INSIDE OUT: THE UNDER-THEORISED OBJECT AND MATERIAL VOICE IN FINE ART PRACTICE

RICHARD COUZINS

A thesis submitted in partial fulfilment of the requirements of The University of Brighton for the degree of PhD following a programme of study at University for the Creative Arts

January 2014
Acknowledgements

I wish to thank my supervisors Professor George Barber, Dr Jean Wainwright and Professor Kerstin Mey for their support during this research, and T. J. Johnson, Conrad Couzins and Reuben Couzins for allowing me to record and use their voices.

I declare that the research contained in this thesis, unless otherwise formally indicated within the text, is the original work of the author. The thesis has not been previously submitted to this or any other university for a degree, and does not incorporate any material already submitted for a degree.

Signed

Dated
Abstract

The human voice is significant to culture and communication and its agency differs across the heterogeneous discourses within which it is listened to and produced. This thesis assesses what the voice does in Fine Art practice where it is under-theorised but often used by artists. The research questions are: how does a voice register as a material and an object (physical material presence), rather than equating only to the subject who produces it? And how can an artist produce a direct address with their voice? The thesis examines the nature of direct vocal address in Fine Art practice with the installation Trialogue (2013) and with the discussion of case studies that privilege the voice. Trialogue uses three screens to emphasise the action of voices and vocal genres. Four single screen video works are played over three screens during which the audience hears a jazz singer, children and my voice.

We are familiar with our voices presenting our selves, but in Fine Art practice the voice is reproduced, and behaves as a material and object. Artworks and theories divide around the reduction and parameters of voice as production of a human subject and as an object in the material world. Therefore the voice is described with a combination of phenomenological, psychoanalytical and cultural theories. The thesis critically examines theories of Dolar (2006), Ihde (2007), Sperber and Wilson (1995), and Bakhtin (1986) in relation to the phenomenon of the voice in Fine Art practice. The thesis describes how the genre, physical space, consequences of reproduction, and action of listening are emphasised by critical Fine Art practice. Bakhtin describes all utterance as having a genre, and phenomenological theories relate voice to perceptual hierarchy and its relationship with the visual realm. The voice is described as a partial object in psychoanalytical theory. The idea of palimpsest is used as a partial space to situate the object voice. The chapters theorise the voice in Fine Art practice as: object and reproduction; the relationship voices have with images in moving image art practice; and the voice and self.
Contents

Prelude ................................. 6
Introduction ............................ 9
1. Voice and Object ..................... 21
2. Object Voice and Case Studies .... 32
4. Voice and Image ...................... 60
5. Voice and Self ....................... 80
6. Voice and Theory ................... 95
7. Voice and Trialogue ............... 113
8. Voice and Listening ............... 130
Conclusion ............................. 143
Bibliography ............................ 152
Works Cited ........................... 158

Illustrations

Illustration 1. Lenny’s Documentary (1978) Ian Bourn ............ 38
Illustration 2. Semiotics of the Kitchen (1975) Martha Rosler ... 39
Illustration 3. The Miller and The Sweep (1897) G. A. Smith ... 116
Illustration 4. Table (1962) Gerhard Richter .................... 116
Illustration 5. Stills from Just Talk ............................. 118
Illustration 6. Stills from The Jazz Singer ....................... 122
Illustration 7. Stills from Filling Space with Rocks ............. 123
Illustration 8. Stills from Forget Me ......................... 126
Audio and Visual Clips

(Audio clips marked ‘A’ are found on CD and video marked ‘V’ on the DVD included in the back of this thesis)

<table>
<thead>
<tr>
<th>Page</th>
<th>Description</th>
<th>Clip number</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>Florence Nightingale</td>
<td>1A</td>
</tr>
<tr>
<td>6</td>
<td>F. D. Roosevelt</td>
<td>2A</td>
</tr>
<tr>
<td>7</td>
<td>T. S. Eliot</td>
<td>3A</td>
</tr>
<tr>
<td>32</td>
<td>clip from Freeing the Voice</td>
<td>4V</td>
</tr>
<tr>
<td>32</td>
<td>clip from Pieces I Never Did</td>
<td>5V</td>
</tr>
<tr>
<td>32</td>
<td>clip from India Shouting Match</td>
<td>6V</td>
</tr>
<tr>
<td>39</td>
<td>clip from Raw Materials</td>
<td>7A</td>
</tr>
<tr>
<td>40</td>
<td>clip from Get Out of my Mind</td>
<td>8A</td>
</tr>
<tr>
<td>43</td>
<td>clip from The Missing Voice (Case Study B)</td>
<td>9A</td>
</tr>
<tr>
<td>85</td>
<td>clip from The Arbor</td>
<td>10V</td>
</tr>
<tr>
<td>85</td>
<td>clip from The Future’s Getting Old</td>
<td>11V</td>
</tr>
</tbody>
</table>

Video works from installation Trialogue included on DVD –

Filling Space with Rocks          2012          (3 mins)
Forget Me                        2011          (2 mins)
The Jazz Singer                  2012          (4 mins 30 secs)
Just Talk                        2011          (6 mins)
Prelude

Audio recordings now stretch back over a hundred years: the sound of voices have become a way of remembering history.¹ Voices currently fill the airwaves and historical discourses. We can hear and re-discover Florence Nightingale by hearing her voice, and in that moment not just reflect on her myths and achievements but actually hear her in person. This thesis describes how artists critique what it is to actually hear a person through their voice. How a voice is multiple: as bodily sound, medium of language and self that overlap in time and space.

In the recording of 1890 Nightingale tentatively announces that:

> When I am no longer even a memory – just a name – I hope my voice may perpetuate the great work of my life [play sound clip 1]

And indeed it has.² Nightingale’s recording reaffirms Thomas Edison’s thought that his phonograph invention might be used to record people’s last words: a timeless summary of their life – right from the horse’s mouth. We hear a ‘life lived’ in a voice; all voices carry their unique experience and personality, and ideas around this became a major issue for late twentieth-century philosophy. Jacques Derrida, for example, questioned Edmund Husserl’s phenomenological account of the voice as pure presence. More than this, as a result of reproduction, a voice also suggests a milieu in a way not possible prior to recording. The milieu has a materiality not obvious in its own moment but emerges with the comparisons afforded by reproduction and the distance of time.

Famous voices are part of historical events and become part of how we remember a milieu. The rich variety of audio recordings helps us picture a decade. The voices of politicians are well known from archives, and their performances for radio and television changed political rhetoric. Although Franklin D. Roosevelt said ‘the only thing to fear is fear itself’ [play sound clip 2] he also instigated what he called ‘fireside chats’ to speak to the American people. Churchill too knew that in order to galvanise the British nation during World War II he had to sound like a Shakespearian King. Mladen Dolar suggests that the lengthy texts of Stalin’s speeches were delivered as a boring drone because they were designed for ‘a future which would supposedly validate them’ (2006:117), whereas Hitler’s voice was a rousing exposition of law in the

---

¹ A contemporary television programme such as America: The Story of Us (2010, Nutopia), restages much oral history.
² Particularly as this is one of the earliest vocal recordings.
present moment. Politicians use their voices in different ways: Ian Duncan Smith described himself as ‘the quiet man’, which came across as an admission of powerlessness. We don’t generally expect leaders to be quiet or tell jokes, although these can succeed as rhetorical ploys. Obama is capable of a very actorly voice; he takes his time, deliberates, and releases the next phrase just when he couldn’t hold on to the previous one any longer. One imagines he has studied his Churchill and John F. Kennedy well. Marshall McLuhan raised the authenticity of mediated voices when he commented on how Nixon’s voice was ‘phony’ in comparison with Kennedy’s.³

A great singer’s voice often arrests us with its unusual or affected power, for example the voices of Bob Dylan, David Bowie, Luciano Pavarotti, Diamanda Galas or Eartha Kitt. Extra-linguistic ticks don’t do any harm either, for example the ‘oohs’ and ‘ahs’ of Michael Jackson, James Brown or Mark E. Smith. The singing voice has been described as a fetish object, but the relationship between Fine Art practice and voice raises theoretical possibilities other than political rhetoric or the aesthetics of song that this research addresses. The poet T. S. Eliot recited his poetry like a ‘speak your weight machine’⁴ [play sound clip 3]. He did this to minimise his subjectivity in a modernist way that is also found in the isolated mouths put on stage and screen by Samuel Beckett or the linguistic bypassing of Antonin Artaud. It also has echoes in Brecht: his ‘distantiation’ technique, attempting to get the audience to experience the ‘ideology’ in the text, to make things alienated so that true meanings are revealed. Our ears can either be entranced by the innate character of a voice, say a Hollywood star like Lauren Bacall, or on the other hand if all character is denied, we are forced to ‘hear’ in a distant fashion – and we think of it as an alienated voice. Literally, somebody is not there. Conversely, HAL, in amalgamating the two, is particularly funny in the sci-fi film 2001 A Space Odyssey (1968, Stanley Kubrick); for a computer, he does sound an awful lot like a young Californian lying on the beach. HAL’s voice is filled with ticks and he is clearly in a very relaxed and mellow place as he decides whether or not to blow everybody up.

Technology can amplify, store and reposition the relatively fragile sound waves of voices. New bits of history can be found or cleaned up. Voices abound and take on different significance when previously they could not be recorded, and this has an

³ http://www.nextnature.net/2009/12/the-playboy-interview-marshall-mcluhan/ (accessed 20/10/13)
⁴ A speak your weight machine now seems like a prescient technology given the increase in such acousmatic voices in our environment.
impact on the discourses of Fine Art practice that, I argue, do many other things that are not theorised.
Introduction

Duplicity & Duality: What do voices do in Fine Art?

The human voice is regularly, and increasingly, encountered in contemporary art, in many different guises, but its presence is under-theorised in this context. The materiality of the voice is passed over in order to better hear the abstractions of language but, like vision, the voice is outside of language. We understand the rhetorical constructions of politicians’ voices more than we do voices heard in Fine Art practice. There is no Pavarotti in Fine Art practice; instead, I am researching the relationships between language, materiality, subjectivity and the presentation of vocal objects in space in Fine Art practice. We read, listen, and respond to the structuring agency of voices as much as what they say. Voices reveal information about the subject speaking that is direct, affective and outside of language. Just as we know the health and frame of mind of someone by listening to their voice, this research listens to voices in order to discover the contemporary situation within which voices communicate in Fine Art. Listening is not a neutral activity but rather involves a strategy and regime. By listening to voices we can engage with the ideas of psychoanalysis, phenomenology and cultural or Marxist critiques, but individually these theories are not able to capture the phenomenon of the voice in its complexity.

By investigating artworks that privilege the voice, and by making new video works in which I use recordings of voices with images, and in which I address the camera directly, I ask: how does a voice register as a material and an object, rather than equating only to the subject who produces it? Also: how can an artist produce a direct address with their voice? Although a direct address may never be possible, I use phenomenological, psychoanalytical and pragmatic theories to ask how artists critique the voice between being invisible or transparent and as an opaque object presence.

And therefore: what are the theoretical contexts in which the voice transforms its materiality into meaning in Fine Art practice? In this way, the research examines what Fine Art practice tells us about voices in both theory and practice and the concept of the voice as object is framed by Fine Art discourse. I emphasise the materiality of voices by describing their reproduction, disposition and relationship with images and the self.

---

5 We always privilege voices in the way that we listen out for them and can mistake other sounds, such as the sound of wind, as vocal sound.
The aim of this research is to identify what the voice does in Fine Art practice:

- To understand why voices have been under-theorised.
- To assess the limit of the voice, experienced in art, as an object unbounded from a subject.
- To consider the effects of reproduction for voice and its reception in the exhibition spaces of contemporary practice.
- To link my art practice with theories of the voice and find a new juxtaposition of theories that will help other artists and researchers to theorise the voice.

In the text, I lay out a combination of theories to account for the use of voices in Fine Art practice: if I present or use a voice, by what discourses is it underpinned? How do reproduction and repetition weaken the voice’s link to subjectivity, where ordinarily this piece of sound stands in for the whole subject? My aim in this thesis, and the practice it accompanies – the three-screen video installation *Trialogue* – is to set out a theoretical template for how the voice is used in Fine Art practice. My approach to the voice as an object and material in this research is not centred in theoretical discussions of ethnicity or gender that join a nexus of disparate factors in the constitution of a voice such as: health, class, ambient temperature or emotional state. Listening to any voice involves ascribing gender to the owner of the voice raising issues around the construction of gender, while dialect and accent relate to ethnic difference that impact on vocal utterance. Gender is constructed in combination with visual information, as can be seen in the case studies of Rosler(1976) and Gibson(2010) that I discuss.

**What is a voice?**

The voice is the subject of many disciplines from phonology to psychology to medical practice, where it is an organ of many synchronised body parts. For psychoanalysis it is a love object, a place for fetish and delight. The voice can switch between registers, adopting a different genre or meaning at any time while – like a balancing act – it facilitates everyday communication. This makes it difficult to attach to a single theory, and therefore the theory I provide is a combination. The human voice has been described as:

An index of the body, a conveyor of language, a social bond, a musical instrument of sublime flexibility, a gauge of emotion, a central component of
the art of acting, and a register of everyday identity. (Jacob Smith 2008:3)

And: ‘a glove, or a wall, or a bruise, a patch of inflammation, a scar or a wound’ (Steven Connor 2000:5).

The voice is duplicitous because: ‘the voice is against \textit{logos}, the voice as the other of \textit{logos}, its radical alterity. (Dolar, 2006:52). The voice becomes different things to different discourses: the partial or transitional object of psychoanalytical theory; the illocutionary force of speech act theory; the texture of a soundscape; the indexical subjectivity of self; the emotion of concrete utterance absent in written text. The voice is partial and asymmetrical in its relationships with other modalities. The voice has its own dialectic that is emphasised if its autonomy is pursued. It is necessary to keep in mind that there is no one theory for the voice.

To account for the voice in Fine art practice it is necessary to develop what Theo van Leeuwan (1999:128) describes as: ‘the re-introduction of the materiality of the sign instigated by Julia Kristeva and Roland Barthes’ and combine this with Bakhtin’s concepts of utterance and genre. If we speculate about the evolution of voice we can imagine a vocal sound before language. Therefore the voice can be thought to come before language, rather than after it, as a residue described in the psychoanalytical theory of Jacques Lacan (1989:306). M. M. Bakhtin also describes the individual style in non literary-artistic speech genres as ‘an epiphenomenon of the utterance, one of its by products’ (1986:63) which I would equate with materiality. This type of materiality is picked out in recent academic discourses within which I am placing the voice. Christoph Cox (2011:145) argues for a material theory of sound art which I am applying to the everyday, recorded voice, whose hierarchy is ordered and tested in artists’ works: the voice used and heard in this thesis and practice. Jane Bennett proposes a rethinking of the boundary between life and nature. Bennett says that the processes of materiality are ignored over our human intentionality because of:

\begin{quote}
The difficulty of theorizing agency apart from the belief that humans are special in the sense of existing, at least in part, outside of the order of natural material. (2010:36)
\end{quote}

Cox has noted that, theoretically, sound falls between stools ‘thwarting musicological
analysis and falling out with art history, which deals with the visual. Sound is at odds with the visual and textual’ (2011:146). This is also the case for the voice, that at once infiltrates all discourses but is often ignored as a material, form and object. These ways of thinking about the voice are counter-intuitive as the voice is an invisible, ephemeral bearer of language. However, as a reproduced element of the art examined and created for this thesis, its duplicity is as an object, material and event that drives communication as much as language. In art practice language is one of the voice’s contexts and point of analysis, but there are others involving its pre-linguistic materiality. Cox has proposed the need for a material theory of sound whereas Kim-Cohen (2009) is wary of sound as pure form.

Norrie Neumark has commented on the duplicity of voice in what he calls the paradox of the voice:

Attention to the ambiguity and paradox of embodied voice is particularly important, because both provide rich veins for media artists themselves and for scholarly understanding of what artists are doing with voice. (Neumark 2010:xviii)

My method is to critically examine the phenomenology of voice provided by Don Ihde and Maurice Merleau-Ponty, and the linguistic theories of Dan Sperber and Deirdre Wilson and M. M. Bakhtin, in an inquiry into the processes of voice in Fine Art practice. Speech genres support a phenomenology of voice for Fine Art practice. Steven Connor and Mladen Dolar have also described the voice from philosophical and cultural perspectives that are important to this analysis.

Dolar uses the Lacanian notion of object voice (object petit a) as a way to account for voice we experience other than language or the aesthetic singing voice. He does not write about the voice in Fine Art practice because he finds that, ‘aesthetic pleasure obfuscates the object voice’ (2006:4). Dolar writes about the voice through the disciplines of linguistics, philosophy and politics to chart its ethics, physics and metaphysics, all of which he says ignore its agency as a material and object. However, the voice Dolar pursues is also found in Fine Art practice, which is a site of critique, but less often a critique of the voice, which this research addresses.
Connor describes the ‘vocalic space’ (2000:12) that we give to a voice and this runs through the voices heard in these works, voices that move between real, symbolic and imaginary. These frictions inhabit the voices in the works I have made and give as examples. I discuss *Trialogue* alongside works by Bruce Nauman, Janet Cardiff and George Bures Miller, Beatrice Gibson, Martha Rosler, Imogen Stidworthy, Ian Bourn, and others. The aim is to describe a nexus of theories for these works as specific utterances of object voices. I have chosen these works because their voices come to our attention and emerge as material objects.

While very common to our experience, the voice is under-theorised because, as sound, it is subservient to vision in our perceptual hierarchy. It is invisible and does not persist long in time and space. As Dolar states, ‘it [the voice] goes up in smoke’ (2006:15) in the moment of utterance. Fine Art has broadened into the widest spaces and places, from the internet to environments other than galleries, cinemas and so on, that have shaped how reproduced voice is presented and represented in contemporary experience. There are historical precedents in Futurism, Dada and performance art that inform ideas about the voice as object and material. Recent practices within the realm of relational aesthetics, as experienced, for example, in the work Tino Seghal focus on the impermanence of relations, of which voice is a prime example, whereas I argue for the object voice that has a partial relationship with the visual.

Kim-Cohen (2009:xx) suggests that a theory of sonic art should be non-cochlear in the way that Duchamp championed a non-ocular art: Š he argues that we should attend to the conceptual context as much as the sound of the voice. Kim-Cohen believes that it is more difficult to re-order sound, in the way artists rearrange visual material, because sonic materiality is always in a flux that is different to the flux of visual objects. Modernism, for example, meant that looking into the painting as an illusionistic window changed to looking at the painting, but the voice is a more complex partial object. If the autonomous artwork is not to be found in sound, I argue that the voice is a particular kind of sound with a partial autonomy.

While voices convey language, they are also abstract sound forms, which can be considered alongside the visual array, which they usually accompany, and this is how

---

6 Duchamp thought that art should address more than the visual realm, and Kim-Cohen extends this to the contextual and conceptual aspects that come with our experience of sound in art.
the voice is approached in the following chapters, as an object and material. The basic and direct address of artists is examined in terms of materiality, subjectivity, the relationship with the moving image and the invisible, ideological structures framing the world and its representation.

Ferdinand de Saussure (1983) described how language involved both an overarching system and individual utterance. Voices are a medium of language but Fine Art practice stretches the voice into new relationships in space and time. These factors are always partial: the voice is a partial object, a partial material, and therefore a theory of voice must be multiple. This thesis, therefore, explains the voice by applying existing theories to Fine Art practice where they have not been previously applied as a contribution to knowledge.

My argument, expanded from Mladen Dolar's book (2006), is that the voice used in Fine Art practice is not just the musical aesthetic of singing, or merely a leftover of speech, but is a material with a structuring agency found in historical avant-garde and contemporary practice. This thesis and practice is an interrogation of that argument that the object voice is theoretically un-regarded, yet a common presence in contemporary art. This structuring agency, for Fine Art, has a discourse around objects; their representation and their sonic materiality, and this has a bearing on the voice as a partial object. I use ordinary voices, but my practice involves an intention different to our everyday use of voices, and this is also true for the other examples I discuss.

Voice and image are related in various ways, sharing duration as synchronous, asynchronous, simultaneous or contemporaneous. Their flows obey physical formulas as fricative motions, between laminar and turbulent flows, noise and coefficients of restitution. I intend Trialogue to activate such aspects and relationships between sound, image and voices, and also between practice and theory.

The voice is used and felt with the familiarity of a part of our body, and this is the starting point for discussing and using the voice in Fine Art practice. My method has been to allow myself free reign to use the voice in the studio and edit suite, a space with an open structure, where disposition between voice and image can be tried out. Doing this garners a phenomenology for the voice in Fine Art practice. We may think the voice belongs in a recording studio rather than an artist’s studio, but my work comes
from a ‘lo-fi’ tradition similar to Art Povera and other avant-garde moving image work. I am investigating the basic disposition of a voice with images. The everyday activity of listening and looking is critiqued in Fine Art practice but this is not often extended to the human voice. How is the agency of the voice different for Fine Art practice? The voice is a pivot in the production of meaning, but Fine Art practice transforms sonic material into more than linguistic meaning.

I have used the recorded voice in a series of previous works that ignited my interest in the differences between the sound of the voice, myself, and moving images. I thought of the voice as having what Gina Bloom (2007:3) calls the ‘Non-verbal meaning of the concrete physical voice’, much like the plays of Richard Foreman, where recorded object voices abound alongside live acting and, as Bloom (2007:6) states:

There are important similarities between stage props and voices. Like props, voices often are imagined as unmanageable, beyond the control of those who ostensibly own and control them.

My PhD research and the genesis of my work, has been driven by the desire to unlock and think through the function of the voice in my video practice. This is extended to the practices of other artists, whose work contributes to the evolution of how voices are used and experienced by their audience. My own experience of using voice in my video practice alerted me to the duality that voices reveal. This ranges from voice standing in for self, to its use as a material and object separated from self. In both my own and others artistic practice the voice is a direct and available tool, but it can also be perceived as an object emphasised by reproduction in the work of artists that I have selected and analysed. My research sets out to interrogate and draw conclusions from the complexities leading from such disparities and dualities.

My rationale for using video was because it electronically encodes sound alongside image. This eases the entry of voices into the work, and even more so with digital

---

7 Hito Steyerl has written about ‘poor image’ – [http://www.e-flux.com/journal/in-defense-of-the-poor-image/(2009)] – but this can be extended to the voice, which is also used across a variety of production formats and sound quality.

8 For example, *Meat* (2003) used an archive recording of an eponymous music hall song and *Language Acquisition in the street* (2006) treated archive voice as detritus in the street. *Otolith* (2003) is a video essay, the result of an artist’s residency in the Yuri Gagarin Cosmonaut Training Centre in Star City in collaboration with The Otolith Group. In *Unovercryable* (2009), I worked with an actress to speak about images of urban streets in order to evoke both walking through the streets and the way in which memory works with actual experience and personal stories.
media. My research focus explored the different situations, dispositions and genres where voice is produced and received. My aim is to describe how the voice as a material object is situated in different discourses, in both my own and other Fine Art practice, and therefore throws new light on how the voice can be theorised.

The relationship between object voices and the discourses through which, I am arguing they are also created, requires the application of multiple theories due to a complex mix of vocal materiality and meaning. This plurality is reflected in each chapter, which places a different emphasis on the voice, its materiality, and its role in Fine Art practice. My literature review is embedded in the relevant chapters that I summarise at the end of this introduction. The rationale of the written thesis is that the object voice is introduced and explored in the case studies, then discussed as reproduction, and in relation to images and bodily self. Phenomenological and pragmatic theories are then introduced. My video installation *Trialogue* receives theoretical reflection throughout the thesis with particular emphasis in Chapter 7, *Voice and Trialogue* that informs Chapter 8, *Voice and Listening*.

Different quantitative methods could be used to track the presence of voices in Fine Art practice across the twentieth century. However, although aware of the different dynamics in art history where the voice became more dominant, I have chosen specifically the practice of Bruce Nauman, Janet Cardiff and George Bures Miller as the case studies for my discussion of genre and disposition that are key to the thesis. A burgeoning of voice can be found in the arrival of 1960s performance art, and I use examples of 1970s video practice by Martha Rosler and others that test the boundaries of a direct vocal performance to a camera. The specificity of the case studies is not to ignore the activities of earlier avant-gardes such as Futurism, Dada and their wide influence on voice in twentieth century art practices. The vocal performances of Hugo Ball and Kurt Schwitters, the concrete poetry of Bob Cobbing and contemporary practice, for example, of Mikhail Karikis all probe the limit of how we define voice. Their works isolate and bracket out voice, whereas my chapters and case studies investigate the voice in relation to discourses of image, space and place. Now, in the twenty first century, interest in relational practice exists alongside the feedback loops of contemporary social media.

Some of my terms shift their meaning. The term *image* is used in a broad sense as the
works I discuss by Bruce Nauman and Janet Cardiff have no images, only voices, but relate to particular spaces. I use the figure of a palimpsest as a site where image and sound meet, and more precisely voice and image. This is an expansion of this term from the visual into the sonic, pertinent to the voice being a partial object. The object voice is partial, and this extends to the tools and terms that surround it. This shifting nature also applies to some of my key terms such as disposition and formant. The formant frequencies of the voice are experienced on the telephone, as disembodied sound, different to their owner’s presence. The idea of disposition is used to investigate how the physical experience of a voice relates to space, screen – or which ever discourse. This term – disposition – opens up the discourses that surround utterance that lead to the creation of vocal genres within which voices are recognised. Rather than the more reductive performance found in the artists listed above the voices in Nauman and Cardiff’s practices allow an investigation of space, time and milieu; for example, in the minimal space of an empty gallery by Nauman and the maximized experience of the outside world in Cardiff’s audio walk. My use of the term disposition comes from recent film scholarship, for example (Albera and Tortajada, 2010), and this follows Michel Foucault’s dispositif discussed by Giorgio Agamben (2009), and explained in Chapter Four.

In Chapter One I introduce the idea of the object voice that I use as the main theoretical description for the voice in Fine Art practice. The object voice is a theoretical construction intended to account for the difference and gap between subjects and the objective world. In this research I am deliberately putting my voice into this division as something also mechanically reproduced.

This thesis brings alive, and theorises, the object voice of everyday use as experienced and objectified in Fine Art practice rather than the aesthetic object of singing or the medium of linguistic messages. Dolar’s book (2006) draws attention to a lack of thought about the voice, which he suggests is ignored as linguistic meaning is received. Dolar argues that the voice is known for both phone and logos but there is also an object voice, ‘an object which functions as a blind spot in the call and as a disturbance of aesthetic appreciation.’ (2006:4). Michel Chion (1999) also identifies this idea of the voice as something other than language or singing in his discussion of the voice in cinema. In Chapter Two, I apply the idea of object voice to two main case studies, Raw Materials (Bruce Nauman, 2004), and The Missing Voice (Case Study B) (Janet Cardiff, 1999).
In Chapter Three I discuss the consequences of using recorded voices. The importance and absolute difference between recorded and live voices are by no means absolute. It is important to point out a symbiotic relationship between voice, reproduction and space. If we think of echoes we can see how the idea of reproducing voices is already pre-figured before the invention of reproductive technology. As Pierre Schaeffer and Chion stress, voices are fundamentally acousmatic (we cannot see their source) and there is confusion in locating their source because they are invisible. The demarcation of an acousmatic voice raises a range of issues about liveness, representation, time and space that are discussed in this chapter.

Chapter Four proposes the idea of disposition to explain the position and meaning of a voice in relation to an image. A basic disposition of voice and image is that we speak about what we see. This is taken as a fundamental way in which voice and image link up in an embodied phenomenological relationship. The presence and absence of image and voice as object stalk the practice like the tip of an iceberg breaking the surface of the ocean. Where a voice is the creation of a set of bodily rhythms and functions, via breath, the image represents the rhythms of the world. Again, the voice is both inside our self and outside in the world, contributing to both orders of rhythm. This is extended to synchronisation between voice and image as simultaneous or contemporaneous.

The relationship between voice and image also relates to the antagonistic relationship between the real and abstract, narrative and non-narrative in Fine Art practice. In Chapter Five I discuss the voice and self and relate this to how first-generation video artists used the tele-present quality of video as a type of mirror to represent the self in real time. These influential works are then compared with contemporary works using verbatim technique. Verbatim technique (sound recording as indexical record) and video tele-presence are both used by artists to question the relationship of voice to subjectivity. These are different approaches to object voices. All of these works actualise the voice as a material presence, but a presence that is articulated by the structures developed in the course of this thesis, such as genre, sonorous envelope, and how the voice moves between a past palimpsest into the future offing.

Erving Goffman (1981:131) wrote about the lack of a ‘substantive unit’ of talk, and this
can be applied to the voice in Fine Art practice that I aim to discover with this research. Voices always have a relationship with context, and this is noted in the work of Sperber and Wilson I discuss in Chapter Six. Bakhtin wrote about the ‘extreme heterogeneity of speech genres’ (1986:60)\(^9\) that we experience in life and art. I am looking for a way to describe the voice in Fine Art practice. The voice can be defined in different ways: for example, its limit could be thought of as earshot, which was perhaps a more stable category before reproduction and amplification. The voice can be in a sonorous envelope as well as a public sphere. I go on to describe how democracy in ancient Greece was born out of the idea of earshot. Don Ihde (2007:26) describes how ‘the aim of phenomenology is to make as precise as possible the shape of the experience being investigated.’ The scope of the voice is wider than the sonorous envelope that protects our personal space, and moves into the public sphere, widened exponentially by reproduction. I have found that voices work across different theoretical positions at the same time. Bakhtin said: ‘we speak in diverse genres without suspecting that they exist’ (1986:78).

To bring together different ideas about the time and space of the voice in Fine Art practice I am using the ideas of palimpsest and offing. Palimpsest means a scroll or page that is scraped clean for further inscription and may retain the marks and structure of previous writing. As a metaphor, palimpsest is familiar within art history, where the erasure of written or printed words extends to images, and I am extending its remit to the sound of voices. I adapt its usage to mean any space that is changeable and I investigate the plethora of spaces created by heterogeneous speech genres, where the moving image screen is a palimpsest and receptacle of object voices and their extra-linguistic sound.\(^10\) I do this because voices permeate experience: the world is pregnant with voices as communication of thought and human presence in space.

The voice is a temporal marking of a situation in space and time whose effects remain in memory: voices have many different effects despite their relational agency. Physical space, memory and expectation attend the voice as a sound overwriting time

---

\(^9\) This wide cultural span is mirrored by the decibel scale that measures the vast volume range of vocal utterance that we can detect. ‘The force acting on the eardrum at the threshold of hearing is 140 million times than the force needed to lift a one ounce weight.’ (Denes and Pinson, 2007:98)

\(^10\) E. K. Brathwaite (1992:7) describes oral poetic traditions in the Caribbean: ‘indigenous languages had to submerge themselves because the conquering peoples insisted that the language of public discourse and conversation, of obedience, command and conception should be in English.’
and space in a perpetually additive and subtractive flux. The broadening of the term *palimpsest* in this thesis accounts for different theoretical frames that contain the voice and structure its presence. By focusing on the sound of a voice we are listening according to experience and knowledge, and Fine Art practice intervenes in this process.

The voice is our habitual tool of communication, but it does more than this in Fine Art practice. Our expectation is sited in the past as *palimpsest* but also the future, which I equate with an *offing*. The *offing* is the part of the sea visible from shore that is distant but into which we expect an approaching vessel to be seen. *Offing* has a colloquial meaning, ‘in the offing’ that broadens to mean anything that will happen in the near future, or that might be expected to follow in a sequence of time. Fine Art practice also intervenes in this process of receiving voices as material objects.

Chapter Six discusses the relationships between theories that can account for the attempt to reduce the voice to material and object alongside theories that provide a context that underpin the voice. I use Merleau-Ponty’s phenomenology to explain how voice and image meet as embodied perception and how, as a partial object, the voice intercedes between different boundaries. Where a phenomenology focuses on the voice, Bakhtin’s theory of speech genre and Sperber and Wilson’s Relevance Theory both stress context for the meaning that I apply to voice.

Chapter Seven describes the four works that make up *Trialogue* and contextualises them within the theorisation I am presenting. In Chapter Eight, ‘Voice and Listening’, I propose that artworks can instigate listening in which a pre-categorical object voice moves between *palimpsest* to *offing*. Jean-Luc Nancy’s (2007) writing about listening is used to find how the attack of vocal sound brings the construction of meaning. The sound before this process can be thought of as pre-categorical and is opened up in the Fine Art practice I describe and produce. Listening can only occur in space and time and the chapter sub-divides into sections about space and time to think about what moving image work does with object voices.

In devising a methodology for describing the voice in Fine Art practice I have had to remember that the voice is a cultural phenomenon that is difficult to reduce to any singular aspect such as timbre, rhythm, or space, as attempted in both modernism and
by empirical methods. The act of listening is itself not objective and is an activity connected to our other senses and the disposition of factors important to a specific utterance. This is what generates the critique of the theorists I discuss, for example, where Merleau-Ponty criticises a scientific reduction to an object in general, exemplified in voice, or where Dolar cites the formal reductions of phonology as killing the voice. There is also a danger in ignoring language, that is, what the voice might be saying, amongst the theoretical paraphernalia with which it can be surrounded as object and material. However, my aim in what follows is to ‘listen out’ for why and how the voice is occurring in Fine Art practice.
Chapter One: Voice and Object

Voice becomes an ambiguous and equivocal name that often ends up indicating the general idea of the verbal sphere instead of indicating a contingent, contextual sonorous articulation that emits from the mouth of someone and this is destined for the ears of another. (Cavarero, 2005:14)

The voice Cavarero seeks to constitute is an individual sound not effaced by the symbolism of language or as a blind spot in psychoanalysis. In this chapter, I describe the theory of the object voice in psychoanalytical theory, which I then apply to the actual voices heard in Fine Art practice in my case studies.\(^\text{11}\) I also investigate the voice as something objectified in the repetition made possible by recording, and as a sonic material object emphasised by Fine Art practice. The nature of such object voices are then investigated in a range of examples culminating in how voices are used in Raw Materials (Nauman, 2004) and The Missing Voice (Case Study B) (Cardiff, 1999). I also discuss the recorded object voice in The Future’s Getting Old Like the Rest of Us (Gibson, 2010), as well as works using extra-linguistic sound by Marina Abramovic, David Critchley, and George Barber. I look at how Martha Rosler and Ian Bourn use their voices in performances recorded on black-and-white videotapes made in rooms in the 1970s.\(^\text{12}\) By examining these works I encounter object voices in action as direct vocal address in different modes presenting plural theoretical discourses necessary to theorise the voice in Fine Art practice. The object voice is a psychoanalytical concept but describes voices in Fine Art practice that are also culturally constructed and objects of phenomenological inquiry that cannot be ignored in this context. The practical and theoretical questions thrown up by these works inform my own vocal address in Filling Space with Rocks and Forget Me and those of the voices in Just Talk and The Jazz Singer, seen on the three screens of Trialogue presented as a multi-layered grouping of object voices.

These works use voices in distinctive ways whose particular effects I describe, raising the contradictory issue of how the object voice can be stripped of language, thereby intervening in the chain of causation between thought and vocalisation. Does focusing

\(^{11}\) Although the idea is created by psychoanalysts, they do not record and analyse the voice itself.

\(^{12}\) These spaces feel like originary spaces for video art. Where Bruce Nauman recorded actions in his studio, or Joan Jonas, in Vertical Roll (1972), conflates monitor/screen/gallery, these domestic spaces resonate as a disposition in video art like the Black Maria used by Edison in early cinema.
on the object voice reverse this process, where the voice can be the starting point in these works? We can hear the sound of the voice and not the language it delivers in various ways: for example when we are half hearing, hearing a foreign language, or if the sound is impaired; when someone is speaking with a mouthful of food or interference is effecting a radio broadcast. What is the distinction between a voice that is making a noise like screaming, or coughing and a voice on whose sound we concentrate\textsuperscript{13} such as the singing voice? Slavoj Zizek goes as far as to suggest:

Music is the last curtain which protects us from directly confronting the horror of the vocal object. (1996:93)

In other words, the voice is structured by rhythms that centre it within communication and culture without which it is a raw sound. The voice is never only a raw sound, yet does not exist without raw sound. Music provides a structure outside of language but there are other structures and rhythms to be discovered for voices in Fine Art practice.

The object voice is a theoretical construction but voices are experienced as real in art practice. Psychoanalytical theory defines the object voice as a blind spot, something we cannot apprehend. I am asking how the extra- and para-linguistic qualities of voice affect moving image art practice. For Lacan, the voice is in the realm of the real more than the imaginary or symbolic. I am questioning whether the object voice of psychoanalysis is the same as other theoretical constructions of extra-linguistic vocal sound, the voice as the undifferentiated phonemes that come before linguistic meaning, or the paradox of timbre (which I explain later). Frederick Kittler discussed the status of reproduced sound, which reproduces the real material qualities of sound, and has not passed through the ‘bottle neck of the signifier’ (1999:4), like language. It follows that vocal sound surrounds speech, but is this sound symbolic itself when removed from an original context by reproduction? Is the real unknowably in opposition to language? A grammar for voices in Fine Art practice intercedes in these categories where different dispositions, genres, sonorous envelopes support object voices.

Fine Art practices, that centre the object voice in questions about the relationship between representation and subjectivity, loosen the voice’s ownership to become an

\textsuperscript{13} It is possible to hum, click the teeth, and slap the flesh with the mouth, but the latter two sounds are considered ‘unvoiced’ for the majority of language groups.
autonomous object which in turn change perceptions and framings of discourse and space. The object voice strikes a chord for my work, but my use of the term is less strictly psychoanalytical than in Dolar’s writing. So to describe the voice in my own practice and other examples I am taking the object voice that passes through psychoanalytical theory from Freud,\textsuperscript{14} Lacan, Winnicott, Zizek and Dolar, and combining it with direct instances in art practice. Dolar writes about an object voice that Lacan develops from Freud’s category of partial object that also includes the gaze (\textit{object petit a}) and comes from the theory of object relations. This posits the voice as something outside of language and related to an impossible relationship with desire. Where Lacan (1994:168 (original 1977)) describes partial objects, Winnicott (1991:4 (original 1971)) uses the term transitional object. These two ideas represent the similar ideas of their authors and are helpful models for theorising the voice as an object in art practice. In Winnicott’s theory, the mental space that a transitional object occupies is neither subjective nor objective but partakes of both:

The precariousness of play belongs to the fact that it is always on the theoretical line between the subjective and that which is objectively perceived. (1991:50)

The voice can be, at once, part of the subject and not the subject, and a transitional process that involves oscillation, or a shifting of the boundaries, between inner and outer – as opposed to stable, fixed structures. Lacan coined the term \textit{extimate}\textsuperscript{15} to describe this quality of being inside and outside, something that is in between: the appearance of the Real in the symbolic.

For me, the voice is a blind spot because it is very much present, yet invisible. The voice has been theorised as both transparent and opaque: Dolar (2006:42) points to the theoretical paradox of the voice being, on the one hand, the measure of self as presence, while at the same time being something we can hide behind, a difference I describe as production and performance. A voice has to be produced, by our bodily viscera beyond our conscious control, at some scale, and is a sound through which others can interpret our subjectivity, like a doctor using a stethoscope, whereas performance requires a conscious self-control. Whether voice is produced or performed, hearing and listening can be compared to the glance as well as the gaze.

\textsuperscript{14} Freud, Sigmund (1905: 222). \textit{Three essays on the theory of sexuality}.

\textsuperscript{15} The term \textit{extimate} – the real expressed in the symbolic – is explained by Dolar (2006:81).
Derrida\textsuperscript{16} critiqued the voice as being the presence of the present, while for psychoanalysis the voice gets in the way and obscures the self but is also a form of legible dissimulation behind which we can connect. Despite these critiques it is difficult to dismiss the feeling that the voice is in a close ordering with the self. In art practice, meanwhile, the mediated voice can be framed as an autonomous object or sound even if it did come out of a mouth with a partial relationship to the construction of subjectivity. In film the mouth is the site of lip synchronisation, but also where the voice moves from the synchronisation of body and mind into synchronisation with space and environment captured and critiqued by Fine Art practice. These aspects of voice are explored in the critical arena of art practice and pushed into many unexpected positions, as a tool and material that is not effaced but is neither altogether originary in the subject.

I use the Object Voice of psychoanalytical theory as the generic form of represented voice in moving image artworks and installations. Instead of being outside of language and meaning, or after the event, the voice becomes a positive material and presence. Analysing its functions in Fine Art practice can then break down the object voice into further categories. A voice comes out of the mouth of its owner, who produced it, but goes elsewhere and is no longer their voice: it then belongs to the listener, who hears or buys it, or just plays it as a recording, or understands something it communicates. It has a currency, and has entered the ambit of the other, passing through their ear and changing something in their mind.

A voice is not an object that solidly extends in space, although we perceive it in space as a sound leaving the body and moving elsewhere. The difference can be compared to phenomena, such as light, explained as both waves and particles. Instead of Shannon and Weaver’s diagram\textsuperscript{17} of communication, the voice as partial object is often found in the intersection of Venn diagrams, for example between subject and other\textsuperscript{18}. The object voice of psychoanalytical theory is an attempt to account for the voice as both subject and object that Lacan describes as a leftover of the signifying process. The object voice experienced in art practice becomes an object that exceeds the subject, or a residue alien

\textsuperscript{17} Shannon, C. and Weaver, W. \textit{The Mathematical Theory of Communication}. Urbana: University of Illinois Press (1949). In this diagram a message is coded and decoded in a process of communication.
\textsuperscript{18} Hopper (1992:25) also shows how Saussure’s diagrams of communication are contemporaneous with the telephone. The diagram he proposes involves the circuit used in telephones where the electricity stands in for the voice.
to the subject. Psychoanalytical theory stresses the noise of the voice as outside of linguistic meaning and noumenal, rather than a sonic material ordered by the practice of art. I wish to emphasise how the environment shapes the sound of the voice before it is perceived. This is evident in *Raw Materials* (Nauman, 2004), although the abstraction of language does a very good job of obscuring, or pushing out of consciousness, these sonic qualities; hence the object voice. The object voice is such excess in Fine Art practice, as Sean Homer notes:

The object voice, in other words, is both the gap that opens up between subject and other and that which resonates within the gap. (2010:45)

This can explain how the voice fills space and defines the path of a relationship as physical presence and this space is structured between subject and material world. Voices without language impart different forms of information: chiefly, they possess affect. It is not possible to discern an accent in coughing in the way a mechanic listens to an engine. There is no range of pitch in whispering, but there is a loss of intonation in shouting. In noises like sneezing and shouting very little subjective information is received. The usual frequency range of voices tells us most about the speaker, such as accent, and health. If I listen to a speaking voice can I ignore its language and treat it as a sound? Is it always the sound of something, or is it an autonomous object and abstract sound? The question of an abstract voice is answered by the structural answer, whereby the voice is partially abstract; its sound is abstract and its language is not. However, the forms of language are also abstract, or arbitrary. This is where the object voice relates to the modernist autonomous art object.

Antonin Artaud advocated a radical vocal performance in which the voice emanates, and is heard by the whole body, rather than funnelled from mouth to ear. In this formation the voice is a whole frequency of presence quite different to how the voice produces and transmits language, which is a particular set of reductive rules and grammars. There is no easy divide as the voice bridges and is embedded in real materiality at the same time as the symbolism of language. In this way the voice is a meaningful noise that Kahn describes as: ‘that constant grating sound generated by the movement between the abstract and empirical’ (1999:26). Dolar shows how Lacan places the voice as: ‘the outcome of the structural operation’ (2006:35), but my research and practice with the voice questions the order of such an operation. The grating sound
is left over rather than something at the beginning of a process that is then ordered, but in art practice the process can start with a voice. Different voices as situated utterances in artworks produce different kinds of object. The voice as aesthetic and phenomenological object is challenged by the post-structural analysis of power structures by Foucault or Derrida’s investigation of the voice as auto-affection\textsuperscript{19}. The first focussing on the authority wielded through law and the second questioning the voice as most authentic of self. Bakhtin (1986:63) cites the identification of genre as the first decision at hearing speech, that I extend to voice, and Chion (1999) argues that the logic of the voice involves placing it with its source, the body from which it was emitted. I discuss this in relation to Trialogue in Chapter Six. The voice is an object that has become divorced from the subject, but is it permanently disembodied as a material? Chion argues that the voice is reunited with a mouth in representational film. The voice cannot go back into the body, but neither can it be wholly disassociated from the subject.\textsuperscript{20} The voice refuses any easy placement, it corresponds only partially with body, place or discourse and this is what elevates the voice in Fine Art practice from normative, everyday tool to something that questions our very being. Dolar suggests that when hearing a voice, the first question is whether it comes from within or outside our selves:

Where does the voice come from? Where do we hear it? How do we tell the external voice from the voice in the head? This is the first ontological decision, the first epistemological break, the source of all subsequent ontology and epistemology. (2006:81)

This decision has to be made by the subject because if we cannot make the decision we are hearing imaginary voices. We accept our inner voice, whereas accepting other imaginary voices as real is delusional. This is never far from the uncanny contradiction of an invisible presence produced by voice.

---

\textsuperscript{19} Auto-affection is our self-awareness that Derrida questions in his critique of logo/vococentrism: ‘Auto-affection is a universal structure of experience. All living things are capable of auto-affection … Speech and the consciousness of speech – that is to say consciousness simply as self-presence – are the phenomenon of an auto-affection lived as suppression of difference.’ Of Grammatology, Baltimore: Johns Hopkins University Press (1976:166).

\textsuperscript{20} An exception to this could be the dubbing of voices in films, a subject on which both Artaud and Borges wrote scholarly critiques – see Mikhail Yampolsky ‘Voice Devoured: Artaud and Borges on Dubbing’ in Stidworthy (2012:155).
A related phenomenon is hearing voices when we should not. We can block out what we hear if we think we should not be listening, and we can hear voices that do not exist caused, for example, by the effects of the wind. Our perceptual system is pre-structured to hear voices in sound in the way we see faces in clouds. It is possible to listen to strings of words and hear other words, sometimes with comic effect.\textsuperscript{21} The effect is not only comic where ‘electronic voice phenomena’ heard in radio signals are interpreted variously as the voice of god, aliens, unknown coded messages, or the voices of the dead.\textsuperscript{22} This is a reversal of artistic interest in abstract sound. Artists put voices in unexpected places, speaking out of turn, and in doing so shed light on the grammar of the voice as object in experience.

The object voice is a positive presence rather than a sound we ignore; it is the material and object of experience. Just as hearing a recording of one's own voice provides distance, and makes us hear ourselves in the way others do (to the extent that we don’t recognise ourselves), so hearing a voice as Fine Art practice demands a focused listening where the audience become conscious of the construction of meaning. A cultural investigation of the voice is all embracing because human culture is so connected to the voice as cultural medium of language or singing. The voice is culture: however, for Fine Art practice the voice is related to the visual realm, and artists have experimented with visual language and visual song (this connection is raised in the phenomenology of Merleau-Ponty, which I visit in Chapter Five). The pre-cultural voice becomes hard to detach from culture. The voice is something other than the speaking subject, although we recognise others, and ourselves, by voice, since it serves as a unique index. The voice is also an event and process, yet we perceive it like the presence of an object, particularly as a reproduced and repeated form in artistic practice. On the one hand the voice is an origin, coming from a subject in space, but at the same time it occurs as the component of a process.

Our own voice, the one we hear as we speak, is too involved in the construction of our subjectivity to become a separate object but we can feel its vibrations in our mouth, and body, as we speak. Our own voice has the weakest relation to the part object. Recorded

\textsuperscript{21} In Bigmouth Strikes Again (The Smiths, 1986), the lyric ‘I was only joking when I said I’d like to smash every tooth in your head’ can be heard as ‘I was only joking when I said I’d like to smash a beetroot in your hair.’ There is a vast amount of such phenomena that also inhabits all mishearing of speech. Another example is the title of a book about the vocoder by Dave Tomkins, How to Wreck a Nice Beach (2010), which sounds the same as ‘How to recognise speech.’

\textsuperscript{22} This is investigated by Joe Banks – rorschauchaudio.wordpress.com/
voices and those of others detach most strongly from the subject as material and object. As an object the voice is strange to us as we are subjects, but this is less the case for people who use their voice as singers or lecturers and are more aware of the voice as an object. In my work *The Jazz Singer*, a singer talks about his voice in these terms, where his voice and his body are his instrument as well as the site of his subjectivity. In talking about his singing voice the singer is discussing another entity, and at one point he breaks into a type of sound like the difference between an aeroplane moving along the runway during take-off and being in the air. The singing voice is like the engine warming up and running smoothly before the flight, which is the sonorous content of language. Connor has described what *The Jazz Singer* does where song is present in all voice: ‘I can break into song, but my voice is already song breaking into, breaking out in me’ (Connor, 2012: 64). In this sense, all voice participates in a rhythmical activity as well as a raw sound.

We can think of intonation, timbre, accent, rhythm as always present in voices producing language: but what is a voice without language? In my case studies I am going to move freely between a number of artists to discuss the voice without language, or voice whose coincidence with language is compromised. Noise is simultaneous with language, the extra-linguistic parts that occur during speech such as ums and errs, coughing, laughing, onomatopoeia, screaming, humming – and the list goes on because we do not have a full vocabulary for such sounds and there are many context-specific examples that will only come to mind within those contexts. ‘The voice of …’ becomes ‘the vocal sound when …’ but Beatrice Gibson answers this conundrum in *The Future’s Getting Old Like the Rest of Us* (2010). I will discuss this work in more depth, as an example of ‘verbatim technique’ in Chapter Five, ‘Voice and Self’. In Gibson’s film, a work based on the radical disjunctions of the B. S. Johnson novel *House Mother Normal* (1971), the voices of the characters are listed like objects and referred to as letters of the alphabet or evidence in a court case.

A – Chummy, cheeky, tone always the same whatever it says
E – Monotone, dull, trite, conscious of femininity, unconscious of repetition, superficial, hearing impaired
F – Slow, stuttering, stammering, shifting volumes, humble, sensitive, unguarded

---

23 ‘Verbatim technique’ – the ability to make an exact copy – can be understood as a way of using and interrogating the indexical representation of sound.
G – Deep, deliberate, polite, pensive, careful not to interrupt, stoic, smoker

Semiologist of sound Theo van Leeuwen (1999) and ethnomusicologist Cornelia Fales (2002) have noted the lack of words available to describe the sound of voices. The list of vocal attributes in *The Future’s Getting Old like The Rest of Us* (2010) range from the scientific to the cultural as if to emphasise the heterogeneous range of different discourses within voices. The voices of elderly subjects are out of sequence, and subjectivity is reduced to the neutrality of a list. Connor (2000:7) links the common dislike of recordings of our own voice to: ‘a loss of the cathected familiar’ 24, and this haunts the voices heard in Gibson’s film. The experience of a long life, heard in the sound of the voice, also causes a stalling of the present by too much of the past. The unique characteristics of each voice forensically listed and enunciated are troubling. As Cavarero comments, the effacement of vocal sonority means that ‘uniqueness does not matter, is epistemologically inappropriate.’ (2005:9)

The voice moves from the human subject to sonic object, as a human production, but the reproduced voice in art practice is received as an object, as something driven by discourses within art about art objects. The voice is also a commodity as the reified sonic material of popular music. The voice used by first-generation video artists such as Joan Jonas, Vito Acconci, Marina Abramovic and Bruce Nauman is involved in a milieu in which the art object has been dematerialised. These voices have equivalence to the supposed neutrality of black-and-white photography of that period and are sometimes intended to have the quality of Gibson’s object voices, listed and labelled like exhibits, whereas object voices in *Raw Materials* are presented like sculptural objects.

However, this is not how we usually receive voices, which pass in a specific envelope of time, place and situation, rather than something persisting at a point in space. All these factors appear to be simultaneous, and once the voice ends its vital production moves into memory. By presenting voices as objects Nauman allows the audience to bring a slower consideration to what they hear, which I also do with my voice in *Forget Me*. James Lingwood describes how the specificity of place entered the universal mode of modernist art objects with minimalism:

---

The most concise articulation of this radical rethinking was Carl Andre’s – “Form=Structure=Place”. The possibility of thinking these categories as equivalent. (1993:23)

In Fine Art practice the reproduced voice is emphasised as object and material and sometimes this means, for example, the percussive vitality heard in *Torture* (Nauman, 1987), featuring shouting clowns or, by contrast, the almost inhuman susurrations heard in (. ) 25 (Stidworthy, 2011), where voices have ‘the impalpability of sonorous vibrations’ (2005:4), as noted by Cavarero. 26 It is necessary to link the sound of the voice to its space and place. How do we account for the relationship between form, structure and place for the object voice in Fine Art practice? The duplicity of the voice can be expressed as sound and structure, the balancing of which covers a wide spectrum of voice in Fine Art practice, as I investigate in the next chapter.

---

25 The title of Stidworth’s installation is this symbol – (. ) – which is used by translators to represent a break in speech.
26 We adjust our hearing for such different voices, the affective tensions caused by an obstreperous drunk encountered on the street, for example, or the gentle sound of happy background chatter at a social event, and audiences perform such focused listening to voices in art practice.
Chapter Two: Object Voice in Case Studies

This chapter explores the object voice in different case studies. The use of the voice as an object by Bruce Nauman in *Raw Materials* (2004) installed in The Tate Turbine Hall has a very different disposition to Janet Cardiff’s voice heard in headphones on the street. The voice is always a sound with a structure that I am describing and theorising for these different works in ways that can apply to other Fine Art practice using voice. The human voice can become effaced to deliver linguistic meaning just as the spatial aspects of a voice are also obscured or become difficult to unravel.

We are familiar with the structure in *Semiotics of the Kitchen* (Martha Rosler, 1975) and *Lenny’s Documentary* (Ian Bourne, 1978) that have a televisual frontal address as ‘pieces to camera’, or documentary, as opposed to the voice as pure sound in *Freeing the Voice* (Marina Abramovic, 1976), *Pieces I Never Did* (David Critchley, 1978) and *India Shouting Match* (George Barber, 2010). In such works, the voice becomes a freer sculptural object unbounded by the performance of character; it is released from normative moving image structures, and pushed to limits of sound with no language.

Ong (1982:121) has commented, that ‘there is nothing as neutral as a list’. This statement can open up the difference between the immanence of vocal production and the rigid orderings of genre, language and writing. Gibson’s list of vocal characteristics in *The Future’s Getting Old* is anything but neutral as it attempts to gauge a salient mixture of prosody, accent, volume and so on for each specific voice: the facets that are often un-noted in language but still absorbed by listeners. Below, I discuss *The Semiotics of the Kitchen* (Rosler, 1975), essentially a list, but definitely not neutral in the manner of its utterance. The list can relate to the utterance in the way the grid is applied to modernism in visual art. The list is a phenomenon of writing, the neutrality of which is easily subverted, for example if its order is changed. Voices don’t really recite lists, which belong to the realm of the written word. However, such stripped back structures are used in many examples of art practice. Once the voice utters a list more meaning is opened up, and Rosler uses this to her advantage. There are also notable list songs (the list song is a genre in itself), jokes and narratives that play with or subvert structure, which, as Ong says, is never more firmly stated than the ‘one thing after another’ of listing. We can compare the vocalised list in Nat King Cole’s *Let there be love* (1961) and *Semiotics of the Kitchen*. Cole sings ‘You, me, oysters, sea, wind, rain,
chilli con carne, sparkling champagne, birds, trees, blessing, sneeze, cuckoos, lark, dove, love,’ whereas, ‘Apron, bowl, chopper, dish, egg beater, fork, grater, hamburger press, ice-pick, juicer, knife, ladle, measuring implements, nut cracker, opener, pan, quart bottle, rolling pin, spoon, tenderizer …’ is spoken by Martha Rosler. Rosler’s list is demonstrably less lyrical than that of Cole, the utensils raw and violent. Cole’s singing produces a pure form of sound that Dolar calls a ‘fetish object’ (2006:4): the objects Cole lists are symbolic and his voice smoothly distant from their sonic realism. This stands in stark contrast to the critique inherent in Rosler’s videotape.

*Semiotics of the Kitchen* is indicative of the simple address to camera used in early video art to interrogate the medium’s tele-presence and the video monitor’s similarity to a mirror. By analysing the simple addresses to camera of Critchley, Bourn and Rosler I contextualise the meanings of my own address to camera in *Filling Space with Rocks* and *Forget Me*.

For some first-generation video art, a concentration on materiality was associated with an apolitical formalism. This view can be found in the work and writing of Martha Rosler (1990:31) and persists in theoretical debates between cultural, psychoanalytical and phenomenological theories. The politics of representation is never far from critical moving image practice in Fine Art and is epitomised by Rosler who expounded video as a practice that should engage critically with the institutions of both art and the media. Her work can itself represent such debates of the 1970s; Rosler suggests that an attempt to define video as a medium would be to deny politics, and issues of video as medium share ground with voice as a medium. The categories of ‘voice’ and ‘video’ have an uncomfortable status as media because they are plural practices. The voice is more than the medium of language, and the nature of video changes with the convergence of digital media. Yet the voice is an important element in *Semiotics of the Kitchen*. More recently, Yvonne Spielmann (2010) has provided a book-length exposition of video as a medium. While digital convergence makes some of these arguments about medium specificity redundant it also renders digital and analogue video two different kinds of medium. Like live and recorded voices there are a range of similarities and differences between analogue and digital video, representing a continuum rather than disjuncture in technological history.

27 Artist Ed Atkins explicitly describes his work as an investigation of the materiality of digital video.
In *Semiotics of The Kitchen*, Rosler (1975) exhibits a restrained violence of action and voice in a performance to camera with kitchen utensils involving an ostensibly neutral, alphabetically ordered utterance of a letter for each utensil to which she mimes actions. The sounds of the utensils have a sonic realism different to that of television cookery programmes. This is a symbolic dry run without food, when utensils would normally join food in the process of cookery. Rosler utters each letter with a curt anger and the reduction to a list of letters becomes a reduced rhetoric full of political meaning.

This is a feminist critique of women’s role in the kitchen, as portrayed by 1950s television commercials and cookery programmes. A discourse around television is present, therefore, although the video is unlike television and relates more directly to the feminist discourse ‘the personal is political’. It is not what she says but what the viewer must infer from her performance that is significant, which ends with a shrug, as if to say ‘what did you expect’. The mode of television is criticised here: women portrayed as objects of commodity fetishism. The measured utterance and slow enunciation of words gives the list format a menace shown in the way she wields the kitchen utensils. This list of words is both blank and loaded with contextual meaning. Presenting the utensils in the style of a teacher telling us something new emphasises the constructions made by a hearer by implication. Modality is at odds here where femininity itself is a construction.

By stripping the information back, Rosler emphasises the processes of communication that are unspoken; those not discussed but already accepted. It would be difficult to read her voice as anything but a very direct communication. She is spelling it out for us, yet not spelling it out for us, as we need to spell it out for ourselves. Her voice is the voice of a feminist critic. Its affective quality is of seriousness, but also has the deadpan delivery of slapstick performance. The speech genre here is a parody of modernist performance. The face does not fall into coincidence with how it often confirms the meaning of the voice.

---

28 Seen on this website on 1/5/2013 – http://www.youtube.com/watch?v=3zSA9Rm2PZA.
29 The work is a sardonic response to structural film and in particular a parody of *Zorns Lemma* (Frampton, 1970), which is attacked for a lack of political critique with the final lunge and flourish of the knife, like ‘Zoro’ describing the letter ‘Z’.
30 The political arena is not isolated from individual action. A disenfranchised group can initiate politics from the bottom up. This is discussed in ‘The Personal is Political’ (1969) by Carol Hanisch and is extended to feminist video work by Catherine Elwes in *Video Art: A Guided Tour*, I.B. Taurus, 2004.
Lenny’s Documentary (1978) is a piece to camera in which the body is framed in a similar way to Semiotics of the Kitchen. Ian Bourn plays Lenny, a man planning a documentary script about his life, and it’s a work prescient of video diaries, social networking websites mixed with ‘kitchen-sink’ realism. Like Semiotics of the Kitchen, it critiques the modality of television but it can also help us think about the contemporary construction of identity through representation on the internet. Is Lenny talking to himself or to an audience? Lenny’s is a marginal voice using vernacular constructions. There is an oppressive banality to these vocalisations; an investigation into how we portray ourselves, the ambiguity of the diarist’s ego. His voice is a channel to directly recycle experience in a performance not unlike contemporary verbatim technique, or the amateurism one encounters on YouTube.

Before moving on to a discussion of the disembodied voices of Nauman and Cardiff, I will discuss the full-bodied voices, fully throated by their performers in front of the camera, in: Freeing the Voice (Marina Abramovic, 1975) [view clip 4V]; Pieces I Never Did (David Critchley, 1979) [view clip 5V]; and India Shouting Match (George Barber, 2010) [view clip 6V].

I have chosen these works as a sample of extra-linguistic sound that expands our experience of timbre and can be used to pursue the voice as object. These works present the object voice as described by Zizek:

I hear myself speaking, yet what I hear is never fully myself but a parasite, a foreign body in my very heart. (1996:103).

Shouting is a quick way to eject this sonic material, foreign body, or object voice at high volume. We cannot shout quietly, although Abramovic seems to be shouting methodically, in line with the minimal and conceptual art historical milieu of late modernism she shares, as an intention, with T. S. Eliot’s mannered performance mentioned in the introduction. Critchley inter-cuts a ‘piece to camera’ with a similarly visceral vocal assault involving loss of voice, and the mediation is emphasised with an attack of jump cuts. Barber arranges and films a shouting match, with rules that focus vocal energy as cinematic spectacle rather than as a pure excess and loss of control. Dolar (2006:103) gives a diagram in which subject and other overlap at the voice, and in the shouting match the overlapping is reduced to the power of who is loudest.
The meeting of breath and genre causes the effects in these different works. The acculturated content of speech provided by teeth and tongue has been shorn off, and the vocal sound of shouting is at different intersections of genre and structure in these works.

When is it usual to laugh or shout? Are these sounds cultural or automatic responses? Barber has tried to get people from different cultures to shout in order to reveal difference in the social constructions of shouting. Where Barber gets people to shout to make a video work, Robert Povine (1999) gets people to laugh for academic research, and in more everyday contexts this can be extended to getting people to talk. As I remark in *Filling Rocks with Space*, getting out a microphone has various consequences that let us know about the power discourses of the human voice. The limelight can repress our words and the voice is stuck in the throat. Zizek says ‘the true object voice is mute and stuck in the throat’ (1996:92), not because we are emotionally choked up, but because the voice is an object pervading experience:

Voice and silence relate as figure and ground: silence is not the ground against which the figure of a voice emerges; quite the contrary, the reverberating sound itself provides the ground that renders visible the figure of silence. Timbre is the canvas. (1996:93)

Feminist film scholars such as Kaja Silverman (1988) have analysed the female voice that is exhausted of language in a scream, what Chion describes as ‘the screaming point [is] where speech is suddenly extinct, a black hole, and the exit of being’ (1999:79). A real scream and a represented scream may have highly variable meanings. This evokes the discourse of liveness in performance discussed by Philip Auslander and in film by Steve Wurtlzer (1992).

These case studies exhibit different ways in which reproduced voice is used in Fine Art practice. Their frames and contexts open up theoretical discussion of object voices that have undergone a representational process evolved through history. For example, the mobile technologies of contemporary life see a loosening of vocal place rather than Dolar’s discussion of the voice as fixed in space.

---

Rather than claiming that an utterance is site specific, I mean that utterance is always situated in space, time, as material and culture. This is where different theories apply: Chion (1999) and Silverman (1988) relate the voice to the image in their psychoanalytically driven theories, where Ihde (2007) and Merleau-Ponty (1996) emphasise the phenomenological, and Jonathan Sterne (2003) the cultural. It is necessary to find where these theoretical methods meet rather than stressing their oppositions.

It can be difficult to get people to use their voices, but we are also familiar with the ‘hyper-fluency’ of radio speakers who deliver a suitably sonorous sound that never stops, which is also received as fake, like a laughter track. Film, television, cinema, theatre, poetry all exert influence on reproduced voices. Therefore, can it ever be accurate to say that voices have any given, singular quality? Erving Goffman32 differentiates what is ‘given’ from what is ‘given off’ and this can be compared to the rift between the production and performance of voices. Voices are always produced, they do not exist in their own right, but reproduction distances the space and time of production with a second space and time of reception.

What is ‘given off’ is the information of which we are not conscious, but others are, such as our health, or frame of mind, and as Kittler (1999) says mediation also ‘gives off’ information. The scratch and hiss behind the voice in a recording evoke the vinyl disk turning on a turntable, like a clock governing pace and overall duration, our mind being drawn through a particular duration.33 In Chapter Four, I go on to discuss Relevance Theory, which suggests that all communication involves us gleaning the intention of information received from another, and that this is not only receiving a readily packaged linguistic message.

Jane Bennett’s (2010) writing about materiality connects to contemporary concerns of art practice and sound also expressed by Cox in calling for a material theory of sound. In this way the cusp of object and subject found in the human voice can be related to sound as extra-linguistic technique and expanded cinematic space. Bennett advises a rethinking of the divide between life and matter (2010) and adds a different split for object voices when she says:

33 This has been explored in the work of Aura Satz that references Rainer Maria Rilke’s ideas in which a gramophone needle is drawn through the fissures of the human skull to reveal a ‘primal sound.’
Objects do not go into their concepts without leaving a remainder, and thus that life will always exceed our knowledge and control. (2010:14)

Just as there is ‘always an excess that is not reducible to its expressive function’ (Homer, 2010:40).

Whereas Bennett is saying that objects should be considered as more vital, and presents ‘this as a liveliness intrinsic to the materiality of the thing formerly known as the object’(2010:Xvi), I am saying that the object qualities of human voices should be given a topology within Fine Art practice where they are not only language.

Below, I discuss the role of the voice in *Semiotics of the Kitchen* (Martha Rosler, 1975), and compare the vocal delivery with a similarly low budget, or amateurish, performance by Ian Bourn in *Lenny’s Documentary* (Bourn, 1978). Do they share any trace of the milieu of their production, as art, critique or documentary, and does a ‘period ear’ emerge? Where some black-and-white documentation of art performance attains an aesthetic in it own right, these black-and-white tapes feel more like kitchen-sink drama.


---

34 Rosler’s contemporary work widens notions of documentary.
35 Goodale suggests a ‘period ear’, developed from Michael Baxandall’s idea of ‘period eye’ (2011:7)
Can these factors be due to vocal performance? These voices have low production values, and are more like the perfunctory voice used in a telephone call. While stripped of narrative, Rosler’s utterance of alphabetical letters is full of meaning, what is cut back to a list is highly intonated, where Bourn produces a self-portrait of an imaginary character. His work is prescient of a return to narrative in art at the time. These works create different critiques of male and female gender construction at the intersection of art and media. If we return to Cavarero’s concern about the rejection of individual voice by theorists we can see that Bourn and Rosler produce works that are different to acting and intended to critique or undo media construction in a way that would not have been possible for Florence Nightingale. Bourn and Rosler emphasise rather than efface genre in their performance.

Duplicitous intentions crowd the voice as an object, medium, practice, performance and production. Cavarero says:

The voice as prior to speech or independent of speech is therefore simply an animal voice – an a-logic and a-semantic phonation. (Cavarero, 2005:34)

This aspect of voices is found in *Pieces I never did, Freeing the Voice and India Shouting Match*. We hear such voices that are prior to speech framed differently, as

---

36 Catherine Elwes has discussed the potential of video diary in feminist art and the idea of video diary is taken up in mainstream television in the 1980s and influences subsequent contemporary blogs.  
37 However, the various and ambiguous meaning of voices is evidenced in 5th century BC when Hippocrates judged the sarcasm of Democritus.
situated utterance by Fine Art practice. Cavarero is referring to the ideas of Immanuel Levinas who differentiated between *saying* and *said*, where what is said is text and saying is always by a subject. ‘The further Levinas gets from the said, the more that animality gets closer’ (2007:32). The absence of language in these works does not mean the absence of the subject, but does put it under pressure because shouting, and a raw vocal noise as something pre-cultural, is received as non-human. However, as Bennett (2010) says, this division can be rethought and is explored in Fine Art practice. Levinas describes the ‘said’ as including the structure and this overshadows the ‘saying’: ‘the logocentric tradition of metaphysics continues to insist on what is said and never asks who is saying’ (Cavarero, 2005:29). Bennett describes a thing-power ‘agency-versus-structure’ debate in the social sciences, according to which structures are described as powerful entities that work with and against human purposes. (2010:29)

The human voice is a partial object constructed by the subtlest conglomeration of movement and meaning where matter meets culture. To be conscious of these processes pushes the limit of voice and subjectivity. Whilst the focus required in listening is at odds or discounts other potential listening. As well as the audience for art practices there are also performers, forensic linguists and others who do listen to such complexities. Added to this, reproduction allows voices to be replayed and displaced. This is clear in the examples explored. In *Raw Materials* and *The Missing Voice* we do not see the source of the voice, but in the former the voices reverberate within the Tate Turbine Hall, whereas in *The Missing Voice* the voice is tied to the ear by headphones as the audience walks around city streets.

I have chosen *Raw Materials* (Nauman, 2004) and *The Missing Voice* (Case Study B) (Cardiff, 1999) as indicative examples of object voices in art practice. These works use no reproduced images on screens but demonstrably centre the voice as object. In these works the visual is supplemental to the voice, but is not withdrawn; we do not close our eyes or sit before a blank screen of a cinematic disposition such as that of *Blue* (Jarman, 1993), which represents HIV related loss of vision. The audience relates to the work without the structures of looking usual in visual art display; instead, looking is ordered by the voice/s in the gallery space, for *Raw Materials*, and in urban streets for *The Missing Voice*. Nauman uses flat speakers whose lack of physical presence mirrors the
invisibility of the voice by discreetly lining the gallery wall while absorbing the monumentality of the architecture. The members of Cardiff’s audience cannot control their reception of the piece in the way that Nauman’s audience can (in *Raw Materials*), in that the audience for the latter can move around to hear/avoid the voices. Cardiff’s voice is pre-structured by the representational codes of cinema, radio and simple verbal instruction collaged together, whereas Nauman presents voices as raw materials that are nevertheless structured by gallery space. A minimalist engagement with situation informs Nauman’s work. The dialectic between recorded voices and those of the audience who are both passing through and standing still to listen carefully vitalises the empty space of the gallery. Cardiff’s voice fixes meaning in the cacophony of the city, where there is no empty space, instead an eternal multi-lane highway of rhythms to which she adds her voice.

Cardiff and Nauman provide different walking and listening experiences. They both successfully release the voice into a space that is contrary to a cinematic disposition because there is no screen. However, Cardiff turns the city into a diegetic space normally framed by the screen, but in her work framed by the voice. In *Raw Materials* voices acquiesce to the protocols of gallery space. Although beyond the remit of expanded cinema both works are combining the voice with images that are prefigured by cinema disposition. In them we can feel a combination of psychological effects that cause our reaction to voice. Van Leeuwen (1999:129) says voices are always a mixture of different features, for example, ‘soft and low’. There are configurations barred to vocal utterance such as shouted whisper, yet such paradox is at the centre of vocal phenomena because the auditor creates meaning. Van Leeuwen is commenting on how the same sound can have different meanings; a tense voice can be ‘metallic’, ‘brassy’, ‘bright’, ‘clear’ and so on, but it is context that gives the meaning. We think of a lax, rough voice, for example that of Louis Armstrong, as exhibiting the wear and gritty reality of life, but this cannot hold as the absolute product of an indexical oral sign. Cardiff’s and Nauman’s works exhibit the friction between a multivalent object voice and its situated utterance.

Cardiff directly incorporates film genre in what we hear in *The Missing Voice*, where Nauman presents the raw materials of voices, but both can be found within Zizek’s analysis of Lacanian partial object ‘object petit a’. Where the voices in *Raw Materials* look to the materiality of the voice, the materiality of Cardiff’s voice is lost in the web
of speech and other genres within her work, such as detective fiction, film noir, ambient sound of a film track, audio book, telephone call, documentary history, and guided history talk, live or pre-recorded. In Nauman we ‘look (listen) at’ and Cardiff we ‘look (listen) through’. But in testing this out we find both phenomena in both works: the voice is extra to linguistic form as the material sonic ballast of its utterance as well as linguistic meaning. Dolar’s (2006:3) example of an army officer ordering ‘charge’ and a soldier replying ‘what a beautiful voice’, directly corresponds with the voice as language and object and the simultaneous potential of screen as object and representational window. As Theo van Leeuwen says: ‘sound never just “expresses” or “represents”, it always also, and at the same time, affects us’ (1999:128).

If _The Missing Voice_ is taken as indicative of structures (like cinema and radio), as it places the voice within an audio walk, a film noir the detective genre and so on, and _Raw Materials_ is read as indicative of voices isolated in time and space, these theoretical angles can be tested as both works, and all voices, have forms and structures. It is just that their structure is often effaced to enable communication, like the muteness of language described by Tim Ingold:

> It seems that, in listening to speech, our awareness penetrates through the sound to reach a world of verbal meaning beyond. And by the same token, that world is absolutely silent – as silent, indeed, as are the pages of a book. In short, whereas sound is of the essence of music, language is mute. (2007:6)

But what of the technological apparatus involved in these works? Both involve a simple use of technology. Cardiff gets the audience to wear headphones and make an audio walk and Nauman installs directional speakers in the space of the Turbine Hall at Tate Modern. In both cases the voice is framed as something invisible that is the simple accompaniment of our experience, but presented in a new way that is a critique of voice in cultural space.

In _Raw Materials_ Nauman [Listen to clip 7A] presents the voice as pure sound by carefully positioning, and playing, various previous works at the same time that the audience have to sonically grasp as an act of listening. The flat-screen speakers are placed like a physical list, in two facing rows: the resulting experience of listening creates an interstitial space through which the audience move.
It can be argued that while *Raw Materials* removes the recordings from an original context to emphasise vocal materiality it also creates a new array of speech genres, particularly in the way that it makes the voices abstract. These voices become raw materials as the noise of recorded chattering mingles with the sound of the public talking, as they move through the space. Such simple staging is typical of Nauman’s engagement with materials. Many of the works are repetitive or have the quality of lists that, by nature, lack narrative. Whereas in Cardiff’s audio walk the sequential logic of narrative cinema is present and rather than taking a line for a walk, the city is structured as a set of points. The audience are mobile in the expanse of the Turbine Hall, their walking tempered by listening out and catching bits of Nauman’s pre-existing works: they may recognise the sound from its previous manifestation. This feels like an aural twist on seeing artists’ works in different places. The situated utterance that Dolar alludes to is evoked as these voices attain the quality of objects.

Nauman’s voices are recordings from many different previous works featuring voices. If we place a voice somewhere, the context will bear down on its meaning: in the gallery these voices become sculptural elements. The audience may have already heard these voices from previous works, but the rawness of their materiality is contextual. These sounds are only half at home in this context and we listen without the closure provided by opera or the theatre, radio or telephone. Unlike in such contexts, the object voices of *Raw Materials* do not ‘go up in smoke’ (2006:15) in delivering meaning as Dolar describes. Instead, the intonation and emotion remains unearthed, as if we are listening to the back of the radio rather than through its speaker.

These are various voices, as Dexter records (Nauman, 2004:21):

Language is spoken by a disparate range of voices; male, female, child, Portuguese, professional actor, amateur, and the artist himself, with variations of pitch, intonation speed, accent, volume, and ranging in expression from whispering to shouting, singing and chanting.

The way Nauman isolates recorded voices and presents them as a palette of raw materials make the audience differentiate between vocal artefacts; however,

---

38 This audio walk is very different to the sound walks envisaged by sound artist Hildegard Westerkamp in which there is no but instead listening. Westerkamp’s idea involves an active listening in which we can train ourselves to listen, focus and unpick the heard environment with a phenomenological method.
A documentary is not the context. Michael Auping comments on this phenomenon in relation to one of the works, *Get out of my Mind get out of this room*: [Listen to clip 10 A]

Such verbal statements have been interpreted as ‘commands’ and a heavy-handed, authoritarian means of manipulating the viewer. This first-layer response is only half of the equation. In fact, Nauman’s voice acknowledges the viewer in a highly intimate way. He has transformed the room into a metaphor of his mind, and we are standing directly inside it. (Nauman, 2004:11)

The exasperated sound of the voice in this work can be compared to how we infer the state of mind of Marina Abramovic, David Critchley, the people shouting in George Barber’s work, and my own vocal performances for *Trialogue*. We construct information about the owner of any voice in more than one way and these works incite this process in the audience to a greater degree. As I have said, the voice is always a mixture, the self-launched into the social, and all of these works bring this to the attention of the audience. Nauman isolates the extra- and para-linguistic aspects of voices. Emotion and intonation as live vocal production remind us that language relies upon the word, the phrase, and the vernacular whose core is the dead letter. *Thank you thank you* (Nauman, 1992) (heard in *Raw Materials*) is repeated and the meaning turned on its head by intonation.

Between disconnected modes of address, the voice of actors and television presenters, we realise that the voice in art has no natural intonation.39 As Auping notes, ‘What you say is a function of how you say it’ (Bruce Nauman, 2004:15). Nauman’s voice becomes grating as the pressure of intonation is ratcheted up and the audience is torn between different reactions, from accommodating its sound to fleeing its noise. If I hear a different voice in each ear of headphones, I choose which one to listen to. I can only be in one space at a time, but I can experience multiple representations of space. I can hear multiple voices, but their relationship with space is less precise, and the listener with a different but partial alignment with vision actively performs the process of accommodation and listening to voices. In Nauman’s installation I can hear different works at different intensities and aural perspectives and foci of listening. This is an

---

39 In a reverse sense, art practices can become hard to represent in other contexts, for example film, theatre and television.
orchestrated overspill rather than the denial of a correct disposition for listening provided by the architecture of concert hall or cinema.\textsuperscript{40} The simplest human action of using the voice occurs in a tissue and network of genre, discourse and history. In my work \textit{Filling Space with Rocks} I also speak simply in a space that immediately opens up the many structures, including how the space affects the timbre of the voice.

In \textit{The Missing Voice (Case Study B)} (1999), Cardiff’s voice directs each participant along the streets of Whitechapel for an hour-long audio walk. The disposition of apparatus, in this work, is the wearing of headphones, placing Cardiff’s voice in the centre of the audience’s head. Cardiff is having a chat with audience members individually as she directs them along the street. This feels like radio, although also a site-specific monologue, and heard via headphones. Cardiff speaks directly to her audience like a radio presenter, but like the radio presenter this is an organised conceit and disposition of voice in time and space. The street is used as a backdrop for its own history and the work is structured by the sonic structures and potential vocal bandwidths that already come with city streets.

In \textit{The Missing Voice} the isolation of sound with headphones distances the audience from the street they walk along by dimming environmental sounds, which remain as background sound. Like Nauman’s \textit{Raw Materials} the work is a variable soundscape, and partial indeterminacy is in its plan. Cities are busy, and have sometimes been busier (and noisier) in historical time than now; they also change (like the voice), in the span of a day, a year, and an hour. The different genres in the work bring with them different temporal modes, from the duration of the walk to historical time. Timing is important on streets and affects the kind of voices experienced. Experiencing \textit{The Missing Voice} makes it possible to think about the grammar of voice in Fine Art practice.

We hear many voices in our lives, but why do they arise? The very familiarity of voices elides this question. Like language, voices have protocols and, although one’s voice feels like a tap to be turned on at any time, the voice is not turned on at any time, but rather involves highly nuanced timing which we take for granted. All Fine Art practice with voice exists in light of such structure and there are many vocal cues occurring in

\textsuperscript{40} A friend experienced a withdrawal of genre and context when, alone and snowed in to an isolated cottage, she went outside and heard the familiar yet unexpected, and very frightening, robotic voice uttering ‘exterminate’ – only later was a toy Dalek found when the snow melted.
The Missing Voice (Case Study B).\textsuperscript{41}

The historical precedent by artists of urban walking that occur with industrialisation lead to the \textit{derive} – the unplanned journey through the city and the modern spectatorship of the \textit{flâneur} who strolls. The wearing of headphones causes the most striking contemporary change in disposition and Cardiff’s walk is prescient of the contemporary use of mobile devices with headphones.

As Jacob Smith (2008: 246) says: ‘the relationship between sound and space is reciprocal’ and ‘voices indicate spaces and frame performances.’ Here, Cardiff’s work brings together genre, space, the walking body and voice. It is a thoughtful, emotional voice. Unlike the chanting voices held in the Turbine Hall, Cardiff’s voice escapes across the city and through the mind like an augmentation of the viewer’s inner voice. If this audio walk is successful it is because it is a critique of guided tours. Visitors commonly use audio guides when visiting exhibitions and this can be taken as a sign of our increasingly augmented reality but we must consider that it is possible to make our own audio tour with the convergence of technologies as mobile and available. It is in this spirit that we receive Cardiff’s works rather than an enforced march round the streets at the beck and call of her voice. Cardiff’s voice is unlike many guided tours, whose voices are more officious and textual than the manner in which she starts the talk: ‘I am standing in the library with you, you can hear the turning of newspaper pages.’ This is said after we hear a librarian answering a phone in the background, confirming the space through speech genre. Cardiff does not efface herself with this recording of her voice, mixed as it is with textual genre.\textsuperscript{42}

There is a friction between the different modes of address, and the voice is in the headphone, in the head, and in the library, and then in the street. Our imaginations are always ready: ‘there will always be more things in a closed, than an open box’ (Bachelard, 1992:88 (original 1958)). Cardiff opens the space like an aural palimpsest, and develops this in the more recent audio walks.\textsuperscript{43} The reproduced voice in the work places us between, or looking into and more aware of, the timescales of our daily life. The voice emphasises the present moment of utterance by implicating many timescales.

\textsuperscript{41} This title also has evokes Lacan’s label ‘object petit a’.

\textsuperscript{42} Ong notes that: ‘Corrections in oral performance tend to be counterproductive, to render the speaker unconvincing. In writing, corrections can be tremendously productive, for how can the reader know they have even been made?’ (1982:103)

\textsuperscript{43} For example, \textit{Alter Bahnhof Video Walk} (2012).
of the detective novel and film: the aural walk as a docu-drama sighted in augmented reality.

Walking precipitates thought differently to that which arises from a motionless state. However, it can be tiring to run and talk. The duration of vocal statements start with the breath, the three-second increment of the phrase in language, but the voice is produced by the same anatomy as the shotgun report of a sneeze, or the enlarged envelope of time in wailing and circular breathing. In *The Missing Voice* [Listen to clip 11A] the voice hot-wires the rhythm of walking to rhythm of thought, although the *flâneur* can easily become discombobulated. The rhythmical locking of voice and moving image by simultaneous image and soundtrack can be loosened by the intermediary rhythm of walking. Walking and talking or listening to a voice creates its own rhythm and timescale that is different to other dispositions of technological apparatus. Although the recording has a determinate duration the experience of *The Missing Voice* feels more indeterminate than, for example, watching a film.

The voices in *The Missing Voice* are fragmentary and dangle among an abundance of structures and frames of address, both first and third person. The voice is recorded on more than one generation of recording material so sometimes the voice is cleanly within our head and between our ears, whereas at other times it retains the trace of other spaces and discourses, for example, different archival media.

Since *The Missing Voice (Case study B)* was made in 1999, Agamben’s discussion of apparatus (2009:4), that we are almost always living through an apparatus, has become more marked by the level of mobile phone exchange and iPod use in the busy city streets in which the audience experience Cardiff’s piece. When listening to *The Missing Voice* through head phones in the street, several people stopped and asked questions and this suggests an increased normalisation of headphone use and therefore augmented reality. Before this, headphone wearing would have signalled insulation to a greater extent, as Cardiff’s voice says: ‘I like talking to nobody all day’. We are less alone in the crowd in an augmented reality.

Cardiff adopts a self-referential and knowing voice. The voice is soft, lilting and friendly – it is the everyday voice, the inner voice. Cardiff says that the *stories* are a
response to the location; ‘as if the site were a Rorschach test’ that she is interpreting. Rather than marking an event, the voice traverses the concerns of the subject on location in the city with affective response. Cardiff began using her own voice when she recorded it and was excited by how it was other than her, one of different voices a subject produces, that it was so intimate heard in a particular location. She has produced a series of audio walks throughout her career. Cardiff uses many frames of address. It is the introspective tone that makes the listener feel they are overhearing, and therefore not focusing on, the sonorous qualities of the voice.

As Cardiff describes her audio walk as a ‘case study’, what we experience is different to Gibson’s more forensic linguistic labelling in The Future’s Getting Old, the actual vocal processes recorded in Stidworthy’s ( ), or the performance of Abramovic in Freeing the Voice. They all have a different methodology of dealing with the voice: different to Nauman’s assault with voice, or Hildegard Westerkamp’s notion of a sound walk conducted without headphones as a purer way of listening to the environment (2007:49).

The audience of Raw Materials and The Missing Voice perform an enhanced type of listening caused by the strongly articulated dispositions of these works. In both, the artists integrate recorded object voices into situations where other voices can be heard in the environment. The audience have to decide how to experience the work and in so doing the difference between a typical auditor/spectator experience in gallery and cinema is underlined. We bring much of the cultural framing of reproduced voices with us from radio and suchlike and the boundary between reproduced and un-reproduced is confused because we are used to disembodied voices as our eyes look forward: this is what cinema accommodates in its disposition of screen and seats. Both works present a plethora of voices: Cardiff’s work is defined, a set of genres to do with film narrative, and Nauman’s voices are materials. Cardiff’s work uses the history of place while Nauman effaces this. This could be reduced to a difference of emphasis on phone and logos in these works. Such differences are explored in the following chapters leading to Trialogue that also presents a plethora of voice but with projected image.

44 http://www.cardiffmiller.com/artworks/walks/missing_voice.html
This chapter has applied the theory of an object voice to reproduced voices used by artists in Fine Art practice. The object voice is a psychoanalytical idea: an object that stands in for a subject, but a partial object on the cusp of materiality that reproduction can re-instigate (this is the subject of the next chapter). The voice also has a phenomenology, outlined by Don Ihde (2007), and a cultural context, stressed by Jonathan Sterne (2003) elaborated in Chapter Six. All of these theories become important for the voice. The critique of phenomenology by Derrida (1967) or the friction between phenomenological and cultural theories is the positive point of actual vocal experience. In this chapter I have shown the voice being coerced into reciting lists by artists as a modernist procedure, where lists mainly work in the realm of the written. Their rhetorical recitation moves through poetic performance and stresses the object voice. Fine Art practice foregrounds spaces, contexts, genres and the mode of address that an object voice generates. The voices used in the examples in this chapter show the voice facing in two directions at once and flourishing into many discourses other than the language and noise split of linguistics and other discourses. In the next chapters, I develop more specific combinations of theories for voices, starting with the consequence of reproducing voices.
Chapter Three: Voice and Reproduction

Reproducing the voice has many consequences, as I have noted in the introduction, and this chapter focuses on its effects in Fine Art practice. Reproduction is of particular consequence in terms of Dolar’s observation about voices that always remain at the place of utterance. Reproduction puts this claim in doubt, which he notes:

at least until the emergence of the good-hundred-year-old technology of sound reproduction, which blurred many lines (Dolar, 2006:59)

Dolar does not pursue this, but it is important to the Fine Art practice under discussion, a practice that mostly uses reproduced voices and within the ‘blurred lines’ Dolar mentions. So, while voices can be heard live in Fine Art performance, I concentrate on recorded voices, although this raises the issue of why, and when, it matters if a voice is a reproduction. Much of what I say applies to both live and recorded voices. Sometimes it matters less if a voice is reproduced, but contemporary theorists, like Jonathan Sterne, are much more inclined to note the process and meaning of reproducing voices than, for example, some film theorists in the twentieth century, such as Christian Metz (1999:357) and Bela Balasz. There is also the issue that there is not a clear divide between recorded voices, and those that are not recorded, and we commonly experience tele-present voices on telephones, radio, television and the internet.

Reproduction allows us to perform the reduced listening put forward by Pierre Schaeffer:

The tape recorder has the virtue of Pythagoras’s curtain: if it creates new phenomena to observe, it creates above all new conditions of observation. (2006:81)

Pythagoras spoke from behind a curtain so his students could better hear what he said, concentrating on his words alone. Schaeffer emphasises and works with the fact that a sound recording separated from its source can be experienced as an abstract sound form.

---

This is contemporaneous with the position John Cage took in using non-musical sound in his compositions. The removal of the visual is not the removal of the space, and we have to be in space to hear sound. Looking at a curtain has a similar bodily disposition to looking at an image, and conforms to the layout of space in a cinema or gallery, as well as with the orientation of conversation. Speaking voices have an optimum disposition in which they are in front of us, and this is retained if a speaker is hidden behind a screen. This stresses a phenomenological, embodied perception where visual and aural space can be found to join in voice.

In Cardiff and Bures Miller’s *Paradise Institute* (2001) a voice is experienced coming from the seat behind us, creating a Brechtian break of theatrical disposition. Cinema speakers surround us to produce a sound perspective and represent space depicted on the screen. Then the voice is not only in front of us, but becomes part of the ‘ambient auditory array’ (Noe, 2004:161). A human subject speaking from behind a screen is different to a loudspeaker behind a screen, but it may not seem like that. Loudspeakers do not have to move around a cinema to represent human voice in space. Performers rarely move around while using their voices because it is harder to speak or sing while moving. Pavarotti’s athleticism was limited to the bodily production of voice rather than any theatrical display involving him moving around. Cardiff and Bures Miller have placed loudspeakers at the exact point in space of the mouths of a choir in *The Forty Part Motet* (2001), but what comparison can be drawn between the movement of a loudspeaker’s diaphragm and a mouth? Bruce Nauman uses loudspeakers that look like sheets of paper. We do not see the insides of our body that are moving in the production of voice. One can imagine a loudspeaker on a boom that could swivel around like the movement of a head in space. The implicit confusion around locating voices in space extends from our body into mechanical representation, and emphasises that there is always a physical process taking place that although invisible is still structured by space.

Where some theorists suggest the voice is unchanged as a recording, contemporary theorists stress reproduction as a process. Kaja Silverman (1988:42) and James Lastra (1992:42) provide lists of theorists on either side of this question, and what follows is a sample of those who see little or no difference between an original and reproduced

---

47 Humans can speak because they walk on two legs and pressure is take off the chest, but if we run we considerably diminish our ability to speak.

48 Goodale (2011:5) notes how industrial and technical noise was depicted as machines having mouths.
voice. Balázsz⁴⁹ says ‘there is no difference in dimension and reality between the original sound and the recorded and reproduced sound as there is between real objects and photographic images’; Metz (1999:356) says ‘auditory aspects, providing that the recording is well done, undergo no appreciable loss in relation to the corresponding sound in the real world’; and Baudry says ‘in cinema – as in all talking pictures – one doesn’t hear an image of the sounds but sounds themselves’.⁵⁰ Boillat says that some voices seem ‘devoid of any type of mediation that could be assimilated to “machinery”’ (Boillat, in Albera and Tortajada (eds), 2010:216). Cavell (1984:253) says that for television there is ‘no sensuous distinction, between the live and the repeat or replay.’

The reason for many of these claims is that the voice is a partial object, so although a copy, it still has to be performed in a physical way we don’t see, whether from within the body or on a flat-screen loudspeaker, and therefore do not question as different. What is implicit in all the quotes above is that meaning is contextual even if it is clearly coming from an embodied subject standing in front of us. It does not mean there is no subjectivity but that context is sometimes effaced to allow subjectivity, especially in the case of voices and their reproduction. Fine Art practice helps investigate the complex boundaries and meaning of reproduced voices.

Although the voice has no image, other than of the moving objects that produce sound, sound does have a perspective. An opposite view about a voice and its reproduction is taken by Alan Williams and Tom Levin, who comment on how cinema can only be a sample of sound specific to location.⁵¹ Jonathan Sterne states that it is more important to recognise the social construction of sound:

The primary ideological effect of sound recording might be the creation of the effect that there is an original independent of its representation. (Sterne, 1992:70)


Rick Altman (1992:42) stresses that sound recording ‘represents’ rather than ‘reproduces’ sound, since choices about such things as microphone placement always make the recording an interpretation of the original sound and James Lastra states that ‘There is no firm ground which is more real than its recording.’ (1992:67)

Where is the limit and boundary, and the differences, between voice and mediated voice? All voices travel though a medium with a signal to noise ratio. There is nothing absolute in effect, if not actuality, in the difference between a voice and its representation by and within media. The boundary is not precise but breaks down more. The meaning of reproduction is itself reliant on context for meaning.

Mediation turns the voice into a material and object that is no longer reverberant in a body, but is captured and inscribed. This widens the artistic potential of what can be done with the live production of a voice. But it is difficult to find the limits of mediation, which is found in all voices, because language is a system of mediation. Live, recorded, virtual and tele-present voices all have a different ontology that impinges on what the artist can do with the voice. They can meld into each other, for example. Steve Wurtitzer (1992:89) writes about the enhancement of live performances in which performers enhance the ‘live’ voice with technology, thereby placing it in between these categories.

<table>
<thead>
<tr>
<th>Spatial Co-presence</th>
<th>Spatial Absence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temporal</td>
<td></td>
</tr>
<tr>
<td>Simultaneity</td>
<td>Live 1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Temporal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anteriority</td>
</tr>
</tbody>
</table>
Nauman and Cardiff explore this by combining recorded performances with live voices. Such practices critique the authenticity of originals by creating a relationship between something historical and something happening now.

The voice is intimately involved in the mediation of communication. We attribute meaning to voices when timbre falls into white noise, at the edge of earshot, or submerged in a sharply crackling static. Sperber and Wilson (1998:29) differentiate between coded and inferential meaning by expanding the arguments of philosophers Paul Grice and Peter Strawson, where the intention to communicate is inferred by a listener with codes of language:

Grice and Strawson use the term ‘utterance’ to refer not just to linguistic utterances, or even to coded utterances, but to any modification of the physical environment designed by a communicator to be perceived by an audience and used as evidence of the communicator’s intentions.

(Sperber and Wilson, 1995: 29)

That is, we can hear a voice more like an environmental sound such as a bird song or a siren and understand it as ‘a context if not a text’ (Nancy, 2007:6). Because of its complex ontology, the always mysterious question of its origin, and the effects of mediation, the voice can be more or less absent, or present to our attention, as in the difference between a phone call to a friend and the sound of Florence Nightingale from the sound archive. Both voices are mediated and quite different from any ‘live’ experience: the telephone voice is reduced to formant frequencies yet this does not become a barrier to communication; the crackly old recording of Florence Nightingale invades our experience of her voice, reminding us that one hundred years have passed. Even though one is a crackly old recording and the other a voice with compressed formant frequencies it is still context more than the actual phonic substance and vocal timbre that dominate meaning. An audience performs a close listening of the voice in certain situations such as singing where we give our attention over to the sound.

There are times when we are not aware of the voice as medium or the effects of its mediation – as Serres (2008) and Connor (2000) point out, we would be in danger of undermining our subjectivity if we had to listen to our body all the time in such a way.

---

52 To call the vocalisations of birds ‘song’ could be considered an anthropomorphic construction.
We are not aware of the mediation of the voice most of the time. In representing voices in cinema and other media, much effort is made to reduce the signal to noise ratio.\textsuperscript{53}

If the divide between the recorded and unrecorded is small or moveable, a more tangible difference is that the recording technology has a structuring principle that can be removed more suddenly than the way we can cease the sound of our voice. The sound produced by pulling a needle from a record crashes the sound into violent noise. Turning a radio off ends the sound of a voice that is created by human thought, and ended with a blunt indifference to Husserl’s ideas of protention and retention.\textsuperscript{54} Such sudden ending emphasises the absence of a pair of lungs. Therefore, technology has a tangible trace: switching the radio off, pulling the needle off the record or instantaneously cutting from one sound to another is not a sound phenomena which our body can produce. To suddenly cut sound is to create a particular ‘attack’ to the silence that follows. We have no ear lids and are subjects of sonorous immanence. Travelling through a soundscape like a noisy city can produce these effects of acousmatic sounds in a subject. But recreating a soundscape with microphones is different to experiencing a soundscape with ears. We privilege the sound of voices but they do not always become figural and are also the background and hubbub of texture.

Such types of sound also acquire meanings, as scratching means aura, the wear of age means timbre, as patina or as the deliberate ‘scratching’ of DJs. Accidental glissando (gradated change between sound) or diminuendo (gradual fading out) are natural sonic processes. Mediation effects communication in particular ways, as the meaning of timbre becomes highly subjective and contextual. To speak a word is to join thousands of years of culture but also to topple into the undetermined future. However, speaking does have a clearly demarcated beginning. Avant-garde art from surrealism to Cage provided different approaches to the abrupt soundscapes of cities and technology. An echo suggests that the voice has already happened and the echo of something is not the peak of its affective attack but its demise and farewell. The echo is the prolongation of the voice as residue of signification. Its situation is undermined by repetition, its linguistic meaning neutralised by the phatic. An echo is just the consequence of speaking in a large space, but can mean the ghostly and disembodied quality of the voice. Reverberant qualities of voices are commonly changed with studio technology.

\textsuperscript{53} The physical presence of a television reporter is more authentic if accompanied and interrupted by the sounds of events they are covering. In film, ‘wild recordings’ are made that can be inserted later. These are in contrast to the studio-bound news anchor or close-miked actor.
Voices are always reverberant and only exist in space. The technique of close-miking removes reverberation to make the voice a textual, unnoticed medium of language, yet the cleanly recorded voice performed in a voice booth cannot be detached from space. Space and time affect sound waves of voices and their cultural meaning. Dub reggae uses echo up front, and not as a leftover signifying ghosts.\(^5\) Aspects of sound recording pre-exist in echo and have a similarly problematic ontology. As discussed in a following chapter, Sperber and Wilson describe linguistic communication as the opening of a channel and the phatic sounding, which means that the speaker is making a noise, which is intended to be meaningful for others.

We often think of reproduced voices as the same as live voices and this is compounded by telephone voices that are both live and reproduced. James Lastra (76:1992) distinguishes between two models of sound representation: the ‘fidelity’ model that concentrates on the sound and the ‘telephone’ model that concentrates on hearing the voice. The tension between these models is found in Raw Materials and The Missing Voice.

The voice is well embedded in culture before reproductive technology. Albera and Tortajada (2010:28) pose the question: ‘what did one call “recreated movement” in the nineteenth century before the appearance of the kinetoscope and the cinematograph and afterwards?’ The same can be asked in considering the ontology of reproduced human voice that was always separated from its owner’s body.

It is difficult to imagine a world without voice and this applies to its spectra of vibrations. Vibration is where sound and matter meet and without it there could be no articulation of things. It is argued by Jonathan Sterne (2003:11) that it requires a human sensibility to perceive sound that might other wise be thought of as vibration.\(^6\) It is human culture that allows us to hear voices as something we bring into focus. Sterne (2003:220) also argues that the representation we hear in reproduction is not the opposite of an ‘original’ but very much part of a process in which the original is not pure, but very much part of the representation. He quotes from The Work of Art in the age of Mechanical Reproduction, arguing that Walter Benjamin (1992) invents the

\(^5\) In Boomerang (1974), Richard Serra exerts pressure between subject and object voice by introducing a delay between what is said and heard.
\(^6\) John Wynne’s Installation no. 2 for high and low frequencies (2012) – uses non-vocal frequencies on the fringe of the audible where we cannot perceive voice (there are no voices in the work).
concept of the ‘aura’ of an artwork because it is reproduced. Sterne has noted that Benjamin’s description of *aura* – that which is lost, or ‘withers’ in reproduction – only becomes available to attention with the invention of reproduction. The *aura* of reproduced voice is always produced and emphasised by its disposition in Fine Art practice as described in this thesis.

Sterne (2003:220) adds to Altman’s distinction between reproduction and representation for the idea of *aura*:

> the very construct of aura is, by and large, retroactive, something that is an artefact of reproducibility, rather than a side effect or an inherent quality of self presence.

We think of the voice as an authentic link to the self, but investigation of timbre shows a process of constructed meaning, in which *aura* relates to a new situation and disposition in the artwork. It is easy to reproduce voices because we are evolved to listen to them and language is a reproducible structure. The voice also maintains structures outside of language for example in stripped back telephonic voices. I may alter my voice when I speak on the phone, but the phone also filters my voice concentrating on formant frequencies that can be seen bunched in dark batches on a spectrograph image.

Where Sterne argues that representation is a process and Kim-Cohen that an attempt to let sounds ‘be themselves’ always fails, the object voice still emerges as the object of hearing. It is also what is before us as meaningful sound. The voice is a unique bodily index but, as Ree (1999:1) points out, ‘voices are also destined for other people: you speak, primarily, in order to be heard.’

As I mention above, the object voice – residue of signification – is used by Nauman like a hammer and tongs in *Torture*, in which a clown repeats the words ‘yes’ and ‘no’. These two seemingly benign words are raised to a level of a threat. Nauman mixes live repetition and reproduced repetition in his works. Nauman is known for isolating intonation in his presentation of written words as sculpture, but recorded spoken words bring other registers of meaning. Bakhtin has commented that: ‘Words belong to nobody, and in themselves they ‘uate nothing’ (1986:85). Repeating something is different to replaying it and Nauman’s works explore this difference. Nauman, in his
use of live and reproduced repetition, interrogates Dolar’s idea that the voice only
happens once, at a certain place and time, as a specific utterance. This is encompassed
in the paradox of voices emphasised by reproduction: while we can repeat things and do
not have to stick with what we say, unlike writing, voices only occur ‘live’, and this is
fixed as a recording. Bakhtin also says:

Objectively neutral styles presuppose something like an identity of the addressee
and the speaker, a unity of their viewpoints, but this identity and unity are
purchased at the price of almost complete forfeiture of expression. (1986:98)

Such a payoff is found in the isolated voices of Nauman’s work, for example Days
(2009), an installation comprising two rows of head height 30cm square paper thin
speakers held, on floor to ceiling wires, like an avenue through which the audience pass,
as many different voices recite days of the week.

The experience of the work evokes many readings. One can start with the extra-linguist
qualities of the voices: the sibilance of different ‘s’ sounds emphasised by recording;
the different paces at which they are listed off by their speakers, where a child lists the
days more emphatically, and adult recitation seems more related to actual days. There
are only seven days but far more speaking subjects, and the names of a lot of days can
be uttered in a day; and this is what happens in Nauman’s installation which evokes
many connections between language, time and form. Nauman’s voices sit between
subject and object, specific utterance sucked into the noise of the world.

A reverse example is the work of Tino Seghal\textsuperscript{57} in which voices are never recorded and
therefore never create secondary orality. Voice falls into insignificance as a medium
that repeats words in a constant process of rephrasing that recording fixes as sound, but
memory returns to the linguistic. Cavarero notes that the uniqueness of voices, all their
extra- and para-linguistic qualities are a threat to the smooth functioning of language.
Language produces recognisable forms with a parsimonious use of materials by
constructing signs but the art practice examined in this research uses the specific and
actual sound of voices. Reproduction fixes what is unique, drawing it closer to the codes
of language. As Kittler has noted:

\textsuperscript{57} Not only are Seghal’s works unrecorded by mechanical reproduction, he also withdraws from catalogue
entries and notices that alert the audience to the time and place of the work.
Both camera lens and microphones recorded indiscriminately and both thereby shifted the boundaries that distinguished noise from meaningful sounds. (1999:xxvi)

However, Fales (2002:58) has pointed out how a recording of a voice is perceived differently to unmediated sound. This is not to do with the fidelity of the recording but to the way timbre is affected by expectation and memory alongside the sound of the voice. Crowder (1993:138) describes memory of timbre as ‘a modest human ability, once it is separated from semantic connotations’. For Zizek, ‘what we experience is not the thing itself, it is always already symbolised’ (1996:112). Timbre is received subjectively as the voice without language but can such a noise be without structure or outside culture? Pre-linguistic voices are shaped by many structures, not least body and environment, and I am describing a grammar and topology for object voice created by, and relevant for, Fine Art practice. Voices invisibly order our presence in the world but are duplicitous and reliant on a context for their meaning. Reproduction fixes this impalpable trace, but like a hall of mirrors the meaning of reproduction is contextual although potentially very different to an un-reproduced voice. Reproduction effects how we hear voices. We are more able to focus on aspects of voices that are unfixed before the invention of sound reproduction and bring the object voice to our attention as part of art practice in a particular milieu.
Chapter Four: Voice and Image

‘One could not be certain [if he had spoken] for his words took form behind his beard as behind a curtain, and it might easily have been another who uttered them.’ Herman Bosch, *The Sleep Walkers* (1988:10)

I have argued that the voice is a part object whose reproduction has a grammar that is connected to the visual, but the connection must be described by different theories. The voice occurs in three-dimensional space unlike the two dimensions of images; it is also only temporal. This chapter describes the meeting of voice and image in Fine Art practice. It starts with the idea of disposition: that the object voice is experienced as meaningful by its disposition in different places and spaces that are brought into the environments and discourses of art. I then go on to discuss how these two entities, voice and image, are put together and brought into Fine Art practice by artists. The voice as nexus of subjectivity is also an object and material that is framed by Fine Art practice within different approaches to space and time.

As well as hearing voices, and listening to what they say or sound like, my method is to describe the disposition of voice in different artworks and apply this to voices that are reproduced in relation to reproduced moving images. Mediation adopts such a disposition because it corresponds with how we perceive the world – the world before our eyes and the voice hovering and sounding out of this perspective. However, art practice creates other dispositions between screens and voices.

In addition to the voice as recording I am fascinated by the meeting between voice and image because it is common to experience; we all speak about what we see, and what we see can make us speak, shriek, laugh and cry. My own practice reworks such experience with reproduced voices and images and I am categorising the different ways voices arise in Fine Art practice. I go on to discuss what kind of ‘meeting’ this is: are voice and image merely co-present like ships passing in the night, or do they hit an embodied perceptual bulls’ eye? These questions are a central paradox to voice and image that can never be separated or are always separated. The process of reproduction radically separates sound and image, but then re-fixes them in perpetuity, whereas in experience voice and image are contemporaneous in a free fall until framed by discourse. The voice has genre, mode of address, and as voice-over it controls the image.
while travelling a parallel course without physical contact, a type of laminar flow without turbulence. However, the voice is turbulence, both causing change in human minds, and caused by friction between air and the inside of the throat and mouth.

The voice and image open up to many different discourses of time and space. We are also conditioned to find meaning in images and voices that we experience, and expect a voice before it happens: we say, ‘I knew you were going to say that’ or ‘I was expecting you to say more than that.’ These ordinary expectations also structure the time and spaces of Fine Art practice; like a palimpsest, this moment contains the past but it is overwritten by the future, yet traces of the past show through. I call this the meeting of a palimpsest that is of the past, and the offing, which awaits the future, through which the voice is uttered in the present moment. Jean-Luc Nancy has commented that:

> Perhaps we never listen to anything but the non-coded, what is not yet framed in a system of signifying references, and we never hear anything but the already coded which we decode. (Nancy, 2007:36)

What we expect and get is framed by the relationship of voice to space, and can be called vocotope. But the voice is also a phenomena of time, and such relationships are created in the street by *The Missing Voice*, The Turbine Hall for *Raw Materials*, and the relationship between gallery space, screens, voices and images in *Trialogue*. Utterance is performed in a situation that usually involves images and it is the potential of this combination that is the focus of my practice.

Reproduced voice and image have met with human thought, conversation and communication and evolved voice-over – a familiar trope of performance that voices perform well; we know the voice is textual and scripted but it is still a voice and not a text. As Chion (1999:24) notes, it has: ‘one foot in the image’. The category of voice-over can be stretched from the stentorian voice in an advert to people passing around, and commenting on, holiday snaps. *The Missing Voice* could be called a voice-over, if it were not for the presence of other voices and speech genres in the piece. The voice-over usually comes from a unified or supposedly neutral authorial position. Voice-overs also cause particular vocal styles, and if we think again of Dolar’s idea that the voice remains uniquely at the time and place of utterance we can think how that particular
disposition also exerts an ideology and relationship with power.\textsuperscript{58} Chris Marker played the same image with three different voice-overings producing different meanings, but he did not experiment with different affective or material qualities of the voice. The same voice could be presented in different dispositions, for example, overheard, coming from a radio, rather than assuming an authorial and therefore authoritative position. The time and place of origin are effaced to synchronise language with meaning, but they cannot be banished from our experience. When voice and image are joined in a present moment they create different potential meanings. The conventional thrust of speech is a narrative meaning, but as an object and material used by artists voices are given other use.

The voice is also a sound repositioned by artists, by the placement of loudspeakers or by emphasising relationships with images. Does the voice hit the image or float over it? Emotion seems physical if we are affected by a voice, particularly with a singing voice, but all voices have physical and affective attributes. They can sound false, for example in the case of political rhetoric. The physics of the voice mingles with its cultural dimension. Pitch, volume, timbre can be joined by thinking of a viscosity between voice and image. The voice is caused by the friction of air leaving the body, but at the same time floats freely around, and within images and physical space, as witnessed in the artworks I am citing. However, the voice is always fricative, its sound wetter or drier. This raises the relative awareness and meaning of viscosity and friction produced by machines and the human body. Machines can break down like bodies for various reasons, and meaning spreads amongst real, symbolic or imaginary frictions. The space between image and sound is that between meta-physics and physics:

Perceiving sounds requires discerning coherent and significant streams of auditory information from an intertwined set of signals bound up with irrelevant noise. (O'Callaghan 2007:17)

Because voices exist in space they can only be contemporaneous with images and exist in space in excess of a two-dimensional image or representation. The voice has one foot in the image but the rest is in space. In a filmic representation this is diegetic space, but I am discussing voices that are not in diegetic space, such as a voice-over or voices that

\textsuperscript{58} This is famously demonstrated by Chris Marker in \textit{Letter from Siberia} (1957) in which a visual sequence is repeated three times with different voice-overings.
may fluctuate in relation to the image. Before describing the voice and image linked in a screen, I shall consider a general disposition that is more available in a gallery space.

Vocal disposition has a difficulty: that the voice is invisible and therefore difficult to place in space by hearing alone, yet can only exist in space. The disposition of audience and apparatus include the relative mobility of the audience, as well as the size of screen, from mobile phone to cinema screen, and this creates particular relationships for the voice with the visual. A screen on a wall is a two-dimensional area or acts as a window to somewhere else, but a voice follows a different logic. There is dialectic of intentionality between its production and performance, its transmission isolates it from other noise and its reception variously requires particular focus by an audience. As Kahn (1999:28) observes, ‘there is no visual equivalent to the utterance of the voice’. However, the rhythm of speech can be reflected in the rhythm of edited images, as a function of time; this relates to the Lacanian real, thought by Kittler to be exposed in reproduction. We might consider the rhythms of editing used by Stan Brakhage to have a materiality of an order comparable to the rhythms of noise beneath spoken language, the kind of anthropomorphism encouraged by Jane Bennett in describing the boundary between matter and life.

In Mediations (Gary Hill, 1986), a speaker cone seen as a voice is heard talking about the relationship between image and voice. Slowly handfuls of sand are dropped into the speaker until full, and the voice speaks about being ‘buried under sand’. We see the sand dancing on the speaker cone before it is engulfed. We are hearing a damaged speaker – a discordant sound, and a different type of sonic breakdown to that offered in I am sitting in a room (Alvin Lucier, 1970).

Alain Boillat (2010:219) differentiates between voices that are live with the speaker in view, produced using a megaphone, microphone or a computer programme that alters the voice, produced by a speaker we cannot see, a voice coming from an apparatus such as a loudspeaker, or a phonograph at which we look. These dispositions produce very different meanings. Dispositions in which apparatus is in view, but is emphasised or whose effects are enacted strongly, is seen in works such as Samuel Beckett’s Krapp’s Last Tape (1957), Acconci’s Face Off (1973) or Kevin Atherton’s In Two Minds

59 In There and then (1982), Gary Hill cuts images to the phonemes of speech.
(1978). In all these works the protagonist engages with a reproduced version of themselves recorded at a much earlier time, the voice as an acoustic mirror.

The idea of disposition and *dispositif* was used by Foucault to discuss power. Power is one of the five aspects that Durham Peters (2004) says surround the voice.\(^6\) Vocal power is exercised by making a noise to change something in the mind of another and ranges from the direct unmediated address sought by Artaud to the ironic elisions of rhetoric, and the switching of meaning by intonation heard in everyday verbal communication analysed by Sperber and Wilson. Some effects are on the periphery of our consciousness. For example, Daniel Stern (2004:xiv) differentiates between awareness and consciousness:

> The experienced micro-world always enters awareness but only sometimes enters consciousness (verbalizable awareness). (Stern, 2004:xiv)

Such a live, ongoing relationship is described by Ihde as a relationship of focus to periphery (2007:126). We can hear voices as noises and not comprehend words for various reasons: they are too far away, lack power, are whispered or mumbled to conceal their linguistic message, or the listener is not listening. In the last case the voice is not brought into focus as ‘verbalizable awareness’, and therefore does not enter language. A voice we are not listening to can become a disturbing noise. Some of the works I am describing deliberately skew the disposition of voice and technology. Our expectation of vocal communication is of understanding what is said whether or not the voice is mediated. By closing off other sonic possibilities we open the potential of language. We think of the voice in art as intentional, rehearsed and recorded, and therefore determined, whereas spontaneous voice can reveal things out of our control that are indeterminate. However, both determinate and indeterminate vocal sounds are used in art practice, so indeterminacy is used, rather than hidden, as both technology and voices become worn and break down.\(^6\)

The voice indicates the position of the speaker’s body before we process the language we receive; this is what Connor refers to as ‘vocalic space’ (2000:12). When hearing a

---

\(^6\) The others are communication, aesthetic expression, physical organ and love object (2004).

\(^6\) This can also be found in the variation between voices in western musical tradition and American Jazz: in the former the voice is pure and in the latter it can be rough.
voice we need to know its relationship with space, but this is less tangible and weakened by reproduction, and even more as networked cyberspace. We don’t pay attention to the owner of such a voice who is safely out of sight and not in our presence, however it still registers as very much real when we speak on the telephone or become interpolated into the narrative of a film. The voice’s noise hovers around its owner as a disturbance of the airwaves, which is resonant in consciousness or awareness, but a talking head emphasises mediation. Therefore, the simple disposition of a talking head is framed by a context and history. The talking head is a common disposition of voice and image and dictates audience expectation. Just as a court needs to investigate silence in a witness box we also would investigate our television if it contained silent talking heads.

Foucault thinks of a dispositive as typically an actual physical structure in which power is exerted and something different to discourses, that are ideational, and this extends to the difference between practice and theory. The aesthetic discourse of the voice in Fine Art practice can be assessed through disposition. We can also think of a dystopian sonic experience outlined by Goodman in which post-human voice short-circuits languages as ‘affective tone’ (2010:xiv).

The power function that Durham Peters (2004) regards as a fundamental aspect of voices is inherent in their relation to images on screens from the basic way in which humans speak about, and therefore control, what they see. Voice actualises language bringing it forth in a dialogical relation with the where, what, when, how and why of the world. The voice stands back from the world as language, like vision, at the same time as being immanent in the viewed scene. This is a primary way in which image and voice are linked. The position of the voice is arguable, constructed by context and the process of representation and cannot be precisely located as a point in space. Sterne has stressed how we use visual points in space such as mouths, ears, microphones or screens at the expense of considering the process of representation obscured by its invisibility and relative speed.

The idea of dispositive is also raised in Agamben’s essay ‘What is an Apparatus?’ (2009). Agamben examines Foucault’s use of the term disposițif, which is translated

---

62 Although the challenge for the sound recordist is to be able to deal with the contingency of sound.

63 This is explored in the video work Power (1980) by Richard Layzell.
into English as apparatus. The *dispositive* describes the physical relationship between objects, events and meanings. While the *dispositive* can be adopted as a theoretical tool to describe any situation it has particular relevance in uncovering theoretical concerns that accompany the physical installation of voices within art practices. Foucault provides a long list of what is included in dispositives: ‘Discourses, institutions, buildings.’ Agamben adds more: ‘computers, cellular telephones, and, why not language itself. The apparatus “is inscribed in power play”’ (2009:2)

To this list can be added the human voice, whose dispositives are always at stake when voice is heard in artistic practice, as situated utterance. Agamben suggests that:

> there is not even a single instant in which the life of individuals is not modelled contaminated, or controlled by some apparatus. (2009:7)

The disposition of a voice and image are wider than the conventions in which we commonly receive reproduced voice in art, cinema, mobile devices or any screen-based media. The relationship between different screens and what they show has become promiscuous, with movies viewed on mobile phones and artists ever stretching the relationship between voice and image. But often the work is manifest in the repositioning of the voice. The disposition of the voice in art practice is how this stream of noise is positioned in a space and combined with visual and other perceptions. The actual sound of the voice often lacks adjectives, and is outside of descriptive language, whether yelping or droning, unfolding its linguistic contents over time, punctuated by intakes of breath. Voices are inherently mobile; going with us, wherever we are, and then from us. The voice travels light, but how light or heavy is the material and object voice? This can be answered by considering the disposition of the voice.

The voice’s first disposition is that it has left the body for some other position. However, this is not the case for our own voice, which remains with our body without reverberation, as inner voice constructing subjectivity. However, a voice must reverberate to be heard, even if this is only through the skull. Hearing our own voice, if we speak to ourselves, breaks the coherence of subjectivity and, in a common sense, speaking to oneself signifies subjective and mental breakdown. This is one of the many reasons given for why people sometimes dislike the sound of a recording of their voice.
Voices appear to have an enormous autonomy and looseness towards the time and space without which they could not exist, because their ontology entails displacement. The voice exists in many places: multiple within the room in which it reverberates, echoing around the football pitch on which it is shouted, but then reframed by mediation. Dolar’s assertion of the voice as a situated utterance can be joined to Connor’s concept of ‘vocalic space’. We use our visual faculties, or our imagination to ‘place’ voices we hear. But the vocal environment can be very busy, with many registers of voices melding together, for example, the vocal cacophony of television, or televised sport with commentary over a cheering crowd. Artistic practice uses voice’s fundamental internal rhythm, fighting or easing into the construction of a soundscape from which it becomes foregrounded as the focus of the listener. Often the sound the voice makes effaces the time and space of its production, for example as a voice-over, but at others is clearly present as a **vocotope**, for example as football crowd:

Any medium of sound reproduction is an apparatus, a network – a whole set of relations, practices, people, and technologies. (Sterne, 2003:225)

Artists’ practice make these networks more explicit, but the reproduced voice that is moving around a network can only be heard at certain points. Signal processing falls under our perceptual bandwidth. A voice itself is the utilisation of a narrow bandwidth to maximum affects. Technology works with what has been evolved by body in an environment. Disposition is where the sound surfaces for our perception, the placement and actual experience of what is heard. We focus on the sound of a voice and think of reverberation as noise in the system, to be effaced, and outside of our focus and hampering communication. The noise of the voice and the noise of the environment are at one but separate when the voice moves into a figure and ground relationship.

Imogen Stidworthy’s work engages with the voice as an object in space. She seeks out unusual vocal performance: ventriloquism, the vocal retraining of stroke victims, back-slang, and the vocal ticks of a blind translator. These vocal phenomena make the voice stand out in their disposition, making the audience listen to the sound of voices where there is a disconnection between language, rather than the normally tight corollary by which subjects are ordered or interpolated into the world. In this way Fine Art practice uses voices to question power structures.
In (Stidworthy, 2011), the audience is confronted with an array of disembodied vocal sounds in a number of new dispositions. The artwork performs a critique that unravels how cultural construction relates to form and sonic materiality. The installation presents recorded voices that are heard as objects of hearing and unwanted noise, the ‘remainders’ described by Dolar and Cavarero, but whose structures we can try to reconstitute.

In describing the relationship between voice and image it is necessary to begin with the contrary idea that voice and image have no relationship. For Jean-Luc Nancy they have no reciprocity. We realise this when seeing a radio presenter for the first time: a voice and a face never go together because, as Nancy points out, I can never see what I am hearing. These are mutually exclusive perceptual modes that do not equate, and remain separate for empirical philosophy, which separates and isolates voice to find its true nature. Voice and image become disconnected, whereas phenomenology links voice and image within the holistic fulcrum of human perception that I am arguing explains voice in Fine Art practice.

The gap between the index of a reproduced sound and the index of a reproduced image is explored in Luke Fowler’s *A Grammar of Listening parts 1–3* (2009). In this work, what we see is accompanied by field recordings, which produce sounds that seem dislocated from their source. The emphasis of the sound recording is so precise that sonic index creates friction with the iconicity of the image – like a voice never really fitting with the image of its owner. In contrast the expectation of a voice is in a much smaller potential zone than the possible spectrum of perceivable sound. We can present a blank screen, empty space, hushed envelope of time in which you could hear a pin drop, but the perception of voice is fixed by cultural expectation that Fine Art practice critiques through practice.

As the medium of language the voice comes after the fact, reporting what is seen or interacting within place, space and location. In the actual complication of real voices a phenomenological reading brings voice and image together in human perception. We speak about what we see in front of us whether as a reproduced image or an environment augmented by our own image as mirror reflection or live video image. Such reflexive possibilities of technology are taken up in first-generation video practice, for example, *Face Off* (1973)Vito Acconci, or *This is a Television Receiver* (1976).
David Hall. The importance of genre in describing voice arises if we think of a commentary being simultaneous with images. Where language is simultaneous with images the voice is contemporaneous, having a more immanent and less fixed relationship with orders of time and space explored in art practice.

Voice and image are a fundamental combination taken for granted as the backdrop to our existence. Their combination is as fundamental as Merleau-Ponty’s discourse about Eye and Mind. If we intensely focus on a voice, our visual awareness can become unfocused, what we see may be more in our awareness than our consciousness. Accepting that consciousness may have a cline or attack, and decay or sudden cessation in an envelope of sound. Our speaking voice puts us in control of the world but is also part of the world and, as Bennett says:

> We need to cultivate a bit of anthropomorphism – the idea that human agency has some echoes in nonhuman nature – to counter the narcissism of humans in charge of the world. (2010:xvi)

As language, the voice floats over the world in voice-over but is at the same time a material sound in the world, and in the disposition of an artwork. The flux of this combination moves outward to ground the voice in art practice, where it exceeds what seem stable genres of voice-over (non-synchronous) or voices emitted from bodies in depicted or actual scenes (synchronous). Voice and image are cemented in these conventions as useful and successful representations that maximise communication. Image and voice can be said to indeed have reciprocity in which images give rise to voice, as both provocation and source. Images exist without voice but human existence and culture bring voice and images together in dialogical relationships theorised by Bakhtin and discussed in Chapter Six.

The voice only exists in the present moment but the language it supplies comes after the event. Like a clutch in an engine the voice also disengages with the visual, sliding into ‘erms’ that keep the sonic channel filled with stop-gap, dummy language, the mind racing as the vocal organ idles. In the example of simultaneous translators, if a speaker stops the halt is translated into language or linguistic explanation – ‘the speaker has stopped speaking.’ Here everything is connected by an acute presence that precedes and overarches sound and vision. The phatic channel is open.
From a phenomenological perspective the question is how we come to think of the voice as an entity separate from its environment. The voice is a tightly delimited phenomenon in comparison to the visual realm, and other sounds, and stands in as the human element where image is the rest of the world perceived visually. The voice is a tightly focused physical and ideational system whereas the visual is what arises in front of our eyes, and this can include our bodies, their reflection and reproduction. The voice locates us in space, and we corroborate this visually; this is the acousmatic procedure whereby, in film, the sound asks a question which the image answers. However, the voice is not just a marker of position as it infiltrates the structures of language and reproduction. Chion (1994:5) advocates the need to move away from the ‘soundtrack’ and the ‘image track’ in cinema as their relationship is not one of discrete and exclusive areas perceived separately but is more complicated. The layout of timeline tracks in video-editing software follows the sensory hierarchy with the image track sitting over the soundtrack, rudely separate and aloof from a phenomenological account of our embodied and perceptually enmeshed experience of moving image. This does fit with a voice floating over the solid ground under our feet, but in experience voice and image are far more enmeshed. We are locked in a present moment synchronisation. The image and soundtrack of the timeline cannot come together using a set of railway points, with images shunted into voices. Rather they are brought together by human perception, when for example, the film is viewed. Whoever is editing images and voices is lining them up in the hope and intention of a particular communication, but there is ‘many a slip between cup and lip’. This slippage is a creative opportunity for Fine Art practice in which the voice drops between numerous speech genres in habitual experience.

But as much as sound travels along the towpath of a visual canal, in which it can control the rhythm of images, it also jumps into the water with a non-linear logarithmic immanence. These transport metaphors bring out the difference between the contemporaneous and simultaneous relationship between image and voice. They also emphasise the temporal over the spatial, where image and voice are irrevocably passing by, and evoke different types of time such as the ancient Greek differentiation of chronos (the measured rhythm of clock time) and kairos (emerging psychological time), and the relationship between the past as memory, history and the future explored in Husserl’s phenomenology.
I look and express what I see with my voice through the translation of language, but I also make other responsive vocal noises arising from other senses like the pleasure from the taste of food or feeling the pain of injury. These sounds are treated as noises, in that we are not listening out for them: they are tears in the fabric of expectation. There is an uncharted territory between linguistic and extra-linguistic sound that is without firm boundaries. Our ordering and understanding of the environment we inhabit arises primarily from the relationship between looking and seeing, listening and hearing. Theoretically this process is between the real and symbolic, noumenal and phenomenal, and phenomenological method does not find a disjuncture between the senses when locating us in the world. The grammar of the voice in Fine Art practice is different for linguistic and extra-linguistic vocal sounds that occur simultaneously.

We de-emphasise the sonority of vocal sound because we use voices to position people in space. As Fales points out, we confuse something for the sound it makes and this is particularly the case for the human voice as well as other objects:

We hear a sad violin – when it is a sound we hear, and we who are sad. Our entire auditory world, it seems, is based on the substitution of the indexed for the indexical and the effect for the cause. (Fales, 2002:63)

If I am talking out of the back of my head, it suggests that what I am saying is cut adrift from reality and disconnected from the frontal disposition and visual communication afforded by speaking. But this isolation of the voice has been used by Chion to describe the voice in cinema, which reveals the speaker like a striptease act. He describes how a film will show parts of the body and pieces of speech, partial acousmetre, just as we hear people at times when we cannot directly occupy the space before them. So in both film and life we only look at synchronised speaking mouths some of the time and construct their space before we know their source.

While it may seem strange to discuss voice and image by removing the image, it has a logic because we cannot bracket out senses and meanings entirely. Dolar brings together the Acousmatics who listened to Pythagoras behind a curtain and the example of Plato’s Cave both of which he describes as ‘philosophico-theatrical strokes of genius’ (2006:196). They are examples that make us rethink what we take for granted in
perception and communication and are not unlike the dispositions of voice created in Fine Art practice. For example, in works such as *Blue* (Jarman, 1993) or the beginning section of *Zorns Lemma* (Frampton, 1970), we hear human voices from a disposition with a screen but with no (or minimal) image. In *Sea Oak* (Wardill, 2008) we hear a plethora of media experts speaking while also hearing a film projector running with no image. Such works not only test the phenomenological limits of perceptual combinations they also emphasise the idea of the voice as something acousmatic, whose source we cannot see. There is no lip-sync because the voice floats in space, but the lips and mouth are hidden rather than absent: they mould a direction for communication because the voice comes out of a mouth, the result of human discourse and a presence which is only missing visually. This is a moot point as seeing a speaker, in fact, increases our understanding with non-linguistic visual clues, rather than making it easier to listen to the voice. Therefore we cannot proscribe an optimum situation for listening to voice: different subjects listen differently as part of the process. Voices structure the space in which they are heard, and spaces affect the sound of the voice.

Voice and image become disjointed in asymmetrical relationships. Where the image is two-dimensional the voice is invisible and exists in three-dimensional space. Yet we easily experience the reproduced voice with the reproduced image. There is a problem in putting the voice in a void because a pure space feels like no space. Just as the voice is the voice of someone, so space is the space of some place. The relationship between voice and space is exaggerated where space is reduced to emptiness creating a difference between visual and vocalic space. ‘Come into my office’ conjures up a particular speech genre, and think of stepping from the airy precincts and gently echoing vocal sounds of a cavernous medieval cathedral straight into the close and stuffy environment of the confessional it contains. We know that we speak differently in different situations. Our voices change according to the time of the day or for the different people with whom we talk. This also holds for the range of voices that can be heard in art practice. The voice also has a poetics of space. It has a complicated relationship with space and place, which is normally ignored in passing on the linguistic message. ‘Nowhere’ is not a real space, a voice is heard in space, and that space, however stripped down, is always on the map, even if at the edge.

There are many dispositions of voice and image in cinema, television, Skype calling, all with different critical discourses. These spaces provide different sonic moulds with
different histories that effect how we listen to and hear the voice. The mould becomes set: the cinema audience sits motionless and is interpolated into action represented on a screen; by contrast, the viewer in the gallery space is mobile. These modes still dominate as models despite the current proliferation of screen and media, where media always contains the trace of earlier media and do not arrive unsolicited. The object voice is a reverberant sound on a stage, a screen, or as the ether of the airways, where in a gallery it might become sculptural, for example in Nauman’s work. However, voice retains the imprint of other media discourse as something extraneous, seeping in from outside. The framing of television, cinema, theatre and other media always includes the extra-diegetic – what is off screen, or stage.

This is apparent in the difference between voice-over, dramatic vocal performance and voices that have an indeterminate relationship with images. The gap between unrelated voice and image and a voice grounded in a situation is explored in Fine Art practice. Stillness exists in relation to the movement of the body that is extended to the voice and its representation. We easily accept the all seeing and all knowing commentary that accompanies television sports events. 64 Different media create the genres that allow us to understand and receive reproduced voices, film theories tend to mark the soundtrack as speech, music and noise. Voices are involved in both immanent and linear aspects of representation.

The voices of film stars are more like fetish objects in comparison with the incessant talking heads and voice-overs of television ‘flow’. 65 Representation and realism are tested by a human voice. The tension between formalist and realist approaches to film is seated in the voice as a sound that, as well as dialogically and dialectically ordering human perception of the world, is just a material noise. These oppositions are not equalised by theory but exert different relationships in specific utterances in artworks. The contemporary turn in cinema theory is discussed by Elsaesser and Hagner (2010:4):

Each type of cinema (as well as every film theory) imagines an ideal spectator, which means it postulates a certain relationship between the (body of the) spectator and the properties of the screen.

64 For example the 4:3 aspect ratio of television screens derives from framing the head.

65 Raymond Williams noted the continuous flow of television: see, for example, his Television: Technology and Cultural Form, London: Fontana (1974).
They propose a study of cinema related to the senses that sidesteps theoretical historicism. This can include the experience of object voices that are reproductions of sound received by a live audience. Where ‘Deleuze sees images on an immanent plane without a perceiving subject’ (2010:125) – something perceived as something other than an embodied phenomenological approach to spectatorship – Elsaesser and Hagner write about haptic theories of cinema.

The voice’s materiality is ubiquitous but unnoticed through most discourses, both academic and everyday, because, as sound, it is evanescent and doesn’t accumulate like other matter. However, as reproduction, the extra-linguistic factors affect image before language. This also applies to the reproduced voice – the sound of the voice leaves the grid of symbolic, like the moving image, as something other than language. The voice seeps under the skin as what Siegfried Kracauer suggests are ‘visceral faculties’ (1960:159) – we experience voice as something direct and before language.

We have ingrained expectations of moving images on screens. A cinema audience sits still in front of a screen whereas television allows greater movement and response from its audience who can change channel. Indeed, the screen itself must compete for attention: contemporary viewing may include the simultaneous use of a second computer screen, laptop or mobile device by the networked individual. Cardiff and Bures Miller update The Missing Voice (Case Study B) (1999) with Alter Banhof Video Walk (2012), a walk on Kassel Station using an iPod; the audience walk around listening to Cardiff’s voice and looking at the screen of the iPod. This is a Proustian walk around the station that joins ideas and material phenomena from different historical periods. The cinema audience may appear to passively perceive the film, but Alva Noe (2004), for example, develops the ideas of Merleau-Ponty and argues that perception is an action. The movement of the audience in Cardiff’s walks perform such actions that intertwine image and voice. We walk and talk to others using a phone; all of these dispositions require accommodation. It is difficult to run and talk. We can talk because we are on two legs and our arms do not bear the weight of our body and in evolutionary terms our larynx has dropped, facilitating speech. Listening is a fulcrum between different movements, the intangible performance of a listener synchronous with the impalpable vibration of the voice. We diminish our ability to listen if we move

---

66 Memory is superimposed on the real via Cardiff’s voice.
because we emphasise our own presence. This is what Dolar describes as the self un-reflected into the other: ‘The moment there is a surface which returns the dimension of the other, it becomes a deferred voice, and narcissism crumbles’ (2006:40). Dolar’s idea is that our own voice is projected and received at the same time, but without any kind of screen, where the otherness of the voice involves reproduction requiring a screen or is an acoustic mirror.

The disposition of voice, screen and audience is the context that forms the object voice. *Trialogue* places reproduced voices in, alongside, or with screens that show images. Within such relationships lie the cultural construction that critical art practice investigates: the default disposition of cinema, television and computer use that equates to our bodily and perceptual interface with the world. We are face to face with the world and the voice follows this logic; in fact, sometimes the voice can be thought to be more frontal in our perception than the image. *Forget Me* uses the front-facing aspect of the talking head that extends from traditions of painting to television. This is a fundamental way the voice aligns with visual perception and is found in the disposition of theatre auditorium. However, in *Filling Space with Rocks* the etymology of ‘auditorium’ involves a stepping back into the three-dimensionality of audition.

Theorists such as Kim-Cohen have focused upon a point in art history when the audience look ‘at’ rather than ‘into’ the space of painting and transfers this to the screen on which the moving image is viewed. Kate Mondloch uses:

> the term ‘screen-reliant’ as opposed to ‘screen-based’ to signal that a screen is a performative category. Almost anything – glass, architecture, three-dimensional object, and so on – can function as a screen. (2010:2)

Nicky Hamlyn has also pointed out that:

> the projector creates the screen, since what is necessary for a film projection is not so much a ‘screen’ per se as a stable interdependency between screen and projection surface. (2011:214)

The surface of a voice could be thought of as timbre, grain, or material quality but is always produced by the action of sound waves.
Mondloch (2010) also thinks of the screen as a space linked to another or virtual space. Similarly, recorded voices have a time of production different to their reception, introducing a plurality that resists modernist reduction. *Triologue* uses a gallery space as a container of screens that both emphasises its space as the surface of a room and breaks it up as windows to elsewhere.

A screen showing a moving image is, on the one hand, the window of cinematic representation and, on the other, the avant-garde, modernist object in itself, as if a metaphor for the voice split into sound and linguistic meaning; but the voice is always linked to another space driven by its compulsive dualism. We do and do not know the position of the voice in space. Whether the screen is emptied in the tradition of modernism or its presence ignored as realist representation allows us to consider both the process of meaning and communication involved as well as the voice’s relationship with the screen.

Early film practices such as *Hales Tours* used the screens literally as windows that surrounded the audience like the windows of a train. This is a type of realism broadened to a window on the whole world and is found in art installation, expanded cinema, simulators, point of view travelling shots and video games.

Cavarero has said that, ‘the visual and the verbal penetrate one another’ and cites Levinas who thought that the human face of eyes, mouth and nose was ‘neither sign which tends towards a signified, nor a mask that hides it’ (2005:27). Deleuze also defines a voice as having an absolute or relative relationship with an image which relates to lip synchronisation or voice-over. Psychoanalyst Kenneth Wright\(^\text{67}\) has noted that for an infant, the mother’s face can’t be apprehended by touch. A face on a screen shares this quality with a real face that is in some way a mask of the subject, itself already a kind of screen. Voices are special kinds of sounds, just as faces are special kinds of image, both being intimately connected to subjectivity. Placing a recording of a person’s voice with a still image of their face emphasises the two-dimensionality of the image and the spatial lack filled by sound for moving images. In our perception,

---

sound surrounds and the visual is in front of us, and because we don’t have eyes in the back of our head sound compensates for this lack in our visual field. Reproduction changes the voice and its relationship with images:

The ontological bond between a sound and its origin that appears so self-evident to us in everyday life is cancelled out and annihilated in the technological set-up of sound cinema. (Elsaesser and Hagener, 2010:129)

Voices are always structured by a disposition before the invention of reproduction: ‘The very familiarity of thinking in a language conceals its shape’ (Ihde, 2007:138).

Voice-over covers the image with invisible exegesis and the lip-synched voice is also the product of ambient deflection and reverberation coming from within a depicted space of the speaking body. This is given critique in Alvin Lucier’s I am Sitting in a Room. In Lucier’s piece the voice becomes the volume of air in the room reversing the normal effacement of vocal materiality in favour of language: here, language disappears in the sound of the voice. This is like expanding the voice’s spectrum to new potential. It is the counterweight to noise reduction such as Dolby, in a way that is similar to Dolar’s description of the voice itself as a ‘counterweight to differentiality’ (2006:36). As Ihde notes, ‘if we hit hollow brass, the instant of the tap reveals little, but the reverberation reveals a great deal’ (2007:99).

The duality of the voice is expressed again in the way, on the one hand, it happily comes with images – we spend much of life speaking about what we see – yet on the other hand, the voice has an uneasy relationship with the reproduced image and its apparatus as something dialogical, evanescent and constituted by genre, inside and out of the apparatus. A characteristic of voice is quick evaporation but without context and placement nothing can evaporate. Immanent rather than singular, as the partial object of space it is also pulled between clock time and emergent time. If the voice is always framed within a context does the context stop when the voice stops or are we still in a state of conversation and therefore discourse created by the expectation of vocalisation? A state of vocal disposition has a fragile existence as the product of a single voice but builds concrete presence as reproduction.
Philosophy has pitched the spoken word against the written word, but less is known about the voice in relation to the image because they come from different, but not necessarily oppositional, phenomenal realms. For example the image is outside language, like the voice, whereas writing and speech are functions of language. Philosophical ideas help expand the meeting of voice and image as a process ranging from a dialogical concordance to an obdurate repulsion like that of oil and water. In arguing for both verbal and visual analysis of art Michael Baxandall (1988:153) quotes Feo Belcari of Florence (1449):

‘The eye is called the first of all the gates
Through which the intellect may learn and taste.
The ear is second, with the attentive word
That arms and nourishes the mind.’

The situation allows us to understand the context in which we hear a voice, where the context cannot be removed from the voice but rather creates the voice as much as the speaking body. The coming together of the sound of the voice and its disposition is relational – arguably, sound is unruly and something a speaking subject tries to consciously contain at the same time as intentionally control. The intentionality of the artists I discuss, in using voices, becomes disjointed from verbal communication. For us to identify a sound as a voice we must place it in a context. Art practice often confounds expectation of context to create different effects such as suspense or dissociation. At other moments the voice is used as a form of realism and context is de-emphasised.

In our experience of the world sound is always immanent but what we see is more stable. Yet the visible has a variable permanence that pertains to different timescales: from the cosmological, geological, historical, the life span, the human breath that supports our voice, and the speed of sound and light that centre us in the here and now. Our bodies are objects that blur the separation of sound and image, and produce voices that are part objects. Our body is an object that is with us all of our lives with an internal environment of moving parts that moves around the wider environment. We can only perceive the world through our bodies.

Where pitch and volume can be measured and timbre described, disposition describes the placement of the voice in the artwork. We unearth the paradox that we want all
voices to have a source but at the same time we cannot define that source. Voice and image are linked if they are in each other’s presence, just as Connor (2000:7) has commented about our voice bringing us into presence with our self. Reproduced voices heard in Fine Art practice provide deferred presence for the audience to understand. If we listen to laughter in the world we might not question its disposition. Robert Provine (2000) has analysed laughter and found its relationship to language and perceptual structure as more than merely a loss of control and ‘flooding out’.

The aspects of the voice that are unconscious are often emotional. In narrative films emotion is often accompanied by a withdrawal of voice because emotion is outside language but also because it is hard to perform what is unconscious. A phenomenology of voice in Fine art discovers grammars for how and why object voices behave the way they do with images.

There must be a rapprochement between attempts to isolate the voice as subjectivity and as an abstract sound. I have described the idea of disposition to account for how Fine Art practice creates meaning with object voices as material form consciously placed in a sonic perspective. Art practice can alter the relationship between voice and image where we think of the image grounded as visual evidence over which the voice is passing. There is a difference between producing a voice and hearing a voice, and between a live and a recorded voice, between hearing our own voice and those of others, but these categories mix up in our experience. This expands into confusion between space, body and voice as disposition. Listening is effected by the visual. The first relationship between voice and image is found in how we speak about what we see. This creates a disposition in which what we see is in front of our bodies and extends to speaking to another before us. This disposition is found in how we view someone we speak with but continues through all screen media. The obvious nature of this frontal disposition structures the meeting of voice and image, but at the same time the voice is freed from this disposition as something wholly mobile, hands-free and invisible.
Chapter Five: Voice and Self

Before moving into the opposing theories of phenomenology and pragmatic theory in the next chapter, ‘Voice and Theory’, this chapter will explore how Fine Art practice critiques the reproduction, and split, between voice and self. I have argued that as an object and material the voice is something other than the self, because the self is a more complicated nexus that Derrida (1973) says is not a foundational self-presence. The voice is always the voice of a self and not the self, yet the self can only be apprehended through appurtenances like the voice. In Fine Art practice, the recorded voice becomes an object and material to be manipulated through the processes of production, transmission, reception, and the dispositions of exhibition. The voice has problems performing a separation in the order of a body and mind split. However, this is a moot point because when we hear the recording of Florence Nightingale we believe that it really is her speaking. Nightingale expresses the hope that her voice will stand in for her life and work, and since then, the authenticity and indexical quality of reproduction has undergone a history of revisions. I have experienced these contradictions in practice when using my recorded voice that, like a photographic image, is weakened as an equation with myself while remaining more than a residue. As well as being a partial object, the voice is also partial in its relationship with self. Nightingale put a recording of her voice into the world that, as a new possibility, looked to the future. We now look back to her actual sound as a historical limit. Since then the recorded voice has become an everyday occurrence providing extended use as a tool. By putting voice into the world as art, and with images, artists emphasise disposition. Although the voice can now be heard anywhere, for example when I receive a telephone call at the top of a mountain in *Filling Space with Rocks*, the meaning is created with implicit knowledge of the situation. This knowledge is undermined in the case of mobile phone calls that precipitate our urge to say where we are, and locate our selves. This indeterminate space is the metaphorical *offing* I use to describe how artists control our expectation of voices.

In *Trialogue*, and other works cited, the audience hears the voice that I suggest assumes an ambivalent position, even when intended as a direct form of address. Directness can be experienced as physical presence, affective absorption of timbre, or as a synecdoche for the self. Unlike the rhetoric of politicians, artists have used their voices to critique

---

68 We have to be told that it is Nightingale speaking. The provenance of some sound recordings in the National Sound Archive have been reviewed when challenged by the heirs, family or friends.
the construction of the self, subjectivity and the materiality of voice. I am attempting to
describe the context for the voice in the specific utterances of critical practice. In this
chapter I begin with Roland Barthes’ concept of *The Grain of the Voice*, which theorises
the voice as something other than self, but rather as a sound that contains the body.
I then discuss video practices that critique the voice and self. The voices heard in
*Trialogue* are produced as antecedents of voices used by first-generation video artists
discussed below. I then discuss contemporary verbatim experiments that critique the
representation of subjectivity by intervening in vocal representation.

Below, I discuss Phillip Auslander’s discussions about self, performance and liveness.
I am interested in the voice that remains at the time and place of utterance raised by
Dolar, and also the idea of the voice as coincidental with self raised by Connor. The
noise of the voice has been variously regarded as inseparable from self-presence, to
being a mediated echo of subjectivity. The voice also provides information outside of
language, over which we lose control, becoming what Goffman describes as ‘given off’.
These are aspects contained within the sound of the voice that seem on one end of a
methexic\(^69\) scale and more like noise, directly linking mind and body, than language.
The noise of the voice reveals information, such as state of mind or physical health,
before the intention of its owner intervenes.

Roland Barthes (1977:179) in *The Grain of the Voice* famously discussed the sound of
the voice, by comparing recordings of singers. For Barthes, language and voice meet,
‘as though a single skin lined the inner flesh of the performer and the music he sings’
(1977:181). The voice is the materiality at which a specific language, such as French,
meets sonorous material. Language is culture and this is vitalised by voices as opposed
to language as a code. Barthes prefers a voice that swings away from a linguistic over-
determination. Scott points out that ‘“Barthes’grain” is not synonymous with timbre’
(Scott, 2003:18). As I have described, timbre has a wider set of meanings. As the singer
in *The Jazz Singer* comments, we can hear the ‘life lived’ in a voice, but we can also
hear its reverberation in space as much as index of the self, when distorted in its
transmission by Alvin Lucier in *I am Sitting in a Room* (1970).

---

[^69]: Jean-Luc Nancy (2007) uses ‘methexic’ to describe listening as something always requiring the
audience as part of the processes.
Barthes’ notion of the grain of the voice is: ‘the body in the singing voice’ (Barthes, 1977:188) and, ‘the materiality of the body speaking its mother tongue’ (Barthes, 1977:182). The body is heard in the voice and is the point where a cultural depth is revealed. For Dolar, this is not enough to fully describe the voice, which always leaves the body for another space, and cannot be reduced to the body. We must also remember that Barthes is describing the singing voice. The operatic voice is an aesthetic object heard on its own, soaring away from its setting, staging and language. The singing voice is an aspect of self that short-circuits language but is also misleading, like the smile of a clown. There is an indiscrete boundary between the idea of a singing voice obviously performed, and a speaking voice that is less demonstrable as performance. The work of Bakhtin and Sperber and Wilson points to how the difference is normally explained by genre or context, where experience is structured by expectation.

Barthes’ idea, that the voice has a visceral materiality supported by the wider culture of a specific language, leapfrogs the self from body to object. In Trialogue, my voice, as represented sonic material within the different genres and meanings, is distanced from my self. The most prevalent thing a voice becomes other than sound is the subject who produced it. We commonly experience the voice as its owner, because if it comes from elsewhere it is ghostly and unexplained: the voice and subject are always separated. It is as if we need voices to be locked to bodies to hold the world together. Lip-synch is only part of how perceptions join to form the present and ameliorate a constant slippage in our perceptual processes. In the past the subject was perceived differently, and Gina Bloom quotes Bruce Smith:

> Bruce smith shows that voice was not equated with self-presence for all early modern authors. In fact, authors emphasize a range of things that “Make the voice strange” imbuing sound with the same deconstructive potential as writing. (Bloom, 2007:4)

Tim Ingold discusses the separation of sound and meaning, and how the sound is taken out of language with the onset of writing, but it was not always so. He asks:

> How did it come about that the essential musicality of song was transferred from its verbal to its non-verbal components of melody, harmony and rhythm? And conversely, how was the sound taken out of language? (2007:8)

---

70 When I interviewed the jazz singer he spoke of singers who lose their speaking voice but can still sing.
Connor (2010)\textsuperscript{71} suggests that the voice literally moves from object to subject as it leaves the body. The journey from nothing to something is from matter to meaning. What starts as noise within the body attains subjectivity as it exits the flesh. The deep subjectivity of the voice starts as animal, abject, visceral matter. However, the voice is still matter and form when it becomes meaning and this is highlighted in artwork, where an audience is engaged with a particular activity in which the extra-linguistic is less a remainder than something that positively forms the work.

The ideas of Sperber and Wilson about human communication show that an intention to communicate meaning is inferred by the hearer. In this sense we are only ever overhearing each other and the accommodation of communication belies the fact that voice is a noise, but one with plasticity suitable to the nuance of language.

Judith Butler places subjectivity between addresser and addressee:

\begin{quote}
Positioned as both addressed and addressing, taking its bearings within that crossed vector of power, the subject is not only founded by the other, requiring an address in order to be, but its power is derived from the structure of address as both linguistic vulnerability and exercise. (1997:30)
\end{quote}

Many first-generation video artists experimented with the tele-presence afforded by placing their image, and voice, in the loop of video camera and monitor. During this period, Rosalind Krauss suggested that rather than a material base, ‘video’s real medium is a psychological situation’ (2008:213) which, because of the screen’s similarity to the mirror, she found to be narcissism. Krauss has gone on to talk about a post-medium condition. Her discussion of video coincides with aspects of voices, and her description of video apparatus ‘becoming an appurtenance’ (ibid) to psychological situations is much like the voice as ‘an effect surpassing its explicable cause’ (Dolar, 2006:8). Krauss also raises the issue of the body in video installation and single monitor work:

\begin{quote}
Chiasmus: A Rough History of the Guttural. A talk by Connor given at the Song, Stage and Screen conference, University of Winchester, 5 September 2010
\end{quote}
In the case of work on tape it has been the body of the artist-practitioner. In the case of video installations it has usually been the body of the responding viewer. (2008:209)

Voice and video both involve a partial materiality. I am extending these questions to the voice in *Forget Me*, in which my own voice is repeated in a loop as self-presence and materiality. A voice is heard with a series of still images, animated to the rhythm of speech, for example, a pavement scratched on purpose or marked by accident, and a cracked window. Such indeterminate visual details are animated with distinctive rhythms until my head slides into shot. This emphatic performance also becomes a rhythmical repetition. The moving mouth falls into synchronisation with the sideways movement of the head as much as its synchronisation with speech. The video effects and transitions smoothly divide the image while the audience notice that this talking head is contradicting itself in ordering the audience to remember and forget speech and speaker. Visual images are presented at the same pace as speech.

Steven Connor (2000:4) says that: ‘the voice brings us into co-incidence with our selves’, but in *Forget Me* reproductive technology intervenes and questions the process, revealing the co-incidence of both place and rhythm. Fine Art practice focuses on voices to perceive of the contradictory nature of human subjects. There is a paradox: the voice is to do with origins, but is also a construct. It reaches far into subjectivity: a voice is a subjective flux (body, self and voice are flux) that is I, but leaves my body and goes into the objective world. For this reason the voice can aptly be described as a part object or transitional object that stands in as both object and subject in particular ways to explain our relationship to the objective and material world. This can also help with where we place the authenticity of voices as a differential ratio between the given and performed.

Philosophy has created a particular discourse around the voice’s relationship with subjectivity and self, where the post-structural subject and its voice are deconstructed and therefore cannot be considered as the originary self-presence of a subject (Derrida, 1973). This remains an important question raised by the contemporary practice of Beatrice Gibson and Clio Barnard even if contemporary theory has moved on from post-humanist debates. These works put pressure on the relationships between the disposition and transmission of voices and as they do so, the audience are made to think
about the construction, and what is invested between voice and self. There is an interrogation of technology and representation as the voice separates from the body of the speaker. Is the reproduced voice less mediated than the image? The voice is always the ‘voice of’ and not a pure source (the unattainable object of desire). As discussed above, Barthes described the friction and force between a body and linguistic culture. Dolar (2006:39) says that hearing our own voice is a production without mediation, that there is no gap between production and reception which therefore: ‘lacks a screen that would return the voice’ (Dolar, 2006:39).

First-generation video artists such as Atherton, Acconci and Abramovic explored the representation of subjectivity, with the video apparatus as a kind of mirror due to its simultaneous play back that included the recording of voice, in order to break and intervene in the pure interiority of vocal production. For example, as I described earlier, Atherton (In Two Minds), Acconci (Face Off) and Beckett (Krapp’s Last Tape) all produced works with a disposition of themselves (or an actor) speaking, with a second representation of themselves heard or seen emitted from an apparatus (tape recorder or video monitor). By contrast, a more recent group of artists explores the authenticity of voices where disposition is tested through recording as indexical record, or sonic portrait of subjects. This is evident in works such as Wearing’s Two into One (1997), Gibson’s The Future’s Getting Old Like the Rest of Us (2011) and Barnard’s The Arbor (2011). These works use forms of verbatim technique that have become popular in contemporary theatre and art, in the work of Alecky Blythe, for example, or Gillian Wearing. The voice is repositioned in different bodies to question the construction of subjectivity.

Barnard gets actors to mouth to the recording of the person they are playing speaking. Gibson gets actors to listen and copy recordings made with people in residential care homes for the elderly. The voice is therefore fully out of coincidence with time and space of its original production, yet re-sutured in a form of critical realism.

Subjectivity is always on-going, to the extent that Derrida concluded that there was no foundation on which the subject is built. This is an example of the voice’s duality as intimately connected to both language and subjectivity. Our voice is constantly changing genre as we live our lives. In everyday experience the examples are wide-ranging and must include Ihde’s phenomenological method, in which a listening subject creates a focus-to-fringe relationship (discussed further in Chapter Six, ‘Voice and
The sound of the voice can change from being the centre of subjectivity to being a semantically exhausted exchange – as in the examples that Goffman gives of vocal activity highly formulated by context and certainly distant from the speaking subject. He gives examples of ‘task coordinated activity outside conversation’ (1981:143), like buying something in a shop, or response cries such as ‘oops’ where ‘our observable plight is not something that should be taken to define us’ (1981:136). Derrida finds our construction of subjectivity relying always on a type of mediation rather than pure originary self-presence.

Auslander has developed Derrida’s theme to question how the acting techniques of Stanislavski, Brecht and Grotowski attempt to uncover a subject to be portrayed in tandem with an actor’s subjectivity:

> An examination of these theories of acting through the lens of deconstructive philosophy reveals that, in all three theorizations, the actorly self is, in fact, produced by the performance it supposedly grounds. (1997:30)

Auslander argues that an actor’s interpretation of a role is ‘determined by the difference between the actor’s emotional repertoire and the character’s’ (1997:31). This process is compounded in Wearing’s *Two into One* (1997), which uses the same verbatim technique as Barnard, where emotional repertoire of mother and sons are swapped as she speaks the words of her sons and vice versa to produce a hybrid form of portrait and documentary.

In *The Arbor* (Clio Barnard, 2010), the separation of voice and image is emphasised as the actors lip-synch the words of real interviewees. This has the counter-intuitive quality of radio ventriloquism but reveals aspects of representation. It is as if we can concentrate better on the characters’ testament if they are not there, like a form of acousmatic listening. This technique puts the voice into a fictional and documentary genre at the same time, meeting at the voice. The supposedly higher objectivity of vision is questioned by the authenticity residing in the sound of the voice as if, perhaps, we are more able to judge the truth of what is being said by partially adopting an acousmatic method of hiding the speaker’s face. This can be related to Brecht’s idea of *gestic* acting that acknowledges and incorporates the self of the actor and of the character being played.
Can we trust the voice? If we can’t trust our own voice our own subjectivity is put under threat, and yet the voice is distrusted by philosophy as a slippery, ephemeral sound – when the written word persists in black and white and vision locates us in space. Recording the voice adds different nuance; for example, only recently have witnesses in court cases been allowed to be cross-examined via video link. A witness in a court has not only to be present but must also affirm their presence by speaking an oath, giving evidence in their own words. If they do not speak, the court needs to know if they refuse to speak because they have lost their voice, are frightened to speak, or have become inarticulate for some other reason: the different meanings of silence that overlap with a phenomenological art practice.

The onus is on the witness to speak about what they may have witnessed silently, translating it into language with the voice. This tension between noise and silence is a characteristic of any voice, which is silent during parts of a conversation. Outside a witness box, silence has even more dispositions. A list of reasons would be endless, but could include: strained voice, fear, lack of opportunity to speak, have nothing to say, are speaking out of turn or at the wrong time to be heard, like a type of ‘esprit d’escalier’. One can think of *Nail Biting Performance* (Ceal Floyer, 2010), performed at a conductor’s podium as a type of silence, a small action sonically emphasising an emotionally charged empty space where we might expect to hear a voice. There is a similar emphasis of space, place and therefore disposition, in the mountaintop oratory in *Filling Space with Rocks*. Speaking on oath puts extra onus on the voice as a structuring agent that points at the visual, offering a preferred meaning or naming of what was seen. All the emphasis is put on the performance of what is said. Although prior, or other than, utterance our inner voice provides an ordering to our experience. If thought is expressed in inner speech, then it too must lag from experience:

> There is always a delay and retroactivity in realization and insight, so there is the inner speech which tries to remedy whatever failed, to correct and to fix one’s own incapacity to be fully alert, present, adequate and equal to the occasion. (Dolar, 2012: 135)

---

72 This French expression refers to thoughts, repressed by the pressure to retort in company, that emerge afterwards on the stairs. It is an example of the importance of place.
Although the inner voice is outside my field of research (as something latently without sonic materiality, or unvoiced) it silently seeps into the context of voices in the sense that:

Words are fools
Who follow blindly, once they get a lead.
But thoughts are kingfishers that haunt the pools
Of quiet (Siegfried Sassoon)\textsuperscript{73}

Dolar describes inner speech as a tape playing in the mind: ‘Things are recorded on this inner tape without one quite being able to steer or filter them’ (2012:135).

The self is already immanent before being put through reproductive apparatus by first-generation video artists. Where these works heavily underline the consequence of disposition for the meaning of voice, the authenticity of voices is investigated in The Arbor by Clio Barnard and The Future’s Getting Old Like The Rest Of Us by Beatrice Gibson with a much purer disposition. These works involve the pure interiority of voice described by Dolar. The Arbor [Listen to clip 10V] tells the compelling and tragic story of the playwright Andrea Dunbar who died aged 28 and the subsequent death of her two-year-old granddaughter two decades later by accidentally taking methadone. Barnard intends to highlight the ‘deprivation, marginalisation and neglect’\textsuperscript{74} that Dunbar writes about and could not move away from.

We are told, with a caption at the beginning, that actors are lip-synching to interviewee’s voices. There is a rubbing together of then and now, fiction and documentary, the same streets the same people in different historical milieu. The now of the film is a reflection by Dunbar’s daughters and others on what has happened. When a subject in a documentary is named with a caption, that usually exhausts our interest – we know who they are. But here it opens up the question of which this is, actor or subject, or voice-body hybrid. The voice also attains a slight structuring quality of voice-over.

A voice-over covers or applies to the whole image, which it can’t do if it comes from the mouth of someone depicted. As Steven Connor points out, we have to ascribe a

\textsuperscript{73} Quoted by Steven Pinker in The Stuff of Thought (2007:151).
\textsuperscript{74} DVD essay notes – fusionmediasales 2011.
body to a voice and here we have tension between body and voice. Also, we are made aware of how lip-synch is not such a percussive point in time and space. *The Arbor* critiques its own claims for, and the limits of, realism, with this subtle technique of exciting the foundation of self and voice.

In comparison, the emotional temperature of the voices in *The Future’s Getting Old* is very different and the timbre of older voices feel like the wider, downstream parts of a river. It raises the question of what is the substantive unit for a voice – there being no sonic freeze frame: this is Bakhtin’s question in defining speech genres.

*The Future’s Getting Old Like the Rest of Us* [Listen to clip 11V] begins with eight disembodied voices heard at the same time against an empty page. The work is adapted from a novel by B. S. Johnson – *House Mother Normal* – but both film and novel are atypical of adaptation or translation between prose and narrative film. The prologue in which we hear voices and see an empty page is a zero degree place for voice – theorists have provided different descriptions of this idea. To Steven Connor’s vocalic space to which we assign voices heard, we may add Adam Phillips’ (1999:53) coinage of an oral stage, the material ballast of sound into which a subject launches their own voice. Gibson conflates visual and oral space, screen or page, where we might imagine a spectrogram of these voices appearing on an empty page, and measured before us in time and space with displaced meaning The sound of these voices are not talking to each other but rather talking to us at the same time, with an undefined mixture of dialogue and interior monologue, as described by Dolar (2012:134):

> words said by other people that one can’t get out of one’s mind, one’s own words previously uttered, words accidentally uttered, words accidentally overheard, words read.

In the novel, eight characters living in the same care home describe their situation – each with a chapter starting and ending at the same point in time, on the same day.

Gibson and Barnard throw attention on the space and place of voices: the room tone as empty space, the blankness and voice:
[sound] affects the ambient acoustic array, that is, the way a medium-filled environment is structured by this sort of disturbance to the medium. (Noe, 2007: 161)

A film allows us to hear more than one voice where we can only read one text. It is difficult for a writer to obliterate a text with another text but the film-maker can play voices at the same time.

A voice coming from a certain direction will have a pattern that is different from one coming from a different direction. The ability to follow one or the other voice at will, or to switch attention rapidly between them, is called “cocktail party effect” 75.

Such cocktail-party effects are acts of focussing and aural equivalents of seeing – glimpsing voices. In psychology experiments, if a different voice is played to each ear through headphones we have to decide which one to listen to. 76

Listening involves a constant refocusing. As The Future’s Getting Older progresses voice and image perform the de-acousmatic routine outlined by Michel Chion but without closure. The frame is filled with parts of bodies and close-ups of the interior of the care home, and focus is pulled across the scene. Rather than conversing with one another the voices join in a clash of vernacular expression and collective memory. They have become simultaneous and not contemporaneous. The sound is structured as a score but evokes a disobediently oral revolt to the narrative economy of parallel cutting.

Voices and images re-join in films but The Future’s Getting Older and The Arbor provide new possibilities in which we can consider the nature of the voice other than something expiring in the moment of utterance. While image and sound can be reproduced with an indexical accuracy, authenticity is always also a human construction, although we do make ethical decisions based on reproduced voices and this is tested in these films.

I believe these films critique issues around reproduced voices by using actors in ways where the para- and extra-linguistic qualities of actual speech actively critique the hidden apparatus of reproduction. In a sense the voice is an invisible apparatus; not something we can see but rather something we attribute and attach to its owner or speaker. These films open out the processes involved by showing the recorded voice as construct and limit of subjectivity. Both achieve a kind of balancing act by giving us very direct vocal performance at the same time as critiquing the voice as something constructed as opposed to something with an absolute value. They give us the chance to experience voice as both performed and produced, acting and actuality. They do this by causing friction between different layers and genres present at one time – in The Arbor layers of realism, and in The Futures Getting Old the voice as object, text and utterance. In their use of the voice they stray only slightly from a normative vocal reproduction yet create a kind of understated extended vocal technique for the actors, with an agency that subtly emphasises reproductive disposition. The friction between the different genres used in these films is emphasised by radically misplacing the voice in the process of reproduction. Its transmission from its source to its destination involves it going into the mouth of someone else: not that surprising for an actor, it’s what they do, speak other peoples words, and the film-maker, who assembles the word, but for a non-fictional documentary subject it raises a different set of ethics. In these works the voice, is unseated between subjectivity and documentary object. Perhaps it is hard to unlock let alone represent inter-subjective relationships between director and actor, artist and subject. These films do not use a full Brechtian method, where we are forced to acknowledge the ideology present in a disposition of a voice, seen, for example, in a film like Nagisa Oshima’s Death by Hanging (1968). Neither do they challenge physical disposition as Cardiff and Bures Miller do in Paradise Hotel. Instead, they allow their audience to question the authenticity of voices between affect and structure, performance and production.

Verbatim technique in films seems counter-intuitive because film already makes an indexical recording of sound and image. However, like The Grammar of Listening, (Fowler, 2009) these films test this kind of indexical voice in the portrayal of its subjects by reversing and questioning where we expect to find the authentic. The

77 In this film, an execution is about to happen when a figure steps in front of the camera and then rolls back time to explain what the accused has done.
verbatim technique is different in these films. They provide different separations of voice and image. In *The Arbor* the actors silently move their lips and out comes a recording. In *The Future’s Getting Old* the actors listen to recordings and make an exact copy. In the first, authenticity lies in all vocal traits (the timbre, rhythm, intonation), whereas copying precisely means using another timbre but obeying the recorded rhythm and timing. This throws up some issues raised by Mary Luckhurst (2010) when analysing actors who play living subjects, such as Michael Sheen playing Tony Blair. Iconicity is as important as indexicality. Luckhurst asks whether impersonation is more authentic than Stanislavski technique – Sheen is more like Blair than Blair himself.

These works delve into the process by which film is used to construct something authentic with representation. They emphasise the meeting of vocal and filmic genre. The problems of authenticity in reproduction have a pertinent fit with the problems of the subjects in these films. There is the further complication that our relationship with our own voice is different to that with any other voice. This encompasses the reasons we often don’t like a recording of our own voice; indeed, the phrase ‘liking the sound of one’s own voice’ is used pejoratively. Most of the art discussed in this research involves listening to other voices, and yet it is our own control of our voice that allows us to listen, accommodate, and respond to voices as special sounds of audition, always in some relation to a subject. This has led Durham Peters to state that:

> Every time I hear my favourite singers or speakers, I am immediately tempted to forget all the compunctions about phonocentrism that Derrida has induced in voice-scholars. (Durham Peters, 2004)

As I mention in Chapter One Connor (2000:7) suggests that we don’t like our recorded voices because a recording removes the ‘cathected familiar’. We are forced to hear a type of ‘mash-up’ by hearing our voice out of context instead of our voice perfectly fitting the emotion and ‘vitality affects’\(^\text{78}\) of the present moment. Extra-linguistic vocal sounds have a meaning and grammar in the artworks in which they are heard. To place one’s concentration on the sound of voice can stop us hearing what it has to say, like the drone of a speaker with whom we are bored. Connor cites a complex analysis of how we are aware of our own voice:

\(^{78}\) Daniel Stern(2004:36): ‘Vitality affects emerge as the moment unfolds. This is captured in terms such as accelerating, fading, exploding, unstable, tentative, forceful, and so on.’
We must listen to it closely and continuously, in order not only to keep it in tune with what we want to say, but also to detect what it might betray of what we do not mean to say – or mean not to say. What is more, we conceal the monitoring from ourselves, in that doubling of the operation, of repression defined by Freud, in which we both censor and censor the awareness of the act of censorship. In speaking, we listen intently to our own speaking voice, in complicated feedback loop, or duet of utterance and response; we eavesdrop on our own speech, but do not, as it were, hear ourselves listening. (2001:8)

Dolar (2006:42) states that:

The voice may be the key to the presence of the present and to an unalloyed interiority, but it conceals in its bosom that inaudible object voice which disrupts both. So if, for Derrida, the essence of the voice lies in auto-affection and self-transparency, as opposed to the trace, the rest, the alterity, and so on, for Lacan this is where the problem starts. The deconstructive turn tends to deprive the voice of its ineradicable ambiguity by reducing it to the ground of (self-) presence, while the Lacanian account tries to disentangle from its core the object as an interior obstacle to (self-) presence.

We differentiate between our voice and that of others, and, as I discussed in Chapter One, Dolar (2006:81) describes this as a fundamental decision, as an: ‘ontological decision’ and ‘epistemological break’. The voice we produce and the voice of others are different creatures. We both speak and listen to our own voices, but only listen to other voices. Yet these distinctions are not clear-cut, our acoustic array is filled with voices: our own, those of others, overheard voices, voices coming from loudspeakers, and our environment is structured by our inner voice.

The work of early video practitioners has been used, in this chapter, to discuss the voice of the self: a re-working that strips back the televisual talking head. Although the audience hear the voice of the other, these works explore the relationship of voice and self through a use of technology with a phenomenological method, and this speaks to the use of our own voice. The Arbor and The Future’s Getting Old, on the other hand, critique the representation of the voices of other selves. The techniques in Barnard’s and
Gibson’s films emphasise both the emotive and pragmatic qualities of voices as objects and affects that obey grammars that can be traced through the work of Bakhtin, and Sperber and Wilson. The mystery of the object voice lies in the contradiction of its invisibility and evanescence coupled with its duplicity and partiality, factors that surround human subjectivity.

A phenomenology for the voice in Fine Art practice must account for the disposition of an object: the sound of the voice always accompanied by a physical disposition that is experienced as memory and forgetting in The Future’s Getting Old and painful memory and evidence in The Arbor. First-generation video art brought the materiality of voice and technology up against self-presence as phenomenological experiment. The materiality of absent bodies effaced in voice-over have a broader relationship in artistic practice. These works operate in the mess of real sound, stretching the framing of genre and context. They mix aspects and push the edges of interview, the protocols of dialogue, monologue or soliloquy. Trialogue creates a variable space for the reception of voice, and emphasises the impossible reality of multiple sources.

The self is centred between silence and voice, the material presence of sound coinciding with self. The voice is a shorthand for self, and as such epitomises the partial object voice. Fine Art practice helps to question and unravel this construction. The range of manifestations in which we perform and pronounce our selves relate to context and disposition for art practice. The noise the voice makes is both in and out of our control, in the body and in the world, although tightly synchronised to our experience. Mis-timing the voice causes it to fall on deaf ears, from the accommodation of a conversation to the power structure between human subjects. The voice remains in a contradiction between the self in self-presence, which passes us by in flux, and a sonic object. Fine Art practice intercedes on the edges of these categories.
Chapter Six: Voice and Theory

‘Communication is not transmission.’ Jean-Luc Nancy (2007:42)

In this thesis, the voice has been described as an object that can be reproduced and exists in relation to the realm of vision. I shall now devote a chapter specifically to theory in order to develop and express different philosophical points of view on voice and to accompany practice. Theoretically (and paradoxically) the voice can never be unequivocal, yet in practice perception focuses it into an object, experienced as a material presence. We think of the authenticity of the voice being in its timbre, like a physical constant, but as I have shown, the voice exhibits behaviours other than this, behaviours in which it is relational with context. I have been arguing that the voice is under-theorised because it is partial and passes by with minimal trace. A combination of theories is needed to explain the specific utterances heard in Fine Art practice. To do this, ‘Voice and Theory’ brings together the pragmatic theories of Bakhtin, and Sperber and Wilson, and the phenomenological theories of Merleau-Ponty and Ihde. These theories account for a voice that boils down to something autonomous, but is at the same time the product or result of many processes. First, I shall preface the discussion of these theories with a wider philosophical description for voice in Fine Art.

As we have seen, the voice is a partial object that is a mixture of personification and objectification, and never purely object or subject. We are so used to the voice that Ihde thinks of the voice structuring our experience even when it is absent.\footnote{For example, in Berthold Brecht’s \textit{Mother Courage and Her Children} (1939), the character performs a silent scream.} Art offers an open space where inner voice is transformed into utterance, what occurs between palimpsest and offing is the marriage of sonic materiality and cultural space and place. The voice also happens in the frontal aspect where vision and voice are embodied, and this relates to Merleau-Ponty’s theory. Fine Art practice creates many different spaces for voices.

The cultural recognition of a palimpsest mutates into the actual experience of what occurs in the offing. We can create auditory measures for the voice such as ‘ear shot’ which, before reproduction, related directly to the political meeting, for example, the orator in the amphitheatre. Since voices have been reproducible they have acquired
more range and their material presence is experienced in different ways. It feels as if the voice does have a constant – I recognise someone’s voice – but more than a physical constant expressible like the speed of light, or Planck’s constant. While the voice is universal for most humans, its qualities, such as timbre, are, as Cornelia Fales says, beyond what we can describe as an absolute value. Roland Barthes also wanted to express such a relationship between culture and body in his idea of grain, as discussed in the previous chapter, ‘Voice and Self’. Mostly the voice is, of itself, a constant in our lives, which paradoxically constantly changes over time: lifetimes, days or utterances.  

Pierre Schaeffer (2006) and John Cage (1973) posit important theories for voice in modernism that ask for a reduced listening to abstracted sound, and that sound is, of itself, explored as art. This means a change in listening, whereby material and objective qualities are less effaced. Voices make us aware of temporality where we might otherwise experience stillness and its sonic corollary, quietness. It is important to remember that a sonorous object is methexic, something completed by the presence of the listener, and this is expressed in the rhetorical idea of the enthyme where the hearer completes the semi-stated message, and this is expanded, as I will discuss, by Relevance Theory.

Derrida’s (1973) critique of Husserl questions whether the voice is an originary self-presence more than writing. This query over the voice underpins differences between the reductions of phenomenology and a pragmatic philosophy that opens the voice up to its context and actual use. The voice that moves between object and subject in psychoanalytical theory is at times in contradiction to a phenomenology of voice, which focuses down rather than opening out.

In The Phaedrus, Plato (1892) puts spoken language in opposition to written words. He was weary of written text losing the vital meaning imbued by speech, the spirit that is lost in the dead letter of writing. This is complicated by the implications of reproduction. Reproduced voices are both received live in the present moment, but are also dead in the medium of their representation. My practice concerns the relationship of voice with image, and this includes all the extra-linguistic factors that make up actual voices, always reactive to place yet made more like the dead letter of writing by

80 Yet its variables are small, particularly when it comes to language, hence Chinese-speakers can have difficulty hearing and making the sounds of English, and vice versa.

This is a theoretical area of dispute: does reproduction make the voice more like an object whose material qualities persist as a repeatable artefact, like the reproducibility of writing? Must we admit, as Dolar suggests, that the voice remains in a time and place of utterance? As Ong says: ‘there is no way to erase a spoken work’ (1982:103), unlike the words of a palimpsest. Voices are always situated utterances because we cannot conceive of a voice without time, space and meaning, but are resituated at the same time whether as coming from you to me, via reproductive technology or as echo and reverberation. We miss voices because we cannot hold onto them; they are rushing past, but art creates a context where the reproduced voice is a type of object. A phenomenology of voice must account for this partiality of the object voice.

Derrida (1973) analysed the phenomenology of Husserl, which is an action parallel to the move from modernity to post-modernity in art theory and practice. Seth Kim-Cohen and Salome Voeglin explore this in writing about sound art. Where Voeglin describes a ‘screaming, screeching, yelling and croaking, shouting the corporeality out of our bodies to expand into the formless shape of sound’ (2010:73), Kim-Cohen cites Krauss as locating:

The modern post-modern rupture in the difference between a concern with a medium’s material versus a concern with its terms between … matter and discourse (2010:82)

Moving-image art practice provides the arena to create, record and observe such affects. The materiality of voice echoes in questions about the materiality of video to which I add Forget Me. In Trialogue, the object voice is subject to aesthetic and theoretical discourses that also emphasise the construction of vocal genre and its relationship with timbre. The voice is not killed off in art practice in the way that Cavarero and Dolar describe its plight in philosophy and phonology.

The voice is a noise that is constantly falling in and out of meaning in real time and its transmission is modified by space, its reception by a listener, and the dance it is led into by the context. Cavarero (2005:13) says the voice ‘is not a leftover but an originary excess’. The voice is distinct from memory and language as something experienced in
the present moment. It is necessary to investigate the actual qualities and functions of
the voice (whispered, shouted, echoed, wet or dry reverberation and so forth), which are
always situated utterances. The affects of the voice connect to its psychological
inclination but it is also a rhythm and prosody of the material world. Where voice is the
most meaningful and controlled human action as language, the voice is also a raft of
sonic motions absorbing the unwanted effects and frictions of the extra-linguistic.

A theoretical approach for voice in Fine Art practice is informed by Chion’s theory of
deacousmatisation in film (where a sound re-joins its initially unseen source), and
Bakhtin’s notion of heteroglossia (many voices) in literature. It can be argued that the
subject is constituted by language but such a constitution also involves the asymmetric
relationship between vision and voice. Yet in our perceptual immersion in the word and
in moving image work, voice and image work together in heterogeneous ways. Unlike
philosophical writing, artworks actualise voices that we can hear, whether emerging
from a throat of flesh or a loudspeaker, where the sound is material object.
Theory overlaps between the cultural and scientific. Where Dolar finds the linguistic
science of phonology falling short in an account of the voice, Roman Jakobson argued
that it is not the job of linguistics to tackle the ontology of voices. The physical
evidence of the voice is different to its cultural meaning.

The dividing line between voice and noise and nature and culture is often elusive and
uncertain but at the heart of practice. Connor\textsuperscript{81} reiterates this theme in showing how we
associate the voice with the animate, but noise with the inanimate. However, there is a
scale between these states that is not absolute. Connor also suggests that the voice is
invested in phono-semanticism where, for instance the ‘v’ in viscous, venomous, vile,
vindicative and vitriolic is snarling; or ‘nt’ has a feeling of shortness in blunt, dent, grunt,
pant and runt. The point here is to realise that the signifier may not always be arbitrary
in relation to the voice, something of which Saussure was also aware.\textsuperscript{82}

The voice performs more directly than just delivering language, and this is why Dolar
stresses the Saussurean linguistic idea of the sounds of speech being constructed of

\textsuperscript{81} Chiasmus: A Rough History of the Guttural. A talk by Connor given at the Song, Stage and Screen
conference, University of Winchester, 5 September 2010.

\textsuperscript{82} This is explored in Starobinski, Jean (1979), Words upon Words: the Anagrams of Ferdinand de
Saussure.
phonemes\textsuperscript{83}, that is purely relational in nature, as not a full account for all vocal phenomena. He claims that phonology goes on to account for prosody, intonation and accent as structural binaries and in so doing ‘does away with its living presence, with its flesh and blood’ (2006:19), an assessment he shares with Cavarero (2005). Timbre does not fall so easily into binary pairing.

By nature voices are not singular, but simultaneously join material and discourse, the quotidian and unexplained. Visual art practice is a position from which to examine the extra-linguistic facets of voice that Dolar (2006) argues are ignored by linguistics and Cavarero (2005) argues are generalised by philosophy. Cavarero locates a gap whereby the voice is more often heard than theorised, or theorised as if it has not been uttered. As Dolar points out, phonology and phonetics deal with the production of sound but only to extract linguistic meaning, but art practice puts the voice in much wider contexts. Factors that artists use in their work include its mediation by technology, power as metaphor, object, material and extra-linguist realm of meaning. When we focus on a voice, hearing it loud and clear, we are bracketing out other potentials of sound and meaning, as the process of receiving voice is normalised as second nature. Such inter-determinacy is at the core of Relevance Theory, phenomenology of voice, and speech genres. These theories provide different ways that are needed to account for the process by which we receive voice.

Argument always flows between two points with the voice, but the interpretation of texts provides an interesting reference. Different schools of interpretation such as romanticism or formalism struggle to reconcile two impossible positions. If we just read what the text says, then we may be misled by lack of context, but if we apply contemporary context to establishing the common sense of a text then we are not really reading it as it was written. In the case of a religious text, nobody can satisfy both camps, which leads to the vicissitudes of exegesis. In other words, determining what the Bible really says is as big a conundrum as determining what constitutes the voice. How can a phenomenological method be used to describe the voice in Fine Art practice? I believe that a phenomenology of voice in Fine Art practice requires the acknowledgement of disposition as part of its operation. Phenomenology has a long

\textsuperscript{83} Other linguists have promoted this view, for example Dwight Bolinger (1965:231) ‘The Sign is Not Arbitrary’ in \textit{Forms of English}, Cambridge: Harvard University Press.
historical trajectory through the writings of Hegel, Husserl, and Heidegger, and in the more contemporary philosophy of Merleau-Ponty, and specifically, for the voice, the writing of Don Ihde. Ihde (2007) has provided a phenomenology of voice developed from the writing of Husserl and Heidegger. Phenomenology seeks to account for the human subject’s involvement in perception in relation to what is perceived. Rather than passively perceiving the object world the perceiver must acknowledge their own effect on perception and they must try to set aside any assumptions that crowd in on how we perceive. The voice is a case in point: our own voice is often used unselfconsciously but also in a heightened envelope of presence. Didier Anzieu\(^4\) wrote about a ‘sonorous envelope’:

> In so far as it confers shape, dimension, and pattern, the sonorous envelope is sound that is half-way to being recruited by the eye, or has anticipated its functions. (Connor, 2000:29)

When experiencing art, the perception of audiences is heightened by the expectation of aesthetic experience and critique. When we hear a voice, or speak ourselves, perception is changed as the imaginary membranes of Anzieu’s ‘sonorous envelope’ are vitalised or re-shaped.

Under the conceptual reduction of phenomenology the object voice becomes more of a thing, as something recorded, repeated, and brought to the focus of our listening within art. Phenomenology is also helpful in understanding and explaining our habitual experience of five separate senses that are at the same time un-separated as our sensorium. Therefore, phenomenology points in two directions by providing a method of focusing on one perceptual mode (the sound of the voice, for example) but also on how the senses are joined in the creation of partial objects like voices. In the artworks discussed in this thesis the voice is privileged, and focused upon, yet it cannot step out of discourse, or space and place. My practice asks what is different and specific in how the voice is concurrent with the moving image and what new relationships are created between voice and image.

The voice also has relationships with smell, taste and touch which all relate to the mouth, the tip of the tongue touching teeth. The voice emerges visually from the

---

moving lips, the mouth is also where we taste, and the voice is associated with touch in
the sense that Merleau-Ponty describes seeing as a form of touching. This visual
touching is different and interstitial with the touch of pressure, pain or temperature.
Raw affective qualities of sound correspond with the visceral orders of taste and smell
or vision, although Aurel Kolnai said there could be no auditory disgust.\textsuperscript{85} The voice is
the ‘voice of’ and not a direct experience, timbre beyond perceptual absolute. The voice
simultaneously holds the sound of the animal or material and the human order of
language in concord with the visual. The voice is a vibration transmitted differently to
its visual placement in the mouth.

The two directions represented in the writings of Ihde and Merleau-Ponty provide two
methods for investigating voices for this research: as a way of isolating the voice, and as
a way of investigating how perceptual modes are mingled and joined in an embodied
perception of the world. Merleau-Ponty’s philosophy emphasises our embodied
experience of the world, our body being part of the world through which we perceive.
Ihde employs the findings of Merleau-Ponty, Husserl and Heidegger to provide
phenomenological analysis of the voice. Ihde (2007) traces Husserl’s method of
bracketing out sound. This reductive method crosses over into the methods of
modernist, minimal and conceptual art. Idhe argues that phenomenology is something
that has to be undertaken in ways that are like an artist producing and an audience
receiving artwork. Artists can experiment with the voice and create ‘new gestalts for
listening’ (Ihde, 2007:189). For example, the object voice need not always be a figure to
a visual background in which our body normally centres us. The voice can become the
ground in the accounts of Zizek and Cage, or, as Ihde points out, in Japanese painting or
Tibetan chanting. The methexic act of listening constructs the voice by focusing
perception.

Phenomenology links the autonomous artwork and its wider field with these two
directions. The critique of Husserl by Derrida (1973) then drives the post-modern
critique of autonomous art, querying whether there is an unmediated self-presence
existing within a raw voice before language; whether Cage’s plea ‘let sounds be
themselves’ is possible. Ihde agrees with Merleau-Ponty that pure auditory experience
is not possible, particularly as we listen through our whole body.

does not present objects like other senses and only gives signs of their presence.
An object voice is perceived as autonomous but has to be connected and produced by a wider set of discourses. This then informs how the object voice relates to the moving image as connection or disconnection, brushing against, tangential to, overheard within the image. Genre is constructed from the physicality of voices. Merleau-Ponty brings the human sensorium to play in artworks. He didn’t write specifically about the voice in art practice, concentrating on vision and painting, but his methods can apply to an expanded thinking of the voice as something reaching out across experience. He treats the forms of a painting as a vital connection between body and image:

To understand these transubstantiations we must go back to the working, actual body – not the body as a chunk of space or a bundle of functions but that body which is an intertwining of vision and movement. (2004:294)

In Merleau-Ponty’s work the relationship between the body and the visual image are opened up by a phenomenological inquiry that questions the relationship between the five senses and can therefore address the voice as a material and form that covers and bridges perceptual modes.

Merleau-Ponty writes about ‘singing the world’ (1996:187): ‘when one vocalises a sound, one gives it to one’s own voice, in order to give it its own voice’ (Connor, 2000:10). Ihde says ‘voice may also be a perspective, a metaphor, by which we understand part of the world itself’ (2007:189). The voice is in relation to the environment, both as affecting the speaker, and as affected by the space when heard by a listener in ways not driven by language.

In examining the voice as something other than speech Ihde opens the voice as a potential in all things. This potential is expressed as a duet. A voice is the sound of air in friction with my larynx, palate, tongue, teeth and lips. It is culture that gives the noise of the voice meaning: we speak of bird ‘song’ when this may be something else to birds.

‘The “who speaks” is never auditorily only a where’ (Ihde, 2007:195). The who and where are strongly linked. For example, we often say where we are when answering a mobile phone and this may be connected to ordinarily being in command of the
synchronisation of our ‘who’ and ‘where’, something that the mobile phone removes. Implicitly we know where we are located as a self in space. Serres describes this: ‘Coenaesthesia says I by itself. It knows that I am inside, it knows when I am freeing myself’ (2008:19).

Ihde questions whether we can isolate sound from our global experience of the world. It is our acceptance of the five senses that prefigures the idea of a reduction to one of the senses. There are inexact relations about the turning off or bracketing out of perceptual modes and pervasive sensations: closing eyelids, wearing noise-cancelling headphones, and wearing gloves. We can only isolate a voice by focusing on it, and this always involves a ratio of ‘focus-to-fringe and the ratio of explicit-to-implicit’ (2007:44). The empiricist tradition in philosophy that attempts to isolate sensory perceptions comes up against how the act of perception is a performance in the way Merleau-Ponty suggests that vision is always linked with movement. We must move through the world to decide what we look at. However, we cannot see a voice: the voice, like ‘the wind is “seen” in its effects … what is done in passing by’ (Ihde, 2007:51). Ihde warns against scientific discourse, from phonetics to acoustics, that removes meaning from its study. Reproductive technology separates voices from sources. Phenomenological inquiry reveals consequences and contradictions of vocal reproduction. Any voice is experienced in a present moment, and in filmic realism we hear voices before seeing their source. However, Ihde says that due to technological expansion of vision with, for example telescopes: ‘what is first seen is later given voice’ (2007:6). Phenomenology is a ‘dialectic between language and experience’ (Ihde, 2007:21) and in our experience ‘vision objectifies and sound personifies’ (Ihde, 2007:21). Fine Art practice critiques the boundary and ordering of object voice and uttering subject.

As an object the voice is put in front of us in the manner of a visual object: ‘More even than my gaze, my voice establishes relations of facing forward’ (Connor, 2000:5). Being brought to attention involves locating the source of the voice. Merleau-Ponty (1996:233) wrote about the reflexivity of a subject who sees their own body and this has a correlative in the subject who hears their own voice. Both are in the world yet part of the self, distant from the world. Merleau-Ponty describes the artist as uniquely able to see the world as a form outside of meaning. Merleau-Ponty sees the body as ‘an intertwining of vision and movement’, where the body is movement that focuses what
we see. But the voice also structures what we see, and vision is prefigured in movement. Merleau-Ponty’s ideas suggest a comparison between the limit of subjective reflexivity for vision and sound. In ‘Eye and Mind’ Merleau-Ponty argues that a painting is not separate from our experience of the painting, and rather than a representation it is something that guides the viewer through a nexus of eye and mind. A voice can be thought of as something that is a nexus of body and mind that guides experience in Fine Art practice.

The success of a phenomenological method rests on whether assumptions can be suspended when hearing a voice and the palimpsest is wiped clean. The modernist experiments of Cage and Schaeffer were instructive, but in no way conclusive, and I think impossible for the voice. I argue that a phenomenology of voice for Fine Art practice must take account of disposition. The voice disperses its energy in the moment of utterance, yet has a fuller presence as affect, shape and form. The meaning of artworks is affected by place; space and the confetti of voices that rain down upon us are not identical, in part because of space and place and their different constitutions. Where Bakhtin divides speech genres into primary and secondary types, Ihde’s phenomenology of voice is built up from philosophical beginnings where:

The atom is the thing reduced to an object. Rather than a thing that shows itself within experience in all its richness, the atom is an object, which has ‘primary’ qualities to which are added as effects ‘secondary’ qualities that are caused by the primary qualities (Ihde, 2007:9)

Art practice widens the way we experience voice, and where we place it in meaning, because it has less firmly set protocols. Voice is less familiar to us in art practice as art object and material where we need to account for it in context. The relationship between context and utterance has been described by Sperber and Wilson as its relevance.

Relevance Theory is a pragmatic linguistic theory which disputes that communication is only the transmission of a message and stresses the implicit contextual processes involved that, it claims, not only underpin but are central to communication.

Such pragmatism is required to unpick the action of the voice in art practice as well as the emotive qualities of voice.

What is the status of the voice as a material amongst the granular movement and integrated flow of particles, the trickling of bodily substances and the flow of human subjects through space, where voice moves from subject to object? How do we account for the dissipated energy of voices? Lost in the wind and unheard, some aspects of messages do not easily pass through the material world, are like water leaking from mains pipes, have an inappropriate signal to noise ratio, or are missed if we are not tuned to the signal. Systems like language are evolved to work through such impediments, and usually succeed to the extent that we can successfully filter the message from all else to support language as the ground of existence.

Relevance Theory emphasises the contingency of communication; a message must be relevant, and we cannot hear what we are not listening out for. Art tests its audience’s hearing of voices, a process that is both physical and cultural. Hearing a voice raises epistemological and ontological questions by introducing a break interceding between self and other, language and body. In a car a driver talks to a passenger, yet outside, a highway of engine noise and air resistance completely masks the sound of a human voice. Hearing a voice requires an aural vantage from which to cut into the network of communication and meaning. Art practice opens the act of listening to voices with particular intentionality, one which moves in and around linguistic meaning and a wider meaning that surround artworks: the narrow meaning of definition, conceptual structures or stereotypes, and wide meanings made in the field of actual experience.

Looking to Relevance Theory provides a theoretical framework on which to build ideas about the context and communication for the human voice in art practice. The quote from Nancy (page 92), with which this chapter opens, accords with Sperber and Wilson’s ideas:

> Verbal communication is a complex form of communication. Linguistic coding and decoding is involved, but the linguistic meaning of an uttered sentence falls short of encoding what the speaker means: it merely helps the audience infer what she means. The output to decoding is correctly treated by the audience as a
piece of evidence about the communicator’s intentions. In other words, a coding-decoding process is subservient to a Gricean inferential process. (1995: 27)

Nancy stresses that communication is a process in which different parts of a process are played out rather than a message being sent and received. The supporting structures are part of the communication and, like the sound of the voice, its material quality, whose noise is effaced for the abstractions and phonemes of language to be understood as words.

Relevance Theory develops the linguistic philosophy of Paul Grice (1989), which noted the difference between the unequivocal meaning of words such as ‘and’ in logical notation, and its general use in language, in which it has various meanings. Therefore communication is a process created from both ends of a process. Although Relevance Theory is referring to linguistic communication it emphasises that communication happens with the tacit alignment of what is outside language. This approach can be used to account for the voice that communicates as an object and material as well as, or at the same time as, its linguistic content.

In doing this, Relevance Theory is a critique of what it describes as the ‘code’ system of communication taken up in semiotics and epitomised in Shannon’s diagram of communication. As Dolar has suggested, the voice is defined as other to the language it delivers. Such object, presence and material is not always outside of meaning, especially the non-verbal meaning and affective object voice in art practices.

The utterance of a word is the momentary expenditure of sonic energy in material space that is sound before it is language: as Nancy puts it, ‘timbre is the first correlative of listening’ (2007:40), or, for Brian Massumi, for raw experience:

The immediacy of visceral perception is so radical that it can be said without exaggeration to precede the exteroceptive sense perception. It anticipates the translation of the sight or sound or touch perception into something recognizable associated with an identifiable object. (Goodman, 2010:70)

As Cox (2011) points out, the aesthetic discourse of sound in art falls between different academic disciplines and the voice is divided even further between its sound and linguistic meaning. According to Sperber and Wilson:

Non-verbal communication tends to be relatively weak. One of the advantages of verbal communication is that it gives rise to the strongest possible form of communication; it enables the hearer to pin down the strongly manifest candidate, with no alternative worth considering at all. On the other hand, what is implicit in verbal communication is generally weakly communicated: the hearer can often fulfil part of the speaker’s informative intention by forming any of several roughly similar but not identical assumptions. (1995:60)

The realm of the visual allows an attention to detail by both audience and artist that is habitual in contrast to sound. We see visual detail where the detail in sound remains a riddle, for example where the sound of helicopter rotor blades resembles the blades of a fan. We are surprised or pleased that the sound is passed because of the visual side of the experience, yet this is the aural equivalent to a conjuring trick where something stands in for something it is not. This is the nub of a contrast between noumenal and phenomenal, real and symbolic, a fulcrum on which the voice exists. Sperber and Wilson argue that ‘in most cases what the communicator wants to communicate is partly precise and partly vague’ (1995: 59), and this can be extended from language to the paradox of vocal timbre.

That communication needs to be relevant to succeed is expressed by Grice as a series of maxims of quantity, quality, relevance and manner. Sperber and Wilson do not believe there is common knowledge between people; instead, we need to summon an expectation of communication for an audience to get a communication. Artworks are more in the business of making things manifest rather than sharing mutually known facts, and are experienced as a partially heuristic process by an audience. Communication is methexic but asymmetrical; that is, the relationship between audience and voice in artworks, and in everyday conversation between two people, is

---

not disconnected. People conversing ‘accommodate’ each other’s voices as material flow that can be thought of as a partial touching, just as Merleau-Ponty suggests that vision is touching. We can feel invaded by the gaze of the other and the voice of the other can have a powerful affect, breaking into our ‘sonorous envelope’. Touching involves a haptic proximity outside of language, notwithstanding systems such as Braille. Daniel Stern gives as an example of lived experience ‘looking at someone in the eyes who is looking at you and taking a deep breath while talking to someone’ (xiii: 2004). To equate the voice with touch is to evoke the subject within a material and object bath of sound, while requiring it to stand out as something that individuates.

Bakhtin (2006:69) says that sentences are grammatical and complete unlike utterances that relate to real situations. We don’t speak in sentences but utterances. Sperber and Wilson’s theory is attuned to the many pitfalls and possible failures of communication – ‘what is mysterious and requires explanation is not failure but success’ (1995:45). They argue ‘that when you communicate, your intention is to alter the cognitive environment of your addressees; but of course you expect their actual thought processes to be affected as a result’ (1995:46). We may associate verbal production with a clarity of communication of the communicative that is in a different realm to the critical discourses of art, which muddies the waters of communication. Is what I say true?

The evidence gleaned from the voice is both pragmatic and emotive, and art fully explores the irony and rhetorical powers of language. People witnessing the same event create:

different representations of it. While grammars neutralise the differences between dissimilar experiences, cognition and memory superimpose differences even on common experiences. A speaker who intends an utterance to be interpreted in a particular way must also expect the hearer to be able to supply a context, which allows that interpretation to be recovered.

(Sperber and Wilson, 1995: 6)

In Relevance Theory different meanings are not those caused by different intonations but instead a meaning is pragmatic and subject to ‘conversational implicatures’ in which meaning is implied by the hearer and not given by the speaker. This fits with Nancy’s description of the sonorous as ‘methexic’ – where the hearer is involved in the creation of the meaning. The voice unfolds in time and this is emphasised by its reproduction.
Artworks can use the voice as a complex temporal object that exceeds language in many ways. Nancy goes on to claim:

> In this resonance, the mutual mimesis of senses, if there is one, is not distinguished from the already evoked methexis: participation, contagion (context), contamination, metonymic contiguity rather than metaphoric transference. (2007:42)

We are pre-programmed to hear not only voice but also the language it delivers. If we hear a word it is hard to think of it as an abstract sound, although it is exactly that, a sound which Saussure said was arbitrary to the given meaning. However, a phenomenology of voice in Fine Art practice must emphasise the difference between voices in their settings and how effects such as repetition can exhaust semantic meaning or create exalted effects. A centre of attention emerges from the complexity of what we listen to and can perceive, surrounded by the immanent, obscured, unheard and unexplained.

Bakhtin suggests a definition of speech genres as being specific utterances, from multivolume novel to single word rejoinder (1986:61), but that these different quantities share thematic content, style, and compositional structure. This range of speech phenomena arises as a result of utterance produced in the material world, including the written or printed word, and this relates to Sperber and Wilson’s observation that:

> By definition, the semantic representation of a sentence, as assigned to it by a generative grammar, can take no account of such non-linguistic properties as, for example, the time and place of utterance, the identity of the speaker, the speaker’s intentions, and so on. (1995:9)

And Bakhtin notes:

> Emotion, evaluation, and expression are foreign to the world of language and are born only in the process of its live usage in a concrete utterance. (1986:87)

---

This can be addressed by using Bakhtin’s further expansion of the idea into primary (simple) and secondary (complex) speech genres. Primary speech genres are taken from unmediated speech communion but receive meaning by being in secondary speech genres. Bakhtin sees that the secondary speech genres are ideological, ‘a one-sided orientation towards primary genres inevitably leads to a vulgarization of the whole process’ (1986:62), and that language and life meet each other in a two-way relationship through concrete utterance. Bakhtin stresses that an utterance is the product of an individual and this is heard in the production of voice. Therefore the voice has to be viewed as reflecting the broader historical, ideological context of the secondary speech genre through the actual utterance we hear. There is much complexity in the relationship. For some voices we are not interested in the speaker, and their voice is the epiphenomenon of the communication; with other voices the subject is central, yet both factors are present. For art practice the object voice is intentionally presented, rather than being a situational epiphenomenon. Bakhtin’s idea raises the question of how these different factors join in speech genre, and can be applied to the voice in Fine Art practice.

Bakhtin identified a problem with speech genres: that little account was given to the relationship between the actual utterance of speakers and the culture, history and ideology of language, and how the two affected each other. As a literary critic Bakhtin did not actually delve into the material qualities of sound, but his ideas of utterance can help in thinking about the voice in art practice. Speech genres can be expanded to include the voice as an object and material incorporated into artworks, and the relationships between utterances and their representation. Bakhtin’s point is that to communicate, speech is a combination of style and genre that link it to a web of previous utterance. For Fine Art practice this web is embodied, material, and involves the disposition of object voices. Art as discussed in this project is an area where the voice is involved in a wide spectrum of activities which can be classified using the framework of speech genres.

To define and explain the theoretical splits found in vocal performance in Fine Art practice I have explained the phenomenology of voice, which both brackets out and locates the voice as a nexus within a context of space and culture. Relevance Theory is a linguistic theory that stresses the context of communication and equates to the notion of speech genres, without which voices are hard to position in discourse.
In *Trialogue*, voices are reaching out for the discourses that will cement meaning. The method of phenomenology is reductive, whereas Relevance Theory and speech genres are expansive. Together these theories come together to account for voices in *Trialogue* and the wider field of Fine Art practice.

Sperber and Wilson state that:

> A context is a psychological construct, a subset of the hearer’s assumptions about the world. It is these assumptions, of course, rather than the actual state of the world that affect the interpretation of an utterance. (1995:15)

The voices in *Trialogue* are not unfamiliar, yet they do not fit the genres with which we are familiar. In some ways the recording of voices enables repetition that removes the power of the voice, as Connor says:

> from being a source of powerfully mingled pleasure and menace, the technologically autonomized voice becomes a source simply of repeatable pleasure, or of the pleasure of repeatability itself. (2000:40)

Meanings vary for the age of a speaking voice, but the children’s voices in *Just Talk* have a pure quality, less complex than an older speaker and augmented by the black-and-white images across which they are heard. They are telling the audience which way to go on a journey, although the chalked images are drawn in response to what they say. Although the audience know these recordings are made for an art installation the voice also retains vitality:

> The technologies of the voice are actualizations of fantasies and desires concerning the voice which predate the actual technologies. (Connor, 2000:40)

How do genre and materiality meet in *Just Talk* in relation to the video being a psychological condition explored by earlier practitioners? The sonorous envelope is not just a psychoanalytical idea: all sound has an envelope – the time in which it occurs:
Envelope determines the shape and texture of a sound through time, from the moment it is first heard to the time it fades out. All sound has a beginning and end. (Krause, 2012:25)

For *Filling Space with Rocks* the voice is obscured by wind sound picked up by the microphone: ‘That is, we cannot record the wind itself – we can only capture its effects’ (Krause, 2012: 51). Bernie Krause is very particular about how he records sound, using the idea of a field protocol to reveal a recognisable milieu, but *Filling Space with Rocks* attempts to construct a vocalic space where rock, voice, wind, body all come together as materials but are also fragile in their status as representations.

To save our voices being confused with the wind Connor introduces the idea of a vocalic body, which, he says, is not described in psychoanalysis or phenomenology (2000:35). We can’t help imagining a body for an autonomous voice, and this applies to a voice with an image. Voice and image always create a frisson, even in the most unconnected physical relationship, such as that of the voice-over and image. Therefore to make a theoretical account of the voice in the artworks I am describing it is necessary to combine phenomenological, psychoanalytical and pragmatic theories that join an object voice to its context.
Chapter Seven: Voice and *Trialogue*

In the installation *Trialogue* the audience listen to the voice critically in its relationship with image and space, and in this chapter I analyse each of the four works of which it is comprised. My own voice is catalogued as a presence in different dispositions for different works to examine how authenticity and subjectivity join materiality and objecthood in the voice. The three screens of *Trialogue* form a hybrid space that alludes to the visual array of the triptych configuration of painting and multiscreen installation, to present, combine and critique discrete voices. Voices can harmonise in singing but how else can they be put together? *Trialogue* is a little-heard type of verbal intercourse compared with the more familiar monologue, dialogue or even polylogue. A triologue is asymmetrical, suggesting vocal interference. Four discrete video works are presented as types of utterance on three screens. The installation sets up the problem of how we reach the limit of utterance and voice, between sound and discourse.

The three screens of *Trialogue* are intended to provide a simple disposition of object voice in space. As Sterne says:

> Whether we are considering Plato’s Phaedrus or Jurgen Habermas’s structural Transformation of the Public Sphere, the idea of face-to-face conversation embodied in mutual spoken dialogue is central to Western philosophy. (2003:342).

Sterne thinks that privileging the single voice creates an authoritarian denial of many voices. The processes of production, transmission and reception of voice are variously effaced, whether recorded or not. In *Trialogue*, the complexities of simple disposition, the creation of a space for voices as neutral or pure, is examined.

The voices augment the presence and tone of the space, as comparative alterations of the ambient auditory array, and emphasise theoretical constructions such as Connor’s ‘vocalic space’ discussed in Chapter Eight, ‘Voice and Listening’. *Trialogue* puts the discrete works, *The Jazz Singer*, *Just Talk*, *Forget Me* and *Filling Space with Rocks* in a figure-to-ground relationship to question the framing of vocal performance and engage the audience’s listening. The space I am creating with voice and image is discussed later in this chapter. My voice fills the space of *Trialogue* and is joined by other voices.
The intention is to produce a number of distinct dispositions with *Just Talk*, *Forget Me*, *Filling Space with Rocks* and *The Jazz Singer* that are created by listening to voices, made object in this reproductive context. The effect is enhanced by a deliberate orchestration of the works within the space, involving slight overflow and sampling between the works. One of the screens is the road sign frame depicted in *The Jazz Singer* and this is intended to stress the gap between real and symbolic so prevalent for voices.

*Trialogue* presents my own voice, framed in different ways, in two video works, alongside two video works using the voices of children and a jazz singer. As a direct address, the voice floats between being something present and presented in different dispositions and modes of address. I am at a distance in a window onto somewhere else in *Filling Space with Rocks*, but I am a talking head in *Forget Me*. In *Filling Space with Rocks* my voice is being blocked by wind. My voice disturbs the air differently than the wind as an interaction with objects. The obtrusive sound of wind makes the audience aware of the unseen microphone that joins the megaphone that can be seen in the transformation of sonic material into meaning on a mountainside. This is an impoverished vocal recording both as sound and meaning:

> If the voice is produced by unstable bodies, transmitted through volatile air, and received by sometimes disobedient hearers how can voice be trusted to convey an individual’s thoughts to a listener? (Bloom, 2007: 4)

Like other artistic practices discussed in this thesis, the genre and mode here are cut loose from what is fully recognisable and expected. Am I talking to myself, is it a diary entry, or intended for the audience of an exhibition? These questions fall back on the mode of address, that extends out from the critique of a direct address to a camera.

In *Trialogue* the voice is reproduced object, presence, materiality, modal configuration and specific utterance that relates to a range of ways the voice is framed: screen, place, vocalic space, genre, and sonorous envelope. The voices in *Trialogue* are talking about voices, a journey, and an act of speechifying from a mountaintop, and an order to the audience about what to think. How persuasive are these voices in changing something in the mind of others?
From within the gallery space a hybrid of inside and outside is created where the screens are both objects and windows onto elsewhere. In *Filling Space with Rocks*, the voice is located on a mountainside because that is where I am seen speaking. Also, as with a news reporter in the field, the environmental sound galvanises the presence of my voice. The sonic perspective also emphasises the presence of the microphone and camera. What is common to a direct address to camera and a direct address to microphone? Only in early cinema do we lack reproduced sound, but other sound can be synchronised with the image. Sound and image become joined in the bracket of what is being addressed, their representation a sample of what they record.

*Trialogue* sets up a three-way relationship between screens and voices where there is some ‘falling and scratching’ and a reaching out for genre where the presence of voices causes both attraction and friction in fields of space and discourse. Three is a minimal crowd, and its vocal sound crosses speech genres with asymmetric turbulence, resonance and dissonance, interrupting and co-opting each other’s signals. For *Trialogue*, the space contains three screens one of which is the wall of the space, another a free-standing screen like an object itself, and the third, smaller, screen held within a metal stand used to support a road sign. These screens have different connotations and suggest the mixing of different utterances and genres.

*Trialogue* extends the voice art installation practices of Nauman, Cardiff and Bures Miller, and Stidworthy. *Just Talk (7 mins)* [Play Video] is a single-screen video that is experienced within *Trialogue* and is a collage of various vocal and visual objects: children’s voices, a laughter track, chalked images, hoarding in a street, a guitar. The images are literally black and white, as the mid-tones have been removed, unlike the sludgy grey of 1970s video seen in *Lenny’s Documentary* or *Semiotics of the Kitchen*. Sean Cubitt has described the replacement of black-and-white television with colour in 1968 as a significant moment for video, like the coming of sound to film. However, the emphasis of black and white is used here to accentuate chalk marks rather than as a way of resonating with the grainy image produced by analogue apparatus. The work also uses black and white tones to both unify the images and simplify their presence to accompany stark, simple voices as tonal fluctuations of sound. The children’s voices are also smooth, unlike the roughness of old voices, and in comparison, the laughter track

---

attains an authoritarian power. Monotone makes us more aware of the construction of the image and this is imported to the voice, whose tone is not monotonous. The work also takes inspiration from G. A. Smith’s film *The Miller and The Sweep* (1897) in which a miller and a chimney sweep collide and start fighting with a bag of flour and soot that mix into grey. The creation of tone is in contrast to the contortions of their bodies as they fight.


*Just Talk* was provoked by Gerhard Richter’s painting *Table* (1962), a black-and-white painting of a table interrupted by the gesture of an expressionistic blur of paint. The image and its paint fall away from each other like the phonic substance and language present in a voice. It is as if the paint is insolently making Magritte’s point ‘this is not a pipe’: the paint is not a table, it occludes the table, rubs it out as if the paint is wiped away – painting as the wiping of surfaces. During the video we hear an unexpected laughter track that sounds like heavy rain. This sound has a multi-stable timbre whose prosody and rhythm can be heard as the sound of different objects, both rain and authoritarian chorus.

Illustration 4. *Table* (1962) Gerhard Richter

---

91 Brandon Labelle (2010:147) describes sound poetry – a sense of how the voice operates *alongside* language without necessarily always arriving at words as the main oral target or goal.

92 A famous example of sound having two meanings is found in Ford Coppola’s *Apocalypse Now* (1979), where the sound of helicopter-rotor blades is faded into the sound of a ceiling fan.
It is within timbre that we find the limit of voice, where we sense the sound is no longer a voice; like a robotic weather front or wash of sonic paint evoked at the flick of a button. Recorded laughter is inauthentic because its production is spasmodic, involving randomness at the edge of representation.

When we use our voice to speak we exert a control that is lost in laughter. Is timbre at the crux and limit of what we can control? Rather than being either given or controlled by the speaker, timbre is also constructed in listening. The laughter heard in Just Talk is heard in the space of the street but the noise it creates effaces the physical trace of the space in which it was recorded. It could be heard as the random drone of rain but retains the ordering of audience response. Unlike the production of our own voice fading into the viscera of our bodies this vocal sound feels like it can be turned on like a water fountain in the street, in other words ‘canned laughter’.

Robert Provine has investigated the grammar of laughter and found that:

If laughter is triggered during speech, its expression must await the passing of phrase-sized chunks of speech-related inhibition. (Provine, 2000:38)

Provine suggests that laughter is an evolutionary precursor of spoken language rather than its opposite. These findings about the position of laughter can give us clues about the phenomenology of voice in Fine Art practice. On the one hand canned laughter is false, on the other it is used because it makes the viewer join the audience, as reproduced laughter is infectious and directly affective.

Jacob Smith quotes an interview with a television writer: ‘real audiences sound phonier than the laugh track. They freeze up and act unnatural’ (2008:43). Laughter is, in fact, a social glue, and we feel that this structure of meaning is retained while the unexpected context of the street in Just Talk make it menacing. This is given more force by the inauthentic nature of a laughter track. These tensions between diegetic voices are what inform Cardiff and Bures Miller’s approach to recorded voices.

---

93 Although more a feature and debate of an earlier period of television, contemporary situation comedies sometimes include audience laughter. If the comedy is shot like a film with one camera, there is no laughter. If it is made in front of a studio audience with a three-camera shoot the audience reaction is heard, except the laughter of the first take is often edited to a later take, creating semi-canned laughter.
The timbre of the sound does not provide its structure. The structure comes from the laughter’s relationship with communication:
Laughter is seen as a pure indication of individual human presence, an expression that becomes particularly significant and contradictory in the context of mechanical reproduction. (Jacob Smith, 2008:41)

We know that the laughter in *Just Talk* has not been recorded in the street: it feels more like something that has come along with the screen that frames performances with laughing audiences. The street as a vocotope is a place where we expect to overhear things and not focus on specific events. Although many important situations arise in streets they are usually places visited on the way to somewhere else. *Just Talk* starts in a real street but becomes more dominated by an imaginary journey described by children’s voices. The work becomes its own journey.

The potential of reproduced voices to be displaced in time and space, is integral to the development of culture and technology from telephones to cinemas, etc. and this affects object voices in life and art. The break between reproduced and un-reproduced voice falls into the origin myths more noted in cinema where audiences were rumoured to have thought images of trains were real. The voice always separates from its body even before mechanical representation and this is central to its partiality. Our habitual accommodation of voices obscures how the relationship between voice and its representation changes over time from the reductive formant frequencies of the telephone voice to the ‘auto tune’ software changing the voice of Cher in *I believe* (1998).

In the last words of *The Jazz Singer* a voice says how we are born into vocality and usually go on to utter our last words, held throughout our lives in sonorous envelopes. The first voices we hear are from inside our mother’s womb. Psychoanalytical theories posit the mother’s voice as a sonorous envelope:

> The image of the infant contained within the sonorous envelope of the mother’s voice is a fantasy of origins. (Silverman, 1988:74)

---

94 *Vocotope* is the relationship between voice and space and is extended from Jacob Smith’s coinage of *Sonotope* – the relation of sound with space adapted from Bakhtin’s *Chronotope*, the representation of time and space.
Silverman questions the constructions of the object voice suggesting, for example, that Chion’s analysis of the object voice leads to ‘the identification of the mother with sound and the father with meaning’ (1988:75).

Last words persist in memory but since the voice could be reproduced, our words outlive us, and not just the last ones. Like the objects of human culture sent into space on spacecraft we can guess that the recorded voice will become like fragmented hieroglyphics: for Sperber and Wilson communication is already like this, and this is how I approach the voice in my practice. In The Jazz Singer a jazz singer speaks about his voice. The voice is in front as a result of the dialogical relationship we have with its owner. This raises phenomenological questions about whether we orientate the world or the world orientates us. [Play The Jazz Singer]

In The Jazz Singer we hear the results of an interview with a singer. He is free and happy to give information about his voice because he feels that he is never usually asked that much about his voice, even though it is ‘my instrument and part of my body’. He is thinking about his voice in a way that many people do not. Dolar (2006:19) argues that phonology kills the voice in reducing it to abstract forms outside of context. This situation can apply to any de-situated voice, but also becomes the material for art practice. The voice is vital and the jazz singer feels it in his body (in opposition to how the speaking body fades out of perception from the neck down). We may think, rather, that the listener is affected by the singer’s voice when it enters their body, reflecting Goodman’s observation of ‘the body as transducer of vibration as opposed to a detached listening subject isolated from its sonic objects’ (2010:47). The voice is a socially cohesive agent that connects our subjectivity to the material of the world.

The voices of jazz singers are known for their distinctive timbres: ‘Frank Sinatra or Billie Holiday are loved because they are recognized’ (Connor, 2000:38). The singer mentions Frank Sinatra. He thinks it not disadvantageous for such singers to ‘produce cracked notes, or be off-key slightly, with some of the words forgotten’. The Jazz Singer shares its title with the famous film The Jazz Singer (Crosland, 1927) in which Al Jolson announced ‘you ain’t heard nothing yet’. The jazz singer states that ‘one thing affects another all the time when you are using your voice, when you are a singer’. This

---

95 In Beckett’s Krapp’s Last Tape (1957), a character listens to recordings of his self at younger ages.
96 Using YouTube we can currently survey the changes in the voice of a celebrity throughout their life.
is clear in Goffman’s idea of what is ‘given off’ rather than ‘given’. Language is given but the voice betrays things about ourselves that we do not want to disclose. The voice holds a plethora of marks and affects.

The recombination of voice and image in the editing of The Jazz Singer speaks of a dissonance between different speech genres. The jazz singer speaks about his voice as if giving a talk, but there is also the trace of an interview. He mentions how he is coming to speak about his voice and how this thought has affected his performance the night before. The work echoes the disposition of Pythagoras whose face was hidden by a screen. This screen is moving and the reason for this agitation is not clear.

The reason is revealed when, at the end, the paper falls to the ground revealing myself crouched behind the frame: but I am not the singer. The singer is aware of the power of his position, and that of his voice when he performs, and cites the example of reducing a noisy pub to a state where you could hear a pin drop. The discourse is shaped by the time and space of disposition. A theatre audience is more focused than those in a pub, the television audience less likely to remain seated than those in the cinema. A gallery audience exhibit different behaviours, and this effects how the object voice arises in art practice. Where the cinema pushes the audience into a chair, the gallery space is less formulated as a site, and art practice flows around how the audience occupy space: 

Trialogue is intended as a hybrid space.

The singer speaks as a blank sheet of paper is animated in a room, scuffing the floor, flapping from between floorboards, and rustling while held on a road sign like an easel. The random noise of the paper scrunching is heard with the voice. Images then appear on the paper randomly hitting its surface like the ‘off-key, and cracked notes’ of the voice being described. These black-and-white images can be thought of as the out of synch accompaniment to the changing genre of inner voice, fragments worrying our consciousness. The Jazz Singer says he prefers such a voice even if it is no longer

---

97 Artist Gary Stevens notes how a theatre audience are ‘with the performer’ more than a gallery audience, who are not – conference paper ‘A problem object’, Exhibiting Performance Conference 3/3/2013 Uiverstiy of Westminster.
Illustration 6
technically proficient. The slight damage and patina reveals the marks of previous actions. We think we hear experience in the voice, as we do in the lines of a face.

In *Filling Space with Rocks* [play *Filling Space with Rocks*] I speak through a megaphone from a mountaintop. Vocalising from a mountaintop suggests that you want to be heard, except there is no audience for this performance, apart from the sheep that
can be made out beneath the noise of wind. The tone of this performance is unclear because there is no audience reaction to anchor it in discourse. The audience are alone in their reaction; the only laughter track is the sound of wind over rock, grass and microphone. The escarpment of grass is a wedge shaped like a ‘louder than’ symbol. Although I say this is ‘good rock’ it is a type of geological aggregate that does not fit the image of a solid rock, although very solid and regularly used by rock climbers. On the one hand the performance is comic, yet it also has a vulnerability that comes from confusion between the physical and conceptual power of the voice.

Steve Wurlzter (1992) uses a quote from a press release by the management of Whitney Houston as the title of an essay: ‘She Sang Live, but the Microphone was Turned Off’. I am in a similar position as the megaphone is not turned on, but the shape of the cliff provides natural amplification. A speech is not a performance in an aesthetic sense. It also has the feeling of an impromptu television presentation. The megaphone suggests the barricade or the hustings. I am making sure I can be heard by using a megaphone, whereas there is no need in a film as I could use a microphone (which is what is happening) or record a voice-over. The speech is falling between a script and spontaneous utterance, as if I am filling space with voice, or have been asked or forced to just say anything. The voice produced and recorded is a product of both mountains and my body. There is no need for a megaphone because, like the amphitheatre of the first democracies, the space naturally amplifies and is like the rocky site at Pingvellir, which is on a tectonic fault line and the site of the first Icelandic parliament.

The solidity and geological timescale of rock houses the ephemeral voice. Early Greek democracy involved one voice amplified to the many in an amphitheatre, like the presence in a witness box. The range of earshot for this assembly represents the correct number of people for live democratic accountability. Filling Space with Rocks and Just Talk allude to different vocotopes: the street and parliament. The voice retains political rhetoric as power of address, a hopeless enactment of Althusser’s theory of subjective interpellation.98

---

98 Althusser describes how by turning to the voice of authority, such as the police, the subject is positioned in hierarchy: ‘the imaginary relationship of individuals to their real conditions of existence’, ‘Ideology and Ideological State Apparatuses (Notes towards an Investigation)’ in Mapping Ideology. London:Verso Books (2012:103).
I discuss speaking alone from the mountain, but the camera operator comments that I am not alone. Then, I answer my mobile telephone and again I am not alone. I enunciate through the megaphone ‘yes’ and then answer the phone ‘yes’, stating I am ‘up a cliff’.

The phone ‘yes’ places me in space more than the megaphone ‘yes’ whose meaning is unclear, as rhetoric requires an audience. Where the phone ‘yes’ feels linked to possible GPS surveillance, the rhetorical ‘yes’ is meant to go out to an audience. I predictably state my position. I end by saying ‘an uncontrollable desire to speak constantly, not only that, but to blurt it out all over the place’. This could be extreme behaviour or could be what humans do all the time, but speaking and listening require timing. In this performance, time and place are out of synch.

The view from the mountains is sublime, the voice less so; it’s comically ephemeral. My speech is lost between genres: it is neither political rally nor television reportage, and is Dolar’s object voice, not song, medium of language, but a monologue displaced from audience response. The simple act of using a megaphone is embedded in a nexus of discourse and disposition. Although this piece focuses on my voice its meaning is quite different to my voice as experienced in Forget Me, which explores self-presence.

For Filling Space with Rocks the romantic the scenery of the background affects the voice. The sound of the space has the grandeur of a work like Nauman’s Raw Materials and is contained within my voice.

We cannot see a voice, but we can see the technology that enables the reproduction of voices: telephone, loudhailer, loudspeaker, and mobile phone headset. With these objects we see the disposition of the reproduced voice in space and history. You cannot tell if a loudspeaker is working by only looking at it but you know if someone is using a megaphone, which cannot speak by itself. You can tell if someone is using a telephone by looking at it. The telephone requires a bodily gesture to bring the voice out of the body. It is an apparatus that stands in and around the voice that cannot be seen. The same holds for the megaphone, but the context is at the opposite end of the scale between public and private from that of the telephone: the gesture and rhetoric are both bigger. The audience witnesses both public sphere and sonorous envelope on this mountain as a result of megaphone and telephone.

In my work Forget Me [play Forget Me], I pose a similar question in asking the audience to differentiate between myself and what I say. Abramovic’s Freeing the
Illustration 8.

Stills from Forget Me 2.5 minutes
Voice (1975) is compatible with the phenomenological methodology of late modernism that I probe in Forget Me, and is a response to the many uses of video as mirror or acoustic mirror.

In Forget Me the audience is thrown into the trap of language that Serres suggests is set by the sound of the voice, with the contradictory instruction to ‘forget what I say and just remember me, forget me and just remember what I say.’ This is uttered by the artist’s head as it drifts across the screen in a video transition that ponders the present moment as if it is attempting an oxymoronic slow cut between shots, in this case the same shot, a long drawn out goodbye recycling as hello, a hybrid of cut and fade. It is ironic that reproduction of the sonorous aids memory of the indexical over what is said.

The repetition of my voice in Forget Me emphasises the quality of the voice as an object in contrast to the absent body and the head and also as an object framed by the screen. The movement of the talking head, like a scrolling text, emphasises the duration of this object in time and space as it recedes out of the present moment into memory.

These kinds of correspondences between voice and image, the attack and decay of sound that can spread from auditory to visual array of the image, are produced in Forget Me in response to Husserl’s idea of protention and retention – a past and future part of the present moment. What I say in this work stays in the indubitable flux of the voice’s present moment, which has been described as a three-second duration that relates to the rhythm of human breath and that Stern has related to ideas such as the ‘specious present’, developed by William James. My speech relates to a form of address in Brecht’s epic theatre, described by Devin Fore as ‘free indirect speech’ (2012:164), in which one becomes confused about who is speaking. Forget Me also uses the rhetorical devise of paralipsis where something is brought to attention by an ostensible passing over of the subject.

The presence of a voice can evoke the absence of something else, the memory of something else:

Verba Volant, scripta manent; Lacan reversed this classical proverb, since it is only the voice which remains there, on the spot where it was emitted and which

---

it cannot leave, where it is born and where it dies at the same moment. (Dolar, 2006:59)

Dolar’s point is that although it seems as if voices fly away in relation to the dead letter of writing, in fact the utterance of a voice is unique to its time and space. Clearly this is important to how voices communicate as shown by Relevance Theory and speech genres. The time and place of utterance are vital to meaning as well as the discourse and disposition that are made present in art practice. Dolar does add that in the twentieth century recording technology has changed this, and recorded voices, ‘from the whirlwind of history’ (2006:59).

Dolar echoes Barthes idea:

Speech is irreversible: a word cannot be retracted, except precisely by saying one retracts it … paradoxically, it is ephemeral speech which is indelible, not monumental writing. (Barthes, 1977:190)

A written word can be rubbed out and forgotten whereas speech remains the gestalt of its context. We get a phenomenal mix of perception when we hear a voice that extends past the words used, and we may be hard pressed to remember exact words. Rubbing out written words creates a palimpsest. The palimpsest retains a trace of the word as image emphasising previous presence like a ghost (Zizek [1996:101] has noted the voice’s ghostliness as something disembodied). The voice is an object engaged in a continuous process of rubbing out, ephemeral yet marking a gestalt of an exact time and space. This is a positive rubbing out, giving life rather than extinction, whereas the voice is an excitation of the air.

The different video works in Trialogue frame voices to explore their relative functions as reproduced objects. Genres of voice are presented in conjunction with depicted and actual place and space. The specificity of place is different to the generic nature of space, and such oppositions are heard in voices. As part objects, voices resist ‘P.F. Strawson’s influential claim that auditory experience, in contrast to visual and tactile-kinesthetic experience, is not “intrinsically spatial”’(O'Callaghan, 2007:29). Seeing and speaking are explored in Trialogue to question how the disturbance of the ambient auditory array by voice is simultaneous with something changing in the mind of the
audience. For voice, genre like timbre, is without absolute value or clearly demarcated limits and becomes ascribed by the tacit agreement of the listener. The next chapter investigates the nature of listening to voices in Fine Art practice.
Chapter Eight: Voice and Listening

‘The world of sight appears to be there; pressing on us without remission, the world of sound is only ever there at the moment of our hearing it.’ (Connor, 2000:17)

In this chapter I discuss how artists have interceded into our listening with an emphasis on the materiality of object voices. Listening is a complex subject to unravel, especially if it includes listening to our own voice, but voice has to be listened to by its speaker to be articulated at all. The voice can travel a long circuit, moving across time and space, or a short circuit that never leaves our head. Our listening to our own voice, transmitted through the substance of our own body, as well as reverberating in space, is different to how anyone else hears our voice:

In acoustic terms, there are, two versions of my speech; the ‘private’ speech I hear inside my body and the ‘public’ speech that others hear outside.’
(Bruce Smith, 1999:7)

First-generation video artists, for example, Richard Serra in Boomerang (1974), Marina Abramovic and Vito Acconci, strayed into these issues, as I have already noted. I am extending Rosalind Krauss’s (2008:215) observation that video installation made the viewer aware of the viewer’s body, and single-screen work the artist’s body, to the voice. The single-screen works investigate the voice of the self and installation works a wider array of voices, the voices of others. Both of these tendencies are present in Trialogue as an installation in which we experience the single-screen work Forget Me, where, among the other three video works, the voice is reduced most. The voice can be broken down into many categories, from a background noise to the central object of perception where the limits of the field, in which we hear voices, are created by different types of listening.

We choose to listen to other voices because our experience and learning pulls us towards voices. Thus listening has a cultural construction that I am pursuing in Fine Art practice. Trialogue asks the audience to listen to voices in an installation, and this chapter investigates the listening experiences provided by Fine Art practice. Jean-Luc Nancy says listening differs from hearing, where listening ‘stretches the ear’ (2007:5) and I try to expose this for the voice in Trialogue: ‘hearing means understanding –
“hearing” is “hearing say” rather than “hearing sound” (2007:6). When experiencing a voice we may favour its linguistic meaning over its sound or the opposite for a singing voice, but there is always a balance between the two in any voice. My practice, and the works I discuss, find common ground for this ‘compulsive dualism’ (Rée, 1999:2) where the voice is embedded in several simultaneous discourses.

The listener must focus their listening, and the gallery joins cinema and theatre as a site with an impetus to hear voices, and where the voice can ride many acoustic storms and impediments within Fine Art practice. In previous chapters I have stressed the context in which voices are heard in Fine Art practice, with the idea of disposition to connect sonic material and meaning. The voice is presented by Fine Art as the partial object of our experience rather than a leftover of the process of communication. As Bruce Smith says:

> sound without logos is, to Plato’s mind, noise. For if sound is without logos there can be no syntax. Our own culture’s anxiety in the presence of the non-verbal sound has a very long history. (1999:135)

A voice exists in space and time, but our own voice, in its protective, sonorous envelope, can feel far removed from reproduced voices that travel through extended time and space. To explore how meaning is created by the voice situated in time and space I will now discuss the relationships, for Fine Art practice, between time, space and listening. I begin with how listening brings timing into genre, as initiated by Bakhtin’s Chronotope (the relationship of space and time in literature).

In *Trialogue*, the audience is in a particular space to experience the work in order to emphasise the relationship between listening and timing. Space and time are joined in an act of listening: if we do not listen at the right time we will not hear, but timing also frames voices in many other ways. For example, timing is expressed in voices as rhythm, which it has in common with rhythms of moving images. The images in *Trialogue* are explicit as medium, material and montage to emphasise the object, rhythm and material of the voice.

---

The screens in *Trialogue* are not just a representation of another space but produce a hybrid space, a chimera that disobeys and contravenes the metaphor of bicycle riding with which Barthes describes the accommodations of communicating with speech:

> We can only make ourselves understood if we maintain a certain speed of delivery. We are like a cyclist or a film obliged to keep going so as to avoid falling or scratching. (1977:191)

The listener accommodates voice in a range of ways in Fine Art practice. Barthes’ example of a film running at a particular speed in its mediation of the image is evident in parts of *Trialogue* where the rhythm of voice and montaged image are equated. Linguistics uses the term *coarticulation* to describe the mix of consonant and vowel sounds in speech in a particular language. This type of rhythm migrates between voice and image in *The Jazz Singer* and *Forget Me*. We hear the voice of a jazz singer, but we do not see him. Instead we see other images that vary in rhythm and duration. In *Forget Me* sections of fast cutting relate to the rhythm of speech. Rhythmic objects affect images and stress the temporal connection of perceptual modes causing friction in how we perform a reduced listening to these voices.

Using and thinking about the voice as object and material raises questions about the duration of objects. Our perception is about timing in the world, our bodies are a set of rhythms amongst the rhythms of the world. Voices structure our consciousness of our environment. Human breath vitalises a time signature that brings about the speed of speech. Our lives are a set of durations, the breath, the heartbeat, the life span and biorhythms amongst the movements of planets, atoms and clocks. The present moment is of a three- or four-second duration related to human breath. The limit and horizon of an object voice is defined by scale, just as planet Earth is an object by which we are swamped, but which also sits easily beneath our feet.

The human voice is the most striking example of the bridge between hearing and listening and this is brought to our attention by the theories of Saussure, Ihde, Bakhtin, and Sperber and Wilson, but also the work of artists such as Nauman, Stidworthy, and Cardiff and Bures Miller. What is effortless and habitual as a tool is also a rich subject for the critical practice of these artists. The act of listening to voices offered in
contemporary art by Philipsz, Cardiff, Aura Satz or Lavinia Greenlaw101 (to give some recent examples) work with an unexpected disposition of voice in space and time. The works in *Trialogue* also create new tensions in the disposition of voice.

Artists question and interrogate how we listen to voices. Although it is hard to speak while running, the physiology of our ears allows us to hear while moving, adding orders of meaning if we wear headphones (anywhere is turned into an auditorium). The macro realm of Doppler effects and wind interference do not usually arise within bodily movement. We ignore the borborygmic (rumbling) sounds made by our heart or stomach, but our stereo hearing creates effects of echolocation. Yet a concentrated listening might suggest the sedentary stillness of the auditorium (whose etymology makes the point – a place, and therefore disposition, to listen). Fine Art practice utilises unfixed relationships that arise from the mobility of a gallery audience to move amongst, or in relation to objects.

Susan Philipsz has placed amplified singing voices in urban streets in *Surround me: a Song Cycle for the City of London* (2010), and under bridges by a canal in Glasgow in *Lowlands* (2010). These voices are unexpected and atomise into parts: urban, song, voice, amplification and place. The work can be seen as the meeting of different temporal registers and types of listening, from the aural glimpse of a reproduced song in the sonic patchwork of the urban environment, to the focus and arrest given over to the listening of a whole song. A voice is brought into focus if we are standing still while remaining a background sound if we are moving (but this is less clear if wearing headphones). The moving listener has a more tangential relationship with a voice they hear if their own movement emphasises its temporality as a representation different to the time in which the listening is happening. How we listen takes on a different presence to how the song, as an existing entity, has been previously heard.

Unlike voices heard via headphones in Cardiff’s audio walks, the precise loudspeaker positioning in Philipsz’s work makes voices pervade the whole available space. The intention is to activate the space by filling it with voice, and a comparison is set up between voice-over in a film and public address in the street. Although both are representations, the dispositions give different meanings. In the street a voice has to

---

fight with other sounds and voices, whereas the cinema auditorium specifically focuses looking at a screen and listening.

In a cinema the sound is created for the stationary, seated viewer. A surround sound recreates the three-dimensionality of a represented space. In cinema the film holds the viewer still, as if the viewer is stunned by the projection. By not moving around, the viewer is avoiding the problems of diegetic space: for example, if the audience had to move nearer to the speakers causing a vocal hot spot on the flat screen, this would break the diegesis.\textsuperscript{102} Just as our ears produce a stereo effect along with our eyes so we perceive three-dimensional space more effectively with both ears and eyes. This is the area of representation probed by Cardiff and Bures Miller. When Bures Miller describes the making of \textit{Paradise Institute} (2001) he explains the use of binaural recording to capture the sound of a cinema in Berlin used in the work. In the presentation of \textit{Paradise Institute} at the museum of modern art in Oxford (2008), the audience lean over a model of a cinema with a tiny screen while wearing headphones. The sonic effect is to represent the effect of being in a cinema, but also to break the diegesis. During the screening an unexpected virtual member of the audience whispers to the viewer wearing the headphones. This instigates the decision Dolar (2006:81) says is needed about the ontology and epistemology of voices arriving from in or outside our head. The uttered voice has a phenomenology as it permeates our bodies, being both inside and outside.

We choose to enter a cinema or gallery but to find a voice issuing from a loudspeaker in a street brings with it a demonstrable disposition of power that is evocative of both police state and the banal structural rules of public announcement. In all listening there is a tacit agreement by the listener to enter the loop of meaning whether by choice, coercion or fear (and note also the imaginary voice of the Lacanian ‘Big Other’ discussed by Slavoj Zizek\textsuperscript{103}). Philipsz’s installation of voices in the street is a different iteration of Dolar’s example (2006:3) of a soldier retorting that a battle order is the sound of a beautiful voice. Philipsz places an aesthetic object voice in a disposition with authoritarian overtones, where listening means obeying.

\textsuperscript{102} In \textit{Power} (1980), Richard Layzell speaks to camera with the microphone turned off, as a Brechtian way to emphasise the control we submit to in listening to reproduced sound and image.

\textsuperscript{103} In \textit{How to read Lacan} (2006:8), London: Granta, Zizek describes words piped into the ears of actors in a Mexican soap opera rather than the use of a script, and how this relates to ‘the Big Other’.
When we listen up we listen through the disruptions of the material world, but the acoustic battle is easily lost in an environment where a voice is swallowed or half heard within the ambient auditory array. The signal to noise ratio extends to our body’s displacement in space and its own noises. To listen requires time, and I would compare this to a comedian ‘stepping on’ their punchline and removing the space in which the audience laugh. An audience constructs a space to listen, but timing is also important.

If the sound of the voice is effaced to get to the linguistic message, then the spatial signature is often even deeper in the mix. It is recording that returns this to our attention. The voice is always a striped-down phenomenon of cause and effect. We are always balancing what we focus on and therefore take from a voice. Reverberations accompany all our negotiations of space and while we do not rely on them as much as bats, echo-locating to gauge space, we are aware of the sound of space and its effects on voice.

Jacob Smith extends Bakhtin’s idea of chronotope to sonotope to describe the ‘intrinsic interconnectedness’ of sound and space (2008:245). In this construction voices indicate spaces and frame performance. This can then be extended to vocotope. Smith gives the concert hall as an example of a sonotope. My practice has used the image of a street, and this is not an obvious sonotope. Voices in streets are more in the zone of the overheard where voices can be hard to hear and are snatched like aural glimpses of people. The empty street precipitates voices such as voice-over and inner voice that are very different to a field recording made in the street. This reflects the uneasy boundary between private and public for voices in streets, but also voices in any space.104

Philipsz is subverting or changing the vocotope of street to allow a consideration of the relationship between voice and street. The street is usually public, open and redolent, in the current milieu, of an endless traffic which does sometimes abate. In this regard it feels like the flux of a voice. Images of streets vary from the quiescence of metaphysical painting, the emptiness of surveillance images, to scenes of riots and carnivals. Streets are public places open to anyone, the plurality of which does not create a single voice, but creates, instead, snippets of conversation or the noise of crowds. Streets are open to

104 City streets used to resound with the voices of hawkers, and rag and bone men. We no longer hear newspaper sellers or the cries of market stallholders. All that is left are the musical chimes of the ice-cream van. In Meat (Couzins, 2003), a song describes the cries of an Edwardian door-to-door meat seller.
the weather and are the stage for street entertainers, hawkers and town criers. To pipe voices into streets invokes authority or control like the subtle ideological structuring of muzak, but phone calls remain in a proximate envelope of time and space. The street, as a place we hear voices, has a two-way theoretical traffic. Barthes describes the body located in viscera physically apprehending the listener in sound. Bakhtin argues from the other direction, where language is always borrowed from the culture outside and brought into the body.

*Just Talk* starts with images in a street, as a boy’s voice is heard. This is a voice-over, rather than a voice located in a particular space as heard in *Filling Space with Rocks*. Neither is the voice piped in by loudspeaker, as witnessed in Philipsz’s works or Cardiff and Bures Miller’s installation of speakers in a gallery in *The Forty Part Motet*. In *Just Talk* the protective sonorous envelope of the boy’s voice is countered by the surging laughter of the crowd, which comes from a much more collective and ideological space. What constitutes an auditorium, a space for listening? The construction of this space can be explored in Fine Art practice, but it is reproductive technology that has allowed the conflation and mixing up of auditoria. The cinema, gallery or virtual online environment are single use auditoria, in comparison with Fine Art practices whose auditoria can be thought to be more multi-use, running from the critique of ideology by Brecht through to the practices of Relational Aesthetics. This is brought about by self-reflexive critique, and the desire to question the boundaries of art and life, where the voice is not anchored in one discourse. It also presents a divide of real or represented space. We could say that in Philipsz’s work the space is real and the voice represented,\(^\text{105}\) that in *Trialogue* we have both space and voice represented, but in some work, such as that of Tino Seghal,\(^\text{106}\) there is no representation. The voice can only exist in space, which is also real even when representing something, or somewhere, else. As I have stated in ‘Voice and Reproduction’ these works test such boundaries.

When we listen to a voice we are at a nexus of time and space, both of which can be effaced, like many other aspects of voices, to facilitate language or the construction of self. This also causes the voice as a partial object, and an object exists in space. How do

---

\(^{105}\) A reproduced image with a live voice, used in some early cinema, brings up an extemporary situation, like a performance of music.

\(^{106}\) For example, *These Associations* (2012), Tate Turbine Hall.
time and space join in a voice? The voice has multiple relationships with types of space in Fine Art practice that effect how we listen.

As has been stated, in realism we might ignore the frame when we look into it, and similarly, Tim Ingold (2007a:6) has suggested that we ignore the sound of the voice while receiving the meaning of words. I have said that in modernism this is reversed, and materiality is not ignored, ‘painting shifts from semiotic absence to phenomenological presence’ (Fore, 2012:193). This is discussed by Merleau-Ponty as an embodied experience and presented as technique in Schaeffer’s ‘reduced listening’ or ‘recovering language as pure acoustic value’ (Fore, 2012:198). Modernism thought voices should speak with their own form. For example, Melissa Trimingham has written about Oskar Schlemmer’s move to formalism, prescient of Schaeffer’s ‘reduced listening’ in his use of sound in performance:

Sounds are simply ‘present’ to us: the way we receive the ‘soundings’ shapes the meaning in a phenomenological direct way that often escapes our consciousness. (2010:138)

I am suggesting that, in the artworks I am making and describing, the audience look and listen at space in the way a sailor looks out to sea, and that this has a phenomenological aspect; that the space thought of as an offing can alter how we experience and think about the voice, an offing of sound and vision whose shape is unfixed and allows focus on the materiality of the space it frames.

We often hear rather than listen to voices, as Jean-Luc Nancy (2007) concludes, but in Fine Art practice different emphasis and manipulation of space can instigate different types of listening. The voice can arrive into experience in critical ways as well as being the familiar sound we hear. The palimpsest of past experience accrues more traces. If I vocalise, I am disturbing the ambient acoustic array, the molecules of air, but because sound only exists in an envelope of presence that falls away into memory, this memory is the palimpsest on to which I am joined. I am joined to vast variations in registers of time, from what I just said, to the milieu of Florence Nightingale and the sounds present in language that predate recording technology.
I have described how the invisible object voice inhabits the space in front of us, due to the frontal disposition of our heads and bodies, and screens mimic this disposition. Sound, however, is experienced all around our body in a spherical zone of space. Voice can only be experienced in space yet the variety of spaces involved and evoked is wide. Unlike the image, the voice cannot exist in two dimensions. The partiality of voice and represented image join and make up for a lack of a third dimension in an image and the invisibility of the voice. These lacks are what give the voice and image their power as representation. There is an opportunity for Fine Art practice to explore the ambiguity of two- and three-dimensional space experienced when voices join screens. Bruce Smith examined the singular act of making the sound ‘oh’ and he found its multiple discourses to be ‘a physical act … a sensory experience, an act of communication, and a political performance’ (1999:3). These discourses exist in Fine Art practice and to them I am adding the construction of space that accompanies voice, complicated by representation: the reality and power of reproduced objects in space.

Our invisible and silent inner voice creates a structure of meaning like voice recordings that replay and cut between genres in our head, but these are heterogeneous. One can think of one’s inner voice as a voice never subject to a microphone’s structuring. Its non-reverberation shares the quality of voice described by Dolar (2006:39) as ‘a pure immediacy where one is both sender and the receiver without leaving one’s pure interiority’. The effects of reproduction are anticipated in our minds. Between recordings and live utterance, the voice rebounds between the real and imaginary. From inside to outside, in these works our inner voice meets an expanded disposition of voice. The relationship between raw being and culture is funneled through the disposition of looking and giving voice in these works. In Nauman’s Raw Materials (2004), we hear the work Get Out of My Mind, Get Out of This Room (Nauman, 1968), which metaphorically suggests a direct correlation between inner voice and the space of a room. By contrast, in I am Sitting in a Room (Lucier, 1969) the artist’s voice is represented as the actual space of a room. Lucier provides an idea of how the partial object voice makes contact with space. The object voice feels autonomous of space, and this is the quality that supports its liveness as a human tool, but the voice always needs space. The reverberation of voice in space has a corollary with silent inner voice.

Some spaces give rise to voices in a manner that recalls Bachelard’s analysis of the poetics of space; that is, Bachelard’s notion of the poetic potential of space may be
useful as a way of thinking about the disposition of place and voice.

Ihde and Nancy have written about listening as a phenomenological activity. My practice presents voices and changes their relationships with space and particularly the two-dimensional space of images. The installation moves from silence to voice, voices in different relations, cacophony, noise in new spatial and temporal phrasing. The voice that is a manic cry can also perform a quiet whisper. How we listen is affected by space psychologically, and what we hear is affected by space physically. The two connect as timbre, which is more of a construction than we realise as it slips into consciousness as an accepted fact.

The flux of actual voices means they are not rooted in one place and time, and art practice interrogates this paradox. Roman Jakobson said that ‘the technology of film allows sound to be visual, sound no longer from the body… voice was never only from particular articulatory sites’ (1978:17). The ‘thenness’ and ‘nowness’ of a reproduced voice equate to when the voice was recorded, and when it is played back. But these two spaces are not as easy to define as an original and a copy.

A palimpsest is not a new surface but rather holds the trace of previous images, whereas voice leaves no trace, only memory. However, voices create spatial structure: what we hear is a function of the space in which we hear it – a good acoustical space, an echoing voice, a windswept mountainside. The empty space is pregnant with discourse and unrealised inner voices. The palimpsest looks backward as a form of retention where we look forward into the offing and into the future as protention. Our subjective immersion in the object world involves a move out of the palimpsest and into the offing. The sound in the street is littered with voices of which we are aware but do not consciously hear because they belong to somebody else’s conversation. They have become part of the everyday background sound, the wild sound of a film recording, or a soundscape. Where Krause (2012:179) extols the virtues of using spectrograms of soundscapes to record environmental changes, Ingold (2007b:10) disagrees with the use of the term soundscape. The term we use to describe sound comes from the discourse we use to record or hear it:
Wild soundscapes are full of finely detailed information, and while a picture may be worth a thousand words, a natural soundscape is worth a thousand pictures. (Krause, 2012:71)

This does not speak of what these sounds mean only that the spectrograph image is a tool for measuring change, probably for the worse in aural ecology. Such diagrams tend to put vision and sound against one another, instead of adopting a phenomenological combination of sensory perception. Ingold cites the ideas of Victor Zuckerkandl:

The sky is not an object of perception, any more than sound is. It is not a thing we see it is rather luminosity itself. In the experience one has of looking up in to the sky, according to Zuckerkandl, lies the essence of what it means to hear. If this is so, then our metaphors for describing auditory space should be derived not from landscape studies but from meteorology. (Ingold 2007b:12)

Although supported by the unstable flux of breath and wind the voice is received as a partial object. The voice instantly sets up a hierarchy of perception that Chion calls ‘vococentrism’ (1992:5). Chion thinks that voices structure sonic spaces in the same way that human bodies structure space. Again, it is not the words but the context: ‘Context is a structural given not of language but of speech and it is the very status of context to be reductive of meaning’ (Barthes, 1977:191). Barthes, like Foucault, writes about how language is literally a law: a teacher adopts a particular voice to be understood just as an artist adopts a genre in their work. This is what is contained in a vocal palimpsest.

The history of technology changes the meaning of space; the cultural and aesthetic possibilities of recorded voices are symbiotic with the history of recording technology:

In the modern age, sound and hearing were re-conceptualized, objectified, imitated, transformed, reproduced, commodified, mass-produced, and industrialized. (Sterne, 2003:2)

Our listening in art practice is shaped by structured intentions, where we are led into new incitements by voices. Trialogue attempts to elevate the status of listening with looking and the voice steps forward in the mix between perception and meaning. In
Filling Space with Rocks and Forget Me we know, or easily believe, I am speaking because I am in view. However, in Forget Me I ask the audience to forget what I am saying. I do not ask them to forget or remember the sound of my voice, the success of which would be difficult to corroborate. Voice is emphasised by a breaking of the rules of Relevance Theory and as a kind of non-productive expenditure. In drawing attention to the tacit work of listening Forget Me clears space to think about the constructions of voices. It also focuses on the act of listening as a possibility: the voices in the installation are presented for listening, something we do all the time, but here they are like objects stalled for examination and reflection.

Space can be construed as empty space waiting to be filled with voices. We can become aware of silence, absence and emptiness in different ways, as the gap – between words and breaths in space – that is filled with expectation, memory and sound. The voice is powerfully affected by place. A spatial spectrum opens for space that can, on the one hand, repress voice or on the other causes voice: where we are forced to be silent or incited to shout or sing loudly, like in a church and sports stadium. Yet the space of these places is changed by voices all the time and the space changes the voices heard: from the tumult of the hymn and the quiescence of the prayer, breaths held for the ‘on your mark’ of the start, to the audience’s vocal crescendo at the finishing line, as if the voices of many have crossed the finishing line as a swarm of sound. Time and space are linked by voice in these utterances.

What we hear when we listen is affected by what we remember, and is constructed in culture. Fine Art practice creates many species of space in which we hear the voice. The jazz singer mentions a space in which we can hear a pin drop. Such silence could ensue because of shock or expectation, but its time is in the order of held breath, not something sustained for long. We do not wait for voices; our lives are more like navigation through voices. In this regard they are like the weather and not the subjects whose feet are on the ground. The nautical origin of the offing links it to Ingold’s proposition of meteorological flux, and the immanence of the sky. The object voice emerges in many different notions of space investigated here. Vocal disposition in Trialogue is informed by Bakhtin’s ‘chronotope’, Smith’s ‘sonotope’, Schaeffer’s ‘reduced listening’, Phillips’ ‘oral stage’, Connor’s ‘vocalic space’, Sperber and Wilson’s ‘cognitive environment’, Goffman’s ‘state of conversation’, Anzieu’s
‘sonorous envelope’, and Noe’s ‘ambient auditory array’. These constructions have different bearings on what kind of voice comes into the offing.

Instead of a spectrogram we could construct a visual representation, like a thermal image, that, instead of representing heat with different colours, shows the hot spots of where we think the voice is located. Speech is located in the speech bubble but the voice is located elsewhere is space. If we think of Dolar’s idea that ‘aesthetic pleasure obfuscates the object voice’ (2006:4) then an artist does not position a voice as an aesthetic decision, rather the position of the voice conforms to genre and context. The position of the voice in time and space connects phenomenology to a physical grammar. If we compare Raw Materials or Days, both by Nauman, we are also free to walk amongst the sound of voices but these voices do not have the structuring of a choir.

The space is empty until the screen fills with image and space, and is vitalised by voices whose reproduction never completely falls into noise, if received by an audience. Voices can be mundane and are therefore ignored, but we can only ignore them for a reason, if our attention is more urgently focused. The voice sits between environmental object and self in particular ways in Trialogue. Trialogue works at the site of that focus, from ‘the moment you can hear a pin drop’ to ‘ears as self protective fortresses’ (Bloom, 2007:127), where figure and ground swap between overhearing, semantic overloading and straining to hear. Phenomenology, Relevance Theory and speech genres join in explaining how the voice communicates in its rich layerings of sonorous envelopes and palimpsests of media screens where material, object and social construction all play parts.

---

107 Pasnau (1999) discusses the incoherence of sound as both surrounding medium and object of hearing – which are found in vocalic space and vocalic body.
Conclusion

This research has investigated what my voice does, and other voices do, in Fine Art practice. By using my recorded voice, and other voices, combined with moving images in art practice, I have investigated the ontology of voice as a connection between genre and disposition. To answer how the voice registers as a material and object requires more than one theory, and I have used phenomenological and pragmatic theories. Sound and linguistic meaning join in the voice, but in Fine Art practice the object and material voice invoke other theories and approaches. The inside out relationship between voice and subject relates to the voice becoming an unobtainable object of desire for psychoanalytical theory, yet the voice is tangible in our experience. The meaning of a voice is structured by its listener, in a relationship with the communicative conventions of a culture, which is critiqued by Fine Art practice. The context of listening links voice and meaning creating gestalt activity where voices can be found inscribed in noises such as wind. The simultaneous layering of vocal meaning privileges language as it passes from a physiological action to acoustic phenomena. But as Ingold (2007a:6) observes we hear through the sound of speech to the silence of language, or we look into a speaker’s words as if their self is expressed in their speech as a point in space. Such phenomena emphasise the rift between the quantifiable aspects of the voice as physiological and anatomical object before the psychology and phenomenology of its perception. Voice connects language to speech and the influence of writing and printing has changed how we perceive the sound of the voice.

The object and material voice is emphasised by reproduction as both part of a subject, and part of the material world. Direct vocal address is the mix of object, subject, disposition and genre. We cannot drill down to the most direct aspect of the voice, but are always part of a relational process,\textsuperscript{108} as when a noise from the viscera of the body is crafted by the articulators (tongue, teeth and lips) as it leaves the body. It is known that context is important for the creation of meaning and this thesis has proposed the phenomenology of Idhe, and Merleau Ponty, Relevance Theory and Speech Genre to underpin the voice in Fine Art practice.

\textsuperscript{108} As Fales (2002) shows, this even applies to timbre.
To describe the time and space of the voice in Fine Art practice I have adopted the figure of palimpsest that, although a still picture plane, hints at the evanescence of what is on its surface.

The voice can be experienced in different discourses, and not just from the perspective of a singer or actor, but as a quotidian object experienced across a number of genres; and, in Trialogue, as a sonic material that investigates meaning in specific utterances. The materiality and objectness of voices is something particular in the discourse of Fine Art generally, and a current issue for contemporary practice. I have assessed artworks by combining theories of speech genres and the phenomenology of voice with spaces and places where artists place utterance.

The use of voice within my own practice generated my interest in this subject, giving rise to these research questions: about self, direct address, reproduction, image, exhibition, and object voice. I have identified these categories in order to analyse voices heard in contemporary art practice. The way voices sound, and are used, change in history; we know this by listening to Florence Nightingale, with whom we opened this field of enquiry about recorded voice.

It has been said that life is one damn thing after another, and it is this everyday approach that I have brought to the voice in my own practice. Therefore, the voices I have discussed are less scripted or scored than those produced by formal methods of acting and singing. Fine Art combines determinate and indeterminate vocal factors that are brought into new balance in emerging genres. The voice is always acousmatic, from its easy relationship with images as voice-over to our synching of voice to lips, a proxy for our inability to look into our bodies at vocal chords or lungs. The slip or mis-registration between vision and sound creates issues for the limit and meaning of form in modernism and post-modernism. A voice on the radio becomes familiar when its owner’s face is unfamiliar. Similarly, we allow sound to stand in for objects. However, I have argued that the voice is an object in that it is a part object. In this context I have queried what kind of thing the voice is, and to do this I have stressed that the voice is always multiple in its agency, without a single condition. This can also be expressed by the relationship between primary and secondary speech genres.

---

109 ‘We hear a door shut in the house and never hear acoustical sensations or even mere sounds. In order to hear a bare sound we have to listen away from things, divert our ear away from them, i.e. listen abstractly.’ Heidegger, Martin (1971:26). Poetry, Language, Thought New York: Harper and Row.
as described by Bakhtin. In art the primary genre can become the secondary; abjection, noise, whispering are not only a primary materiality behind a linguistic meaning delivered by a voice.

I believe that by examining the voice in examples of Fine Art practice I am fleshing out how the voice can be thought of as an object and material in communication. This is how I account for my voice being an aspect of myself as well as a material with a sculptural potential in my practice. My contribution to knowledge is the definition of the actual make-up of partial vocal objects in artworks; how this psychoanalytical construct connects noise, self and language as disposition. I have also used a combination of theories to describe the fields of space, screens, discourses and places to explain the function of voices in Fine Art practice. The evolution of media and hybridity of artistic practice make the voice increasingly present as a type of object in art practices. This project has researched the contexts of genre, milieu, time and space for voices.

We are familiar with media structures that give rise to talking heads, voice-overs, narrative dialogue, or background crowd noises. Listening is categorising, and by intervening in these existing contexts, Fine Art models new practice. Our environment is infected by voices, the recorded voice is something we turn on rather than produce, yet telephonic tele-presence has a foot in both live and recorded sound.

The voice registers as a material and object in the field of Fine Art practice when it is reproduced, and brought into focus in that context. Voices are connected to, and driven by, language yet are received into the spaces of art and the world outside of language. The familiarity of voices is questioned by artistic practice, and I have manipulated the voice as a visual practitioner. It is not only when listening to singing that focus falls upon the sonic form of the voice. In Trialogue, and the case studies described in this thesis, the voices are sonic forms that come with underlying structures other than, or as well as, language. The meaning and power of the voice lies in its disposition. The voice is duplicitous in its meanings and physical presence, filling a whole space or emanating from a point in space.

Fine Art practice actualises the vocalic space described by Connor and uses the remainder that Dolar describes where the sonic materials of speech acts positively lodge
in vocalic space. Artists use voices and test out the relationship between spaces and objects, and this is the intention of Trialogue and the other artworks I have discussed.

I have related the voice to the realm of the visual, for in space we experience one thing in front of another, and this is caused by the orientation of our body and gaze that lead to a flattening on to a screen in representation. A shadow on a surface has the strange quality of being in front of our gaze but also emanates from an object elsewhere, like the effect of a disembodied voice. Although invisible the voice is strongly linked to our embodied experience of space. A voice can feel slight, as material and metaphor, as I suggest in the title of Just Talk, but it is also the seat of subjectivity. Where Lacan presents the voice as a leftover of a process of signification, the voice, in Fine Art practice, can have a positive presence. The most directly subjective vocal attributes are extra-linguistic and can be on the threshold of the unvoiced and silence. All sound falls into the categorical: my argument has been that abstract sound is always related to a disposition and cultural structure. The voice is never only a sound, but a process of communication and representation. Even as a cry or pure note, the voice is already activating a circuit of communication more powerful than the linguistic content it imparts. My conclusion is that artists revitalise the extra-linguistic sound of voice more usually effaced by communication in a more apparent relationship with context.

I have considered the experience of voices in artworks and placed them in different categories, from sonic material to partial object. These categories are activated by how we listen. The argument revolves around whether a voice can exist without genre as a pure type of sound. The work operates in this relationship. The specificity of materials and places are not wholly hidden by communication, and this is borne out in contemporary academic discourse and art practice. The cultural constructions found in Bakhtin’s ideas would place the minimal condition of voice in dialogue, rather than a reduction to sound. In Raw Materials (Nauman, 2004) the voices are semantically overloaded by their disposition. Yet the work structures these voices rather than creating a collapse into cacophony. There is difficulty in reducing the sound of a voice in a modernist or phenomenological way. Bakhtin (1986:61) says a single word utterance has a speech genre. Cornelia Fales (2002) posits timbre as always being relational. If we apply John Cage’s suggestion to voices, that sounds should be

---

110 A voice is rarely a pure note, but rather a set of resonant frequencies caused by the irregular sounding cavities of the body, and seen as formant frequencies in a spectrogram.
themselves, we find a multiplicity of vocal meaning. The voice is the least abstract of sounds, in partnership with language, while imparting the abstract phonemes of language.

However, I have described how sonic materiality is not jettisoned but is used positively by artists in the process of communication with voices in Fine Art practice. The voice puts us in a process in which another subject is trying to communicate, as explained by Relevance Theory. To use voice is to interfere or act within the process of communication. Utterance is at the cusp of uncertain physical and conceptual turbulence but at the same time the vibrant potential of the voice has evolved through history. Sperber and Wilson, and Bloom, express surprise at how efficiently the voice succeeds, given the environmental and physical barriers that it encounters. This suggests a constant, and highly evolved process of focus and refocusing of sound and other perception. The efficiency of our evolved body in its environment, augmented by technology, can lead to the acceptance of a semiotic model of messages being delivered with no attention to the noise of the system. However, *Trialogue* emphasises the materiality of the object voice; a breath from inside, that at the same time goes outside containing the subjectivity of a speaker at the same time as a plethora of human history and culture.

To differentiate how artists use voice I have explained the need for a combination of theory. Real and representational space is infused with voices that I have described with the figure of *palimpsest* that is accompanied by the different theories. *Palimpsest* and *offing* receive theoretical constructions found in my research such as: ‘vocalic space’ (Connor, 2000:12), ‘ambient auditory array’ (Noe, 2004:161), ‘aural stage’ (Phillips 1998:43) and ‘vocotope’, expanded from Jacob Smith’s ‘sonotope’ (2008:245) amongst others. A palimpsest is neither a blank sheet of paper nor finished page of text, instead its surface retains the trace of previous marks, like the scrunched paper surface in *The Jazz Singer*, and I have related this to the temporal nature of voices. The palimpsest does not normally move but records the evidence of previous markings. A visual work of art, such as a painting can be seen as a palimpsest, a record of its own making, whereas voices expire in the moment and process of utterance, and are invisible. This is more like the surface of a screen in moving image work. The surface of a palimpsest displays its own presence like the emphatic quality of a voice, and also shares the directness of frontal address. The process of erasure and layering that is expunged from
the finished visual work is always present, but effaced in our listening to voices. I have used this term to describe how the voice has different layers.

Palimpsest also stands in as a temporal equivalent of figure and ground. Figure and ground are determined by visual perception and become less stable for the temporal flux of sound where a voice can be both figure and ground. Connor describes vocalic space (2000:12) thus: ‘the voice takes place in space, because the voice is space’. Such vocalic space is created by Nauman, Cardiff, Rosler, and Bourn in their practice. I place my own and other voices in a gallery space in Trialogue to loosen their relationship with the self so that they may become objects and materials of artistic practice. My vocal address in Filling Space with Rocks and Forget Me allows a consideration of the relative constructions and limits of direct address. Some direct address, such as the recorded instructional address, is about the easy relay of information, a different construction to that which runs deep into the construction of self. A conflict of the two is presented in the character of HAL, as mentioned in the prelude.

Voice in Fine Art practice brings about philosophical arguments about presence and speech genre. Fine art since twentieth century modernism has also adopted directness with objects and materials and this has been carried through to the way in which voice is used in moving image work and performance. Fales’ (2002) idea, that timbre is not an absolute value but a relational factor, is important here for the voice in Fine Art practice. Artists navigate the limit of direct voice. There is a process of listening, where the voice is recognised and importance afforded, into which artists intercede. The amplitude of a scream by Abramovic, the interrogation of ideology by Rosler, or the raw bodily timbre, unbridled of language by Barber, have been variously explored in my case studies. When we hear a voice we differentiate it from a background and assign it a body. This sound becomes relevant to the listener, and this is what is described by Relevance Theory and Ihde’s phenomenology of voice.

Trialogue has been made in relation to what I identify as the use of object voice in the case studies I have cited, and adds critical and practical mass to this use. In Raw Materials, Bruce Nauman presents voices as raw material that also bring their own sonorous envelope or any combination of the theoretical constructions mentioned
above. By emphasising materiality verbatim techniques also intercede within realism by questioning the authenticity and hierarchy of recorded voices.

Bruce Nauman suggested that:

Misunderstanding is the basis for how a lot of us relate to each other – trying to figure out what the misunderstanding means. Linguists and therapists use misunderstanding as material to work with.  

This quote encapsulates how an audience listens, where what we already know about voices becomes something new in the pressure between the pre-categorical and what it goes on to mean.

The voice is also objectified by reproduction where it can be repeated and retain what is otherwise lost in the present moment of utterance. The lack of a third dimension in the image is compensated by the spatial quality of the object voice whose actual position in space is ambiguous. What we see is in front of us and what we hear surrounds us, but as a partial object the voice intercedes between two and three dimensions. This ambiguity leads to the screen being energised in a direct relationship with voice. Voice and image are also joined by rhythm, which is part of our embodied perceptual experience. Listening to the voice requires time and space, both of which art practice creates in new measure: while the orderings of phonology, and the narrowness of vocal wavelengths allow voice to communicate with language, art strays into the unused parts of available spectra.

My method has been to account for specific artworks by using phenomenological and pragmatic theories for the voice as object and material. While the voice is a fundamental aspect of humankind and culture, voice is less of an unwavering constant like a fingerprint: its sonic print must always be reconstituted in the duration of listening. The voice is a constant flux at the nexus of many factors, such as meaning, sound and place. We think of the voice as an absolute of subjectivity because we can only live in the present moment. However, artistic practices that have developed with the invention and development of reproductive technology reveal the construction of voices. The voice is

---

112 For reproduction, Benjamin noted: ‘film captures the image at the speed of an actor’s speech’ (1992: 213).
something of an uncharted territory in Fine Art practice, where I have argued it has an even less stable phenomenology, requiring multiple theoretical approaches. I have argued that the voice is a part object whose reproduction has a grammar that is connected to what we see.

My research has investigated how the object and material qualities of voices in Fine Art practice create meaning at the same time as being taken as a given of self and authenticity. It has proved important to state how technology does not impose an absolute boundary between live and recorded, original and copy.

The reproduced voice becomes integral to how sound is recycled from archives as the raw material of art practice. This would be a further subject for other researchers, developed out of this thesis. Appropriation receives renewed historicity as reproductive technology and its cultural consequences mutate and this involves object and material voices. How constant does the voice remain across history? The measurements of the physical world were evolved from the human body, but are now derived from the constants of the physical world in metric measurement. The voice is constant only as the framing of a sonic material between object and subject. As Bennett (2010:4) says, it is important to remember that ‘we are also nonhuman and that things, too, are vital players in the world’.

As my research has progressed there has been an increased or resurgent interest in materiality\(^\text{113}\) between Fine Art and material culture and this impacts on the object and material voice. The continued development of verbatim techniques brings new ways of thinking about materiality and this is also a potential area for further research. The prevalence of digital technology has caused an interest in materiality as something lost to the analogue. Yet, while we are embodied beings in the world, materiality is as present as ever and this is manifest in the intangible materiality of voices.

The voice is a constant in our experience of the world between language and material, yet paradoxically never itself constant as a flux of materials in time. By listening we focus on some factors of the voice at the expense of others. In this sense the voice is a fulcrum of perceptual balance, either by intention or accident, but one always framed by

\[^{113}\text{Materiality has a contemporary urgency for art and theory found in the writing of Jane Bennett and Object Orientated Ontology, and the video practice of Ed Atkins.}\]
discourse and context, as described in this thesis for the context of Fine Art practice. As an object the voice does not only stand in for the subject, it also leaves the body as a cultural object evolved by human history, and this is emphasised by reproduction. Direct vocal address must be regarded as a nexus in a multiple structure where directness relates to concepts such as milieu and genre I have opened up in this thesis. We hear the direct qualities of Florence Nightingale’s voice in a recording but experience both the distance and the loss of connection with the public sphere or sonorous envelope of her milieu. The voices I have investigated in other artists’ work and in *Trialogue* are direct modes of address with object and material voices whose performance and rhetoric provide their own discourses and theoretical questions for Fine Art practice.
Bibliography


Albera, Francois and Tortajada, Maria (eds) (2010). *Cinema Beyond Film: Media epistemology in the modern era.* Amsterdam: Amsterdam University Press.


Press.


Works Cited

Abramovic, Marina. *Freeing the Voice* (1975) [Single channel video].
Acconci, Vito. *Face Off* (1973) [Single channel video].
Atherton, Kevin. *In Two Minds* (1978) [Two channel video installation].
Barber, George. *India Shouting Match* (2010) [Single channel video].
Cardiff, Janet and Burres-Miller, George.
   *The Missing Voice: Case Study B* (1999) [Audio walk].
   *Alter Bahnhof Video Walk* (2012) [Audio/visual walk].
   *The Paradise Institute* (2001) [Mixed media installation].
Cardiff, Janet.*The Forty Part Motet* (2001) [Audio installation].
Couzins, Richard.
   *Just Talk* (2011) [Single channel video].
   *The Jazz Singer* (2012) [Single channel video].
   *Filling Space with Rocks* (2012) [Single channel video].
   *Forget Me* (2011) [Single channel video].
Critchley, David. *Pieces I Never Did* (1978) [Two channel video installation].
Lucier, Alvin. *I Am Sitting In A Room* (1970) [audio].
Nauman, Bruce.

Days (2009) [Audio installation].

Get Out of My Mind, Get out of This Room (1968) [audio installation].

Torture (1987) [Two video installation].

Oshima, Nagisa. Death by Hanging (1968) [35mm Film]. Japan: Art Theatre Guild Shochiku.

Philipsz, Susan.


Lowlands (2010) [Audio installation].

Richter, Gerhard. Table (1962) [Oil paint on canvas].

Rosler, Martha. Semiotics of the Kitchen (1975) [Single channel video].


Stidwothy, Imogen. ( ) (2012) [Multi-channel video installation].

Smith, G. A. (dir) The Miller and The Sweep (1897) [35mm film]. London: BFI.

Wardill, Emily. Sea Oak (2008) [16mm film installation].

Wearing, Gillian. Two into One (1997) [Single channel video].