, G., 2006, Experimental investigations of the influence of room acoustics on the teacher's voice, *The Journal of the Acoustical Society of America*, 120, 3359-3359.

Kob, M., Kamprolf, A., Neuschaefer-Rube, C., Goldschmidt, O., Behler, G., 2006, Experimental investigations of the influence of room acoustics on the teacher's voice, *The Journal of the Acoustical Society of America*, 120, 3359-3359.

Krause, J. C., Braida, L. D., 2002, Investigating alternative forms of clear speech: The effects of speaking rate and speaking mode on intelligibility, *The Journal of the Acoustical Society of America*, 112(5), 2165-2172.

Kühne, L, 2012, *TVISÖNGUR* [sound sculpture], Seydisfjördur, Iceland.

Kühne, L., 2022, *"79,6G1*85 - Lukaskuehne.com." TVISÖNGUR Construction Seydisfjördur*, East Iceland, http://www.lukaskuehne.com/TVISONGUR-Construction.pdf (last viewed May 30, 2022).

Lacey, J., 2016, Sonic Placemaking: Three approaches and ten attributes for the creation of enduring urban sound art installations, *Organised sound*, 21(2), 147-159.

Lamberty, D., 1980, Music practice rooms, *Journal of Sound and Vibration*, 69(1), 149–155.

Lane, H., 1963, Estimates of autophonic level are not estimates of loudness, *The Journal of the Acoustical Society of America*, 35(5), 777–777.

Lane, H. L., Catania, A. C., Stevens, S. S., 1961, Voice level: Autophonic scale, perceived loudness, and effects of sidetone, *The Journal of the Acoustical Society of America*, 33(2), 160-167.

Lane, H., Tranel, B., Sisson, C., 1970, Regulation of voice communication by sensory dynamics, *The Journal of the Acoustical Society of America*, 47(2B), 618-624.

Lane, H., Tranel, B., 1971, The Lombard sign and the role of hearing in speech, *Journal of speech and hearing research*, 14(4), 677-709.

Lane, H., Wozniak, J., Matthies, M., Svirsky, M., Perkell, J., O'Connell, M., Manzella, J., 1997, Changes in sound pressure and fundamental frequency contours following changes in hearing status, *The Journal of the Acoustical Society of America*, 101(4), 2244-2252.

Lazarus, H., 1986, Prediction of verbal communication is noise—A review: Part 1., *Applied Acoustics*, 19(6), 439-464.

Lefebvre, J., 1999, Physical basis of acoustics, *In Acoustics*, 1-39, Academic Press.

Lenth R., 2020, Emmeans: Estimated marginal means, aka least-squares means, *R Package 1.4.5*, https://github.com/rvlenth/emmeans (Last viewed March 2, 2021).

Levine, M. W., Shefner, J. M., 1981, *Fundamentals of sensation and perception*, Random House, New York.

Liu, A., Tonkin, M., 2022, *Sining Ringing Tree*, from https://tonkinliu.co.uk/singing-ringing-tree (Last viewed May 31, 2022).

Lucier, A., 1990, *I am Sitting in a Room* (1980 recording) [sound recording], Lovely Music, New York.

Macnamara, B. N., Hambrick, D. Z., Oswald, F. L., 2014, Deliberate practice and performance in music, games, sports, education, and professions: A meta-analysis, *Psychological Science*, 25(8), 1608–1618.

Mayas, M., 2019, *Orchestrating timbre: unfolding processes of timbre and memory in improvisational piano performance* (Doctor dissertation), University of Gothenburg.

McLean, J., 2018, *A new way of moving: developing a solo drumset practice informed by embodied music cognition* (Doctoral dissertation), The University of Sydney.

McMahon, M., 2022, *Graded constraints: a practise methodology for solo piano improvisation* (Doctoral dissertation), The University of Sydney.

McMinn, T., 2013, A-weighting: Is it the metric you think it is?, *Acoustics*, 1-4.

McPherson, G. E., Zimmerman, B. J., 2002, Self-regulation of musical learning, In *The New Handbook of Research on Music Teaching and Learning* (eds Colwell, R., Richardson, C.), 327-348, UK: Oxford University Press, Oxford.

McPherson, G. E., Renwick, J. M., 2011, Self-regulation and mastery of musical skills, In B, *Handbook of self-regulation of learning and performance* (eds Zimmerman J., Schunk D. H.), NY: Routledge, New York.

McPherson, G. E., Nielsen, S. G., Renwick, J. M., 2013, Self-regulation interventions and the development of music expertise, In *Applications of self-regulated learning* (eds Bembenutty, H., Cleary, T. J., Kitsantas A.), 355-383, NC: IAP, Charlotte.

Meyers, R., 2002, *Encyclopedia of physical science and technology*, Academic Press, California.

Michael, D. D., Siegel, G. M., Pick Jr, H. L., 1995, Effects of distance on vocal intensity, *Journal of Speech, Language, and Hearing Research*, 38(5), 1176-1183.

Miranda Jofre, L. A., Cabrera, D. A., Yadav, M., Sygulska, A., Martens, W. L., 2013, Evaluation of stage acoustics preference for a singer using oral-binaural room impulse responses, *The Journal of the Acoustical Society of America*, 133(5), 3402–3402.

Monson, B., Hunter E., Story, B., 2012, Horizontal directivity of low- and high-frequency energy in speech and singing, *The Journal of the Acoustical Society of America*, 132(1), 433–441

Moorer J. A., 1979, About this reverberation business, *Computer Music Journal*, 3, 13–28.

Nakagawa, S., Johnson, P. C. D., Schielzeth, H., 2017, The coefficient of determination R2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded, *Journal of the Royal Society Interface*, 14(134), 20170213–20170213.

Newell, K. M., 1986, Constraints on the development of coordination, In *Motor development in children: Aspects of coordination and control* (eds M.G. Wade & H.T.A Whiting), 341-361, Martinus Nijhoff Publishers, Amsterdam.

Newell, K. M., Liu, Y.-T., Mayer-Kress, G., 2001, Time scales in motor learning and Development, *Psychological Review*, 108(1), 57.

Newman, W. I., 2012, *Continuum Mechanics in the Earth Sciences*, Cambridge University Press.

Nijs, L., Rychtáriková, M., 2011, Calculating the optimum reverberation time and absorption coefficient for good speech intelligibility in classroom design using U50, *Acta Acustica united with Acustica*, 97(1), 93-102.

Nilsen, R. B., Lu, X. J., 2004, Retroreflection technology, In *Optics and Photonics for Counterterrorism and Crime Fighting*, International Society for Optics and Photonics, 5616, 47-60.

NVIDIA, 2022, *CUDA Toolkit*, https://developer.nvidia.com/cuda-toolkit (last viewed May 30, 2022)

Olsen, W. O., 1998, Average speech levels and spectra in various speaking/listening conditions: A summary of the Pearson, Bennett, & Fidell (1977) report, *American Journal of Audiology*, 7(2), 21–25.

Orth, D., van der Kamp, J., Button, C., 2019, Learning to be adaptive as a distributed process across the coach–athlete system: situating the coach in the constraints-led approach, *Physical Education and Sport Pedagogy*, 24(2), 146-161.

Park, C. E., 2003, *Voices from the Straw Mat: Toward an Ethnography of Korean Story Singing*, University of Hawaiì Press, Honolulu.

Pearsons, K. S., Bennett, R. L., Fidell, S., 1977, *Speech levels in various noise environments*, Report no. EPA-600/1-77-025, Environmental Protection Agency, Washington DC.

Pedgley, O., 2007, Capturing and analysing own design activity, *Design Studies*, 28(5), 463-483.

Pelegrín-García, D., Brunskog, J., 2012, Speakers' comfort and voice level variation in classrooms: Laboratory research, *The Journal of the Acoustical Society of America*, 132(1), 249-260.

Pelegrín-García, D., Brunskog, J., Rasmussen, B., 2014, Speaker-oriented classroom acoustics design guidelines in the context of current regulations in European countries, *Acta Acustica united with Acustica*, 100(6), 1073-1089.

Pelegrín-García, D., Fuentes-Mendizábal, O., Brunskog, J., Jeong, C. H., 2011a, Equal autophonic level curves under different room acoustics conditions, *The Journal of the Acoustical Society of America*, 130(1), 228-238.

Pelegrín-García, D., Smits, B., Brunskog, J., Jeong, C. H., 2011b, Vocal effort with changing talker-to-listener distance in different acoustic environments, *The Journal of the Acoustical Society of America*, 129(4), 1981-1990.

Pelegrín-García, D., 2011, Comment on "Increase in voice level and speaker comfort in lecture rooms" [J. Acoust. Soc. Am. 125, 2072-2082 (2009)] (L), *The Journal of the Acoustical Society of America*, 129(3), 1161–1164.

Pelegrín-García, D., Brunskog, J., Lyberg-Åhlander, V., Löfqvist, A, 2012, Measurement and prediction of voice support and room gain in school classrooms, *The Journal of the Acoustical Society of America*, 131(1), 194-204.

Pelegrín-García, D., Brunskog, J., Rasmussen, B., 2014, Speaker-oriented classroom acoustics design guidelines in the context of current regulations in European countries, *Acta Acustica united with Acustica*, 100.

Peters, R., 2020, *Uncertainty in acoustics: measurement, prediction and assessment*, Taylor & Francis Group, Milton.

Pick Jr, H. L., Siegel, G. M., Fox, P. W., Garber, S. R., Kearney, J. K., 1989, Inhibiting the Lombard effect. *The Journal of the Acoustical Society of America*, 85(2), 894-900.

Pinder, R. A., Davids, K., Renshaw, I., 2012, Metastability and emergent performance of dynamic interceptive actions, *Journal of Science and Medicine in Sport*, 15(5), 437–43

Puglisi, G. E., Astolfi, A., Cantor Cutiva, L. C., Carullo, A., 2017, Four-day-follow-up study on the voice monitoring of primary school teachers: Relationships with conversational task and classroom acoustics, *The Journal of the Acoustical Society of America*, 141(1), 441-452.

R Core Team, 2018, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.

Rapp, M., Cabrera, D., Yadav., M., 2021, Effect of voice support level and spectrum on conversational speech, *The Journal of the Acoustical Society of America*, 150(4), 2635–2646.

Rapp, M., Cabrera, D., Lu, S., 2022, A polytrihedral dome for acoustic retroreflection, and its application to creative-arts practice-led research, *Applied Acoustics*, 195.

Reinhardt, D., Martens, W., Miranda Jofre, L., 2013, Solving acoustic performance of curved surfaces by interfacing parametric design, structural engineering and acoustical analysis, in *Open Systems:Proceedings of the 18th International Conference on Computer-Aided Architectural Design Research in Asia* (CAADRIA 2013), Singapore.

Renshaw, I., Davids, K., Phillips, E., Kerherve, H., 2011, Developing talent in athletes ascomplex neurobiological systems, in *Talent identification and development in sport: international perspectives* (eds Baker, J., Cobley, S., Schorer, J.), 64-80. Routledge, London.

Renshaw, I., Davids, K., Newcombe, D., Roberts, W., 2019, *The constraints-led approach: Principles for sports coaching and practice design*, Routledge, London.

Rindel, J. H., 1986, Attenuation of sound reflections due to diffraction, in *Proceedings of Nordic Acoustical Meeting*, 257-260.

Rindel, J. H., 2010, Verbal communication and noise in eating establishments, *Applied Acoustics*, 71(12), 1156-1161.

Roberts, W., Newcombe, D., Davids, K., 2019, Application of a constraints-led approach to pedagogy in schools: embarking on a journey to nurture physical literacy in primary physical education, *Physical Education and Sport Pedagogy*, 24(2), 162-175

Roy, N., Merrill, R. M., Thibeault, S., Parsa, R. A., Gray, S. D., Smith, E. M., 2004, Prevalence of voice disorders in teachers and the general population, *Journal of Speech Language, and Hearing Research*, 47(2), 281-93.

Schneider, J., Wagner, C., Broschat, S., 1998, Implementation of transparent sources embedded in acoustic finite-difference time-domain grids, *The Journal of the Acoustical Society of America*, 103(1), 136–142.

Schober, M. F., Clark, H. H., 1989, Understanding by addressees and overhearers, *Cognitive psychology*, 21(2), 211-232.

Schroeder, M. R., 1962, Natural sounding artificial reverberation, *Journal of the Audio Engineering Society*, 10(3), 219–223.

Seifert, L., Button, C., Davids, K., 2013, Key properties of expert movement systems in sport: An ecological dynamics perspective, *Sports Medicine*, 43(3), 167–178.

Shan, X. J., Yin, J. Y., Yu, D. L., Li, C. F., Zhao J. J., 2013, Analysis of artificial corner reflector's radar cross section: a physical optics perspective, *Arabian Journal Geosciences*, 6, 2755–2765.

Shaw, R. P., Courtine, D. F., 1969, Diffraction of a plane acoustic pulse by a free orthogonal trihedron (three-dimensional corner), *The Journal of the Acoustical Society of America*, 46(5B), 1382–1384.

Sheaffer, J., van Walstijn, M., Fazenda, B., 2014, Physical and numerical constraints in source modeling for finite difference simulation of room acoustics, *The Journal of the Acoustical Society of America*, 135(1), 251-261.

Sjölander, P., Sundberg, J., 2004, Spectrum effects of subglottal pressure variation in professional baritone singers, *The Journal of the Acoustical Society of America*, 115(3), 1270-1273.

Slater, P., 2020, *The dark pattern: Towards a constraints-led approach to jazz trumpet* (Doctoral dissertation), The University of Sydney.

Sluijter, A. M., Van Heuven, V. J., 1996, Spectral balance as an acoustic correlate of linguistic stress, *The Journal of the Acoustical society of America*, 100(4), 2471-2485.

Smiljanic, R., 2021, Clear speech perception, in *The handbook of speech perception* (eds Pardo, J., Nygaard, L., Remez, R., Pisoni, D.), Wiley, New York, 177-205.

Smith, H., Dean, R. T., 2009, *Practice-led research, research-led practice in the creative arts*, Edinburgh University Press, Edinburgh.

Somerstein, S. F., Fuller, J., Kennedy R., 1989, Retroreflector field of view properties for open and solid cube corners, In *Current Developments in Optical Engineering and Commercial Optics* (eds Fischer, R. E., Pollicove, H. M., Smith, W. J.), 7(1168), 10-11 August 1989, Society of Photo Optical, San Diego, California.

Song, S. R., Kim, Y. J., 2015, Effect of a self-evaluation method using video recording on competency in nursing skills, self-directed learning ability, and academic self-efficacy, *Journal of Korean Academy of fundamentals of Nursing*, 22(4), 416-423.

Summers, W. V., Pisoni, D. B., Bernacki, R. H., Pedlow, R. I., Stokes, M. A., 1988, Effects of noise on speech production: Acoustic and perceptual analyses, *The Journal of the Acoustical Society of America*, 84(3), 917-928.

Sundberg, J., 1990, What's so special about singers?, *Journal of Voice*, 4(2), 107–119.

Szmit, R., 2017, Geometry design and structural analysis of steel single-layer geodesic domes, in *2017 Baltic Geodetic Congress* (BGC Geomatics), 205–209, IEEE.

Teune, B., Woods, C., Sweeting, A., Inness, M., Robertson, S., 2021, The influence of environmental and task constraint interaction on skilled behaviour in Australian football, *European Journal of Sport Science*, 1–8.

Titze, I. R., Hunter, E. J., 2015, Comparison of vocal vibration-dose measures for potential-damage risk criteria, *Journal of Speech, Language, and Hearing Research*, 58(5), 1425–1439.

Tonkinson, S., 1994, The Lombard effect in choral singing, *Journal of Voice*, 8(1), 24–29.

Toyota, Y., Komoda, M., Beckmann, D., Quiquerez, M., Bergal, E., 2021, *Concert halls by Nagata Acoustics: Thirty years of acoustical design for music venues and vineyard-style auditoria*, Springer International Publishing AG.

Traunmüller, H., Eriksson, A., 2000, Acoustic effects of variation in vocal effort by men, women, and children, *The Journal of the Acoustical Society of America*, 107(6), 3438-3451.

Tsunoda, T., 1997, *Extract from field recording archive #1* [sound recording], Kawasaki, WrK.

Tuominen, H., Rämö, J., Välimäki, V., 2013, Acoustic retroreflectors for music performance monitoring, in: *Proc. Sound and Music Computing Conf. SMC2013*, Stockholm, Sweden.

Uehara, L., Button, C., Falcous, M., Davids, K., 2016, Contextualised skill acquisition research: a new framework to study the development of sport expertise, *Physical Education and Sport Pedagogy*, 21, 153-168.

V. Bohm, 2018, *freeverb_64bit*, https://github.com/v7b1/freeverb_64bit (Last viewed May 2, 2021).

Vercammen, M., 2013, Sound concentration caused by curved surfaces, *The Journal of the Acoustical Society of America*, 133, 3348.

Vilkman, E., 2000, Voice problems at work: a challenge for occupational safety and health arrangement, *Folia phoniatrica et logopaedica*, 52(1-3), 120-125.

Wang, T., Wang, W., Du, P., Geng, D., Kong, X., Gong, M., 2013, Calculation of the light intensity distribution reflected by a planar corner-cube retroreflector array with the size of centimeter and above, *Optik*, 124(22), 5307-5312.

Ward-Steinman, P. M., 2014, The vocal improviser-educator: An analysis of selected American and Australian educators' influences and pedagogical views, *International Journal of Music Education*, 32(3), 346–359.

Warren, R. M., Bashford, J. A., Lenz, P. W., 2005, Intelligibilities of 1-octave rectangular bands spanning the speech spectrum when heard separately and paired, *The Journal of the Acoustical Society of America*, 118(5), 3261–3266.

Watts, C., Barnes-Burroughs, K., Estis, J., Blanton, D., 2006, The singing power ratio as an objective measure of singing voice quality in untrained talented and nontalented singers, *Journal of Voice*, 20(1), 82–88.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., et al., 2019, Welcome to the tidyverse, *Journal of Open Source Software*, 4(43), 1686.

Williams, C. E., Stevens, K. N., 1972, Emotions and speech: Some acoustical correlates, *The Journal of the Acoustical Society of America*, 52(4B), 1238-1250.

Withagen, R., Michaels, C., 2007, Transfer of calibration between length and sweet-spot perception by dynamic touch, *Ecological Psychology*, 19(1), 1-19.

Xu, J., Buchholz, J. M., Fricke, F. R., 2005, Flat-walled multilayered anechoic linings: optimization and application, *The Journal of the Acoustical Society of America*, 118(5), 3104-3109.

Yadav M., 2011, *A system for room acoustic simulation for one's own voice* (M.Phil. thesis), University of Sydney.

Yadav, M., Cabrera, D., Martens, W. L., 2012, A system for simulating room acoustical environments for one's own voice, *Applied Acoustics*, 73(4), 409-414.

Yadav, M., Cabrera, D., 2017, Autophonic loudness of singers in simulated room acoustic environments, *Journal of Voice*, 31(3), 388-e13.

Yoo, M. S., Yoo, I. Y., Lee, H., 2010, Nursing students' self-evaluation using a video recording of foley catheterization: Effects on students' competence, communication skills, and learning motivation, *Journal of Nursing Education*, 49(7), 402-405.

Zimmerman, B. J., 2000, Attaining self-regulation: A social cognitive perspective, in *Handbook Of Self-Regulation* (eds Boekaerts, M., Pintrich, P. R., Zeidner, M.), 13-42, UK: Academic Press, Oxford.

Zimmerman, B. J., Bonner, S., Kovach, R., 1996, *Developing self-regulated learners: Beyond achievement to self-efficacy*, American Psychological Association, Washington D.C.

Photo gallery

Figure G.1. Photos of the polytrihedral dome.

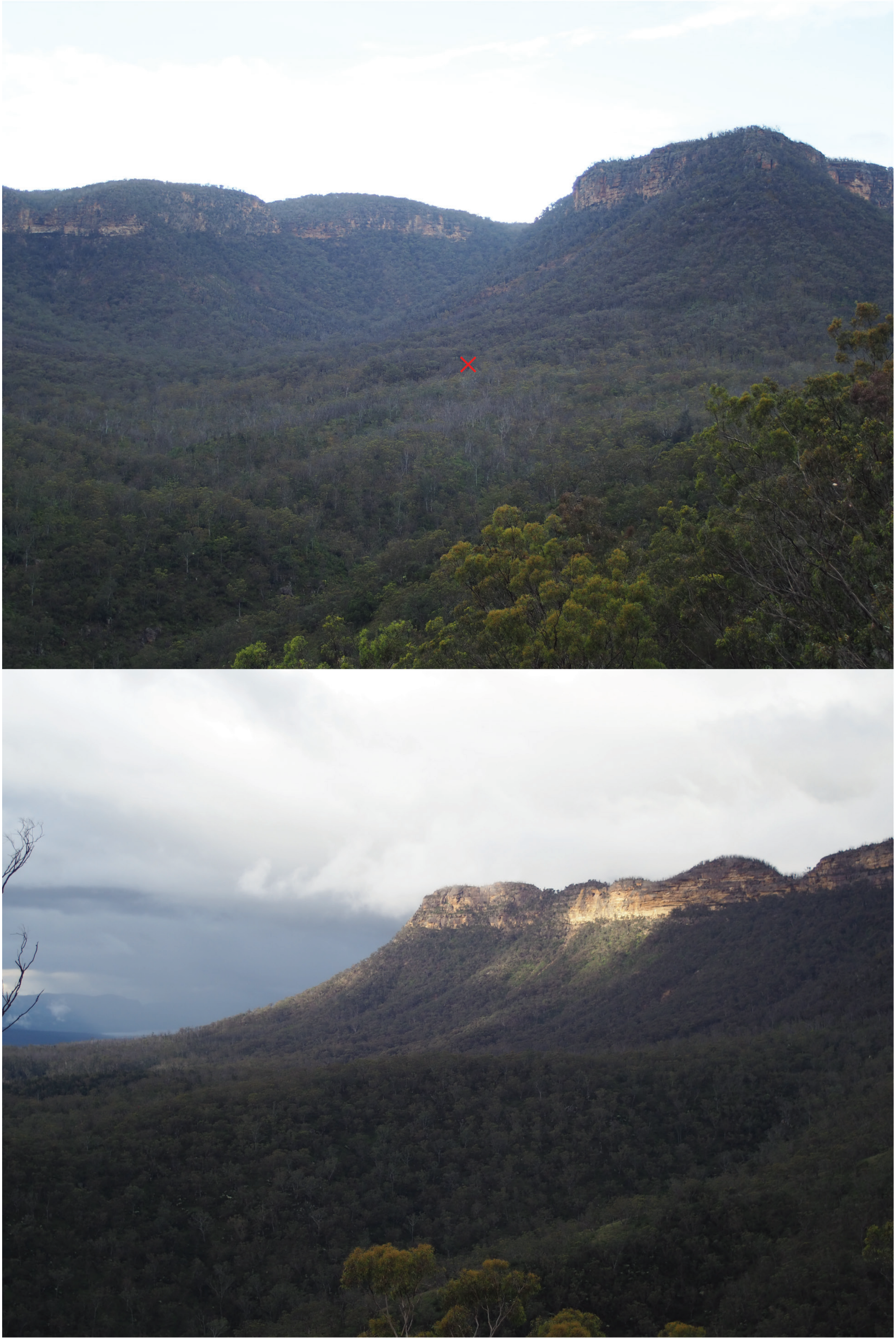


Figure G.2. Photos of the Burragorang Valley. Top: the red cross marks the approximate location of the polytrihedral dome.

Figure G.3. Top: a photo showing the method of attachment of the polytrihedral dome to the wooden posts. Middle: a photo of the polytrihedral. Bottom: a photo of myself singing in the polytrihedral dome.

Figure G.4. Photos of the polytrihedral dome in an anechoic room and with the HATS for voice support measurements (the meausurement positions with the HATS horizontal were not used in the analysis of voice support).