Human or AI?

Exploring the Impact of AI Tools on Audio Content Production and Perception

Barbara Brandstetter
Department of Information Management
Neu-Ulm University of Applied Sciences
Neu-Ulm
e-mail: barbara.brandstetter@hnu.de

Abstract— **There is a growing trend of people consuming audio content in Germany. As a result, many media companies have invested in audio content in recent years. With the help of Artificial Intelligence (AI) tools like Elevenlabs or MurfAI, producing high-quality sound has become relatively easy. The first part of the study aims to determine if media users can differentiate between AI-generated and human voices and how they perceive AI-generated audio content compared to human-generated content. In the second step, the analysis wants to determine how AI influences the content's credibility and the users' willingness to pay for audio content.**

Keywords-AI; Audio content; Podcasts; Credibility; Willingness to pay

I. INTRODUCTION

Germany has a growing trend of using audio and video content for information and entertainment. In Germany, nearly 53 million individuals aged 14 and above used at least one audio service every working day in 2023. On average, users listen to audio services for more than four hours a day [1]. The younger demographic has shown a particular interest in podcasts, enabling media to reach well-educated, affluent target groups who are moving away from traditional news consumption [2]. Some experts observe a shift from written to spoken content, as people value the ability to listen while engaging in other activities.

Publishing houses are increasingly investing in audio content, taking advantage of the growing popularity of audio formats. Many media companies in Germany are now offering podcasts and audio versions of their written articles [3]-[6], and the trend of audio content will continue. Advancements in AI technology are also driving the surge in audio content. AI tools, such as text-to-speech technology, have made it possible to create more natural-sounding speech, improve audio quality, and enhance personalized recommendations [7]. This has allowed media companies to work more efficiently, reducing production costs and time. AI tools like Elevenlabs or MurfAI have made adding high-quality audio to content easier, enabling the replication of emotions, tones, accents, and even translation into different languages [8]. Many media companies and podcasters now rely on various AI tools for content conceptualization, production, post-production, and marketing. Despite these advancements, more research is needed on how audiences perceive AI-generated voices. The study aims to address the research gap by answering the following research questions:

1. Can participants distinguish whether a voice is human or synthesized by an AI tool?
2. Does the use of AI tools impact the credibility of content or the willingness to pay for it?

Section 2 of the paper focuses on related audio production and AI literature. Section 3 explains the methodology. Section 4 considers the first results of the study. Section 5 provides a conclusion, and the last section addresses the limitations of the study.

II. RELATED LITERATURE

Many newsrooms have used artificial intelligence for various purposes, such as personalized content, fact-checking, and content production [9]-[13]. AI tools have helped media companies save costs and time. Before the introduction of ChatGPT, some media companies used algorithms to report on stock market developments and weather forecasts. In recent years, the focus of using algorithms in legacy media has been on automated texts and research comparing texts written with the help of algorithms with those written by humans [14].

A recent comprehensive analysis by Thurman et al. examined how media users in the UK perceive human-made, partially automated, and highly automated short-form videos. The researchers found that the participants did not detect huge differences between the differently produced videos [15]. A representative study conducted in the USA, Germany, and China, covering audio, image, and text, shows that test subjects need help distinguishing human-generated content from AI content [16]

With the constant improvement in the quality of text-to-speech tools, an increasing number of media companies, such as *Neue Züricher Zeitung*, *Süddeutsche Zeitung*, and even regional newspapers like *Rheinische Post* are offering the option of reading articles aloud [17]-[19]. Additionally, the emergence of AI tools for creating and optimizing audio content, such as Elevenlabs or MurfAI, has led to many media houses using these tools for audio content production. These tools can be used to optimize audio recordings and even to clone voices.

In this study, we will focus on human, cloned, and artificial voices used in podcasts and for the read-aloud

function on media company websites. The study aims to discover how AI affects audio content perception and whether people can detect humans from cloned or artificial voices.

Studies by industry services, such as Bitkom, show that it is essential to media users that journalistic content notes whether AI has been used [20]. However, the effects on willingness to pay and credibility in the audio sector still need to be determined. The study also wants to fill this research gap.

III. METHODOLOGY

In the first step, we conduct a within-subjects experiment. We ask participants to listen to different audio files and determine whether the voice was produced by an AI tool or a person. After each test, we conduct individual-focused interviews based on the experiment results. This method allows for detailed and profound questioning [21]. According to Mayring, content analysis is used to categorize and analyze the interviews [22]. In the last step, we will inquire in a brief survey about the participants' socio-demographic aspects, audio use, and willingness to pay for audio content, such as podcasts. The study also aims to understand the importance of test subjects knowing whether AI was used in creating journalistic content and how this information affects the perceived credibility of the content and the willingness to pay for it.

A. Stimulus materials

The study required test subjects to listen to audio content (human, cloned and artificial voices). The experimental stimuli were divided into podcasts and audio voices, which offered the service of reading articles published on media websites. Ten different audio files from various areas, such as politics, business, sports, and regional affairs, were selected for the study. The order of the examples presented to the participants was altered to prevent potential learning effects. The following files were played for the test subjects:

1) *Podcasts:*
- The Episode about Russia and Ukraine - the cloned voice of the host
- The Episode about Russia and Ukraine - the human voice
- The Episode about the search for a new trainer of FC Bayern - the cloned voice of the host
- The Episode about the training of FC Bayern - the human voice
- The Episode about new AI tools - the cloned voice of the host

2) *Spoken Articles:*
- Salaries at RWE - the cloned voice of a reporter
- Queer people in Hamburg - artificial voice
- Here I come - an article about reckless people - the cloned voice of a reporter.

Additionally, the participants were asked to compare the human voice and the cloned voice generated by the AI tool Elevenlabs from two different podcasters.

B. Participants

When selecting test subjects, we ensured a balanced ratio of men and women. The test subjects were required to have experience listening to a podcast or using the option of having a text read aloud on a website. Nine test subjects took part in the first test run, which was conducted via Zoom in May 2024 (see Table 1).

TABLE 1. LIST OF PARTICIPANTS

Number	Participants			
	Sex	***Age***	***Podcast use***	***Audio use***
1	Female	45-54 years	yes	yes
2	Male	55-64 years	no	stopped using it
3	Female	55-64 years	yes	stopped using it
4	Male	55-64 years	no	stopped using it
5	Female	35-44 years	yes	stopped using it
6	Female	25-34 years	yes	stopped using it
7	Male	45-54 years	no	stopped using it
8	Male	55-64 years	yes	stopped using it
9	Male	35-44 years	yes	stopped using it

Some of the test subjects listened to podcasts regularly. All of them had tried having an article read to them at least once. However, the respondents had one thing in common: everyone except one respondent no longer used this service. The unanimous argument was that the audio output quality needed to improve, and listening to the artificial voice was challenging. Some also mentioned that they preferred scanning a text for interesting passages rather than listening to an audio recording. One test subject utilized the read-aloud feature to have articles in foreign languages read out loud. Nevertheless, all test persons were surprised at how the quality of AI-generated voices improved.

IV. FIRST RESULTS

The initial results have shown that none of the test subjects could identify all AI-generated voices. This result is consistent with those of the study by Frank et al. Media users need to be informed about the use of AI tools in producing media content. The test subjects even felt that the information that AI was used needed to be increased. They would like to know precisely for which production steps the editors or podcasters have used AI. For example, the test subjects find listening to an AI-generated voice less problematic - if they like the voice and intonation. However, the situation is different when AI is used to research content. Respondents are particularly skeptical about journalists using AI for research. For instance, respondent 3 mentioned, "I experiment extensively with AI tools and therefore know

that the answers are not always perfect. That is why I would not trust AI-generated content in journalism." However, as our first results show, providing information about the use of AI tools can lead to lower credibility. Nevertheless, respondents are divided when it comes to their willingness to pay. Many would not be willing to pay the same price for journalistic content if it were generated with the help of AI tools.

V. CONCLUSION

In the first step of our research project, we wanted to find out if people can detect humans from cloned and AI-generated voices. This is relevant because many media offer audio content using AI tools. Our first results show that the probands could not say if a voice were human or artificial. Even though people could not detect differences in the audio examples provided in the test, people said that media should indicate if and for what steps in the value chain media used AI tools. However, the information on the use of AI tools generally affects the content's credibility and willingness to pay for it.

VI. LIMITATIONS

It is important to note that our study is ongoing. With nine test subjects, the sample is still tiny. We will expand our study by analyzing audio and video content and testing it with more test persons. We will ask probands to listen to audio and video content produced with the help of AI tools and produced by humans. We will use the usability lab of HNU conducting an eye-tracking test, and a facial expression analysis using the software iMotions. We will meticulously analyze the emotions evoked during audio and video consumption. Even if initial results show that hardly anyone succeeds in distinguishing AI-generated voices from human voices, people may react differently emotionally to the content or fixate on other content with their eyes in AI-generated videos.

While we have initial results, a comprehensive analysis and further testing are still underway. We look forward to sharing these additional insights shortly.

REFERENCES

[1] K. Gattringer, Audio usage 2023 in Germany: A look at the latest ma audio figures. Media Perspectives 5/2024, Available from: https://www.ard-media.de/media-perspektiven/publikationsarchiv/detailseite-2024/entwicklung-der-audionutzung-in-deutschland [accessed 30/05/2024]

[2] Podstars OMR, Our big podcast survey 2024. Available from: https://podstars.de/blog/podcast-umfrage-24/ [accessed 30/05/2024]

[3] R. Fletcher, K. Eddy, C. Robertson, and R. Kleis Nielsen, Digital News Report 2023. Available from: https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2023-06/Digital_News_Report_2023.pdf [accessed 30/05/2024]

[4] J. Behre, S. Hölig, and J. Möller, Reuters Institute Digital News Report 2023: Results for Germany. Hamburg: Verlag Hans-Bredow-Institut. https://doi.org/10.21241/ssoar.86851

[5] B. Domenichini, Podcast perspectives - from hype to sustainable trend. Media perspectives 7/8 2022. Available from: Media_Perspektiven_07_2022.indd (ard-media.de) [accessed 30/05/2024]

[6] L. Frühbrodt, and R. Auerbacher, Hitting the right note. The podcast boom in Germany. Otto Brenner Stiftung, 2021. Available from: https://www.otto-brenner- stiftung.de /fileadmin/user_data/stiftung/02_Wissenschaftsportal/03_Publikationen/AH106_Podcasts.pdf [accessed 30/05/2024]

[7] S. Hodgkins, How Artificial Intelligence is Transforming Podcasting, 2023. Available from: https://www.linkedin.com/pulse/how-artificial-intelligence-transforming-podcasting- simon-hodgkins [accessed 30/05/2024]

[8] R. Simmonds, Audio AI: How AI Is Changing Podcasts, Audiobooks & More. Hubspot, 2024. Available from: https://blog.hubspot.com/marketing/audio-ai#:~:text=Audio%20AI%20is%20the%20use,AI%20is%20breaking%20new%20ground [accessed 30/05/2024]

[9] S. Chan-Olmsted, "A Review of Artificial Intelligence Adoptions in the Media Industry", The International Journal on Media Management, vol. 21, no. 3/4, pp. 193-215, 2019, https://doi.org/10.1080/14241277.2019.1695619

[10] B. Wilczek and M. Haim, „How artificial intelligence can increase the efficiency of media organizations?" MedienWirtschaft 4/2022, pp. 44-50.

[11] J. Heesen, C. Bieber, A. Lauber-Rönsberg, and C. Neuberger, Artificial intelligence in journalism. Potentials and challenges for media professionals. Whitepaper from the Learning Systems platform, München. Available from: https://doi.org/10.48669/pls_2023-1

[12] A. Schmidt, AI in the media industry: content creation, media design and production. MedienWirtschaft 1/2023, pp. 65-73

[13] WAN-IFRA / Schickler Report. 2023. Gauging Generative AI's impact on newsrooms. Survey: Newsroom executives share their experience so far. Available from: https://www.schickler.de/2023/05/generative-ki-erobert-die-newsrooms-neue-studie-mit-umfassenden-einblick/ [accessed 30/05/2024]

[14] A. Graefe and N. Bohlken. 2020. "Automated journalism: a meta-analysis of readers' perceptions of human-written in comparison to automated news", Media and Communiation, vol. 8, no. 3, pp. 50-59, 2020

[15] N. Thurman, S. Stares, and M. Koliska, "Audience evaluation of news videos made with various levels of automation: A population-based survey experiment", Journalism, https://doi.org/10.1177/14648849241243189, 2024

[16] J. Frank et al., "A Representative Study on Human Detection of Artificially Generated Media Across Countries". Cryptography and Security. https://doi.org/10.48550/arXiv.2312.05976, 2024

[17] Süddeutsche, Have you heard? The SZ articles are now read aloud, 2021. Avialable from: https://www.sueddeutsche.de/service/in-eigener-sache-schon-gehoert-die-artikel-der-sz-werden-jetzt-vorgelesen-1.5451594 [accessed 30/05/2024]

[18] NZZ, Neue Züricher Zeitung now also available to listen to, 2019. Available from: https://unternehmen.nzz.ch/2019/04/neue-zuercher-zeitung-ab-sofort-auch-zum-hoeren/ [accessed 30/05/2024]

[19] Rheinische Post. How the new read-aloud function works in the RP portal, 2023. Available from: https://rp-online.de/redaktion/rheinische-post-ki-klont-autoren-stimmen-fuer-vorlese-funktion_aid-100037205 [accessed 30/05/2024]

[20] Bitkom. News overload on the internet: Every second person feels overwhelmed, 2024. Available from: https://www.bitkom.org/Presse/Presseinformation/Online-

Journalismus-Nachrichtenflut-ueberfordert [accessed 30/05/2024]

[21] R. K. Merton, and P. L. Kendall, The focused interview. In: Hopf, C. (eds.): Qualitative social research. Stuttgart: Klett-Cotta,1984, pp. 171-204 .

[22] P. Mayring, Qualitative content analysis: basics and techniques, 2015. Beltz Pädagogik.