

Audio ‘Sonic’ Ecology under the Resurgence of Auditory Culture: A Case Study of Chinese Podcasts (Postprint)

Authors: Li Yao

Date: 2023-03-24T00:00:00+00:00

Abstract

For a long time, the dominant development of visual symbols has constrained the evolution of auditory culture, as people pursue intense visual sensory stimulation under the attention economy. However, the advancement of mobile communication technologies has revitalized audio media, with Chinese podcasts representing an emergent force that cannot be overlooked within new sound media. The inherent mobility and concomitance of Chinese podcasts enable podcast audio to provide users with companionship on both psychological-emotional and physical-spatial dimensions.

Full Text

The Acoustic “Sound” Ecology Under the Return of Auditory Culture: A Case Study of Chinese Podcasts

Beijing Institute of Graphic Communication, 102600

Abstract: For a long time, the dominant development of visual symbols has constrained the growth of auditory culture, as people pursue intense visual sensory stimulation within the attention economy. However, advancements in mobile communication technology have revitalized audio media, with Chinese podcasts representing a significant emerging force among new sound media. Chinese podcasts possess inherent mobility and companionship, offering users psychological and physical spatial companionship through sound.

Keywords: auditory culture; auditory return; Chinese podcasts; acoustic companionship

Canadian media theorist and thinker Marshall McLuhan proposed the concept of “tribalization—detribalization—retribalization,” arguing that the greatest trans-

formation in returning to the tribal era is the transition from eyes to ears—the return of acoustic space.

With the mature development of mobile communication technology, humanity has entered the mobile communication era, gradually awakening the subjectivity of sound and accelerating the return of auditory culture, which has propelled the rapid growth of the online audio industry. According to iMedia Research’s “2020-2021 China Online Audio Industry Report,” China’s online audio user base reached 640 million in 2021 and is expected to rise to 690 million in 2022[1]. Podcasts are products of internet-era media technology development, and Chinese podcasts are podcasts that primarily use Chinese as their language. Leveraging mobile communication technology and employing audio formats, Chinese podcasts disseminate accessible, useful, and engaging information, creating a unique information “sound” ecology.

The Return of Auditory Culture: Breaking the Status Quo of Audio-Visual Imbalance

In the 1960s, Marshall McLuhan proposed the concept of “returning to the auditory” in his work *Understanding Media: The Extensions of Man*. McLuhan divided human social development into three stages from the perspective of media evolution history: the tribal age, the detribalized age, and the retribalized age. The “tribal age” was an era biased toward oral communication, that is, an era relying heavily on auditory transmission; the “detribalized age” was an era biased toward text, that is, an era relying heavily on visual transmission; and the “retribalized age” is the era of electronic transmission, which McLuhan also called the era of “returning to the auditory.”

In the era of oral communication, the ear was the archetypal mode for individuals to receive external information. With the invention of phonetic writing and printing technology, a culture centered on visual transmission formed. Vision became an important perceptual method in the communication process, and the development of modern technology and civilization in the 20th century further strengthened the dominant position of vision, leading to an imbalance between auditory and visual senses in information transmission. As Dewey believed, “the connection between ear and ear (that is, ears bearing important and extroverted thoughts and emotions) is far closer and more diversified than that between eye and eye”[2]. This not only demonstrates his emphasis on auditory culture but also illustrates the irreplaceable role of hearing in communication activities.

1.1 Perceptual Imbalance Under Visual Hegemony

Since the advent of modern civilization, visual media has consistently occupied a dominant position in information transmission, and people have gradually formed the habit of accepting visual information, creating a cultural industry where vision reigns supreme. With the derivation and renewal of visual symbols, people have continuously intensified their control and pursuit of vision,

making visual supremacy the mainstream. In daily life scenarios, “seeing” has almost replaced the body’s other sensory organs, becoming the primary perceptual pathway for people to understand and transform the world. Even for media like television and film that incorporate both sound and images, people generally use the expressions “watching TV” and “watching movies,” which, while emphasizing the importance of visual symbols, simultaneously neglect the role of hearing[1]. Consequently, in audio-visual media, the auditory sense is regarded as an auxiliary method to vision.

Guy Debord proposed the concept of the society of the spectacle, arguing that the development and popularization of electronic media have enabled images to form a system of their own, with social life itself being preferentially presented as a world of images. The universalization of images constructs a society of the spectacle. In an era where everything is media, the entire world has become highly visualized, with visual senses being filled with various postmodern meaningless symbols. Visual symbols have become important chips for capturing traffic in the attention economy era, enriching people’s aesthetic experiences[2]. However, these visual symbols are driven by commercial interests, and as the visual symbols flooding platforms like Douyin and Kuaishou demonstrate, they can easily lead to the decline of the public’s aesthetic perception and judgment capabilities.

1.2 The Development of Mobile Audio Awakens the Auditory Constitution Under Visual Culture

Radio and silent film were the two dominant media in 20th-century America, but today, only radio media continues to exist and is constantly integrating into people’s lives. This relates to how humans perceive information: people can accept listening without seeing, but cannot accept seeing without listening. The different fates of these two media also demonstrate the importance of hearing for human information reception. Moreover, compared to vision, sound control can better activate human physical perception and emotional responses.

In the 1960s, high visualization led to sensory imbalance in humans, limiting human development and gradually reducing people to “one-dimensional man.” Scholars including McLuhan, Heidegger, and Plato have all called for the return of hearing, believing that the return of hearing can also balance the excessive expansion of vision. The development of contemporary internet technology and mobile communication technology has propelled the arrival of the mobile communication era and provided an opportunity for the return of hearing.

In the mobile communication era, digital technology has eliminated the boundaries of time and space. While imaging technology has gradually reached its extreme, audio content has migrated to portable, editable mobile devices, completely liberating the spatiotemporal limitations of hearing[3]. As the pace of life accelerates and mobile communication technology continues to permeate daily life, people’s life scenarios have increased and time has become more frag-

mented. The emergence of mobile audio enables people to obtain information through a single auditory sense without relying on visual symbols and allows listening in multiple scenarios. Therefore, mobile audio can accompany people through various fragmented periods without consuming excessive attention resources. Consequently, in the mobile communication era, mobile audio can fully integrate into people's lives, awakening the auditory constitution that has been suppressed under the dominant development of visual culture.

Podcast—A New Form of Audio Media

“Podcast,” known in English as Podcasting, primarily refers to downloadable aggregated audio files published via the internet using Really Simple Syndication (RSS) technology[4]. Podcasts were born in the United States, with the term coined by British *Guardian* journalist Ben Hammersley in 2004 when introducing internet audio. As podcasts were introduced to China, various Chinese podcast websites were established, and Chinese podcasts, leveraging their trans-spatiotemporal and mobile advantages, sparked an early wave of podcast development. However, due to immature development models, narrow audience reach, and the impact of the rise of video media, Chinese podcast development experienced stagnation and even contraction.

2.1 Current Development Status of Chinese Podcasts

Affected by the COVID-19 pandemic in 2020, the reduction in offline activities stimulated the rapid growth of the stay-at-home economy, and Chinese podcasts have also encountered new development opportunities. Currently, Chinese podcast development is concentrated in two aspects. First, podcast channels on audio platforms. Online audio platforms such as Ximalaya, NetEase Cloud Music, QQ Music, and Lizhi FM have all launched podcast channels. Second, vertical audio podcast platforms. In March 2020, the Jike team launched China's first podcast app—Xiao Yuzhou—which can provide vertical, aggregated podcast content, gathering a large number of podcast creators and loyal listeners, breaking the dilemma of Chinese podcasts being marginalized on audio platforms.

Compared with other audio-visual media products, Chinese podcasts demonstrate strong competitiveness. According to the “2020 Chinese Podcast Listener and Consumption Survey,” Chinese podcasts are mostly leisure entertainment and casual chat types, while industry-specific and practical podcasts are relatively rare. Second, as an “auditory” product, podcasts can support “multi-tasking,” providing physical companionship and psychological companionship, allowing users to listen without conflicting with their current activities, exercise, or work. In terms of audience reach, podcast user groups are concentrated among people under 35 years old, living in first-tier cities, and mostly holding bachelor's degrees or higher. According to PodFest China's “2020 Chinese Podcast Listener and Consumption Survey,” 86.4% of Chinese podcast listeners have bachelor's degrees or above[1]. In the same survey, 59.2% of respondents indicated that they reduced their music listening time after listening to podcasts,

and consumption duration for other audio content such as audiobooks, audio knowledge payment products, and audio live streaming also decreased to varying degrees, demonstrating that Chinese podcasts are currently experiencing strong momentum and possess considerable competitiveness.

2.2 Product Characteristics of Chinese Podcasts

In the online audio market, whether audiobooks, host programs, or voice live streaming, most adopt UGC+PUGC as the primary creation model. From the perspective of Chinese podcast creators and content, “UGC+PGC+PUGC” is the main creation model. Podcast content production generally involves individuals or small teams of several people and possesses strong IP attributes, with creating vertical IPs being the biggest characteristic of this content production model. Additionally, podcast language styles tend to be colloquial, with topics generally appealing to popular interests. Therefore, podcast program atmospheres are generally relaxed, witty, and humorous, capable of transcending the singular dimension of sound, bringing listeners a tangible, immersive experience.

Currently, Chinese podcast program formats are primarily group chats and conversations, which can be divided into four categories: multi-host chat, guest interview, solo narration, and non-fiction storytelling. These four interview-style programs each have distinctive features. Multi-host chat programs have no fixed questioner and interviewee; each episode centers on a topic, with hosts expressing viewpoints and insights on corresponding issues. Guest interview programs interview a guest in each episode, where the program host becomes the questioner, and the interviewed guest can share their professional knowledge and experience on the program. Solo narration programs feature only the host delivering a monologue, with the program’s presentation effect mainly depending on the host’s personalized characteristics. Non-fiction storytelling programs are like “non-fiction writing” presented through sound, often involving one or multiple people sharing a unique experience. According to surveys, regarding podcast formats, more users tend to choose multi-host chat programs and guest interview programs, with proportions reaching 81.4% and 77.6% respectively. Next are solo narration programs and storytelling programs, accounting for 47.8% and 33.1% respectively[1].

The Acoustic Companionship of Chinese Podcasts

In research on Chinese podcast platforms, companionship is an important aspect, which encompasses two primary dimensions. On one hand, it refers to psychological companionship—the degree of psychological accompaniment Chinese podcast platforms provide users. On the other hand, it refers to physical spatial mobility—users can use Chinese podcast platforms unrestricted by time and space. In the development process of modern society’s individualization, networking, and urbanization, individuals’ “atomized” lifestyle patterns have become increasingly prominent, with empty-nest populations in urban life growing and expanding, and more people experiencing lonely living conditions amidst

busy work schedules. The emotional companionship and spatiotemporal companionship inherent in Chinese podcast sound can penetrate the crevices of life, filling and soothing modern people's sense of emptiness and anxiety.

3.1 Emotional Companionship: Oral Narrative Constructs an Emotional Community in Sound

Auditory interaction can develop empathetic capacity in the most primitive way, and language communication can better receive information, consensus, and beliefs within the relevant social system through hearing, promoting the construction of collective social sentiment[2]. Chinese podcast platforms not only possess the physical mobility characteristic of mobile terminals but also, as sound media, bring emotional companionship.

Communication has evolved through five stages: oral communication, written communication, print communication, electronic communication, and network communication, among which oral communication represents the most primitive form in human society, the foundation of all communication methods, and one that remains in constant use.

In Chinese podcast programs, the main formats are group chats and conversation-based podcasts, and both forms transmit information through sound as the medium. Each podcast program episode transmits information through oral communication forms. Sound itself possesses strong individuality—text may be repetitive, but each person's voice, timbre, and expression methods are all different. From this perspective, Chinese podcasts' use of oral forms to arrange programs breaks the high-temple narrative style, making them more conducive to understanding and listening, catering to current users' media usage habits.

The oral communication method can not only effectively transmit information, but hosts' voices also have distinctive characteristics. Even though topics in various podcast programs may be similar, their language styles are unique, capable of vividly bringing the described images before listeners' eyes through speech, allowing people to feel the subtle emotional temperature. The highly individualized oral communication dependent on hosts constitutes podcasts' unique sound expressiveness. Moreover, from the listeners' perspective, beyond expressing viewpoints on topics, hosts in podcast programs also share their lives before episodes and provide care for listeners during gaps between episodes, bringing listeners electronic companionship and generating emotional bonds. Therefore, for many listeners, podcasts have become invisible friends and emotional outlets, consequently enabling them to form emotional communities with others[1].

3.2 Spatiotemporal Companionship: Covering Fragmented Life Scenarios

Podcasts currently represent the fastest-growing audio medium within mobile audio platforms, naturally carrying the “companionship” gene, which is also

consistent with the media characteristics of traditional radio. “Companionship” has two meanings: on one hand, it means people can do other things while listening to podcasts, where podcasts act as a “background medium”; on the other hand, it means podcasts can be listened to privately and personally, allowing users to listen whenever they want, giving podcasts the function of a “personal medium.” Mobile listening, close listening, and anytime listening fully demonstrate the companionship advantages of mobile audio media and also give podcasts extremely high adaptability to space and environment.

The mobile internet’s all-encompassing intrusion into people’s daily lives makes modern people need to frequently switch between different contexts and spaces, and “multitasking” has also become the norm for most people’s lives. Podcasts can perfectly serve as the life background sound during “multitasking.” Scenarios such as commuting, doing mechanical repetitive work, doing housework, exercising and running, bathing, and before sleep are all scenarios where users utilize Chinese podcast platforms, where podcasts bring users a sense of companionship and reassurance without affecting other things they are doing. Therefore, when facing the dual pressures of fragmented time and scattered attention, Chinese podcasts with companionship characteristics can fully adapt to the randomness of user behavior and deeply integrate with various life scenarios.

From the perspective of physical mobility, Chinese podcast platforms rely on mobile terminals, allowing users to access them anytime and anywhere. From the perspective of psychological companionship, Chinese podcast program types are diversified, gathering various audio content program types covering multi-host chats, guest interviews and conversations, storytelling, and other forms, and the colloquial narration methods are highly personalized to hosts, with more relaxed and free expression methods. Podcasts have gradually become the life background sound accompanying “young people living alone.”

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv — Machine translation. Verify with original.