

Department of Media and Communications, Goldsmiths, University of London

IMPORTANT

1. This template must be used for the MA DISSERTATION
2. The Dissertation must be submitted by the deadline

Late submissions will not be marked unless there are acceptable mitigating circumstances according to College procedures

Insert your personal details below

Student Number	33474454
College Email	dtuff002@gold.ac.uk
Degree Programme	MA Film and Screen Studies
Word Count; Incl. notes, excl. reference list & appendices (NB: Appendices are <u>not</u> assessed)	14,980
Deadline	25 th August 2017

Dissertation Title

Type in box below

Broken Silence:

How the Acoustic Language of Sound and the Voice Shaped and Transformed the Visual Language of Cinema

Definition of Plagiarism

Plagiarism is an attempt (deliberate or inadvertent) to gain advantage by the representation of another person's work, without acknowledgement of the source, as the student's own for the purposes of satisfying formal assessment requirements.

Recognized forms of plagiarism include

1. the use in a student's own work of more than a single phrase from another person's work without the use of quotation marks and acknowledgement of the source;
2. the summarizing of another person's work by simply changing a few words or altering the order of presentation, without acknowledgement;
3. the use of ideas or intellectual data of another person without acknowledgement of the source, or the submission or presentation of work as if it were the student's own, which are substantially the ideas or intellectual data of another person;
4. copying the work of another person;
5. the submission of work, as if it were the student's own, which has been obtained from the internet or any other form of information technology;
6. the submission of coursework making significant use of unattributed digital images such as graphs, tables, photographs, etc. taken from books/articles, the internet or from the work of another person;
7. the submission of a piece of work which has previously been assessed for a different award or module or at a different institution as if it were new work;
8. a student who allows or is involved in allowing, either knowingly or unknowingly, another student to copy another's work including physical or digital images would be deemed to be guilty of plagiarism.
9. If plagiarism is suspected students will be required to supply an electronic copy of the work in question so that it may be subjected to electronic plagiarism detection testing. Therefore students are required to keep work electronically until after they receive their results as electronic detection may be part of the investigative process.

Source: Assessment Handbook 15f.

In submitting this work I

- 1) Confirm I have read and understood the regulations relating to plagiarism and academic misconduct when I confirmed my **assessment confirmation form** online.
- 2) I agree to have my dissertation entered in my program's Dissertation Archive if it is selected.

SIGNATURE: Darrell Tuffs

DATE: 30/05/17

Broken Silence:

How the Acoustic Language of Sound and the Voice Shaped
and Transformed the Visual Language of Cinema

Submitted by Darrell Tuffs 33474454 in partial requirement for the degree of MA Film and
Screen Studies, Goldsmiths, University of London.

CONTENTS

ix. The Death of Silence	1
1. Literature Review	6
1.1 Classic Sound Theory	7
1.2 Modern Sound Theory	18
2. ‘M’ and the Silent Sound Film	25
2.1 Noise and Cinematic Language	26
2.2 Acoustic Landscapes: The Arrival of Aural Space	36
3. ‘Dr. Mabuse’ and the Voice in Cinema	46
3.1 Silent Voices: ‘Deaf’ Cinema and the Voice	47
3.2 Voice as Meaning and Presence	51
3.3 The Voice and Devices of Power	60
4. Cinema Transformed	67
5. Bibliography	75
6. List of Illustrations	79

ix. The Death of Silence

Writing in his work, 'Theory of the Film', Bela Balazs recalls his first reactions to sound cinema, encapsulating also the mood many theorist and filmmakers shared upon the arrival of sound to the cinema in the late 1920s.

“Two decades ago I wrote: ... The silent film was just about to develop a psychological subtlety and creative power almost unrivalled in any other art. Then the invention of the sound film came down on it like a landslide ... The whole rich culture of the silent film (is) now in danger.” (Balazs, 1952, p.195-196)

The concerns of Balazs were widespread among theorists, including staunch critics such as Rudolf Arnheim, who stressed that, “Not only does speech limit the motion picture to an art of dramatic portraiture, it also interferes with the expression of the image.” (Arnheim, 1957, p.228) The two writers speak similarly to the ‘rich culture’ and ‘expressive image’ of the silent era as elements that made the art form a unique experience; a silent world of expression depended on moving visual images rather than the static world of painting or the staged literature one of the theatre. In

these comments, Balazs and Arnheim allude to the vulnerability of images and cinematic language when occupied by sound and speech, a fear that the expressive rhythms and motions of the silent film may fall victim to the conflicting communicational demands made by sound and dialogue. These fears are summarised by critic and filmmaker René Clair, who boldly discredits the 'talking film' in stating that, "the screen has lost more than it has gained... It has conquered the world of voices, but lost the world of dreams." (Clair, 1985, p.95) In a similar fashion, Marguerite Duras claimed that the 'imaginary' within cinema had become "closed-off" during the arrival of sound. "Something about the silents is lost forever. There is something trivial (in) the unavoidable realism of direct dialogue (and) the inevitable trickery it involves." (Duras, 1975, p.9) Duras, like Clair, speaks to an artistic freedom found within silent film, an ability to imagine sound, to dream of its sonic possibilities while removed from the acoustic landscape of our own reality. Speaking to this effect, Michel Chion adds, "If for some people the talkie still seemed vulgar by comparison to the silents, it is because the real voices heard in it came into conflict with the imaginary voices that everyone could dream to their heart's content." (Chion, 1999, p.8) The arrival of sound, for many early critics, cast the dream state and imaginary qualities associated with silent filmic language into jeopardy by providing the spectator with sustainable amounts of reality, thus disrupting the visual language created by images to instead

reflect life, rather than imagine it. Clair concludes this point by describing a film audience he witnessed after leaving a sound film screening.

“they showed no sign of the delightful numbness which used to overcome us after a passage through the silent land of pure images. They talked and laughed, and hummed the tunes they had just heard. *They had not lost their sense of reality.*” (Clair, 1985, p.95)

To create a ‘sound film’ during the early stages of its birth meant changing the process and environment of filmmaking; to create and discover the correct conditions needed for sound to be accurately recorded. “Providing these conditions heavily influenced image recording as well as sound” claims Rick Altman. (Altman, 1980, p.5) And yet, the pioneers of cinema were determined to make sound work, determined to combine artistically the acoustic and visual languages of sound and image to extend the creative possibilities of the cinema. How did the inclusion of ‘noise’ and ‘the voice’ together as sound transform the aesthetic qualities of the visual silent cinema? How does sound inform filmic space, create presence and meaning, or even establish a connection from the film’s spectator and characters to emerging sound technologies as devices of influence and power? Upon the death of silence, in defiance of early critics and technical complications, how did the language of cinema grow and transform from

an art form dependent on visual imagery to one constructed of audio-visual elements working collectively to form a cause of expressive unity?

In this paper, I shall attempt to answer these encompassing questions within two smaller filmic case studies, each directed by the pioneering German film director, Fritz Lang. Lang began his filmmaking career within the silent cinema, directing influential and innovative films such as *Metropolis* (1927) and *Woman in the Moon* (1929), yet what makes Lang essential to this study is his highly successful transition to sound cinema, which he not only found success within, but also helped expand creatively.

In my discussion of Lang's 1931 film *M* as a 'silent sound film', I account for Lang's creative use of sound, examining the film's successful unification of visual and acoustic language. I also detail the ability of sound in the film to suggest the presence and absence of on and off-screen characters, while also establishing a sense of affect to unsettle the spectator through sound's sonic ability to appear as an acoustic signifier whilst concurrently concealing the visual source of its origin. The narrative of *M* is founded upon a murder mystery, which is made interactive for the spectator by offering sounds as traces of the unseen. The film is made unsettling by the murderer's ability, via way of sound, to infiltrate the onscreen space of the film while remaining undetected visually. I explore Lang's use of an acoustic landscape in *M*, using sound to spatially connect

two or more visual spaces with aim of uniting the temporality of their occurrence by deploying sound as a medium existing only in time and space.

In *The Testament of Dr. Mabuse* (1933), Lang provides the sound sequel for a cinematic world created in silent cinema. Over the ‘Mabuse’ series, I examine the first, *Dr. Mabuse the Gambler* (1922), not as a ‘silent’ film, but as a work of ‘deaf’ cinema, in which voices are present and audible within the world of the narrative, yet never directly heard by the spectator. I go on to study the voice in *The Testament of Dr. Mabuse*, exploring the persona of Dr. Mabuse as a mute character able to dislocate voice from body, and image from presence. Mabuse never directly speaks onscreen, yet his voice, through a range of visual and technological methods, haunts the edges of the film’s frames to deceive and defy not only the film’s characters, but also its spectators. Finally, I detail the film’s use of sound apparatus as devices of power, accounting for their hypnotic qualities when compared to the ideological audio-visual conceit of cinema itself.

I will begin this work by exploring the work of both classic and modern sound theorists by studying the theoretical arguments of early and late literature on film sound.

PART ONE

Literature Review

1.1 Classical Sound Theory

The early theorists of film are theoretically united by their inherent suspicion of synchronised film sound; a subject to which they cautiously approach with varying degrees of uncertainty. In 1928, Sergei Eisenstein, Vsevolod Pudovkin, and Grigori Aleksandrov released a joint statement discussing the emerging popularity of film sound. Although not directly opposed to sound in general, the statement attacks early sound films for advancing the invention of film sound in an “incorrect direction”. The statement fears “a misconception of the potentialities within this new technical discovery (that) may not only hinder the development and perfection of the cinema as an art but also threatens to destroy all its present formal achievements.” (Eisenstein, 1977, p.257) The statement is somewhat typical for classical sound theory, pronouncing hesitation towards sound as threatening the language of cinema as a visual form of communication in likening its emerging use to “photographed performances of a theatrical sort.” (Eisenstein, 1977, p.258)

The Eisenstein controlled statement attempts to protect the aesthetic achievements of montage editing, identifying the visual technique’s “powerfully effective strength” as far outweighing the “simple curiosity” of early film sound. For Eisenstein, the cinematic montage is composed of single material elements that create meaning within a constructed whole.

Restoring the natural sound to any given object reinstates its independent autonomy, thus destroying its potential use as a “montage element”. The statement proclaims that “only a contrapuntal use of sound in relation to the visual montage piece will afford a new potentiality of montage development and perfection.” (Eisenstein, 1977, p.258) Eisenstein advocates “the creation of an orchestral counterpoint” that considers sound to be an element within cinematic montage. (Eisenstein, 1977, p.258) Sound deployed in such a way will “introduce new means of enormous power to the expression and solution of the most complicated tasks that now oppress us.” (Eisenstein, 1977, p.259) This new expression is found within sound acting as a counterpoint to visual images; rather than dictating the rhythm of montage, sound must heighten and intensify its communicative tempo.

While Eisenstein emboldens contrapuntal uses of sound as extending the creative possibilities of montage editing, his statement co-signer Vsevolod Pudovkin later drew upon a somewhat different conclusion. In 1929, Pudovkin spoke of ‘asynchronism’ as a creative possibility for film sound. He believes “the first function of sound is to augment the potential expressiveness of the film’s content.” (Pudovkin, 1954, p.156) Yet, arguing against a more naturalistic use of sound, Pudovkin goes on to observe that a “deeper insight into the content of the film cannot be given

to the spectator simply by adding an accompaniment of naturalistic sound: we must do something more.” (Pudovkin, 1954, p.156) This ‘something more’ included the asynchronism of film sound; “the development of the image and the sound strip each along a separate rhythmic course.” (Pudovkin, 1954, p.156) This distribution of sound would see the visual and the aural “connected as the result of the interplay of action” rather than being interwoven by “naturalistic imitation”. (Pudovkin, 1954, p.156) Pudovkin provides a filmic example of this interplay by describing a scene expressing the inner content of a ‘town-bred’ man standing in the middle of an empty desert.

“In silent film we should have had to cut in a shot of the town: now in sound film we can carry town-associated sounds into the desert and edit them there in place of the natural desert sounds.” (Pudovkin, 1954, p.157)

Whereas Eisenstein promotes a contrapuntal use of sound to counteract visual material used within montage editing, Pudovkin endorses the use of associative aural and visual elements working asynchronously to further express the creative rhythm and tempo of both the film’s created world and the character’s objective perception of that world simultaneously.

Another endorsement of asynchronous sound comes with director Alberto Cavalcanti in 1939. Cavalcanti's theories of film sound are similar to those of Pudovkin, yet he grounds his central observations within the distinctive qualities that separate both sound and image. For Cavalcanti, "while the picture is the medium of statement, the sound is the medium of suggestion." (Cavalcanti, 1939, p.109) The suggestive quality of sounds further extends the expressive nature of silent films by preventing cinema from becoming "primarily a medium of statement". (Cavalcanti, 1939, p.110) While examples of suggestive visuals and statement sounds are given by Cavalcanti, his central argument rests on the fact that images primarily state the real signified presence of objects, while unidentified noises "do not inevitably suggest what made them", thus "can be used incognito." (Cavalcanti, 1939, p.108)

Cavalcanti's willingness to embrace film sound comes with the notion that, even within the early silent period "some sort of sound accompaniment" was externally tied to visual images; "In other words, the silent film never existed." (Cavalcanti, 1939, p.98) Yet, he remains critically distance from the overuse of film speech, warning that "Long stretches of dialogue inevitably cancel movement and visual variety... Emphasis and emotional effect must of necessity be left to the care of the visuals." (Cavalcanti, 1939, p.103) Thus, Cavalcanti concludes that film

speech should be “economical, and balanced with other elements of the film” in order to keep the visuals moving and creatively varied. (Cavalcanti, 1939, p.104)

Whereas Cavalcanti argued for an economically balanced use of film speech, early theorist Rudolf Arnheim remained critical of film dialogue, claiming that its use, “narrows the world of the film”. (Arnheim, 1957, p.226) As a formalist, Arnheim’s critical theories run opposed to the idea of film reproducing reality, thus speech was feared as a tool for decreasing the space between reality and film as both a language and an art form. Arnheim accuses speech of limiting “the motion picture to an art of dramatic portraiture”. (Arnheim, 1957, p.228) This is because the dramatic use of dialogue “limits the action to the human performer”; rather than representing “the world animated by man”, dialogue instead presents “man set off against his world.” (Arnheim, 1957, p.227) Arnheim stresses the importance of images to confirm the existence of man within natural settings, arguing that the dominance of speech could only be achieved by the performers within the film, thus pushing “all other things (into) the background.” (Arnheim, 1957, p.227)

Arnheim also attacks speech for limiting the expressiveness of visual images. He declares that, “It was precisely the absence of speech that made the silent film develop a style of its own, capable of condensing the

dramatic situation.” (Arnheim, 1957, p.229) The talking film limited the movements of bodies in action, of communication via gestures and facial expressions, instead reducing this movement to the restricted “monotonous motions of the mouth”. (Arnheim, 1957, p.228) Like many other early theorists, Arnheim ultimately compares the space of the talking film to that of the theater stage, blaming film speech for “replacing the visually fruitful image of man in action with the sterile one of the man who talks.” (Arnheim, 1957, p.229)

René Clair also opposed the overuse of speech within cinema. As a filmmaker, Clair’s frustrations towards the talking film are industrial. He accuses the industrialists of American cinema for financing talking films above all other works of cinema, stating that “so many thousands of million dollars have been invested” in order to guarantee the success of the talking film, thus rendering the production of silent cinema a close impossibility. (Clair, 1985, p.92) Clair upholds the sound film; “on which the last hopes of the advocates of the silent film are pinned.” (Clair, 1985, p.92) Clair argues that the asynchronous use of sound could spare the cinema from falling into the “stereotyped patterns” of the talking film, leading to tiresome and theatrical films lacking in creativity. (Clair, 1985, p.93) An example from the American film *Broadway Melody* (1929) is given by Clair in order to highlight a successful use of asynchronous

sound, one within which, “an appropriate form” is found, one that reflects “neither theater nor cinema, but something altogether new.” (Clair, 1985, p.93)

“We hear the noise of a car door being slammed and a car driving off while we are shown Bessie Love’s anguished face watching from a window the departure which we do not see. This short scene in which the whole effect is concentrated on the actress’s face, and which the silent cinema would have had to break up in several visual fragments, owes its excellence to the ‘unity of place’ achieved through sound.” (Clair, 1985, p.94)

Clair’s example is similar to that of Pudovkin’s desert scene; sound is used in such a way as to reflect the emotional state of Bessie Love’s character without the need for speech, intertitles, or a cut. The visual language of film is heightened by the suggestive use of sound, supplying its filmic language with a greater depth of complexity, thus enhancing the unique poetic and artistic qualities that cinema may have lost after the emergence of speech.

Writing in his 1945 book, *Theory of the Film*, Bela Balazs exclaims frustration towards the uncreative and inactive first uses of sound within

early sound cinema. He points out sound's ability to "reveal to us our acoustic environment, the acoustic landscape in which we all live". With sound, so too comes the powerful and emotional voice of the world itself "from the roar of machinery to the gentle patter of autumn rain on a windowpane." (Balazs, 1952, p.197) Balazs bestows upon the sound film a more meaningful and complex task of "redeem(ing) us from the chaos of shapeless noise by accepting it as expression, as significance, as meaning." The central aim of these films must be to project noises as "intimate voices ... speak(ing) separately in vocal, acoustic close-ups". (Balazs, 1952, p.198) These acoustic close-ups supersede experiences of reality because they "make us perceive sounds which are included in the accustomed noise of day-to-day life, but which we never hear as individual sounds because they are drowned in the general din." (Balazs, 1952, p.210)

Balazs advocates the use of 'sound-montage' in purposefully ordering the noises of the acoustic world to create new forms of visual and aural art. Balazs's theory incorporates the ability of cinematic sounds to re-educate the ear, establishing dramatic effects through their suggestive nature. "We are not accustomed (to) draw conclusions about visual things from sounds we hear. This defective education of our hearing can be used for many surprising effects in the sound film." (Balazs, 1952, p.212) The hissing of an unseen object within a darkened room could be that of a

dangerous snake, yet by applying visuals to the already established layer of sound, we may discover the noise is nothing more than the hiss of a boiling hot kettle. (Balazs, 1952, p.212) Through these effects, Balazs favours the creative use of asynchronous techniques. He describes how noise alone holds a separate spatial awareness to that of sound and images woven together; “the sound wave is different if we see (an object’s) movement.” (Balazs, 1952, p.205) In particular, Balazs highlights the mysterious quality of hearing a sound before the visuals are able to establish its source, concluding that “this handling of picture and sound provides rich opportunities for effects of tension and surprise.” (Balazs, 1952, p.209)

Writing much later in 1960, Siegfried Kracauer represents the most historically informed view of sound within classical film theory, yet Kracauer’s hopes for the sound film were similar to those of earlier theorists in attempting to establish a cinematic use for sound, one that would enrich and enhance film’s images, rather than restrict them. The way Kracauer propose films do this is to ‘de-emphasise’ film speech, to decrease the importance of the single words and sentences themselves in order to “play down dialogue with a view to reinstating the visuals.” (Kracauer, 1997, p.106) Kracauer views a cinematic use of dialogue as “verbal statements which grow out of the flow of pictorial communications

instead of determining their course.” (Kracauer, 1997, p.106) This may also include ‘undermined’ speech, which works to represent the inclusion of dialogue without drawing focus away from visual images. Kracauer writes about the opening scene of Charlie Chaplin’s *City Lights* (1931) as an appropriate example.

“In the opening sequence of *City Lights* the orators celebrating the unveiling of a statue utter inarticulate sounds with the grandiloquent intonations required for the occasion. The sequence not only makes fun of the inanity of ceremonious speeches but effectively forestalls their absorption, thus inviting the audience all the more intensely to look at the pictures.” (Kracauer, 1997, p.107-108)

With this, Chaplin shifts “emphasis from the meaning of speech to its material qualities.” (Kracauer, 1997, p.109) This use of speech informs the visuals, placing cinematic importance on the image of a character talking rather than that of the words they are speaking.

In closing the classical sound theory period, Kracauer appropriately calls back to both Eisenstein and Pudovkin, accusing them of failing to recognise the differences between using counterpoint and asynchronous sound as filmic techniques that may serve any purpose, rather than methods with independent significance. Kracauer endorses these sound

methods as tools with which the director may choose to enhance his or her film, yet he argues against the independent use of one or another. Eisenstein and Pudovkin's "obsession with the montage principle" is blamed for their shortcomings. (Kracauer, 1997, p.115) Kracauer adds the cinematic tool of de-emphasising speech to theoretical discussion for its ability to "(stir) the visuals to become eloquent", before emphasising that both Eisenstein and Pudovkin were correct to encourage contrapuntal uses of sound, yet "they did so for the wrong reasons..." (Kracauer, 1997, p.116)

1.2 Modern Sound Theory

In exploring theories of modern film sound, immediately apparent is a critical shift away from film language, and a move towards the modern film spectator/audience. No longer are central theoretical arguments based upon the contrapuntal or asynchronous uses of sound, but instead treat sound as a given, exploring instead the material and affective qualities of sound with regard to their concern for the film's space, ideology, and spectator. These concerns accumulate appropriately within an interview given by Jean-Marie Straub and Daniele Huillet, in which they discuss 'direct sound'.

Huillet and Straub condemn the Italian cinema for its enthusiasm towards film dubbing. Not only does the effect of dubbing create a false sense of cinematic space and environment, but it crushes imaginative freedom within audiences; "The great silent films gave the viewer the freedom to imagine the sound. A dubbed film doesn't even do that." (Huillet & Straub, 1985, p.150) They label the dubbing of films an ideology, attacking the technique as "the cinema of lies, mental laziness, and violence, because it gives no space to the viewer and makes him still more deaf and insensitive." (Huillet & Straub, 1985, p.150) This ideology also plays into the visual rhythm of dubbed films; 'direct sound' recorded within the same space as a film's visuals limits the manipulative use of editing

techniques since “each image has a sound and you’re forced to respect it.” (Huillet & Straub, 1985, p.152) Thus, direct sound acts as a safeguard against the visual manipulation of truthful into false imagery.

As opposed to classical sound theory, Huillet and Straub treat the deployment of direct sound as a moral and ideological decision as much as a technical one. With respect for the viewer as well as the film itself, Huillet and Straub warn against “transform(ing) a real space into a confused labyrinth.” (Huillet & Straub, 1985, p.152) They blame an international film aesthetic for favouring films with which authentic voices and spatial noises may be removed to suit the particular language and ideological requirements of any given country; “The international aesthetic is an invention and weapon of the bourgeoisie. The popular aesthetic is always a personal aesthetic.” (Huillet & Straub, 1985, p.153)

Approaching film sound from the viewpoint of semiotics, Christian Metz attempts to show that sound, in the form of a perceptual aural object, “is a constructed unity, socially constructed, and also a linguistic unity.” (Metz, 1980, p.31) Metz is also interested by the states within which audiences experience the qualities of sounds as objects in their own right – as ‘aural objects’. Metz deconstructs the sound of an object from the visual signifier of its source, contemplating, “How is it possible that we are capable of recognizing and isolating the sounds of ‘lapping’ on the sound

track of a travelogue ... even when we don't know its source.” (Metz, 1980, p.24) Metz proposes the answer is that, “‘lapping’ exists as an autonomous aural object, the pertinent traits of its acoustic signifier corresponding to those of linguistic signified”. (Metz, 1980, p.24) His observations work against the notion of sound as a suggestive medium, identifying its use to be as much a unit of perceptible signification as images themselves. These conclusions lead Metz to expose sound as, not a characteristic of an object, but as a social construct informed by any given listener’s knowledge of the world itself.

Metz’s theory is grounded upon what he labels the “ideological undermining of the aural dimension”. (Metz, 1980, p.25) To explain, Metz states an example using the identification status of the words ‘rumble’ and ‘thunder’.

“If I perceive a ‘rumble’ without further specification, some mystery or suspense remains ... however, if I perceive ‘thunder’ without giving any attention to its acoustic characteristics, the identification is sufficient.” (Metz, 1980, p.26)

Metz understands sound within a superior ideological structure that prioritizes phenomenological traits such as sight and touch as socially superior to that of the aural. Warning against this structure, he states that,

“Ideologically, the aural source is an object, the sound itself a ‘characteristic’ (but) there is nothing natural in this situation: from a logical point of view, ‘buzzing’ is an object, an acoustic object in the same way as a tulip is a visual object.” (Metz, 1980, p.26-27) Metz’s challenge to this structure attempts to reposition the hierarchical status images held over sounds during the classical sound period, unveiling the constructions to which film sound as a device is inherently defined by social knowledge.

Mary Ann Doane also warns against the ideological constructs that film sound operates within. Doane describes the combination of image and sound within the sound film as a ‘heterogeneous medium’; this heterogeneity must disguise itself as a unified whole in order to survive. One method used within the sound film in achieving this survival is by constructing what Doane labels the ‘fantasmatic body’. Within the silent cinema, “The absent voice reemerges in gestures and the contortions of the face” in order to form an order of expression. (Doane, 1980, p.33) For Doane, this order becomes endangered by the arrival of film sound, thus the ‘fantasmatic body’ acts as a ‘point of identification’ for the spectator; “the body reconstituted by the technology and practices of the cinema”. (Doane, 1980, p.33) Although sound and image are two separate mediums, Doane describes the cinema’s attempts to form an ‘organic’ unity through the ideology of the ‘fantasmatic body’, stating that “Sound

carries with it the potential risk of exposing the material heterogeneity of the medium; attempts to contain that risk surface in the language of the ideology of organic unity.” (Doane, 1980, p.35) Doane states sound synchronization as one technique used by the cinema in order to construct this ideology. Synchronised sound recording for the cinema aims for an authentic ‘presence’ within the aural space constructed by the film’s soundtrack with aim of reducing the material differences between sound and image, “concealing the work of the apparatus, and thus reducing the distance perceived between the object and its representation.” (Doane, 1980, p.35) This authentic spatial presence provides a point of anchorage by which the body is supported; “The fantasmatic visual space which the film constructs is supplemented by techniques designed to spatialize the voice, to localize it, give it depth, and thus lend to the characters the consistency of the real.” (Doane, 1980, p.36)

The ‘voice-off’ affords a distinctive challenge to the sound cinema’s organic unity of space. The notion refers to an unseen space, one constructed by the diegesis of the film itself, thus attempting to establish an unobserved visual space exclusively with the deployment of aural material. Doane marks traditional uses of the voice-off as constituting “a denial of the frame as a limit and an affirmation of the unity and homogeneity of the depicted space.” (Doane, 1980, p.37-38) Thus, not only does the voice-off construct virtual off-screen space referred to by

on-screen characters, but provides a reformation of synchronized sound as organically depicted within a unified world of visual and aural material. Doane outlines three types of space created by any given 'cinematic situation'. (Doane, 1980, p.39)

1. The space of the diegesis (virtual space)
2. The visible space of the screen as a receptor of the image
3. The acoustical space of the theater or auditorium

These three spaces are perceived by the spectator, yet the characters of the film may only give acknowledgment to the first in order to preserve the cinema's unified sense of space. Doane uses these spaces to describe how classical narrative films work to "deny the existence of the last two spaces in order to buttress the credibility of the first". (Doane, 1980, p.40)

In accordance with this spatial structure, uses of the voice-off within cinema work to support and extend the construction of diegetic space, "deepen(ing) the diegesis ... and thus support(ing) the claim that there is a space in the fictional world which the camera does not register. In its own way, it accounts for lost space." (Doane, 1980, p.40)

An alternative view to which Doane offers the voices of cinema comes in the form she labels, "the pleasure of hearing." In a similar matter to cinematic voyeurism, Doane supposes that, "in the fiction film, the use of

synchronous dialogue and the voice-off presuppose a spectator who overhears and, overhearing, is unheard and unseen himself.” (Doane, 1980, p.43) This affective pleasure also forms connections inside the psychoanalytical boundaries of childhood; the audible voice of the mother or father defines a child’s sense of space more sufficiently than that of the visible. The voice “traces the forms of unity and separation between bodies. The mother’s soothing voice, in a particular cultural context, is a major component of the ‘sonorous envelope’ which surrounds the child and is the first model of auditory pleasure.” (Doane, 1980, p.44)

PART TWO

***'M'* and the Silent Sound Film**

2.1 Noise and Cinematic Language

Something rather atypical for a sound film happens about 20 minutes into Fritz Lang's 1931 film *M...* visual action on screen, yet complete and utter silence on the soundtrack. The preparations of a police raid on the streets of Berlin are accompanied by the cold sound of silence; no words, no acoustic layers of the film's world, not even a musical arrangement to complement the visual movement of the images. After 20 minutes of introducing audiences to the transformed world of his sound cinema, Lang abruptly mutes the cinematic image, resulting in a disruptive rhythmic shift that leaves the visuals lacking in texture; a temporal alteration that decreases the visual integrity of Lang's film, yet increases its symbolic significance. Within the shift, Lang draws direct comparison between the sound and silent cinema, using silence to reduce the textural individuality provided by acoustic noise into the emblematic and uncanny visual realm created and enhanced by a lack of aural substance. The police lining the streets in formation are no longer individual bearers of noise who can affect the space of the film with isolated sounds of distinctive physical movements, but are now a collective visual signifier unified by their shared symbolic appearance and mutual silence. Like the Tsarist soldiers of Eisenstein's Odessa steps sequence in *Battleship Potemkin* (1925) the police here are united as a faceless 'group force' in their muteness; an

effect further emphasized by Lang's removal of an already established layer of aural content. A possible reason for the reduction of visual autonomy within this silent sequence relates to the viewer's perception of time and movement, to which Michel Chion claims that "sound endows shots with a temporal linearization... sound *vectorizes* or dramatizes shots, orienting them towards a future, a goal, and creation of a feeling of imminence and expectation." (Chion, 1994, p.13-14)



Images 1 & 2: Lang unifies the police group as a force by removing the autonomy given to the visuals by sound.

Lang disturbs the temporality of the sequence, resulting in an uncanny visual aesthetic that renders time and movement infinite and unsystematic; yet he also removes the added sense of autonomy so far given to his images by the introduction of sound, instead arranging the sequence as blocks of visual material that purely 'state' the symbolic arrival of a police raid without provoking, as Eisenstein labels, "the independence of its meaning". (Eisenstein, 1977, p.258) *M* manages to cinematically ground itself to the symbolic tendencies of the silent cinema,

yet incorporates this influence within the altogether new and revolutionary possibilities of the sound cinema.

Within this scene, as well as many others throughout the film, Lang helps to outline a new type of cinema, one literally definable as 'sound cinema' via its acoustic soundtrack, yet one also sharing many of the same linguistic cues as that of 'silent cinema', particularly due to its use of montage. Noel Carroll clearly states that "*M* is what might be called a silent sound film." (Carroll, 1978, p.16) To expand this, Carroll outlines one sequence of the film possessing two temporally paralleled spaces. In them, the conflicting gangster and government groups in the film separately discuss the topic of how to capture an unidentified serial killer who roams the streets of Berlin while both kidnapping and killing innocent children. Lang assigns similar stylistic visuals between the spaces, including the matched atmosphere of murky cigarette smoke, and the silent bodily movements of introspective pacing. Yet, what is more important is the way he edits by sound "the editing almost elides the two spaces; a criminal could be seen as addressing an official and vice versa", thus the soundtrack, in part, dictates the visual language of cinema and how it is presented onscreen. (Carroll, 1978, p.16) As an off-screen speaker suggests an idea within the government's space, the listeners look outside the frame of the film to listen, thus seeing what the film's audience does not, yet hearing the same aural signifier. This audio-visual

technique is continued into the gangsters' space, we see the shadowed silhouetted figures of the men as they further discuss the issue. The sequence draws ethical parallels between the groups through the matching of visual cues, yet these cues, such as the silhouetted figures and government officials looking off-screen, create new montage possibilities because of the inclusion of sound. Were it not for sound, the only way Lang might have suggested similar lines of conversation would have been to break the sequence apart using intertitles. Here, Lang deploys sound as a montage element, using its presence to enhance and further shape the visual elements of the sequence. As stated by Carroll, "this pejorative equivalence is emphasised through an elaborate set of comparative juxtapositions ... The underlining presupposition of Lang's system is that the nature of film is to reconstitute reality, not to record it." (Carroll, 1978, p.17-23) Lang structures his filmic language to interplay sound and image within a conversation of both symbolic and perceptive rhythms, echoing Vsevolod Pudovkin's demand for a cinema in which "The image may retain the tempo of the world, while the sound strip follows the changing rhythm of the course of man's perceptions, or vice versa." (Pudovkin, 1954, p.158)



Images 3 & 4: Lang deploys sound as an element independent of the films visuals.

Throughout the sequence, Lang deploys the use of counterpoint sound, consummating in an audio-visual language in which sound and image “constitute two parallel tracks, neither dependent on the other.” (Chion, 1994, p.36) In doing so, Lang allows himself the freedom to experiment widely with the cinematic form, including the way he establishes his characters while visualising their mental perceptions. One famous sound element of the film comes in the form of child killer Hans Beckert’s obsessive whistling of Grieg’s ‘In the Hall of the Mountain King’ while attempting to calm his murderous impulses. The tactile sound of the whistling itself is exhibited as an acoustic leitmotif, an invisible signpost that bypasses the visual elements of the onscreen film to affect the physical space of the viewer’s auditory vicinity. This technique is made more effective as a counteracting sound element while the signified sound of the whistling itself disconnects from the visuals, thus forming what Christian Metz may label an ‘aural object’, one aesthetically constructed by the film as an independent filmic object used to provoke a mental image

of the child murderer without his physical presence ever needing to be made clear. In fact, looking at the plot of *M*, this is indeed how the murderer is caught, by a blind man who lacks visual evidence of Hans, yet recognises his physical presence by the acoustic description of his whistling as an aural object independent of visual reference, yet definable by a social understanding. Alberto Cavalcanti describes Hans's whistling in the film, labelling the element most threatening "when you could hear the noise, but could not see the murderer." (Cavalcanti, 1939, p.108) This description speaks to Lang's successful asynchronous approach to sound, to which Noel Carroll consummates with the investigative themes of the film itself. Carroll suggests that "by consistently presenting scenes where the audience must infer off-screen agents from their traces Lang continues to make the spectator's relation to the events of the film analogous to a detective's relation to his clues." (Carroll, 1978, p.18) Hans's whistling, upon his first victim, is provided as an off-screen clue while the presence of his shadow obscures a public wanted poster. On another occasion, the same whistling is heard while Hans buys a balloon for his next attempted victim, his back turned to camera, thus concealing the visual source of the noise. On both occasions, the sound, rather than the visuals, provide clear investigative clues that foreshadow the events of the film, yet also include and invite the spectator into the film's thematic investigation. In summary, Carroll concludes that "the use of sound to

present off-screen traces (induces) an investigatory attitude on the part of the audience –thereby simulating ... the fictional experience of the characters in the viewing experience of the spectators.” (Carrol, 1978, p.18) Similarly, Michel Chion argues that, for much of the film, Hans’s presence is constructed of acoustical elements acting as clues of his manifestation, while visually preserving the anonymity of his character; “even though we hear his voice and his maniacal whistling from the very beginning ... Lang preserves the mystery of the character as long as he can, before ‘de-acousmatizing’ him.” (Chion, 1994, p.72) This process is completed by revealing to the audience Hans’s identity, in which the uncanny and suggestive qualities of his dark and murderous aural persona are extinguished as traces and transformed into visual statements. This deconstructs Hans’s frightening facade, instead presenting him in true image, as a lonely and obsessed man, isolated by the curse of his own convictions.

Throughout *M*, Lang uses sound in a dramaturgical manner, deploying the use of sound when needed to affect the film’s images, rather than overlapping a continual stream of sound that abides to the image’s every manoeuvre. Lang himself claims that the inspiration for such a technique was in part inspired by his own real world perceptions.

“I found, for example, that when I was sitting alone in a sidewalk café, of course I heard the noises from the street, but that when I was immersed in an interesting conversation with a companion, or when I was reading a newspaper that totally captured my interest, my organs of hearing no longer registered these noises.” (Lang, 2003, p.35)

These comments become pertinent when focusing once more on Lang’s use of sound in *M*. During one scene, Hans walks a busy sidewalk while, unknowingly to him, a police inspector investigates his house. Pausing here, it may be noted that the visual space of the street itself is in fact not as busy as projected; the noise of constant car traffic, public transport trams, and footsteps, creates the illusion of a chaotic space, while the visuals of the scene themselves remain relatively static and unpopulated. In wide shot, Hans stops to browse a local shop window; shortly thereafter, a cut takes the film’s viewpoint inside the store, looking out into a medium shot of Hans as he glances. In cutting into Han’s personal space, the noises of the street halt completely, forming an acoustic long take of silence as Hans spots a potential child victim in the reflection of a mirror. Like Lang in no longer registering the noises outside his sidewalk café, the language of the film combines an individual medium close-up of Hans with the abrupt use of silence to infect the viewer with his disturbed

sense of voyeuristic fascination and obsession. Upon viewing the child leave, Hans's stunned aural perceptions slowly returns to normality; once more, the chaotic sounds of street movement fade up into the foreground of the film's soundtrack. This moment is made cinematically powerful through Lang's use of selective silence, a silence that compresses the individuality of all visual elements to create one cohesive world of acoustical balance within which Hans's impulses remain firmly at the centre.

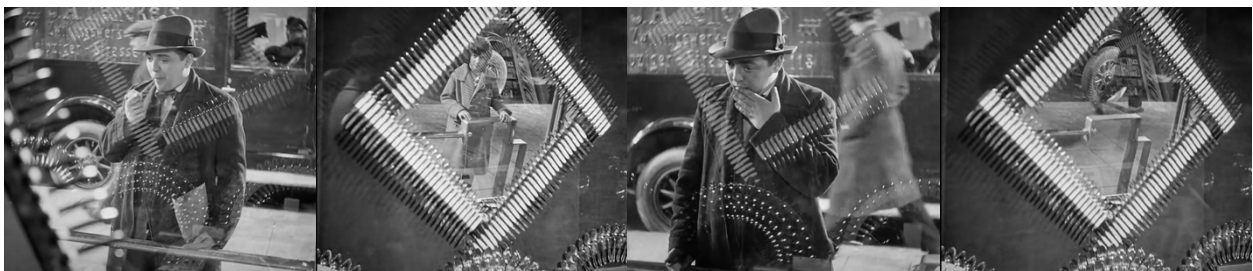


Image 5: As Hans's mental state weakens, the sequence's silence extends the affective possibilities of communicating his distress.

The cinematic language of the film is shifted by silence, jumping from the tumultuous street to Hans's silently impulsive mind; a jump in which the signified layers of acoustic street elements are flattened under the amalgamated levelling of unifying silence. To this, Bela Balazs speaks to the power of silence to merge visual filmic elements; "Sound differentiates visible things, silence brings them closer (making) them less dissimilar." (Balazs, 1952, p.206) Balazs concludes that, within silent close-up, an isolated characters or object "takes on a significance and tension that seems to provide and invite the event which is to follow." (Balazs, 1952,

p.207) Similarly, Fred Camper states that “One of the many uses of silence is to remove an image from its attachment to the natural world.” (Camper, 1985, p.380) Thus, free from the aesthetically natural effects of sound “the work can now operate solely for the eye and, through them, address the mind.” (Camper, 1985, p.372)

2.2 Acoustic Landscapes: The Arrival of Aural Space

Upon the very first scene of *M*, Lang makes clear his use of sound not only to create unseen virtual space, but to render the existence and presence of any objects/characters within these spaces. The film opens on a high angled establishing shot of a group of young children playing an elimination game. One child sings a tune, pointing at each member of a circle of children, metaphorically eliminating them while warning of the roaming child murderer. As the camera pans away to reveal an adult ordering the children to stop singing, they continue to do so off-screen, thus conceptually provoking the already established space as an unobserved mental space via the connective tissue of sound. The spectator is visual witness to all but a banister and washing line within the panned away image, yet the child's singing reduces the visual significance of this onscreen space, establishing an acoustic landscape within which the visual and aural combine cinematically to create new meaning and significance. Bela Balazs unpacks this further, affirming that "If the sound or voice is not tied up with a picture of its source, it may grow beyond the dimension of the latter." (Balazs, 1952, p.209-210) These remarks of sound traveling 'beyond the dimension' of visual images suggests the influence of space, in which both 'audio' and 'visual' elements of the film work cooperatively to expand. Michel Chion also speaks to this audio-

visual phenomenon in declaring that “sound’s ‘offscreenness’, in monaural cinema, is entirely a product of the combination of the visuals and aural. It is in this *relation* of what one hears to what one sees, and exists only in this relation.” (Chion, 1994, p.83) To lose the visuals within this scene would be to misplace that the second round of singing is in fact happening off-screen; thus, the singing connects the spaces temporally while confirming the unified existence of each.



Images 6 & 7: The children’s singing as an acoustic element unifies the spatial dimensions of the scene’s visual sections.

This effect is confirmed narratively by the film within the next scene, in which, after one adult complains about the children’s singing, a second replies, “Leave them alone. As long as we can hear them singing, at least we know they’re still here.” With this, Lang confirms his intentions to deploy sound spatially, using its acoustic qualities to confirm both the presence and absence of off-screen elements, but also to combine visual elements within two contrasting spaces. For example, soon after the opening of the film, a mother awaits the homecoming of her child from school. Within her onscreen space, a church bell is heard from a distance,

acoustically symbolising the end of the school day. The film then cuts to an establishing shot outside a public school, within which the bell continues to ring, only louder and with more timbre and spatial proximity.



Images 8 & 9: The sound of the church bell aurally unifies the mother's emotional relationship to the school, while allowing the spaces to exist concurrently.

Not only does this sequence suggest the unseen presence of a church bell that is very much an object of the film, yet never visualised, but the sound also cinematically connects the mother to the school, and the school to the mother. Through the sound of the bell, the film establishes a communal space in which these events occur at the same time, rather than by one shot after another in filmic sequence. This use of sound establishes an acoustic landscape that discontinues the spatial language of the silent film, in which signified narrative progressions of visual blocks would constantly move forward; 'mother' *then* 'school'. Instead, Lang's use of off-screen sound universalises the film's signified visual space; 'mother' *meanwhile* 'school'.

Describing the spatially unifying effect of film music within the silent era, Theodor Adorno and Hanns Eisler state that “Its social function is that of cement: (music) binds together elements which would otherwise oppose each other unrelated – the mechanical product and the spectators as well as the spectators among themselves... It seeks, after the fact, to breathe into pictures some of the life that photography has taken away from them.” (Adorno & Eisler, 2005, p.59) The ability of sound to ‘breathe life’ into pictures can be explained within its relationship to time and space. Adorno and Eisler allude to the process of mechanical reproduction undertaken by film images, in which the reality of their visual content is physically bound by the medium they exist within. Tom Levin, in defining the ‘Acoustic Dimension’ of film sound, affirms that “Sound exists only in time (there is no acoustic equivalent to the freeze frame) and sound exists only in space.” (Levin, 1984, p.62) From this, it could be established that sound, in existing only in time and space, is able to bypass the confinement of its own reproduction, and thus holds the ability to reaffirm the spatial dimensions of images when used collectively with visual content. Indeed, Levin goes on to state that “In its three-dimensionality sound thus seems more directly tied to the space of its occurrence, its ‘here and now’, even when sound is mechanically reproduced.” (Levin, 1984, p.62) The ability of sound to travel within air-bound acoustic vibrations defines its use as a spatial object able to unify visual spaces;

the church bell heard by the mother in *M* is the same bell heard by the film's audience, which is in turn the same bell heard by the sound engineer when capturing its sound. Jean-Louis Baudry notes that "In cinema... one does not hear an image of the sounds but the sounds themselves... they are reproduced, not copied." (Baudry, 1980, p.47) Within the scene, the spatial aspects of the film's editing in communicating the filmic and sonic rhythm of 'mother' *meanwhile* 'school' can be recognised as a product of sound as a time and space bound medium, one in which the spectator experiences the same aural effect as the film's characters, yet also one able to defy its restrictions as a work of reproduction, thus reapplying to the film's visuals a sense of their three-dimensional foundations of which would have been abandoned during their transformation from spatial reality to visual reproduction.

Another important scene in the film when examining Lang's use of film sound to establish space arrives a short time later, in which the mother, now worried about the whereabouts of her child, desperately calls out her name in the hope of a response. She begins onscreen, calling somewhat calmly from a window, "Elsie! Elsie!"; yet the camera soon cuts away, inhabiting several empty spaces in which the mother's increasingly distressed calls begin to acoustically penetrate, calling attention away from the spaces as individual locations, and instead identifying each site as a unified representation of Elsie's lack of physical presence. The

mother's voice begins to explore each location, separating from the body of her visual existence to autonomously encounter and explore each space. These spaces share no visual similarities; a staircase, an attic, an empty plate and chair at a dinner table... the locations are connected not by what they hold, but what they do not.



Image 10: The mother's voice spatially connects the spaces, allowing them to represent what is visually absent, rather than what is visually present.

The mother's voice, uncaged by the virtual limitations of the film's narrative space or the material limitations of the film's framed space, leaps the visual confinements of the cinema to exist within its acoustic landscape, a space not inherently definable by the frame, but by time, both the spectator's and the film's. Even critics such as Rudolf Arnheim, who largely opposed the use of sound in cinema, acknowledge the ability of sound to establish an extended layer of cinematic space; declaring that, "Sound arouses an illusion of actual space, while a picture practically has no depth." (Arnheim, 1933, p.232) Here, Arnheim speaks to the spatial priority of sound given over vision, one that allows its sonic vibrations to travel around walls and into unseen spaces, both physical and mental. Siegfried Kracauer later described film sound as possessing "the quality

of bringing the material aspects of reality into focus”, while also outlining sound’s ability to evoke mental imagery; “any familiar noise calls forth inner images of its source as well as images of activities, modes of behaviour, etc.” (Kracauer, 1997, p.124) From this, we can extract that the mother’s call within Lang’s sequence holds and sustains a segmented mental image of her physical presence, evoking the manifestation of her space while concurrently visualising another separate space, thus cinematically merging the sensibilities of the two locations; we hear and mentally envision the mother’s suffering and desperation, without once needing to cut back to any visual evidence of the event. The spatial dimensions of the empty locations expand under the weight of the acoustic time-bound medium when amalgamated, thus enhancing their symbolic meaning within the film’s audio-visual editing, while also revealing and emphasising the emptiness of their depicted visual space.

Lang’s spatial use of an acoustic landscape also holds advantages in terms of narrative, within which, the film’s space is constructed to provoke fear and suspense from an emotionally invested spectator. Balazs notes that “a sound that seeps into the shot sometimes impresses us as mysterious, simply because we cannot see its source.” (Balazs, 1952, p.209) This effect is used throughout *M* when concealing the identity of the child murderer, most notably upon his first appearance. As young Elsie plays on the street, innocently throwing a ball against a wall, Lang’s

camera pans up to reveal a wanted poster, which in turn is obscured by the shadow of a mysterious man who soon begins talking to Elsie. This short conversation is heard off-screen while the wanted poster remains visible, yet the vulnerability of the situation is amplified within its spatial restrictions.



Images 11 & 12: Together, the murderer's visual shadow and unseen voice spatially penetrate the scene, allowing him to become present while also remaining anonymous.

The murderer's sinister voice intrudes upon the film's imagery, piercing an already established space with a new aural dimension, one lacking in visual reference. Because of this, the murderer's voice is disconnected from a visual body, thus creating an uncanny psychological effect in which the spectator can register the presence of 'a' man, yet cannot intellectually identify his individual attributes. The mental possibilities of this man's spatial presence and appearance are far more haunting and disturbing than any possible visual reality, thus the film provides the audience with his acoustic space, using this to provoke nervous presumptions about his unseen visual space. Sound deployed to such an effect works to manipulate the spectator's sense of presence and absence; the cinema

in sound no longer required visual cues and reference to signpost the presence or absence of characters, but could suggest haunting presence with the obsessive whistling of Hans, or lonely absence with the echoed call of Elsie's despairing mother.

This spatial phonology communicates a sense of emotional affect to the viewer, who perceive off-screen danger via sonic traces, yet are restricted from the identity of its source. To this, Alberto Cavalcanti notes that "noise seems to by-pass the intelligence and speak to something very deep and inborn." (Cavalcanti, 1939, p.138) Indeed, Mary Ann Doane theoretically links this notion to infancy, stating that "Space, for the child, is defined initially in terms of the audible, not the visible" (Doane, 1980, p.44) Guy Rosolato extends this argument in suggesting that "It is only in the second phase (of development) that the organization of the visual space insures the perception of the objects as external." (Rosolato, 1974, p.80) Both Doane and Rosolato situate Cavalcanti's 'deep and inborn' associations with sounds to human development, in which noise establishes the spatial foundations of a child's perceptions, before vision then confirms such noises as happening externally from the human body. Thus, within the murderer's disconnected looming voice, the spectator becomes engulfed within a confused space of pre-development, one in which the visual evidence of his physical presence is absent. Because of this, we cannot entirely confirm his physicality outside our own mental

perception; his voice takes on the ghostly privileged position of infiltrating the spectator's viewing space, while in turn, the spectator is completely blind to the attributes of the murderer's space. Through this use of sound, the spectator becomes the spectacle of the murderer, as their spatial position is exposed, while his is not; the voyeuristic tendencies associated with cinema are reversed. This occurs within a visually restricted viewpoint that increases a sense of vulnerability on the part of the spectator, forcing them to rely solely on the instinctive and emotional attributes of their own sonic perception. In manipulating the film's acoustic landscape in such a way, Lang is able to heighten tension and surprise, provoking viewer discomfort through the creative and collaborative use of off-screen noise and counterpoint sound.

PART THREE

‘Dr. Mabuse’ and the Voice in Cinema

3.1 Silent Voices: 'Deaf' Cinema and the Voice

After losing 150,000 marks while gambling, Edger Hull attempts to pay off his debt by visiting the man he believes his previous night's game of cards was lost to. The scene soon cuts to an intertitle, '*The room next door*'. Upon this visual cue, we cut to Cara Carozza listening in on the conversation next door, her body positioned to communicate her acoustic experience as she leans one ear and a hand pressing against the offending wall. The scene itself belongs to Fritz Lang's 1922 silent feature, *Dr. Mabuse the Gambler*, in which Dr. Mabuse, the infamous cinematic criminal, uses hypnosis and confusion to cheat unsuspecting victims into complying to his every demand. The scene itself works against the basic notions of a 'silent' film, for the spectator does not hear Edger's conversation, yet they understand that Cara is able to, even without the advantage of visually witnessing the scene. The setup is presented as possessing an acoustic landscape, only this landscape is definable exclusively by what is seen, rather than what is heard. Lang uses a range of subtle visual techniques to construct the 'sense' of a sonic atmosphere, including the scene's visual intertitle, parallel editing, and the bodies of the performers to suggest symbolically the acts of hearing and listening. The spectator is led to assume that this visually depicted world does include sources of sound within its virtual space, yet the sounds these sources

admit are unheard, contained within the edges of the film's frames, and are only made cognitive via the cinematic language of images.



Image 13: An intertitle suggests the acoustic possibilities of the film's unheard digesis.

Micheal Chion reflects on this point in stating that “The silent film may be called deaf insofar as it prevented us from hearing the real sounds of the story. It had no ears for the immediate aural space, the here and now of the action.” (Chion, 1999, p.7) Thus, Lang’s scene is restricted technologically of the sonic connective tissue that the church bell ringing provides in *M*; yet this in turn does not mean that its sound is ‘silent’, or, as some suggest, ‘mute’, but rather that the scene’s sound is contained beyond the film’s ability to display it, evident in the fact that Cara is able to hear and acknowledge Edger’s conversation outside her visual capabilities. Chion goes on to conclude, “it’s not that the film’s characters are mute, but rather that the film was deaf to them.” (Chion, 1999, p.8)

For filmmaker Robert Bresson, there never was a ‘mute’ or ‘silent’ cinema, but rather a mute cinematic style. “For the characters did in fact talk, only they spoke in a vacuum, no one could hear what they were saying. Thus it should be said that the movies had found a mute style.”

(Bresson, 1966, p.301) Bresson is referring to an aesthetic style in which the body and facial expressions become exaggerated bearers of an imagined sound. This effect is apparent within our chosen scene, in which Edger, confused and bewildered by the psychopathic mind-games of Dr. Mabuse, attempts to pay his gambling debt off to a hungover man who, has not only never met Edger before, but is just as confused as he is. The pair engage in a perplexed exchange that sees the hungover man attempt to explain the situation within a range of quickly conflicting expressions and gestures, each separated by a pained facial movement and the symbolic rubbing of his clearly aching head. The man's expressionistic gesturing attempts to override the film's suppressed sonic atmosphere, thus the film's characters seem to acknowledge the fact of their own perceived muteness via their efforts to overdramatically express themselves by visual means. This visual expression accumulates into Bresson's 'mute style', but is also explained by Mary Ann Doane, who supposes that "The uncanny effect of the silent film in the era of sound is in part linked to the separation, by means of intertitles, of an actor's speech from the image of his/her body". (Doane, 1980, p.33) The scene's 'mute style' creates an aesthetic in which the intertitles become the words of any given character's voice, yet not their voice itself. The complete sense of a 'voice' within Lang's scene is thus constructed by three essential visual building blocks:

1. *Intertitles* – which carry the signified ‘meaning’ of individual silent words, ensuring the correct message has been communicated.
2. *Exaggerated Expressions* – which attempt to override the film’s technological disadvantages by exaggerating and extending the affective impact of silent sounds and voices.
3. *Silent Aural Environments* – which connect silent spaces in the film by depicting characters able to hear voices within the space of its diegesis, even while lacking visual evidence of the sound’s source.

These elements combine to create perceived voices occurring within the imagination of the spectator. On this, Chion notes, “Still, this spectator who is forced to be deaf cannot avoid hearing voices – voices that resonate in his or her own imagination.” (Chion, 1999, p.9) Lang’s film, through its construction of ‘deaf sound’, lives within a world of acoustic aspirations; its visual aesthetic attempting to emulate the future advantages of filmic sound, yet its technological capabilities unable to produce the desired effect. *Dr. Mabuse the Gambler* instead attempts to construct an imagined deaf sound via the visual aesthetic of its silent form, using the language of silent film to construct perceived acoustic landscapes, proving relevant the comments of André Bazin in suggesting that “not all silent films want to be such” (Bazin, 1971, p.138)

3.2 The Voice as Meaning and Presence

As four criminals enter a secured darkened room, they catch a glimpse of a mysterious silhouetted figure from behind a giant curtain stretching from one side of the room to the other. As the lights switch on, the silhouette disappears, yet is replaced by a voice. This voice is no longer silent, no longer constructed of visual elements; the world of Dr. Mabuse in *The Testament of Dr. Mabuse* (1933) is now acoustic, a solidified world in which bodies and voices form spatially to reflect a unified audio-visual presence... that is, all apart from Dr. Mabuse himself. The man behind the curtain during this point in the film is believed to be Dr. Mabuse, yet the spectator already knows this cannot be the case since the voiceless body of Dr. Mabuse has already been seen locked away within a mental hospital earlier in the film. The voice behind the curtain appears to represent Dr. Mabuse's presence within the room since its words expose the actions of the criminals, even shutting down the speech of one man who indicates with visual body language that he is about to interrupt. Thus, this bodiless voice holds the ability to see the men, while in turn, they, like the spectator, cannot see it. The bodiless voice in the scene fits the description given by Chion of what may be labelled the "*acousmêtre* – the one who is not-yet-seen, but who remains liable to appear in the visual field at any moment." (Chion, 1999, p.21) Chion goes on to describe the

acousmètre as harbouring unique 'all-seeing' powers within cinema, including "Being in the film and not, wondering the surface of the screen without entering it... He must, even if only slightly, have *one foot in the image*". (Chion, 1999, p.24) Within the film, Dr. Mabuse's visual body remains silent, yet his aural presence permeates the entire film. This is achieved, in part, by safeguarding the tone of Mabuse's voice; we never see him speak directly, and was deaf to his muted voice within the previous silent films. The acoustic sound and colouring of Mabuse's voice remains anonymous, thus every occurrence of a bodiless voice within the film holds the possibility of being his; the physical embodiment of his presence is able to move in, out, and between filmic spaces fluidly while imitating the omnidirectional nature of sound.



Images 14 & 15: Mabuse's voice and body appear to exist independently of each other.

The film's acousmètre effect is also created by Mabuse's surrogate voices, of which there are two. One of these is the film's namesake 'testament', which is obsessively written by Mabuse during his silent stay

at the mental hospital. Like the intertitles of the silent films, Mabuse uses written visual communication to give meaning and instruction to his words. These words are then verbalised and mass communicated by Mabuse's other surrogate voice, of which, is situated behind the curtain, and, as suspected, does not belong to Dr. Mabuse at all. Reflecting on Chion's concept of the *acousmètre* voice created by such an effect, Mladen Dolar notes, "The voice behind the screen not only fuels our curiosity, but also implies a certain disavowal epitomized by the formula ... "I know very well that the voice must have some natural and explicable cause, but nevertheless I believe it is endowed with mystery and secret power." ... it presents a puzzling causality, as an effect without a proper cause." (Dolar, 2006, p.66-67) Dr. Mabuse's physical embodiment permeates the space of the room via the sound of his surrogate voice, alluding to his impossible presence, while in turn segmenting the images of his silent body from the audio of his bodiless voice. This scene, as well as the images of Dr. Mabuse before it, tear apart the fundamental aspects of his existence, thus creating, in essence, two Dr. Mabuse's, each representing an uncanny 'sense' of his complete manifestation, yet neither able to portray a coherent acceptance of his presence. Unlike the film's other characters, Dr. Mabuse is never entirely 'present' within the diegesis of the film's virtual world. Instead, Mabuse lives within the outside edges of the film's frames, looking in.

To further understand Mabuse's verbal surrogate voice, it should first be noted the way in which its source is revealed to both the film's characters and its audience. Within the scene, two of the film's main heroes, both Kent and Lily, are locked within the aforementioned curtain room after being captured by Dr. Mabuse's team of criminals. Once again, the bodiless voice assumed to belong to Mabuse begins to speak; it threatens Kent and Lily, informing them that they will "not leave this place alive". In defiance, Kent shoots towards the silhouetted figure behind the curtain, before discovering, on the opposite side, nothing more than a human-shaped board and a mechanical loudspeaker, which has until now acted as the physical presence of Dr. Mabuse, yet is in fact a mechanical mass communication device.



Images 16 & 17: Lily and Kent expose the apparatus behind Mabuse's acousmêtre voice.

Lily and Kent expose the perceived unity of body (the board) and voice (the speaker) by revealing the hidden apparatus of its constructed aesthetic form; this in turn disbands Mabuse's sense of presence by

eliminating the effect of which it has relied on to sustain its fundamental conceit of authenticity. Michel Chion discusses the relevance of this scene, linking its themes to the cinematic and audio-visual conceit of cinema itself.

“What the couple finds in tearing down the barrier of the curtain is what the film spectator would find if he tried to take the projected image as a concrete material reality ... if he were to tear apart the screen, this curtain that conceals itself – in order to enter the space of its false depth.” (Chion, 1999, p.42)

Chion draws comparison between the apparatus of Mabuse’s concealment and the hidden apparatus of the film itself. Mabuse’s voice and perceived physical presence become spatially congealed within the anonymous presence of such apparatus, as are the voices and bodies of Kent and Lily in the scene by the mechanical process of Lang’s audio-visual filmmaking. The process of constructing such concealed audio-visual ideology is labelled by Mary Ann Doane as the ‘fantasmatic body’, of which, film’s such as *The Testament of Dr. Mabuse* attempt to challenge in their handling of voice to suggest physical presence and symbolic meaning. Doane works from the notion of Pascal Bonitzer, who suggests that, within the cinema, uses of the ‘voice-off’ are, “submitted to

the destiny of the body". (Bonitzer, 1975, p.25) Doane takes this statement further, claiming its concept to inherently be the case since off-screen voices belong to "a character who is confined to the space of the diegesis, if not to the visible space of the screen. Its efficacy rests on the knowledge that the character can easily be made visible by a slight reframing which would reunite the voice and its source." (Doane, 1980, p.41)

The described scene from *The Testament of Dr. Mabuse* works against Doane's claim, as the film's 'reframing' of Mabuse in exposing the source of his surrogate voice only works to confirm his disassembled materiality within the film, as opposed to reuniting it. The mystery created within the acousmêtre voice is exposed; so too is Dr. Mabuse's status as an on and off-screen 'fantasmatic body' within the film, yet his voice, in this moment, still resists returning to Dr. Mabuse as a material possession, but rather vanishes from the confinements of the vocal apparatus in an attempt to re-establish itself somewhere else. This is apparent in the fact that Dr. Mabuse dies shortly before this scene, having never uttered a single word while onscreen, but is also found in the reveal that another character in the film, Professor Baum, was in fact the 'voice' of Dr. Mabuse all along. Thus, the 'voice' of Dr. Mabuse is established as, less the acoustic communication device of speech, and more the aural manifestation of his persona via a surrogate speaker. This persona is established within the perception of the listener, and so it could be claimed that Mabuse's true

‘voice’ is not one single signifier, but rather, means different things for different characters... or perhaps even, does not exist at all. As Chion explains, after the ‘de-acousmatization’ of Mabuse’s voice, “it would appear that Mabuse is nothing – nothing more or less than what people construct him as – and that he can exist at all because none of his properties are fixed. ... he is in this name without an identity, this body without a voice, this voice without a place”. (Chion, 1999, p.36-37)

The limitless possibilities of Mabuse’s aural presence are extended by his death, by the detaching of voice and body as two separate elements able to move fluidly around the space of the film, one never reliant on the attendance of the other. This concept is visualized in one scene depicting Dr. Mabuse after his death, as a ghost. Unlike any other moment of Mabuse onscreen, during this scene, he is depicted speaking directly on camera; an apparent unified manifestation of voice and body that could be viewed as a submission to the ideology of the fantasmatic body. Yet, Mabuse’s ‘body’ in this scene does not represent the physical space of the film’s diegesis, but is rather, as suggested earlier, a visual perception, a mental construction of Mabuse’s persona. In the scene, Mabuse reveals himself as a ghost to Professor Baum, who witnesses the ghost while reading Dr. Mabuse’s left behind testament. Mabuse begins to speak within a menacing whisper, one echoing the sounds of an old witch as he struggles to force strained words from his deteriorating voice. Important

to note is the fact that Mabuse does not address Baum directly, but rather continues reading out loud from the testament at the exact point from which Baum left off. This would suggest the hypnotic state of Baum, one induced from the written words of the testament, resulting in the constructed manifestation of Mabuse's persona, one assembled from the deepest and darkest fears of Professor Baum, rather than the accurate attributes of Mabuse's true vocal colouring.



Image 18: Baum's voice is used as a bridge by which Mabuse can access his body.

Chion describes the scene, noting the symbolic progression of the voices in suggesting a gradual increasing of Mabuse's presence. "In turn this reading voice (of Baum) summons *another voice*, an I-voice that inhabits him and then becomes the voice of a ghost that becomes visible, immobilizing him with his hypnotic eyes." (Chion, 1999, p.138) Starting from the source of the testament, Mabuse's written words use Baum's voice as an acoustic connection upon which to cross into visual form. Mabuse has no acoustic voice, rather his voice and presence in this scene are mentally constructed by Baum via the written communication of his testament and the visual manifestation of his ghostly image. More than

any other acoustic voice or visual representation in the film, the testament acts as the source of Dr. Mabuse's many surrogate voices; be it the voice of Professor Baum or the eerie voice of Mabuse's ghost, the most accurate sense of Mabuse's presence within the film generates from his testament. This is accumulated within the film's very last scene, in which Baum, now driven insane by Mabuse's evil, is seen tearing up the testament. Although Mabuse physically died less than halfway through the film, the tearing of his testament symbolises his true onscreen death. The destruction of the document is the destruction of Mabuse's voice, the removal of his acousmêtre presence, and the death of his silent communication established within early cinema in favour of the new and altogether more technologically advanced arrival of the onscreen spoken word.

3.3 The Voice and Devices of Power

In establishing the themes of possession within *The Testament of Dr. Mabuse*, Michel Chion details a shift from the hypnotic gaze to the hypnotic voice.

“During the silent era, (Mabuse) was a hypnotist and acted through the look. *The Testament* refers to that only in brief moment of close-ups on blankly staring eyes. But on the other hand, the film deploys the new power exercised by texts and above all *by the voice*.”
(Chion, 1999, p.138)

To understand this progression, it is first useful to understand how Lang constructs the hypnotic look within Dr. Mabuse’s silent episodes. During one sequence in *Dr. Mabuse the Gambler*, Mabuse attempts to hypnotise state prosecutor Norbert von Wenk during a game of cards. The scene is filmed from point-of-view vantage points as both men look directly to camera as they communicate; the effect creates the cinematic form of back and forth communication, but also allows the film’s audience to become complicit in the visual act of hypnosis about to unfold. Mabuse reveals a pair of Chinese spectacles, which he claims are from “TSI-NAN-FU”. Like a magic spell, these words begin to haunt von Wenk, visually

appearing on his cards and on the table before him; Mabuse's intertitles bypass the boundaries of the cinema to infiltrate the physical space of the film via way of von Wenk's hypnotic state. Mabuse also uses his eyes to disorientate von Wenk, who sits confused in a wide shot as darkness begins to fade out both sides of the screen, eventually trapping him in isolated darkness, before cutting to a huge darkened close-up of Mabuse's menacing eyes.



Image 19: Before sound, Mabuse deployed visual manipulation through the look and gaze.

This scene is personal, a two-way interaction between Mabuse and von Wenk, one told within isolated gestures and the direct visual gaze. Stefan Andriopoulos picks up on this sequence as an allegory of cinema itself, claiming that "Fritz Lang's reliance on point-of-view-shots forces the spectator to identify with Mabuse's hypnotized victim. The cinematic representation of hypnosis is thus transformed into a celebration of the hypnotic power of cinema". (Andriopoulos, 2008, p.107) If the silent films of Dr. Mabuse celebrate the visual technologies of early cinema, then *The Testament of Dr. Mabuse* does the same for, not just sound cinema, but for all aural communication devices and technologies of the 20th century.

Jonathan Crary details the changing nature of Mabuse's hypnotic manipulation by framing his argument within the progression of cinematic technology. "Like the optical modalities in the earlier film, hypnotic forms of influence here proceed by the isolation of a single sense, in this case hearing rather than vision... Lang was a close observer of the ways in which different technological networks permeated a densely layered social space". (Crary, 1996, p.273-274) For Crary, the progression from Dr. Mabuse's silent world to that of sound marks an evolution of technological ability, evident in the film by both Mabuse's and Baum's use of technology as devices of power to manipulate the free will of individuals and large groups of characters. Recording and listening devices of power are not only used to trick and deceive individuals and groups within the film itself, but are also able to imitate human personas, seamlessly embodying the presence of characters to such an extent that even the film's spectator is completely fooled. For example, throughout the film, Professor Baum vocally denies access to his office from multiple characters. This denial is generated from an intruder moving Baum's office door handle, to which he replies from beyond the door, "I am not to be disturbed." The control of this command overpowers the actions of the intruder, forcing them away from the office door. The narrative will later discover that Baum's instructions are in fact generated from a recording

set to a gramophone, which has been manipulated to play each time the door handle is moved.



Image 20: The gramophone acts as a device of power by providing to Mabuse a surrogate voice.

Using this apparatus, not only can Mabuse be absent from the scene, but so can his surrogate voice. The film uses the gramophone in such a way to perfectly replicate the human voice, its tones, shape, and colouring. The accuracy of the device in the scene would have been superior to commercial gramophones at the time of the film's making, yet media critics such as Friedrich Kittler would later foresee such advanced devices of power as transforming the real, pulling the phantom of the spectre towards the authentic, while concurrently blending the line that separates their aesthetics; "As if there were no distance between the recorded voice and the listening ears, as if voices travelled along the transmitting bones of acoustic self-perception directly from the mouth into the ear's labyrinth, hallucinations become real." (Kittler, 1999, p.37) Mabuse's many acoustic mechanisms bring to life the physicality of his ghostly presence. The hypnotic powers of this device are no longer found within the direct visual

hypnotic gaze, but rather within a completely voluntary desire to believe the signified meaning of Baum's recorded words; as Chion explains, "We learn from this that the only obstacle to restoring the acousmatic voice to its source is our own voluntary blindness, the desire to believe in (Mabuse's) power." (Chion, 1999, p.34) Mabuse cheats and manipulates in much the same way as in *Dr. Mabuse the Gambler*, only this time the gaze towards one individual has been replaced by a sonic communication able to reach multiple victims. In exploring the film further, Stefan Andriopoulos goes on to note the hypnotic possibilities of such devices with reflection to their use in Lang's film.

"The rapid emergence of radio as a mass medium ... simultaneously displaced the anxiety of an irresistible hypnotic influence on to another medium capable of addressing a spatially dispersed but potentially infinite number of listeners at the very same time." (Andriopoulos, 2008, p.119-120)

Andriopoulos makes connections between Mabuse's use of mass suggestion and propaganda/psychological advertising techniques, all of which were made possible through the technological advancements of listening, speaking, and recording devices such as the radio, microphone, or gramophone. This effect can result from removing the voice from its

original source and placing it within another, especially when that source is unknown, hidden, or misrepresented, as is the case in *The Testament of Dr. Mabuse*. Within Mabuse's devices of power, an uncanny contrast between the 'visible' and 'audible' is created. Mladen Dolar goes on to define this contrast, stating, "the visible world presents relative stability (and) a location at a distance; the audible presents fluidity, passing, a certain inchoate, amorphous character, and a lack of distance... The acousmatic voice is so powerful because it cannot be neutralized with the framework of the visible." (Dolar, 2006, p.79) In hearing the unknowable voice from behind the curtain, the characters of *The Testament of Dr. Mabuse* face an amorphous opponent, one whose hypnotic power and control are increased by an anonymous image and an ungraspable voice. The ability of Mabuse to create this sense of aural confusion is foundational in establishing his filmic power. The possessive voice in the scene flows fluidly through the space of the room, controlling and commanding the large groups of criminals who inhabit the space through its ungrounded nature. Removed from the confinements of the body or a fixed visual signifier, the authoritative power of Mabuse's many voices move undetected around the space of the film, always present, yet selectively heard.

The Testament of Dr. Mabuse allows Mabuse to progress as the cinema does; from a media of pure visual hypnosis, to one combining the

powers of the audio and the visual in order to create the ideology of a material heterogeneity. Mabuse is never united as body and voice, yet creates the allusion of unity for the purposes of spectatorial control and manipulation. As does cinema, Mabuse uses sound to embody and enhance his voiceless image; the film, “expose(s) the very structure of sound film, based on an off-screen field inhabited by the voice, which is the inevitable corollary of the onscreen field” (Chion, 1999, p.150) As the cinema increases in its technological powers, so too does Mabuse in haunting the screen from behind its ideological curtain. Mabuse succeeds in the silent world by acknowledging the visual power and significance of the media he lives within; likewise, Mabuse succeeds in the sound cinema by manipulating the very technological apparatus needed to create its mask of audio-visual unity. By doing so, Mabuse’s ‘voice’ as a surrogate voice, as a testament, and as a device of power, is able to uncover the mask of the sound cinema, using its acoustic advancements for his own criminal and sinister thirst for influence; even after his physical onscreen death, Mabuse’s voice lives on.

PART FOUR

Cinema Transformed

As detailed at the start of this essay, early film theorists such as Balazs were, in first witnessing the gradual approach of sound as a cinematic standard, somewhat apprehensive to say the least. However, hope in the cinema and cinematic language was never lost by Balazs, who would go on to state:

“in history there are no moral tragedies, only crises ... Only after such an awakening does the dialectical process of development begin, in which the now conscious purpose seeks more new technical possibilities of expression.” (Balazs, 1952, p.196)

The arrival of sound, a purely sonic element within what is often considered a primarily visual medium, transformed the aesthetic qualities of cinema in ways more profound than any other cinematic development in history. Cinema now possessed an entirely new aesthetic layer to expand creativity and communication possibilities, one in which on and off-screen space could be enlarged, one that provided the unity of body and voice within the integration of the visual and audio, and one that could exist outside the visual frames of the film to harbour and conceal mysterious and haunting elements with the aim of increasing spectatorial participation and affect. But had the ‘sound film’ lived up to its critical expectations? In fear of the theatrical ‘talkie’ that would reduce cinema to

the literature style of the theatre stage, the sound film, as discussed within classical sound theory, presented cinema with the 'technical possibilities of expression' noted by Balazs. Lang deployed sound economically, using the sonic qualities of sound technology to create new psychological spaces that could mentally expand the visual possibilities of his films. *The Testament of Dr. Mabuse*, as well as *M*, provide aesthetic solutions to many theoretical concerns of early critics, yet also go further in their experimentations with sound to expand the arguments of later theorists, thus providing a map with which to contrast and debate the formalistic principles of both, while in turn offering new ideas and filmic possibilities to the newly emerging form of audio-visual cinema.

To begin with *M*, Lang makes immediately clear sound's ability to be cinematic, to extend and elevate the visual language of cinema. Returning to the statement on film sound led by Sergei Eisenstein, it becomes clear how Lang overcomes Eisenstein's concerns. The statement makes clear the importance of segmented visual images in editing as montage elements, while also stressing the degrading effects of sound to restore autonomy to such images, thus disrupting their constructed cinematic meaning. Without offering a specific example, Eisenstein broadly claims that "the first experimental work with sound must be directed along the line of its distinct nonsynchronization with the visual images." (Eisenstein, 1977, p.258) Eisenstein acknowledges the ability of sound to work in

counterpoint to images, yet his aims are more aligned to preserving the symbolic quality of silent images than that of using sound to establish spatial connections and create new mentally perceived ones. Lang does not follow Eisenstein's suggestions directly; sound connects the film's spaces, yet it stops short of using sound "as a factor divorced from the visual image." (Eisenstein, 1977, p.259) Instead, to return to the off-screen haunting of Hans's murderous voice, the film separates sound from its signified source, yet the acoustic attributes of that sound are still able to affect the visual world of the film. Hans's voice remains, 'in sync' with his presence and speaking, as shown in the moving lips of his disembodied shadow on the wall, yet the spectator cannot directly see Hans as the sound's source, thus he remains neither synchronised to the images or in counterpoint to them, but instead inhabits a halfway point in which neither sound nor image removes or confirms the aesthetic qualities of the other. Here, as with the many other noted examples taking place in Lang's film, sound and image work to create an equal balance of aesthetic relevance. An image is present, yet we only see a shadowed trace of the film's main visual focus; a voice is also present, yet its visual source is unseen and unidentified. Lang displaces both sound and image in order to equalise them in unity, forming an audio-visual cinematic language in which each accomplish a balance symbolic role in creating narrative communication and emotional affect. Sound in the film is used to suggest presence and

absence, meaning that, via sound, Lang creates off-screen threats and unidentified narrative traces, some of which are only ever present in sound. *M* thus fulfils the theoretical expectations of the sound film, while also withholding the inherently incomplete and indexical aesthetic qualities of the silent film in providing audio-visual traces with which the spectator can dream or perceive unquantified virtual spaces. The sound of *M* does not aesthetically overpower the film's visuals or plainly explain its narrative, rather, it deepens them.

In *The Testament of Dr. Mabuse*, the spectator never once sees Professor Baum interact with his sound apparatus; in fact, the voice of the machine talks to the criminals in the curtain room even before Baum is inhabited by Mabuse's ghost. The question of who is operating the machine before Baum becomes involved is never asked or needed because the sound of the voice creates an unseen visual dimension, one that is assumed to make sense, yet is only trusted because of its altogether lack of visual evidence in the first place. With sound, Lang creates the possibility of a fluid and roaming voice without the need for complex plot explanations or visual storytelling. This issue is noted by Michel Chion, who notices that "In one such scene (the) editing gives the impression that Baum is simultaneously in his office reading, and the source of the voice behind the curtain." (Chion, 1999, p.34-35) Film sound works with the visuals, but can also exist within its own layer of filmic

reality, one independent of film's motions, attributes, or even narration. Thus, returning once more to Mary Ann Doane, who states that the voice-off, such as the one emanating from Mabuse's machine, "supports the claim that there is a space in the fictional world which the camera does not register ... It validates both what the screen reveals of the diegesis and what it conceals." (Doane, 1980, p.40) Through sound, the presence of Dr. Mabuse and Professor Baum are able to exist in two places at once, the insane asylum and office as bodies, and the criminal's hidden room as an assumed voice. Each source of presence is validated as a logical space within the film by visual and acoustic signifiers, yet each are established in the theatre space of the spectator as both a voiceless body and bodiless voice. Mabuse's previous silent ventures could not provoke such a disembodied effect of sound as an independent filmic tool, but rather relied on visual effects such as superimposing images to create spiritless bodies and bodiless spirits. Silent films have always been interested in the haunting prospect of a disembodied presence, yet only with the inclusion of sound was this curiosity fully solidified within a virtual and unseen cinematic landscape. The voice in cinema enhanced the hypnotic powers of Mabuse as a silent character, yet in causing a rupture in Mabuse's personification, the film also ruptures the cinematic language of cinema, exposing the technological equipment with which the film itself has transformed upon. With sound, Lang breaks the ideology of visual

montage cinema and the unity of voice and body, exposing these as technological constructs, while in turn deploying this same technology in order to heighten and enrich Mabuse's haunting power over the film's digesis and virtual spaces.

From my research and analytical work on Lang's films, it becomes clear that early concerns of film sound were, within the wider scope of cinema, mostly warranted. Many films did transform within sound to become inexpressive, trivial, and static. With sound and the voice also comes the temptation to communicate plot and character motivations by dialogue; yet, neither does Lang 'tell not show' nor 'show not tell', but creates innovative structures upon which to transform sound and image into one unified audio-visual language, one that resists placing either medium in aesthetic importance above the other, yet also one managing to avoid the 'talkie' issue of using one medium to aesthetically rely on the onscreen presence of the other. *M* stands as a classic example of a creative sound film, and *The Testament of Dr. Mabuse* maps an important shift in the technological advances of cinema, noting a progression and transformation from the look and gaze to the concurrent and disconnected amalgamation of sound and image, body and voice, as well as presence and absence. This is cinema transformed, transformed from the silent world of symbolic imagery; from one in which the spectator's relationship to the gaze and the look is central as an outwards expression, to one

formed of audio-visual elements that draw the film's created world inwards towards a vast virtual environment of sonically connected space and the possibility of both embodied and disembodied presence through a newly established relationship between body and voice. Never had such a dramatic shift in cinematic language taken place; through Lang's work, sound was finding its warranted place in the language of film, bringing with it an entirely new spectrum of creative possibilities.

Bibliography

Adorno, T, Eisler, H. (2005) *Composing for the Films*, London: Continuum.

Altman, R. (1980) 'Introduction' *Yale French Studies*, no. 60.

Andriopoulos, S. (2008) *Possessed: Hypnotic Crimes, Corporate Fiction, and the Invention of Cinema*, Chicago: University of Chicago Press.

Arnheim, R. (1933) *Film*, London: Faber and Faber.

Arnheim, R. (1957) *Film as Art*, Berkeley: University of California Press.

Balazs, B. (1952) *Theory of the Film: Character and Growth of a New Art*, London: Dennis Dobson LTD.

Baudry, J, L. (1980) 'The Apparatus', in Theresa Hak Kyung (ed.) *Apparatus*, New York: Tanam Press.

Bazin, A. (1971) *What is Cinema?* Vol. 1, Los Angeles: University of California Press.

Bonitzer, P. (1975) 'Les silences da la voix', *Cahiers du Cinéma*, no. 256 (February-March).

Bresson, R. (1966) Interview by Michel Delahaye and Jean-Juc Godard, *Cahiers du Cinéma*, no.178, reprinted in *La Politique des auteurs*.

Camper, F. (1985) 'Sound and Silence in Narrative and Nonnarrative Cinema', in Elisabeth Weis and John Belton (ed.) *Film Sound: Theory and Practice*, New York: Columbia University Press.

Carroll, N. (1978) 'Lang, Pabst and Sound', in *Cine-Tracts*, vol. 2, no. 1 (Fall).

Cavalcanti, A. (1939) 'Sound in Film' *Films* 1, no 1: 25-39; reprinted in Elisabeth Weis and John Belton (ed.) *Film Sound: Theory and Practice*, New York: Columbia University Press.

Chion, M. (1994) *Audio-Vision: Sound on Screen*, New York: Columbia University Press.

Chion, M. (1999) *The Voice in Cinema*, New York; Columbia University Press.

Clair, R. (1985) 'The Art of Sound', in Elisabeth Weis and John Belton (ed.) *Film Sound: Theory and Practice*, New York: Columbia University Press.

Crary, J. (1996) 'Dr. Mabuse and Mr. Edison', in Kerry Brougher (ed.) *Art and Film since 1945: Hall of Mirrors*, New York: Monacelli Press.

Doane, M, A. (1980) 'The Voice in the Cinema: The Articulation of Body and Space' *Yale French Studies*, no. 60.

Dolar, M. (2006) *A Voice and Nothing More*, Massachusetts: The MIT Press.

Duras, M. (1975) *Marguerite Duras*, Paris: Editions Albatros, quoted by Michel Chion in *The Voice in Cinema*, New York; Columbia University Press.

Eisenstein, S. (1977) 'A Statement', in Jay Leyda (ed.) *Film Form: Essays in Film Theory*, New York: Harcourt Inc.

Huillet, D, Straub, J, M. (1985) 'Direct Sound: An Interview with Jean-Marie Straub and Deniele Huillet', in Elisabeth Weis and John Belton (ed.) *Film Sound: Theory and Practice*, New York: Columbia University Press.

Kittler, K. (1999) *Gramophone, Film, Typewriter*, Stanford: Stanford University Press.

Kracauer, S. (1997) *Theory of Film: The Redemption of Physical Reality*, Princeton: Princeton University Press.

Lang, F (2003) 'Fritz Lang on M: An Interview', in Barry Keith Grant (ed.) *Fritz Lang: Interviews*, Jackson: University Press of Mississippi.

Metz, C (1980) 'Aural Objects' *Yale French Studies*, no. 60.

Pudovkin, V. (1954) *Film Technique and Film Acting*, London: Vision Press Limited.

Rosolato, G. (1974) 'La voix: entre corps et langage', *Revue française de psychanalyse*, 38 (January).

Levin, T. (1984) 'The Acoustic Dimension: Notes of Cinema Sound'. *Screen* 25, 3 (May-June).

List of Illustrations

Figure 1. Still from *M* (1931)

Figure 2. Still from *M* (1931)

Figure 3. Still from *M* (1931)

Figure 4. Still from *M* (1931)

Figure 5. Sequence from *M* (1931)

Figure 6. Still from *M* (1931)

Figure 7. Still from *M* (1931)

Figure 8. Still from *M* (1931)

Figure 9. Still from *M* (1931)

Figure 10. Sequence from *M* (1931)

Figure 11. Still from *M* (1931)

Figure 12. Still from *M* (1931)

Figure 13. Sequence from *Dr. Mabuse the Gambler* (1922)

Figure 14. Still from *The Testament of Dr. Mabuse* (1933)

Figure 15. Still from *The Testament of Dr. Mabuse* (1933)

Figure 16. Still from *The Testament of Dr. Mabuse* (1933)

Figure 17. Still from *The Testament of Dr. Mabuse* (1933)

Figure 18. Sequence from *The Testament of Dr. Mabuse* (1933)

Figure 19. Sequence from *Dr. Mabuse the Gambler* (1922)

Figure 20. Still from *The Testament of Dr. Mabuse* (1933)