City Research Online

City St George's, University of London

Citation: Chadburn, L. (2023). Captions, Characters, Self-Portraits: Compositional Approaches to the Disembodied Speaking Voice and the Voice-Text-Music Relationship. (Unpublished Doctoral thesis, City, University of London)

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/30127/

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to City Research Online policy.

City Research Online: http://openaccess.city.ac.uk/ publications@citystgeorges.ac.uk

City, University of London

Doctoral Thesis

Captions, Characters, Self-Portraits: Compositional Approaches to the Disembodied Speaking Voice and the Voice-Text-Music Relationship

Author:

Leo Chadburn

Thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

City, University of London

School of Communication & Creativity

Department of Performing Arts

January 2023

Contents

Acknowledgements i
Abstract ii
List of portfolio compositions iii
List of figures viii
List of tables viii
Chapter 1: Structure of thesis and scope of the portfolio 1
1.1 Introduction to chapter 1 1
1.2 The portfolio: concert works with disembodied voices and studio works derived from vocal recordings 4
1.3 Structure of thesis 5
1.4 Scope of the use of disembodied voices in my practice / limitations of research 7
Chapter 2: Context (1): perceptions of the disembodied speaking voice 11
2.1 Introduction to chapter 2: drawing on models from extra-musical artforms and media 11
2.2 Cinematic notions of the disembodied voice: gods and ghosts 12
2.3 Radiophonic notions of the disembodied voice 17
2.4 Vox ex machina: talking computers and the vocalic body 28
2.5 Electroacoustic and sound art models: uneasy sonic materiality and authorial presence 30
2.6 Ramifications of the perception and function of cinematic, radiophonic and other acousmatic voices on my practice 35
Chapter 3: Context (2): categories of speech-music relationship 37
3.1 Introduction to chapter 3: formal-structural interactions 37
3.2 Type A: discrete text-music relationship 40
3.3 Type B: coexistent text-music relationship 44
3.3 Type C: cognate text-music relationship 47
3.4 Type D: rhythmic text-music relationship: chant or rhythmised speech 51
3.5 Type E: speech melody and speech rhythm: “speechified” music 53
3.6 Type F: speech as music 56
3.7 Type G: speech as sonic material 58
3.8 Applicability of the categories to disembodied speech 58
Chapter 4: First approach: found text (1): list texts, litany structures 60
4.1 *The Indistinguishables* (2014) 64
4.2 *Freezywater* (2016) 72
4.3 *Affix Stamp Here* (2016) 79
Chapter 5: Second approach: found text (2): radiophonic and narrative character voices and documentary texts 84

5.1 *Red and Blue* (2015) 85
5.2 *Five Loops for the Bathyscaphe* 89
5.3 *Anticlock* 94
Chapter 6: Third approach: stream of consciousness and montage: notions of the disembodied voice as self-portrait 103
6.1 *The Subject* (2020) 104
6.2 *The Object* (2020) 108
6.3 Notions of self-portraiture: authentic and manipulated portraits 111
Chapter 7: Conclusion 119
Appendix: *Red and Blue*: table of text sources, structure, voice-music relationships and treatments 123
Bibliography 130

Acknowledgements

I would like to express my deep gratitude to the following:

- The ensembles and organisations who instigated, commissioned, and funded the works comprising the portfolio of works: Apartment House, Britten Sinfonia, Quatuor Bozzini, Decibel, EXAUDI, Fondation Hoffmann, PRSF Composer's Fund and The Wigmore Hall.

- The performers whose insight and dedication brought this music to life, in both live performances and recordings, including Alissa Cheung, Anton Lukoszevieze, Bridget Carey, Caroline Dearnley, Clemens Merkel, Ed Bennett, Eliza McCarthy, Gemma Saunders, Gordon MacKay, Huw Watkin, Isabelle Bozzini, Jack McNeill, Jacqueline Shave, James Weeks, Jimmy Holliday, John Corban, Jonathan Bungard, Kerry Yong, Lucy Goddard, Mark Knoop, Martyn Sanderson, Merlin Jones, Michelle Holloway, Mira Benjamin, Neil McGovern, Paul Norman, Philip Thomas, Sebastiano Dessanay, Simon Limbrick, Simon Whiteley, Stéphanie Bozzini, Tom Williams and Xavier Lepage-Brault.

- The producers and audio engineers whose expertise and technical advice made possible the recordings of these works: Eric James at Philosophers Barn Mastering, Leafcutter John, Laurent Major, Vincent Blain and especially John Lely, who produced the majority of the tracks that make up the album *Slower / Talker*.

- Jennifer Walshe and Martin Arnold for support and discussion during the development of *The Indistinguishables*, as part of Quatour Bozzini's "Composers' Kitchen" project, and Jennet Thomas for inspiration in the development of *Red and Blue*.

- Luke Nickel for reading early drafts, and especially to Cassandra Miller for continuing support and invaluable advice on the writing.

- Claudia Molitor, for supervising this thesis.

Finally, thanks and love to Frederic Chadburn, for the generous support and patient encouragement.

Abstract

In this thesis and accompanying portfolio, I show how my compositional practice proposes the disembodied speaking voice as musical material that yields new compositional avenues.

Unlike the singing voice, the speaking voice has an ambiguous connection to music. Does it suggest, even demand new musical material and forms, which deviate from dialectical and teleological musical structures? When this voice is no longer physical, but rather disembodied, does this compound the ambiguity?

I explore three divergent approaches to selecting, generating, and structuring text for disembodied speaking voices, which unmoor the voice from the musical paradigms of literary, lyrical, and poetic texts; and from the linear narratives typically associated with music. My practice eschews these, investigating found text and lists; verbatim and paraphrased documentary material; and stream-of-consciousness free-writing; all of which might break with narrative or semantic meaning.

The deployment of these texts calls into question the function of the disembodied voice. I thus trace an evolving view of the perception of the voice as a cinematic narrator, a radiophonic character, an apparently supernatural presence, an audio caption for the accompanying music, a material sonic object, or as the reflexive presence of the composer themselves. This leads towards a notion of sonic self-portrait in my most recent work, brought about through specific kinds of uses of recordings of my own voice.

These approaches are preceded and contextualised by a framework in which I consider extra-musical models and theories from cinema and artists' film, radiophonic work and the everyday sphere of *vox ex machina* automated voices, with reference to a wide range of analogous and contrasting approaches in works by other composers and sound artists.

A second layer of contextualisation is provided by my proposal of a categorisation of types of macro- and micro-structural relationships between speaking voices and music. These range from an entirely separate, discrete presentation of voice and musical material to scenarios in which the voice alone is presented as musical material.

List of portfolio compositions

Works are listed in the order in which they are discussed in chapters 4-6.

The Indistinguishables (2014)

Instrumentation and format:

> Concert work for string quartet and pre-recorded voices.

Documentation provided:

> Score (.PDF) and sound recording (.WAV) of studio version of the work, performed by Quatour Bozzini and taken from the album *Slower / Talker* (Library of Nothing Records, 2021). Recording copyright held by Leo Chadburn.

Duration of sound recording:

> 14′49″

Performance history:

- 17 June 2014, Maison des Jeunesses Musicales du Canada, Montreal: Quatour Bozzini.
- 25 November 2014, HCMF, Huddersfield: Quatour Bozzini.
- 18 October 2016, Brunel Museum, London: Phaedra Ensemble.
- 12 March 2019, Milton Court, London: Quatour Bozzini.
- 7 March 2017, Café Oto, London: Phaedra Ensemble.
- 5 November 2019, Café Oto, London: Apartment House.

Broadcast history:

- 15 April 2017, BBC Radio 3, *Hear and Now*.
- 21 September 2019, BBC Radio 3, *The New Music Show*.

Freezywater (2016)

Instrumentation and format:

> Concert work for ensemble (seven players) and pre-recorded voices: strings (11110), piano, reed organ, hand bell (C4).

Documentation provided:

> Score (.PDF) and sound recording (.WAV) of studio version of the work, performed by Apartment House and taken from the album *Slower / Talker* (Library of Nothing Records, 2021). Recording copyright held by Leo Chadburn.

Duration of sound recording:

> 17′29″

Awards:

> Winner of the 2016 BASCA British Composer Award for Chamber Music.

Performance history:

- 27 February 2016, Wigmore Hall, London: Apartment House

Broadcast history:

- 29 October 2022, BBC Radio 3. *The New Music Show*.

Affix Stamp Here (2016)

Instrumentation and format:

> Concert work for live voices (at least 4 performers), pre-recorded voices and analogue synthesizers, with optional projections.

Documentation provided:

Score (.PDF) and sound recording (.WAV) of live performance of the work by EXAUDI. Included with the kind permission of James Weeks.

Duration of sound recording:

17′48″

Awards:

Nominated for a 2017 BASCA British Composer Award.

Performance history:

- 25 May 2016, Performance Space, City, University of London: EXAUDI
- 23 October 2016, Jerwood Hall, LSO St Luke's, London: EXAUDI

Red and Blue (2015)

Instrumentation and format:

Fixed media electronic work. Digital release only (Library of Nothing Records, 2015). Recording copyright held by Leo Chadburn.

Documentation provided:

Sound recording (.WAV).

Duration of sound recording:

16′37″

Broadcast history:

- 22/23 December 2015, BBC Radio 3, *Late Junction* (the complete work was split over two episodes).

Five Loops for the Bathyscaphe (2018)

Instrumentation and format:

> Concert work for piano trio and pre-recorded voices.

Documentation provided:

> Score (.PDF) and sound recording (.WAV) of studio version of the work, performed by Apartment House and taken from the album *Slower / Talker* (Library of Nothing Records, 2021). Recording copyright held by Leo Chadburn.

Duration of sound recording:

> 9′56″

Performance history:

- 23 January 2018, West Road Concert Hall, Cambridge: Britten Sinfonia
- 24 January 2018, Wigmore Hall, London: Britten Sinfonia
- 26 January 2018, St Andrew's Hall, Norwich: Britten Sinfonia
- 24 July 2021, Walter Hall, Toronto: Philip Chiu (piano), Jessy Je Young Kim (violin), Braden McConnell (cello)
- 13 November 2021, Jerwood Hall, LSO St Luke's: Eliza McCarthy (piano), David Alberman (violin), Laure Le Dantec (cello).

Anticlock (2019)

Instrumentation and format:

> Concert work for amplified ensemble (9 players) and pre-recorded voices: violin, recorders (one player: soprano, alto, tenor), baritone saxophone, bass clarinet, trombone, percussion (one player: 2 dog clickers, woodblock, vibraphone [two bows], marimba, large crash cymbal, large bass drum), piano, electric guitar, electric bass.

Documentation provided:

> Score (.PDF) and sound recording (.WAV) and video recording (.MP4) of live performance of the work by Decibel. Included with the kind permission of Ed Bennett.

Duration of sound recording:

> 14′49″

Performance history:

- 1 April 2019, Café Oto, London: Decibel
- 2 April 2019, The Lab, Royal Birmingham Conservatoire: Decibel

---

The Subject / The Object (2020)

Instrumentation and format:

> Fixed media electronic work. Cassette tape and digital release (Library of Nothing Records, 2020). Recording copyright held by Leo Chadburn.

Documentation provided:

> Production scores (.PDF) and sound recording (.WAV).

Duration of sound recording:

> 40′00″ (two tracks: 20′00″ each)

Broadcast history:

- 4 July 2020, BBC Radio 3. *The New Music Show*.
- 19 July 2020, BBC 6 Music, *Stuart Maconie's Freak Zone*.
- 12 August 2020, Dublin Digital Radio.
- 20 July 2020, Resonance FM.

List of figures

Figure 1: Frederic Rzewski's *Coming Together*, measures 1-6 of the score 49
Figure 2: Fundamental musical material accompanying each of the seven groups of names in *The Indistinguishables* 68
Figure 3: Example of use of approximate speech-rhythm in *The Indistinguishables* 70
Figure 4: Map of locations surrounding London named in *Freezywater* 74
Figure 5: Comparison of phrases 16 and 17 in *Freezywater* (short score) 78
Figure 6: Structure of *Affix Stamp Here*: drones and spoken intertitles 81
Figure 7: Tessituras of each section of *Five Loops for the Bathyscaphe* 91
Figure 8: Voice cues for piano gestures in the opening section of *Five Loops for the Bathyscaphe* 92
Figure 9: Example of crossfade editing on sibilant consonants in *The Subject* 110

List of tables

Table 1: Categories of speech-music relationships 39
Table 2: Sequencing of text in *The Indistinguishables* 65
Table 3: *Freezywater* location placenames, numbered in correspondence with Figure 4 74
Table 4: Section-by-section speech-music relationships and instrumental coordination in *Five Loops for the Bathyscaphe* 93
Table 5: Movement titles, text materials and voice-ensemble coordination in *Anticlock* 98
Table 6: Example of text mutation in the first line of all six iterations of the text for *The Subject* 106
Table 7: Text iterations by timestamp in *The Subject* 106

Chapter 1:
Structure of thesis and scope of the portfolio

1.1 Introduction to chapter 1

"Don't Talk": song dissolves into musical speech: a demonstration

When physical song is replaced by disembodied speech, a deep aesthetic ambiguity arises in music. Below is a personal, hypothetical situation, outlining this view.

Imagine that I am with you now and singing to you. I have chosen to perform the Beach Boys' "Don't Talk (Put Your Head on My Shoulder)" from *Pet Sounds* (1966) and I am accompanying myself at the piano. You will hear that the rhythm of the music matches the natural metrical stresses of Brian Wilson's lyrics:

˘ ˘ / ˘ / ˘ /
There are words we both could say

˘ ˘ / ˘ ˘ / ˘ ˘ / ˘
But don't talk, put your head on my shoulder

(The Beach Boys, 1966)

Each syllable of the lyric is sung to a pitch. The melody ascends to the word "say", where it lingers expressively, before tumbling down again to "shoulder". The relationship between voice, text and music is clear. They are symbiotic and commensurate in importance—it is plausible to envisage that the words and music were composed

simultaneously. The music highlights the meaning of the words, with phrases that rise and fall like a sigh, while the words articulate the melodic contour of the music.

Now imagine that I continue to play the chords of the song at the piano—those same harmonies turning through 6th, 7th and 9th chords—but instead of singing, I am speaking the words of the song, quite naturally, quite undramatically. The melody and rhythm of the vocal line are gone, and the words drift out of alignment with the accompaniment. The relationship between voice, text and music is suddenly less clear. It might seem as if I were delivering a reading of poetry, accompanied by music or, conversely, that I were narrating the music. Music and words still might seem to be illustrating each other, or they might simply be coexisting, unconnectedly.

Now imagine something else: the chords cease, and I recite only the body parts named in the song. Slowly, metronomically, I speak one word at a time:

eyes… shoulder… eyes… hand… heart… head… eyes… hand…
heart… head… shoulder… eyes… head… eyes… head… shoulder

Without harmony, melody, or even clearly defined pitch apparent, this might seem much less obviously like music. At least there is some semblance of rhythm: a formal structure in time, through the periodic delivery. Any immediately perceptible semantic meaning or narrative is now eliminated too. What is this mysterious litany expressing? If you didn't know the song, you might ask how these words were chosen and arranged, and what their source was. Is it a ritual, prayer-like, or more symbolic of the physical realm: a spoken taxonomy, analogous to a display of physical objects in a museum? Perhaps the meaning of

the list will ultimately coalesce into something resembling the original song: a concatenation of verbal images relating to the paucity of the physical senses to express love.

Finally, my spoken list of body parts continues; it is audible, but I am no longer physically present. My voice is disembodied. You are listening to a recording. How do you perceive this? In the pivotal scene of David Lynch's film *Mulholland Drive* (2001), a disembodied (sung) vocal performance is a signifier of an unsettling slippage in what is real and illusory: the singer Rebekah Del Rio collapses to the floor, yet her performance of Roy Orbison's "Crying" (translated into Spanish as "Llorando") continues after she is physically absent. Lynch's nightclub emcee (played by actor Gino Silva) announces, "*No hay banda*. There is no band. It is an illusion". Very sinister. Del Rio's vocal disembodiment suggests the supernatural; an uncanny paradox: simultaneous presence and absence.

Yet, disembodied voices are also familiar to us in utilitarian and functional settings in our everyday lives, and may not be freighted with any supernatural significance: the objective voices of Siri, Alexa, voice mail, satnav, and automated supermarket check-outs, of BBC Radio 4's *The Shipping Forecast*. Do my disembodied, spoken fragments of "Don't Talk" more closely resemble these kind of quotidian vocalisations?

Definitions of terminology: "disembodied" and "recorded" voices

As can be seen from the scenario above, in this thesis I use the terms "disembodied" and "recorded" with a particular nuance, in reflection of the formats and intentions of the work in the portfolio.

The term "**disembodied**" refers to voices which do not emanate from physical bodies: their human sources are not visible to the listener, and those human sources are indeed absent from the performance or listening experience of a work. In this thesis, I privilege "disembodied" over the term "unseen". The latter term implies an experience in which the vocal performer might, at some point, appear: for example, a vocalist obscured behind a curtain, or temporarily hidden from view. As will be seen in chapter 2, the notion of vocal disembodiment has been characterised by certain writers on the subject as uncanny, or even inherently violent, but my use of "disembodied" in this thesis is intended—and reclaimed—as a neutral term, simply describing the situation of physical absence.

"**Recorded**" refers to the means by which this disembodiment is achieved: a technological mediation. In all my works, the effect of disembodiment is enabled in this manner: there are no works in which the theatrical curtain is haunted by a live performer.

The disembodied voice is thus made possible *through* being a recorded voice. Its human source remains absent, and it is thus rendered *acousmatic*: "a situation wherein one hears a sound without seeing its cause" (Chion, 1994, 32).

1.2 The portfolio: concert works with disembodied voices and studio works derived from vocal recordings

The portfolio—which this written thesis accompanies—comprises works of mine that are unified by the use of recorded speech as a primary sonic and structural element of the music.

Five of these explore an atypical musical scenario, in that they feature disembodied, recorded voices in conjunction with live instrumental ensembles. These are *The Indistinguishables* (2014), *Freezywater* (2016), *Affix Stamp Here* (2016), *Five Loops for the Bathyscape* (2018) and *Anticlock* (2019).

Two further works exist as studio compositions (i.e., fixed media, acousmatic works, intended to be heard as audio recordings, without a live performance element): *Red and Blue* (2015) and *The Subject / The Object* (2020).

1.3 Structure of thesis

This thesis, which includes a commentary on my practice, puts these works into context with other music that explores conceptually contrasting approaches. It also considers my work's relationship to examples in other media and artforms, which exhibit analogous approaches to text and structure, some of which I have transferred to my own practice.

Chapter 2 considers the reasons for this transferral of the aesthetics of cinematic, radiophonic, and everyday occurrences of the disembodied speaking voice to my own work, with attention to how these are perceived and have been theorised. The chapter ends with a consideration of a particular kind of sonic material: the authorial presence enabled by recordings of the artist's voice.

Chapter 3 proposes a categorisation of voice-text-music relationships, when the material is the speaking voice. This includes how new musical forms are unlocked or demanded by this use of the voice.

Chapter 4 introduces my practice (the portfolio works), considering the compositional techniques and reasons for my transferal of the aesthetics and structural relationships outlined in chapters 2 and 3 to my own work. This critical commentary considers three different lines of enquiry to selecting, generating, and presenting texts, and my evolving view of how these function. Chapter 4 considers *The Indistinguishables* (a string quartet built around a list of 70 species of moth sighted in the UK), *Freezywater* (an ensemble piece built around a list of 49 topographical features around Outer London), and *Affix Stamp Here* (a vocal ensemble piece using locations and messages taken from picture postcards and UK placenames).

In these earlier works, through selecting specific kinds of textual material (found texts in the form of lists), I emphasise object-like, non-dialectical musical forms and structures, derived from visual arts models, and introduce my concept of the "caption voice".

Chapter 5: found text (2): radiophonic and narrative character voices and documentary texts. This chapter considers *Red and Blue* (a solo recording using verbatim and paraphrased correspondence between Ronald Reagan and Margaret Thatcher and their memoirs), *Five Loops for the Bathyscaphe* (built around paraphrases of the recollections of oceanographer Jacques Piccard) and *Anticlock* (which employs the recorded voice as an antagonistic "speaking metronome"). In these works, I view the recorded voice as a means of including documentary material and radiophonic characterisations, alongside original, fictive texts in non-linear narrative forms.

Chapter 6: stream-of-consciousness and montage work, which employs free writing (a technique of generating written material semi-consciously, or subconsciously)[1] and

[1] See chapter 6.1 for a more thorough explanation of this term.

other systematic processes to generate extended texts for the speaking voice, which are arranged in montage forms in the studio. The resulting work is primarily the album *The Subject / The Object*. This chapter goes on to consider how I have developed a notion of the audio self-portrait, through my use of my own voice in these later works.

1.4 Scope of the use of disembodied voices in my practice / limitations of research

The recorded, disembodied speaking voice is present in my work as elemental sonic and musical material. This acousmatic voice breaks with archetypical expectations we, as listeners, have of the voice in music, as commonly encountered as song:

The singing voice...	is replaced by...	the speaking voice.
The physical voice...	is replaced by...	the disembodied voice.
The live voice...	is replaced by...	the technologically mediated voice (recordings).

Use of this unanticipated mode of vocal delivery allows the composer to:

- Link voice and music in unfamiliar ways, building new relationships between the two.
- Introduce new musical forms and structures, including forms borrowed from other media in which the disembodied speaking voice is more familiar (such as cinema, radio, and installation art).
- Include text material in unfamiliar ways (narratively and semantically).
- Exploit the perceptual ambiguity of the disembodied voice, which might be viewed as uncanny or quotidian; diegetic or non-diegetic.

Inclusion of text

The voice in my practice is primarily a vessel for the transmission of text material. It appears as "bearer of an utterance [...] of linguistic expression" (Dolar, 2006, 14). Voice and text are coupled throughout the works in my portfolio, with the exception of the second part of *The Subject / The Object*, wherein the absence of text serves to highlight and contrast with the intentional verbosity of the first part of the piece.

My work thus excludes more abstracted uses of the speaking voice, such as non-verbal sounds, quasi-glossolalia, or invented languages, exemplified by Kurt Schwitters' *Ursonate* (1932) or the work of avant-garde and improvising vocalists such as Phil Minton, Shelley Hirsch and Diamanda Galás.[2] Nor does it focus on the "prelinguistic" and "postlinguistic" phenomena of the voice (Dolar's terminology), such as coughing, hiccups, laughing, screaming.

Similarly, while my work has utilised recordings of the voice, my focus on intelligibility has seen my work eschew heavily processed vocal recordings, as might be expected from electroacoustic composition using vocal material.

A final limitation of my research is the emerging field of AI-generated voices, and other ongoing developments of artificial, synthesised, or similarly technologically mediated

[2] Representative recordings by these artists include Phil Minton's first solo album, *A Doughnut in Both Hands* (1981), which includes "everything [...] from overtone singing of the most exotic kinds right through yodelling to childish or animalistic noises" (Cochrane, n.d.); Shelley Hirsch's, *Singing* (1987), which takes an expanded approach to the singing voice, including laughter-like sounds, glossolalic speech and exaggerated breath sounds; and Diamanda Galás's anguished response to the AIDS crisis, *Plague Mass* (1991), which is characterised by her terrifying screams and wide vocal range.

and generated voices. These are outside my current practice, and space precludes a consideration of these here.

Impact on text and form

This focus on the linguistic voice allows music—a generally abstract form when lacking the voice, with the rare exception of onomatopoeic music—to refer to the extra-musical directly.[3] Typically, this might be narrative (as in opera, musical theatre) or lyrical, literary or poetic (as in most traditions of song). My use of the disembodied speaking voice, however, is a vehicle for breaking with those text sources. Hence, narrative, lyrical, and literary texts are replaced with found text: verbatim written and spoken text from a variety of sources: technical or scientific data; non-narrative original material and free writing. An examination of my sources for these texts is a thread running through chapters 4 to 6.

Disembodied speech for non-teleological forms

An objective of the work comprising the portfolio is to exploit these alternative sources of performed texts in new forms. Strophic forms and perceptible linear, goal- or climax-oriented musical narratives are often the realm of the singing voice; in my work, I am concerned with non-linear narrative, non-teleological forms, collage, monolithic forms, and process music. The choice of texts guides these forms and musical structures, and allows me

[3] Onomatopoeic meaning music that attempts to recreate extra-musical sounds, such as Heinrich Biber's *Sonata Representativa* (c.1669), in which the violinist mimics the sounds of various birds, frogs, cats, etc., or the many composers who have utilised transcriptions of birdsong, exemplified by Olivier Messiaen. This is a separate idea to *programmatic* music, which attempts to express, through personal or cultural association, a narrative, or literary or visual idea.

to translate and borrow forms from the visual arts: installation art, artists' film and cinema, and radio.

These non-teleological forms spring from a perceived non-temporality or stasis (Pearsall, 2006). Although naturally subjective, I will demonstrate how I have derived forms from specific non-musical artworks, so that the music resembles the object-like qualities of these. For example, I demonstrate in chapter 4 how I have translated Susan Hiller's conception of curated collections of discrete, physical objects into fleeting musical figures separated by silence in *The Indistinguishables*.

Chapter 2:
Context (1): perceptions of the disembodied speaking voice

2.1 Introduction to chapter 2: drawing on models from extra-musical artforms and media

Despite the formal-structural and textual (i.e., text-source) implications of use of the disembodied voice (as discussed at the end of the previous chapter), its appearance in my practice is fundamentally personal, as an expressive, dramatic and sonic tool. This has led me to consider some of the perceptive modes of the disembodied voice; that is to say, how it functions and how it is received by the listener/audience.

This chapter provides an overview of the breadth of the field, through examples from cinema, radio, and everyday contemporary technology, besides sound art and music. It should be noted that these models are not intended to be comprehensive, and the many accompanying examples do not represent a canonical view of the field, or represent an attempt to construct a canon, or strict taxonomy.

Rather, the examples chosen are indicative of the research that has fed into my own work. The idiosyncratic breadth of style, form and genres in the examples in both this chapter (2) and the following chapter (3) therefore seek to establish a matrix of influences, inspirations and precedents against which my own practice has reacted.

The uncanny versus the quotidian quality of the disembodied voice

The perception of the disembodied voice has often been characterised by theorists of film, narrative and ventriloquism as uncanny: suggesting omnipotence or the supernatural. This view of the disembodied voice originates in the unfamiliarity and the concomitant strangeness of the novel recording technologies of one hundred years ago, and in ancient, hieratic interpretations of unseen human voices.

This theoretical view of the uncanny acousmatic voice is persistent. However, I argue that we now absorb more benign and quotidian examples of the disembodied voice in our daily lives, through radio, through automated, computerised voices. These contradictory perceptions of the disembodied speaking voice provide three key models on which I have drawn: the cinematic, the radiophonic, and the *vox ex machina*.

2.2 Cinematic notions of the disembodied voice: gods and ghosts

Steven Conner, in his studies of ventriloquism, notes that live instances of vocal disembodiment date back at least as far as ancient Greece, where it may have had a religious function (Connor, 2000). Indeed, the origin of the term *acousmatic* may be traced back to ancient Greek culture: legendarily (and possibly entirely mythically), Pythagoras lectured behind a curtain, hidden from his silent, listening disciples, the *akousmatikoi*. The *musique concrète* composer Pierre Schaeffer adopted this term (in French: *acousmatique*) and promulgated its use in the early 1960s.[4]

[4] Brian Kane, in *Sound Unseen*, writes at length about the credence lent to this story by various theorists and historians of sound and music, despite its mythic origin (Kane, 2014, 41-72).

Disembodiment as a kind of theatrical illusion or entertainment, in the form of ventriloquism, also dates from at least as early as ancient Greece,[5] and was persistent in Europe, reaching widespread popularity in the 18th and 19th centuries. It is notable that many practitioners prior to the late 19th century did not employ dummies, or similar props to personify their voices, such as the tremendously popular Joseph Askins, who appeared at Sadler's Wells Theatre in London in 1796 and 1797, performing "dialogues between himself and his *invisible* familiar, Little Tommy" (Hodgson, 2009).

The occasional appearance of *musical* disembodied voices prior to the 20th century tends to belong to the realm of the religious, or quasi-supernatural. Brian Kane describes the 16th century Milanese tradition of unseen convent choirs, instigated by Archbishop Borromeo (1538–1584), whose purpose was to suggest "angelic voices", in addition to keeping the nuns cloistered from view (Kane, 2014, 108).

Despite these early examples, it is only with the invention of technological media that facilitates the separation of voice and body that everyday occurrence and theorisation of the disembodied voice begins in earnest. The developments of the phonograph, cinema and radio explode the artistic exploitation of the disembodied voice, both in the frequency of its utilisation, but also in the number of interpretive modes in which it is perceived by artists and listeners. The uncanny perception of this voice remains, but the wide variety of modes in other media has expanded how we might interpret the presence of a disembodied voice.

[5] For a detailed examination of this, see C.B. Davis's article "Distant Ventriloquism: Vocal Mimesis, Agency and Identity in Ancient Greek Performance" (Davis, 2003). Davis notes that the earliest written explicit reference ventriloquism occurs in Aristophanes' *The Wasps* (first produced in 422 BCE), in the context of a discussion on spirit possession.

It is inevitable that we might turn firstly to models from cinema. The cinematic voiceover is easily appreciated as analogous to earlier narrative devices in literature, which might feature an omniscient or first-person narrator, whose presence might be diegetic or non-diegetic.[6] Cinematic voiceover (as opposed to synchronised onscreen dialogue) might be simply functional, as in narrated factual documentaries, or express a particular effect through its disembodiment, such as representing an internal monologue.

Despite the century-long,[7] and now totally commonplace presence of synchronised sound, cinema continues to be envisaged as primarily a visual medium. Thus, the categories of interaction between (sounding) voice and image have been described and analysed by film theorists beyond a simple binary of diegetic and non-diegetic, or extra-diegetic, but almost always from the standpoint of the relationship of sound to image.

For example, Mary Ann Doane (1980) outlines modes of interaction that hinge on a bodily presence or absence, ranging from synchronous voice, 'voice-off' (i.e., a diegetic character who is simply outside the frame), interior monologue, to voice-over in a flashback.

[6] In narratology, "diegetic" describes storytelling or sound which occurs within the world of the action or story, such as (in cinema or radio) a character's dialogue or interior monologue, as opposed to non-diegetic voiceover by an unseen narrator. Gérard Genette (1980) extrapolates this idea into a complex theory of narrative levels, using the terms "intradiegetic" (for characters within a story) and "extradiegetic" for modes of storytelling outside the action. Genette and Mieke Bal (Bal & Tavor, 1981) additionally use the terms "narrator" and "focaliser" to describe storytelling which describes and "sees" the action respectively, a concept which is reflected in the multi-perspective voices in works of mine such as *Five Loops for the Bathyscaphe* (chapter 5).

[7] Alan Crosland's 1927 feature film *The Jazz Singer* is commonly thought of as the beginning of synchronised sound on film, featuring musical numbers and Al Jonson's first spoken lines, "Wait a minute, wait a minute, you ain't heard nothin' yet." However, synchronised sound film technologies had been patented in Germany prior to this in 1919, by Joseph Engl, Hans Vogt and Joseph Massolle, and public screenings of short films with synchronised sound had taken place at the Alhambra-Lichtspiele, Kurfürstendamm, Berlin in 1922 (Crafton, 1999, 46).

The acousmêtre

In cinema, when the disembodied voice remains "never physicalised" (that is, never identified by a corresponding image at any point in the action) Michel Chion describes this as a very particular mode of the disembodied narrator: the *acousmêtre*.

> When the acousmatic presence is a voice, and especially when this voice has not yet been visualised—that is, when we cannot yet connect it to a face—we get a special being, a kind of talking and acting shadow to which we attach the name *acousmêtre*. (Chion, 1999, 21)

Chion goes on to assign a very particular character to this kind of disembodiment: "the [acousmêtre's] powers are usually malevolent, occasionally tutelary" (Chion, 1999, 23); "the powers are four: the ability to be everywhere, to see all, to know all, and to have complete power" (Chion, 1999, 24).

Chion's clearest and most significant example of this is the character of the Wizard in *The Wizard of Oz* (1939), who exhibits all these characteristics until the point in the action at which the curtain is drawn back and the voice is de-acousmatised by its placement within a physical body. The acousmêtre's power is thus imbued by its disembodiment alone and can be rescinded by a subsequent embodiment. This quasi-supernatural mode of the disembodied voice, and its reliance upon its unseen-ness recalls the ventriloquistic illusions of the ancient Greeks. Chion regards the human unease with hearing a voice disconnected from its bodily source as deep-seated and psychological; it resembles an hallucination.

These theories operate on the assumption that a viewer feels a constant desire to locate a voice in a body, to see the source from which the voice emanates. The tension

created by a disembodied voice arises from the expectation of synchronisation in the visual medium of cinema.

In musical works, however, does the voice without a physical source automatically become a kind of acousmêtre? There are certainly examples suggesting this, in operatic works well predating the medium of cinema. On the subject of disembodied operatic voice, Carolyn Abbate writes:

> What force do we associate with voices whose source is obscured, seeming without ground in physical reality? Omniscience, veracity, the power to convince, it would seem. (Abbate, 1998)

Abbate notes several examples of acousmêtre-like uses of operatic disembodied voices in works by (for example) Richard Wagner (the prophetic "Forest Bird"[8] in *Siegfried*, 1876) and Giuseppe Verdi (the *voce dal cielo* [voice from heaven] in *Don Carlos*, 1867).

Moreover, the *speaking* disembodied voice, when it occasionally appears in opera, seems destined to assume the acousmêtre's powers. In Benjamin Britten's community opera *Noyes Fludde* (written in the post-cinematic year of 1958), the role of the Voice of God is typically presented as an off-stage speaking narrator (as, perhaps, it might have been in its mediaeval mystery play source). Here, the effect is resolutely tutelary. Britten's Voice of God derives its power from the text, dramatic delivery and perhaps a cinematic perception of the disembodied voice in the mind of the audience.

[8] Abbate's translation of the role of the *Waldvogel*, more often translated into English as the "woodbird".

2.3 Radiophonic notions of the disembodied voice

Functional radiophonic voices

A cinematic perception—and Chion's widely cited notion of the acousmêtre in particular—is only one potential model for the effect of the disembodied voice in music.[9] For greater plurality of perception, we turn to the medium of radio. As listeners to radio broadcasts, we do not generally experience a "radio-acousmêtre", on account of the simple fact that is it "is a different thing since there is no possibility of seeing [the person speaking]" (Chion, 1999, 21).

This means that actors, narrators and characters in radio drama and serials have less possibility of being automatically perceived as omnipresent, omniscient, and omnipotent, despite radio being "*a fortiori*, the acousmetric [sic] medium, where the sound always appears without a corresponding image" (Weiss, 1995, 32). We are not unsettled by the disembodiment of the speaking voices of the characters in *The Archers*;[10] nor did children find the disembodied storytellers of *Listen with Mother* discomforting.[11] We accept them at face(less) value, since radiophonic voices possess an inherent credibility: they are never expected to present themselves as bodies, so the dislocated, uncanny quality that may be present in the cinema is less immediate, or nullified, or absent.

[9] Chion's bias towards a perception of the disembodied voice as acousmêtre reflects his own acousmatic work, such as *Requiem* (1973), a haunting audio collage of voices, noise, and music. Chion describes this work as "cinemascope music". Its inspiration is innately unsettling: the "troubled minority of the living, rather than the silent majority of the dead" (Chion, quoted in the liner notes for the 2007 re-release of *Requiem*).
[10] *The Archers* is the world's longest-running serial radio drama (soap opera), broadcast since 1951 on the BBC Light Programme, BBC Home Service, and currently on BBC Radio 4. Its plot concerns the interpersonal relationships of the inhabitants of a fictional village in rural England.
[11] *Listen with Mother* was a BBC children's radio programme, broadcast on the BBC Light Programme, BBC Home Service, and BBC Radio 4 between 1950 and 1982. The well-known catchphrase, with which each story began, was "Are you sitting comfortably? Then I'll begin."

Moreover, radio drama is but a small part of the output of the majority of broadcasters internationally, so we encounter a multitude of other modes of the radiophonic voice: continuity announcers and DJs, interviewers, interviewees, participants in radio phone-ins, documentary and archive material: modes that radio can move swiftly between, without provoking a sense of dislocation in the listener. The same goes for those formats that are formally derived from, or related to radio broadcast, such as podcasts and audio books.

Ambiguous perceptions of radiophonic voices: multivalent associations

However, radiophonic voices do still have the potential to generate an uncanny perception in the listener. Very occasionally, this is extreme, such as the case of Orson Welles' infamous 1938 radio production of *The War of the Worlds*, which provoked panic when a small minority of listeners mistook the dramatised news reports of a Martian invasion for reality. Radio possesses the malleability to move swiftly between modes of speaking voice, from described place-to-place, time-to-time, or identity-to-identity in a disorienting manner. In Welles' case, this also included moving between the modes of veracious news report and fictional drama, in a manner that was apparently credible.[12]

The medium also holds potential for *obscuring* identity, place or time. For example, in the radio work of Samuel Beckett, narrative is often disrupted through ellipsis. In *Rough for Radio I* (1961), the identities of the speaking characters are ambiguous, and we hear

[12] This infamous moment in broadcasting history, apparently demonstrating radio's calamitous credibility, is widely described as provoking mass hysteria, but research indicates that the listenership was actually small, and the number of listeners misled by the programme was minimal. Subsequent reports of the incident are considered to have exaggerated the public reaction (Schwartz, 2015).

single sides of telephone calls (Esslin, 1977); in *Words and Music* (1961), both the identities of the speakers and location are ambiguous. In the words of Kim Conner:

> When we go to a stage play, we develop our understanding of it from a comprehension of the interaction between what we see and what we hear [...] we develop an image, more precisely, a visual memory of them. In contrast, and obviously, when we listen to a radio play, we develop our understanding of it solely from what we hear [...]. Therefore, the disembodied voice becomes a remarkably flexible, if inconstant, agent for the communication of identity.
>
> Radio drama offers Beckett a form that, on one hand, is intimate, direct, and expansively imaginative [...] With radio, [Beckett] gains full access to the multivalent associations and possibilities of the human voice, and is able to explore the relationship between that voice and identity, [...] that voice in time and space, [...] and that voice and materiality. (Conner, 1997, 303)

These "multivalent associations" of the disembodied radio voice produce deep perceptual ambiguities, not just in terms of narrative, but also in the emotional effect on the listener. Two apparently contradictory examples are the Maritime and Coastguard agency's *Shipping Forecast* (broadcast four times daily on BBC Radio 4) and the so-called "Numbers Stations" prevalent on shortwave radio bands since the Cold War period. Both are somewhat similar, in their delivery of lists of information generally not understood by a casual listener, but provoke widely diverging emotional responses. The *Shipping Forecast* is a technical description and forecast of the weather in the sea areas surrounding the British Isles, read as a systematic list (clockwise around the coast). It has been widely described as having comforting qualities by aficionados, such as the writer Carol Ann Duffy, who referred to it as "the radio's prayer" in her poem "Prayer" (Duffy, 2013), and its broadcasters, such as Zeb Soanes, who described it as "Dependable, reassuring and never hurried [...] a still small voice of calm across the airwaves" (Conner, 2020, 7-8) and Kathy Clugson, who described it as

"like a lullaby almost [...] There's something about the sound of it and the rhythm—it's so repetitive—that is so soothing" (quoted in Hudson, 2012).

In contrast, the Numbers Stations are broadcasts of what is assumed to be encoded messages for espionage purposes, transmitting lists of numbers or phonetic letters from several international locations. These have built a similarly wide and enthusiastic casual listenership,[13] but the fascination is due to the perceived unsettling and enigmatic qualities of the broadcasts: "There is a shroud of mystery surrounding Numbers Stations, a disorientating feel [...] dark and sinister" (van Peer, 2000). Accordingly, the musical works inspired by Numbers Stations adopt a dark tone, such as Yannis Kyriakides' *a conSPIracy cantata* (1999), or Boards of Canada's track "Gyroscope" (from the album *Geogaddi*, 2002). Clearly, the cultural associations of the Numbers Stations affect our emotional perception of them, but the general sense of unease generated by these particular disembodied voices reading a litany of text could not be further from the sense of solace provided by the disembodied voices speaking the litany of the *Shipping Forecast*.

Vocal montage: speaking voices as radiophonic music

Conner's multivalent associations of the voice and vocal materiality also reflect examples of experimental radio broadcast that exploit the sonic material of disembodied speaking voices in montage forms, which move beyond linguistic meaning in works that are

[13] For a range of articles and resources relating to Numbers Stations, see https://www.numbers-stations.com/. The stations received wider public attention following the release of a four-CD compilation album of recordings in 1997 on Irdial-Discs, *The Conet Project: Recordings of Shortwave Numbers Stations*. These recordings were subsequently reused as samples by a number of musicians.

sonic in emphasis. The examples below form important precedential examples to my work *The Subject / The Object*, as considered in chapter 6.

Glenn Gould's broadcast work *The Idea of North* (1967)[14] uses overlapping edits of interviewees' voices, which become sonic material in a manner that is both radiophonic and inherently musical. Gould refers to his use of overlaid, simultaneous speaking voices here as "contrapuntal radio" (Gould, 1984, 393). A similar approach is taken by the collage artist Vicki Bennett (also known by the pseudonym People Like Us) in works intended both for broadcast and gallery exhibition as fixed media audio works. In Bennett's six-channel sound installation *First Person* (2020), more than seventy-five different speaking voices (recorded independently by the speakers themselves) are heard reciting fragments of text drawn from online user profiles.[15] Bennett's edit of the source recordings switches from voice to voice moment by moment, creating a kaleidoscopic effect of "a chorus of people [...] a symphony of voices" (Bennett, 2020).

It is notable that both Bennett and Gould use musical terms to describe their work ("contrapuntal", "chorus", "symphony"), despite the only sonic material being spoken word, which is presented without manipulation beyond a meticulous edit, and the medium being radio and installation. Indeed, Gould noted that his montages and collages of speaking voices employed "a number of techniques which I would be inclined to identify as musically derived" (Gould, 1984, 393). The musicality of the speaking voice is inherent here: in chapter

[14] *The Idea of North* is the most well-known of the three radiophonic works realised by Gould for CBC (Canadian Broadcast Corporation), which were later collected for release under the title *The Solitude Trilogy*. The other two works are *The Latecomers* (1969) and *The Quiet in the Land* (1977). All three use the voices of people living in isolated Canadian communities (in Northern Canada, the Newfoundland outports and the Russian Mennonite community of Winnipeg, respectively).

[15] One of the voices in this work is my own. By coincidence, I was working on the final edit of *The Subject / The Object* when Bennett put out an online call for vocal contributions to *First Person*. Both finished works utilise recordings of speaking voices, but their text, sound, atmosphere and media are highly differentiated.

3.6 I will explore further examples of this approach from the perspective of the voice-music structural relationship. It may well be, for the listener, that the density of the material in both Bennett and Gould's work causes attention to drift from the linguistic utterances and become purely sonic in perception. The meaning of the words and phrases dissolves; what is left is the pure sound of the voice: the phonological aspects of intonation, pitch, and vowel and consonant sounds. When utilising a chorus of voices, even these elements may become less important to the perception of the work; idiolectal features are subsumed in a general impression of the phonology of the human voice.

> Few opera composers have been deterred from utilizing trios, quartets, or quintets by the knowledge that only a portion of the words they set to music will be accessible to the listener—most composers being concerned primarily about the totality of the structure, the play of consonance and dissonance between the voices [...] I would like to think that these scenes can be listened to in very much the same way that you'd attend the [fugue at the end of Giuseppe Verdi's opera] *Falstaff*. (Gould, 1984, 393)

An example of an artist using a single voice in this manner, rather than a multitudinous "chorus" occurs in Robert Ashley's fixed media audio work *In Sara, Mencken, Christ and Beethoven There Were Men and Women* (1974). Here, Ashley himself speaks the entire text of an experimental book-length poem by outsider artist John Barton Wolgamot. Ashley's work presents his recitation with every pause edited out, so we hear a constant stream of his voice.[16] The poem's cyclical structures, with many repetitions of phrase, means that the listener's focus is likely to drift away from the text, to the sonic characteristics (the phonology) of Ashley's voice. Again, this deemphasising of the linguistic function of the

[16] This "wall of voice" effect is usually contingent on technological mediation, to achieve a genuinely pause-less, or overlapping effect, but there are examples of something approaching it in theatrical works by Samuel Beckett, such as *Not I* (1972), with its almost continuous high-speed monologue, or the character of Lucky's speech in Act 1 of *Waiting for Godot* (1953), and in the dance-theatre performances of Nigel Charnock, such as his hyperactive improvised monologues in DV8 Physical Theatre's *Strange Fish* (Hinton, 1992).

speaking voice, in favour of the sonic impression, is a key precedent to my work *The Subject* (chapter 6).

Composers engaging with radiophonic forms: *Neues Hörspiel* and other examples

A comprehensive examination of music written for the radio is obviously outside the scope of this thesis, since incidental music for radio dramas (e.g., introductory music for radio programmes, jingles and sound design) has existed since the inception of the medium and this work does not usually directly concern itself with a structural relationship to the disembodied voice, or use the disembodied voice as sonic material.

However, there are a number of composers who have worked within a radiophonic framework (producing works intended for broadcast), or who have otherwise adopted a radiophonic approach.

A particularly fertile milieu for composers working within radio was 1960s to 1980s Germany, when the development of the so-called *Neues Hörspiel* movement of radio drama allowed composers to engage with the form on a more experimental basis (Cory & Haggh, 1981), producing works that often employ a montage of material, either similar to the approach described above, or featuring multiple voices in non-narrative forms, alongside music and noise. This kind of work is especially associated with the broadcasts commissioned by Klaus Schöning, writer, director, producer at Westdeutsche Rundfunk Köln

(WDR: *West German Radio, Cologne*), and theorist of this genre of broadcast.[17] These works provide a direct model for works of mine such as *Red and Blue* (see chapter 5.1), which use a similar montage of voices, sound and music.

Mauricio Kagel was a particularly prolific writer of neue Hörspiele during this period, bringing a composerly approach to radio drama in works such as *(Hörspiel) Ein Aufnahmezustand I, II, III* (1969), *Rrrrrrr: Radio Play on a Radio Fantasy* (1982) and *Der Tribun* (1979). Akin to my own practice, many of these works by Kagel combine his original music with non-linear, or non-narrative spoken texts, often delivered in the composer's own voice, such as in *Der Tribun*, in which Kagel himself portrays an incoherently ranting dictatorial politician, whose words interrupt, or are interrupted by Kagel's march-music score and crowd noise sound effects.

Neues Hörspiel works (predominantly in German) have remained relatively little-known to the English-speaking world, and so their direct influence is limited. An English language example[18] is John Cage's *Roaratorio* (1979), which was produced by Schöning not for WDR but for IRCAM (*Institut de recherche et coordination acoustique/musique*) in Paris (Perloff, 2004, 133). Composers taking an experimental approach to the radio form have been less prevalent elsewhere, but include, for example, certain works produced under the auspices of the BBC Radiophonic Workshop in London. Delia Derbyshire's *The Dreams* (1964) is one

[17] The term *Hörspiel* simply refers to a radio play, or radio drama of any kind. Schöning promulgated the term *Neues Hörspiel* to refer to the more experimental works he was writing and producing, in which spoken word, music and noise are more freely deployed, in non-narrative and non-linear narrative forms. Schöning's published writings use the term *Neues Hörspiel* widely to denote this kind of work (e.g. Schöning, 1969). Subsequent non-German language artists using the term *Hörspiel*, such as Momus & Laplantine (2003) and myself (Bookish, 2007) inevitably do so to denote a more experimental, musical approach. The nature of this experimental approach by Schöning was to consider radio as residing "closer to the contemporary genres of open theatre, art film, postmodern dance, happenings and performance art" (Frost, 1991, 284).

[18] This might be more accurately described as a *predominantly* English language example, since *Roaratorio* is based on James Joyce's *Finnegans Wake* (1939), which is characterised by its idiosyncratic contortions of the language, full of word play, neologisms and multilingual puns.

such work, produced in collaboration with the writer Barry Bermange, as part of their series of "Inventions for Radio" for broadcast on the BBC Third Programme. Here, again, Derbyshire takes a composerly approach to the material: a meticulously edited collage of interviewees' speaking voices, describing their dreams in five categories of imagery (running, falling, land, sea, and colour), combined with Derbyshire's slowly unfolding musique concrète and synthesised drones. A more recent example is Gavin Bryars and Juan Muñoz's *A Man in a Room Gambling* (1992), a series of ten pieces intended for broadcast, each lasting five minutes (the format and structure of the entire work dictated by the confines of the broadcast time available), in which Bryars' music for string quartet accompanies Muñoz's spoken descriptions of sleight of hand tricks and methods for cheating at card games. The pieces were broadcast on BBC Radio 4 in September 1997, without introduction from continuity announcers or credits, to deliberately engage with the same ambiguous listener perception of broadcasts such *as The Shipping Forecast* (see chapter 2.4). According to Bryars: "Juan [Muñoz] imagined a listener driving along a motorway at night being bemused by this fleeting and perhaps enigmatic curiosity, in fact precisely the way in which most listeners encounter *The Shipping Forecast*" (Bryars, 1995).

The works described above are radiophonic not just in their approach to voice, text, and structure, but also in their medium, having been written explicitly for broadcast. In Momus and Anne Laplantine's *Summerisle Hörspiel* (2003), a radiophonic approach is clear from the title, but the work was distributed online. Here, the composers interleave fragments of original spoken word and song, alongside electronic music, to create an impressionistic, non-narrative work inspired by the film *The Wicker Man* (1973). An earlier work of my own, the album *Trainwreck / Raincheck* (2007, released under my pseudonym Simon Bookish)

takes a similarly radiophonic approach, in a collage structure of spoken word and song, based on dream transcriptions. However, despite the initial conception of the work as a broadcast Hörspiel, it subsequently evolved into a live performance work, so is not a direct precursor of my research in works such as *Red and Blue*, as described in chapter 5.1.

Adoption of these radiophonic modes allows an extension of the text material in musical works to include documentary material. This is an important aspect of my own works *Red and Blue* and *Five Loops for the Bathyscape* (see chapter 5), which both include verbatim material from the historical figures who provide the music's subject matter. Use of documentary material, as recorded voices, in recent experimental music and sonic art has fallen into two categories:

1. "Real" documentary voices

These are works that use archive recordings, or primary source interview material as sonic and textual material. Notably, Steve Reich has used both, in work such as *Different Trains* (which uses archive interview material) and *The Cave* (which uses interviews conducted by Reich himself as material). More recently, Sarah Hennies' *Contralto* (2017) took as its source material interviews (again, conducted by the composer) with a group of trans women about their personal experience of vocal therapies, alongside demonstrations of vocal exercises. In both Reich and Hennies' case, the interview material is treated, through fragmentation and repetition, in a way that is obviously musical, while maintaining a link to the modes of documentary radio (in the case of *Different Trains*) or documentary cinema (in the case of *Contralto* and *The Cave*, both of which use video recordings).

2. Narrated documentary voices

In this case, the documentary material is presented as narration, rather than using primary source material. This is the case in Janet Cardiff's *The Missing Voice (Case Study B)* (1999), in which the artist's voice narrates a soundwalk around the streets of London, and Hildegard Westerkamp's *Kits Beach Soundwalk* (1989), in which the composer's voice narrates a field recording. In both these works, the artist appears as disembodied narrator-observer. A manifestly musical example of this is Philip Venables' violin concerto *Venables plays Bartók* (2018), a work that includes the composer's recorded voice as narrator of both the process of writing the work and his own personal history of meeting the violinist Rudolf Botta as a teenager (the voice of Botta also appears, as a character played by a voice actor). Here, the radiophonic quality of the work is explicit: Venables describes it as a "radio music drama" (Venables, 2018).

This narrator voice enables a blurring of fact and fiction; documentary and drama: in the manner of a voice actor, the character they inhabit may be naturalistic, or include fictional elements, or be a heightened version of their persona. Indeed, it is notable that all three of examples above feature the artist/composer themselves, signifying an authorial presence that will be explored further at the end of this chapter. The blurred factual, fictional and autobiographical tenor is an aspect of the dramaturgy of *The Missing Voice*, in which Cardiff's narrator superimposes a mysterious, quasi-cinematic plot over her soundwalk recordings. This ambiguous narrator role finds an archetype in the work of Laurie Anderson, whose work frequently includes her own voice *in character*, in songs such as "From the Air" (from the album *Big Science*, 1982) in which Anderson "plays" an airline pilot: "Good evening. This is your Captain. We are about to attempt a crash landing." In other works,

Anderson's narrating voice is autobiographical, or strays between fact and fiction. As early as 1977, Anderson reflected on her use of narratorial identity in this way:

> I used myself as subject matter pretty much—personal anecdotes, things that had happened to me. Then I found that I had ran out of them, and the second thing was I found I had two pasts, one was what happened and the second was what I had said about what happened. I'm sure you've had that experience, of not being able to tell the difference. (Anderson, quoted in Reed, 2022, 5)

This slippage between reality and fantasy informs both *Red and Blue* and *Five Loops for the Bathyscaphe* (chapter 5) In both these works, my own voice is present as narrator, but the spoken text moves freely between verbatim material (i.e., my voice "in character"), reflective material, or my own original commentaries on this material (i.e., my voice as third person omniscient narrator) and fictive original material.

2.4 Vox ex machina: talking computers and the vocalic body

Personification

Beyond cinematic and radiophonic modes of disembodied voices, I propose an additional mode (and concomitant perception) that feeds into my practice: one that is constantly with us in our everyday lives in the 21st century: that of talking machines and computers, in which a disembodied voice is given body by an object and personifies it. These are functional voices, programmed as audio assistants, such as Siri, automated supermarket check-outs, GPS systems and automated telephone messages. These voices may also act as

accessibility aids (akin to audio description for television and cinema). Their utterances are generally non-narrative: captions that are short phrases, or single words.

We interpret these as literal voices of the machine: *vox ex machina*. The voice appears to emanate from the machine itself, in a ventriloquistic manner: “For voice is not simply an emission of the body; it is also the imaginary production of a secondary body, a body double: a ‘voice-body’” (Connor 2004, 158).

Connor goes on to describe this effect as the “vocalic body”, in which the disembodied voice appears to inhabit a surrogate physical presence. In this way, the disembodied voice gives life to an inanimate object; it personifies it:

> Voices are produced by bodies: but can also themselves produce bodies [...] It may then appear that the voice is subordinate to the body, when in fact the opposite is experientially the case; it is the voice which seems to colour and model its container. (Connor, 2000, 35)

Our quotidian experience of this personification effect means we generally do not perceive it as uncanny. We typically experience these voices as emotionally neutral (and the quality of voice acting tends to emphasise this), without ascribing them the psychological power of Chion’s acousmêtre. Moreover, our familiarity with these voices as plausible, audible representations of the technology to which they are assigned sets up the potential for the *reverse* effect, in rare instances when their voices jump into the real world, into a human body. For example, artist Tyler Coburn employed the voice actor Susan Bennett, the original voice of Siri, in a live performance of his text work *NaturallySpeaking* (2013-2015). In this instance, for anyone in the audience familiar with her voice, the uncanniness of the situation

would be from the impression of a machine come to life: the vocalic body of the iPhone and its software *made flesh*.

This kind of machine or computer voice also commonly appears in science fiction, where its effect can be either benign, such as the voice of the computer in the television and film franchise *Star Trek*, or a superlative acousmêtre when it deviates from its functional role, such as the HAL 9000 in Stanley Kubrick's film *2001: A Space Odyssey* (1968): an omnipresent, omniscient computer voice which ultimately acts antagonistically to endanger human lives. In fact, the direct origins of contemporary voice assistants are found in science fiction, since technologists drew inspiration for them from fictional examples.[19]

Transferring these captioning, functional voices from utilitarian settings to works of art can present perceptual ambiguities, depending on their deployment. In chapter 4, I demonstrate how I have applied captional voices not to machines, but to musical ensembles and musical material so, for example, the live string quartet becomes a vocalic body. In chapter 5, I demonstrate how the science fiction perception of the malevolent talking computer feeds directly into the voice of *Anticlock*.

2.5 Electroacoustic and sound art models: uneasy sonic materiality and authorial presence

There are other models for use of the disembodied speaking voice in music and the wider field of sound art, primarily in the field of electroacoustic music and musique concrète.

[19] For example, in 2011, Google used the codename "Majel" for development of their voice assistant Google Now, in tribute to Majel Barrett, the actor responsible for the voice of the computer in *Star Trek* from 1966–2009 (Faber, 2020).

Here the voice often becomes purely sonic material, as it would if the composer were using any other kind of field recording, studio recording or synthesised material.

In Luciano Berio's *Thema (Omaggio a Joyce)* (1958-59), a reading by vocalist Cathy Berberian from James Joyce's *Ulysses* is fragmented and treated beyond intelligibility, akin to the use of sung vocals in works by Karlheinz Stockhausen (*Gesang der Jünglinge*, 1955-56) and Jonathan Harvey (*Mortuos Plango, Vivos Voco*, 1980). In these works, the electronic manipulation of the vocal recordings prioritises the sonic result, rather than clear transmission of text, or naturalistic representation of a performer (or performing body).

Within this field, the disembodiment of the voice has met with criticism by some theorists, notably R. Murray Schafer, who coined the pejorative term "schizophonia" for the perceived artificiality of the "split" of sound and sounding body:

> We have split the sound from the maker of the sound. Sounds have been torn from their natural sockets and given an amplified and independent existence. Vocal sound, for instance, is no longer tied to a hole in the head but is free to issue from anywhere in the landscape. (Schafer, 1994, 90)

Simon Emmerson similarly describes his unease with a disembodied human presence, in the context of the voice:

> Human presence in acousmatic music is often fundamentally frustrating even when joyous or celebratory rather than threatening or cruel. It represents a displaced 'other' – the other side of an impenetrable curtain. We hear (and hence observe) but we cannot communicate back. (Emmerson, 2007, 80)

Emmerson's invocation of an Ozian *man behind the curtain*—whether intentionally or not—recalls the epitome of Chion's acousmêtre. There is no artistic reason, of course, that this uncanny perception is invariably undesirable. In fact, as Edmund Hunt notes, this is available for creative exploitation:

> Does the disembodied, electroacoustic voice distance the audience from the communicative power of the words that are heard? Although Simon Emmerson argued that the disembodied human voice in acousmatic music can often seem frustrating, this sense of disembodiment might be turned to the composer's advantage, as the basis of a methodology for creative practice. (Hunt, 2020)

This perception of the disembodied voice in electroacoustic music as frustrating, unnatural or "schizophonic" has been widely propagated, but if the voice is "de-acousmatised", not through physical revealment, but through acknowledgement of the speaker's identity, it can be perceived quite differently, as seen below.

The voice as authorial presence in electronic music

As already noted, the voice in my work is "bearer of utterance", so a purely material and abstract approach—that is, using vocal sounds as sonic material for processing beyond intelligibility, rather than as a vessel for text—is generally outside the scope of my practice.

However, in many electroacoustic musical and sound art works, the voice is not only unprocessed and intelligible, but identifiable as the disembodied voice of the composer themself. This *authorial presence* is an aspect of my practice that has gained increasing importance, through the appearance of my own voice in most works comprising the portfolio.

Emmerson (2007) identifies several musical works in which the authorial presence is vocally signified, including Karlheinz Stockhausen's *Hymnen* (1966-1969), in which the composer's voice is heard intermittently in conversations with studio technicians, and—as already described—Westerkamp's *Kits Beach Soundwalk*, in which the composer "both presents and represents herself along with the soundscape" (Emmerson, 2007, 79) and Cardiff's *The Missing Voice*. Laurie Anderson's aforementioned quasi-autobiographical works similarly feature a vocal authorial presence, as does Anderson's collaboration with the artist Sophie Calle, *La visite guidée* (1994), in which Calle's voice (accompanied by Anderson's music) describes a series of personal possessions (a bed, a wedding dress, a coffee cup, etc.). In these examples, Westerkamp's work is descriptive in nature, and Calle-Anderson's is descriptive/biographical, while Cardiff's work also contains obvious cinematic modes (a mysterious, semi-fictionalised narrative echoing film noir tropes).

Significance of the inclusion of the artist's voice

These inclusions of the artist-composer's voice in a work bring the reasons for its presence into question. The artist's own voice is the most immediate vocal tool at hand, so its use may be a simple case of "availablism" (to use the performance artist Kembra Pfahler's term).[20] In my own work, the immediacy and direct artistic connection of using recordings of my own voice has certainly been one aspect of my employing it.

[20] Pfahler defines her term *Availablism* as "making the best use of what's available… It's not a celebration of poverty, but of abundance. It's having the willingness to be an interdisciplinary artist. In my case, performance was about using the available tools that I had, which was my body" (quoted in Dempsey, 2014).

However, it also becomes signifier of a personal perspective, and even self-portrait in sound, through its distinctive phonological and idiolectic features. In describing Stockhausen's *Hymnen*, Emmerson notes, "If we know the composer's voice, then the breathing is clearly Stockhausen's with his characteristic vocal tract formant features" (Emmerson, 2007, 80). So, this impression is contingent on the listener knowing who is speaking, but once that occurs, the material quality of the voice is that which guides our appreciation of the impression of portraiture.

When the authorial voice is both identifiable and carrier of narrative, this has the effect of humanising what might otherwise have been an austere acousmatic work, and be signifier of a reaction against the uncanny impression of purely electronic works, in contradiction of the unease felt by Emmerson and Schafer. In works by Luc Ferrari, the composer's voice acts as guide, caption, narrator and intradiegetic protagonist, as heard in *Far West News* (1998-99), an extended audio diary of edited field recordings. Rather than being "displaced", Ferrari's voice here invites us, vicariously, into a specific location and set of experiences.

The inclusion of the artist's voice thus becomes an act of autoethnography, as they place themselves within the subject matter of their work, rather than remaining an objective onlooker: the absent modernist composer becomes an attendant post-modernist, as can be heard in the works by Calle-Anderson, Cardiff, Ferrari and Westerkamp above; the artist's voice becomes a reflexive tool.

John Levack Drever also notes that a reflexive approach to music opens the possibility that "compositional processes and audio manipulations that previously lay transparent could be exposed and discussed within the work" (Levack Drever, 2002), with

especial reference to *Kits Beach Soundwalk* and its voiceover by the composer (Westerkamp). This is one aspect of my work *Anticlock* (chapter 5), in which my own speaking voice makes transparent the musical structures and material. But beyond this, Westerkamp's voice in *Kits Beach Soundwalk* is a literal portrait of her own voice, providing personal and emotional content. This notion of the voice as sonic self-portrait gains increasing importance in later works in my portfolio, especially *The Subject / The Object* (as explored in chapter 6).

This is not to say that the artist's voice cannot also be perceived within the cinematic and radiophonic descriptions above. For example, in Bruce Nauman's sound installation *Get Out of my Mind, Get Out of This Room* (1968), the artist's disembodied voice is clearly a terrifying, omniscient acousmêtre, hectoring the listener with the work's eponymous imperative phrase in an otherwise empty space. Similarly, in works of my own, my voice becomes radiophonic character actor (in *Red and Blue*, discussed in chapter 5) and even antagonistic acousmêtre (in *Anticlock*, discussed in chapter 5).

2.6 Ramifications of the perception and function of cinematic, radiophonic and other acousmatic voices on my practice

This wide variety of perceptions of the disembodied voice, in visual arts, radiophonic work, music and everyday life, leads me to consider the effects and functions of the disembodied voice in my own practice as composer. In chapters 4-6, I will consider these on a work-by-work basis, since the musical context, structural deployment and textual content all have ramifications.

I consider the disembodied voice in my musical work to possess both cinematic and radiophonic aspects. In works featuring live musicians, there is an explicit, albeit understated visual element (performers onstage), for which the speaking voice might be perceived as a quasi-cinematic voiceover. Yet, at the same time, as a primarily acoustic medium, the presence of the disembodied voice in music might also suggest radiophonic modes.

To further complicate matters, my deployment of text material such as lists intersects with the functional qualities of the *vox ex machina*. As noted above, this leads to a situation in which the voice produces a vocalic body in an otherwise non-verbal, yet physical, human entity: the personified voice of a string quartet, or other instrumental ensemble.

The perception of the disembodied voice, then, is rich with contradictory signification:

- Uncanny — quotidian
- Diegetic — non-diegetic
- Signifier of bodily absence — signifier of the artist's presence
- Functional —poetic
- Distancing device — personification device.

This ambiguity and variety is the source of its attraction and fascination.

Chapter 3:
Context (2): categories of speech-music relationship

3.1 Introduction to chapter 3: formal-structural interactions

This chapter proposes seven categories of relationships between the speaking voice and music, with examples drawn from works whose aesthetic or structural approach to text and music has fed into my practice.

My work is concerned with this formal-structural relationship between speech and music, so this musical-organisational description serves to situate my work within the field. Akin to the selected models and diverse range of examples in chapter 2, the categories presented here are not intended to represent a prescriptive taxonomy, or stringent method for the analysis of all music featuring the speaking voice. Rather, these categories represent a fluid grid: a lens through which this type of art can be viewed. The choice is musical examples, again, is not intended to represent a canon, but a personal selection that has contingency on my practice.

In chapters 4 to 6, I demonstrate how these categories of relationship have been applied to text in my own music. I also demonstrate the multifarious hybrid interactions between these categories of structural relationships and the perceptual and modal models from chapter 2.

It should be noted that in my practice, the formal relationship between speaking voice and music strays between the categories outlined below, from work to work, and sometimes within a single work. Clearly, the same is true of many artists who have used the speaking

voice in their work: it is not my intention to present these categories as necessarily singular and exclusive within any particular work. Nonetheless, formalising these categories in a non-contiguous manner is important for understanding the potential for dramatic and semantic variety as well as the potential for formal ambiguity of the speaking voice in music.

Unlike the singing voice, which is always perceived as musical, speech may be presented as extra-musical: a separate entity to a musical presentation, as in my first category (A), where voice and musical material are entirely discrete. It can also be deployed with structural interdependence to music, to a point where speech and music are indistinguishable, as in the categories of Speech as Music (category F) and Speech as Sonic Material (category G).

A summary of my categories is tabulated below, beginning with the most structurally separate (A) moving sequentially to the most structurally interdependent (F and G).

These offer a broadly architectural way of understanding the relationship between speaking voice and music.

A	**Discrete Text-Music Relationship**	Music and speaking voice are entirely separate.
B	**Coexistent Text-Music Relationship**	Music and speaking voice are concurrent, but do not structurally interact on a micro-level.
C	**Cognate Text-Music Relationship**	Music and speaking voice share macro-structural organisation.
D	**Rhythmic Text-Music Relationship: Chant or Rhythmised Speech**	The speaking voice is organised on a micro-level according to musical structures.

E	**Speech Melody and Speech Rhythm: "Speechified" Music**	The music is organised on a micro-level according to speech structures.
F	**Speech as Music**	Speech material is indivisible from musical material.
G	**Speech as Sonic Material**	Speech is treated as sonic material, to be manipulated according to musical structures on a micro- and macro-level.

Table 1: Categories of speech-music relationships

In each of these, the voice might assume any of the modes outlined in chapter 2: the omniscient narrator (acousmêtre), character, caption, etc. This chapter focusses on the structural relationships between those voices and their musical setting.

My categorisation differs from previous studies of the text-music relationship in some significant ways. Peter Stacey (1989) lists six basic techniques of "relating music and text" in contemporary composition, focussing primarily on semantic aspects, in use of sung and spoken voice. His six types are "direct mimesis, displaced mimesis, non-mimetic relationship, arbitrary association, synthetic relationship and anti-contextual relationship" (Stacey, 1989), which consider the extent of a mimetic relationship, which might include musical allusions to subject matter, imagery, or the "sound image" of individual words or phrases (word painting).

Cathy Lane lists nineteen different "compositional techniques used in spoken word work" (Lane, 2006), which again consider the varying techniques of using voice on a semantic, sonic, and linguistic level. For example, identifying the concepts of "dissolution of semantic meaning through processing", "accumulation of meaning by massing of voices or "montage" and "nonsense juxtaposition and permutation including the inventing of new

words". Lane's study is especially useful in analysing the work of electroacoustic composers whose music features the voice, but also reflects on my own varied use of the semantic and sonic power of the voice that I will explain in chapters 4-6.

Unlike the studies above, my categorisation below is concerned solely with the overarching structural relationship between the speaking voice and music: the architectural aspect, rather than the semantic.

Additionally, deployment of the speaking voice in my work is within structures that deviate from strophic song and linear narratives (through-composed song forms), so it will be seen below that I have deemphasised my investigation into speaking voice-text-music relationships that are typically associated with these forms—for example, I have generally not utilised Type D (chant, which is close to the singing voice) except in specific instances in my work (such as *Red and Blue*: see chapter 5).

3.2 Type A: discrete text-music relationship

This is the most independent category of relationship: that of speaking voices pre-empting the musical material, or voices interleaved with episodic musical material. Speaking voice and music occur exclusively and, in most instances, may be considered to have equal primacy through never sounding simultaneously: neither voice nor music borrow structural material from each other. The voice is utilised in a naturalistic manner: the speech is not rhythmised, or overtly made musical, although it may well appear in a dramatic register. This voice often assumes the role of narrator, and this relationship is keenly associated with

mixed-media dramatic work. This kind of relationship is long-standing—ancient, even—and persistent.

Many forms of historic dramatic entertainment have exhibited this relationship, including the parallels and precursors to opera, such as the masque form,[21] *comédie-ballet*,[22] and *Singspiel*. Quasi-operatic forms exhibit this use of voices well into the classical era, with characters who speak, or passages of unaccompanied spoken dialogue between arias or instrumental music, in such well-known examples as Wolfgang Amadeus Mozart's *The Magic Flute* (1791), a *Singspiel*. Mozart was undoubtedly influenced by yet another form of spoken word and music theatre popular in 18th century Germany, the *melodrama*, as exemplified by Georg Benda's *Medea* (1775). In this work, brief passages of orchestral music and unaccompanied dramatic narration are tightly interleaved throughout.[23] Later conventional operas might replace these speaking voices with recitative, which stands in for the narrative function of speaking characters. However, many contemporary musicals (i.e., musical theatre works) are similarly not "sung-through" and continue to exhibit this voice-music relationship, with spoken character dialogue interleaved with song.

Later operatic or oratorio works exploit the dramatic separateness of this relationship. In Igor Stravinsky's *Oedipus Rex* (1927), the narrator is a speaking role, whose words almost never coincide with music, emphasising the character's extra-diegetic distance from the action, in contrast to the named characters in the drama, who sing. Here, speech disconnected from music signifies a distance from the stage action (narrator role), as opposed to song,

[21] Exemplified by the Jacobean and Caroline theatrical works of Ben Jonson.
[22] Such as Voltaire and Jean-Philippe Rameau's *La princesse de Navarre* (1745).
[23] Describing his enthusiasm for Benda's *Medea* in a 1778 letter to his father, Mozart writes, "[I] was absolutely delighted...there is no singing in it, only recitation, to which the music is like a sort of obbligato accompaniment to a recitative... this produces the finest effect" (quoted in Drake, 1971).

which signifies characters implicated in the action. Concert works also employ this use of a musically discrete narrator for dramatic purposes. Sergei Prokofiev's *Peter and the Wolf* (1936) is a popular example, in which the role of speaker is clearly understood as narrator of the programmatic musical action. Again, in this work, the music generally ceases for the interleaved spoken narration.

This relationship is not the exclusive domain of musico-dramatic or linearly narrative works, however. Recent albums across genres have employed linking dialogue between songs, for example in the form of comedic skits as popularised by De La Soul's *3 Feet High and Rising* (1989). Saint Etienne's album-length work has taken diverging approaches to this. On *So Tough* (2003), the interleaved spoken word sequences are snippets of pre-existent film dialogue, while on *Finisterre* (2002), the linking spoken word is scripted and performed by actor Michael Jayston. In both cases, the relationship is not an overtly dramatic one: the nature of the commentary on the musical tracks provided by the voices is left semantically open for the listener: the voice is not overtly a dramatic narrator, commenting on the action of the songs.

An unusual example of this discrete interleaving of speech and music occurs in Arthur Russell's track "The Name of the Next Song" (1986).[24] Here, Russell's speaking voice introduces each section of music, assigning each a different title, yet the music remains fundamentally unchanged (rhythmic, harmonic, and sonic material, and sung lyrics remain the same, with variation in length of section and improvised melodies).

[24] This track is available on the remastered versions of Russell's album *World of Echo* (1986). The track was not included on the original release, but is included on the versions reissued by Audika Records (2004 and 2005) and Rough Trade Records (2005).

[Spoken introduction:]	**[Sung/chanted lyric:]**
The name of the next song is "Anti-American"	*California, here I come...*
The name of the next song is "Painted Box"	*California, here I come...*
The name of the next song is "Dodo"	*California, here I come...*

(Russell, 1986)

Here, the speaking music is not an expository device, as in the dramatic works above, but rather serves to obfuscate the meaning of the music and lyrics that follow. It sets up a paradoxical relationship between speaking voice and music, in which the titles might be perceived either: as arbitrary and unrelated to the music, since the sung lyrics and music are essentially unchanging, or as altering the listener's perception of the music, if one's appreciation of any music is affected by knowing its title. This paradox is a device exploited further in my work *Freezywater* (chapter 4).

Russell's use of a clearly perceivable, repeating structure (speaking voice alternating with song), is a voice-music relationship that comes close to the cognate relationship described in chapter 3.3 below.

Works by Tom Johnson, such as *Illustrated Music #8: Narayana's Cows*, also exhibit this quasi-cognate voice-music structure, but in *Narayana's Cows*, the text performs a very different function from "The Name of the Next Song". Rather than provide a lyrical-poetic (albeit paradoxical) frame for the interleaved music, the spoken passages describe the ensuing formal processes in the music. The work is based on a conundrum by 14th century mathematician Narayana, which calculates the total number of a hypothetical herd of cows, if steady birth rates continue. Johnson's spoken text describes the mathematical permutations, for example:

> The first year there is only the original cow and her first calf. The second year there is the original cow and two calves. The third year there is the original cow and three calves. The fourth year the oldest calf becomes a mother, and we begin a third generation of Narayana's cows. (Johnson, quoted in Beyst, 2016)

The musical material (always discrete from the spoken text) "describes" this process: "…pitch correlates with generation, and duration with the difference between a cow (long note) and a calf (short note). We thus come to hear a motif, which, by extension with ever new analogies, grows into an accumulative melody, a special variant of 'repetitive music'" (Beyst, 2016). So, both text and music augment, in analogy to the mathematics. Beyst views the primacy of music and text as equal in this example: "the speaker is, hence, no narrator - his words do not conjure up mental images - at best, his transposition in words refers to the consecutive transposition in sounds" (Beyst, 2016).

3.3 Type B: coexistent text-music relationship

In the coexistent relationship, speaking voice and music sound together. However, the speaking voice remains in no way "musicalised". The natural rhythms and pitches of speech are generally retained. The music too, does not respond to the pitch or rhythmic implications of the speech, but might provide illustrative underscore (in the cinematic sense) to the music. Otherwise, music and speech may simply coexist.

This relationship coincides, to some extent, with Stacey's (1989) categories of Direct Mimesis (word painting, or filmic underscore), Displaced Mimesis (music imitating more oblique aspects of the text), Non-mimetic Relationship (in which "the composer makes no

attempt to imitate any aspect of the text") and Arbitrary Association (in which there is "no resemblance between sound image and object referred to, but an association has been established by continual usage"). However, my category refers specifically to the use of speaking voices in a musical setting, in contrast to Stacey's wider view of the voice in music, much of which is sung.

The most obvious example of coexistent relationship is the accompanied literary reading: poetry or prose accompanied by incidental or background music. Laurie Anderson's *The Ugly One with the Jewels* (1995) is an album-length work comprising primarily spoken prose: storytelling, accompanied by beds of sound that do not directly illustrate the words or voice, or respond to the text musically. In this way, the album is a purer example of the coexistent relationship than other albums by Anderson, such as *Big Science* (1982), or *Mister Heartbreak* (1984). In these works, Anderson moves freely from cognate speech-music relationships (chapter 3.3), chant (chapter 3.4) and song from phrase to phrase, or even word to word, analogous to alternating typographical emphasis (e.g., use of italicised or bold characters). In contrast, the musical material in *The Ugly One with the Jewels* serves only to add mood to the recitations. Music and spoken voice do not interact on a micro-structural level.

Gil Scott-Heron's recordings frequently feature the coexistent relationship applied to poetry, rather than prose, or storytelling such as Anderson's. In the recorded versions of "The Revolution Will Not Be Televised" (1970 and 1974),[25] Scott-Heron's recitation remains structurally separate from its musical accompaniment. The prosody of the vocal performance

[25] The more widely known version was included on the album also called *The Revolution Will Not Be Televised* (1974), but this was preceded by a version with a sparser arrangement on the album *Small Talk at 125th and Lenox* (1970). These two, completely different arrangements underline the structural separateness of the words and accompanying music.

has a rhythmic profile that the music does not imitate in any way and, as such, is also an example of Stacey's Arbitrary Association.

This is not to say that the music does not illustrate a general mood. Moreover, the musical accompaniment contextualises the text in a specific milieu: the bongo and conga backing of the original version and jazz-funk rhythm section and flute of the later version both identify a place and time by cultural association: New York in the late 1960s and early 1970s. For politically charged text such as Scott-Heron's, this locational aspect of the music was important for contemporary listeners (compare this approach to John Cooper Clarke's use of rhythmised poetry with musical backing in chapter 3.4 below).

The coexistent relationship of music to speech is also widespread in cinema, where underscore is likely to be applied to dialogue as well as non-verbal actions. In contrast to the above examples, this may well be illustrative (Direct Mimesis), as well as atmospheric, or suggestive of place and time.

In some instances of text-music coexistent relationship, music has primacy, and the voice assumes the role of audible caption, subtitle or intertitle, to borrow terminology from cinema again. Here, the recitation is not continuous, and the music does not score the text. Instead, the spoken text describes the music, while it is sounding.

Karlheinz Stockhausen's *Der Jahreslauf* (1977) is a theatrical work, accompanying a staged ballet with dancers representing the passage of time (years, decades, centuries, millenia). The music is largely instrumental, but the occasional speaking voices caption, describe, or clarify the stage action. For example, during a passage in which the character of a chef arrives, the recorded voice (Stockhausen's own, in a mode that is perhaps more *availablism* than self-reflexivity, or deliberate marker of authorial presence) announces, "Ein

Koch mit exquisiten Speisen" (*a cook with exquisite dishes*). The effect is highly cinematic, akin to a subtitle, or intertitle (title card). It should be noted that in a revised version, *Jahreslauf* (2001), prepared for inclusion in the opera *Dienstag aus Licht*, Stockhausen replaces the disembodied speaking voice with physical singers, performing sung versions of the same text. The effect is lost: the voice no longer functions as narrator-caption, but as operatic character.

Use of the recorded speaking voice as caption in Stockhausen's work occurs elsewhere in his *Licht* cycle of operas. *Mittwoch aus Licht* (1995-1997) opens with a 52-minute sequence of multichannel, instrumental synthesizer music, but at thirty-eight minutes into the work, a voice (again, Stockhausen's own) announces the title of the scene (in both German and English): "Mittwochs-Gruß; Wednesday Greeting". The effect is again cinematic: an audible intertitle.

3.3 Type C: cognate text-music relationship

In this relationship, the rhythm and pitch of the speaking voice is again natural, so still not musicalised on a micro-level, similar to the coexistent relationship above. But now the voice and music share a single musical structure: they are cognate: obviously linked at points in time. This structure may be an iterative one (such as a sequential, episodic, or repetitive structure, in which the voice enters periodically), but the speech is not rhythmised, as in Type D (chant).

In Frederic Rzewski's *Coming Together* (1971), a speaking voice performs a fragment of text from a letter written by Sam Melville, who had been convicted of bombing

US government buildings in protest at American imperialism and the Vietnam War and was killed in the Attica prison uprisings of 1971. Rzewski segments the text into short phrases, sometimes as brief as single words, and sets them against music for ensemble that responds to a bassline in constant semiquaver motion in 4/4 throughout the piece. The text begins thus:

> I think the combination of age and the greater coming together is responsible for the speed of the passing time. It's six months now and I can tell you truthfully few periods in my life have passed so quickly. I am in excellent physical and emotional health. there are doubtless subtle surprises ahead, but I feel secure and ready.

Rzewski manipulates the division of the text as shown below, with the voice speaking at the beginning of each bar: "The first fragment of text ("I think") should be spoken as an upbeat to the beginning. Thereafter every fragment of text is spoken at the beginning of each measure" (Rzewski, [performance procedure to score] 1971):

Bar 1: "I [upbeat into first bar] think"
Bar 2: "the combination"
Bar 3: "of age"
Bar 4: "and a greater coming together"
Bar 5: "is responsible"
Bar 6: "for the speed"
Bar 7: "of the passing time [...etc.]"

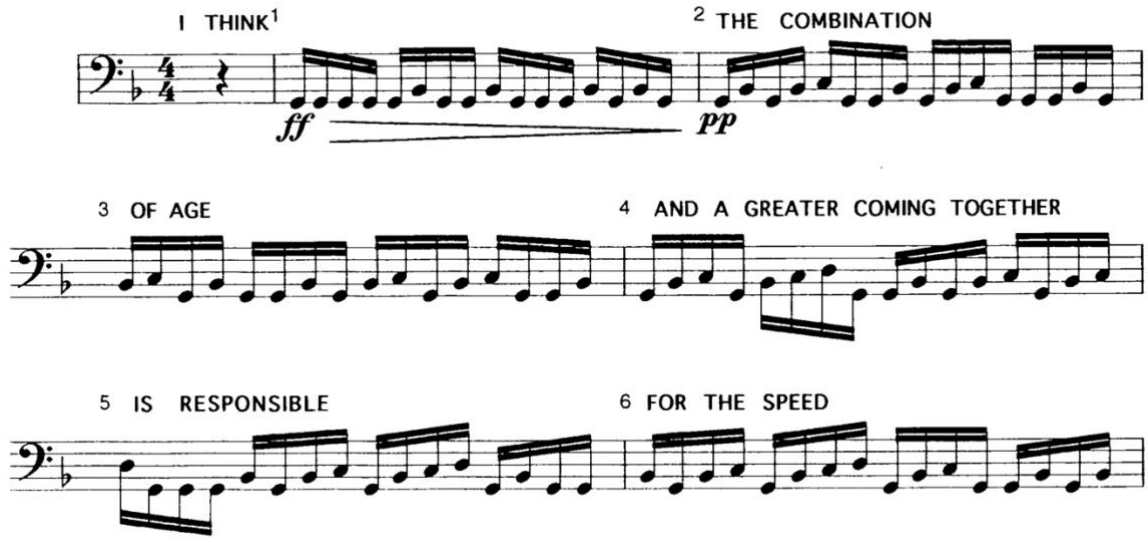

Figure 1: Frederic Rzewski's *Coming Together*, measures 1-6 of the score[26]

In this way, the speaking voice and accompanying music are structurally linked, giving the impression of them being integrated throughout the work. Although the speaker maintains perceptibly natural speech rhythms and pitch within each bar, the pacing of the text and music are matched.

The cognate relationship cements the structural interdependence of music and text. In the closing 6-minute sequence of Mike Oldfield's "Tubular Bells Part 1" (from the album *Tubular Bells*, 1973), the speaking voice of Vivian Stanshall (credited as "Master of Ceremonies") announces which of Oldfield's instruments is playing the melody. This section of the work is a chaconne form, with a repeating ten-bar melody and phrase-length, so Stanshall's voice enters roughly at bar ten in every phrase:

[26] The complete score is available via https://imslp.org/wiki/Coming_Together_(Rzewski%2C_Frederic), provided by the composer under a Performance Restricted Attribution Non-commercial No Derivatives 1.0 license.

	Speaking voice (Stanshall)
Phrase 1:	“Grand piano”
Phrase 2:	“Reed and pipe organ”
Phrase 3:	“Glockenspiel”
Phrase 4:	“Bass guitar”
Phrase 5:	“Double-speed guitar”
Phrase 6:	“Two slightly distorted guitars”
Phrase 7:	“Mandolin”
Phrase 8:	“Spanish guitar and introducing acoustic guitar”
Phrase 9:	“Plus, tubular bells”

The voice here is implicated in the music’s structure, in a highly audible manner. It should be noted that there is an obvious humour to the use of this subtitular cognate voice-music relationship here, in Stanshall’s crisp enunciation of instrument names. This is a reflection of a similar use of the voice on Stanshall’s own (earlier) track “The Intro and the Outro” (from the Bonzo Dog Doo-Dah Band’s album *Gorilla*, 1967), from which Oldfield took inspiration. On this track, the introductions are absurd (“Princess Anne on sousaphone [...] Digging General de Gaulle on accordion”) and the pacing of the voice is freer, following short, irregular phrase lengths (two or four bars), but still matching the structure of the instrumental entries.

This use of list and litany-type texts is well suited to deployment in this cognate relationship between speaking voice and music, and is one I have explored extensively in my practice: in chapters 4 and 5, I will demonstrate ways I have exploited it in *Freezywater*, *Affix Stamp Here* and *Anticlock*.

3.4 Type D: rhythmic text-music relationship: chant or rhythmised speech

Here, speech is musicalised by application of rhythm, on a micro-level (i.e. syllable to syllable rhythmic and metrical settings) as well as a macro-level (following more large-scale musical structures, as in the cognate relationship).

Rhythmised speech—chant—is the most prevalent form of speaking voice in music, occurring in a vastly wide number of contexts and traditions. Moreover, there is grey area of what might be termed "true chant" (rhythmised speech without musicalised pitch, stress pattern and intonation), and "talk-singing", or half sung performances (as seen across a wide variety of pop and rock, from Bob Dylan and Serge Gainsbourg to Mark E. Smith (of band The Fall) and the current proliferation of British post-punk bands using this style of vocal delivery.[27] Thus, the boundaries of chant also spill over into song: Schoenberg's use of *Sprechgesang* in *Pierrot Lunaire* is a pitched voice, for example, albeit one that "immediately abandons [the notated pitch] by falling or rising" (Schoenberg [introduction to the score of *Pierrot Lunaire*], 1912).

The popularity and diversity of contemporary hip hop and rap has made the rhythmised speaking voice ubiquitous across international contemporary popular music. Since its origins in late-1960s Jamaican toasting cultures, arguably related to earlier post-colonial West African griot traditions of chanted storytelling, rap has embraced huge stylistic

[27] For a journalistic critique of contemporary British rock and indie bands using chanted or spoken vocals, see Fergal Kinney's article, "Landfill Sprechgesang? Yard Act's The Overload Reviewed" (Kinney, 2022).

variations and degrees of nuance between chant and song, putting a discussion of the form outside of the scope of my research.

This overlap between chant and song has made this type of relationship a generally less fruitful tool for use in my own practice, although it appears in the collage structure of *Red and Blue* (chapter 5) as only one mode of vocal delivery in a purposefully diverse set. Additionally, the quasi-spoken forms of musical delivery in song, exemplified by recitative and parlando, exist outside of my categorisation. My interest is in escaping familiar song forms and structures, such as strophic forms and linear narrative forms; chant often lends itself to these forms. However, in works such as *Anticlock* (chapter 5), rhythmised speech is utilised as a coordinating device, while the text material it carries is highly abstracted, which avoids any suggestion that the chanting voice is bearer of any familiar song form, or narrative material (e.g., strophic text, or lyrical song text writing).

Rhythmised speech also overlaps with poetry. In the performances of poet John Cooper Clarke, the metrical stresses and structures are highly emphasised. In fact, simply adding a musical backing transforms this poetry into song. On the track “The It Man” (1980), Cooper Clarke’s text is strongly revealed as having a 4/4 musical metre and rigidly four-bar phrase lengths, by the addition of drums, synthesizers and piano sharing his rhythmic structure, and the metrical feet of the verse are thus transmuted from poetic rhythm to musical rhythm.

The act of applying music to poetry on “The It Man” also situates the words in the late-70s post punk scene, through the sounds and styles of its production. According to writer-comedian Stewart Lee, “it makes explicit the rhythms that are implicit in the poems,

which seems a shame… also there's something about the drum sounds and the production on some of those records that really fixes it in time" (Lee, quoted in Ross, 2012).

Compare this chant relationship of poetic words to the more fluid and equivocal—possibly more nuanced—coexistent relationship of voice and music in Scott-Heron's "The Revolution Will Not Be Televised", which similarly situates a text in time and place through its musical arrangement but leaves an expressive ambiguity in the rhythmic (dis)connection between voice and band.

3.5 Type E: speech melody and speech rhythm: "speechified" music

In this relationship, the musical-mimetic role of chant, or rhythmised speech, can be said to be reversed and instead music is "speechified". Here, the inherent rhythms and/or melodies of speaking voices are used as models for imitation by the music, on a micro-structural level. The music is thus derived from the contours of the speech, rather than a rhythm being imposed on the speech, as in chant.

In Steve Reich's *Different Trains* (1988, for string quartet and pre-recorded audio), snippets (very short phrases) of recorded testimonials from Holocaust survivors, and reminiscences from Reich's private tutor and a Pullman porter[28] are looped. The pitch contours and rhythms of the vocal snippets are transcribed and imitated by the strings in canon.

[28] Pullman porters were staff hired as sleeping car attendants by the American railways Pullman Company from the 1860s till the dissolution of the company in 1968. Those employed were almost exclusively African American men, and the profession led to the formation of the first African American labour union in 1925.

Maltman and Vickery (2018) trace the origins of Reich's inspiration for this technique to Scott Johnson's *John Somebody* (1980-82), which is an obvious precedent in its use of looped recorded speech imitated in pitch and rhythmic unison by electric guitars.

Speech melody in these works hinges on the speech-to-song illusion identified by Diana Deutsch: "speech is made to be heard as song, and this is achieved without transforming the sounds in any way, or by adding any musical context, but simply by repeating a phrase several times over" (Deutsch, 2003).

Deutsch's examples focus on pitch, rather than rhythm, and stress the importance of repetition in this illusion, but speech melody transferred from recorded speech to instruments in non-repetitive works is also found in Peter Ablinger's *Voices and Piano* (1998-ongoing), a series of pieces wherein recorded interviews and speeches by various cultural and political figures are analysed for pitch content, which is played in rhythmic unison by the pianist, as if the melodic-rhythmic analysis were occurring in real-time.

A more nuanced example of instruments following the rhythm and stress patterns, and timbre of recorded speech occurs in Vickery's own *The Semantics of Redaction* (2014), in which any recorded topical news broadcast generates a score for a percussion, who either mimics or obliterates these patterns.

David Fennessy's Piano Trio (2012) is a notable example of a speech rhythm work that inverts the typical scenario. This work uses an archive recording of a conversation between the composers John Cage and Morton Feldman. Here, rather than the instruments mimicking the syllables of their speech, the instruments play punctuating gestures in the *pauses* in their conversation, creating a negative sonic image of their speech rhythms, using

the pauses as "a kind of mould into which I have poured my own musical material" (Fennessy, n.d.).

It might seem that this imitative technique is endemic to use of recorded speaking voices. However, examples of a live speaking voice also exist, albeit more rarely. In Robert Ashley's *Tap Dancing in the Sand* (2004), an ensemble attempts to mirror the rhythms of a speaking soloist in real-time. An imitation of the pitches here would be virtually impossible, and Ashley acknowledges in the score that the instruments will play "in sychronization (*as closely as possible*) with the syllables of the English text" (Ashley, 2004, emphasis mine). A similar situation occurs in Thomas Kessler's choral-orchestral work *...said the shotgun to the head* (2003-2004), which features a text and spoken solo performance by the poet-rapper Saul Williams. At the outset of this work, a solo cello imitates the rhythm of Williams's recitation, and later the entire orchestra moves in passages of near-rhythmic unison with the speech. Here, the orchestra's material is fully notated, so the effect hinges on the rigour and constancy of Williams's technique and performance.

It should also be noted that a separate, but related form of generating musical material exists, in which the contours of natural speech rhythms are imitated instrumentally, while any voices remain unheard. This occurs in the work of John Coltrane, such as on the album *A Love Supreme* (1964). On the track "Psalm" Coltrane sets a poem instrumentally (the text is reproduced on the album's sleeve), so that a "comparison of the poem with Coltrane's improvisation reveals that his saxophone solo is a wordless recitation of the poem, one note to each syllable" (Porter, 1985).

This kind of instrumental setting of speech rhythm also occurs concurrently with vocalised settings of a text in Liza Lim's opera *Tree of Codes* (2016): "At the end of Act 3,

Adela, played by soprano Emily Hindrichs, swaps role with a plant-creature… and relates a version of Goethe's "Der Erlkönig" as told by Bruno Schulz. What you hear is both her chanting whisper and the chanting of rasping woodblocks. The speech patterns of Goethe's poem are transferred into these froggy, insectoid scraping sounds" (Lim, 2017).

An even more rigorous transformation of speech to instrumental music occurs in works that resynthesize vocal sounds through spectral analysis. In Jonathan Harvey's *Speakings*, this analysis results in a score for orchestra, in which the music clearly seems to reproduce non-verbal, human vocal sounds. In Peter Ablinger's deeply uncanny *Speaking Piano* series of works (such as *A Letter from Schoenberg*, 2006), this is carried further: a computer controlled piano appears to "speak" intelligible words, through careful mapping of vocal formants to the keys of the instrument.

3.6 Type F: speech as music

In this type of relationship, the speaking voice becomes the sole sonic element, and the musicality of the vocal sounds are accepted as inherent. This approach to the speaking voice is often the domain of sound poetry (for example, in the works of *poésie sonore* poets Henri Chopin and François Dufrêne in the 1950s). Here, the distinction between the prosodic elements of performed poetry and music become blurred, through a focus not necessarily on the linguistic features of the spoken language, but through phonemes and non-linguistic vocal sounds chosen for their sonic or musical qualities. Naturally, practitioners of this kind of voice-music relationship often find their work categorised as ambiguously existing between the artforms of music and poetry, so artists such as Jackson Mac Low are described (or self-describe) as "poet-composer", or similar terms. Others emphasise the musical content of this

kind of work by preferring terms such as "text sound composition" (coined by artists Bengt Emil Johnson and Lars Gunnar Bodin, 1967). For some practitioners, such as the sound poet Bob Cobbing, this form even represents a quasi-primeval form of music:

> They [Chopin and Dufrêne] had gone back beyond the word, beyond the alphabet to direct vocal outpourings which completely unified form and content. They were back where poetry and music began. (Cobbing, quoted in Krogh Groth, 2015)

This perception of the non-linguistic speaking voice as a quasi-precursor to all music recalls Kurt Schwitter's *Ursonate* (1932), which is structured according to a resolutely musical form (sonata form), rather than a poetic form.

Certain artists, such as poet-composer-performer Chris Mann, include both linguistic and non-linguistic elements in their speech as music work, choosing the linguistic elements both for their sonic and semantic characteristics. As such, Mann's work "straddles the boundary between music and language" (Schwartz, 2013), further clouding the distinction between literary and musical forms.[29]

This category of speech-music relationship informs my work *The Subject / The Object* (chapter 6), in which a spoken text provides the primary musical material, without the adornment, contextualisation, of augmentation of any significant additional pitch or rhythmic content.

[29] The perception of "speech as music" as (n)either poetry or music is further obfuscated by the number of artists who perform this kind of work (as poet-composer-performers) within musical ensembles. For example, Bob Cobbing's work with the free improvisation ensemble BirdYak and Chris Mann's work with the ensemble Machine for Making Sense. The latter's album *On Second Thoughts* (1994) represents the uncategorisable result of this poetic-musical approach, on which "the music is structured around a text which is written and performed by Mann. The other members of the group improvise around his reading". (Gallacher, 1995)

3.7 Type G: speech as sonic material

In this category, recorded speech becomes sonic object, or source material for manipulation in electronic, electroacoustic and musique concrète works, such as those already discussed in chapter 2.5.

On a micro- and macro-structural level, the music might be guided by the inherent rhythms of the speech. In this way, the category overlaps with the Type E and F categories, especially when the sole material is speech, such as in Steve Reich's tape pieces *Come Out* (1966) and *It's Gonna Rain* (1965), in which "by using recorded speech as a source of electronic or tape music, speech-melody and meaning are presented as they naturally occur" (Reich, 2002), but the speech might also be electronically manipulated in a way that imposes a musical structure upon it. This is especially the case in music using sampled speech, such as Paul Hardcastle's track "19" (from the album *Paul Hardcastle*, 1985), where the rhythmisation of the voices is achieved through studio sequencing.

3.8 Applicability of the categories to disembodied speech

As demonstrated by the variety of sources quoted in this chapter, the speaking voice can be present in musical works as either physical or disembodied, and I contend that all the above categories can contain either kind of voice. A category such as speech melody / speech rhythm (Type E) might appear to demand disembodiment, but that is not the case: Steve Reich followed *Different Trains* (utilising disembodied voices) with further works in which the voice is physical, through use of the same techniques applied to video material, rather

than audio recordings, in *The Cave* (1990-1993) and *Three Tales* (2002) (both in collaboration with the video artist Beryl Korot).

However, as I will demonstrate in the succeeding chapters, which are case studies of my portfolio, several of the categories are less applicable to my own work, due to the intersection of the modes of disembodiment (chapter 2), intelligibility of the texts, and my concern for fusing speech to music, as outlined in chapter 1.4. For example, the discrete relationship (Type A) is not generally present in my work, since it does not seem a fertile medium outside of physical, theatrical presentations of the voice, and the heavily manipulated electronic treatment of Type G is not used, due to my desire to present the voice as "bearer of an utterance" (Dolar, 2006).

Chapter 4:
First approach: found text (1): list texts, litany structures

This chapter (4) and the following two (chapters 5 and 6) describe the development of my use of the disembodied voice in my practice between 2014 and 2020, as practically demonstrated through the works that comprise the portfolio.

These chapters describe the sources for the texts I have employed, and the nature of their deployment. The combination of a highly selective and specific use of text, alongside an awareness of the listener's potential perception of the voice (as described in chapter 2), and the structural relationship between voice and music (as described in chapter 3), characterises and delineates my practice.

The works have taken three approaches to text source:

1. Chapter 4: Found text (1): lists, litanies (i.e., musicalised list structures) and captions.
2. Chapter 5: Found text (2): radiophonic and narrative character approaches.
3. Chapter 6: Stream of consciousness and montage: notions of the disembodied voice as self-portrait.

Anti-lyrical text: defining the term "found text"

A chief concern in my selection of text sources for all the works described is avoiding typically "musical" texts, such as settings of poetry, song lyrics, operatic libretti, or other forms of writing that exhibit a metrical or prosodic quality. As previously noted (in chapter 1.4), the focus on the speaking voice, rather than the singing voice in music opens an aesthetic door to atypical texts, as well as atypical musical scenarios and structures.

As such, I frequently use the term “found text”. This should be understood as a broad description of my methods of deriving texts from non-lyrical sources, such as scientific texts, examination of maps, and my edits of verbatim or interview material. It is not necessarily demonstrative of the text having been discovered by chance (although that is literally the case in the postcard messages that form the basis of the text of *Affix Stamp Here*). The term “found text” in my practice is directly analogous to poets working with pre-existing text, presented in collage, montage or cut-up forms, or otherwise edited and re-organised to create “found poetry”, such as Howard Nemerov’s emblematic *Found Poem* (1987), which is constructed from a montage of phrases from a newspaper.

Translations of forms from the visual arts: “object-like” music

Another thread running through the portfolio of works is a concern with deploying text in a manner that suggests a translation of forms from visual arts practice. This springs from an impulse to create musical works that are more “object-like” and possess a quality of “boldness”. This boldness is manifested as a simplicity of form, or an audibly perceptible form, such as the works written under the first approach (chapter 4), or the monolithic forms of *The Subject / The Object*, which are non-dialectical or non-teleological, or otherwise exhibit a complex montage or collage form, in the case of *Red and Blue*.

As will be seen in the discussion of *The Indistinguishables* and *Freezywater*, this concern with object-like music leads me to draw on models of visual artworks that comprise collections of curated objects, besides employing musical material that is likely to provoke a sense of familiarity, such as triadic chords, presented in a manner suggestive of found objects.

List texts in music: precedents

The first group of works I will consider all utilise text in the form of lists: single words or short phrases, arranged systematically. These are *The Indistinguishables*, *Freezywater* and *Affix Stamp Here*. In all these works, the lists are spoken by a disembodied, pre-recorded voice, in conjunction with live musicians.

There are several precedents for list texts in musical works, in works both destined for live performance and in studio works. Notably, several British composers associated with the "English Experimental" school[30] have set lists as song. Michael Nyman's *Bird List Song* (written in 1979 and prominently featured in the soundtrack to Peter Greenaway's film *The Falls*, 1980), as its title suggests, comprises a list of obscure bird species, sung on a single pitch against an accompaniment of cycling dominant sevenths. Laurence Crane's song cycle *Events* (1997) sets lists of notable birthdays, monetary exchange rates, and British weather statistics, all taken from a single edition of *The Guardian* newspaper, echoing the found poetry approach of Nemerov.

Crane has also employed a spoken list in the studio work *492 Telecom Announcements* (1993), which uses pre-recorded *vox ex machina* automatic telephone announcements. This work is an hour-long litany of these announcements in sequence, set against cyclic music for electronic keyboards and percussion. Similarly, John White's "Finale – Drum Song" from *Fashion Music* (recorded 1990, for commercial release in 1992) uses a

[30] In this context, I use the term to refer to composers directly associated with the Scratch Orchestra (which performed from 1969–74) and the circle around Cornelius Cardew, such as Michael Nyman and John White, and composers a generation younger who have acknowledged the influence of these musicians (Pisaro, 2001).

spoken list of fashion designers' names to punctuate a drum machine rhythm track ("Katherine Hamnett... Jean-Paul Gaultier... Yohji Yamamoto... Gianni Versace..." etc.). The speaking voice in both Crane and White's works, combined with the sounds of cheap electronic instruments, lend the music a wry quality.

This humorous use of the spoken list is also found in pop. In the song "The Booklovers" (from the album *Promenade*, 1994) by the band The Divine Comedy, the spoken lyrics comprise a list of authors, with glib one-liner jokes succeeding each name. The accompanying music, with its cyclic harmonies, echoes Michael Nyman's. Amanda Lear's "Alphabet" (from the album *I am a Photograph*, 1977) features an alphabetised list of Lear's phrases reflecting the anxious-yet-louche milieu of the late 1970s "for the children of my generation" (for example, "F is full-frontal and friends; G of course, stands for getting a divorce; H now stands for hijacking"), recited over a disco version of J.S. Bach's prelude in C from the *Well-Tempered Clavier* (BWV 846).[31]

The presence of a spoken list in a musical context might often be perceived as absurd (in the case of The Divine Comedy), or absurdist (in the cases of Crane and White), on account of its unfamiliarity and the potential for a comedic delivery, but it is not necessarily the case that a list inherently has this effect. The song "Imperfect List" (originally released in 1990) by the recording project Big Hard Excellent Fish[32] features a spoken list of dire societal issues ("The Bomb... Heysel Stadium... police harassment... the death of the rain

[31] Use of the alphabet as a structuring device appears in many other musical and sonic art works, such as Bob Cobbing's sound poem *ABC in Sound* (1965), Louis Andriessen's *M is for Man, Music, Mozart* (1991) (text by Peter Greenaway) and The Tiger Lillies' "ABC" from *The Gorey End* (2003) (text based on writing by Edward Gorey).

[32] This project was primarily led by Pete Wylie, who was associated with various punk and post-punk bands in Liverpool in the late 1970 and early 1980s, the best known of which was Wah!. Wylie was not credited on the initial releases of "Imperfect List", although he wrote the music and text.

forest.. the Troubles...") and enervating inconveniences and "pet hates" ("Neighbours... lost keys... phony friend..."). Here, the vocal performance by Josie Jones has a grim, splenetic delivery that pushes the tracks towards the tragic. However, the choice of spoken, as opposed to sung vocals imbues the track with a kind of *Verfremdungseffekt*: there is a distanced, objective quality, rather than a melodramatic effect, as may have been the case is these words were sung. It is this ambiguous, emotional distancing aspect of the spoken list that I have exploited in the following works in this chapter.

4.1 *The Indistinguishables* (2014)

Text source and structure

The Indistinguishables is a work for string quartet and pre-recorded voice. This is the earliest example in the portfolio of a performance work for live musicians and a disembodied voice. In this instance, the text is a list of seventy common names of moth species sighted in the UK: a demonstration of both the list/found text approach as text source and an inherent litany approach as musical structure. The choice of moth names opens the work to a rich range of evocative text-imagery, since the names of moths have a wide variety of allusions: some of them fanciful or folkloric in origin, some simply descriptive of the visual appearance of a species.

The choice of names for the *The Indistinguishables* was arrived at through a curatorial approach to the text. Initially, a long list of several hundred names was compiled, drawn from

online taxonomies[33] and entomological field guides. This was then edited by exclusion of any name including the word "moth" (except for the final name – "Ghost Moth") to create an ambiguity of source. This shortlist was then edited and sequenced, through grouping the names into seven loose categories of allusion, with ten names in each category. This process was instinctive, as can be seen from the final edit of the text, arranged in the groups below:

Group 1. Association / allusion: geometry / the astronomical
The Satellite; Strange Wave; Rhomboid Tortrix; The Many-lined; Clouded Border; Lunar Double-stripe; The Passenger; Gold Triangle; Small Black Arches; Sorcerer
Group 2. Association / allusion: journeying / natural features
The Traveller; Grass Emerald; Chalk Carpet; Small Rivulet; Clouded August Thorn; Pine Processionary; Frosted Green; Bordered Pearl; Scarlet Tiger; Map-winged Swift
Group 3. Association / allusion: beauty / prestige
The Forester; Merveille du Jour; Satin Lutestring; Peach Blossom; Beautiful Golden Y; Mother of Pearl; The Festoon; Yellow Horned; Feathered Beauty; The Gem
Group 4. Association / allusion: the abject / threat
The Nonconformist; Pistol Case-bearer; The Shark; Dark Dagger; Blood-vein; Slender Burnished Brass; Heart & Club; The Drinker; Silver Hook; The Seraphim
Group 5. Association / allusion: impermanence / the ephemeral
The Exile; Old World Webworm; The Confused; Large Dark Prominent; Feathered Ear; The Lackey; Silver Cloud; Scarce Forester; The Triangle; Water Veneer
Group 6. Association / allusion: magic / the alchemical
The Alchymist [sic]; Spotted Sulphur; The Geometrician; Tamarisk Plume; The Vestal; Figure of Eighty; Chinese Character; Fiery Clearwing; The Flame; The Phoenix
Group 7. Association / allusion: the deathly
The Gothic; Apple Leaf Skeletonizer; Pale Stigma; Thistle Ermine; Silvery Arches; White-speck; Old Lady; Small Angle Shades; The Vapourer; Ghost Moth

Table 2: Sequencing of text in *The Indistinguishables*

[33] Especially the comprehensive systematic lists on the website https://www.ukmoths.org.uk/systematic-list/ .

Transferal of visual arts aesthetics: curated objects

This curatorial approach to the text springs from a transferal of the aesthetic of a curated collection of physical objects, which in turn suggests a structure of discrete, fleeting moments of music separated by fermatas in a grid-like structure: musical objects, which should exhibit some self-similarity: The list provides not only the spoken text, but also provides a clearly audible musical form.

> Each discrete chord or gesture in *The Indistinguishables* is like a moth pinned to a board. It is a whole, complete object, but designed to be viewed and compared with many other similar objects. (Mayo, 2015)

Visual artworks in which the artist acts as curator of a collection of found physical objects are numerous, but *The Indistinguishables* draws specifically on the works of Susan Hiller, such as *Rough Seas* (1982-2019, a large collection of dry prints of wave imagery) and *Dedicated to the Unknown Artists* (1972-1976, which comprises a large number of found postcards and other documents mounted as a collection). Besides physical objects, Hiller's work has also utilised extensive agglomerations of recordings of spoken interviews, in works such as *Witness* (2000) and *Channels* (2013).

Hiller's view of these works comprising collections of discrete objects is that a narrative inevitably emerges from their juxtaposition when viewed as a whole, which nonetheless possesses an ambiguity of meaning:

> I take it that any conscious configuration of objects tells a story [...] If you think about the narrative that collections of assemblages of things make, the interesting thing is that there are always at least two possible stories; one is the story that the narrator, in

> this case the artist, thinks she's telling—the story-teller's story—and the other is the story that the listener is understanding, or hearing, or imagining on the basis of the same objects. (Hiller and Einzig, 1996, 226-35)

Here, the term "conscious configuration" signifies the artist's curatorial manipulation of the objects into a form that suggests overall meaning. In approaching an analogous work in music, the choice of text is obviously important to the work overall, but the sequencing of text from one phrase to the next provides a sense of meaning: while an assemblage or collection of objects and their collective meaning (depending on physical extent and layout of the objects) might be perceived at a glance, music unfolds in time, so is never perceivable in its entirely in a single moment. This is the purpose of the curatorial sequencing (or "conscious configuring") of the names in *The Indistinguishables* in seven groups of association, rather than (for example) presenting them in a random, or in alphabetised sequence. In this way, each moment of the text provides both intellectual and emotional meaning *and* ambiguity in relation to the text which precedes or succeeds it, in the manner of Hiller's perceptual "two stories".

Text-music relationship

The quartet's material in this work is generally a single gesture (a staggered chord of a bow's length, or timed to end with the vocal). Each of the seven groups of names has an associated chord or harmonic area, technique, or effect (see figure 2).

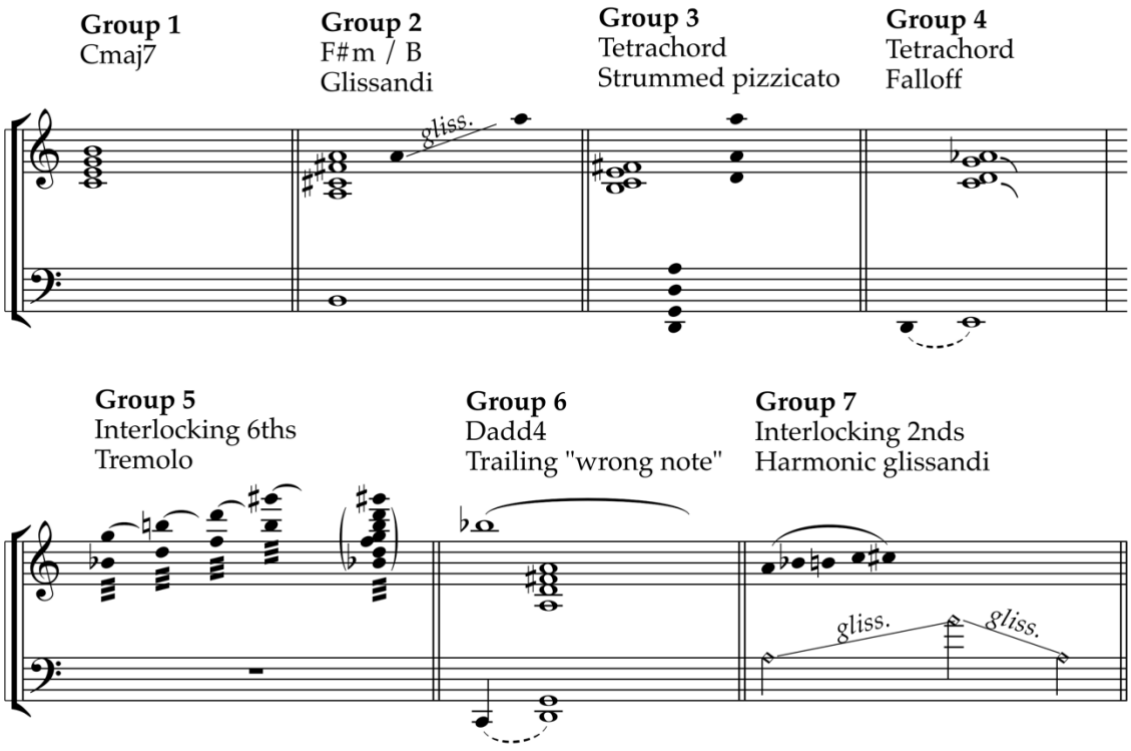

Figure 2: Fundamental musical material accompanying each of the seven groups of names in *The Indistinguishables*

As can be seen, this material is a combination of extended triadic harmonies (groups 1, 2 and 6), in which the inversion of the chord is clear; and more ambiguous harmonies that nonetheless suggest the possibility of a resolution into a triadic harmony. These harmonies are likely to provoke a sense of familiarity in the listener, in a manner intended to reflect a group of curated, found objects: for example, I view the opening C major 7th chord as a "harmonic object". This would be less obviously the case if the materials were atonal and in avoidance of pitch centres. Accordingly, the material is not treated to any developmental manipulation, such as a motivic development, functional harmonic motion, or organic transition from group to group.

The variations to which the harmonies within each section are treated are chordal inversions and re-voicings, which are freely composed, bar to bar, suggestive of an object viewed from several positions. The persistent use of "one full bow long" throughout the score as a maximum unit of durational value also reflects the object-like use of this inherent aspect (limitation) of string technique. This is only occasionally subverted, in what might be termed freely composed "magic moments", such as the extended, pulsing diminuendo/crescendo at bar 39 ("Silver Hook") and in the tremolandi of group 5, wherein the players might reasonably be expected to maintain the general pacing of the work, established through the bow's length chords.

On a micro level, the work is organised according to the inherent speech rhythms of the stressed syllables of the voice (Type E), with the approximate rhythms of the voice notated in the score for imitation by ear (as shown in figure 3), rather than according to a more precise notation in the manner of Kessler's *...said the shotgun to the head*, for example. The speech melody is not imitated: the work is a freer setting of the recordings, and the recordings are not looped or repeated.[34]

[34] As previously noted, the meaningful employment of speech melody is generally contingent on some kind of looping effect, to highlight Deutsch's speech-to-song effect, as in the works by Reich and Johnson cited in chapter 3.5.
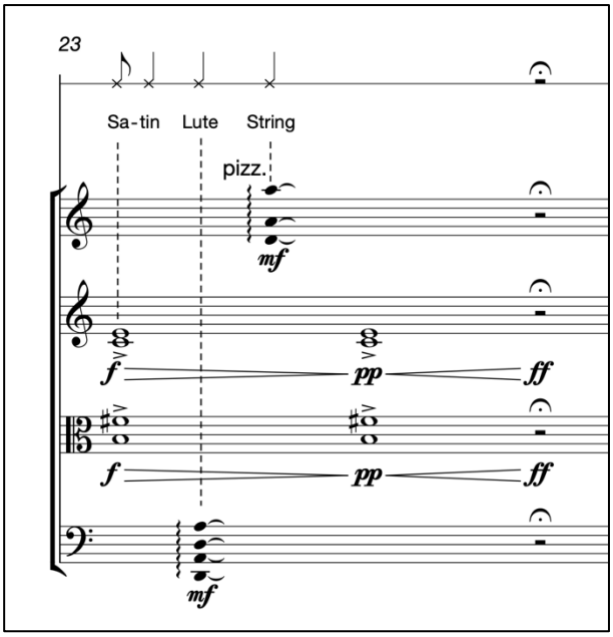
Figure 3: Example of use of approximate speech-rhythm in *The Indistinguishables*

Perception of speaking voice: the caption effect

Since the list structure of the text informs the musical structure on a macro-level (utterances of each name are separated by a fermata; each name is heard alongside a fleeting chord or gesture from the quartet) and a micro level (through speech rhythm), this sets up an ambiguous hierarchy between live musicians and pre-recorded voice. The quartet might seem to be providing a soundtrack, or framing the names with musical material, or conversely, the musical gestures by the quartet may seem to be "captioned" by the names. The voice might also be perceived as the collective speaking voice of the quartet (i.e., a personification effect).

These nuances of hierarchy between voice and musicians are compounded by the practicalities of the performance: the quartet have control of the pacing of the playback of the names (through use of a foot pedal and the theatrical cueing software QLab, so one name can

be triggered at a time), so the live musicians' sense of ensemble and choice of how long any fermata should last is theirs: this is a string quartet work including a voice, rather than a soloistic work for a voice accompanied by the quartet. The disembodiment aids the perception of the voice not as a soloist; a physical vocalist would lend the work a distinct theatricality and shift the perceived hierarchy towards that particular performer.

This perception of the voice as caption is also produced through specific choice of vocalist and intonation. The voice of *The Indistinguishables* is actor Gemma Saunders, using a deliberately neutral tone and inflection, which is close to the typically dispassionate delivery of *vox ex machina* (as considered in chapter 2.4), rather than a radiophonic, dramatic, or overtly poetic reading.[35]

Published reactions to the work by Richard Glover and Chris Mayo indicate that this intended ambiguity is, at least in-part, successful:

> Each section [of *The Indistinguishables*] inhabited slightly different harmonic and melodic environments, and the audience was left to reflect upon the combination of the two elements being presented. (Glover, 2015)

[35] Saunders had provided spoken voiceovers on works of mine predating those in the portfolio, before I began the more methodical investigation of the speaking voice in music under discussion in this thesis. These included the electronic pop track "Richard of York", from the album *Unfair / Funfair* (2006, released under my pseudonym Simon Bookish). Saunders' career includes stage appearances with (e.g.) the RSC (Royal Shakespeare Company) and National Theatre (London), but she is more extensively credited with voiceover appearances, in over 100 BBC radio dramas, in documentary and commercial voiceover work. My impulse for featuring her as the voice of *The Indistinguishables* came from a desire to work with a "disembodied voice expert" on this project, in addition to the enjoyment of continuing our working relationship. Saunders suggested and performed many subtle variations of nuance in the studio sessions for the work, enabling me to achieve exactly the desired vocal quality.

Glover's personal perception of the work is not one of quartet and voice being intermeshed and inseparable, but rather that the quartet is "an entirely separate force playing their own material". (Glover, 2015)

In contrast, Mayo perceives the hierarchy of the work as more in favour of the voice, with the quartet providing accompaniment:

> The name of each moth is accompanied by a single chord or gesture from the quartet. Each chord or gesture provides a subtle halo, illuminating some tiny speck of detail from the image evoked by each name. (Mayo, 2015)

4.2 *Freezywater* (2016)

Freezywater is a second work for pre-recorded voices and a live instrumental ensemble (in this case, string quartet, piano, reed organ and hand bell), and is an extension and variation of the ideas in *The Indistinguishables*: employing a list text and litany structure, with individual utterances and musical gestures separated by fermatas.

Text source and structure

The work again takes a quasi-curatorial approach to the selection of found text; the material is a list of place names and topographical features (settlements, street names, rivers, woods, etc.) found in an approximately sixty-kilometre diameter circle around Outer London, roughly following the route of the M25 London orbital motorway.

Rather than impose a secondary level of semantic content on the list items (as per the allusive groups of names in *The Indistinguishables*), the place names are presented in a

geographical sequence, moving from the Northernmost point, clockwise, to the Southernmost point and back again. The piece ends with a reiteration of the first list item (and title), "Freezywater".

Transferal of visual arts aesthetics: structuralist film and psychogeography

The use of geographical names as the text source anchors the piece to the overarching subject of "place" (i.e., physical location, environs). This opened the composition of the work to the transferal of several formal and generative extra-musical ideas, the first of which was use of the situationist *dérive* technique. The specific focus on the M25 motorway in *Freezywater* suggested itself to this kind of approach, due to the attention paid to it as a physical and symbolic boundary around London by UK psychogeographers, such as Iain Sinclair (in his book *London Orbital*, 2003) and the self-described practitioner of "deep topography" Nick Papadimitriou, in the documentary *The London Perambulator* (Rogers, 2009).

Guy Debord defines the *dérive* as a kind of unguided journey, undertaken by one or more people in the course of a day:

> Dérives involve playful-constructive behaviour and awareness of psychogeographical effects, and are thus quite different from the classic notions of journey or stroll.
> In a dérive one or more persons [...] let themselves be drawn by the attractions of the terrain and the encounters they find there. (Debord, 1958)

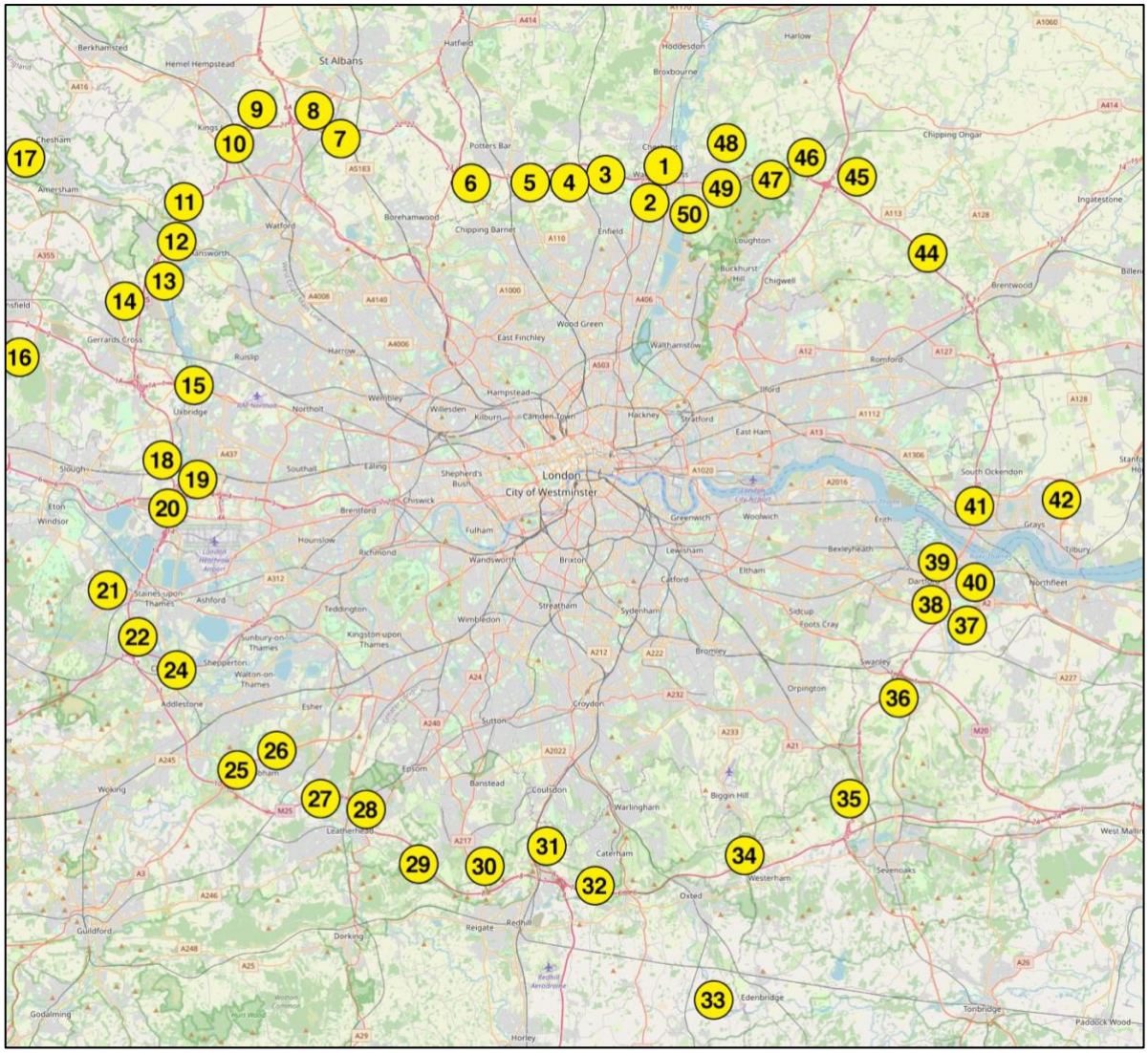

Figure 4: Map of locations surrounding London named in *Freezywater*

1. Freezywater	11. The Clump	21. Spring Rise	31. The Shrubbery	41. Dolphin Way
2. Unity Road	12. Mill End	22. Anchor Copse	32. The Conduit	42. Hangmans Wood
3. Cinder Ash	13. Maple Cross	23. [Traveller] *fictive	33. River Eden	43. [Lodestar] *fictive
4. Tilekiln Osiers	14. Copper Ridge	24. Stepgates	34. Wet Wood	44. Horseman Side
5. Pond Wood	15. Grand Union	25. The Broom	35. Star Hill	45. The Rough Patch
6. Dancers Hill	16. Happy Valley	26. The [River] Mole	36. Ashen Bank	46. Bell Common
7. Smug Oak	17. Devil's Den	27. The Mounts	37. Lane End	47. Sunshine Plain
8. Spielplatz	18. Old Slade	28. Uplands	38. Dark Trees	48. Winters Way
9. Long Spring	19. Saxon Lake	29. Herring Grove	39. Stone	49. Gunpowder Park
10. The Unicorn	20. The Island	30. Babylon Lane / Majority Plantation	40. High Trees	50. Tumbling Bay

Table 3: *Freezywater* location placenames, numbered in correspondence with Figure 4

Analogously, the list-text for *Freezywater* was generated by a dérive-like examination of physical and online maps (rather than a physical journey, as envisaged by Debord), looking for attractions of the text rather than "attractions of the terrain". The resulting map of named locations, plotted in OpenStreetMap,[36] can be seen in figure 4.

In addition to the forty-eight real placenames, the text also includes two extra words ("Traveller", name no. 23 and "Lodestar", name no.43), which were added as mysterious, poetic, fictive elements.

Freezywater also engages with the structural and material tropes of experimental film to generate its formal architecture. Specifically, the piece draws on two examples of British avant-garde film, John Smith's *The Black Tower* (1985-87) and Peter Greenaway's *Vertical Features Remake* (1978), both of which employ a highly restricted palette of filmic material in some way: in Greenaway's case, a strictly ordered sequence of static shots of natural and man-made verticals, arranged according to temporal grids. In Smith's case, the illusion of multiple locations—and states of mind—created through diverse shots in multiple angles of a single geographical feature: the sinister, titular black tower, which was actually a "found" location, standing on the site of the now closed Langthorne Hospital in Leytonstone, London (Walker, 2010).

Vertical Features Remake is both a demonstration and parody of the UK "structuralist film" movement, which grew out of the work of US filmmakers such as Tony Conrad, Paul Sharits and Hollis Frampton, whose work was identified as "structural film" by P. Adams Sitney:

[36] OpenStreetMap is an open-source mapping project, available at https://www.openstreetmap.org/.

> Theirs is a cinema of structure wherein the *shape* of the whole film is pre-determined and simplified, and it is that shape that is the primal impression of the film [...] The structural film insists on its shape, and what content it has is minimal and subsidiary to the outline. (Sitney, 1969)

This minimalism of content throws the structural devices of a film to the fore. In the words of Bart Testa:

> The term [structural film] serves as a generalization to mean that Greenaway foregrounds formal systems as constructs rather than imbedding films within mimetic ends, like storytelling. (Testa, 2008, 80)

Freezywater explores ways in which to map Greenaway's structures to a musical work. In Greenaway's film, sequential shot lengths increase or decrease by a single frame. Analogously, the musical material of *Freezywater* is subject to a process of gradual augmentation (a single note is gradually augmented to a phrase of music) and subsequent diminution till the end of the piece (ultimately, again, a single note). In this way, the work "insists on its shape" through a clearly audible approach to material and structure.

Following the example of *The Indistinguishables*, the harmonies of *Freezywater* are intended—at least at the outset of the work—to provoke a sense of familiarity, reflecting an object-like approach to the material. In the same way that Greenaway's vertical features comprise images of recognisable objects in a landscape (trees, the stems of flowers, fence and telegraph poles, equestrian show jumps, etc.), the fundamental musical material of *Freezywater* is a recognisable harmonic object: a C major triad, unsettled by an added second. The intended sense of familiarity was inferred by Richard Fairman, reviewing the first performance, who wrote that, "the instruments sank into deep, warming C major chords,

as if coming home" (Fairman, 2016, emphasis mine). As the work progresses, this chord is augmented into a rocking figure, before the apparently unyielding constancy of the C major harmony is undermined by a flattened supertonic anacrusis in the piano (a "wrong note"), and the harmony gives way to an ambiguous (through lacking a third) D minor/major 9th harmony. In this way, the C major "object" is gradually *made strange*, analogous to the sensation of Greenaway's everyday vertical objects becoming increasingly dislocated and unreal as his film progresses, through his use of montage and rhythmic edits.

This use of a highly restricted palette of material, and the use of a gradual process of musical development is additionally intended to be reflective of the experience of a motorway journey, in which the visible landscape changes as one travels, but certain features (i.e., the road itself, street furniture, the sky) remain constant.

Text-music relationship

In contrast to the Type E (speech rhythm) relationship of *The Indistinguishables*, *Freezywater* employs the cognate Type C relationship. In this way, the treatment is simpler, reflecting the pared down character of the structuralist model. The structural relationship is very strict: each iteration begins with utterance of a place name, accompanied by a sustained reed organ dyad or triad, after which the strings, piano and (later in the piece) the hand bell play their phrase.

Perception of speaking voice: captions as recontextualisation

This use of the Type C relationship draws attention to the voice as a captioning device. Since each iteration of the music is so similar to that which precedes or succeeds it, the voice here has the effect of "recontextualising" the music. In this way, *Freezywater* shares the perceptual effect of Arthur Russell's "The Name of the Next Song" (chapter 3.2), in that the text captions might affect the listeners' programmatic perception of the music.

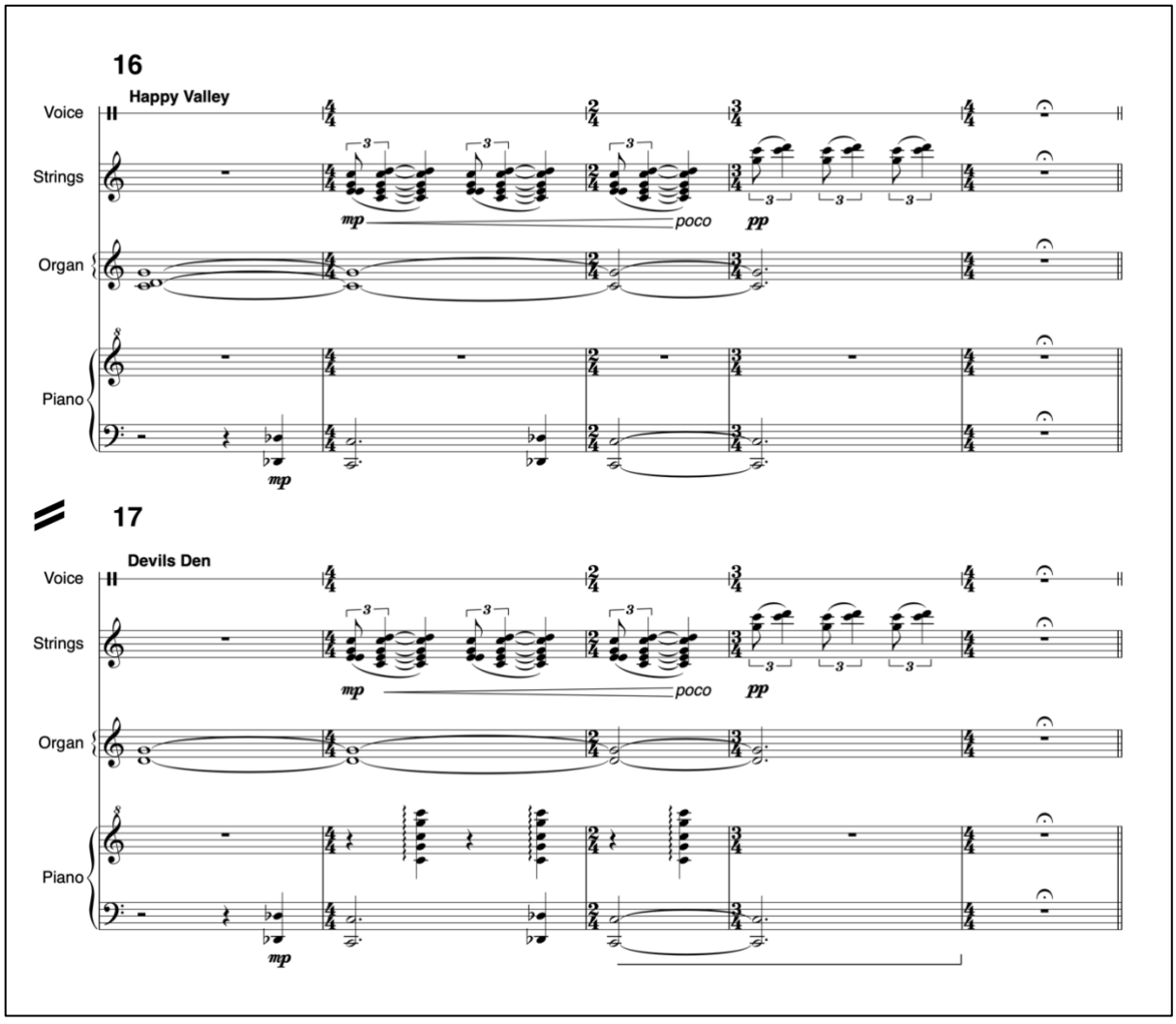

Figure 5: Comparison of phrases 16 and 17 in *Freezywater* (short score)

For example, two consecutive place names are “Happy Valley” and “Devil’s Den” (the short score of this passage is shown in figure 5); although the musical material for these is extremely similar (only differing in the addition of an arpeggio in the piano part), might the listener perceive the music very differently? Do the emotional and cultural associations of the place names psychologically colour the musical material?

Again, the effect of “voice as caption” is heightened through the specific use of a neutral and non-dramatic intonation and inflection in the vocal performance, although this time the voice is my own, adding a reflexive aspect to the work and pointing towards the more self-portrait-like work later in the portfolio (*The Subject / The Object*, chapter 6).

4.3 *Affix Stamp Here* (2016)

Affix Stamp Here is a work for live vocal ensemble (four or more singing voices) and pre-recorded speaking voice. It is the third work in the portfolio to use a disembodied speaking voice and a list text, and again employs a curatorial approach to found text. It differs from *The Indistinguishables* and *Freezywater* in that disembodied speaking voice and physical singing voices are juxtaposed, and their functions within the work are clearly delineated: more than in the previous two works, this draws attention to the capability of the disembodied speaking voice to be perceived as “caption”.

Text source and structure

The text of the work is drawn from verbatim messages transcribed from picture postcards sent by twentieth century holidaymakers. The writing of the work was therefore

preceded by a pre-compositional phase of the acquisition of a large number of second-hand postcards (through second-hand shops and online). Over five hundred postcards were catalogued and transcribed, then subsequently reduced to a final thirty "scenes" through editing the texts. The curatorial selection of these was carried out through identifying succinct narratives, good geographical and chronological spread (the selected texts date from 1944-2001) and textual signifiers of the (anonymised) writers' personalities and motivations.

This choice of verbatim text for a sung work, rather than a literary text or dramatic libretto, enables the work to engage with its key themes of nostalgia, notions of identity, place and culture, while drawing attention to itself as a literal example of found text.

Transferal of visual arts aesthetics: audio testimonies

The use of found postcards makes an obvious connection to the curated collections of postcards by Susan Hiller, but also to audio-visual works of Hiller's such as *Witness* (2000). This work comprises the recorded audio testimonies of hundreds of people who claim to have witnessed a UFO, exhibited as an installation of a field of suspended loudspeakers. *Affix Stamp Here* also reflects my own experience of performing Jennifer Walshe's *A Folk Song Collection* (2011), which comprises very brief song settings of apparently overheard conversations. My work echoes both these sources as examples of first-person accounts, since the postcard texts inevitably chronicle personal experiences.

Text-music relationship and perception of the speaking voice

Affix Stamp Here sets each postcard text as a freely composed song for a solo singer, answered momentarily by an interjection by the remaining ensemble of a single word. Highlighting an object-like approach to text, each song is accompanied by a dyad drone, alternately of a major 2nd, minor 3rd, major 3rd, perfect 4th, perfect 5th, minor 6th, major 6th and minor 7th.
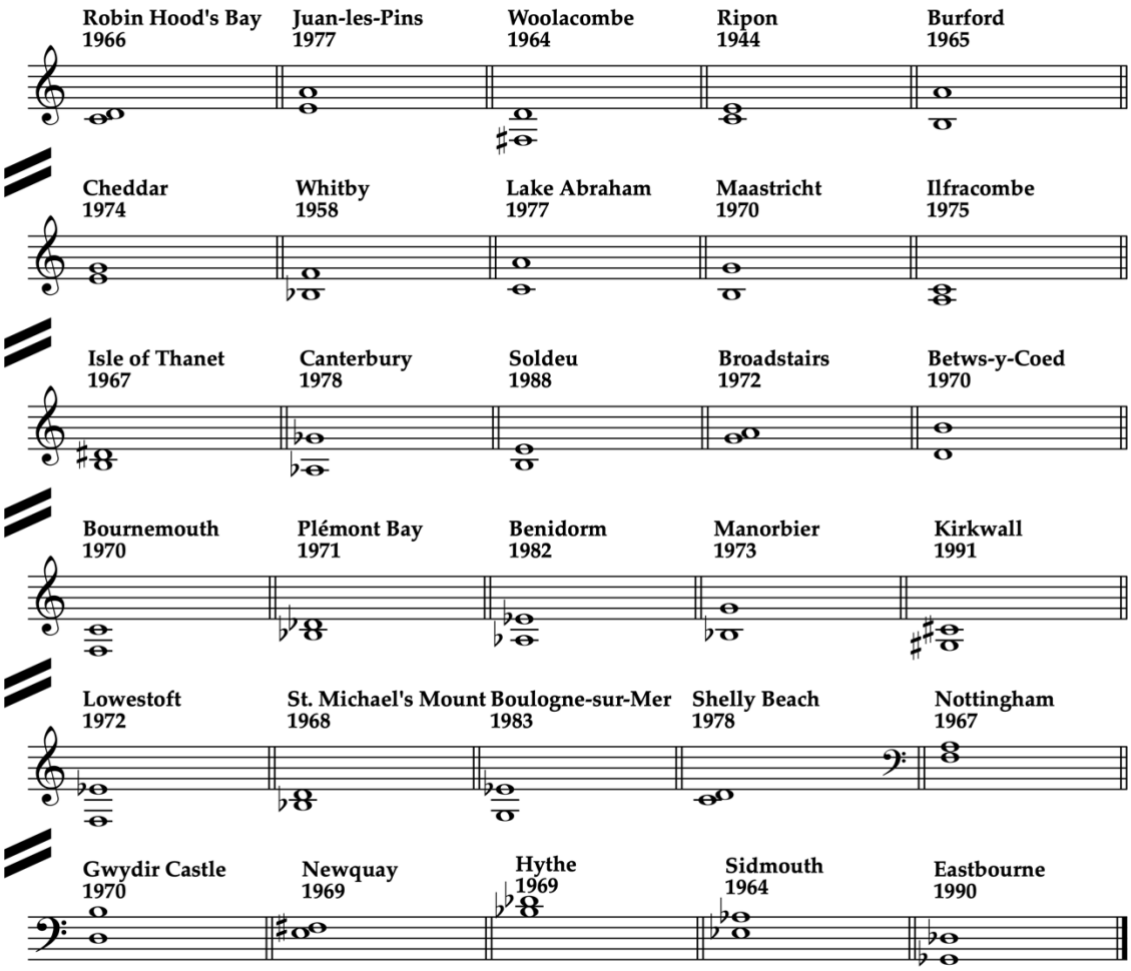
Figure 6: Structure of *Affix Stamp Here*: drones and spoken intertitles

The intention of these drones is to reflect “moments frozen in time”, so each song is perceived by the listener as discrete, rather than as part of narrative though the work. The complete structure of the work can be seen in figure 6, which maps the drones and spoken intertitles.

Harmonically, these dyad drones are arranged so no adjacent two form simple triads, tonic chords, or form cadential pairs. This breaks any sense of traditional musical narrative, since no drone leads harmonically to the next. The choral interjections also avoid reinforcing any harmonic centres that might be implied: they never double any notes of the drones (except for the very final chord of the work: “Remember”), and the vocal harmonies form extended 7^{th}, 9^{th} and 11^{th} chords in inversion, when heard against the drones. In this way, the work lacks any sense of harmonic arrival throughout, so no single postcard text is lent any importance over the others: this emphasises that this is a sequence of narrative fragments, rather than a work with a through-narrative, as might be implied by use of functional harmonic progressions.

The disembodied speaking voice introduces each song with the location and year of the postcard, in a Type C relationship (akin to *Freezywater*), entering with the change of drone, thus delineating the structure of the piece. In this way, the speaking voice sets the scene, in the manner of a cinematic intertitle. In the initial, workshop performance of the work, the music was combined with genuine intertitles on two video screens during the spoken captions, and images of the postcards during the songs. This visual element was removed for subsequent performances, as the perception of the performers and audience was that this very literal use of the source material made the work less interesting, through making the implicit explicit.

The perceived roles of the physical singing voices and disembodied speaking voice are clearly differentiated in this work: the solo singing voice is diegetic, narrative and first person; the speaking voice is non-diegetic, non-narrative and third person. The choral voices are a third narrative layer, providing a commentary through highlighting a single "keyword" from each postcard text, which is ambiguous in its diegetic relationship. The perception of the speaking voice in this work hinges on our cultural acceptance of the benign omniscient narrator voice, familiar from both cinema and radio, opening my line of investigation into openly dramatic uses of the disembodied speaking voice in the works of the following chapter.

Chapter 5:
Second approach: found text (2): radiophonic and narrative character voices and documentary texts

The second line of approach (and thus the second group of works comprising the portfolio) examines a radiophonic approach to the voice in a musical context. This is not to say that the use of list texts, litany structures and the notion of the caption voice from the first approach (chapter 4) are abandoned entirely: they become one compositional tool among a more diverse set. The works in this category are the studio work *Red and Blue* (2015), and the ensemble works *Five Loops for the Bathyscaphe* (2018) and *Anticlock* (2019).

As seen in in chapter 2, the voice in radio exists in a multiplicity of modes. The roles assumed by a voice on the radio might be those of news reporter, DJ, continuity announcer, interviewer, interviewee, storyteller, voice actor, and more, and the inherent disembodiment of these voices is culturally accepted, rather than being perceived automatically as disturbing or uncanny. The modes and roles of the radiophonic voice may switch quickly and seamlessly, or exist simultaneously (imagine, for example, a typical news report on a political interview, in which reportage and archive material are superimposed).

Approaching the disembodied voice with these radiophonic vocal roles and tropes in mind suggests the presentation of a multiplicity of textual content, interleaving narrative fragments, found text, documentary material and fictive writing. The radiophonic model also suggests a possibility of adoption of a multiplicity of sonic materials, interleaving spoken word, song, music and sound design.

Narratively, the possibility of presenting verbatim text and found text alongside my own fictive writing informs the works *Red and Blue* and *Five Loops for the Bathyscaphe*, leading to slippage between fictive and factual material, following the example of experimental radio, such as the Neue Hörspiel works described in chapter 2.3.

This results in a freer use of my original texts in *Anticlock*, which ultimately leads to the exclusive use of original texts in The *Subject / The Object* (chapter 6), as the concept of the voice as self-portrait take precedence.

5.1 *Red and Blue* (2015)

Note: this section should be read in conjunction with the Appendix: Red and Blue*: table of text sources, structure, voice-music relationships and treatments, which is the full annotated text of the work.*

Red and Blue is the first portfolio work to engage with the modes of the radiophonic voice, rather than the caption voice of the preceding works in chapter 4. It is a sixteen-minute fixed media work, assembled in the studio, rather than a work involving live musicians, or intended for live performance. In its form and function, it is a hybrid of electronic music album (in its sectionalised structure) and radio play (in its use of narrative fragments and vocal characterisation). It was released as a digital EP and broadcast on BBC Radio 3's *Late Junction* in its entirety.

Text source and structure

The theme of *Red and Blue* is the relationship between two significant international politicians of the 1980s: Margaret Thatcher and Ronald Reagan, and their so-called "special relationship" as Prime Minister of the UK and President of the USA, against the backdrop of the late Cold War period. It takes a curatorial approach to its text, which is assembled from a large number of publicly accessible, online documents, recording their written correspondence and their telephone conversations, together with passages from both politicians' recollections of each other in their respective autobiographies, and a list of the USA's nuclear weapons tests through the 1980s.

Much of the verbatim Thatcher-Reagan correspondence is in the public domain only in redacted versions, in accordance with US and UK security policies. This suggested a fragmentary, episodic form, in which the texts appear neither chronologically nor in a linear through-narrative. Thus, the texts appear as fleeting sonic images, in twelve brief sections, each of which is derived from a single piece of correspondence or, in the case of sections no. 5 and 10, from several pieces of correspondence, the texts of which are superimposed, or interpolated.

These source texts are sometimes presented verbatim, and at other times in my reflective paraphrases. My original (paraphrased) texts are generally written in the first person, to match the verbatim texts. Hence, the piece is an interleaving of "authentic" and fictive text, creating a multi-layered impression of the subject matter and material.

Text-music relationship

The twelve sections of the piece are highly varied in mood and non-vocal sound sources (including a wide variety of field recordings, synthesisers, drum machines, prepared piano, trombone improvisations, sampled brass instruments and sampled percussion). These sounds and musical episodes are derived from a "sibling" work: my score for artist-filmmaker Jennet Thomas's narrative short film *The Unspeakable Freedom Device* (2015). This film and *Red and Blue* share some thematic concerns: the plot of *The Unspeakable Freedom Device* concerns a quasi-mediaeval group of characters who journey to Blackpool's Winter Gardens to witness a meeting of a "cargo cult of Margaret Thatcher", in a "fantastic, primitive-future [cinematic] world, [where] the difference between technology and magic has become incomprehensible" (Thomas, 2015).

Red and Blue comprises music cues and sound effects composed for, but not ultimately used in the finished edit of the film, which are restructured and "recycled". In this way, the collage of sonic material parallels the collage of text fragments. Within these sonically varied sections, the text-music relationship is varied even further, moving fluidly between Type B (coexistent), Type C (cognate) and Type D (chant) and song. Additionally, the spoken vocals are subjected to treatments through delay effects, multitracking (several layers of voice added in the studio), and through recordings made with the use of Dictaphones. The voices are also placed freely within the stereo field. These variations occur from section to section or, more frequently, from line to line of text (see the appendix for details of how these variations are applied to the text).

Perception of speaking voice

This highly changeable use of text-music relationship and vocal treatment is intended to reflect the dialogue between two or more voices (i.e., those of Thatcher, Reagan and my own), even though the only recorded voice is mine and no archive recordings are used. Sometimes, this is a relatively representative treatment, such as in Section 7, where the two sides of a transcribed telephone conversation appear as left and right stereo channels, but generally the treatments are non-realistic. Additionally, the constant variation is intended to be analogous to a printed text that is marked up with a multitude of formatting and typographical emphasis: italic, bold, underline, changes in typeface, etc.

In contrast to previous works in the portfolio, such as *The Indistinguishables*, in which the neutrality of the tone of voice indicates the caption voice, in *Red and Blue* the voice is highly characterised and expressive in its variety of tone (adopting the modes of a radio voice actor), including passages that move between whispering, speaking, shouting and song.

This multiplicity of vocal modes is reflective of a radiophonic mode of perception, influenced by collage-type works of Neues Hörspiel. Although the format of the work is a monologue, my voice here represents characters who are both intradiegetic (through the verbatim material) and extradiegetic (through my inserted original material and paraphrase material). The rapidity of the shifts through different kinds of vocal material is intended to be disorienting, analogous to the scenes in Nicolas Roeg's film *The Man Who Fell to Earth* (1976), in which the character of Thomas Jerome Newton (an alien, played by David Bowie), watches a wall of television screens simultaneously displaying different channels. This

reflects my own childhood experience of viewing Thatcher and Reagan on broadcast news as "televisual spectres", not fully understanding the significance of these two figures.

5.2 *Five Loops for the Bathyscaphe*

The multivalent, collage approach of *Red and Blue*, in which disembodied voices might be perceived as radiophonic documentary voice actors, informs the succeeding portfolio work, *Five Loops for the Bathyscaphe*. Here, however, the use of radiophonic voices is applied to a work intended for live performance, rather than a purely fixed-media work: it is scored for a piano trio (piano, violin, cello), with pre-recorded speaking voices. *Five Loops* also adopts a more restrained approach to material, with no use of vocal treatments or theatrical vocal delivery, a sparser use of spoken text, and musical material that is considerably more understated and slower in tempo, resulting in a work that is atmospheric in tone, rather than overtly dramatic.

Text source and structure

Again, *Five Loops* is based on paraphrased documentary material: in this case, the story of oceanographer Jacques Piccard's pioneering descent to the bottom of the Challenger Deep in the Mariana Trench in the Western Pacific Ocean (the deepest known part of the ocean, at a little over 6.8 miles down). Together with fellow oceanographer Don Walsh, Piccard made this journey inside a US Navy-owned vessel, the bathyscaphe *Trieste*, on 23 January 1960.

Piccard's account of the experience was published as a reflective, journalistic article entitled "Man's Deepest Dive" in the August 1960 issue of *National Geographic* magazine. This article forms the basis of the text of *Five Loops* but, in the same manner as *Red and Blue*, the original text is paraphrased and fragmented, to create an elliptical, mysterious impression of the source material, rather than being a straightforward verbatim setting. For example, the following is an excerpt of Piccard's text:

> Outside the water is magnificently limpid—no trace of life, no plankton. Switched on for an instant, the searchlight casts its rays deep down beneath the bathyscaph [sic]; it seems as if nothing can stop this light. We are in the void, the void of the sea, which squeezes upon the wall of our sphere with nearly 150,000 tons. (Piccard, 1960)

This paragraph is paraphrased in *Five Loops* as a single short phrase: "Searchlights trailing the limpid water". Elsewhere in the piece, the text comprises a single word or phrase taken from Piccard's article (or example, "horizonless"), and only the penultimate utterance is a direct quote, "I feel as if these hundreds of gallons of water are passing into my veins".

Besides the paraphrases of Piccard's article, the text of Five Loops comprises several other strands of material:

- The typical chemical composition of sea salt ("Sulphate, 8%... Magnesium, 3%... Calcium, 2%...").
- The scientific names of families of deep-sea, bioluminescent fish ("Melanocetidae... Gonostomatidae...").
- Countdowns of depths in metres, from 200m to 6000m.
- The names of the five pelagic zones (the layers of the oceanic water column, by increasing depth): Epipelagic Zone, Mesopelagic Zone, Bathypelagic Zone, Abyssopelagic Zone, Hadopelagic Zone.

The latter of these text elements, the five zones, also provides the impetus for the five-section structure of the work, and their five different kinds of musical material. This material is based on non-developmental, repetitive, or "static" (i.e., drone) material. This has a programmatic intent, reflecting the preternatural stillness of the deep ocean environment. The general tessitura of the work also descends as the work progresses, programmatically illustrating a journey downwards through the first four sections, before the strings glissando to the top of their ranges in the fifth, illustrating a return to the surface (see figure 7).

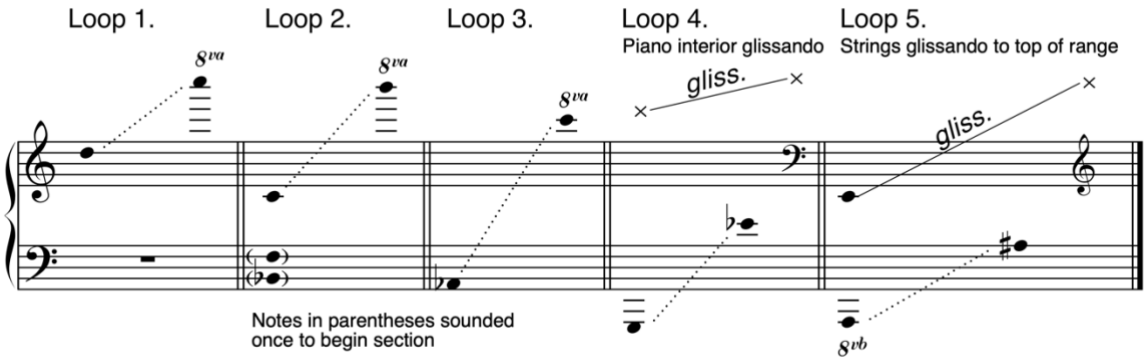

Figure 7: Tessituras of each section of *Five Loops for the Bathyscaphe*

Within these sections (loops), the music material is differentiated as follows:

- Loop 1: strings play sustained artificial harmonics; piano plays arpeggios approximating harmonic series.
- Loop 2: all instruments play descending legato lines.
- Loop 3: piano plays disjunct chorale material; strings play "flurries" of pizzicato.
- Loop 4: strings hold drone chord; piano plays gestural material on interior.
- Loop 5: piano plays pulsing chord; strings tremolo glissando upwards.

Text-music relationship

The spoken text and musical material played by the trio are presented in a combination of coexistent (Type B) and cognate (Type C) relationships. In contrast to *The Indistinguishables* and *Freezywater*, in which the live musicians cue the recorded voices with foot pedals or through cueing software (QLab), in *Five Loops*, the playback of voices is a continuous audio file (with timestamps marked in the score), which provides audible cues for the musicians. In this way, the voices act not just as vessels for transmission of the text material, but also as timekeepers. A clear example of this can be seen in the first section of the work, wherein the voices cue the timing of the piano's gestures (see figure 8).

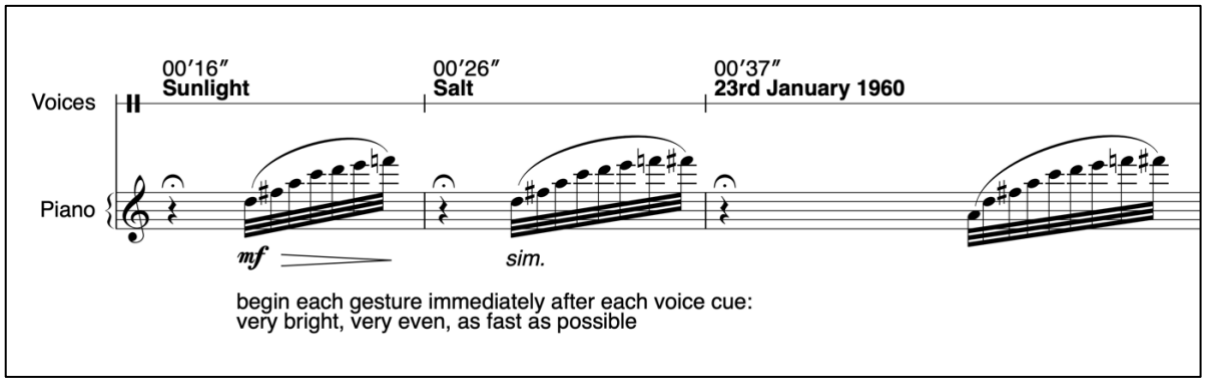

Figure 8: Voice cues for piano gestures in the opening section of *Five Loops for the Bathyscaphe*

However, this relationship is not rigorously maintained throughout the work. Rather, the instrumentalists are required to alternately observe the cues provided by the voices and play in a coexistent relationship (that is, "ignore" the voices). In the example above, while the piano strictly follows the voices' cues, the strings play in their own tempo. This variety of independent and dependent relationships between voices and instruments, and also between

the instrumentalists themselves (sometimes all three musicians play in their own independent tempi), creates a more fluid, drifting impression of the material, which is radiophonic in its collage of texts and music. This approach is in opposition to the strict, grid-like structures and consistent text-music relationships of the earlier works (chapter 4). Table 4 demonstrates the changing relationships from section to section in *Five Loops*:

Loop 1.	• Violin and Cello play independently (Type B relationship with voices). • Piano follows voice cues (Type C relationship). • All three listen for final cue (Type C).
Loop 2.	• All three instruments play in independent tempi (Type B). • All three listen for final cue (Type C).
Loop 3.	• All three instruments follow voice cues (Type C), although each plays musical gestures in their own tempi.
Loop 4.	As in Loop 1: • Violin and cello sustain chord (Type B). • Piano follows voice cues (Type C). • All three listen for final cue (Type C).
Loop 5.	• All three are in the same tempi, ignoring vocal cues (Type B).

Table 4: Section-by-section speech-music relationships and instrumental coordination in *Five Loops for the Bathyscaphe*

Perception of speaking voice

Five Loops highlights its use of voices as radiophonic by employing two voices (the actor Gemma Saunders and myself), suggesting the possibility of a representation of the two

protagonists descending in the bathyscaphe (Walsh and Piccard, as described in Piccard's account of the journey) and the possibility of a dramatic dialogue.

However, the voices are deliberately abstracted: Saunders and I do not assume characters, or overtly dramatise the text in our performances, in the manner of *Red and Blue*. The fragmentation of the text (including lists and statistics alongside the documentary material), avoidance of passages of dialogue between the voices, and the neutrality of the tone of voice used in the work suggests a hybridity with earlier works (those in chapter 4). The voices are (n)either character (n)or caption, and their functionality as timekeepers may not be immediately obvious to the listener. It may be that the music is perceived as a soundtrack to the text, or the text as captioning the music. Again, this reflects the impressionistic intent of the work, which prioritises an atmospheric, programmatic aural quality over the systematic construction of the works in chapter 4. This appears to have been conveyed successfully to some listeners, such as critic Fiona Maddocks, who described the work with adjectives in accord with the "mysterious" atmosphere intended: "*Five Loops* recalls the treacherous descent to the deepest ocean floor, dark and near freezing [...] eerie, almost motionless harmonics, are offset by ghostly flourishes at the top of the keyboard and low, muffled booms from the bottom, with a few unearthly plucked sounds from inside the piano: in all, mysterious, floaty, chilling" (Maddocks, 2018).

5.3 *Anticlock*

The concept of the disembodied speaking voice as audible timekeeper in an ensemble work destined for live performance is fundamental to *Anticlock*, which extrapolates this idea from *Five Loops*, employing the speaking voice as a presence that the ensemble alternately

follows, disregards, or contradicts. In this way, the voice is not only functional (it acts as unseen, yet audible "conductor" of the live musicians), but also becomes an antagonistic dramatic character. *Anticlock*, therefore, engages directly with the notion of the acousmêtre, since the voice presents itself as an apparently omniscient, sometimes omnipotent narrator that literally *tells the ensemble what to do*.

The work is scored for an amplified, mixed ensemble of nine players (the line-up of the commissioning ensemble, Decibel: recorders, bass clarinet, baritone saxophone, trombone, percussion, piano, electric guitar, bass guitar and violin)[37] and the pre-recorded voice is again my own. This timbrally incongruous instrumentation suggested a work that is more animated in character (brash, even) than *Five Loops* and those described in chapter 4, with a preponderance of loud material and fast tempi, but also more pluralistic in style and material, in the manner of *Red and Blue*.

The material of the piece also draws on the English Experimental composers' notion of musical "machines", or apparently rigorous systems music or procedural process music, as in certain works by John White[38]. The music therefore resembles an absurd parody of minimalist music and process music. This absurdism is also reflected in the choice of thirteen

[37] All the players also double on dog clickers. These are handheld devices, typically with a curved metal and plastic flange that makes a loud click when depressed. They have been used in positive reinforcement animal training since B.F Skinner's studies of animal behaviour in the 1940s but have been especially popular in dog training since the 1980s. In *Anticlock*, they fulfil two functions: firstly, they mimic the sound of a ticking analogue clock. Secondly, they are an object representing one entity telling another entity what to do, which is obviously a theme of the work.

[38] Several works of the 1960s and 70s by White use a compositional or performance process (akin to later minimalist music), or else employ "extended repetition as the sole means of generating movement" (Nyman, 1974, 8) and include the word "machine" in their title, such as *Cello and Tuba Machine*, *Knee and Desk Machine*, *Newspaper Reading Machine* and the well-known *Drinking and Hooting Machine*, in which performers gradually drink and "hoot" bottles of beer, resulting in music that gradually lowers in pitch.

short movements (each around one minute long), in reference to the 'Pataphysical Calendar of thirteen months in a year (i.e., an absurd marking of the passage of time).[39]

Text material and structure: the speaking metronome

Anticlock abandons both found text and documentary material in favour of an original text. However, this text is primarily the simplest possible use of language, limited to counted numbers in all but three movements. The other textual material in the piece is the alphabet (in the third movement, entitled "ZEE"), a list of rhyming, single syllable words (in the eleventh movement, "DICHOTOMIES") and a brief passage of spelled-out hidden messages in the ninth movement, "SERIES". Analogous to the use of numbers and letters, those basic building blocks of textual communication, the musical material is based around simple chords (major triads and dominant sevenths) and the chromatic scale, representative of a kind of sonic alphabet of equal-tempered tonality.

This focus on counted numbers reflects real-world examples of recorded voices that act as timekeepers: telephonic speaking clock services, vocalists guide tracks, and spoken count-ins used by musicians in a multitude of live and studio settings. Indeed, the work uses the recorded voice as a count-in in several movements, and also as a spoken click track, or speaking metronome. The absurd dramaturgy of the piece is to represent this voice "malfunctioning", so in certain movements it ceases to act as an obvious timekeeper, and instead the spoken numbers indicate chord inversions, partials in the harmonic series, pitches

[39] 'Pataphysics is the pranksterish "science of imaginary solutions" instigated by playwright Alfred Jarry. The Perpetual 'Pataphysical Calendar is a 13-month alternative system to the Gregorian calendar, devised by the Collège de 'Pataphysique in 1949. In accordance with the indecorous tone of Jarry's writing, the months bear names such as Haha, Décervelage ("Debraining"), Merdre (usually translated as "Pshit"; the first word of Jarry's play *Ubu Roi*, 1896), and Phalle ("Phallus") (Hugill, 2012, 21).

in the chromatic scale, or indeed simply become sonic material. In this way, the presence of the voice could be thought of as analogous to the architecture of buildings such as the Centre Pompidou in Paris (1977, designed by Richard Rogers and Renzo Piano) or the Lloyds Building in London (completed 1986, designed by Richard Rogers & Partners), with its elevators, pipework and service ducts on the building exterior.[40] In *Anticlock*, the temporal and metrical structures of the music are surface material: they are on-display and personified through the voice. An overview of the text material and structure of the work is shown in table 5.

Movement 1 "ENTRY"	
Text material: Spoken count-ins ("1, 2, 3, 4") in seven different tempos: (q =60, 120, 70, 140, 90, 180, 110).	Ensemble follows voice in tempo.
Movement 2 "ACCELERATE"	
Text material: speaking metronome, in 4, occasionally "malfunctioning" in bars of 3, 5, 6 and 7.	Ensemble follows voice in tempo.
Movement 3 "ZEE"	
Text material: Spoken alphabet.	Ensemble follows voice, but with timing that is "falling apart".
Movement 4 "VOICELESS"	
Text material: speaking metronomes (in stereo, whispered).	Ensemble follows voice in tempo, as additional vocalists.
Movement 5 "NON-RISSET" Note: title refers to the Risset Scale auditory illusion of constantly ascending tones.	
Text material: sung numbers (1-12), assigned to the notes of a chromatic scale.	Ensemble follows voice in tempo, in plain harmonies and canon.

[40] This "inside-out" style of 20th century architecture is sometimes described as "bowellism", with Rogers considered its most significant practitioner. Simon Sadler notes that the style is a "micromovement", that originated in the late 1950s with the architect Michael Webb, of the "hypothetical" avant-garde, British architecture practice Archigram. (Sadler, 2005, 23)

Movement 6 "DOG"	
Text material: speaking metronome.	Ensemble follows voice in tempo at first (doubling on dog clickers), but gradually accelerates independently.
Movement 7 "REALCLOCK"	
Text material: speaking metronome (continuous count up, gradually crossfading to clock sound).	Ensemble follows voice in tempo, with percussionist playing increasing subdivisions of beat.
Movement 8 "OFF (DJ SENSIBILITIES)"	
Text material: speaking metronome (with delay effect).	Ensemble follows voice in tempo (in canon), but never plays on first beat of bar.
Movement 9 "SERIES"	
Text material: spoken numbers 1-12 (non-repeating), naming partials.	Ensemble: violin solo: violin plays harmonic partials 1-12 of a G string, following voice (cognate relationship).
Movement 10 "VOICED"	
Text material: sung numbers 4-1 (non-repeating), naming inversions of a dominant seventh chord.	Ensemble: follows voice, with answering phrases (cognate relationship)
Movement 11 "DICHOTOMIES"	
Text material: rhyming single syllable words (chant).	Ensemble follows voice in tempo.
Movement 12 "PRINT-THROUGH" Note: "print-through" is a contact-transfer fault of magnetic tape, resulting in audible artefacts of other sections of a recording.	
Text material: multiple simultaneous speaking metronomes.	Ensemble: independent from voice (coexistent relationship)
Movement 13 "DECELLERATE"	
Text material: spoken numbers 1-99 (non-repeating).	Ensemble follows voice, but gradually decelerates independently (chant—coexistent relationship).

Table 5: Movement titles, text materials and voice-ensemble coordination in *Anticlock*

The text material in movement 11 ("DICHOTOMIES") is representative of the speaking metronome malfunctioning. In this case, the numbers are replaced by rhyming list

of words that continues to fulfil the role of timekeeper. This text (my own) is inspired by Bruce Nauman's text-based artworks that contrast oppositional or contradictory phrases, such as *Good Boy Bad Boy* (1985) and *One Hundred Live and Die* (1984). A sample of the text of the latter work, which is displayed in multicoloured neon, is as follows:

> COME AND LIVE COME AND DIE
> GO AND LIVE GO AND DIE
> KNOW AND LIVE KNOW AND DIE
> TELL AND LIVE TELL AND DIE
> (Nauman, 1984)

This format, of pairs of phrases, is transferred to *Anticlock* as pairs of single words. The strongly accented duple time metre of the movement (with instruments alternately playing off and on beats) is thus maintained, as if the disembodied voice were finding alternatives for repeatedly saying "one–two". The full text of the movement is as follows:

> One–Two, One–Two, One–Two, On–Off, Two–One, Two–One, Two–One,
> Off–On, One–Two, One–Two, One–Two, Left–Right, One–Two, A–B, Yes–No,
> Fast–Slow, In–Out, Back–Front, Up–Down, East–West, High–Low, To–Fro,
> There–Here, Far–Near, North–South, Loud–Soft, Branch–Root, Source–Mouth,
> Live–Dead, Say–Said, Hot–Cold, Dry–Wet, Love–Hate, Walk–Wait, Push–Pull,
> Eat–Ate, Sweet–Sour, Salt–Sharp, Plane–Train, Drive–Park, Day–Night, Dim–Bright,
> All–None, Wrong–Right, Kling–Klang, Hush–Bang, Can't–Can, Ape–Man,
> This–That, Steep–Flat, Food–Drink, Swim–Sink, Dark–Light, Black–White,
> Red–Blue, When–Who, Me–You, Thou–Thee, Her–She, Him–He, One–Pair,
> Boy–Girl, Butch–Femme, What–Where, Thin–Thick, Well–Sick, Crawl–Run,
> Stand–Sit, Foot–Head, Rope–Thread, Point–Line, Yours–Mine, Old–New,
> Short–Tall, Big–Small, Most–Few, Pre–Post, Now–Then, Read–Write, Pad–Pen,
> Fire–Ice, Once–Twice, Dogs–Cats, Rats–Mice, Whole–Part, Life–Art, Head–Heart,
> Dumb–Smart, Catch–Throw, Kick–Punch, Fight–Flight, Reap–Sow, Cut–Heal,
> Rip–Mend, Take–Lend, Straight–Bend, Pro–Con, Bad–Good, Will–Won't,
> Shan't–Should, Sun–Moon, Base–Top, Wake–Sleep, Go–Stop.

The hidden messages in movement 9 (“SERIES”) are three lines of original material and the title of the piece. These are referred to in the score as “Alien Ciphers”:

M.E.N. / A.N.D. / W.O.M.E.N. / S.I.N.K.I.N.G. / S.L.O.W.L.Y.
W.H.O. / I.S. / I.N. / C.O.N.T.R.O.L. [?]
T.H.E. / F.O.R.G.O.T.T.E.N. / O.N.E.S.
A.N.T.I.C.L.O.C.K.

The first of these cryptic phrases, “men and women, sinking slowly”, is a recurring easter egg (a term for a hidden message, or hidden feature, arising from software development) that also appears in several other of my works: the songs “Houseboats” (from my album of dream transcriptions, *Trainwreck / Raincheck*, 2007) and “Dumb Terminal” (from the album *Everything / Everything*, 2008), and in truncated form in the text of *Five Loops for the Bathyscaphe*.

The impulse to include these messages, and the reason for the appellation “Alien Ciphers” is to highlight another perception of the voice in *Anticlock*. This is the tenor of science fiction: a speaking computer, speaking in code, with a possibly malevolent consciousness. The use of the cryptic voices at what might be the eeriest musical moment of the work (the sonically fragile and technically unstable violin harmonics, accompanied by “shimmering” cymbal) is intended to suggest an unknown or unknowable character.

Text-music relationship

The count-ins and speaking metronomes, used throughout much of the work, are rhythmically spoken, in order to be functional. As a result, these passages exhibit a chant

relationship (Type D) to the music. However, the usual hierarchy is inverted: rather than a physical voice performing in time to the music, here the disembodied voice is carrier of the temporal and metrical structures, which the live musicians must follow.

As seen in table 5, at certain points in the work the musicians move away from the voice's tempi into an independent tempo of their own (movements 6 and 13). The Type D (chant) relationship here degrades into a Type B coexistent relationship. This demonstrates how easily (and swiftly) one type of relationship can move to another, through the simple device of timing drifting apart (or together).

Perception of speaking voice

The effect of *vox ex machina*, the talking machine, or talking computer (chapter 2.4) is enhanced in many movements throughout the work, through use of a neutral tone of voice. The disembodiment is clearly essential to this effect of an automated voice, recalling the disembodiment of (e.g.) the speaking clock; the work would be perceived quite differently if the music were coordinated by a physical vocalist.

The impression of this machine "malfunctioning" is similarly characterised by tone of voice in movements such as "ZEE" (movement 3) and "DECELLERATE" (movement 13), where the vocal performance becomes aggressive and theatrical. This links *Anticlock* to Nauman's acousmêtre in *Get Out of my Mind, Get Out of This Room* (as described in chapter 2.5). Nauman's voice here is perceived as potent (if not omnipotent) through its aggression, and omniscient in that it "sees" the visitor, while we remain certain that the voice is automatic: it is that of a machine, since it draws attention to itself as a recording (Nauman's voice is audibly a recorded loop). Similarly, the voice in *Anticlock* signposts its artificiality:

in movement 4 (“VOICELESS”) the voice is doubled in stereo: clearly the result of studio manipulation, while in movement 7 (“REALCLOCK”) the voice morphs (through use of a studio cross-fade) into the sound of a ticking clock.

When the voice of *Anticlock* shifts from its functional role, with its neutral voice, into what seems to be a character, or narrative acousmêtre, with its emotive tone of voice, this provides the theatrical tension of the work. The voice appears to jump into a vocalic body, become personified, become character. The listener might ask themselves “who, exactly, is speaking?”. This ambiguity is compounded by textual material such as the “Alien Ciphers”, which suggest that an otherwise machine-like, seemingly automated voice might possess a consciousness, be supernatural, or extra-terrestrial, imbuing the work with the tropes of science fiction.

Factually, however, the voice of *Anticlock* is that of the composer. If the Ozian curtain were to be pulled back, I would be behind it. It was me all along: Leo Chadburn. In a work such as this, it is the composer guiding the music, and we hear that explicitly in *Anticlock*. It is only natural that the choice of voice would be my own. This autobiographical insertion of myself into the work, which may have been an unexamined shadow theme throughout works in the portfolio so far, becomes a key topic in the next work, *The Subject / The Object*.

Chapter 6:
Third approach: stream of consciousness and montage: notions of the disembodied voice as self-portrait

The third and final approach to use of the disembodied speaking voice in my portfolio springs from an exploration of stream of consciousness texts (generated through free writing and subsequent editing) and montage (that is, a splicing together of vocal recordings in the studio, to create seamless, monolithic forms). The result of this is represented in the portfolio by a single, extended work, *The Subject / The Object* (2020).

The is a fixed-media piece, released as an album. The work is impossible to perform live, due to the complexity of the studio edit, so it exists as a recorded artefact. Through the making of this work, and in reflecting on the contrasting use of the voice in earlier portfolio works, I developed a notion of *The Subject / The Object* as a form of audio self-portraiture.

The Subject / The Object comprises two tracks, each of exactly 20 minutes duration. The first (*The Subject*) is derived from recordings of my speaking voice, reading at speed, accompanied by backgrounded improvisations on spring drums and synthesizers; the second (*The Object*) is a drone piece, utilising only multitrack recordings of my wordless singing voice.

The two tracks are intended to be sonic mirror images of each other: words, unpitched noise and simultaneity (*The Subject*) versus wordlessness, pitched sound and harmony (*The Object*). It follows that the listener might also interpret these dichotomies as signifying programmatic or psychological opposites, for example exterior versus interior, night versus day, right brain versus left brain. The album was released digitally, but also on C40 cassette

tape. The availability of a physical version (I considered either a vinyl or cassette release) was part of the original planning of the realisation of the work, since a two-sided format emphasises the duality of the two tracks, each of which occupies one side of the tape (or would have occupied one side of a 12″, 33⅓ rpm vinyl record, if that format had been used).

6.1 *The Subject* (2020)

The composition of *The Subject / The Object* is fundamentally grounded in the conception of speaking voice as music: the Type F relationship. Specifically, it echoes ideas found in sound work by other artists who use overlapping recordings of voices to create dense beds of sound: the vocal montages by Ashley, Bennett and Gould described in chapter 2.3, in which the density and continuity of material is intended to produce the effect of linguistic meaning dissolving, highlighting the purely phonological aspects of the voice.

These works are contingent on studio editing for their effect. The pause-less edit of Ashley's *In Sara, Mencken, Christ and Beethoven There Were Men And Women* is obviously mirrored in the sonic textures of *The Subject* but, as I will demonstrate below, my original technique developed for the recording of my work creates a very different impression, through creation of an apparently seamless edit.

Text source and structure

The text for *The Subject* was initially generated through sessions of stream-of-consciousness free-writing. This is a technique whereby I write without pre-meditation,

allowing words and images to pass quickly through my mind.[41] These texts were then extended through a grid-like process of careful editing, using collocations of words matching the tone of the free-written texts and synonyms.

Although my initial impulse was to create a text that never repeated a word throughout the duration of the work (a *hapax legomenon*), this idea was ultimately rejected in favour of a cyclical text, a type of variation form with six different iterations of the text, in which the same semantic structures recur with half the vocabulary replaced from iteration to iteration. This structure is intended to create a sense of *déjà entendu*, in which textual images mutate throughout the work.

As an example of how the "text mutation" technique described above works in practice, table 6 shows the first line of the text of all six iterations. The nouns and adjectives that mutate in the subsequent iteration are marked with "↓changes". In general, the determiners, prepositions, and conjunctions are left unchanged from iteration to iteration.

As shown, the entire text of iteration 6 is constructed in such a way that it loops back to iteration one. In order to make this possible, the original iteration, which most resembles the unedited free-writing sessions, is iteration 3—the other iterations were built around this (extending the text *outwards from the centre*).

[41] The term "free writing" denotes the technique of writing quickly, without predetermined planning or structure. It was described as a literary technique as early as 1934 by Dorothea Brande, who suggests that one "write… until you feel you have utterly written yourself out" (Brande, 1934). In my case, the resulting texts resemble the forms of "stream of consciousness" literature, as exemplified by the writings of Virginia Woolf, James Joyce and Marcel Proust, but the sensation of writing in this manner also recalls the separate notion of "automatic writing", which is a parapsychological term applied to clairvoyants.

Iteration of Text:	**First Line of Text:**					
Iteration 1:	***"billowing*** ↓*changes*	***veil***	***of the***	***salute*** ↓*changes*	***to the***	***sphere"***
Iteration 2:	***"hazy***	***veil*** ↓*changes*	***of the***	***offering***	***to the***	***sphere"*** ↓*changes*
Iteration 3:	***"hazy*** ↓*changes*	***feel***	***of the***	***offering*** ↓*changes*	***to the***	***slide"***
Iteration 4:	***"grating***	***feel*** ↓*changes*	***of the***	***shuttle***	***to the***	***slide"*** ↓*changes*
Iteration 5:	***"grating*** ↓*changes*	***transit***	***of the***	***shuttle*** ↓*changes*	***to the***	***station"***
Iteration 6:	***"billowing***	***transit*** ↓*changes*	***of the***	***salute***	***to the***	***station"*** ↓*change*
LOOP: Iteration 1:	***"billowing*** ↓*changes*	***veil***	***of the***	***salute*** ↓*changes*	***to the***	***sphere"***

Table 6: Example of text mutation in the first line of all six iterations of the text for *The Subject*

The finished work is a reading of the entire text (Iterations 1-6) twice, with a final repetition of iteration 1 that fades out in the final minute, as below (table 7).

Timestamp	**Text**
00′00″	Iteration 1
01′39″	Iteration 2
03′15″	Iteration 3
04′42″	Iteration 4
06′21″	Iteration 5
07′59″	Iteration 6
09′34″	Iteration 1
11′16″	Iteration 2
12′48″	Iteration 3
14′18″	Iteration 4
15′51″	Iteration 5
17′26″	Iteration 6
19′05″	Iteration 1 (Fade Out)
20′00″	(End)

Table 7: Text iterations by timestamp in *The Subject*

The exchange of synonyms and collocations of words between iterations has the effect of maintaining an atmosphere, while any specific narrative meaning of the text is avoided. For the writer Tim Rutherford-Johnson, the effect is dystopian:

> While the monologue of *The Subject* appears at first to wander without form or direction, certain key words and phrases can be heard to return in varying contexts [...] Some of these returning words are modern and/or scientific, referring to data, isotopes, nylon, classification. Others, mixed equally within the same stream of consciousness, are ancient: obelisk, bronze, offering, hessian. The simultaneity of ancient and modern, or the vision of a future that has been remade as though ancient, calls to mind (or to my mind at least) the peculiarly English dystopian fiction of Alan Garner or Russell Hoban. This impression is enhanced by a third category of phrases that caught my ear, referring to the aftermath of some terrible event: sunken, destroyed, the sorry state of things, the town where we sheltered for the night, the dogs by the encampment. (Rutherford-Johnson, 2020)

It is not surprising that the tone of a book such as Russell Hoban's *Riddley Walker* (1980) is subconsciously reflected in *The Subject*: its setting in a quasi-Iron Age future provided direct inspiration for Jennet Thomas' *The Unspeakable Freedom Device* (see chapter 5.1), and I had previously read it several years before this project. The sense of dystopian unease in *The Subject* is also enhanced by the non-vocal element of the work: the accompanying improvisation with spring drums[42] and synthesizers (producing low tones and gong-like noises). This improvisation was added during the editing process, to envelop the voice in an understated, artificial acoustic, and is a simultaneity (that is, a Type B: coexistent

[42] This instrument is also known as a thunder tube (on account of the variety of low-pitched, thunder-like tones it produces): it comprises a spring fixed to a membrane, attached to a drum-like tube. It has been cited as an uncommon example of a shaken idiophone (rather than a membranophone), since the sound is typically generated by shaking the body of the instrument, causing the spring to move and vibrate chaotically, while the membrane and drum merely amplify and resonate the sound (Knight, 2016, 11). However, the spring can also be struck and agitated manually, as in *The Subject*.

relationship) alongside with the voice: it is not deliberately coordinated with the vocal recordings, except for a single moment of crescendo at the textual halfway point of the work (timestamp: 09′34″), i.e., at the beginning of the first repeat of iteration 1.

6.2 *The Object* (2020)

Material and structure

In contrast to the stream of spoken text in the preceding track, *The Object* consists of a multitracked choir of my singing voice, sustaining an E minor chord, which fades into a C major chord after 10 minutes (both in second inversion, to create a sense of harmonic tension, and both as close to just-interval tuning as my voice could sustain). The pitch material is arranged to cover close to the full range of my non-falsetto voice.

In deliberate contrast to *The Subject*—and in contrast to all other works presented in the portfolio—no words are used. Instead, the chords morph through vowel sounds, moving sequentially from closed vowels to open vowels [uː, ɛː, iː, ɔː, ɑː], with a fully closed [mː] acting as a linking sound. In the final moments, the vowel sounds morph into sustained sibilants [ʃː], as if suggesting a return to the unpitched voice and noise of *The Subject*.

The intention of *The Object* is to provide a foil for *The Subject*, offering a sense of relief or resolution, with obvious dichotomies between their materials. However, the two tracks share a novel recording and editing technique.

The Subject / The Object: recording and editing techniques

In *The Object*, each studio take of a tone within a chord was held for a full breath. In this way, the pitches waver fractionally as my breath support runs out and the imperfections of my voice become part of the surface activity of the music. No digital pitch correction or time-stretching was used, for this reason.

In the studio, takes of each note were recorded at overlapping time stamps, so in the edit cross-fades between them eliminated any breaths. This overlapping of takes is also employed in *The Subject*: here, the recordings of my speaking voice are staggered so the performance of each subsequent take begins halfway through the preceding one, each time with me speaking for as long as one breath allows, so the exact tone and pitch of voice could be maintained. These takes in *The Subject* are edited with extremely short crossfades on sibilant syllables, rather than between words. In this way, the seamlessness of the edit is maintained throughout.

As an example of how precise these edits were, in figure 9, the crossfade edit between two takes is shown (just before timestamp 00′45″). In this screenshot from Logic, with the text overlaid, it can be seen that the “s” of the word “beside” is the edit point. The whole crossfade is less than one frame long (i.e., a twenty-fifth of a second).

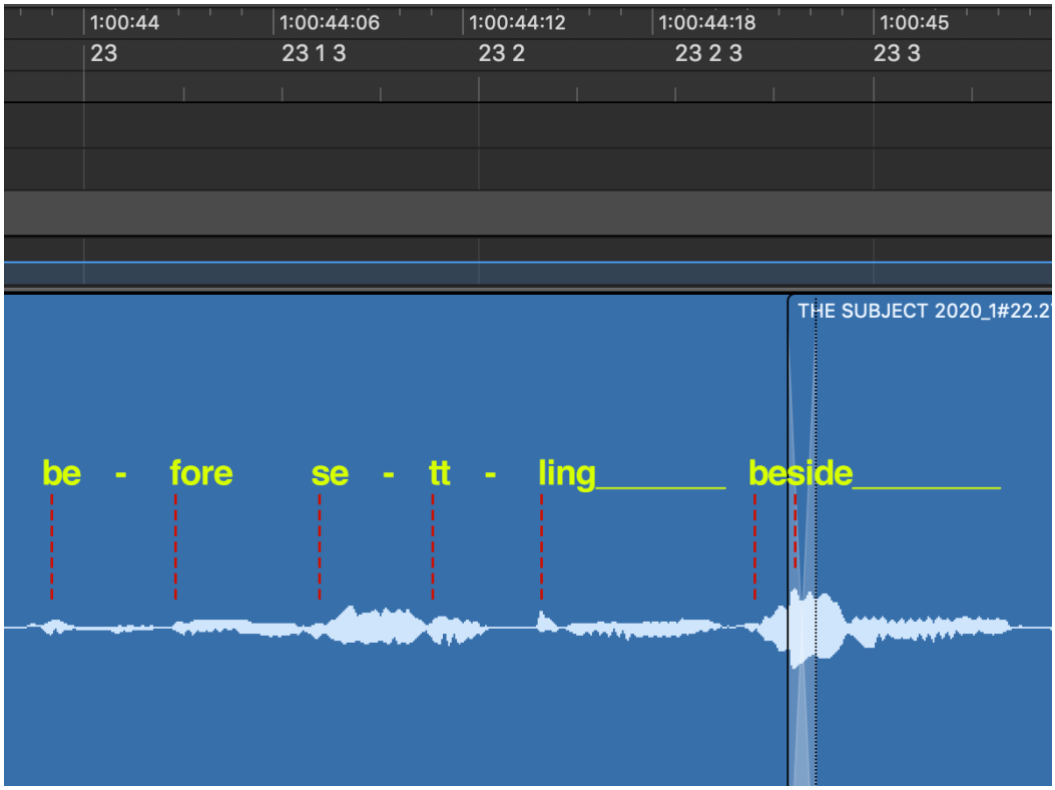

Figure 9: Example of crossfade editing on sibilant consonants in *The Subject*

In this way, all breath sounds (inhalations) are removed from both tracks. Inevitably, this editing technique also produces an uncanny quality in the work, through the unnatural, or apparently supernatural effect of having a voice speak without interruption.[43]

The exceptional moment is at the beginning of *The Object*, which commences with a single, theatricalised intake of breath. The effect of this is described by Rutherford-Johnson as follows:

[43] Anecdotally, several listeners reported to me that they found themselves unthinkingly, sympathetically holding their breath while listening to *The Subject* for the first time. Incidentally, before beginning work on *The Subject*, I had previously tested out this technique of pause-less editing of monologues in a short work entitled *Opinion Piece* (2016). I had written this for inclusion in composer Neil Luck's Resonance FM series of radio programmes, *Drivetime Underground*. *Opinion Piece* comprises transcriptions of radio phone-ins, translated and re-translated through Google Translate (https://translate.google.co.uk/), until their semantic meaning was lost, then performed by me and edited in a similar manner to *The Subject*. In this way, it is a prototype of the later work. An archive recording of *Opinion Piece* is available here: https://soundcloud.com/simonbookish/opinion-piece.

> Like the hinge between two panels of a diptych the breath joins *The Subject* and *The Object* and marks the division between them. It is both subject and object: the body of the singer and the material of the composition. (Rutherford-Johnson, 2020)

This perception of the breath as vocalic body of the artist recalls Emmerson's identification of Stockhausen's characteristic breathing as a quasi-reflexive presence (chapter 2.5), and similarly points towards an interpretation of *The Subject / The Object* as a work of self-portraiture.

6.3 Notions of self-portraiture: authentic and manipulated portraits

This notion of the work exhibiting the qualities of self-portrait emerged during the making of *The Subject / The Object*. Naturally, we associate the idea of the self-portrait primarily with the visual arts, rather than the non-visual, sonic realm of music and sound-art, so in this subchapter I will delve more deeply into my transferal of aesthetics from the visual arts.

Self-portraiture might be thought of as both bodily (an image of the artist's appearance) and psychological: an exposing of the subconscious through depiction of their image. Moreover, there exists a tension between apparently authentic aspects of self-portraiture (i.e. the artist attempting to create an accurate visual depiction of themself) and projected, or manipulated aspects (i.e. the artist either depicting an idealised, even heroic image of themself, or using their body as a metaphoric blank canvas). The artist Holly Marie Armishaw writes that this "projection" results from the "conscious element of control", which is present even "within the most candid and spontaneous works [of self-portraiture]"

(Armishaw, n.d.). That is, the artist's inevitable and numerous aesthetic and practical decisions in making a work of art.

These projected aspects of self-portraiture date back as far as Albrecht Dürer, the first prolific self-portraitist. In Dürer's 1500 self-portrait, he portrays himself in a pose that would have been immediately read as Christ-like by viewers at the time, with his monogram prominently displayed, reflecting his self-awareness as an internationally successful artist.[44]

In visual arts culture of the late 20th and 21st century, the notion of the artist's body as site of investigation has informed the work of artists such as Frida Kahlo, whose paintings are loaded with autobiographical, cultural and psychological symbols and signs. Kahlo's works have an obviously psychotherapeutic aspect of making the internal external, but the artist's body can also be used to project social or cultural subject matter, such as Samuel Fosso's theatricalised self-portraits as Martin Luther King and Angela Davis. Fosso's work is self-expressive, but also treats his own image as a an immediately available conduit for his intellectual concerns. The most notorious self-portraitist using their own body as a blank canvas in this way is Cindy Sherman, whose oeuvre almost solely comprises images of herself as stereotypical portrayals of fictive characters.

This tension between authentic and manipulated/projected image is more subtly demonstrated by a work such as Jeff Wall's photograph *Picture for Women* (1979): at first glance, the picture appears to be a candid, casual self-portrait of the artist and a model. However, we swiftly realise the photograph is a literal mirror image, in which the

[44] Woods-Marsden (1998) argues that the blossoming of the self-portrait as an art form during the renaissance reflects the change in social status of artists, together with a wider promulgation of philosophical ideas of self-examination and persona, while noting that earlier Western artists tended towards exclusively portraying themselves "in character" as heroic or mythological figures.

photographic apparatus (camera, tripod and shutter release cable) is foregrounded, drawing attention to our gaze on both model and photographer, and to the act of looking itself. Finally, we may realise that the picture is carefully posed, recreating the composition of Édouard Manet's *Un bar aux Folies Bergère* (1882). This authentic-projected-manipulated aspect of self-portraiture is even demonstrated by modern "selfie" culture, evidenced by the number of smartphone apps available to retouch, enhance or animate photographs.

Self-portraiture in the visual arts is thus multifarious in its expression, but unified in that—almost without exception—it includes actual, physical representations of the body of the artist in some way. The artist is present, at least in image.

The nature of a musical self-portrait is more precarious in definition and form–and examples are comparatively thin on the ground. Music exhibits a high degree of abstraction; how could a composer produce an analogy to the self-portrait in sound alone?

A small number of composers have titled a work in a manner that suggests self-portraiture, as early as François Couperin's *La Couperin*, a movement from his fourth book of pieces for harpsichord from 1730. Perhaps there is some psychological self-depiction or significance to this piece's chromaticism, indicating some kind of *Affekt* of inner turmoil, but if so, it is too abstract to be absorbed by the listener. More recently, in György Ligeti's *Selbstportrait mit Reich und Riley (und Chopin ist auch dabei)* (1976),[45] there are allusions to the characteristic phasing and close canonical techniques of Steve Reich, and Terry Riley and Chopin's melodic patterning, in music that otherwise recalls Ligeti's own "pattern-meccanico" idiom (Clendinning, 1993): the cloud-like rhythmically obscured material and

[45] The pieces were published under the collective title *Monument · Selbstportrait · Bewegung* (Monument · Self portrait · Movement) (1976).

the clock-like fast interlocking material.[46] In both Ligeti and Couperin's cases, it transpires that the music only portrays them in that it is obviously, stylistically, written by them (but only to those familiar with their music, naturally). This notion of artistic-stylistic traits as self-portrait echoes Michel Beaujour's interrogation of "metaphysical" literary self-portraits:

> [The literary self-portraitist] proclaims, "I am this appearance"; I am, for instance, my "styles", my "writing" my "text"; or even, more radically, I am style, writing, text; I am a textual and stylistic histrion [i.e. actor], wild in my dispersal." (Beaujour, 1980, 338)

Similarly, Ligeti and Couperin's works appear to announce, "I am style", but the composers' bodily presence is not represented in any tangible manner.

An example of a more literal self-portrait in music is to be found–or rather, might be searched for–in the track by Aphex Twin (Richard D. James) commonly known as "Equation"[47] (the B-side to his 1999 single *Windowlicker*). In the last twenty seconds of the track, any listener viewing the spectrogram of the waveform of the music would see an image of the composer's face, in a characteristic demonic grin, programmed into the music. Any casual listener, however, would be unaware of its presence. The music does not seek to portray the composer in any audible manner, but is a well-hidden, visual easter egg: the actual sound of the image of the face is the arbitrary noise resulting from encoding the image.

The Subject / The Object displays more direct transferals from the aesthetics of visual arts, which go beyond matters of compositional style and abstracted self-reference. In

[46] The analogies of the "clock" and "cloud" to these musical materials are Ligeti's own, exemplified in his choral-orchestral piece *Clocks and Clouds* (1972-3).

[47] The actual title of the track, as listed on the sleeve of the release is $\Delta M_i^{-1} = -\alpha \Sigma_{n=1}^{N} D_i[n]\,[\Sigma_{j \in C[i]} F_{ji}[n-1] + F\mathrm{ext}_i[n^{-1}]]$, reflecting the quasi-mathematical obscurity and sense of humour found across James' work.

general, from its conception, the work was intended to exhibit characteristics that recalled a work of visual art, in the monolithic, object-like forms.

The use of my voice projects a vocalic body; I "authenticate my presence" in a manner similar to Emma Hart's description of contemporary video artists who feature themselves in their own work:

> They are all in their videos. [John] Smith and [Laurie] Prouvost are voice-overs, authenticating their presence behind the camera. [Ryan] Trecartin and [Monster] Chetwynd[48] are in their videos, covered in strange make up and performing. To the viewer it speaks of a desire to collapse distance and connect with us eyeball to eyeball. (Hart, 2013, 48)

From a textual point of view, although *The Subject* is derived from my semi-conscious episodes of free-writing, these are heavily edited and manipulated, so they are not unmediated representations of my subconscious. Rather, the texts use my voice as blank canvas, and I become an ambiguous narrator, projecting the dystopian subject matter as described above. The continuity of the wall of voice, however, has the effect of causing words to gradually lose their meaning, while the listener's focus is drawn to the sound alone. The wilful verbosity of the text is impossible for a listener to retain as a narrative thread, so the work becomes a portrait of the idiolectal features of my physical voice: its phonology, intonation and accent. This is followed in *The Object* by a portrait of the range of my voice, and the characteristic sounds of my breath running out, which gives the surface of the music

[48] Chetwynd has previously exhibited work under the names Spartacus Chetwynd (as quoted in Hart, 2013) and Marvin Gaye Chetwynd, and has used the moniker Monster Chetwynd since 2018.

its subtly unstable quality. This focus on the phonological aspects of my voice is enhanced through the recording technique of a very close placement of the microphone.

In its use of the unique sonic qualities of my voice, the work echoes Alvin Lucier's *I am sitting in a room* (1969). In the composer's own recordings, his vocal "imperfection" (sic)—his stammer—becomes a key sonic element of this piece. Since the material of Lucier's work is solely the speaking voice, gradually "smooth[ed] out" by the resonant frequencies of a space, until just the general pitch material and sibilants remain, I propose that this is a self-portrait of his voice, both physically (the recording becoming a vocalic body), but also as an image of his vulnerability. The "imperfections" reveal themselves like the fingerprints of the artist on the canvas, proof of the humanity at work.[49]

The Subject / The Object thus recalls the tension in visual art between authentic and manipulated self-portraiture: authentic in its exploitation of the sound of my voice, vocal imperfections, and text emerging from the subconscious; manipulated in its use of a carefully edited text, theatricalised delivery of that text (talking at speed), and the seamless audio edit. In this way, the microphone becomes my mirror and camera, capturing a seemingly realistic image; the text becomes an ambiguous costume, reflecting both the real me and a narrative character; the studio becomes my darkroom, where these raw images are both enhanced and made enigmatic.

[49] When *I am sitting in a room* is realised by other performers (as it frequently has been), the work similarly becomes a performed portrait of that speaker's voice, with their unique characteristics. Though no longer a self-portrait, the work continues to act as a mirror to the performer. This was demonstrated effectively by the 2021 online performance of the work by ninety different musicians, organised by ISSUE Project Room in celebration of Lucier's ninetieth birthday, in which the wide variation in the character of the performances revealed the performers' personalities.

Aspects of self-portraiture in earlier portfolio works

Inevitably, my conception and realisation of *The Subject / The Object* as a work of audio self-portraiture causes me to reflect, in hindsight, on the extent to which this notion also informs my earlier works, as described in chapters 4 and 5, albeit subconsciously in the making of these works.

With the exception of *The Indistinguishables*, my voice features in all these works. In some respects, this echoes Pfahler's idea of "availablism" (Dempsey, 2014), in that they exhibit a positive embrace of my voice as an immediately available, fundamental sonic material. However, I also acknowledge that I have persistently and deliberately sought to "authenticate my presence" (Hart, 2013) through featuring my own voice as narrator, and as voice actor, and in this way, there is absolutely an element of self-portraiture in these works: I present myself, through my voice, as a key component of the subject matter and material.

Hart's description of the increasingly regular occurrence of the (visual) artist presenting themselves in their work is not matched in contemporary classical music: it remains unusual, for example, for concert works for chamber ensembles to include the composer's voice, so in this way the presence of my voice in *Five Loops for the Bathyscaphe* and *Freezywater* signifies self-portraiture.

This is contingent, of course, on listeners knowing that it is me, on recognising my voice, in a situation similar to the familiarity of Stockhausen's voice to audiences, as described by Emmerson (2007, 80). While my own voice may not be as familiar to a general listenership as Stockhausen's, the reappearance of my voice in consecutive works reinforces my authenticated presence: I become the recurring protagonist, and become increasingly

familiar and identifiable. This also links the music in this portfolio to the experimental pop albums from earlier in my career,[50] strengthening the aural image of my voice for listeners familiar with a range of my work.

My works exhibit an ambivalent and nuanced approach to the reflexivity of voice as described by Levack Drever (2002) in the works by Westerkamp and Ferrari discussed at the end of chapter 2.5. Generally, I do not reveal compositional approaches, or provide a personal perspective through the spoken narrative. The exception is *Anticlock*, wherein the voice almost exclusively functions in revealing metrical, rhythmic or pitch structures, but in this work I present myself as an *antagonist* rather than benign observer: I am *in character* as an alien presence. Therefore, the portfolio works express themselves generally as manipulated rather than authentic self-portraits, and my persona is revealed and perceived through my vocal phonology and idiolect, rather than through narrative content.

[50] These three albums, all featuring my singing and speaking voice, were *Unfair / Funfair* (2006), *Trainwreck / Raincheck* (2007) and *Everything / Everything* (2008). Their release under my pseudonym Simon Bookish reflects a manipulated approach to self-portraiture, of course.

Chapter 7: Conclusion

The fundamental aims in my utilisation of the disembodied speaking voice were to connect to a wide range of textual sources, including those atypical in musical works, and to develop new forms in my work. The disembodied speaking voice has proved a fertile source of material in this regard, as demonstrated by the variety of the text sources, and curatorial approaches to its deployment: verbatim texts that are both globally political (*Red and Blue*) and intimately idiosyncratic (*Affix Stamp Here*); found texts that hold within them both a poetic and quotidian quality (*The Indistinguishables* and *Freezywater*); documentary material (*Five Loops for the Bathyscaphe*); and material sourced from the subconscious (*The Subject*). Use of the voice and text material in these works has allowed them to make topical and thematic connections to the extramusical: the works *look out into the world*.

The adoption of modes of voice derived from cinematic, radiophonic and everyday models have resulted in not only new forms in my work, but divergent forms, from the highly structured, austere forms of *Freezywater* and *The Indistinguishables* to the collage and non-teleological dramatic forms of *Red and Blue* and *Anticlock*. The disembodied speaking voice connects these works to extramusical formal tropes and a complex web of allusions. My work has thus forged structural-organisational links to visual art works and other media, so that works such as *Red and Blue* exhibit a musical-radiophonic hybridity, while *The Subject / The Object* exhibits a direct resemblance to visual art works, in its monolithic form and function as a work of self-portraiture. Elsewhere, I have shown that the disembodied speaking voice provides a cultural and emotional resonance in my work, through the deliberate adoption of perceptions of the voice as cinematic, radiophonic or derived from the everyday.

My development of a new means of categorising the relationships between speaking voice and music (chapter 3) has highlighted the flexibility of its use, through the highly varied sources cited, demonstrating that the speaking voice can be entirely separate from musical content, or enmeshed in music, or provide the sole sonic or musical content. In this way, I have shown how multifaceted the operations of the relationships between speaking voice and music can be, in comparison with the singing voice.

A Changing Field

It is this flexibility that inevitably means there are numerous routes for further research, and many avenues of investigation that have fallen outside the scope of this thesis and my own practice. Besides the emergent technologies mentioned in chapter 1 of this thesis, a limitation of my current research and practice precluded detailed discussion of the concurrent culture of research into gendered voices and vocal identity. My research has drawn my attention to this changing field, and the possible intersections with my interest in self-portraiture and characterised (radiophonic) voices, and in the appreciation of the disembodied speaking voice as a musical object.

It has also highlighted a changing attitude and an increasing interest with the disembodied speaking voice that has recently arisen in the wider field of music. As an anecdotal example of this current musical engagement, while I was completing this thesis in August 2022, an audio recording was made publicly available online by the UK railway

operator ScotRail, apparently under a Freedom of Information (FOI) request.[51] This recording is a two-hour long file comprising every automated spoken announcement used on board ScotRail trains and at stations, predominantly recorded by voiceover actor Alison McKay. As of late August 2022, this recording had attracted dozens of musical uses, reinterpretations, and remixes in a multitude of styles, and had been shared on Twitter over 5000 times.[52]

Naturally, this use of a *vox ex machina* caption voice, and an example of spoken found text, echoes my own work, so it was intriguing to see the immediate response to this kind of material by a large number of musicians. This is surely indicative of the disembodied speaking voice—and of non-lyrical, non-poetic texts—having not just a musical use, but a cultural relevance and a musical *desirability*.

Future Directions

My development of the notion of the audio self-portrait, through manipulated recordings and vocal disembodiment, is also one I intend to apply to future work. Research into this idea, while writing *The Subject / The Object* also cast my thinking about my earlier work in a new light and altered my thinking about the work of other composers (such as can be seen in my discussion of Alvin Lucier's *I am sitting in a room* in chapter 6.3). Thus, using this notion as a model, I can see potential for new ways of analysing and thinking about the aesthetics of the wider field of vocal music and sound art, alongside development of my own

[51] The United Kingdom's Freedom of Information Act 2000, commonly abbreviated to FOI, is legislation to provide access to information held by public authorities. Under this act, public authorities are obliged to publish certain information about their activities and members of the public are entitled to request information, subject to certain limitations.

[52] See https://twitter.com/jonbradyphoto/status/1560630522002321408 (Jon Brady, Twitter post, 19 August 202, 3:11 p.m., accessed 30 August 2022). The audio file is available on the ScotRail website here: https://files.scotrail.co.uk/ScotRail_Station_Announcements_June2022.mp3.

body of work. Conceivably, this model could lead to works that portray a wide number of other vocal collaborators, exploring new avenues of vocal identity and approaches to participation.

My employment of musical-radiophonic hybridity of perception of voices as characters and narrators in works such as *Five Loops for the Bathyscape* and *Red and Blue* also points the way to future forms of acousmatic theatrical work: opera as podcast, as radio and across future media.

Since completing the works that comprise the portfolio, I have already embarked on projects that utilise the divergent palette of techniques developed and described in this document (structural, perceptual and textual), but with new freedom. For example, my 2021 work *The Halogens* utilises a combination of texts derived from lists, scientific data and free-writing, on the subject of the Group 19 chemical elements. Unlike earlier works, the curatorial approach to text is freer, and the voice-music relationship is unpremeditated and improvisatory in deployment, since the work was created in the studio with live musicians, resulting in a fixed media work. Informed by my research, my ongoing approach to the disembodied speaking voice has embraced this spontaneity: a musical, technical and linguistic fluency.

Appendix: *Red and Blue*: table of text sources, structure, voice-music relationships and treatments

Key:

Blue text: Verbatim material by Margaret Thatcher (MT)
Red text: Verbatim material by Ronald Reagan (RR)
Green text: Paraphrase and original texts by Leo Chadburn
Purple text: Lists taken from
https://en.wikipedia.org/wiki/List_of_United_States_nuclear_weapons_tests

SECTION 1
Text Source:
MT letter to RR, 5 May 1982: MT expresses concerns about US/Peruvian peace proposals

Time stamp:	Text:	Voice-Music Relationship:	Delivery:	Treatment:
00′00″	*Red and Blue*	Type D (chant)	Spoken	None
00′11″	*I am writing to you*	Type B (coexistent)	Spoken	Dictaphone
00′16″	*I am writing to you because I think*	Type B (coexistent)	Spoken	Dictaphone
00′23″	*I am writing to you because I think you are the only one*	Type B (coexistent)	Spoken	Dictaphone
00′30″	*I am writing to you because I think you are the only one who will understand*	Type B (coexistent)	Spoken	Dictaphone
00′41″	*I am writing to you because I think you are the only one who will understand the significance*	Type B (coexistent)	Spoken	Dictaphone
00′52″	*I am writing to you because I think you are the only one who will understand the significance of what I am trying to say*	Type B (coexistent)	Spoken	Dictaphone
01′08″	*Stop*	Type B (coexistent)	Spoken	Dictaphone

SECTION 2
Text Source:
RR letter to MT, 12 October 1984: RR sends sympathy following Brighton bombing

Time stamp:	Text:	Voice-Music Relationship:	Delivery:	Treatment:
01′18″	*Vistas, mountains, weapons, shopping*	Type D (Chant)	Chant	Multitrack
01′25″	*I declare this business closed!*	Type B (coexistent)	Spoken	None
01′28″	*Violence gathers at the border*	Sung	Sung	None
01′31″	*Indiscriminate*	Type D (Chant)	Chant	None
01′33″	*And brutal*	Type D (Chant)	Chant	Delay
01′34″	*A threat to our democracy* [original: “threat to all democracies”]	Type D (Chant)	Chant	Multitrack
01′36″	*We must work together*	Sung	Sung	Multitrack

01′40″	*To thwart this scourge against humanity*	Type B (coexistent)	Spoken	None
01′44″	*In the context of our special*	Type D (Chant)	Chant	None
01′49″	*special relationship*	Type D (Chant)	Chant	Multitrack
01′52″	*Water, iron, money, intelligence*	Type D (Chant)	Chant	Multitrack
01′59″	*Our experts are red-hot in anticipation of a call from your experts*	Type B (coexistent)	Spoken	None
02′08″	*A scourge against humanity*	Sung	Sung	None
02′10″	*We'll bring the perpetrators down* [original: "bringing the perpetrators to justice"]	Sung	Sung	None
02′16″	*Vistas, mountains, weapons, intelligence*	Type D (Chant)	Chant	Multitrack

SECTION 3
Text Source:
RR letter to MT, 30 April 1975: RR writes following the Fall of Saigon

Time stamp:	Text:	Voice-Music Relationship:	Delivery:	Treatment:
02′25″	*Black as a starless sky*	Type B (coexistent)	Spoken	Dictaphone
02′39″	*The news arrived of Saigon's surrender*	Sung	Sung	Multitrack
02′58″	*And suddenly the shadows seem to have lengthened*	Sung	Sung	Delay
03′27″	*Black as a starless sky*	Type B (coexistent)	Spoken	Dictaphone

SECTION 4
Text Source:
MT letter to RR, 5 March 1981: MT thanks RR for hospitality in Washington

Time stamp:	Text:	Voice-Music Relationship:	Delivery:	Treatment:
03′31″	*The photographs are beautiful*	Type D (Chant)	Spoken	Delay
03′31″	*The sea, the sun*	Sung	Sung	Multitrack
03′32″	*Memory, guilded*	Sung	Sung	Multitrack
03′34″	*Picture, perfect*	Sung	Sung	Multitrack
03′36″	*Landing, handshake*	Sung	Sung	Multitrack
03′37″	*Smile for the cameras!*	Type D (Chant)	Spoken	Delay
03′39″	*I declare this business closed!*	Type D (Chant)	Whispered	Multitrack
03′43″	*I declare this business closed!*	Type B (coexistent)	Spoken	Delay
03′45″	*In every sense a memorable occasion*	Sung	Sung	Multitrack
03′46″	*And the hospitality and the generosity*	Sung	Sung	Multitrack
03′48″	*Stop!*	Type D (Chant)	Spoken	None
03′58″	*The photographs are beautiful*	Type D (Chant)	Spoken	Delay
04′00″	*Shining, transaction*	Type D (Chant)	Spoken	Delay
04′02″	*The sun, the sea and*	Sung	Sung	Multitrack
04′04″	*Unity*	Type D (Chant)	Spoken	Multitrack
04′06″	*Festive, greetings*	Sung	Sung	Multitrack
04′07″	*Vacation, meetings*	Sung	Sung	Multitrack
04′09″	*Sign here! Sign here!*	Sung	Sung	Multitrack
04′11″	*Why won't you look at me?*	Type D (Chant)	Spoken	Multitrack
04′13″	*Look at me*	Sung	Sung	Multitrack
04′20″	*And the splendid glass eagle by Steuben on my desk as I write*	Type B (coexistent)	Spoken	None

04′31″	*The photographs are beautiful*	Type D (Chant)	Spoken	Delay
	[entire text of section, played in reverse]	Type B (coexistent)	Spoken	Multitrack
04′50″	*Sir Winston's portrait is just superb*	Type B (coexistent)	Spoken	None
04′55″	*Sign here, sign here! Why won't you look at me?*	Sung	Sung	Delay
05′02″	*The photographs are beautiful*	Type D (Chant)	Spoken	Multitrack

SECTION 5
Text Sources:
RR letter to MT, 24 January 1984: plans for future space missions
MT letter to RR, 15 March 1984: response to the above
List of US Nuclear Weapons Tests

Time stamp:	Text:	Voice-Music Relationship:	Delivery:	Treatment:
05′05″	*US nuclear weapons tests 1981-1991*	Type B (coexistent)	Spoken	Dictaphone
05′19″	*Praetorian: devices fired: 20*	Type B (coexistent)	Spoken	Dictaphone
05′19″	*We are floating*	Sung	Sung	Multitrack
05′26″	*Phalanx: devices fired: 19*	Type B (coexistent)	Spoken	Dictaphone
05′32″	*Thank you for your message*	Sung	Sung	Multitrack
05′35″	*Fusileer: devices fired: 16*	Type B (coexistent)	Spoken	Dictaphone
05′44″	*Grenadier: devices fired: 16*	Type B (coexistent)	Spoken	Dictaphone
05′53″	*Blast him into space*	Type B (coexistent)	Spoken	Multitrack
05′55″	*Charioteer: devices fired: 16*	Type B (coexistent)	Spoken	Dictaphone
06′00″	*And proceed*	Type B (coexistent)	Spoken	Delay
06′03″	*Musketeer: devices fired: 16*	Type B (coexistent)	Spoken	Dictaphone
06′03″	*With a manned space station programme*	Type B (coexistent)	Spoken	None
06′08″	*Manned space station programme*	Sung	Sung	Multitrack
06′10″	*Touchstone: devices fired: 15*	Type B (coexistent)	Spoken	Dictaphone
06′15″	*British astronauts*	Type B (coexistent)	Spoken	Delay
06′17″	*Cornerstone: devices fired: 17*	Type B (coexistent)	Spoken	Dictaphone
06′19″	*British*	Sung	Sung	Multitrack
06′23″	*Aqueduct: devices fired: 13*	Type B (coexistent)	Spoken	Dictaphone
06′28″	*Thank you for giving us these wonderful opportunities*	Sung	Sung	Multitrack
06′31″	*Sculpin: devices fired: 9*	Type B (coexistent)	Spoken	Dictaphone
06′37″	*We accept*	Sung	Sung	Multitrack
06′39″	*Julin: devices fired: 9*	Type B (coexistent)	Spoken	Dictaphone

SECTION 6
Text Source:
MT letter to RR, 14 January 1985

Time stamp:	Text:	Voice-Music Relationship:	Delivery:	Treatment:
06′43″	*The alliance marches together into 1984* [original: 1985]	Type D (Chant)	Chant (whisper)	Spring reverb
06′48″	*In the hope that, as you have said yourself*	Type D (Chant)	Chant (whisper)	Spring reverb
06′51″	*It should become the year of dialogue*	Type D (Chant)	Chant (whisper)	Spring reverb

06′54″	*The alliance marches together into 1984*	Type D (Chant)	Chant (whisper)	Spring reverb
06′59″	*In the hope that, as you have said yourself*	Type D (Chant)	Chant (whisper)	Spring reverb
07′03″	*It should become the year of dialogue*	Type D (Chant)	Chant (whisper)	Spring reverb
07′08″	*The alliance marches together into 1984*	Type D (Chant)	Chant (whisper)	Spring reverb
07′14″	*Echoing your tribute to members of your delegation*	Type D (Chant)	Chant (whisper)	Spring reverb
07′20″	*Creating a more stable and safer world*	Type D (Chant)	Chant (whisper)	Spring reverb
07′26″	*The alliance marches together into 1984*	Type D (Chant)	Chant (whisper)	Spring reverb

SECTION 7
Text Source:
MT telephone conversation with RR, 5 March 1981

Time stamp:	Text:	Voice-Music Relationship:	Delivery:	Treatment:
07′36″	*Your voice is faint*	Type C (cognate)	Spoken	Dictaphone
07′41″	*I can hear you*	Type C (cognate)	Spoken	Dictaphone
07′46″	*Afraid of the hail of gunfire*	Type C (cognate)	Spoken	Dictaphone
07′51″	*Be not afraid*	Type C (cognate)	Spoken	Dictaphone
07′56″	*We will pull [original: put] together a government*	Type C (cognate)	Spoken	Dictaphone
08′01″	*Those people on the islands are remarkable*	Type C (cognate)	Spoken	Dictaphone
08′04″	*The announcement to the press that our forces are onshore*	Type C (cognate)	Spoken	Dictaphone
08′11″	*She's a wonderful person*	Type C (cognate)	Spoken	Dictaphone
08′16″	*She's captured our city by storm*	Type C (cognate)	Spoken	Dictaphone
08′21″	*There's a lot of work to do yet*	Type C (cognate)	Spoken	Dictaphone
08′26″	*Oh yes*	Type C (cognate)	Spoken	Dictaphone
08′31″	*But, we are colloids*	Type C (cognate)	Spoken	Dictaphone
08′36″	*We are an emulsion*	Type C (cognate)	Spoken	Dictaphone
08′41″	*A foam*	Type C (cognate)	Spoken	Dictaphone
08′46″	*WE GEL*	Type C (cognate)	Spoken	Dictaphone
08′51″	*I love to hear your voice*	Type C (cognate)	Spoken	Dictaphone
08′56″	*I feel pleasure*	Type C (cognate)	Spoken	Dictaphone
09′01″	*I must return to debate in the house*	Type C (cognate)	Spoken	Dictaphone
09′06″	*Go get 'em*	Type C (cognate)	Spoken	Dictaphone
09′08″	*Eat them alive*	Type D (Chant)	Spoken	Dictaphone

SECTION 8 Text Source: MT memorandum to RR, 25 July 1984				
Time stamp:	**Text:**	**Voice-Music Relationship:**	**Delivery:**	**Treatment:**
09′17″	*Tell him I've got my fingers crossed*	Type D (Chant)	Whispered /Spoken	Delay
	Tell him I've got my toes crossed	Type D (Chant)	Whispered /Spoken	Delay
	Tell him I've got everything crossed	Type D (Chant)	Whispered /Spoken	Delay

SECTION 9 Text Source: RR letter to MT, 2 February 1981				
Time stamp:	**Text:**	**Voice-Music Relationship:**	**Delivery:**	**Treatment:**
10′17″	*This is a premonition*	Type A (discrete)	Spoken	None
10′20″	*We exist in a time of saturated images*	Type C (cognate)	Whispered /Spoken	Multitrack
10′25″	*We observe ourselves*	Type D (Chant)	Spoken	None
10′27″	*Playing ourselves*	Type D (Chant)	Spoken	Multitrack
10′29″	*On the television*	Type D (Chant)	Spoken	Delay
10′31″	*Amorphous*	Type D (Chant)	Spoken	None
10′33″	*Fashioned in latex*	Type D (Chant)	Spoken	None
10′37″	*Colour-coded*	Type D (Chant)	Whispered /Spoken	Multitrack
10′39″	*red and blue*	Type D (Chant)	Spoken	Delay
10′43″	*This is a premonition*	Type C (cognate)	Spoken	None
10′46″	*A record of our mutual concern* [original: "special concern"]	Type C (cognate)	Spoken	None
10′50″	*For democracy*	Type D (Chant)	Spoken	Multitrack
10′53″	*And for liberty*	Type D (Chant)	Spoken	Multitrack
10′55″	*The essence of our* [original: "the"] *special*	Type D (Chant)	Spoken	Multitrack
10′58″	*special*	Type D (Chant)	Spoken	Delay
10′59″	*relationship*	Type D (Chant)	Spoken	Multitrack
11′03″	*This is the inauguration*	Type D (Chant)	Spoken	Multitrack
11′05″	*The beginning of our cooperation*	Type D (Chant)	Spoken	None
11′08″	*And consultation between your government*	Type D (Chant)	Spoken	None
11′12″	*And mine*	Type D (Chant)	Spoken	Delay
11′14″	*This is a premonition*	Type D (Chant)	Spoken	None

SECTION 10 Text Source: RR, "An American Life" 1990 MT, "The Downing Street Years, 1979-1990" (MT and RR;s autobiographical recollections of each other)				
Time stamp:	**Text:**	**Voice-Music Relationship:**	**Delivery:**	**Treatment:**
11′16″	*Fall...*	Sung	Sung	Multitrack
11′17″	*I could speak to her for hours...*	Type B (coexistent)	Spoken	None
11′23″	*Fall into a swoon...*	Sung	Sung	Delay
11′26″	*Tell me everything about him...*	Type B (coexistent)	Spoken	None
11′29″	*But tell me*	Sung	Sung	Delay
11′30″	*I liked her immediately...*	Type B (coexistent)	Spoken	None
11′32″	*Who's that woman..*	Sung	Sung	Delay
11′33″	*She was warm, feminine*	Type B (coexistent)	Spoken	None
11′37″	*...on the other side of the room?*	Sung	Sung	Delay
11′42″	*Intelligent...*	Type B (coexistent)	Spoken	None
11′43″	*Looking at me from the corners of your eyes*	Type D (Chant)	Spoken	Multitrack
11′45″	*A kind of mocking distain*	Type B (coexistent)	Spoken	None
11′49″	*1, 2, 3...*		Spoken	None
11′51″	*WHAT?*		Shouted	Spring reverb
11′51″	*Warm, buoyant*	Sung	Sung	Low Pass Filters
11′55″	*Feminine, Intelligent*	Sung	Sung	Low Pass Filters
11′57″	*We were soul mates, elected*	Sung	Sung	Low Pass Filters
12′01″	*Like movie stars, make speeches*	Sung	Sung	Low Pass Filters
12′05″	*The barrier failed, I've stopped the clock*	Sung	Sung	Low Pass Filters
12′09″	*The snare opens, it's losing its identity*	Sung	Sung	Low Pass Filters
12′12″	*And adjusted for GMT and coming from nowhere*	Sung	Sung	Low Pass Filters
12′15″	*He was risen from poverty*	Type D (Chant)	Spoken	None
12′32″	*Let the sky OPEN*	Type D (Chant)	Spoken/ Shouted	Multitrack
12′33″	*Let the silky ship DESCEND*	Type D (Chant)	Spoken/ Shouted	Multitrack
12′35″	*Let the helicopter pull the leaves from the lawn*	Type B (coexistent)	Spoken	Delay
12′37″	*LET'S MEET*	Type D (Chant)	Shouted	Multitrack
12′39″	*At Aspen, Colorado*	Type D (Chant)	Spoken	Multitrack
12′40″	*Handshakes*	Type D (Chant)	Spoken	None
12′41″	*Let the sky CLOSE*	Type D (Chant)	Spoken/ Shouted	Multitrack
12′42″	*Let the silky ship ascend*	Type D (Chant)	Spoken	Multitrack
12′44″	*LET'S MEET*	Type D (Chant)	Spoken/ Shouted	Multitrack
12′45″	*Let the ice freeze on Moscow*	Type B (coexistent)	Spoken	Multitrack
12′47″	*Mountains, vistas, money, shopping*	Type B (coexistent)	Spoken	Dictaphone
12′49″	*Weapons, Leningrad, Moscow*	Type B (coexistent)	Spoken	Dictaphone
12′56″	*Water,iron, intelligence*	Type B (coexistent)	Spoken	Dictaphone

SECTION 11 Text Source: MT letter to RR, 30 March 1981: MT expresses shock at assassination attempt on RR				
Time stamp:	**Text:**	**Voice-Music Relationship:**	**Delivery:**	**Treatment:**
13′20″	*Let this be the last attempt*	Sung	Sung	Multitrack
13′56″	*This nation's injury was my distress*	Sung	Sung	Multitrack

SECTION 12 Text Source: RR letter to MT, 29 December 1978				
Time stamp:	**Text:**	**Voice-Music Relationship:**	**Delivery:**	**Treatment:**
14′35″	*By the way, since we're here*	Type D (Chant)	Spoken	Dictaphone
14′39″	*I should say what I heard*	Type D (Chant)	Spoken	Dictaphone
14′42″	*At the time of the strike*	Type D (Chant)	Spoken	Dictaphone
14′46″	*The bread strike, there in Britain*	Type D (Chant)	Spoken	Dictaphone
14′50″	*Let them starve*	Type D (Chant)	Spoken	Dictaphone
14′53″	*No wait, I'm losing the plot*	Type D (Chant)	Spoken	Dictaphone
14′57″	*You said, as the barrier failed*	Type D (Chant)	Spoken	Dictaphone
15′01″	*As I stopped the clock, as the snare opened*	Type D (Chant)	Spoken	Dictaphone
15′04″	*It's not so common here [original: Even though the expression isn't so common here]*	Type D (Chant)	Spoken	Dictaphone
15′06″	*It's not an expression*	Type D (Chant)	Spoken	Dictaphone
15′08″	*We would use*	Type D (Chant)	Spoken	Dictaphone
15′09″	*There's an ocean between us*	Type D (Chant)	Spoken	Dictaphone / Multitrack
15′12″	*Get to the point*	Type D (Chant)	Spoken	Dictaphone / Delay
15′14″	*Quickly now*	Type D (Chant) / Sung	Spoken / Sung	Dictaphone / Multitrack
15′15″	*You told them, "use your loaf"*	Type D (Chant)	Spoken	Dictaphone
15′17″	*"You must use your loaf"*	Sung	Sung	Multitrack
15′17″	*"You must use your loaf".*	Sung	Sung	None
16′19″	*Stop!*	Type B (coexistent)	Spoken	Dictaphone

Bibliography

Abbate, Carolyn. 1998. “Debussy's Phantom Sounds.” *Cambridge Opera Journal* 10, no.1: 67–96. https://doi.org/10.1017/S0954586700005334

Ablinger, Peter. n.d. “Voices and Piano.” Peter Ablinger (website). Accessed 21 June 2022. https://ablinger.mur.at/voices_and_piano.html

Ablinger, Peter. 2016. “Speaking Piano.” Peter Ablinger (website). Accessed 23 January 2022. https://ablinger.mur.at/speaking_piano.html

Altman, Rick. 1980. “Moving Lips: Cinema as Ventriloquism.” *Yale French Studies*, no. 60: 67–79. https://doi.org/10.2307/2930005

Anderson, Laurie. 1982. *Big Science*. USA: Warner Bros.

Anderson, Laurie. 1984. *Mister Heartbreak*. USA: Warner Bros.

Anderson, Laurie. 1995. *The Ugly One with the Jewels*. USA: Warner Bros.

Andriessen, Louis. 1991. *M is for Man, Music, Mozart*. London: Boosey and Hawkes.

Aphex Twin. 1999. *Windowlicker*. Sheffield: Warp Records.

Armishaw, Holly Marie. n.d. “The Philosophy of Self-portraiture in Contemporary Art.” Accessed 1 August 2022. http://www.hollyarmishaw.com/the-philosophy-of-self-portraiture-in-contemporary-art---essay.html

ArtAngel. n.d. “A Man in a Room, Gambling.” ArtAngel. Accessed 1 July 2022. https://www.artangel.org.uk/project/a-man-in-a-room-gambling

Ashley, Robert. 1974. *In Sara, Mencken, Christ and Beethoven There Were Men and Women*. Milan: Cramps Records.

Ashley, Robert. 2000. *Music with Roots in the Aether*. Cologne: Editions MusikTexte.

Ashley, Robert. 2004. *Tap Dancing in the Sand*. New York; Lovely Music.

Ashley, Robert. 2007. *Tap Dancing in the Sand*. Netherlands: Unsounds, sound recording.

Ashley, Robert. 2009. *Outside of Time: Ideas About Music*. Cologne: Editions MusikTexte.

Aylmer, Olivia. 2016. “The Feminist Provocation of Kembra Pfahler.” Another Magazine. Accessed 25 June 2022. https://www.anothermag.com/fashion-beauty/8908/the-feminist-provocation-of-kembra-pfahler

Bal, Mieke, and Eve Tavor. 1981. “Notes on Narrative Embedding.” *Poetics Today* 2, no.2: 41–59. https://doi.org/10.2307/1772189

Beaujour, Michel. (1980) 1992. *Poetics of the Literary Self-Portrait*. Translated by Yara Milos. New York University Press.

Beckett, Samuel. 2006. *The Complete Dramatic Works*. United Kingdom: Faber & Faber.

Bennett, Vicki (People Like Us). 2020. "First Person, Fourth Wall with People Like Us: Playlist from September 11, 2020." WFMU. Accessed 7 July 2022. https://wfmu.org/playlists/shows/95422

Berio, Luciano. (1958-59) 1998. "Thema (Omaggio a Joyce)." On *Many More Voices*. USA: RCA Victor Red Seal.

Beyst, Stefan. 2016. "Tom Johnson: Narayana's Cows: Musical Design as Statement." Stefan Beyst (website). Accessed 1 May 2022. http://d-sites.net/english/tomjohnsoncows.html

Big Hard Excellent Fish. 1990. *Imperfect List*. London: One Little Indian.

Boards of Canada. 2002. *Geogaddi*. Sheffield: Warp Records.

Bolaños Chamorro, Gabriel José. 2020. "An Analysis of Jonathan Harvey's Speakings for Orchestra and Electronics." *Ricercare* 13: 72-107. https:///doi.org/10.17230/ricercare.2020.13.4

Bonzo Dog Do-Dah Band. 1967. *Gorilla*. United Kingdom: Liberty Records.

Bookish, Simon [Leo Chadburn]. 2008. *Everything / Everything*. Cologne: Tomlab Records.

Bookish, Simon [Leo Chadburn]. 2007. *Trainwreck / Raincheck*. London: Library of Nothing Records.

Bookish, Simon [Leo Chadburn]. 2006. *Unfair / Funfair*. London: Library of Nothing Records.

Branigan, Kevin. 2008. *Radio Beckett: Musicality in the Radio Plays of Samuel Beckett*. Bern: Peter Lang.

Britten, Benjamin. 1958. *Noye's Fludde.* London: Boosey and Hawkes.

Bronzwaer, William. 1981. "Mieke Bal's Concept of Focalization: A Critical Note." *Poetics Today* 2, no.2: 193–201. https://doi.org/10.2307/1772197

Bryars, Gavin. 1995. "A Man in a Room, Gambling." *Parkett* 43: 52-55.

Calle, Sophie, and Laurie Anderson. 1996. *La visite guidée.* Sound Recording and Book. Rotterdam: Museum Boijmans Van Beuningen.

Campany, David. 2011. *Jeff Wall, Picture for Women*. London: One Work, Afterall Books.

Cardiff, Janet. 1999. *The Missing Voice (Case Study B)*. Binaural recording, CD player. Whitechapel Library, London. https://www.artangel.org.uk/project/the-missing-voice-case-study-b/

Catz, Caroline, director. 2020. *Delia Derbyshire: The Myths and the Legendary Tapes*. London: Anti-Worlds.

Chadburn, Leo. 2016. "Opinion Piece." Soundcloud. Accessed 14 September 2022. https://soundcloud.com/simonbookish/opinion-piece

Chion, Michel. (1973) 2007. *Requiem*. Brussels: Sub Rosa Records.

Chion, Michel. (1994) 2019. *Audiovision: Sound on Screen.* Translated by Claudia Gorbman. New York: Columbia University Press.

Chion, Michel. 1999. *The Voice in Cinema*. Translated by Claudia Gorbman. New York: Columbia University Press.

Clendinning, Jane Piper. 1993. "The Pattern-Meccanico Compositions of György Ligeti." *Perspectives of New Music* 31, no. 1: 192–234. https://doi.org/10.2307/833050

Cobbing, Bob. (1965) 2015. *ABC in Sound.* Eighth edition, with an introduction by Robert Sheppard. London: Veer Books.

Coburn, Tyler. 2015. "NaturallySpeaking." Tyler Coburn (website). Accessed 21 June 2022. http://www.tylercoburn.com/naturally.html

Cochrane, Richard. n.d. "Phil Minton: A Doughnut in Both Hands." Metropolis. Accessed 1 September 2022. https://metropolis.free-jazz.net/a-doughnut-in-both-hands/general/2406/

Coltrane, John. 1964. *A Love Supreme*. New York: Impulse! Records.

Conner, Kim. 1997. "Beckett and Radio: The Radioactive Voice." *Samuel Beckett Today / Aujourd'hui,* no. 6: 303–312.

Connor, Alan. 2020. *The Shipping Forecast Puzzle Book*. United Kingdom: Ebury Publishing.

Connor, Steven. 2000. *Dumbstruck: A Cultural History of Ventriloquism*. Oxford University Press.

Connor, Steven. 2004. "The Strains of the Voice." In *Phonorama: Eine Kulturgeschichte der Stimme als Medium*, edited by Brigitte Felderer, 158-72. Berlin: Matthes and Seitz.

Cooper Clarke, John. 1980. "The It Man." On *Snap, Crackle and Bop*. Manchester: CBS.

Cope, Julian. 1994. *Head-on: Memories of the Liverpool Punk-scene and the Story of the Teardrop Explodes (1976-82)*. United Kingdom: Head Heritage.

Cory, Mark, and Barbara Haggh. 1981. "Hörspiel as Music: Music as Hörspiel: The Creative Dialogue between Experimental Radio Drama and Avant-Garde Music." *German Studies Review* 4, no.2: 257-279. Baltimore: John Hopkins University Press. https://doi.org/10.2307/1429272

Crafton, Donald. 1999. *The Talkies: American Cinema's Transition to Sound, 1926-1931*. University of California Press.

Crane, Laurence. 1993. "492 Telecom Announcements." SoundCloud. Accessed 21 June 2022. https://soundcloud.com/laurence-crane/492-telecom-announcements

Crane, Laurence. 1997. *Events.* London: n.p.

Crane, Laurence. 2016. *Sound of Horse*. Olso: Hubro Records.

Davis, C. B. 2003. "Distant Ventriloquism: Vocal Mimesis, Agency and Identity in Ancient Greek Performance." *Theatre Journal* 55, no. 1: 45–65. https://doi.org/10.1353/tj.2003.0015

De La Soul. 1989. *3 Feet High and Rising*. New York: Tommy Boy Records.

Debord, Guy. (1958) 2006. "Theory of the Dérive." In *Situationist International Anthology*, edited and translated by Ken Knabb. United States: Bureau of Public Secrets.

Dempsey, Dean. 2014. "Artist Highlight: Kembra Pfahler." *International Art and Culture*, 1 March 2014. https://www.sfaq.us/2014/03/artist-highlight-kembra-pfahler

Derbyshire, Delia. (1964) 2022. *The Dreams*. Italy: Fantôme Phonographique

Deutsch, Diana. 2003. *Phantom Words and Other Curiosities*. La Jolla: Philomel Records.

Deutsch, Diana, Trevor Henthorn, and Rachael Lapidis. 2011. "Illusory transformation from speech to song." *Journal of the Acoustical Society of America* 129, no. 4: 2245-2252. https://doi.org/10.1121/1.3562174

Doane, Mary Ann. 1980. "The Voice in the Cinema: The Articulation of Body and Space." *Yale French Studies,* no. 60: 33-50. https://doi.org/10.2307/2930003

Dolar, Mladen. 2006. *A Voice and Nothing More*. Cambridge, MA: MIT Press.

Drake, John D. 1971. "The 18th-Century Melodrama." *The Musical Times* 112, no. 1545: 1058-60. https://doi.org/10.2307/954911

Duffy, Carol Ann. 2013. "Prayer." In *Mean Time*. London: Pan Macmillan.

Dürer, Albrecht. 1500. [Self-portrait]. Oil on panel, 67.1 x 48.9 cm. Alte Pinakothek, Munich.

Emmerson, Simon. 2007. *Living Electronic Music*. Farnham: Ashgate.

Esslin, Martin. 1977. "Beckett's Rough for Radio." *Journal of Modern Literature* 6, no. 1: 95–103. http://www.jstor.org/stable/3831022

Faber, Liz. 2020. *The Computer's Voice: From Star Trek to Siri*. United States: University of Minnesota Press.

Fairman, Richard. 2016. "Apartment House [Review]." *Financial Times*, 2 March 2016.

Fennessy, David. 2012. *Piano Trio: Music for the Pauses in a Conversation Between John Cage and Morton Feldman*. Vienna: Universal Edition.

Fennessy, David. n.d. "David Fennessy: Piano Trio: Work Introduction." Universal Edition. Accessed 25 June 2022. https://www.universaledition.com/david-fennessy-4082/works/piano-trio-13862

Ferrari, Luc. n.d. "Analysis / Thoughts: Far West News." Luc Ferrari (website). Accessed 24 January 2023. http://lucferrari.com/en/analyses-reflexion/far-west-news

Fitch, Claire. 2022. *Sounding Emerging Media*. Abingdon: Routledge.

Fleming, Victor, director. 1939. *The Wizard of Oz.* USA: Metro-Goldwyn-Mayer.

Fox, Christopher. 2003. "Which Side Are You On?" *The Musical Times* 144, no. 1882: 65–68. https://doi.org/10.2307/1004713

Frost, Everett C. 1991. "From Wild Strawberries to Strawberry Fields Forever: A Polemical Introduction." *Theatre Journal* 43, no. 3 (Radio Drama): 278-287. Baltimore: John Hopkins University Press. https://doi.org/10.2307/3207582

Galás, Diamanda. 1991. *Plague Mass*. London: Mute Records.

Gallacher, Lyn. 1995. "Enough Said: Chris Mann and the Kabbalah." Paper presented at *Religion, Literature and the Arts, 2nd Australian International Conference, 5-8 January 1995*. North Sydney.

Genette, Gérard. 1980. *Narrative Discourse: An Essay in Method*. Translated by Jane E. Lewin. Ithaca: Cornell University Press.

Glover, Richard. 2015. "Huddersfield Contemporary Music Festival: 24–26 November 2014." *Tempo* 69, no. 272: 65-67. https://doi.org/10.1017/S0040298214001089

Gould, Glenn. (1967-77) 1992. *Glenn Gould's Solitude Trilogy*. Canada: CBC Records.

Gould, Glenn. 1984. *A Glenn Gould Reader*. New York: Knopf.

Greenaway, Peter, director. 1978. *Vertical Features Remake*. London: Peter Greenaway.

Greenaway, Peter, director. 1980. *The Falls*. London: British Film Institute Production Board.

Hardcastle, Paul. 1985. *Paul Hardcastle*. USA: Chrysalis Records.

Hart, Emma. 2012. *Monument to the Unsaved #2 (M20 Death Drives)*. Video and mixed-media sculptural installation. Whitstable Biennale, Whitstable. https://whitstablebiennale.com/project/monument-to-the-unsaved-2-m20-death-drives/

Hart, Emma. 2013. *How to Do Things With Cameras.* PhD thesis, Kingston University. University of the Arts, London, Research Online. http://ualresearchonline.arts.ac.uk/8744/

Harvey, Jonathan (1980) 2008. "Mortuos Plango, Vivos Voco." On *Tombeau de Messiaen.* Sargasso Records.

Heile, Björn. 2006. *The Music of Mauricio Kagel*. London: Ashgate.

Hennies, Sarah. n.d. "Contralto." Sarah Hennies (website). Accessed 27 June 2022. https://www.sarah-hennies.com/contralto

Hiller, Susan, and Barbara Einzig, eds. 1996. *Thinking About Art: Conversations with Susan Hiller.* Manchester University Press.

Hiller, Susan. 2000. *Witness*. Audio-sculpture; 350 loudspeakers, 10 CD players, amplifiers, wiring, lights, etc. The Chapel on Golborne Road, London. https://www.artangel.org.uk/project/witness/

Hinton, David, director. 1992. *DV8 Physical Theatre: Strange Fish*. Halle: ArtHaus Musik.

Hirsch, Shelley. 1987. *Singing*. Eindhoven: Apollo Records.

Hodgson, John A. 1999. "An Other Voice: Ventriloquism in the Romantic Period." *Romanticism on the Net*, no. 16. https://doi.org/10.7202/005878ar

Hudson, Alex. 2012. "The Lull of the Shipping Forecast." BBC News. Accessed 27 June 2022. https://www.bbc.co.uk/news/magazine-17065521

Hugill, Andrew. 2012 *'Pataphysics: A Useless Guide*. Cambridge, MA: MIT Press.

Hunt, Edmund. 2020. "Composition as Commentary: Voice and Poetry in Electroacoustic Music." *Journal for Artistic Research,* no. 20. Amsterdam.

ISSUE Project Room. 2021. “I am Sitting in a Room: Alvin Lucier's 90th Birthday Celebration.” Accessed 1 August 2022. https://issueprojectroom.org/event/i-am-sitting-room-alvin-luciers-90th-birthday-celebration

Jack, Adrian. 2008. “Obituary: Mauricio Kagel.” *The Guardian*, 19 September 2008. https://www.theguardian.com/music/2008/sep/19/obituary.mauricio.kagel

Jefferson, Peter. 2011. *And Now the Shipping Forecast: A Tide of History Around Our Shores*. Cambridge: UIT Cambridge.

Johnson, Scott. (1980-82) 1986. *John Somebody*. USA: Nonesuch.

Johnson, Scott. n.d. “John Somebody.” Scott Johnson Composer. Accessed 17 June 2022. https://scottjohnsoncomposer.com/compositions/johnsomebody.html

Joughin, Kyran, and Jennet Thomas. 2015. “Jennet Thomas in conversation with Kyran Joughin: 25 January 2016, Nunhead, London.” *MIRAJ: Moving Image Review & Art Journal* 10, no. 1-2: 186-203. https://doi.org/10.1386/miraj.5.1-2.186_7

Kagel, Mauricio. 1979. *Der Tribun*. Mainz: WERGO.

Kahn, Douglas, and Gregory Whitehead, eds. 1992. *Wireless Imagination: Sound, Radio and the Avant-Garde*. Cambridge, MA: MIT Press.

Kane, Brian. 2014. *Sound Unseen*. Oxford University Press.

Kessler, Thomas. 2004. *...said the shotgun to the head*. London: Boosey and Hawkes.

Kinney, Fergal. 2022. “Landfill Sprechgesang? Yard Act's The Overload Reviewed.” *The Quietus*, 2022, January 26 2022. https://thequietus.com/articles/31041-yard-act-the-overload-review

Knight, Roderic. 2016. “A New Look at Classification and Terminology for Musical Instruments.” *The Galpin Society Journal*, 69 (April).

Kolber, David. 2002. “Hildegard Westerkamp's *Kits Beach Soundwalk*: Shifting Perspectives in Real World Music.” *Organised Sound* 7, no. 1: 41–43. Cambridge University Press. https://doi.org/10.1017/S1355771802001061

Krogh Groth, Sanne. 2015. “Going Beyond the Word: From Sound Poetry to Text Sound.” In *Boooook: The Life and Work of Bob Cobbing*, edited by Rosie Cooper and William Cobbing, 150-158. London: Occasional Papers.

Kubrick, Stanley, director. 1968. *2001: A Space Odyssey*. USA: Stanley Kubrick Productions; Metro-Goldwyn-Mayer.

Kyriakides, Yannis. 2002. *a conSPIracy cantata*. Amsterdam: Unsounds.

Kyriakides, Yannis. n.d. "a conSPIracy cantata." Yannis Kyriakides (website). Accessed 27 June 2022. https://www.kyriakides.com/a_conspiracy_cantata.html

Lane, Cathy, ed. 2015. *Playing with Words: The Spoken Word in Artistic Practice*. Axminster: Uniformbooks.

Lane, Cathy. 2006. "Voices from the Past: Compositional Approaches to Using Recorded Speech." *Organised Sound* 11, no.1: 3-11 https://doi.org/10.1017/S1355771806000021

Lear, Amanda. 1977. *I am a Photograph*. Munich: Ariola Records.

Levack Drever, John. 2002. "Soundscape Composition: The Convergence of Ethnography and Acousmatic music." *Organised Sound* 7, no. 1: 21-27. https://doi.org/10.1017/S1355771802001048

Ligeti, György. 1976. *Monument · Selbstportrait · Bewegung*. Mainz: Schott Music.

Lim, Liza. 2017. "Time, Possession & Ventriloquism in my Operas." Liza Lim Composer. Accessed 2 May 2022. https://lizalimcomposer.com/2017/07/08/time-possession-ventriloquism-in-my-operas

Lowndes, Sarah. 2016. *The DIY Movement in Art, Music and Publishing: Subjugated Knowledges*. United States: Taylor & Francis.

Lucier, Alvin. (1969) 1981. *I am sitting in a room*. New York: Lovely Music.

Lynch, David, director. 2001. *Mulholland Drive*. USA: Universal Pictures.

Mac Low, Jackson. 2009. *Thing of Beauty: New and Selected Works*. Berkeley: University of California Press.

Machine for Making Sense. 1994. *On Second Thoughts*. Australia: Tall Poppies.

Maddocks, Fiona. 2018. "The Week in Classical: London Sinfonietta 50th Anniversary Concert; Britten Sinfonia - Review." *The Observer*, 28 January 2018. https://www.theguardian.com/music/2018/jan/28/london-sinfonietta-50th-anniversary-concert-review-britten-sinfonia-leo-chadburn

Maltman, Jonathan, and Lindsay Vickery. 2018. "The Foundations of Speech-Melody Composition And The Dawn Of The Digital Music Era." Paper presented at *The Australian Computer Music Conference 2018*. Edith Cowan University, Perth.

Mayo, Christopher. 2015. “Composer Christopher Mayo on Leo Chadburn.” British Music Collection. Accessed 21 June 2022. https://britishmusiccollection.org.uk/article/composer-christopher-mayo-leo-chadburn

McLuhan, Marshall. 1964. *Understanding Media*. London: Routledge and Keegan Paul.

Mills, Billy. 2013. “Poster Poems: Found Poetry.” *The Guardian*, 9 August 2013. https://www.theguardian.com/books/booksblog/2013/aug/09/poster-poems-found-poetry-cutup-collage

Minton, Phil. 1981. *A Doughnut in Both Hands*. New York: Rift Records.

Momus, and Anne Laplantine. 2003. “Summerisle Hörspiel.” UbuWeb. Accessed 21 June 2022. https://ubu.com/sound/momus_summerisle.html

Mozart, Wolfgang Amadeus. (1791) 2006. *Die Zauberflöte* [The Magic Flute]. Kassel: Bärenreiter-Verlag.

Nauman, Bruce. 1984. *One Hundred Live and Die*. Neon and metal, 299.7 x 335.9 x 53.3 cm. Collection of Benesse House Museum, Naoshima.

Nauman, Bruce. 1985. *Good Boy Bad Boy*. Video, two monitors, mono audio. Collection of Tate Gallery, London. https://www.tate.org.uk/art/artworks/nauman-good-boy-bad-boy-t06853

Nauman, Bruce. 2005. *Please Pay Attention Please: Bruce Nauman's Words: Writings and Interviews*. United Kingdom: Penguin Random House LLC.

Nemerov, Howard. 1987. “Found Poem.” *The Missouri Review* 10, no. 3: 71. https://doi.org/10.1353/mis.1987.0005

Neumark, Norie, Ross Gibson, and Theo van Leeuwen, eds. 2010. *Voice: Vocal Aesthetics in Digital Arts and Media*. Cambridge, MA: MIT Press.

Niebur, Louise. 2010. *Special Sound: The Creation and Legacy of the BBC Radiophonic Workshop*. Oxford University Press.

Nyman, Michael. (1974) 1999. *Experimental Music: Cage and Beyond*. Fifth edition, with a foreword by Brian Eno. Cambridge University Press.

Nyman, Michael. (1979) 1981. “Bird List Song.” On *Michael Nyman* (album). Japan: Victor Records.

Oldfield, Mike. 1973. *Tubular Bells*. United Kingdom: Virgin Records

Perloff, Marjorie. 2004. “John Cage's Dublin, Lyn Hejinian's Leningrad: Poetic Cities as Cyberspaces.” In *Architectures of Poetry*, edited by Craig Douglas Dworkin and Maria Eugenia Díaz Sánchez. Amsterdam: Rodopi.

Piccard, Jacques, and Thomas J. Abercrombie. 1960. “Man’s Deepest Dive.” *National Geographic* 118, no. 2: 224-239.

Pisaro, Kathryn Gleasman. 2001. *Music from Scratch: Cornelius Cardew, Experimental Music and the Scratch Orchestra in Britain in the 1960s and 1970s*. United States: Northwestern University.

Porter, Lewis. 1985. “John Coltrane's A Love Supreme: Jazz Improvisation as Composition.” *Journal of the American Musicological Society* 38, no. 3: 593-621. University of California. https://doi.org/10.2307/831480

Potts, Kate. 2017. *The Blown Definitions: Towards a Poetics of the Multi-Vocal Poetic Radio Play*. PhD thesis, Royal Holloway, University of London. https://pure.royalholloway.ac.uk/portal/files/28311280/The_Blown_Definitions_Dr_Kate_Potts_PhD_thesis..pdf

Prokofiev, Sergei. 1936. *Peter and the Wolf*. London: Boosey and Hawkes.

Ravy, Tawnya, and Eric Forcier, eds. 2014. *Words, Worlds and Narratives: Transmedia and Immersion*. Freeland: Inter-Disciplinary Press.

Reed, S. Alexander. 2022. *Laurie Anderson's Big Science*. Oxford University Press.

Reeves, Philip. 2013. “The Shipping Forecast: From Britain's Seas into Its Soul.” National Public Radio. Accessed 27 June 2022. https://www.npr.org/2013/12/16/249722733/the-shipping-forecast-from-britains-seas-into-its-soul

Reich, Steve, and Beryl Korot. 2002. *Three Tales*. London: Boosey and Hawkes.

Reich, Steve, and Beryl Korot. 1993. *The Cave*. London: Boosey and Hawkes.

Reich, Steve. 1987. *Early Works*. USA: Elektra Nonesuch

Reich, Steve. 1988. *Different Trains / Electric Counterpoint*. USA: Elektra Nonesuch.

Reich, Steve. 1988. *Different Trains*. London: Boosey and Hawkes.

Reich, Steve. 2002. *Writings on Music, 1965-2000*. Oxford University Press.

Roeg, Nicolas, director. 1976. *The Man Who Fell to Earth*. United Kingdom: British Lion Films.

Rogers, John, director. 2009. *The London Perambulator*. London: Vanity Projects.

Ross, John, director. 2012. *Evidently... John Cooper Clarke*. BBC Four

Russell, Arthur (1986) 2005. "The Name of the Next Song." On *World of Echo*. Remastered edition. London: Rough Trade.

Rutherford-Johnson, Tim. 2021. "Leo Chadburn, The Subject/The Object. Library of Nothing, Bandcamp." *Tempo* 75, no. 295: 111-112. https://doi.org/10.1017/S0040298220000819

Rzewski, Frederic. 1971. *Coming Together*. Brussels: Sound Pool Music. https://imslp.org/wiki/Coming_Together_(Rzewski%2C_Frederic)

Sadler, Simon. 2005. *Archigram: Architecture Without Architecture*. Cambridge, MA: MIT Press.

Saint Etienne. 1993. *So Tough*. United Kingdom: Heavenly Recordings.

Saint Etienne. 2002. *Finisterre*. United Kingdom: Mantra Records.

Schafer, R. Murray. 1994. *Soundscape: Our Sonic Environment and the Tuning of the World.* Vermont: Destiny Books.

Schlinkert, Guido and Maria Francesca Bonetti. 2004. *Samuel Fosso*. Verona: Harry N. Abrams.

Schoenberg, Arnold. 1912. *Pierrot Lunaire*. Vienna: Universal Edition.

Scholz, Christian. 1997. "A Historial Survey of the German Sound Poetry." Dr Christian Scholz (website). Accessed 1 August 2022. http://www.fen-net.de/christian.scholz/lecture1.html

Schöning, Klaus. 1969. *Neues Hörspiel: Texte, Partituren.* Frankfurt: Suhrkamp.

Schwartz, A. Brad. 2015. *Broadcast Hysteria: Orson Welles's War of the Worlds and the Art of Fake News*. United States: Farrar, Straus and Giroux.

Schwartz, Oscar. 2013. "unAustralian English: Oscar Schwartz Visits Chris Mann in NYC." Cordite Poetry Review. Accessed 25 June 2022. http://cordite.org.au/essays/unaustralian-english/

Schwitters, Kurt. 1932. "Ursonate." *Merz,* no. 24. Hanover.

Scott-Heron, Gil. 1970. *Small Talk at 125th and Lenox*. USA: Flying Dutchman Records.

Scott-Heron, Gil. 1974. *The Revolution Will Not Be Televised*. USA: Flying Dutchman Records.

Sinclair, Iain. 2003. *London Orbital*. London: Penguin.

Sitney, P. Adams. (1969) 2000. “Structural Film.” In *Film Culture Reader.* Lanham: Cooper Square Press, 326-348.

Smith, John, director. 1987. *The Black Tower*. London: Arts Council of Great Britain.

Stacey, Peter F. 1989. “Towards the Analysis of the Relationship of Music and Text in Contemporary Composition.” *Contemporary Music Review* 5, no. 1: 9-27. https://doi.org/10.1080/07494468900640501

Stockhausen, Karlheinz (1969) 1995. *Hymnen*. Kürten: Stockhausen Verlag.

Stockhausen, Karlheinz (1998) 2003. *Mittwochs-Gruss*. Kürten: Stockhausen Verlag.

Stockhausen, Karlheinz. (1955-56) 1991. “Gesang der Jünglinge.” On *Stockhausen Edition no. 3: Elektronische Musik 1952-1960*. Kürten: Stockhausen Verlag.

Stockhausen, Karlheinz. (1977) 1993. *Der Jahreslauf*. Kürten: Stockhausen Verlag.

Stockhausen, Karlheinz. (1977/1987-91) 1996. *Dienstag aus Licht*. Kürten: Stockhausen Verlag.

Stockhausen, Karlheinz. 1977. *Der Jahreslauf.* Kürten: Stockhausen Verlag.

Stockhausen, Karlheinz. 1977/1991. *Jahreslauf.* Kürten: Stockhausen Verlag.

Stravinsky, Igor. 1927. *Oedipus Rex*. London: Boosey and Hawkes.

Testa, Bart. 2008. “Tabula for a Catastrophe: Peter Greenaway's *The Falls* and Foucault's Heterotopia.” In *Peter Greenaway's Postmodern / Poststructuralist Cinema*, edited by Paula Willoquet-Maricondi and Mary Alemany-Galway. Lanham: Scarecrow Press.

The Beach Boys. 1966. *Pet Sounds*. USA: Capitol Records.

The Conet Project. 1998. *The Conet Project: Recordings of Shortwave Numbers Stations*. London. Irdial-Discs.

The Divine Comedy. 1994. *Promenade*. London: Setanta Records.

Thomaidis, Konstantinos, and Ben MacPherson, eds. 2015. *Voice Studies: Critical Approaches to Process, Performance and Experience*. Abingdon: Routledge.

Thomas, Jennet. 2015. “The Unspeakable Freedom Device.” Accessed 1 July 2022. https://jennetthomas.wordpress.com/videos/the-unspeakable-freedom-device

van Peer, René. 2000. [Review of] “The CONET Project: Recordings of Shortwave Numbers Stations.” *Leonardo* 33, no. 1: 67–69.

Venables, Philip, and Béla Bartók. 2018. *Venables plays Bartók*. London: Ricordi.

Venables, Philip. 2018. “Venables plays Bartók.” Philip Venables (website). Accessed 27 June 2022. https://philipvenables.com/venables-plays-bartok

Vickery, Lindsay. 2014. “The Semantics of Redaction (Ferguson).” On *sacrificial zones: chamber music XI*. Bandcamp. https://lindsayvickery.bandcamp.com/album/sacrificial-zones-chamber-music-xi-2014

Walker, Ian. 2010. “Revisiting the Black Tower: Notes Based on a Talk Given at The Royal College of Art.” Ian Walker (website). Accessed 2 September 2022. http://ianwalkerphoto.com/wp-content/uploads/2016/03/Revisiting-The-Black-Tower.pdf

Wall, Jeff. 1979. *Picture for Women*. Cibachrome transparency mounted on lightbox, 204.5 cm × 142.5 cm. Centre Georges Pompidou, Paris.

Walshe, Jennifer, presenter. 2022. *Fümmsböwö (or What is the Word)*. Aired 3 January 2022, BBC Radio 4. https://www.bbc.co.uk/programmes/m001327w

Walshe, Jennifer. 2011. *A Folk Song Collection*. Dublin: The Contemporary Music Centre.

Weiss, Allen S. 1995. *Phantasmic Radio*. Durham: Duke University Press.

Westerkamp, Hildegard. (1989) 1996. “Kits Beach Soundwalk.” On *Transformations*. Canada: Empreintes DIGITALes.

White, John and Gavin Bryars. 1978. *Machine Music*. London: Obscure Records.

White, John. 1993. *Fashion Music*. Austria: LondonHALL.

Whitehead, Gregory. 1989. “Who’s There? Notes on the Materiality of Radio Art.” *Art & Text* 31: 10-13.

Woods-Marsden, Joanna. 1998. *Renaissance Self-portraiture: The Visual Construction of Identity and the Social Status of the Artist*. New Haven: Yale University Press.

Young, Miriama. 2016. *Singing the Body Electric: The Human Voice and Sound Technology*. United Kingdom: Taylor & Francis.