Historical Study of Recorded Speech: Voiceover's Evolution

Dr. P.V. Rajlakshmi
Assistant Professor of English
Kongu Engineering College, Erode

OPEN ACCESS

Volume: 12

Special Issue: 1

Month: December

Year: 2023

P-ISSN: 2320-2645

E-ISSN: 2582-3531

Received: 24.10.2023

Accepted: 05.12.2023

Published: 14.12.2023

Citation:
Rajlakshmi, P. V. "Historical Study of Recorded Speech: Voiceover's Evolution." *Shanlax International Journal of English*, vol. 12, no. S1, 2023, pp. 333–39.

DOI:
https://doi.org/10.34293/rtdh.v12iS1-Dec.111

Abstract
Voiceover covers a wide spectrum from radio ads to video game character. But our journey through history, it is defined as the voice of an unseen narrator speaking. This is the initiation of voiceover's evolution. It begins in the 19th century when the concept of recording sounds was in infancy and imagine, where the sound was a mystery waiting to be unveiled. All of the requiring professional voiceover services to convey messages effectively. It is a time when voices reach global audiences from the comfort of home studios. The ease of access to voiceover opportunities, coupled with the growth of online freelance networks, has made it simpler than ever to start a career in voiceover. It was a moment that expanded the possibilities of voice manipulation. The rise of online video platforms, especially YouTube, has opened new avenues for voiceover, at the COVID – 19 pandemic times. The internet's vast landscape offers countless opportunities, from explainers' videos to independent animations. This paper focused on the chronological evolution of voiceover from the era of mime shows and also the current technological advancements due to the influence of Artificial Intelligence. This study is not detailing about the algorithms, machine learning and deep learning models works. Instead, it is explained with real time living beings that involves into the evolution of voice cloning technology.
Keywords: Phonautograph, Broadcast, Evolution of Voiceover, Voice Cloning Technology.

Voiceover covers a wide spectrum from radio ads to video game character. But our journey through history, it is defined as the voice of an unseen narrator speaking. This is the initiation of voiceover's evolution. It begins in the 19th century when the concept of recording sounds was in infancy and imagine, where the sound was a mystery waiting to be unveiled. We will kick the things off with Edmund-Leon Scott de Martinville's1860 rebdition of 'Au Clait de Lune', a scientific narration received on a phonautograph. Then, we have Frank Lambert's 'Experimental Talking Clock' from 1878, which marked an improvement in audio quality as an technological advancement. There is a haunting beauty to these early recordings, a glimpse into a future filled with voices. By 1900, inventor Reginald Fessenden took a leap by broadcasting a weather report using radial signals. Although it is lacked quality, it was a significant step in the world of voiceover – the ability to cast audio across immense distances was just becoming a reality. Broadcast is born.

By 1906, Fessenden had redefined his technology, by transmitting live radio messages across an 11-mile radius during Christmas. This moment cemented voice-over's connection with radio broadcasting a medium that would become the playground for voice actors.

The Electrical Era

Fast forward to the 1920's, the British Broadcasting Corporation (BBC) was born. It was revolutionising broadcast, and reaching audiences worldwide. It was the dawn of a new voiceover and the establishment of a broadcast institution that has lasted over 100 years. In 1926, we encountered 'My Old Kentucky Home', a sound-based cartoon. The first of its kind, to feature, a recorded voice acting performance, accompanied by visual animation while revolutionary its places history's marred by its extremely prejudiced subject matter and terminology.

In 1927, 'The Jazz Singer' arrived, effectively ending the era of silent movies with its sound-on-disc technology. A new bar was set; sound and film would be married forever more. The electrical era also brought us a quintessential animated short of its time; Walt Disney's 'Steamboat Willie'. In 1920, Steamboat Willie was the first fully synchronized cartoon using sound, with Disney himself lending his voice to Mickey Mouse. Disney emerged as a leader on the frontier of animation something that continues today. 1932 introduced Tarzan's iconic call, while its origin is contested, this may possibly be the earliest example of apply effects to a voice and showcasing advancements in audio editing. It was a moment that expanded the possibilities of voice manipulation.

The Golden Age of Animation

The rise of animation in the 1930's gave birth to voiceover as entertainment. Mel Blanc, known as the "Man of a Thousand Voices" revolutionized the industry with his vocal work in "Looney Tunes". Mel Blanc's talent and performances legitimized the voice acting craft: a turning point, where voice actors could become stars in their own right.

Orson Welles and The War of the Worlds

Orson Wells is known for delivering "The War of the Worlds", a radio drama aired live at 8pm Eastern Time on October 30, 1938, via the CBS (Columbia Broadcasting System) Network. This particular episode gained notoriety for sparking panic among some of its listeners, who were led to believe in a Martian invasion. The extent of the panic remains a subject of debate since the programe had relatively small audience. Nevertheless, it marked a pivotal moment when voiceover artistry blurred the boundaries between fiction and reality, demonstrating the profound influence of spoken words, for good and for ill.

Voiceover Vs. The Nazis

The Second World War was a worldwide humanitarian disaster and radio broadcasts were used in propaganda by many nations during the conflict. One such radio host, Mildred Gillars, a Nazi sympathizer, used her voice to unnerve and demoralize American troops. earning the nickname, Axis Sally. Fortunately, voiceover was used for good too. Sefton Delmer's use of Black Propaganda helped undermine the Nazi war machine. The actors would pretend to be Nazi supporters to subtly spread misinformation about the state of the enemy and sow doubt in the minds of German forces and citizens. Needless to say, voice acting played a crucial role in the war over hearts and minds won. Voices have power that much is clear.

Voiceover on the Small Screen

The 1950's saw the proliferation of TV advertising, featuring voiceover in the form of tags at the end of product commercials, a milestone in the advertising industry. Commercial work has since become a cornerstone of voiceover work. Also, Disney's ability to dominate animation continued. The television made it possible for voices to be right there in our living room and integral to our daily lives. Voice actors became household names.

Don LaFontaine and the 1960s

In a world, where voices are still emerging as legitimate stars, the man, the myth, the legend, Don LaFontaine could not be ignored. He bridge the gap between commercials and cinema with his iconic movie trailer voiceover style, cementing voiceover as a respected profession in the entertainment industry. His voice is now synonymous with movie trailers, and has been parodied and copied countless times.

2001: A Space Odyssey

Arthur C. Clarke and Stanley Kubrick's 2001: A Space Odyssey released in 1960 – a work that delves into the realms of human evolution and Artificial Intelligence (AI). At the heart of the narrative stands HAL 9000, a sentiment computer brilliantly brought to life by the talented actor Douglas Rain. In the film, Rain flawlessly delivered a neutrally accented English vocal performance, making an iconic line, "I am sorry Dave, I am afraid I can't do that," both a sinster and convincing portrayal of an intelligent AI. Perhaps, there is a suspicion if AI does offer such a threat to our own future.

The Magnetic Era

Moving to the 1970s, the advent of the Sony Walkman brought about a revolution in personal audio consumption. The term 'Walkman' became synonymous with personal audio devices, and this era witnessed the rise of audio books. Audio cassettes became the ideal medium for abridged audio books, and most importantly, music. The Walkman allowed people to enjoy literature and their favourite tunes on the go, personal and portable media players took a huge step forward into our palms of our hands. Olympic Gold medalist, Duvall Hecht, played a pivotal role in this revolution by founding Books on Tape in 1975, specializing in putting books on tape. Audiobooks opened up literature to a broader audience. Including those with disabilities and non-native speakers, making written content more accessible. Novels were no longer confined to the printed page.

Darth Vader

The late 70s and early 80s were marked by a pivotal moment in cinematic history: the rise of Darth Vader. The "Star Wars" franchise, beginning with Star Wars in 1977, later renamed 'A New Hope', introduced audiences to the iconic character – and voice of Darth Vader. Vader was designed to intimidate but the on-set performance of David Prowse was not enough, and his lines were revoiced with the dark, distinctive voice of James Earl Jones, creating an indelible cinematic villain. James Earl Jones' dark, imposing voice fit Darth Vader like a glove and remains one of the iconic voices in film history. It is a testament to the impact a well-choosen voice actor can have on a character's legacy – not to mention that this was a huge step for voiceover dubbing. With a voice actor defining a live action role for decades, film's quintessential, most recognizable villain's most defining feature is provided by a voice actor.

The Digital Era

As we entered into 1980's voiceover made the mark in the burgeoning world of video games. "Castle Wolfenstein", released in 1981 for the Apple II, featured voiceover that added intensity to the game play. The use of speech, despite the limitations of technology at the time, marked a signigicant step forward. Silas Warner, an innovative game developer, "Voice", a digital audio recorder that set the standard for voiceover in early video games. In the following decades, professional voice acting became essential to engage gamers and enhance storytelling in video games.

The Digital Revolution and the Role of Unions

The 1990s witnessed the digitalization of information and the internet, marking a significant revolution in voiceover history. Digitalised democratized content production, making it more accessible and affordable and Digital Audio Workstation (DAWs) made it easier for small studios. Independent artists to create high-quality content while internet provided a platform for content creation and distribution, early live-directed voiceover sessions often used ISDN (Integrated Services Digital Network) technology, to transmit audio. Acting unions played a crucial role in the industry's early days, with voice actors recording lines, receiving payments through agents or unions, and focusing solely on their vocal talents. It was a time when voice actors relied on traditional structures. However, as technology advanced, prices dropped, and recording quality improved, voice actors gained more autonomy. They began recording in home studios, negotiating rates, and handling their own editing, reflecting the changing landscape of Voiceover. The Digital Era gave birth to a new bread of voice actors. It suddenly became easier than ever to become a voice actor.

Anime and Otaku

In the 1990s, Japanese animation, or "Anime", experienced a resurgence. Hits like "Neon Genesis Evangelion" required exceptional voice actors to bring complex characters to life. This era marked the dominance of character-driven stories and mecha anime, with series like "Cowboy Bebop" setting high standards for animation, writing, direction and voice acting. Anime transcended borders and languages. The success of Anime led to the rise of Otaku culture, a dedicated fanbase of Anime enthusiasts. "Sub vs. Dub" debates, discussing Japanese voice with subtitles versus English dubbing, became intense discussions within Otaku communities became intense discussions within Otaku communities. The popularity of Anime continued to grow, and today, many voice actors have Anime localization and dubbing work in their portfolios. Voiceover has been so pivotal in crossing boundaries, especially between Eastern and Western cultures.

Voiceover in the 21st Century

In recent decades, voiceover has gained recognition as a promising career with celebrities increasingly lending their voices to commercials, video games, and audio books. The involvement of celebrities like Tom Hanks, Patrick Stewart, and Samuel L.Jackson has added prestige to the profession. Motion capture or MoCap technology has further blurred the lines between actors and digital counterparts, allowing animators to capture actors; visual likeness and facial expressions. Video games, Anime and Western animation, now feature celebrity voices as a standard marketing strategy. It is a testament to the power of voices in modern entertainment.

Emerging Voiceover Trailblazers

In addition to celebrities, established voice actors like Nolan North, Troy Baker, Jennifer Hale, Laura Bailey, and Steve Blum have risen to celebrity status within the gaming community. These professionals have become synonymous with iconic characters and performances, reflecting the growing prominence of voiceover in entertainment. They are the unsung heroes behind our favourite characters.

The New Digital Frontier

The rise of online video platforms, especially YouTube, has opened new avenues for voiceover. The internet's vast landscape offers countless opportunities, from explainers videos to independent animations, All requiring professional voiceover services to convey messages effectively. It is a time when voices reach global audiences from the comfort of home studios. The ease of access to

voiceover opportunities, coupled with the growth of online freelance networks, has made it simpler than ever to start a career in voiceover. However, this accessibility has also led to an abundance of talent, making competitions fierce.

Influence of Artificial Intelligence

Modern AI voices require a combination of algorithms, machine learning and deep learning to create the system that makes it comprehend a query and respond with an answer as well as only a few seconds of audio to replicate sound. But AI voices struggle to accurately portray emotions, which, means they regularly fail to connect with audiences. If you are a content creator who cares about your global audience human voices are always better.

AI voices are said to have a profound impact on the creation of industries from video games, Education to Youtube videos and advertisements. So we need to take a look how they're made. Some AI voices are already highly recognizable, like Bev Standing for TikTok, Susan Bennett as Siri. To get an AI voice to have the quality of Siri was difficult. Susan Bennett had to record thousands of lines over many years in a paro studio to capture all the intonations and individual phonemes. Apple's algorithm for Siri then strings these recordings together and concentrates the audio to recreate realistic sounding speech. Creating Siri was a complex process compared with earlier text-to-speech machines, such as robotic voice of the CallTaxi 5010 which was popularized by Stephen Hawking. But the original, in 1976, Kurzweil Reading Machine sounds very inhuman. Now-a-days, AI voices sound much more realistic thanks to their foundation.

AI is structured to mimic the human brain, the most complex biological machine in the known universe, in what is called artificial neural networks. They are the powerhouse of modern AI. The only issue that we face is sophisticated AI requires an immense amount of data to anticipate and react to user inputs effectively. For an AI designed to identify photo of cats, it must be fed pictures of cats and other animals like dogs. So it can learn the difference. What a cat looks like may seem obvious to us but an AI is completely blank state. It has no point of reference or ability to differentiate between the two animals. So it needs to learn from scratch. The artificial neural network starts by predicting whether a picture is a cat or not. Then it answers are cross-referenced against data providing the correct answer. The AI with the highest success rate the one that is the best telling what a cat is and it will be used as a template for the AI's deep learning model. The template will be continually refined through immense amount of trial and error each iteration of AI system becomes better and better at identifying what a cat looks like compared to a dog, Incrementally improving the efficiency; eventually, evolving to the point of instantly identifying a cat in busy photos. This process is called backpropagation. It allows neural networks to learn and improve the more data fed into machine. The better of AI becomes at completing its task combination of these artificial neural networks, machine learning and deep learning methods allows the AI to clone us the way we look, create and even sound.

Creating artificial voices used to take weeks or even years of recording time. Now only a few seconds of audio is needed to recreate the sound of our voice. This is achieved in three main stages:

- Break down
- Natural Language Processing
- Speech Synthesis

Firstly, AI analyses audio data of someone's speech to break down voice into its core pronunciations. This requires another side of the AI to use natural language processing to comprehend the way language works. Comprehending language means shifting through immense amounts of audio data to develop a system capable of navigating speech and the way it should sound. Speech synthesis is the final stage, which pulls together the previous two steps to create a

cloned voice pairing up the voice template from the vocal data and the language processing model to synthesis the speech as an AI voice-over. It has been never so easy to create an AI version of you or someone else. However, there are many dangers with voice cloning technology which is why tools like Meta's voice box have been held back from the market.

References

1. Aaltonen, R. (2002). An overview of documentary - typical translation concerns and strategies: the translation performance in focus. Retrieved from: Studies at the Interface of Translation and Culture Web site: http://www.eng.helsinki.fi/images/kuvituskuvat/sitcu_rosa.rtf
2. Agost, R. (2001). Aspectos generales de la traducción para el doblaje. In J. D. SANDERSON (ed.), ¡Doble o nada! Actos de las I y II Jornadas de doblaje y subtitulación (pp.13-26). Alicante: Universidad de Alicante.
3. Aleksonyte, Z. (1999). Comparative analysis of subtitles and voice-over in Danish and Lithuanian Respectively as Compared to English (based on the Danish Film Breaking the Waves). Unpublished manuscript. Vilnus: University of Vilnus.
4. Aufderheide, P. (2007). Documentary film: A very short introduction. Oxford: University Press. Ávila, A. (1997). El doblaje [Dubbing]. Madrid: Cátedra.
5. Ballester Casado, A. (1995). La política del doblaje en España (número monográfico de Eutopías, 94), Valencia, Episteme.
6. Bartrina, F. (2001). La investigación en traducción audiovisual: interdisciplinariedad y especificidad. In J. D. SANDERSON (ed.), ¡Doble o nada! Actos de las I y II Jornadas de doblaje y subtitulación (pp. 27–38). Alicante: Universidad de Alicante. BESAS, P. (1985). Behind the Spanish lens. Denver: Arden Press. BBC Broadcasting Research. (1985). Spoken translations and subtitles in documentaries. Special report provided by the BBC Written Archives Centre. 116
7. Bogucki, Ł. (2004). The Constraint of relevance in subtitling. Journal of Specialized Translation, 1, 71-88. BRAVO, J. (2003). La investigación en traducción audiovisual en España: los textos cinematográficos. In M. A. García Peinado, & E. Ortega Arjonilla (eds.), Panorama actual de la investigación en traducción e interpretación, vol 2, (pp. 235–252).
8. Granada: Atrio. Cattrysse, P. (2001). Multimedia and Translation: Methodological considerations. In Y. Gambier, & H. Gottlieb (Eds.), (Multi) Media Translation (1-13). Amsterdam, Philadelphia: John Benjamins.
9. Chaume, F. (2004). Synchronization in dubbing: A translational approach. In P. Orero (ed.), Topics in Audiovisual Translation (pp. 35—53). Amsterdam and Philadelphia: John Benjamins.
10. Chaves, M. J. (2000). La traducción cinematográfica. El doblaje. Huelva: Universidad de Huelva.
11. Daly, A. (1985). Interpreting for internatuinak satellite television. In H. Bühler (ed.), Translators and their position in society / Der übersetzer und seine Stellung in der öffentlichkeit. Xth Wold Congress of FIT / Kongressakte X. Weltkongress der FIT (pp. 203—209). Vienna: Wilhelm Braunmüller.
12. Delabastita, D. (1989). Translation and Mass-Communication: Film and TV Translation as Evidence of Cultural Dynamics. Babel 35 (4), 193–218.
13. Díaz Cintas, J. (2003). Audiovisual Translation in the Third Millennium. In G. Anderman, & M. Rogers, (Eds.), Translation today: Trends and perspectives (pp.192-205). Clevedon, Buffalo, Toronto, Sydney: Multilingual Matters. 117 Díaz Cintas, J; ORERO, P. (2006). Voice-over. In Encyclopedia of Language and Linguistics. (2nd edition, pp. 477-479). Oxford: Elsevier.

14. Díaz cintas, J., & Anderman, G. (Eds.). (2009). Audiovisual Translation: Language Transfer on Screen. London: Palgrave Macmillan Díaz Cintas, J. (1998). La labor subtituladora en tanto que instancia de traducción subordinada. In P. Orero (ed.), Actes del III Congrés Internacional sobre Traducció (pp. 83–89). Bellaterra: Servei de Publicacions de la Universitat Autónoma.
15. Diaz-cintas, J. (2003). Audiovisual Translation in the Third Millennium. In G. Anderman, & M. ROGERS (ed.), Translation Today: Trends and Perspectives (pp. 192—204), Clevedon: Multilingual Matters.
16. Dries, J. (1994). Slightly out of synch. Television Business International, 62-63. DRIES, J. (1995). Breaking Eastern European Barriers. Sequentia, II (4), 6. Eguíluz, F. (1994). Transvases culturales: Literatura, Cine. Vitoria-Gasteiz: Universidad del País Vasco.
17. Espasa, E. (2004). Myths about documentary translation. In P. Orero (Ed.), Topics in Audiovisual Translation (pp. 183-199). Amsterdam, Philadelphia: John Benjamins Publishing Company European Broadcasting Union (EBU). (2004). EBU Report on Access Services. EBU Technical-Information 144-2004. Fawcett, P. (1983). Translation modes and constraints. The Incorporated Linguist 22 (4), 186-190.