

---

Dmitri Zakharine

## **Preface: Electrified Voices. Medial, Socio-Historical and Cultural Aspects of Voice Transfer**

### **Subject 'E-voices'**

The aim of this book is to explore the phenomenon of the 'electrified voice' through interdisciplinary approaches such as media and technology studies, social history, and comparative cultural studies. The book will focus on the following problem clusters: reflections on the societal level about the task of electronic voice transmission; the mediation of gendered voices in audio and audio-visual formats as well as the genesis of vocal stereotypes in national radio and film cultures. Such a historicizing approach to societal experience in the field of voice mediation is of great relevance today in light of the collective learning processes currently triggered by rapid advances in technology.

A greater awareness of electro-acoustic voice media increasingly defines the day-to-day experience of modern man. The contemporary audience is able to calmly read about catastrophes while eating breakfast, but will quickly lose its appetite when being confronted with the voices of catastrophe victims in audio or audio-visual real-time formats. Ever since the spread of electricity-based methods of *voice recording*, *voice reproduction*, *voice transmission*, *voice reception*, *voice montage* & *voice synthesis* at the beginning of the 20<sup>th</sup> century, an abstractly understood "vocality" has gradually replaced concrete human voices. The voice of the speaker has become separated from the person of the speaker, just as the voice of the singer has become detached from the singer's person. Calculated strategies of "letting hear" and "making hear" can nowadays manipulate the individual's sense of perception and guide it in various directions. And vice versa: by accessing electro-acoustic means of production, the contemporary individual seizes the possibility of manipulating the perception of others. By modifying amplitudes, frequencies and phase shifts, electronic voice transformers enable the user to model new communicative frames and new identities by refining shrill voices or suppressing pronunciation mistakes.

## E-voice Recording

The development of electricity-based methods of voice *recording* has advanced public interest in aspects of voice that were not organized and filtered through the grid of language. In contrast to written language optical sound recording enabled people to capture roaring, sighing and hissing as well as the elements of the so called paralanguage which comprised intonation and stress. And the reverse is true as well: the development of electricity-based recording techniques in the early twentieth century was prepared and prefigured by liberal critique of the elevated speech of upper classes. Philosophers (Friedrich Nietzsche) and sociologists (Georg Simmel) referred to screaming and yodeling as the purest and most primal manifestation of the human voice before the advent of optical sound recording. E-voices were welcomed by different intellectual interest groups to mediate the primordial acoustic expression of popular masses.

The electricity-based voice *recording* made it possible to determine and analyze specific acoustic features of male and female voices. The feasibility to construct gender on the basis of audible characteristics of voice inspired debates about 'gendered' voices of radio hosts. And the reverse is true as well: determining specific features of male and female voices was anticipated by the differentiation of social norms related to salon pronunciation. Salons mainly led by women became hotbeds of female creativity in the 19<sup>th</sup> century. The impact of salons on the literary and artistic public triggered a substantial public discussion about the right of women to speak about political matters. Traditional social roles allocated to women, men and couples on the basis of gender started to waver during the first wave of European proletarian revolutions of the 1920s which were supported by revolutionary technologies of sound recording.

## E-voice Reproduction

The *reproduction* of e-voices was inseparably tied to the idea of sound amplification that extended speech in the space. Beginning from the third quarter of the nineteenth century speaking in a microphone established itself as an efficient technique of crowd manipulation. With a loudspeaker, a single person could speak simultaneously to hundreds and thousands of listeners. Hitler was credited with a famous saying "without the loudspeaker we could not have conquered Germany" (1937). The techniques of e-voice reproduction determined the nature of totalitarian state power from the ground up.

And the reverse is true as well: the advent of electronic voice reproduction was anticipated by the new character of state power that faced a rapid influx of rural population into towns in the second half of the 19<sup>th</sup> century. Local authorities

were constrained to develop crowd-control methods in order to prevent unlawful acts such as rioting and looting. This explains the steadily rising demand for powerful and combative male voices designed to defuse crowds of irrational mobs easily influenced by wrongdoers. Early theories of propaganda in which voices of public speakers are compared to those of tamers of wild beasts (cf. Le Bon, *Psychologie des Foules* (1895) and Edward Bernays, *Propaganda: The Formation of Men's Attitudes* (1923)) have predicted the use of reproduced e-voices for the purpose of crowd manipulation.

## E-voice Transmission

The electronic *transmission* of voice is based on the effect of 'schizophonia', which means the splitting of original voice and its electroacoustic reproduction (in terms of Murray Schafer). Schizophonia destabilized a well embedded relationship between voice and human reference. From now on audio media produced and transmitted "acousmatic" voices, whose sources could not be seen by audience. The unsolved question as to whether acousmatic voices are genuine and trustworthy led to a fierce discussion about the ethical justifiability of broadcasting. The deep mistrust of indigenous and peasant communities towards e-voices (to be observed in the early days of radio in the patriarchal Russia) might be interpreted in terms of collective reaction to the frightening effect of schizophonia. E-voices threatened to split traditional communities.

And the reverse is true as well: e-voices were given a task to hold supranational communities together. From the very beginning, disembodied e-voices were mainly in demand within the anonymous urban contexts. In these contexts neighbors had not become or only vaguely became acquainted with each other. Natural voices speaking in various dialects appeared there overlapped with other voices speaking with various accents. The turmoil of voices increased the demand for standardization of pronunciation to guarantee secure communication on the basis of a shared acoustic experience.

## E-voice Reception

E-voices have sensitized the listening audience for a precise manner of expression, for discerning subtleties and nuances in language and speech. The regular listening to e-voices helped to recognize deviations from the standard of radio pronunciation. From now on, unambiguous identification of someone's birthplace was ensured due to the possibility to compare acoustic data from every day interactions with norms of acoustic communication mediated by

electronic sound media. Thus one can suggest that the advent of e-voices changed the whole system of habitual identification markers.

And the reverse is true as well: the advent of e-voices was anticipated by the text-based research in the field of acoustic identification markers. This research comprised studies in dialectology and onomatopoeia. Both types of studies accumulated experience in dealing with graphical representation of exotic acoustic patterns, such as dialectal peculiarities in the language of country folk, sounds of animals and noises of machines. The increasing mobility of ethnic groups required even greater efforts to put mixed languages down in writing. The rising density of the urban population made it necessary to manage “split identities” whose carriers were people living between many cultures. Faced with the impossibility to provide reliable information on its members through visual channels, urban communities were forced to update identification details. Acoustic parameters of voice that are mainly beyond conscious control were raised to the status of reliable identity markers.

## E-voice Montage

The advent of electricity-based methods of sound fixation enabled the *montage* of voice, noise and music on the one hand and the montage of voices and images on the other hand. In contrast to audio montage, audiovisual montage was used as a dramaturgic strategy that left open the question of whether the voice source would be shown or hidden on the screen. The removal of the voice source from a frame generally increased the difficulty of matching a voice to a character and stimulated the viewer’s imagination. The audio and audiovisual montage provided an easy method to display antagonistic semantic relationships between voices and images. From now on social diversities could be depicted in the form of contrapuntal voice lines broken through towards a supposed multi-dimensional reality. The montage promoted consequently an adequate presentation of a multi-faceted modern society.

And the opposite is also true: the advent of audio montage and audiovisual montage was prefigured by the trend to polyphony, which gradually replaced primordial monophonic singing and speaking. The polyphonic art was increasingly growing and shaping the European music, literature and film by the end of the 19<sup>th</sup> century. In music art, polyphony exerted a strong influence on the modern opera, which was given a task to mediate voices of nation, including voices of civil servants, workers and peasants. In literature, polyphonic narratives presenting various points of view in the form of inner monologues, archetypal dream motives and indirect citation have advanced to become key trends. A classic example for such type of narratives is provided by Dostoevsky’s

prose, which was characterized as “polyphonic” by the literary critic Mikhail Bakhtin. In the field of silent film, montage was conceived in terms of editing short shots into sequences condensing time, space and information. The Soviet filmmaker Sergei Eisenstein defined montage as an idea that arises “from the collision of independent thoughts”.

The term ‘counterpoint’ which conjures up the idea of different motifs used against each other took on universal quality in the early theories of sound film montage. So far it had been used mainly in the context of music art, being applied for the superimposing of different rhythmic grids in the Middle Ages, for displaying contrasts between instruments and voices in the Baroque, for combining leitmotifs in the Romantic. Being used in a broader sense by the theoreticians of audio-visual montage, the term ‘counterpoint’ inspired the viewer and listener to delve into the mysterious connection between political charisma and art. Politically engaged musicians of the ‘Third Reich’, such as Ernst Hanfstaengel, compared the combination of thundering crescendos, pitch falls and spectacular gestures in Hitler’s speech with a contrapuntal byplay of leitmotifs in Wagner’s operas.

## E-voice Synthesis

The electricity-based methods of voice *synthesis*, which were invented in the early 1930s, pertain to the same matters as cloning of genes or as artificial intelligence. The phenomena now addressed make it impossible to differentiate between real and unreal. This type of indistinguishability would be usually rendered today by the term ‘virtual reality’ related to the simulation of physical presence in the real world as well as in imaginary worlds. The electroacoustic synthesis made it possible to imitate voices of people who had died long time ago as well as to simulate voices of human-like machines that did not yet exist but that could be created in the future. One of the completely new types of experimentation was synthesizing voices of imaginable beings coming from other planets, stars and heavenly streams. The e-voice synthesis led to a splitting of natural soundscapes and sound-designs.

And the reverse is true as well: the splitting of natural soundscapes and sound-designs as well as the forming of new soundscapes on the basis of mediated sound-designs can be explained through the growing gap between the biological and social rhythms of human beings. In contrast to the biological day-night rhythm, which determines physiological processes in plants and animals, social rhythm is culturally and industrially determined. Both rhythms drifted increasingly apart in the modern era. People living in the urban context were no longer dependent on the acoustics of natural soundscapes. In contrast to the

country folk that had to wake up to bird songs to feed the cattle, factory workers were awoken to the night shift by howling sirens. Self-employed freelancers, such as journalists, architects and engineers, determined their own work schedules using alarm clock rings that sounded to the set time. The drifting apart of biological and social rhythms of life increased the demand for autonomous sound media to program everyday life and leisure.

## Approaches

It is always popular and, with certain reservations, even sensible to express surprise or to express alarm about the power of technological innovations. However, scholarship has thus far failed to provide an empirical basis for demonstrating the interplay of acoustic patterns and patterns of societal communication. It is stating the obvious to say that acoustic communication is not restricted to the perception of words and phrases. Choosing special acoustic parameters (like intensity, frequency, timbre and duration), sound and voice producers have always taken into account the collective acoustic experience of the community. The argument for the social importance of the audio media is valid only if it can be proven that both listeners and those who let others listen to something, use and interpret these media in a certain culturally coded manner. Different research methods are to be applied in the present volume in order to investigate (1) the interplay between sociological problems of voice recording and voice transmission, (2) implementing the technical know-how in sound & voice design and (3) public debates about the social impact of sound & voice designs. This requires greater insight into:

1. psychoacoustic aspects of sound and voice perception
2. methods of creating sound and voice designs
3. listening practices, e. g. the set of social understandings of the use of acoustic media
4. structures of social communication, in which senders and receivers of messages are involved.

To put it in a nutshell, the present volume bridges a gap between socially objectifying and hermeneutic approaches by providing media analysis as a missing link. The present volume that comprises 24 contributions of different authors aims to provide a comprehensive but easily accessible overview of media-technical, socio-historical and cultural aspects of voice transfer. The book is not exhaustive, nor does it claim to solve all the mentioned questions from a theoretical point of view. Instead it seeks to combine the theoretical perspective of media studies with empirical approaches presented by practicing sound de-

signers. Barry Truax, David Sonnenschein, Philip Brophy, Andrey Smirnov and Lydia Kavina who have written their contributions for the present volume are known both as scholars and as outstanding musicians, who perform electronic compositions, or as outstanding sound designers.

## The Content of the Volume

The content of the present book is divided into *four* thematic blocks. Within the *first*, contributors will address the problem of dynamic interaction between sonic environment and sound-designs, laying a special emphasis on the contrast between voice-as-environment and environment-as-voice. The *second* thematic block is devoted to the interpretation of ‘gendered’ e-voices in terms of the type of information conveyed by voice. A more extensive block *three* will enable us to trace back the history of audio and audio-visual voice formats, whereas contributions in the *fourth* block will examine human voice cultures, which resisted electrification.

### Introduction

The first contribution *Voice—Emotion—Personality* by *Walter Sendlmeier* opens up new research prospects for the human voice which is considered as a cluster of physiological, psychoacoustic and sociolinguistic meanings to be interdependent and intertwined with each other in the process of communication. The article summarizes acoustic parameters which determine our perception of voice by addressing the following questions: what orientation patterns enable us to differentiate between typically male and typically female, typically young and typically old, typically angry, happy and sad voices in the situations in which the speaker cannot be seen by the listener? Can it be assumed that the same or similar distinctions enable us to recognize peculiarities of voice in different contexts, e.g. while communicating face-to-face or listening to the voices of radio hosts?

### E-voices: From Soundscape to Sound-Design

Both the questions and the answers presented in the introduction act as a prologue to the discussion in the *first block*. It is related to e-voices at the interface between soundscape and sound-design. *David Sonnenschein’s* “psycho-spatial model” called Sound Spheres summarizes cognitive and behavior-oriented

recognition patterns that enable us to create a corresponding relationship between a single heard sound and a certain group of sounds or voices. The “psycho-spatial model” is split in different sub-models (such as ‘I think’, ‘I am’, ‘I touch’, ‘I see’, ‘I know’, ‘I don’t know’) that set sound perception in relation to memories, motor skills and various types of sensory processing.

Barry Truax’s contribution *Voices in the Soundscape* contrasts voices being “disembodied” on the one hand (by means of electronic transmission) and being re-embodied on the other hand (via the loudspeaker). The paper shows that both the primary and secondary forms of orality (to use Walter Ong’s term) are designed to promote memorability, though for different cultural goals. In contrast to traditional audio media, contemporary audio media rely on exact repetition, and the position of ads within the highly structured programming flow. Repetitiveness has been proved to be more effective with audiences that have shorter attention spans. Experimental acoustic art tradition of the second half of the 20<sup>th</sup> century made use of sounds recorded in nature trying to restore the broken connection between soundscape and sound-design. It has been experimenting with variously termed text-sound, phonetic art and sound poetry, aiming to create art forms free of the perceived hegemony of music.

The famous experiments by Luciano Berio who used a reading by Cathy Berberian of the first 40 lines of chapter 11 of Joyce’s *Ulysses* as sound material to display a continuum between speech, music and noise (as quoted by Barry Truax) undergoes a detailed analysis in *Friederike Wissmann’s* contribution *About the Generation of Affects in Serial Music*. The paper demonstrates the whole scope of Berio’s fancy experiments that range from describing the spoken *s* sound as comparable to “white noise” to playing up the double ‘a’ within the name Martha as a counterpoint to the opening theme. The contribution encourages the assumption that any recorded text should not be treated otherwise than approaching it through a combination of word sound and speech rhythm. The electronic processing, says the author, enables discourses on meaning and meaninglessness of words to dive up and down in disparate entanglements.

### Gendered E-voices and Listening Community

The systematic analysis of female and male voices (view *Sendlmeier’s* contribution) enables us to switch to a current discussion about the worldwide perception of masculine and feminine frequencies (view the *second block*). Female voices are generally assumed to be higher pitched than male voices. The reason for this is the fact that women’s vocal cords are smaller (vocal cords are two bands of muscular tissue within the larynx). Male cords are thicker and longer, thus they are able to vibrate at lower frequencies, producing sounds with lower

pitch. Following on this biological difference one should not refuse to take into account that male and female speakers stylize their voices to perform specific gender roles. Girls imitate their mothers' high pitched breathy voice, while boys copy low pitched hoarse voices of their fathers. In addition to other nonverbal markers of gender, the perception of female and male voices constitutes the basis for gender-specific role behavior. The fiercely ideological interpretation of gender-specific role behavior provides the foundation for genderism that refers to a belief that one gender is superior to the other. Fierce debates about gendered e-voices gained significance and strength in the 20<sup>th</sup> century due to the spread of the new sound media that claimed to represent the whole community by means of electronic voice recording.

*Philip Brophy's* contribution *Revoicing & Gendered Vocalization* develops a model of "evaporated music" to show that on one hand, electronic 'vocality' is inevitably and inescapably gendered. And on the other hand, the 'e-vocality' is not based on gender identities of the performers. It is also not important whether performers are biologically male or female. The question is rather how vocal recordings uncontrollably reveal or disclose their ingrained gender. The author applies alternative models of analysis—'evaporation', 'phantom operation', 're-voicing', 'analytic inversion', 'vocal unbecoming' and 'gender oxidization'—to demonstrate a large variety of stereotypes which have nothing to do with the sex or with the gender of physical people. The e-voice is a voice which is 'given up' in the act of mediation. The e-voices of 'Elton John—the animated corpse: dried, withered, corporeal utterance', 'Phil Collins—the mirrored ghost: phased modulated communiqué' or 'Celine Dion—the crone witch: raspy, cat-coughing screech' pertain less to the world of finished vocal structures, but rather more to the world of vocal options, that excite the imagination and invite the listener to actively construct the voice's gender.

*Nicola Dibben's* paper *The Intimate Singing Voice* provides a look at a little-researched field of e-voices radiating intimacy. Under which circumstances can the fore-grounded singer's voice be heard as an intimate sound through which the listener gains the illusion of access to the inner thoughts and feelings of the performer? The author applies Edward T. Hall's theory of proxemics to determine how spatialization in recordings impacts on the perception of interpersonal distance. Given the mutual visual perception, decreases in interpersonal distance are associated with gaze avoidance. Male-male dyads have greater interpersonal distance than female-female dyads. Does this rule apply to the case of acoustic or better for the case of 'acousmatic' communication between singers and listeners? The effects of reverberation and breath sound seem to be crucial with regard to mediating distance and proximity. Small spaces specified by a short time of acoustic reverberation are preferred and considered more intimate and more pleasant than larger spaces.

In contrast to *Nicola Dibben's* paper, which is devoted to electroacoustic mediation of interpersonal and intimate distance, *Kate Lacey's* contribution *Speaking Up and Listening Out* deals with e-voices in the public sphere. The paper explores a particular moment in German history when the idea of 'voice' was at the intersection of three critical democratic developments—the founding of the Weimar Republic, the emergence of women into the public sphere, and the rise of mass communication in the form of public broadcasting. The author shows that on one hand the debates about women's voices on the air were couched in technological terms. It was widely claimed that voices with a higher pitch tended to suffer from distortion in the process of conversion of sound waves to electrical signals. On the other hand technological arguments centered less on what was technically sympathetic to the ear, and more on what was culturally most in agreement with convention. And the convention was, simply, that the male voice was more authoritative and more versatile. The critical attitude of the community to female e-voices demonstrates the state of mind of the listening public in the early days of radio. The radio audience has predominantly been conceptualized in terms of 'listening in' to a particular show or station, rather than 'listening out' in the politically active role of a critical public.

*Jason Loviglio's* contribution *U.S. Public Radio, Social Change and the Gendered Voice* enhances the discussion about the interrelationship between 'gendered voices' and collective expectations of listening audience. The case of the American "National Public Radio" (NPR) is interesting due to the fact of the latent animosity between different types of political elites, which all use 'gendered e-voices' to manipulate public opinion, and the listening competence of the audience. According to the NPR website, the author says, there are currently seven women news presenters as compared with only four men; there are nine women hosts of news and information programs compared to only six men. All three "senior correspondents" are women and the voices of women hosts, presenters, and correspondents of NPR's most widely heard programs are uncommonly low in pitch. The author claims that NPR's flexible attitudes towards gender roles obscure the ways in which the network has come to match the increasingly conservative spirit of the times. As NPR lurched rightward in the 1980s it had to find the sounds that would simultaneously appeal to an audience of affluent listeners—a group more likely to attract corporate underwriting and convey a sense of "publicness."

## E-voice Synthesis and E-voice Montage. Techniques and Social Meaning

The goal of the third thematic block is to historicize the e-voice phenomenon and to integrate it contextually. This block is devoted to discussions of the 1930s-2000s about the vocal stereotypes that have developed in the context of national radio and film cultures. This discussion seems relevant in view of the increasing unification to which international voice transmission formats are currently subjected. As a result of this unification in film and radio, the national vocal stereotypes that have emerged over the past century under the influence of audio media are exhibited less and less by themselves. Instead, they are shifted within and into the incorporated interpretive patterns of the actors and thus seem far from obsolete. Based on the public reception, imitation, and parodying of media voices, every culture developed its own tradition of the use of voice recording technology as well as the own concept of identity, which reflected the vocally constituted nature of the respective collective identity. Technological differences as well as differences in the interpretation will be elucidated based on individual presentations on American, British, German, and Russian radio and film voices.

### 1920s: The Roaring Twenties

*Liubov Pchelkina* directs her attention to the social and technological utopias of the 1920s which were boosted by the advent of electroacoustic media and spurred by the experience of belated modernization in Russia. Her contribution called *The Biomechanics of Voice and Movement in the Solomon Nikritin's Projection Theatre (1920s)* describes the typology of human voices developed by the Russian painter and philosopher Solomon Nikritin (1898 – 1965). The typology that also comprises classifications of movements and gestures, emotional states, sounds and color palettes was based on the principles of biomechanics, musical harmony and acoustics. Nikritin's drawings portray human beings as well-functioning "socially-engineered Human-Machines" of the future. In these drawings voices were divided into small segments and syllables, which are used, as granolas to create sound textures, forming a dense stream of sound, in accordance with a special score. From Pchelkina's study one can conclude that methods of electronic voice-image montage were anticipated by the universalistic approaches to the use of body language. These approaches based on acoustics aimed to maintain its standardizing effect in a rapidly changing future.

*Andrey Smirnov's* paper *Synthesized Voices of the Revolutionary Utopia* enhances the discussion about the interplay between social and technological utopias. The contribution deals with the scientific estate of the Russian composer and musical theorist Arseny Avraamov who proposed (1932) to use the

optical methods of voice synthesis for reviving Lenin's voice. Avraamov's project aimed at the use of the synthesized timbre of Lenin's voice for vocalizing Lenin's works (that were never read out loud by Lenin himself). The experimental methods developed by Avraamov and other scholars (cf. Yankovsky) were based on audio computing techniques. They implied discretization and quantization of spectral data, followed by the synthesis of ready-made parts. These methods can be considered as a kind of proto-computer, briefly reminiscent of modern digital technology which enables the composition of music via computer sequencer.

The article *Thereminvox*, which was written by one of the leading performing musicians on the theremin, *Lydia Kavina*, is devoted to the other breakthrough invention of the 1920s. The electronic instrument which bears the name of its inventor was developed by the Russian physicist and musician Lev Theremin (1919) who experimented with high frequency circuits. Unlike many other electronic instruments the theremin reminds listeners of the human voice. That is why it is better known as '*Theremin-Vox*' in Russia. It appears symptomatic that the Soviet espionage tool called "The Thing" designed for tapping human voices from a distance was developed by the same inventor. While composing music for the theremin, it is important to sing along to let the body adapt to the wavelength of the instrument. Melodies, Kavina says, are traditionally notated for the theremin in the same way as they are noted for voice, or for the flute. Nowadays there are only few acknowledged experts who are able to write music for the theremin with a guarantee that this music is playable. The theremin seems to be indispensable for performing alienated, cold, melancholic, etheric voices such as the voice of Little Mermaid in the Ballet by Lera Auerbach.

### 1930–1940s: The Shouting Thirties

The basic discussion about the interrelationship between the voice-in-soundscape and the voice-in-sound-design invites us to pose questions and obtain answers on more specific questions such as the interrelationship between voices, e-voice-designs and e-voice communities. The relationship between electro-acoustic technologies and political power at the beginning of the era of electronic voice reproduction has been a matter of some debate in contemporary scholarship. Many of the relevant publications define radio and early documentaries as instruments of propaganda, imposed by the Nazi and Soviet dictatorship with the purpose of assembling far-reaching surveillance structures. Conversely, other researchers increasingly see repressive regimes as rooted in scientific inventions in the field of telecommunications. In this second research corpus, the radio, the loudspeaker and sound film are assigned a structuralizing function

within society. Scholars in anthropology and the philosophy of media have made the case that acoustic and audio-visual media have a particular capacity to exercise power. One of their main arguments holds that humans have two eyes, but no “earlids”. Since man cannot evade auditory impact, the source of political power is said to inhere in the involuntary acoustic guidance of attention. In the 1960s, media studies pioneer Marshall McLuhan established the “monopolistic effect” of the radio by bringing concepts such as “auditory space” and power into a close relationship with one another. The famous analogy between the radio and the tribal drum elaborated in McLuhan’s writings has produced a number of recent studies on acoustic communication.

*Dmitri Zakharine’s* contribution *Voice—E-Voice—E-Voice Community* addresses issues related to the perception of electrified voices in the 1930s and 1940s. The study deals with the interrelationship between the voice stereotypes of radio and TV announcers on the one hand and the voice stereotypes of classical horror films (monster, scream queen) on the other hand. The author assumes that electronic sound effects which were initially designed for fictional genres gradually migrated into the genre of the newsreel. The synthesis of animal-monster voices in fictional film projects on the one hand, and the choice of news speaker voices in documentary film projects on the other hand, were both subject to similar selection criteria. The social tendency to find male voices more authoritative and credible blocked the critical sociological interpretation of imperfect technical solutions, such as distorted tuning of high-pitched voices. And vice versa, the imperfect technical solutions strengthened the existing prejudices against women announcers.

The paper titled *The Pitch of Historic Sound Films as a Sociolinguistic Variable* by *Folke Müller* uses qualitative and quantitative methods of pitch analysis to convince the reader that e-voices in sound film of the 1930s might be considered the prototypes and functional analogues of later technologies of mass media manipulation whose objective was to affect the aesthetic impressions of the audience by acoustic means. Most roles in German movies of 1930s correspond to specific stereotypes of social behavior that manifests itself in the pitch of actors’ voices. Women with a lower socioeconomic status (such as chambermaids and midwives) possess voices with a lower pitch. High pitch voices of courageous lads who represent the young combative heroes of the time (cf. ‘Hitlerjungen’) are being aesthetically upgraded in comparison to the voices of civil servants that pertain to the public administration of the Weimar republic.

The paper titled *The Exploited Recordings: Czech and German Voices in the Film ‘Theresienstadt’* by *Natascha Drubek* aims to enrich the initiated discussion about the use of the e-voice montage in the sound film for the purpose of propaganda. The paper provides an analysis of the role of voices in the original documentary ordered by the SS and shot by the inmates of a concentration camp

in the last year of the Second World War. The film had to show the prosperity of Jewish inhabitants of 'Theresienstadt'. The documentary transmits conversations, shouting and singing in German and Czech. However, one voice does not belong to an inmate—it was added to the soundtrack in the final cut in March 1945 and belongs to a German narrator. The voice-over informs the spectators that Jews are living in good conditions in German camps. In comparison to the incessant talking on 'Deutsche Wochenschau' the voice-over and the audiovisual style of "Theresienstadt" was much more reminiscent of German newsreels for abroad. In particular the "Panorama-Monatsschau" (in color) of September 1944 has long intervals of silence—as if the German commentators had nothing more to say to their audiences.

### 1950s: The Icy Fifties

The time after the Second World War was characterized by a new competition evolved between political systems. The Cold War which had an enormous influence on education (including the activities to conquer space) was transferred to the Cosmos in the fifties. The use of space for security needs became a central element in strategies, doctrines and planning. The change of priorities explains the increasing demand in vocalizing the extraterrestrial life. From now on acoustic art innovations were boosted less by the blatant propaganda directed against terrestrial enemies. Instead, the focus was placed more on the modeling voices of the hostile humanoid beings coming from other planets. Needless to say, that the virtual e-voices kept all the traces of political and technological change.

The paper titled *The Voices of the Cosmos* by *Konstantin Kaminskij* deals with the electronic synthesis of special sound effects in Soviet vs. American science fiction movies. The paper demonstrates how the connection of science fiction plots and acoustic media-semantics overlays the ideological message implied in science fiction movies that intended to prefigure the socialist future in the present of the Cold War era. To put it more briefly, the author shows how the Soviet science fiction film of the early 1960s was given a "voice" and how that "voice" was conveyed in Hollywood. The voice of the planet Venus in the Soviet film *Planet of Storms* (*Planeta Bur'* (1959)) was performed through the montage of the voice by a noted Peruvian soprano Yma Sumac (Hollywood's "Inca Princess" and "Queen of Mambo"). Sumac's voice was integrated into a melodic composition accompanied by theremin. Kaminskij describes in the following text an exploitive integration of socialist science fiction film into the American entertainment market. To adapt the alien material, the (more or less) neutral video-track of these films was separated from the soundtrack with its ideo-

logical-saturated content. As clearly illustrated by the American versions of *The Silent Star* (1960), *Planet of Storms* (1962), and *Ikari XB-1* (1963), even the 'non-ideological' electronic music of the socialist films was replaced by classical Hollywood orchestrations.

### 1960s: The Storming Sixties

*Hans-Ulrich Wagner's* paper *Sounds like the Sixties* defines this time as a period of medial and societal transformations. 'Radio Luxemburg' and its new sound attracted the younger generation. The 'Norddeutscher Rundfunk' from Hamburg had to take into account that young people were bent on listening to pirate stations in the North Sea. Some observations of the author are inspired by his personal memories of his early radio listening experience: "As a child in the Sixties and the Seventies I listened to the voices of which I *now* think that they are characteristic of this time. [...] Persons belonging to the generation of my parents remember Mr. Zimmermann's shouting voice telling the radio listeners that Germany wins the World Championship in 1954 [...]. But I remember the weekly teleconferences. In these teleconferences radio stations switched from one stadium to the other". The early fascination with listening grew into the author's interest for the roots of acoustic memory. The paper deals mainly with conceptual issues as to *why* and *how* should one analyze voices of the past.

*Susan Smith's* contribution *Vocal Sensations in Boom!* addresses the phenomenon of Elizabeth Taylor's voice, whose reception might be considered in the context of the movement of the sixties to get back to nature, to disregard for the boundaries of modesty and, among other things, to revise critically the artificiality of amplified e-voices. More specifically, the paper refers to the vocal chemistry between movie stars Elizabeth Taylor and Richard Burton, using their performances in the film *Boom!* (Losey (1968)). Taylor's own long-standing association with animals and nature from child stardom onwards adds a specific logic to the film's final merging of her character with the surging storm swells. The turbulent properties of Taylor's voice find their extension in the crashing sounds and restless energy of the ocean. Mike Nichols (who directed Taylor in *Who's Afraid of Virginia Woolf*) allegedly had wanted Elizabeth Taylor to take voice lessons before filming *Virginia Woolf*, even though ultimately her voice could be enhanced by the sound-track. But Elizabeth insisted that she would manage fine without lessons.

## 1970–1980s: The Metallic Seventies

The symbolic content of intensive coarse and husky voices associated with crowd manipulation has been re-evaluated in the context of the pacifist movement of the seventies. In the thirties, sounds of roaring airplane engines and wheel motors as well as the loud voices shouting political slogans and military commands maintained the legitimacy of state power. In the seventies, the acoustic manifestations of the military state spirit were reinterpreted in terms of menacing signals bringing about a real threat to culture and peace. In the thirties metallic sounds had inspired politicians to take onomatopoeic pseudonyms that evoked associations with the firmness of steel: e.g. “Stalin”. In the seventies, the steely voice became a characteristic feature of illiterate army instructors. The national audience in Europe and the US split into the listening public and listening counter-public. The latter one increasingly found like-minded people among the transnational listening audience striving for an acoustic space to be free from state control. Experiments with the use of metallic voices appeared to be aesthetically upgraded. They were shifted into the newer music trends such as thrash metal, speed metal, death metal, which used combinations of high-pitched screaming and extremely fast double bass percussion as well as satanic themes in poetry to convey the nihilistic views of the post-war youth generation.

*Nils Meise's* paper *Seven-six-two millimeter. Full, metal, jacket* draws the reader's attention to the semantics of voices and sound in (anti-)war films, such as *Full metal jacket* (1987, directed by S. Kubrick) or *Apocalypse Now* (1979, directed by F. Coppola). The title of Kubrick's film refers to the full metal jacket bullet used by infantry riflemen. The heroes of the film are recruits of the U.S. Marines. After having their heads shaved, the recruits meet their senior drill instructor Gunnery Sergeant Hartman who employs draconian tactics while preparing soldiers for combat. According to the author's observation, the voice of the drill instructor is loud and low pitched, while the voices of recruits are mostly high pitched. The main sound effect is based on the combination of silence, echoing sounds of footsteps and the drill instructor's characteristic shouting. The sound designer did not use any music to underscore this film scene. The same imaginative sound montage is used in the last scene (of the first part of movie) whose impact is likewise primarily sustained by the combination of voices and everyday noises. The roles are changing in this scene. One of the recruits is shouting against the background of the turmoil of sounds, while we can also hear a rifle being loaded. The drill instructor, who has heard someone shouting, enters the scene, only to be shot down by the mentally unstable recruit, who has brought himself to protest against the system of shouting.

## 1990s-2000s The Calculating Late Modernity

The socio-acoustic model which postulates the differentiation between “listening public” and “listening counter-public” is related to the seventies. It appears to be inadequate to depict the increasing complexity of acoustic space of the nineties. The individualization and accelerated differentiation of the economy, society and media affected all aspects of life by the end of the millennium. The sound media of the 2000s mirror the hybridism between the indigenous and the foreign. At a single time there is no doubt that the artificial homogenization of cultural landscapes and soundscapes (controlled by globalized media) favors the formation of parallel societies and parallel listening audiences in the background. An important question of the 2000s is whether a particular imagined community (whose integration is mainly realized by symbolic means) can be directly identified through its vocal performances. The constant necessity to consider the differences between own and other, between one’s own fiction and alien reality, between attractive alternatives and inadmissible deviations moves into the foreground of the modern society and shapes its acoustic culture.

The contribution titled *Theoretical-methodical Approaches to Radio Aesthetics* by Golo Föllmer argues that acoustic information comprising subtle aesthetic details of speech is more relevant for identifying social groups as previously believed. The so-called “stationality” or “channel identity” provides us with a key to understand listening preferences of audiences. Föllmer claims that the channel identity comprises differences in speech style, timing, rhythm, including the so-called teasers, etc. Some of the results presented in the article are derived from the research project “Radio Aesthetics—Radio Identities”, which was co-initiated by the author and realized under his supervision. The project shows that e. g. a recorded announcer’s voice can be transformed into the ‘radio voice’ representing channel identity by merely compressing the dynamics of the signal. There is a certain degree of risk that the channel identity remains unstable. It may disappear and become ‘invisible’. In contrast to newspapers, film and television, whose appearance is continuously displayed by various visual elements, the identification of radio channels is not guaranteed. This is all the more so when the primary element of most programs, namely music, is played. The famous saying assigned to Socrates “Speak, so that I am able to see you” is gaining importance nowadays with regard to the unique property of radio channels, which can be identified only through e-voices on the basis of speech performances.

The paper titled *Alien Voices: the Sonic Construction of Foreignness in Science Fiction* by Philip Preuß and Steffen Lepa enhances the opened discussion about the distinction between own and foreign. The authors claim that the figure of an extraterrestrial being should be interpreted as a narrative device to highlight the

legitimate as well as the illegitimate ways of treating foreigners. In a broader sense, “alien voices” constitute a cinematic language to handle the problem of cultural diversity. The article analyses three blockbuster movies that were shot in different historical epochs: 1) *The Day the Earth Stood Still* (1951), 2) *ET-The Extra-Terrestrial* (1982), and 3) *District 9* (2009). There are both similarities and differences to be observed with regard to the form of alien voices in the aforementioned films. Similarities are determined by the need to display the semi-human source of acoustic manifestations that represent foreignness. Differences are mainly related to the grade and type of the electronic alienation of voice. To some extent these differences might mirror the changing listening competence of the audience. According to the observation of the authors, the late modern audience (infected by skepticism) would probably not take the naïve sonorous peace message of Klaatu seriously (*The Day the Earth Stood Still*). Klaatu’s voice radiating calmness and superiority stands in contrast to the swirling sound of the electronic musical instrument theremin, which seems to express feelings of the ordinary people during their confrontation with the UFO (cf. also *Kaminskij*’s contribution in this volume). Likewise the audience of the fifties would find it difficult to interpret the chirring-knocking sounds of the disturbing “prawns” in *District 9* in terms of a peace message.

The series of essays analyzing techniques and social meaning of e-voices in late modern sound-designs is followed by *Chistofer Jost*’s contribution *Computer Based Analysis of Audiovisual Material*. The software called “TrAVis”, which was developed in the context of the research project *The Analysis of Popular Cultures*, aims at exploring the interrelationship between images and the smallest imaginable elements of sound-design which comprise syllables and thirty-second notes. The implementation of this software makes it possible to dissolve the human voice in the ocean of computer byte codes and to recover it on the basis of the similar byte codes. The software enables scholars to analyze the pre-programmed and planned as well as the intuitive and spontaneous handling of sound material by sound designers. The statistical processing of data, which comprise voices, noises and music, from various sources would provide insights into the subconscious logic of brain activities. Besides that, it would give people an insight into configurational logic of cultural symbols which are perceived with eyes and ears.

#### Human Voice Cultures That Resisted Electrification

The presented history of e-voice montage should obviously not convey the impression that communication based on the perception of human voices was completely replaced by the area-wide use of e-voices. Extensive field research

and data collected in exotic regions between America, Europe and Asia would allow the reader to escape the deeply linear logic of Eurocentric media history. The main question of the fourth block is the following: how are social norms, local superstitions and everyday mythology implemented in the development of artificial electronically recorded voices?

The contribution titled *A Medium of Magical Power: how to do Things with Voices in the Western Amazon* by Bernd Brabec de Mori analyzes the functions of voice within the indigenous community Shipibo-Konibo which dwells in the Western Amazonian rainforests. The problem of voice electrification is treated in the article by a twofold approach: firstly, the question arises as to why recorded and reproduced voices are considered by the indigenous to not work properly while mediating magic power; secondly, the author asks how electric devices are summoned by trained singers in order to enhance the magic power. It is to be assumed that the voice is viewed by Shipibo as a substantial item to transmit animistic beliefs. According to these beliefs it appears impossible to reproduce magic power by means of an electronically reproduced voice. Records are believed to be able to imitate the voice only. The animistic beliefs related to sound do not stop Shipibo people from using radio programs, telephone and tape, CD or DVD recorders or players in their everyday life.

Tanja Zimmermann enhances the discussion about voice cultures, which resisted electrification, by tracing the history of the *Voice of Gusle*. A gusle is a single-stringed instrument used by the Southern Slavs. It was often played with a bow to accompany epic narration and it was compared with the voice of Slavs already by the founders of the modern historical science (such as the German historian Leopold von Ranke). For ages playing gusle has been associated as political weapon, which represented the national spirit of the Serbs. The instrument ultimately participated in the recent electoral campaigns in Serbia. The documentary *Serbian Epics gusle* (1992) which was shot in the Republika Srpska in 1992 presented Radovan Karadžić and his associates as passionate *gusle* players standing in front of the icon of St. Michael. Mixing the sound of gusle with the sound of electronic instruments was prohibited until the recent past. The sound of gusle was recorded live and transmitted unchanged via television and radio.