

The Voice as A Transcending Power:  
The Female Singer, Body, and Political Discourse in 1930s Shanghai

by  
Ziyang Li  
Critical Asian Humanities Duke University

Approved:

Carlos Rojas, advisor

---

Leo Ching

---

Eileen Chow

---

Thesis submitted in partial fulfillment of the requirements for the degree of Master of Arts in the  
Department of Critical Asian Humanities in The Graduate School of Duke University  
2020

ABSTRACT

The Voice as A Transcending Power:  
The Female Singer, Body, and Political Discourse in 1930s Shanghai by  
Ziyang Li  
Critical Asian Humanities Duke University

Approved:

Carlos Rojas, advisor

---

Leo Ching

---

Eileen Chow

---

An abstract of a thesis submitted in partial fulfillment of the requirements for the degree of Master of  
Arts in the Department of Critical Asian Humanities in The Graduate School of  
Duke University  
2020

Copyright by  
Ziyang Li 2020

## Abstract

After the introduction of the synchronized sound technology, China entered the sound-film era in the 1930s. This period features many female singers' cross-over performances in films. This paper mainly focuses on *Two Stars in the Milky Way* (银汉双星 1931) and *Street Angel* (马路天使 1937) to analyze how the female voice engaged with film images, the foreign Hollywood mode, and the indigenous political discourse.

First, sound and visual technologies facilitate a Hollywood-style synchronization rule, which unites the female voice with the body and reduces the songstress to an object of male gazing and listening. However, this masculine rule of embodying voice is not immutably successful. Revisiting the singing skills and the various song-image-narration relations in films, the paper argues: 1) the narrative and camerawork occasionally detach the voice from the body, indicating a more fluctuating body-voice relationship; and 2) the technological mediation provides the female voice with more provocative and interpretative possibilities, turning the voice to an open site that accommodates different subjectivities and transcends the masculine aural and visual containment. In this way, the female singer finally gains self-reflexive consciousness.

Additionally, such multipotential qualities of the female voice moved beyond film and formed an ambiguous relationship with the official discourse: On one hand, as an intimate friend of the official power, the female voice was constructed as the vehicle of national salvation and the ideal of New Woman. On the other hand, because of some songs' foreign origin and emotive timbre, the voice was also regarded by domestic anti-imperialism and nationalism as decadent noise that should be censored and banned.

However, again, the voice's multipotentiality obscured its nature and ultimately escaped the censorship, dismantling the containment of the official knowledge framework.

In sum, combining visual studies, sound studies, musicology, psychoanalysis, gender and sexuality, nationalism, and anti-imperialism, this paper analyzes how the female voice in 1930s Shanghai films becomes an inclusive site containing traces of simultaneous intimacy and dichotomy with, obedience and resistance to patriarchal containment in film and the masculine political discourse off screen.

**Key words: the female voice; singer; body; political discourse; 1930s Shanghai**

## Table of Contents

<i>INTRODUCTION</i> .....	1
<i>CHAPTER 1:</i> .....	7
<i>The Voice as An Eclectic Site: The Female Singer, Body, and Sexuality in 1930s Shanghai Films</i> .....	7
I. The Synchronisation Rule that Embodies the Female Voice .....	7
II. Revisiting the Voice-Image-Narrative Relation: The Detachable Voice and Female Self-Awareness.....	13
III. The Voice as a Multidimensional Site.....	21
<i>CHAPTER 2</i> .....	31
<i>A Consolidating or a Destructive Force? — The Paradox of the Replayed Female Voice Outside of Film Order out of</i> .....	31
I. Recording and Replying: The “Freed” Female Voice as a Versatile Tool.....	31
II. Individual or Collective Listening? — The Paradox of Listening via Radio .....	39
III. A Helper or a Threat? — The Ambiguous Female Voice in Social Discourse and Governmental Censorship .....	44
<i>CONCLUSION</i> .....	55
<i>WORKS CITED:</i> .....	60

## Introduction

In 1932, a full-page ad for the Pathe Record Company (百代唱片) appeared in the June 15<sup>th</sup> edition of the *Movie & Radio Weekly* (*Diansheng* 电声) (see Figure 1). The two sides of the ad displayed four pictures of different cross-over female singers. The central section featured an eye-catching, large-sized slogan that read, “These Movie Stars Have Their Voices Recorded for Pathe Record Company” (这几位电影明星是专为百代唱片而灌音的), accompanied by a framed text about the record company’s advantages:

The Standard-playing Record, a new type launched by Pathe, had the best quality—clear and precise sound, and flawless acoustics. In particular, with the seamless embedment with the brainchild of the brilliant singers, it was preceded by no other counterparts. Hence, it was no exaggeration that records from Pathe represented the top in this industry(百代新法灌制的钢针唱片，质地最好，发音清亮准确，收音完美，毫无瑕疵，且尤最优等歌曲家所灌，珠联璧合，尽善尽美，非别家唱片可比，所以要听最好唱片，非用百代出品不可).

The emphasis on “质地最好，发音清亮准确，收音完美，毫无瑕疵” presented the reader with a familiar philosophical claim about sound reproduction: sound could be perfectly replicated, preserved, and replayed without any distortion.



Figure 1. Advertisement of Pathe Record Company

This emphasis on perfect replication is what Jonathan Sterne calls the philosophy of “sound fidelity”, a term that has gained traction since the age of modern technology (217). Those pursuing sound fidelity reflected a strong desire to technologically transform the original sound into a copy while completely eliminating the artificial process, making it as though the relation of the copy to the original were “transparent” or simply “not there” (218). Such ideology of sound fidelity actually ontologises sound (including the voice), reducing its production and reproduction to a mere result – a fixed,

static material outcome. However, a host of relations orbiting sound reproduction is almost-if not completely- neglected. Sound is merely a physical phenomenon that can be taken for granted. With such sonic reduction, as Nina Eidsheim says, “a thick sonic event may be reduced to a static one, and in the process of this reduction we identify an object, a stable referent” (2). As a result, the complex and flexible sonic event is oversimplified as a set of restricted terms, such as pitch, range, rhythm, singing styles, genres, musical forms, etc.

This ontology of sound as a vibrational affect, according to Stephen Goodman, is a “reductionist materialism” that reduces sound to a “qualifiable objectivity” without further meaning (71). However, aside from sound’s sonic physicality (its corporeal effects), “the semiotic significance of its symbolic composition or content”, that is, sound’s incorporeal affects, or metaphorical effects, is by no means less important (Goodman 71). The metaphorical level of sound makes it so that it cannot be simply understood through its material surface. Moreover, the voice, as a subset of the broader term “sound”, presupposes a strong relation to human action, not only because of humans’ vocal organs but also due to ideologies that determine the production of voice. The ideologies of a singer or of a speaker are both internal (i.e.: the vocalist) and external (i.e.: social discourse). For example, a singer’s musical content, singing style, and musical persona are all consciously determined by multiple elements. Even the ostensibly objective technology is about ideology, because the technical mediation, modification, and transmission of voice all reflect embedded traces of ideologies. For instance, the way in which 1930s female singers made use of modern sound

technologies to adjust their voices to match the “politically required” melody or the way musicians redesigned songs for political reasons, and the kinds of voice that were allowed to be reproduced and transduced are all far from natural. Through material transmissions and ideological redesigns, new sonic qualities and meanings were produced.

This all demonstrates that production and reproduction of voice make up an ongoing process involving human listening and reception, technology, socio-cultural practices, and other aspects. During this process, the boundaries between the internal and external, the self and the other, the active and the passive, and the individual and the collective are constantly blurred and questioned. The voice is not “the result of an action”, but is itself “the action” (Eidsheim 24). It is always unfolding, coming into being rather than standing still as an established being. This thesis aims to free voice from its ontological status and reframe it as an ongoing activity, through which a set of complex relations can be deciphered.

What is more, the dynamic, flexible, undecided, and multidimensional elements of voice are always interwoven, which also complicate gender relations. In 1930s Shanghai, the female voice contributed to creating a diverse acoustic landscape, including the speaking voice in film soundtracks, the singing voice in songs, the broadcasting voice on the radio, and so on. The voice worked both within and beyond the medium of film and both existed in and fled from image. This dynamic and unstable attribute of voice prompted us to reconsider the female singer as belonging to a special category that cannot be simply subsumed into the visual group of actresses. Nor can the

songstress be confined to the phrase “female star”, given this term’s implication of vision-centred iconicity. When the voice became an independent, mobile object that carved out its own discursive space, what new meaning did this sonic space generate? What kinds of relationships did the voice forge with the star image in the 1930s Shanghai music and cinematic fields? How did the voice’s intrinsic ambivalence work with a set of other related material as well as ideological relations to reshape gender relations on and off screen, inside and outside the ether? Is it possible that the voice’s ability to obscure borders between the self and alterity, subject and object, enriched gender interaction that was until that point dominated by masculine politics? The many possibilities for voice and its interaction with other elements ultimately prompted me to select the voice as an alternative lens through which to explore gender issues within the rich visual and sonic environment of 1930s Shanghai.

Chapter one mainly focuses on two films, *Two Stars in the Milky Way* (银汉双星 1931) and *Street Angel* (马路天使 1937), to analyse how the female voice as an unstable force engaged with patriarchy within the film. Sound and visual technologies facilitate a Hollywood-style synchronisation, which ties the female voice to the body and reduces the songstress to an object of male gazing and listening. However, this masculine rule of embodying voice is not immutably successful. In revisiting the singing skills and the various song-image-narration relations in films, this paper argues the following: 1) narrative and camerawork occasionally detach the voice from the body, indicating a more fluctuating body-voice relationship; and 2) technological mediation provides the female voice with more provocative and interpretative possibilities, turning the voice

into an open site that accommodates different subjectivities and transcends masculine-imposed aural and visual containment. In this way, the female voice becomes an index of blurred active-passive boundaries and develops a space for the songstress' self-reflexive consciousness.

Chapter two examines how the multipotential qualities of the female voice move beyond film, forming an ambiguous relationship with official discourse. The female voice's tender and sentimental quality is a construct developed by the Nanking National Government as a vehicle for national salvation and the ideal of New Woman. This construct was facilitated by the rise of radio broadcasting, which enabled the government to gather individuals together through a mode of collective listening. In this way, the female voice on the radio became an extensive megaphone of patriarchy. However, the female voice's uncertainty and ambiguity challenged patriarchal instrumentalization. First, it is important to note that the mode of listening to a female voice on the radio presupposes prior privatisation and individualisation, which questions the governmental ideal of a shared listening media. Second, the voice's emotive timbre and unclearness eluded the rhetorical framework stipulated by the government. Threat and anxiety caused by these elements of the female voice pushed the government to illegalise it as a decadent noise that should be censored and banned. However, again, the voice's multipotentiality obscured its nature, worked with multi-media channels, and ultimately escaped the censorship.

## Chapter 1 The Voice as An Eclectic Site:

### The Female Singer, Body, and Sexuality in 1930s Shanghai Films

#### 1. The synchronisation rule that embodies the female voice

A young teahouse singer occupies the centre of the scene, her eyes open and her mouth a thin line. Her sorrowful stare reflects how reluctant she is to believe that the male customer forcing her to sing is her beloved boyfriend. However, she has no choice but to follow the order. The camera moves as she walks downstairs, and then zooms in on her face when she starts singing “Wandering Songstress (天涯歌女)” (Figure 2). The close-up clearly shows her knit brows, her braided hair, the crossed fingers, the tears falling down her oval face, and the hurt look in her doe eyes – all of which depict an innocent, weak girl (Figures 3 and 4). As she sings, the camera shifts between the weeping girl and the angry man goggling at her (Figure 5).



Figure 2. Xiao Hong walks downstairs.

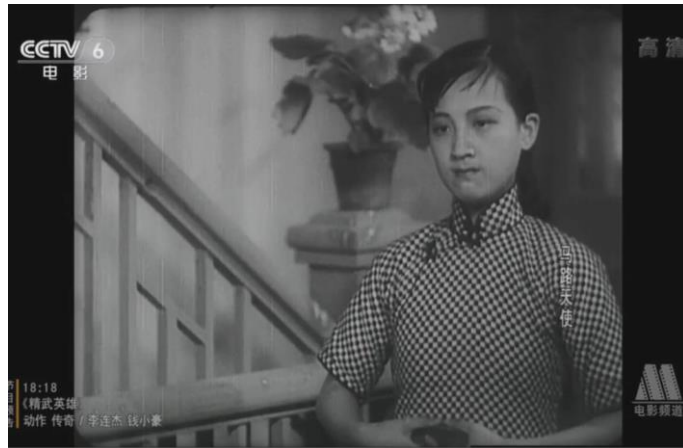


Figure 3. The camera zooms in on her face.



Figure 4. A close-up of her crying face.



Figure 5. Xiao Chen gazes at Xiao Hong angrily.

The above descriptions depict the third song scene in Zhou Xuan's 1937 masterpiece *Street Angel* (马路天使). Due to a quarrel, the male protagonist, Xiao Chen,

angrily rushes to the teahouse and orders a song from the female protagonist, Xiao Hong, a singer there. The camera's shifting pattern shows how the dominant man gazes at the female body and how the powerless girl is looked at. This is a common narrative strategy mirroring the traditional Hollywood cinema style because of the China-West contact, which included the introduction of Hollywood narrative movies and its subsequent influence on Chinese filmmaking and even social life. Many 1930s Chinese films are actually a synthesis of foreign and indigenous modes, as many scholars note.<sup>1</sup>

Considering the great impact of Hollywood films on 1930s Chinese filmmaking, we can say that the shooting practice in *Street Angel* follows what Laura Mulvey in *Visual Pleasure and Narrative Cinema* calls the traditional Hollywood representational logic. According to Mulvey, in traditional narrative films, the man always acts as an active gazer (voyeur) while the woman is always forced into an "exhibitionist role" (837). This sexualised female fantasy, following "a narrative centred on the hero's point of view", makes women "mysterious, ultimately elusive, fascinating and destructive" (53). Mulvey believes that all cinematic representations of women do not signify women themselves but exhibit "a phallic woman", and spectators simply take this representational logic for granted (12).

---

<sup>1</sup> Many works and articles have already addressed the continuity between China's filmmaking in the 1930s and Hollywood traditions in terms of character persona, narrative strategies, and camerawork techniques. For example, In *Hollywood and Shanghai Cinema in the 1930s*, Adrian Song Xiang stress that western female star Mary Pickford's "innocent and lively" star persona was adapted by Chinese filmmakers in *Song of the Fisherman* (1934) as the character of Xu Xiaomao, a rustic songstress with unspoiled innocence, and was later inherited by the teahouse songstress Xiao Hong in *Street Angel* (3). Ma Ning's famous piece *The Textual and Critical Difference of Being Radical: Reconstructing Chinese Leftist Films of 1930s* discussed about how Chinese filmmakers in the 1930s viewed cinema as a western intervention while striving to adopt indigenous art forms to pander to Chinese audience. Paul Pickowicz's chapter *The Theme of Spiritual Pollution in Chinese Films of the 1930s* also deals with Hollywood's residual influence on China's 1930s film production.

While Mulvey mainly focuses on female bodies to reveal the whole masculine meaning-making system behind images, Kaja Silverman introduces another significant layer: the voice. According to Silverman's *The Acoustic Mirror: The Female Voice in Psychoanalysis and Cinema*, traditional filmmaking adheres to a strict synchronisation rule that requires a perfect unity between characters' images and voices. This rule involves at least two key technologies: the synchronised sound technology, which perfectly merges sound with images; and the dubbing technology, which avoids delay or mismatch caused by live band accompaniments. Both technologies were mature in China in 1937; therefore, all the three song scenes in *Street Angel* – be it the first scene where the female singer sings for several local gangsters; the second scene where the hero and heroine provide a duet performance, with the man playing an instrument and the woman singing; or the last scene that shifts between the singing woman and the irritated man – seem to synchronise sound with images, and thus conform to people's expectation of what male and female voices should be.

Once synchronised with the body, the female voice is anchored in the image, and an automatic, simultaneous transferring system is ultimately formed. As Kaja Silverman states, the female voice is closely related to “a stable visual representation” and “a predictable cluster of attributes”, as exemplified by the correspondence between Xiao Hong's girlish voice, innocent appearance and miserable displacement due to Japanese invasion (166). The female voice in the film, therefore, becomes what Silverman calls the “embodied” voice, that is, the voice localised and anchored in the body (image). This embodied female voice, far different from the “disembodied male

voice”,<sup>2</sup> ultimately excludes the female singer from “discursive power” and constricts her to a safe position within in the film, that is, under the male character’s visual-auditory double surveillance (Silverman 164). In *Street Angel*, we can say that the female singer Xiao Hong experiences double gazing and listening from both the characters in the film and the spectators outside the film. Her voice is embodied by the synchronisation rule, and her body becomes a semic code. Accordingly, the female singer herself becomes the synonym of corporeality and specularity.

In a word, if Mulvey affirms that the body is forever culturally mediated, Silverman includes the vocal aspect and reveals how the embodiment process establishes “the diegetic containment of the female voice... [the] confinement to the body, to claustral spaces, and to inner narratives” (Silverman 45). Mary Ann Doane in *The Voice in the Cinema: The Articulation of Body and Space* also criticises the strict synchronisation rule in the motion picture industry, stating “[T]he combination of sound and image is described in terms of ‘totality’ and the ‘organic’.... In the discourse of technicians, sound is ‘married’ to the image and, as one sound engineer puts it in an article on post-synchronization, ‘one of the basic goals of the motion picture industry is to make the screen look alive in the eyes of the audience...’.” (35). Silverman and Doane’s

---

<sup>2</sup> Silverman endeavours to demonstrate that male voices in films are always bodiless. For example, male voiceover “occurs most frequently in police thrillers and prison dramas of the ‘B’ variety” (163). This disembodied male voiceover represents “superior knowledge” and “eventual justice”, and thus possesses an “invisibility, omniscience, and discursive power” (163). However, the female voice seldom functions as voiceover, and it is always “matched up in some way, even if only retrospectively, with the female body” (Silverman 164). Even when “it is transmitted as a voice-off, the divorce is only temporary; the body connected to the female voice is understood to be in the next room, just out of frame, at the other end of a telephone line. In short, it is fully recoverable (Silverman 164).” In sum, in Silverman’s opinions, male voices in traditional narrative cinema are disembodied, authorial and exterior, while female voices are embodied, representing impotency and (diegetic) interiority.

psychoanalytical model resonates with many other scholars, such as Guy Rosolato<sup>3</sup> and Jean-François Lyotard.<sup>4</sup> In the case of *Street Angel*, we see how soundtrack technologies (e.g. synchronisation and post-dubbing technologies) facilitate the voice-image union to pander to the male characters' and the spectators' needs.

However, if we re-examine the above song scene's visual and sonic elements, we can draw a different conclusion from that of Mulvey and Silverman's. On the visual level, Xiao Hong's tight frown, pout, and sad tears show that she fully knows that she is being made an object of male gazing and listening and that she condemns this masculine power. On the vocal and aural levels, in contrast to her mellifluous voice and ascending tone in the first song scene when she is still in love with her boyfriend, Xiao Hong's voice in the third song scene is trembling, hoarse, and desolate, expressing her reluctance and resistance. By devoting clear close-ups of the girl's powerlessness and muting singing skills, the scene not only displays masculine gazing and listening but also simultaneously resists them. Based on this alternative understanding, we can say that the scene is more a disavowal than a paradigm of Mulvey and Silverman's

---

<sup>3</sup> Rosolato argues in his essay *La voix: entre corps et langage*: "The harmonic and polyphonic unfolding in music can be understood as a succession of tensions and releases, of unifications and divergences between parts which are gradually stacked, opposed in successive chords only to be resolved ultimately into their simplest unity. It is therefore the entire dramatization of separated bodies and their reunion which harmony supports (Cited by Doane 44)." Here I refer to Doane's translation of Rosolato. For the original French version, see Guy, Rosolato: "La voix: entre corps et langage." *Revue Française de Psychanalyse*, 38.1 (1974): 75.

<sup>4</sup> Jean-François Lyotard in *The Unconscious as Mise-en-scene* stresses the importance of bodies to aggregate "multi-sensory potentialities": "...the mise-en-scene turns written signifiers into speech, song, and movements executed by bodies capable of moving, singing speaking; and this transcription is intended for other living bodies-the spectators-capable of being moved by these songs, movements, and words. It is this transcending on and for bodies, considered as multi-sensory potentialities, which is the work characteristic of the mise-enscene. Its elementary unity is polyesthetic like the human body: capacity to see, to hear, to touch, to move. (24)" Based on this, he argues that performances are "linked to the idea of inscription on the body" (Lyotard 24).

“persistent male dominance” models.

Following such a relational perspective on vision and sound, we can further question Mulvey and Silverman’s male-coded spectatorship: Is Silverman’s synchronisation rule always perfectly conducted? Is the voice easily matched to the image without any independent function? Is the female singer completely suppressed by the impregnable male gazing and listening without any possibility of agency? Section 2 of this chapter re-scrutinises the relationship between voice, image, and narrative, and argues for more dynamic, multidimensional gender relations in films other than the singular logic of repression.

## **2. Revisiting the Voice-Image-Narrative Relation:**

### **The Detachable Voice and Female Self-Awareness**

#### **1) The fluctuating relationship between the voice and the image**

Some theorists (e.g. Mulvey) often neglect the voice. Others (e.g. Silverman and Doane), although alleging to discover the audible part of cinema, still perceptually refer back to the image and finally corporealize the voice. The latter’s emphasis on how the voice is synchronised with the image in films implies that the voice can be easily and perfectly fused with the image, without considering the voice’s uniqueness that can complicate the voice-image relationship.

However, the synchronisation rule proposed by Silverman and Doane is not always successful; rather, the voice often escapes corporeal confinement and becomes a flowing, independent element. For instance, the ancient Western legend of the Sirens’

enchancing voice luring sailors to destruction and modern operas like the lingering voice in *The Phantom of the Opera*<sup>5</sup> are prime examples of the intrinsic detachability of voice. This phenomenon of voice transcending body also appears in Chinese literature and films where male protagonists first hear the voice of female protagonists before seeing them. The asynchronous voice, free from the limitation of vision, denaturalises the automatic and simultaneous sound-image transferring system and epitomises vocal power over image.

Silverman does admit the existence of asynchronous female voices in films, but she also stresses that such voices “are often coded as a partaking in memory or imagination”, and their separation from bodies are always temporary. Indeed, on some occasions, the divergence of female voices from their bodies is quite transient. For example, in *Two Star in the Milky Way* (银汉双星 1931), a(n) (all-male) film crew is attracted by a distant, melodious singing voice. They follow the voice and finally locate its owner: a beautiful girl named Li Yueying (李月英). Only after seeing the female singer’s striking beauty do they decide to cultivate her as a professional singer and actress.

However, when it comes to the ending, the narrative transgresses the synchronisation rule: Yueying breaks up with the actor Yang Yiyun (杨倚云) because he is already married. Heartbroken, she moves back to her house in the countryside with her father. The ending shows the male protagonist standing in front of Yueying’s

---

<sup>5</sup> It is noteworthy that the idea of the lingering voice without being seen in *The Phantom of the Opera* was utilised by the Chinese film *Song at Midnight* (夜半歌声 1937), which also deals with a singing voice separated from the specular body.

house and listening to her singing (Figure 6). The camera zooms in to make a close-up: the man stares at the house and frowns, his lips quivering and his eyes saturated with deep grief (Figure 7). After listening for a long time, he finally turns around and leaves, with the camera panning his fading image from behind (Figure 8). During this process, several scenes of the interior space of the house set in, presenting the female singer sitting beside a table and singing (Figure 9). Notably, compared with the long take on the man, scenes of Yueying are much shorter. Moreover, the camera makes clear close-ups of the man's facial expression to show his strong emotions, but scenes of the songstress are almost static—only the movements of her mouth indicate that she is singing. These differences show that the narrative concentrates more on the female voice, which is accompanied by a live band, and its impact on the man. The man's final decision to leave without letting the woman know indicates that the female voice has become inaccessible to the male gaze. This woman, singing without being seen, disrupts the specular regime of the female body. The ending illuminates how the narrative and the camerawork split the female voice with the image.



Figure 6. The hero stares at the songstress' house.



Figure 7. The man turns his head away.



Figure 8. Close-up of the man from behind.



Figure 9. The songstress sits and sings in her room.

Aside from the narrative and the camerawork, the use of sound technologies, which according to Silverman and Doane should conspire with the camera's synchronisation rule, sometimes separates the voice from the body. In the second half of *Two Stars in the Milky Way*, the film company holds a party to celebrate the commercial success of the film starring Li Yueying and Yang Yiyun, and one man turns on a huge radio placed in the centre of the hall and asks guests to listen to the song sung by Yueying. Yueying also listens to her replayed voice without surprise, suggesting that she is fully aware that the technical operation can separate her voice entirely from the body (Figure 10). This moment of the female singer becoming her own listener, as Jean Ma convincingly analyses, strongly demonstrates that the voice "never really belong[s] to this body in the first place" (118); rather, the voice is reproducible and detachable. In this way, the film opens a fissure within the audio-visual unity and powerfully denies the standard synchronisation that "binds the voice to a body in a unity whose immediacy can only be perceived as a given" (Doane 47).



Figure 10. The man and woman listen to the songstress' voice via the radio.

The voice's departure from the body in *Two Stars in the Milk Way* even becomes a

legacy for later films. For example, in *Songs of the Peach Blossom River* (桃花江 1956), there is a scene where the female singer, Little Wildcat (小野猫), sings to a microphone, while the male protagonist, Li Ming (黎明), records her voice. This scene demystifies the recording mechanism to the audience, showing that the voice they hear is “an effect of mediation” (Ma 118). More importantly, it demonstrates the split nature of the voice. Just as Jean Ma says in her analysis of this film:

No longer the property of the visible body, the voice announces its machinic origins and lays bare the composite nature of song performance as a trick of the playback process, along with the status of the songstress’s body as a chimeric sign, divided across the visual and audial registers (119).

From the above analyses, we realise that audio-visual unity is by no means self-evident. Firstly, by showing the man listening to the songstress without seeing her, the film’s narrative avoids localising the voice, and thus violates the synchronisation rule. Secondly, the scene where people listen to the songstress’s replayed voice demonstrates that the voice is not inextricably embodied or imaged. Either the narrative or the scene presenting the mechanical operation of the voice proves that the voice-image relation is much more fluctuating, dynamic, and diverse than Silverman and Doane think. The two scholars’ problem lies in their overemphasis on the vulnerability of the female voice to “corporeal encroachment” and enslavement (Silverman 49). As such, they neglect “the most intriguing and provocative characteristics of the voice—its fluidity, its non-specularity, its heterogeneity”, as stressed by Britta Sjogren in her book *Into the Vortex: Female Voice and Paradox in Film* (13). In her book, Sjogren criticises that Silverman, Doane and Mulvey all adhere to a problematic visualist model. Under such model, “the

'female' is always only a body, even when she seems represented as a voice. In Silverman's schema, then, the voice is always referred back to some visual signifier, reduced to it through a logic that understands the image as the ontological centre of cinema, one which determines and rules its aural 'accompaniment' (48)." In doing so, Silverman and Doane ontologise the image and specularise the voice, thereby reinscribing as much visual constraint on the voice as Mulvey does.

## **2) The self-reflexive consciousness of the female singer**

In Mulvey, Silverman, and Doane's model, patriarchy in traditional Hollywood cinema is so monolithic that female subjectivities seem impossible. However, when re-examining 1930s Chinese films, mainly gechang pian(歌唱片) or song films, we find portrayals of female consciousness, which sometimes even becomes strong enough to reverse the male-female hierarchy.

This nascent female consciousness often sprouts from the character's self-awareness of their singing voice and its detachability. For instance, at the celebration party in *Two Stars in the Milky Way*, the female singer listens to her own voice from the radio. Such self-awareness of her detachable voice becomes stronger when she later criticises her own voice, saying "This singing makes one feel sad... don't listen to it anymore". She then asks the man to go outside with her, leaving all the audience behind in the hall. Even in the garden, the man is still immersed in the desolate song and asks the woman, "Why don't you listen to that song?" The woman retorts, "Is it not a song of disappointed love?" This plot displays the songstress's strong self-awareness of her

own musical performance, star persona, and individual identity. The moment Yueying actively rejects her voice from the radio marks her return to her own personal identity. The framework of the character in the film is far from a cohesive whole that can completely absorb the female singer who knows quite well her freedom to resist it.

With such a self-reflexive awareness of her own extradiegetic identity, the female singer plays a more active role in her intimate relationship with the man. In the garden, when the two lovers hug and get closer to kiss, the man, thinking of his own wife, suddenly turns his head away, even though the woman is eager for his kiss (Figure 11). In this scene, the woman's desire is more genuine and stronger, while the man is much more self-constrained, especially after listening to the songstress's sorrowful song. This film also devotes more scenes showing how the woman "counter-gazes" at the man instead of the contrary. For example, during their first encounter, the woman curiously fixes her eyes on the handsome man, but the man does not see her. This female counter-gaze again reflects strong female agency and desires.



Figure 11. The man refuses to kiss the girl.

The narrative is far from overriding or coherent, nor does it impose a solid framework onto the woman. The female singer's free oscillation between her star persona and individual identity, her knowledge of the detachability of the voice and her ultimate abandonment of her stardom<sup>6</sup> all reflect female self-reflexive awareness, which renders the patriarchal hierarchy unstable and incoherent.

### **3. The voice as a multidimensional site**

The first two sections of this chapter demonstrate the unstable voice-image relationship that can breach enduring patriarchy. This section delves deeper into the essence of the voice itself to reveal its multipotentiality.

Silverman and Doane align the voice with vision (i.e. the body) and later centre vision into masculine subjectivity, as described by Mulvey. Following this logic, they suppose that subjectivity is disembodied/bodiless, in Sjogren's words, "imaginary/visual" (27). However, the characteristics of this ideal "subjectivity" are almost tantamount to their depiction of the man. Accordingly, they can never decipher female (i.e. singer and spectators) subjectivity, since their premises on subjectivity are already more germane to masculine subjectivity. It is this perspective that prompted Sjogren's critique of Silverman and many other feminists' theories, referring to them as a "masculine-coded" template of subjectivity.<sup>7</sup>

---

<sup>6</sup> Yueying firstly temporarily abandons her voice when she leaves the celebration party, and finally relinquishes her star career after breaking up with the man.

<sup>7</sup> Sjogren states that Silverman and many other theorists rest on "the model of sadomasochism suggested by theories of spectatorship centered on the cinematic image alone" (77). Their understanding of subjectivity is "based on the primacy of the phallus as a visible signifier, one that seems unlikely to account for an experience of feminine subjectivity, or for the effect or experience of the acoustic register of film (13)."

The underlying assumptions of this model are as follows: (1) Within in the voice, the relationship always and only involves one subject dominating one object; (2) one's subject can only either be the self or the other, the active or the passive; and (3) only when the active triumphs over the passive can the relationship be kept in a relative harmony. This logic of opposition and repression contains at least two problems: Firstly, the definition of "subjectivity" is ruled by an "either-or" thinking. Secondly, the voice is regarded as only allowing one type of relationship of repressing and being repressed, leaving no possibility for contradiction nor difference. On the contrary, I argue that subjectivity can transcend the subject-object opposition, and the voice is a site where multiple relationships between different subjectivities coexist.

The voice is not situated on any single side of either subject or object; instead, it has an "in-between" status, a "middle voice". The term "middle voice" was explicated by Hayden White and later nuanced by Sjogren. In his 1990 essay *Historical Emplotment and the Problem of Truth in Historical Representation*, White proposes to recover the "middle voice" in modern historical writing.<sup>8</sup> This "middle voice" dissolves the opposition between "agency and patiency, subjectivity and objectivity, literalness and figurativeness, fact and fiction, [and] history and myth" (49). As such,

---

<sup>8</sup> White's argument relies heavily on Roland Barthes's concept of the "middle voice". In *To Write: An Intransitive Verb?* (1970), Barthes insists that our modern writing is always intransitive, without a direct object. If "to write" is transitive with a direct object (like "I write an essay") "the agent is not interior but anterior to the process of writing: here the one who writes does not write for himself, but as if by proxy, for an exterior and antecedent person." However, Barthes states that this is a dichotomy that assigns interior/exterior to writer/text respectively. To refute this binary, Barthes asserts that modern writing always uses "middle voice", a tense between active and passive, which suggests that "the subject is constituted as immediately contemporary with the writing, being effected and affected by it" (Barthes 18-19). From this perspective, "to write" is no longer transitive but intransitive, because the writer cannot separate the "writing self" with the "written self". The subject is always affected by the action of writing. What is written (the text) is no longer the exteriority and object, and the writer is no longer the interiority and subject. In this way, Barthes dissolves the interiority/exteriority and subject/object binary.

the writer is no longer the self, and the text is no longer the object. Writing in the “middle voice” means the writing subject transforms its relation to the world, that is, the writer remains inside the action of writing such that the subject-object dichotomy no longer exists.

White’s contribution is that he proffers a new method to understand the status of the subject. His concept of the “both active and passive” writing subject is later situated by Sjogren into the cinematic field to analyse the female voice:

As opposed to an expressly passive/active split, where “either the subject or the object remains outside the action,” in the middle voice “the distinction between subject and object is obliterated.” Thus, a subject need not, as in the “voyeur” model of spectatorship, be “exterior” to the object or “distanced” (nor indeed, is the “exterior” narrative region thus privileged as the arena of agency), for a model of subjectivity does exist in which “interiority” connotes a special heterogeneous agency in which “subject and object . . . are in some way conflated (Sjogren 71-72)

“Neither passive nor active, but between—the voice is both simultaneously”  
(Sjogren 26)

Sjogren negates a longstanding active/passive dualism in which passivity is always translated into “objectification, impotence, non-(self)expressivity” and non-agency, and females are always attributed the passive “diegetic interiority” (Sjogren 72; Silverman 142).

In contrast to this binary, the “middle voice” suggests two potentials of the female voice: Firstly, the middle voice fuses subject with object, and generates what Sjogren calls a “dialectical self-consciousness”, which “makes possible a fusion of mental entities, which would be judged merely contradictory by the standards of ordinary logic”

(cited by Sjogren 73). The problem of self-consciousness or subjectivity becomes a problem of synthesis. In other words, the female singer no longer needs to pursue a “pure”, “unadulterated”, and completely autonomous female subjectivity to gain agency; rather, her subject inhabits different consciousnesses not necessarily emanating from herself. These different consciousnesses cannot be simply categorised as either object or subject, as they interact in a dynamic relationship. Once the subject and the object become indistinguishable, women are able to gain their own subjectivities.

Furthermore, the voice becomes a broader site of synthesis, which preserves whilst transcending difference and contradiction. This synthetic “middle voice” does not obliterate difference as Silverman or Doane assumes; rather, it invokes and allows multiple subjectivities to coexist within the same univocal space, conflicting, conflating or interacting with one another. Just as Steven Shaviro argues, “the voice always stands in between: in between body and language, in between biology and culture, in between inside and outside, in between subject and Other, in between mere sound or noise and meaningful articulation” (cited by Whittaker and Wright 3). From this perspective, the voice is not singular but plural, eclectic, open and even ambiguous, providing a plurality of views in films. Such multipotential qualities establish the voice as means to carve out ruptures within the seemingly firm patriarchy.

Following this dialectic “middle” voice and its new definition of subject, we can reconsider song scenes in *Two Stars in the Milky Way*. As previously mentioned, both male and female protagonists display strong desire for each other, and oftentimes, the female singer shows an even greater desire than the man who constantly struggles with

an ethical dilemma. While listening to the female voice, their desires converge and conflict within the univocal space. For example, the songstress's desolate voice from the radio triggers a violent inner struggle for the man: his responsibility towards his wife and his desire for the songstress. However, to the songstress, the melancholy embedded in the voice serves as a foil to her current happiness, and thus should be expelled from their intimate world. The lovers' different emotions and consciousnesses are firstly evocated by the voice, and then overlap and clash within in the same sonic space. The voice invokes both the internal (i.e. female desire) and the external (i.e. male desire), and later incorporates them into its own discursive zone, demonstrating the strong discursive capability of the voice and the female agency in it.

Such plural, inclusive, and dynamic qualities of the female voice are more apparent in the third song scene of *Street Angel*, which has been described at the beginning of this chapter. At first glance, the stark contrast between the ireful man and the ostensibly compliant female singer, to some degree, reflects Mulvey's male-coded spectatorship. Re-examining the relationship between the song and the images, however, we learn that the voice contains plural consciousnesses. As the female singer starts singing, several scenes reminiscent of their romantic love in the second duet scene, including that of the girl singing while feeding the bird and the man playing the instrument, are superimposed on the girl's weeping image (Figures 12 and 13). These same reminiscent scenes are later superimposed on the man's angry and contemplating face (Figures 14, 15, and 16). Here, the voice not only links the present with the past, but also bridges the conscious levels of the man and the woman. Their thoughts and emotions overlap

and conflict simultaneously within in the sonic zone. Additionally, the same scenes superimposed on their background images make both female and male consciousness more indistinguishable, and thus endow spectators with more possible identificatory processes. The voice's concurrent evocation of both male and female subjectivities indicates the coexistence of difference and contradiction, impairing the necessity of one-sided containment, ascendance, and obedience. For this reason, we can transform the previously unidimensional male gazing and listening template into a more multidimensional, dialectic, and dynamic site where the possibility of female agency lies.



Figure 12. Xiao Hong's crying face superimposed by her memory of feeding the bird.



Figure 13. Xiao Hong's crying face superimposed by the man's image.



Figure 14. Xiao Chen's contemplating face overlaid by the songstress' image.



Figure 15. Xiao Chen's contemplating face overlaid by his own image.



Figure 16. Xiao Chen's face superimposed by scene of his past with Xiao Hong.

Both female voices in *Two Stars in the Milky Way* and *Street Angel* manifest their ability to evoke “subjectivity in difference”, activating “two (or more) subjectivities in sustained, contradictory signification” (Sjogren 74). In representing the interior and exterior, active and passive, subject and object, subjectivity in difference transcends Mulvey's voyeur model of spectatorship and the singular paradigm that confines men to the active role and women to a receptive position.

I do not advocate for a distinctive female subjectivity nor total empowerment of women in these films. What I emphasise is the multiple gender interactions and interplays expressed by the combination of images, sounds, and narratives. It is from these multi-layered aspects of gender relations that we can envision female agency and interpret females' position in films from a more positive perspective.

In sum, the psychoanalytical epistememes of Mulvey, Doane, Silverman and many other theorists (e.g. Jean-Louis Baudry's exposition of the cinematic ideology) still tend to restrict sound to images. This tendency is exacerbated by sound technologies that

emphasise perfect replication of sound without distortion. The pursuit of sound fidelity prompts people to regard sound as the reality, the present, the here and now. Sound is viewed as a mere aural accompaniment—a supplementary to narratives and characters’ images—without its own autonomous arena. Such a vision-centred notion results in the omission of sound in films. Moreover, since the visual realm is perceived as exclusively masculine, this vision-based and male-coded episteme tends to conclude with a monolithic, male-coded spectatorship where females can never speak or be heard.

However, just as Jonathan Sterne says in *The Audible Past*, given that “intricate technology serves sound recording, editing, and re-recording”, sound is never neutral but consciously crafted (5). The recorded and replayed female voice in films, therefore, reflect not the reality or the present but people’s impression of reality. A cultural artefact rather than a natural outcome, sound goes beyond its secondary role in films as mere affective stimulant and thematic support. Its unique characteristics, such as fluidity and ambiguity, enable it to exceed its corporeality, or we may call the “flesh” of the voice.

Given the artificial appropriation and the intrinsic essence of sound, I decide to reconsider the sound-image relation in films. By extending the visual language of films in combination with the female voice, this chapter argues that (1) image depiction, vocal quality, and narrative work together in films to show females’ self-awareness of and resistance to male objectification; (2) the way films display the mechanical operation of the female voice indicates its detachability from the body, which violates the (masculine) synchronisation rule; and (3) in cooperating with the text and vision, the multidimensional and dynamic female voice, which reflects the “middle voice”,

ruptures coherent patriarchy and complicates gender relations in films.

All these analyses reveal how 1930s Chinese films combine the story, image, and voice, transcend a male-coded, binary model of spectatorship/subject, and form a new dynamic relationship. Borrowing Sjogren's words, multiple cinematic elements work "in a vortexical fashion" (15). Mulvey and Silverman do envisage such a relational perspective in cinema, but they believe that this amalgam can only be found in alternative cinematic practice beyond traditional Hollywood films (e.g. avant-garde and experimental feminist films). I argue that in China in the 1930s, a transitional period from the silent era to the sound-film era, such organic combination of vision, voice, and narrative and subsequent female agency already surfaced. What I try to do in this chapter is to utilise a relational perspective to explore how the voice, once adopted as a new layer in cinema analysis, can complicate our understanding of gender relations in films.

## **Chapter 2 A Consolidating or a Destructive Force?**

### **— The Paradox of the Replayed Female Voice Outside of Film**

The first chapter has analysed how the multipotential female voice functions within film and forms a dialectical relationship with filmic images (the body in this case). However, the voice is not restricted to the cinematic realm; rather, it is a mobile, dynamic object that can flee from the medium of film, gaining an independent position in society. In 1930s Shanghai, the emancipation of the female voice from film was mainly facilitated by radio broadcasting and the record business. Once set free from film, the female voice could be redesigned, remixed, reproduced, and replayed, all of which transformed its original interiority and amplified its elusive nature. The reproduction and replaying of the female voice did not only involve problems of source and replica; rather, as this paper demonstrates, vocal reproduction became a new kind of production, triggering a host of new political, economic, and gender issues that could stretch beyond the original patriarchal framework.

### **Recording and Replaying:**

#### **The “Freed” Female Voice as a Versatile Tool**

The emphasis on sound actually started before the sound-film era. Many exhibitions of films in the 1920s included singers' live performance accompanied by live bands, and only after these singers finished their performances would the films continue. For example, for promotion of the film *A Crushing Blow* (刘关张大破黄巾 1927), an advertisement on *Shun Pao* (申报) particularly stressed the female

protagonist and songstress Hu Die (胡蝶)’s “melodious singing and soft and beautiful voice (歌喉之婉转, 舌音之柔美)” in her live performance to attract more of the public. The huge commercial success of these films “interrupted” by songstresses’ live performances demonstrated the audience’s fascination with the listening experience and heralded the voice’s later liberation from film.

In the 1930s, the maturation of sound recording and preservation technologies set the voice free from film and motivated producers to sell songs as independent products. Recording companies prospered and competed with one another to recruit the most popular female singers, the majority of whom were cross-over stars who both sang and acted. Record companies even cooperated with cinemas, erecting stalls at the entrances to convince moviegoers to buy records of the theme songs. In this boom of sonic products, female singers began to eclipse their male counterparts in terms of record releases, sales, and popularity. According to *The Inventory of the Template of Old Records of Chinese Recording Factory* (中国唱片厂库存旧唱片模板目录) published in 1964, with 150 records, Zhou Xuan became the most recorded songstress in the first half of the 20<sup>th</sup> century (Chinese Records Club 247-251). Other examples include Yao Li, who recorded 130 songs (253-256), and Gong Qiuxia, who recorded 67 (264-266).

Concomitant with songstresses’ record releases were the emergence of wireless transmission technology and the subsequent rise of the broadcasting station. Between February 1934, July 1937, and January 1939, Shanghai had 41, 29, and 35 broadcasting stations, respectively, that were officially approved by the government (cited by Ge 152). Apart from these legal broadcasting stations, many underground stations also

flourished, creating the busy “traffic condition” of airwaves. Recorded songs accounted for a large part of the programs of these radio stations. According to the statistics provided by Tao Ge in his *Records and Modern Shanghai Social Life*, in 1937, the 28 radio stations in Shanghai broadcasted a total of 341.25 hours, and the time spent playing recorded songs was 49.8 hours, which took up 14%6 of the total broadcasting time (153). Among these radio stations with music programs, two of them broadcasted four hours per day, three broadcasted three to four hours, and six stations broadcasted two to three hours (Ge 153). In the early 1930s, some stations like the Chinese Broadcasting Company even specialised in music programs. This overwhelming radio boom eventually replaced the gramophone as the public’s major listening medium.

It is notable that the entanglement of cinemas, record companies, and radio stations to promote the newfound power of sound and song was mainly embodied in female singers. Besides the sheer number of records by female singers that existed, many advertisements in newspapers and magazines foregrounded the songstress’ beautiful silver-toned voice in film, and radio stations invited popular female singers to perform on their programs. This multidimensionality improved the marketability of female singers and ushered in new types of audio-visual contracts. As a result, the female voice was endowed with an independent position and a spectral afterlife beyond film.

Once extracted from film, the singing voice could be pinpointed, redesigned, and reconstructed more easily and independently. The voice’s sonic characteristics, be it the pitch, timbre, or rhythm, though still related to the original film story, were reoriented for a wider audience, which was not necessarily made up of moviegoers. The separation

of senses and the isolation of listening triggered a reconstruction of the whole acoustic space. Listening, therefore, became a discrete event that attracted extra technological, commercial, and, most importantly, political attention.

In the 1930s, due to the Japanese invasion and China's domestic turbulence, the national government started to use sonic culture to promote national salvation and personal development. Female singers were then utilised as vehicles of such propaganda. Many of them appeared in films as powerless young girls with miserable childhoods and tragic destinies, such as the aforementioned teahouse singer Xiao Hong in *Street Angel*, the tragic rustic songstress Xu Xiaomao in *Song of the Fisherman* (渔光曲 1934), and the poor songstress Red Peony in *Sing-Song Girl Red Peony* (歌女红牡丹 1931). These innocent, weak young characters became emblems of Japanese imperialism and vicious urban corruption. Their voices were recorded and repeatedly played on the radio to reach a wider community. For example, Wang Renmei (王人美)'s *Singing Girl Downtrodden* (铁蹄下的歌女), the number of *Children of Trouble Times* (风云儿女 1935), describes a songstress who is displaced by cruel war and cries for peace. Her *Song of the Fisherman* describes a helpless girl's tenacity in the face of catastrophic adversities. Chen Yanyan (陈燕燕) and Li Lili (黎莉莉)'s duet, *New Woman* (新女性 1935), the theme song of film of the same name, is a suite composed of six songs narrating the oppression of 1930s female workers and the awakening of their self-consciousness. Additionally, the two numbers of the 1935 film *The Big Road* (大路), *Song of Yanyan* (燕燕歌) by Chen Yanyan and *Song of New Fengyang* (新凤阳歌) by Li Lili, both convey females' soft sentiments and reflect resistance against the

war. In April 1932, the renowned musician Li Jinhui (黎锦晖) published *Songbook of Patriotic Songs* (爱国歌曲集), whose pieces were mainly sung by popular songstresses like Zhou Xuan, Wang Renmei, Li Lili, and other female singers.

In accordance with the image of a poor little girl, singing techniques at this time aimed to help singers imitate the innocent, genuine, tender, and lively qualities of a lass. Female singers squeezed their voices forward and kept their larynx at a high and unstable status. In this way, their voices were made to sound high and thin, and they were able to raise their vocal volume as the tone heightened, mimicking young girls' "natural shouting" (Wu and Wu 61). Many female singers in the 1930s, like Li Lili, Li Minghui (黎明晖), Wang Renmei, Hu Die (胡蝶), Xue Lingxian (薛玲仙), and Hu Jia (胡笳), shared these singing characteristics. Their uniquely syrupy voices were called the "little sister voice (小妹妹腔)". For example, in her prose *On Music*, Eileen Chang wrote:

In the past, the popular women singers in China tended to have thin and scratchy voice in a high pitch, due to the public worship for the sound of "little sister." That explained the "unusual" performance in the song *Peach Blossom River* often played by the radio amplifier, so weird to foreigners that they could not help asking why Chinese women possessed voice like that (中国的流行歌曲，从前因为大家有“小妹妹”狂，歌星都把喉咙逼得尖而扁，无线电扩音机里的《桃花江》听上去只是‘价啊价，叽价价叽家啊价……’外国人常常骇异地问中国女人的声音怎么是这样的) (199).

The "little sister" (小妹妹) here refers to Wang Renmei and Li Lili, the two original singers of *Peach Blossom River* (桃花江 1929).

Another major example of using the "little sister voice" for nationalistic

propaganda is Zhou Xuan's "golden voice". While singers like Li Lili and Wang Renmei tried to imitate young girls' natural daily singing and talking, Zhou Xuan went further in making her voice even more girlish and cloying. For instance, her two numbers in *Street Angel*, *Wondering Songstress* (天涯歌女) and *Song of Four Seasons* (四季歌), both depict a weak young girl who suffers because of brutal war and male dominance to evoke people's sympathy and promote resistance against the Japanese army. In these songs, Zhou Xuan used a nasal and breathy voice (or murmured voice) and a combination of up-and-down slides to make her voice sound whispery, mellow, affective, and even cloying.<sup>9</sup> In addition to using singing techniques to imitate a young girl, Zhou Xuan skilfully leveraged the microphone to manipulate her voice. For example, by adjusting her distance from the microphone and fine-tuning her phonation according to the mixing effect, she created a whispery and floating quality while hiding some of the weaknesses of her original voice (i.e.; the shrill quality). As a successful example of the combination of singers' singing skills and the techniques of sound technologies, Zhou Xuan laid out a clear roadmap for many later singers, such as "the silver voice (银嗓子)" Yao Li (姚莉) and the "singing queen of nasal voice (鼻音歌后)" Wu Yingyin (吴莺音) (Sun 137).

The high-pitched, saccharine "little sister voice" was partially related to the legacy of traditional Chinese opera and folk tunes as well as singers' lack of Western singing training, but the main reason that this singing technique was adopted was due to female

---

<sup>9</sup> Zhou Xuan did incorporate Western bel canto techniques to make her voice thicker in the 1940s, e.g., in her film songs *Song of Triumph* (凯旋歌 1947) and *Blossom Year* (花样的年华 1947), but her 1930s songs were nearly all characterised by the typical "little sister voice".

singers' wish to maintain a sense of girlish innocence and simplicity. This desire also aligned with mainstream masculine standards of “‘true character’ (*bense*, literally ‘originally color’) [本色] and the ‘good girl’”, all of which required starlets to exude “‘beauty (*mei*)” and “‘authenticity (*zhen*)”, as Michel G. Chang states in his article *The Good, the Bad, and the Beautiful: Movie Actresses and Public Discourse in Shanghai, 1920s-1930s* (129). Chang argues:

This “‘authenticity” was, in turn, grounded in the socially constructed ideal of women as “‘good girls” who were “‘natural” (*ziran*), “‘innocent” (*tianzhen*), and “‘genuine” (*zhen*). The construction of and adherence to this standard by directors, screenwriters, critics, and audiences as well as actresses themselves reflected an unwellness to accept and praise women for being anything except centered, knowable, and thus harmless or nonthreatening subjects (153).

The masculine influence demonstrated by Chang also applied to the sonic field, resulting in the prevalence of the “‘little sister voice” among 1930s female singers. Admittedly, powerful songs belting out national struggles did gain popularity, such as *March of the Volunteers* (义勇军进行曲) and *Lovers over the Yellow River* (黄河之恋 1937), and some female singers indeed strived to make their voices more sonorous by increasing their vocal volume, thickness, and range. However, the Nanjing National Government held complex and sometimes ambivalent attitudes towards these leftist, masculinised voices. On one hand, they could not neglect sweeping patriotic sentiments and supporters. On the other hand, they worried that the massive distribution of these songs might irritate the Japanese government (Ge 201). They also dreaded the negative allusions to the ruling government that were embedded in these leftist progressive songs. Such mixed thoughts drove the national government to ultimately carry out a strategy

of overall suppression of these songs (Ge 201). In response, progressive composers and lyricists found loopholes by integrating explicit, masculine (pre-war or wartime) discourse into their music through a more docile, nonaggressive female voice (Steen 461). Musicians also incorporated or adapted gentle folk songs into their melodies and covered up their ideologies with feminine, sentimental stories.

In brief, the confinement of 1930s songstresses' vocal qualities can be divided into at least three stages. First, the record business and radio stations peeled the female voice away from film so that it could be more easily redirected and reconstructed. Second, the government and leftist politicians soon noticed the manipulability of the female voice, and they leveraged the recording and broadcasting business to establish the female voice as a megaphone for political ideologies. Third, the socio-political and gender standards imposed themselves onto the female voice, turning it into a sentimental sonic symbol, and rendered listening a means of keeping pace with the national rhetoric. Both singing and listening became techniques that could be instrumentalised towards political ends.

It is through this three-step process that the female voice was ultimately transformed into a representational, rhetorical, pedagogical, supervising, and, ironically, deafening tool. By mass producing a syncretic form of the female voice, auditory politics transformed the nation into a gigantic music device that regulated and monopolised singing and listening practices. This overwhelmingly resonant device legalised and then infinitely amplified certain kind of songs and voice while silencing other types of music, approved certain kinds of listenership while deafening others. Just

as Jacques Attali argued in *Noise: The Political Economy of Music*, modern musical distribution techniques actually contribute to “the establishment of a system of eavesdropping and social surveillance”, because “[t]he monologue of standardized, stereotyped music accompanies and hems in a daily life in which in reality no one has the right to speak any more (34).”

In the expansion of the nation’s overarching power, the female voice played an indispensable role, permeating the construction, production and reproduction, transmission, and reception of sound. Female singers were confined to a genre-specific musical persona and a new vocal style that adamantly exuded nonaggressiveness. Localising and specifying power, female voices travelled across borders to knit together a nationalistic, collectivist sonic culture. This strategy relied heavily on modern transmission and receiving technologies, which, as the government hoped, suggested a principle of collective listening and sharing. However, this principle of collectivism is not axiomatic. On the contrary, as is addressed in the following section, the use of modern sound technology and broad-based media networks might actually reflect metaphorical counter-force to collectivism.

## **Individual or Collective Listening?**

### **— The Paradox of Listening via Radio**

The 1930s were marked by a fast-expanding broadcasting infrastructure and the domestication of radio, especially in the Shanghai area. According to *Annals of Shanghai Radio and Television*, by September 30, 1935, more than 68,000 radios in the

Shanghai area had been registered with the Shanghai International Telecommunications Bureau (上海国际电信局) (cited by Zhao, 21). In this period, the radio, previously a beeping and screeching gismo, became a common element of most people's houses.

The overwhelming boom of radio broadcasting prompted the government to take broadcasting styles, content, and voices and people's listening practices under their grasp. There was gender-based labour division in the governmental regime of broadcasting voices. The male voice was encouraged to convey passionate, inspirational content, such as an explanation of the Kuomintang Party (KMT hereafter)'s programme, while the female voice was often assigned a softer role. For example, from July to August in 1936, in order to "enhance the national thought (提高民族思想)" and "enrich the broadcasting content (充实广播内容)", the KMT bought records of "songs of the Party" and "Speeches of Celebrities" (mainly male politicians like Wang Ching-Wei) wholesale and required radio stations to play them (cited by Ge, 195). However, in contrast to the male voice's generally authentic status, the female voice mainly pervaded the entertainment field, with popular film songs describing the poor girl in the midst of a devastating war.

In order for the sentimental female voice to reverberate relentlessly in people's ears, three major aspects were necessary: the reproducibility of voice, the technical possibility of distribution, and an experience of collective listening. Both the KMT and the communist politicians believed that the technopolitical regime of broadcasting voices and the wide domestication of radio guaranteed that a voice could be fully controlled and people could listen to the same content collectively. Just as Rudolf

Arnheim proposed in his *Radio: An Art of Sound from 1936*, “An apparatus which technical peculiarity simply consists in enabling sounds made at a particular spot to be simultaneously reproduced in as many and as far removed places as one wishes by disrespectfully breaking through boundaries of class and country, signifies a spiritual event of primary importance (Arnheim 1972, 226).” This ideal of homogenising voices as monolithic “voice media” presupposed a collective principle embedded in the listening manner of the radio. It first required that radio became private property of people and then posited an all-encompassing media network that could interlock scattered household radios together across temporal-spatial borders. By emphasising that all people listened to the same voice and content together regardless of time and space, the government believed that an all-inclusive “ear of the world” was finally constructed. In other words, both the KMT and the leftist politicians believed that modern distribution and receiving technologies carried with them a self-evident rationale of collectivism.

However, such emphasis on the intrinsic collectivism of listening to the radio had the potential to backfire on a discursive level because listening via modern sound technologies actually comprises a mixture of individuality and collectivity. Indeed, the nationwide distribution and reception of identical content and voice could form an enclosed auditory space, but the precondition of this macroscopic space was the commodification, and therefore the fragmentation and privatisation, of the sonic whole. This was because the commodification of voice and radio came much earlier than the governmental interference. This can be explained further. First, after the voice was

recorded as a product and radio became a popular commodity, all individuals could buy and own these products. Only after the voice and the listening apparatus gained ownability could they participate in the process of commodity exchange, being sold, bought, and re-owned. Second, after the voice and radio were individualised as private properties, they could be distributed, exchanged, and finally re-organised by a collective power. This reasoning reflects that the mode of listening through radio first and foremost signifies individuality and privacy rather than collectivity. Thus, the government's focus on collective sharing and listening was paradoxically predicated on individualisation. As such, the more the government trumpeted a solid community formed by a collective listening mode via the radio, the more this rhetoric might have been dismantling it on its pre-existing metaphorical basis.

Here one thing worth clarified is that I do not argue that the voice and the radio necessarily connote a primary sense of individuality. What this text emphasises is a combination of the two. When the voice is transduced and replayed by the radio, these two elements form a modern listening mode. This listening manner, once popularised by the commercial field, contained pre-existing privative and individual attributes. It is not about the voice or the radio as independent media. Rather, it is their combination (the mode of listening to the voice via radio) that points to a prior individuality. It is also based on this cooperative functioning that we can see how voice and modern sound technology, when interwoven, can mutually reshape each other's properties and how such nuance brought by their combination can challenge the political rhetoric of collectivism.

Based on the above analysis, it is possible to say that the governmental use of radio to broadcast a certain kind of female voice did not create an unquestionable sense of collectivism. Instead, as Jonathan Sterne says, the radio “facilitated listening in a collective yet individual manner” (Sterne 161). In his *The Audible Past: Cultural Origins of Sound Reproduction*, Sterne starts from a pre-existent individuated acoustic space formed by the instructional stethoscope, and extends this dialectical relationship between individuality and collectivity to more advanced technologies like radio:

The whole process of technicization operated according to a logic somewhat analogous to a bastardized version of “social contract” mythology: economic, social, and cultural forces produce property-owning individuals who then perceive themselves as voluntarily entering into a collective and later participating in a “general will” (161).

William Howland Kenney aptly described this subtle individuality-collectivity relationship with the phrase “alone together”. According to Kenney, “large numbers of individuals around the country and indeed the world, ‘alone together,’ [were] actively using their phonographs to replay as they wished commercially mediated musical messages (4).” Such a new vision of the individual-collective relationship allows us to make a departure, revealing that “phonographs, far from promoting only ‘ceremonies of the solitary,’ paradoxically encouraged widely shared patterns of popular behavior, thought, emotion, and sensibility (Kenney 4).”

Borrowing this phrase, we can say that in the 1930s, Chinese people listened to “little sister voice” via radio “alone together”. This process was not merely a collective auditory experience, but rather a dialectical relation wherein an individualised sonic

space came before it was experienced collectively. It entailed a politically regulated “general will”, the technical mediation, the commercial entertainment industries (the record, broadcasting, and star training business, etc.), and the public’s individual needs. Amidst all these elements, the female voice served as paramount nexus, whose marriage with modern sound technologies formed an ostensibly collective listening mode via the radio. However, this collective auditory space was based on the segmentation and individualisation of the auditory field. From a metaphorical perspective, the preliminary nature of the combination between voice and radio subtly fractured the patriarchal political discourse of a cohesive collectivism.

## **A Helper or a Threat?**

### **— The Ambiguous Female Voice in Social Discourse and Governmental Censorship**

As has been discussed above, aside from assuming a collective sharing and listening experience through the radio, the government determined an acoustic field that was naturalised through distinct sonic, singing, performative, and listening styles. In terms of songs, especially film songs, this stipulation generated a sentimental “little sister voice” to evoke the public’s nationalism and patriotism. However, efforts to fully control the physical properties of voice sometimes misfired. This is because the voice’s physical (or corporeal) aspects, whether timbre, pitch, rhythm, or others, did not correspond to a singular, certain meaning. Rather, the physical aspects of the voice reflected another metaphorical level, which significantly increased the voice’s emotive

and provocative possibilities. Hence, the sentimental and saccharine qualities of the “little sister voice”, though meant to reference an innocent and pure nature, could also be thought to suggest a sense of amorous feeling and flirtation, which incurred harsh critique. Such opposite opinions regarding the “little sister voice” demonstrated that due to its multiple possibilities and ambiguity, the female voice conveyed dual meanings: the positive, encouraging voice and the negative, decadent tone. The latter side was regarded by nationalists and anti-imperialists as a feminine threat that should be purged.

One example is Li Minghui’s *The Drizzle* (毛毛雨 1927), the first piece of Chinese modern popular song. This smash hit earned Li Lili overnight fame, with her Hunan-dialect and small voice resonating from gramophones and radios across the country. However, comments on this ancestor of the “little sister voice” diverged greatly. Some people praised Li’s voice paying homage to beautiful life and people’s positive attitudes. For instance, in one essay appearing in *Shun Pao* on June 9, 1933, the author describes being surprised to see a ragged and penniless old man singing *The Drizzle* happily and thus connected this song with a sanguine and carefree view of life (cited by Steen 164). In this prose, Li’s voice was regarded as a cure for citizens who had become fatigued by urban life. However, some were averse to such a voice, criticising it as demoralising, tawdry metal opium. One representative of these fervent opponents was the famous writer Lu Xun. In his *Toy* (玩具), *The Drizzle*, together with “姨太太、鸦片枪、科学灵乱、金刚法会”, was listed as Chinese adults’ “toys” (753). Besides, Lu Xun extremely detested Li Minghui’s singing style, saying that even “姨娘’s” favourite “奇

葛隆冬强” kind of sound of buskers was “much better than the hung-cat-like *The Drizzle*” (753). These contradictory thoughts of *The Drizzle* proved that the female voice could be regarded as both beautiful and morbid.

Zhou Xuan’s *When Will You Return* (何日君再来), the theme song of the 1937 advertising film *Stars Moving around the Moon* (三星伴月), was even more controversial. As a phenomenal hit single, this song was rebuked as decadent music damaging the public’s moral health. For instance, an article titled *Suggestions on Revising the Lyrics of Popular Songs* (修改流行歌曲歌词的建议) was published in *Shun Pao* on February 16, 1940 criticised this song as so obscene and depraved that it should be totally banned in Shanghai “Isolated Island”:

In short, we should use songs as an educational tool, while never allowing them to anesthetize people. As for the creation of new songs, we should pay close attention and take the responsibility of opening up a bright and easy path. In this way, decadent voices will become extinct in the space of an isolated island. This will indeed greatly benefit youth, children, society. and the nation (总之，我们应该用歌曲作为教育工具，绝对不允许借它来麻醉人们，至于新歌的产生，更应严密注意，负责开辟一光明坦途，如此则靡靡之音绝迹于孤岛空间，于青年儿童社会国家，诚大有裨益。) (Chen 592).

Although the politically controversial lyrics were the main triggers of public opinions, the affectedly sweet voice Zhou Xuan used in this song was also a major target of attack. Many other songs sung in the “little sister voice”, such as Wang Renmei and Li Lili’s *The Express* (特别快车) and *Peach Blossom River*, were associated with amorous talk and a depraved lifestyle. Such “decadent voice” lingered in Shanghai, making it an area where “people could hardly meet their true love or overcome

countless places of entertainment and brothels, and men had to be criticised because they were enticed to forget about their responsibilities of fighting for the national crisis” (Steen 165).

The negative equation of the female voice with salacity escalated to an overall attack of “decadent” recording and broadcasting. For instance, an article titled *The Morals that Radio Broadcasters and Receivers should Maintain* (广播无线电播音者与收音者应有之道德) published on October 10, 1932 in *Collection of Questions and Answers on Radio* (无线电问答汇刊) put forth that “some broadcasters indeed lack morals, frequently including several vulgar and nauseating records as well as intolerably obscene words, or wantonly railing during the broadcasting (播音者实在是缺乏道德, 于播送节目时间, 往往加入几张粗俗而肉麻的唱片与不堪入耳的言词, 或者竟肆口漫骂) (Wu 362).”

The underrepresent popularity of “the little sister voice” and the controversies stirred by it soon rang the alarm among KMT politicians. Since its early establishment in 1927, the Nanking National Government quickly instated surveillance of the publication, broadcasting, and cinematic fields (Steen 332). As Chiang Kai-shek stated in 1929:

Disregarding the positive effects of the national anthem, school song, and military song, theatrical performances and movie screenings have caused huge and crazy consequences for public entertainment, posing a serious threat to the physical and mental health of citizens (无视国歌、校歌和军歌所产生的有益影响, 戏剧演出和电影放映给大众娱乐造成了巨大的疯狂后果, 对公民身体和精神健康形成了严重的威胁) (cited by Steen 133) (cited by Steen 133).

In 1934, the KMT launched a “New Life Movement” (新生活运动), a “modern

counterrevolution” that applied traditional Chinese moral values to modern problems to fulfil its goal of “totalitarianism” (Dirlik 945). With the release of a series of official principles that normalised people’s behaviours in order to cultivate “ideal nationals” (理想国民), the condemnation of improper opera performances and film music finally deteriorated into the comprehensive censorship of “feminised”, decadent voices (Chiang 41).

The supervision of songs started with “四明文戏”. In April 1931, the Shanghai Municipal Government listed records of “四明文戏” as “salacious records” (淫片), “decadent, unlistenable sound” (淫靡之音, 不堪入耳) and prohibited the sale of such records (cited by Ge 166). Later, the censorship extended to Pingtan (评弹), Xiaodiao (or local tune, 小调), and popular songs. For instance, the July 2 and July 15 editions of *Xinwen Bao* (新闻报) reported three lists of banned programs, including a wide range of genres like comic talk (滑稽), songs, Xiaodiao, folk ditty (小曲), and so on<sup>10</sup>. What is noteworthy in these lists is that much of the prohibited music was either about romps with voluptuous women or about young girls’ audacious expression of their desires. For example, the first list included banned songs such as *A Sixteen-year-old Girl* (十六岁姑娘), *Her Face of Flower Made* (花想容), and *The Kiss Delivered* (寄来的吻). List 2 included banned songs like *The Express*, comic talks like *Modern Girls* (现代摩登女) and *Ruan Lingyu* (阮玲玉), and folk ditty like *Send A Pair of Flowers to Shilang* (对花送十郎). List 3 included banned stories like *Gao, the Woman who Sacrifices her Life for Chastity* (高烈妇) and *Crazy Woman Suffering from Leprosy* (麻

---

<sup>10</sup> See *Shanghai Radio Broadcasting in Old China* (旧中国的上海广播事业). Beijing: Archive Press, 1985: 224.

疯女), and amusing talks (趣谈) like *Women Participate in the New Reform* (妇女新改良) (224). In other words, the censorship of music was tied to females, whether those labelled as perverted women or female singers described as having decadent voices.

The mixed attitudes reflected by social discourse and the KMT government pointed to the polysemy of the female voice. On one hand, its affectively sweet qualities represented purity and powerlessness, which enabled it to be constructed as the representational tool of national salvation. On the other hand, the female voice's kittenish timbre could also merge with sentimental melody and emotive lyrics, thus rendering it a threatening force that could challenge the male-dominated public discourse. The ambiguity of the female voice uncovered a masculine sonic framework in which the female singer was represented a fetish sonic icon through which man could establish his more complete, masculine, and powerful ideal ego. However, this sonic icon was not only coded for strong acoustic pleasure and affective impact, but it also threatened to evoke the anxiety of men. This anxiety stems from the Freudian castration complex and is signified in the female (voice). In this masculine sonic framework, the tender, lingering, and sentimental female voice became indices of absence of (a penis and all other symbols derivative from it). As such, the female voice elicited both acoustic pleasure and anxiety over castration. It was both a supportive force (as the vehicle for the nationalist cause) and a destructive force (as an underlying threat of emasculation). Correspondingly, 1930s Shanghai (masculine) discourse kept oscillating between a fascination with female singers (i.e.: the female voice's popularity despite the governmental ban) and misogyny (i.e.: social contempt for the "decadent" voice)

However, despite the state's censorship and moral campaigns, the female voice still evaded legal punishment and became even more popular among the urban masses. As Ge says: "Although the public opinion was strongly opposed to all kinds of radio records of the 'decadent voice' (靡靡之音), it was difficult to take effects. Radio stations played such types of records to attract a large amount of audience and to make profits for itself (52)." Andreas Steen also stresses the following:

However, no matter the press criticism or the censorship failed to change the circulation of records and different audience groups or consumer groups' tastes. As what occurred next, despite the increase in criticism, nationalism, and the number of revolutionary records, the entertainment function of records prevailed (339).

"On a whole, since the release of "Rules of Shanghai Municipal Bureau of Education on the Examination of Opera and Records" (上海市教育局审查戏曲·唱片规则) in June 13<sup>th</sup> 1931, the actual implementation effects of the censorship of records was quite unsatisfactory [to the government] (Ge 202)." The female voice still predominated among the urban masses, outperforming the male counterpart.

The female voice's survival of the political censorship and ban can be partially attributed to the difficulties associated with the actual execution of these regulations<sup>11</sup>, "the popular taste of the urban masses", and the ineluctable expansion of the "demand-oriented and consumer demand- evoking recording industry" (Ge 202; Beus 69; Steen

---

<sup>11</sup> As Ge says, "The reason [for the failure of the censorship system] is that the record makers almost completely ignored the rule, making it a mere scrap of paper. The first step to implement the record censorship is to strictly implement the submission of records for examination, but none of the Shanghai record manufacturers complied with this legal procedure. As for the so-called 'obscene songs' that the Municipal Education Bureau was banned from being transmitted, produced, and sold. However, manufacturers ignored the ban and continued to make and sell these records(究其原因, 在于各唱片制造商完全无视该规则, 致使其成为一纸空。唱片审查制度若要落实, 首先在于严格执行唱片送审, 但对这项法定手续, 沪上各唱片制造商竟无一家遵照执行。而对于市教育局已经明令禁止传播、制售的所谓“淫词秽曲”, 制造商们也全然不予理睬, 继续制造、销售该类唱片)." See Ge Tao. *Records and Modern Shanghai Social Life*. Shanghai: Shanghai Lexicographical Publishing House, 2009: 202.

339). One essential reason for the female voice's escape from purge is rooted in the KMT government's continuous ambivalence towards the evaluation and categorisation of it. The state's hesitance was, again, due to the female voice's polyvalence and ambiguity, which enabled it to float between decadent tunes and traditional folk music and opera singing. Many female singers inherited skills from Xiaodiao (小调), Pingtan (评弹), and different opera performances. For example, the skills of using a breathy voice and nasal sound, straining one's voice to reach high notes, warbling, and using a combination of short glides and lengthened syllables all come from traditional folk music. These singing techniques made the female voice more protean and elusive because the voice, unlike vision, which has a much clearer and richer vocabulary, could only be described by unclear adjectives of mood and atmosphere. Just as Tom Whittaker and Sarah Wright state, the voice "has lent itself vividly to a broader and ever more bewildering array of functions and metaphors" (1). Due to the "slippery" qualities of the voice, interpreters might have simply found it difficult to confine the voice to any fixed category.

Moreover, the already polyvalent female voice was further complicated with the help of modern sound producing and editing, which placed the voice in an "in-between" position. For example, Zhou Xuan's two masterpieces *Wondering Songstress* and *Song of the Four Seasons* were adapted from two Suzhou folk tunes *Cry for Qiqi* (哭七七) and *Zhixin Ke* (知心客), both of which expressed women's feelings for men. The composer Tian Han (田汉) then rewrote the lyrics for them, and the musician He Luting (贺绿汀) followed with rearrangements with more modern instruments, nuances in the

tone colour, and sound effects. The redesigned song merged with Zhou Xuan's voice, skilfully hiding political metaphors while successfully evading the KMT's censorship.

Just as Yifen Beus analyses:

Ironically, thanks to cinematography and sound editing, Zhou's songs on- and offscreen were mechanically mass-produced and helped create an imagined community or an "abstract utopia" (Attali 1985, 145) where ambivalent sounds like Zhou's, which bordered between and at times blended yellow music and local tunes (xiaodiao, 小调) with occasional patriotic allusions, coexisted with the state mandate and the concurrent leftist nationalism (68).

To the KMT government, the female voice stood between 1) progressive leftist political references (mixed); 2) people's sentimental love (neutral); 3) "a simple, vivid, and direct longing for life, which characterises traditional folk songs" (neutral); and 4) an association with prostitutes' kittenish singing due to the residual "courtesan ethics"<sup>12</sup>(negative) (Huang 29; Chang 147). These multiple interpretative possibilities further challenged the state's censorship of female voices.

In addition, such ambiguous female voice was ultimately mystified and rendered even more elusive by the sonic world and other related arenas. The 1930s music industry formed a complex interior, including internal acoustic spaces without the interruption of outside noises and gazes, sleek machines with scientific precision, musical workers with technical expertise, and artists with aural sophistication. All these elements casted a mysterious aura on the music industry, coding it as a desirable but unknowable sonic system. This system closed itself to the public, but also constantly

---

<sup>12</sup> According to Michael G. Chang, though 1930s female stars had carved out a new profession and gained fame and more respect, people still constantly associated female stars with 1920s courtesans (singers). See Chang, Michael G. "The Good, the Bad, and the Beautiful: Movie Actresses and Public Discourse in Shanghai, 1920s-1930s." *Cinema and Urban Culture in Shanghai, 1922-1943*. Stanford, Calif: Stanford University Press, 1999: 150.

revealed parts of itself by cooperating with the print media. For example, on October 15, 1935 *Radio* (无线电) magazine published an article titled *The Production of Records* (唱片之制造), which described the internal workings of the machines and the studio itself (59).

The structural and artistic enigma of the music industry was later epitomised by the female singer. An article titled *The Secret Inside Story of the Record of Song of The Fisherman: Wang Renmei's Initial Refusal to Sing Resulted in Two Versions of the Records* (“渔光曲”唱片之秘密内幕，王人美一度拒唱，结果唱片有两张) appeared on September 14, 1934 in *Diansheng* (Shanghai) magazine, detailing both the songstress's career and private stories (Maigererfeng 686). These types of articles were becoming common and popular. They helped readers connect the internal mystique of the music studio, the female singer's on-screen performance (i.e.: her professional musical training), and her off-screen presence (i.e.: rumours about the songstress' private relationships). The print media, the cinematic field, and the sonic world ultimately converged to the composite wholes, and this mixture maintained a balance between the revelation and concealment of its internal secrets. It remained an ever-unknown, yet exportable lure. The female singer, as an inspirational yet inaccessible icon, aurally as well as visually, freely wandered across different media within the mysterious wholes.

As the female voice's intrinsic polyvalence worked with musicians' mediating practices to evade political purge, multi-media complexes allowed the female voice to vacillate between different areas and meanings, further amplifying its ambiguity. The

inherent nature of the voice and the external environment ultimately transformed the female voice into an elusive mystique. The female voice's mysteriousness as a phenomenon indeed reflected a male-centred symbolic system, which subordinated the female voice to a fetish object of pleasure and a manifestation of the original castration complex. However, we should also note the voice's creative engagement with this male-dominated coding. Its multipotentiality and ambivalence evolved into an unpredictable, uncontrollable, and threatening force that refused to be pinned down. Regardless of the controversy among social intellectuals caused by the female voice, the state's equivocal attitudes and the KMT's failure to eradicate the "decadent" voice demonstrated that the songstress was not so much a passive receiver as an active game player. The female singer could use the technological media and political discourse that were originally used by masculine socio-political powers to indirectly avoid being completely subsumed into the patriarchal narrative. In this process, the voice increased the uncertainty of patriarchy, triggering a fission within it. From this perspective, we can outline an alternative gender history distinct from the official narrative of the representation, education, construction, and censorship of women.

## Conclusion

In 1930s Shanghai, popular female voice first emerged as accompanying musical properties of films, but later promised an alternative entertainment world outside of cinema. As a mobile object, the voice kept shifting its relations to film and image (the body), occupying a space within it and yet fleeing from it. From both inside and outside film, this dynamic nature of voice destabilised the patriarchy that was trying to overdetermine and fix everything.

Inside film, scenes presenting a songstress listening to her replayed voice disjointed the voice from the body, dissolving the synchronisation rule that originally locked the voice into the male-constructed female body. These scenes also displayed how the songstress could perceive her recorded voice (the commodity) as an exteriority with no internal awareness, demonstrating her insistence on the discursive agency of her own voice (emanating from her own body). Moreover, the voice's essential ambiguity worked with images and led to a great increment of identificatory possibilities. Becoming a rich site of intersubjectivities, the voice obliterated the previously fixed position of the female as a passive object, while also dismantling subject-object, self-other oppositions. In brief, filmic representations revealed flaws in the dominating patriarchal ideology and enabled the female voice to develop its own space.

Outside film, masculine political discourse soon took advantage of the prevalence of popular female voices as independent products. Disciplined as a nasal, high-pitched, saccharine "little sister voice", the constructed female voice was then sewn into a vast broadcasting network that normalised the public's behaviours, mobilising the masses to

participate in a national movement of moral enhancement. In this context, the mode of listening to the radio was turned into an instrument of collectivisation, and the female voice was established as a pedagogical and domesticating tool. However, both this mode of listening promoted by politics and the policed voice revealed weak points in the patriarchy. Firstly, the promoted mode of listening to the radio was based on a prior segmentation of acoustic space, the individualism of which kept dismantling the collectivity stressed by the KMT government. Secondly, the voice's polyvalence and multipotentiality obscured its meaning. Via other media, the voice transformed into a mysterious fetish, ultimately countering against the official censorship that was stimulated by an anxiety originally signified by the feminine voice.

These analyses demonstrate the instability of the patriarchy and the potential agency of the female voice. No matter inside or outside film, the patriarchy betrayed its intrinsic unstableness and even self-deflation when endeavouring to manipulate the female voice. The holes in the logic of the patriarchy were dug out and in turn utilised by the female voice with its innate ambiguous, and thus destabilising and destructive, nature. The double destruction of the patriarchy both by the patriarchy itself and the female voice raises a broader question about our methodology of examining gender relations at different historical junctures: How should we perceive the role of the voice when analysing relationships between politics, social cultures, technologies, gender norms, and other elements? How should we reconsider the voice-image-text relation to transcend the paradigm centred on film images and literary texts when exploring Chinese gender relations?

This thesis has tried to demonstrate that the way in which voice reaches our ears is a nonlinear process of craftsmanship, rather than a static result itself. The 1930s female voice invited people to embark on either a journey within the filmic world or an imaginary air voyage. During these processes, the meanings of singing and listening shifted according to multifarious political, technical, and gender implications. Audition became a vast site for power relations and gender hierarchy to commingle and manifest themselves. In other words, the extensive sonic space became a meaningful place, where acts of listening, tuning, reproducing, replaying, and imagining should be understood as symbolic appropriation of the world.

The voice has never operated in isolation, but rather forms part of larger social and political complexes. As Sterne ascertains, what really matters is “not the breaking down of borders between sound and not-sound” but rather “the continuous constitution and transformation of the two” (337). The emergence of modern sound technology in China expressed a simple yet paramount philosophy: the ear is indeed a mechanism that can be utilised and oriented towards numerous ends. Behind the technology was a set of shared, compulsory political and operational principles. All activities organised around these principles, including singing, composing, and post-syncing, therefore, embodied ingrained norms and coming social tendencies. The prevalence and mutation of the female voice in 1930s Shanghai not only underlined technological possibilities, commercial conspiracies, and political strategies, but also unveiled a nascent modern auditory politics that revaluated listening as a producer of cultural values.

It should also be emphasised that this text does not distinguish the voice in a

singular, monodimensional sense. Instead, it conveys the rationale that a sonic event is a multisensory practice. Though this text argues against a universal body in Chapter 1 and reckons that the voice is a complex event deserving a more rounded description, this does not mean to suggest that the voice be privileged or completely separated from other senses. Just as the female voice that resounded in 1930s Shanghai ether cannot get rid of its association with the body, an audience's pleasure of listening to voice cannot be completely split from the aesthetic and imaginary delight caused by their familiarity with the songstress images and gossips on different media platforms. The audience cannot be described as comprising mere listeners, but rather "eye listeners", or "ear watchers". Throughout this paper, the voice has been leveraged as an entry point to explore how it was produced by different domains, moved among them, and ultimately reconstructed them anew.

This text-vision-sound relational perspective reminds us not to reduce either vision or sound to a singular sensory activity. As Eidsheim points out, "Academic departments each claim a single perceived sense as their domain: music has claimed audition, dance covers touch and movement, art and art history focus primarily on vision (although this has changed as artists have broadly challenged the confines of that domain), and so on (4)." Sensorial studies have started to challenge these traditionally agreed-upon divisions of knowledge. For instance, Christopher Small's redefinition of music as "musicking", an encompassing concept germane to "all people involved in music making and perceiving", counters against pre-notions of divides of sensorial domains whilst proffering us an fresh relational model to confront not only voice practices but

also other related problems (Eidsheim 4). It is based on this relational model that this text argues that using the voice as a lens enables us to blend epistemological rupture and confront issues of gender, politics, and other challenges through novel angles.

## Works Cited

- Arnheim, Rudolf. *Radio: An Art of Sound*. Boston: Da Capo Press, 1936.
- Barthes, Roland. "To Write: An Intransitive Verb?" in *The rustle of language*. Trans. Richard Howard. New York: Hill and Wang, 1986: 11-21.
- Beus, Yifen. "On Becoming Nora: Transforming the Voice and Place of the Sing-Song Girl through Zhou Xuan." *Vamping the Stage: Female Voices of Asian Modernities*. Eds. Andrew N. Weintraub and Bart Barendregt. University of Hawaii Press, 2019: 65-82.
- Chang, Eileen 张爱玲. *Prose of Eileen Chang (张爱玲散文)*. Hulun Buir: Inner Mongolian Cultural Press, 2003.
- Chang, Michael. G. "The Good, the Bad, and the Beautiful: Movie Actresses and Public Discourse in Shanghai, 1920s-1930s." *Cinema and Urban Culture in Shanghai, 1922-1943*. Eds. Zhang Yingjin. Stanford, Calif: Stanford University Press, 1999: 128-159.
- Chen Shu 陈述. "Suggestions on Revising the Lyrics of Popular Songs (修改流行歌曲歌词的建议)." *Shun Pao*, 1940(2). 14: 592.
- Dirlik, Arif. "The Ideological Functions of the New Life Movement: A Study in Counterrevolution." *Journal of Asian Studies*. 34(4): 945-980.
- Doane, Mary Ann. "The Voice in the Cinema: The Articulation of Body and Space." *Yale French Studies*, 60 (1980): 33-50.
- Eidsheim, Nina Sun. *Sensing Sound: Singing & Listening as Vibrational Practice*. Durham: Duke University Press, 2015.
- Gao Mingfang 高明芳, eds. *Archives of President Jiang Zhongzheng: Documents and Scripts Volume 14 April to May in 19332 (蒋中正总统档案: 事略稿本 14 民国二十一年四月至五月)*. Taipei: Academia Historica 国史馆, 2004.
- Ge Tao 葛涛. *Records and Modern Shanghai Social Life (唱片与近代上海社会生活)*. Shanghai: Shanghai Lexicographical Publishing House, 2009.
- Huang Qizhi 黄奇智. *The Age of Shanghainese Pops, 1930-1970 (时代曲的流光岁月 1930-1970)*. Hong Kong: Joint Publishing, 2000.
- Kenney, William Howland. *Recorded Music in American Life: The Phonograph and Popular Memory, 1890-1945*. New York: Oxford University Press, 1999.
- Liotard, Jean-Francois. "The Unconscious as Mise-en-scene." in *Performance in Postmodern Culture*, ed. Michel Benamou and Charles Caramello. Madison: Coda Press, 1977.
- Lu Xu 鲁迅. "Toy(玩具)." *Collection of Lu Xun's Essay(鲁迅杂文全集)*. Henan: People's Publishing Press, 1994.

Mai, Geerfeng 麦格尔风. "The Secret Inside Story of the Record of Song of The Fisherman: Wang Renmei's initial refusal to Sing Resulted in Two Versions of the Records(渔光曲" 唱片之秘密内幕, 王人美一度拒唱, 结果唱片有两张)." *Movie& Radio Weekly*(无线电). 1934 (9). 3.35: 686.

Ma, Jean. *Sounding the Modern Woman: The Songstress in Chinese Cinema*. Durham; London: Duke University Press, 2015.

Mulvey, Laura. *Visual and Other Pleasures*. Bloomington: Indiana University Press, 1989.

Ning, Ma. "The Textual and Critical Difference of being Radical: Reconstructing Chinese Leftist Films of the 1930s." *Wide Angle* 11.2 (1989): 22-31.

Pickowicz, Paul. "The Theme of Spiritual Pollution in Chinese Films of the 1930s." in *China on Film: A Century of Exploration, Confrontation, and Controversy*. Lanham: Rowman & Littlefield Publishers, 2012: 43-72.

Steen, Andrea. *Between Entertainment and Revolution Gramophone, Records, and the Early Age of Shanghai Music Industry 1878-1937* (在娱乐与革命之间 留声机、唱片和上海的音乐工业的初期 1878-1937). Trans. Wang Weijiang 王维江 and Lu Shu 吕澍. Shanghai: Shanghai Lexicographical Publishing House, 2015.

Sjogren, Britta H. *Into the Vortex: Female Voice and Paradox in Film*. Urbana: University of Illinois Press, 2006.

Sterne, Jonathan. *The audible past: Cultural origins of sound reproduction*. Duke University Press, 2003.

Sterne, Jonathan, eds. *The Sound Studies Reader*. New York: Routledge, 2012.

Silverman, Kaja. *The Acoustic Mirror: The Female Voice in Psychoanalysis and Cinema*. Bloomington: Indiana University Press, 1988.

Sun Rui 孙蕤. *A Brief History of Chinese Popular Music*(中国流行音乐简史). Beijing: Chinese Circles Publishing House, 2004.

"Screening Today: First Great Movie of Pure Historical Costume and War Story (今日开映 轰传已久第一伟大纯粹古装历史战事影片)." *Shun Pao*, 1927 (7). 22: 44.

Wu Chunming 伍春明 and Wu Yinfei 伍音菲. "Exploring Star Singers' Singing styles in Shanghai Popular Songs in the National Period (民国时期上海流行歌曲的歌星演唱风格初探)." *Qilu Realm of Art* (齐鲁艺苑). Shandong: Shandong College of Art. 2009. 6: 60-66.

White, Hayden. "Historical Emplotment and the Problem of Truth in Historical Representation." in *Figural Realism: Studies in the Mimesis Effect*. Baltimore: Johns Hopkins University Press, 2000: 27-42.

Wu Shizhong 吴侍中. "The Morals that Radio Broadcasters and Receivers should Maintain (广播无线电播音者与收音者应有之道德)." *Collection of Questions and*

*Answers on Radio* (无线电问答汇刊). Eds. Su Zuguo 苏祖国. Shanghai: Radio Research Press (无线电研究社), 1932 (10). 19: 362-363.

Whittaker, Tom and Sarah Wright, ed. *Locating the Voice in Film: Critical Approaches and Global Practices*. New York, NY: Oxford University Press, 2017.

Xiang, Adrian Song. "Hollywood and Shanghai Cinema in the 1930s." *CLCWeb: Comparative Literature and Culture*, 15.2 (2013). *Gale Academic OneFile Select*, [https://link-gale-com.proxy.lib.duke.edu/apps/doc/A346926251/EAIM?u=duke\\_perkins&sid=EAIM&xid=bb0043e0](https://link-gale-com.proxy.lib.duke.edu/apps/doc/A346926251/EAIM?u=duke_perkins&sid=EAIM&xid=bb0043e0). Accessed 29 Nov. 2019.

Zhao Kai 赵凯, ed. *Documents of Shanghai Radio and Television* (上海广播电视志). Shanghai: Shanghai Academy of Social Sciences Press, 1999.

*The Inventory of the Template of Old Records of Chinese Recording Factory* (中国唱片厂库存旧唱片模板目录). Chinese Record Press, 1964.