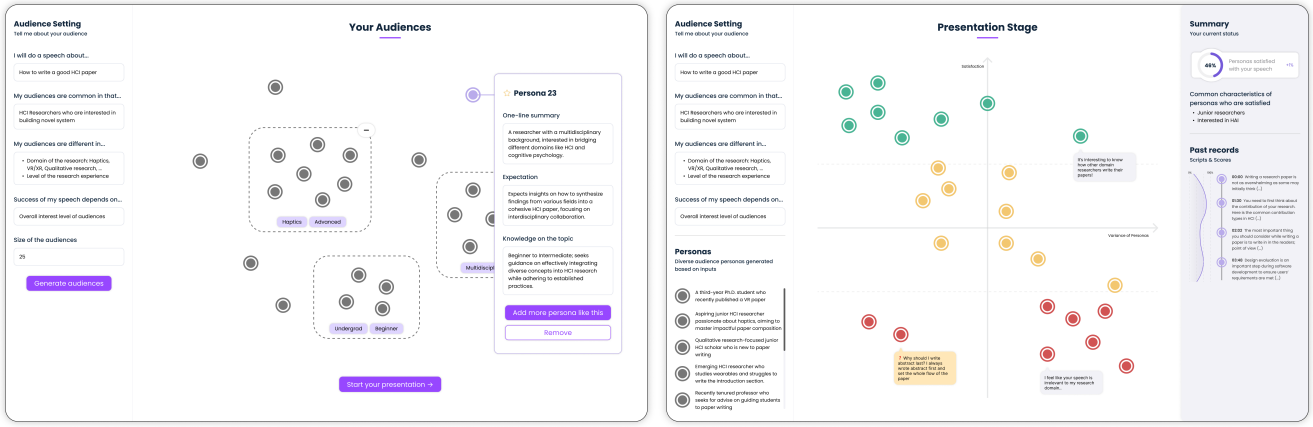
AudiLens: Configurable LLM-Generated Audiences for Public Speech Practice

Jeongeon Park*
School of Computing, KAIST
Daejeon, South Korea
jeongeon.park@kaist.ac.kr

DaEun Choi*
School of Computing, KAIST
Daejeon, South Korea
daeun.choi@kaist.ac.kr



(a) (b)

Figure 1: (a) Audience Generation & Configuration (b) Practice Speech Session with Real-Time Summary

ABSTRACT

AudiLens is a large-language model (LLM)-based audience simulator for public speech practice that allows speakers to generate and configure a group of generated audiences, and use them to receive feedback on their speech during and after the practice in multiple aspects. AudiLens leverages the capability of LLMs in being able to generate a diverse set of personas and being able to simulate human behavior, and provide flexibility to the speaker in terms of practicing their speech with multiple sets of audience groups in multiple speech formats. We demonstrate the use of AudiLens in two scenarios—giving a tutorial and debating.

CCS CONCEPTS

• **Human-centered computing** → **Interaction design**.

KEYWORDS

public speech, audience analysis, multi-agent interaction, LLM

*Both authors contributed equally to this research.

UIST '23 Adjunct, October 29-November 01, 2023, San Francisco, CA, USA

ACM ISBN 979-8-4007-0096-5/23/10.
https://doi.org/10.1145/3586182.3625114

ACM Reference Format:
Jeongeon Park and DaEun Choi. 2023. AudiLens: Configurable LLM-Generated Audiences for Public Speech Practice. In *The 36th Annual ACM Symposium on User Interface Software and Technology (UIST '23 Adjunct), October 29-November 01, 2023, San Francisco, CA, USA.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3586182.3625114

1 INTRODUCTION

The audience is an essential component of public speech [4]. In preparing for and executing the speech, many aspects are tailored to the audience [3, 4], including the structure, the type of content to include, and the level of engagement with the audience. For example, when explaining the field of human-computer interaction (HCI) to a group of elementary school students, the structure should be more straightforward, the contents should be explained with more metaphors, and the speaker should engage more frequently by asking questions along the way.

However, having a sufficient understanding of the potential audience and their reactions is challenging. While the speaker could have an understanding of the overall characteristic of the audience (e.g., students of an HCI course), variations could exist in a more complex way, such as their prior knowledge in the field and their motivation. This additional variance is often difficult to be taken into account by the speaker. In addition, it is difficult to predict how the audience would react to the speech, which imposes an additional challenge for the speaker to decide on the overall structure as well as specific details (e.g., where to explain further or what examples to include). To support a better understanding of the audience,

previous work has proposed guidelines on audience analysis [5] or suggested methods to understand and predict audience reactions [7, 9, 10]. But even with the existing supports, there still remains a challenge in replicating the real setting which sometimes is more complex and large-scale.

Inspired by successful applications of large-language models (LLMs) in simulating human behavior [1, 2, 6], we adopt the capability of LLMs in the context of public speech practice where we generate simulated audiences. We propose AudiLens, a system that enables generating different audience personas and utilizing personas through generated reactions during and after the simulated speech. The system generates a spectrum of audience personas based on the user's provided characteristics and allows users to flexibly configure the composition of personas. During the user's speech practice, each persona uniquely engages in the conversation. Their collective responses and on-demand details are presented, helping users to understand the efficiency of their speech and receive multi-level guidance on how to improve their speech.

2 AUDILENS

2.1 Audience Generation & Configuration

The initial step involves audience configuration, where users are prompted to define their intended listeners (Fig 1. (a)). Users can define the topic of their speech, the number of the audience, common attributes and diversity among the audiences, and the primary speech objective—such as the level of the audience's interest. Subsequently, the LLM generates diverse audience personas in response to the input. These personas are visually represented as individual dots within the main panel. Users can interact with these personas by clicking on them to access detailed information. This includes details commonly employed in audience analysis [8], such as a one-line description outlining the persona's attributes, their expectation of the speech, and their knowledge level of the topic at hand. After reading the details, users can either include more personas similar to the current one or remove the current persona from the audience. Additionally, the interface shows a cluster of similar audiences, represented by a tag denoting their common attributes. Users can adjust the overall audience using this cluster, such as by enlarging or reducing the cluster's size or opting to exclude the entire cluster from the audience composition.

2.2 Speech Session with Real-time Summary

When the user starts the speech, personas are visualized as dots on the main panel, arranged according to two key dimensions: *score* and *attribute variance* (Fig 1. (b)). The score shows how much of the goal previously set by the user is accomplished from each persona's point of view. The user can modify the variance dimension, enabling them to explore the impact of various attributes on the distribution of scores across personas. Furthermore, the user can uncover the rationale behind each persona's score by selecting the respective persona. Each persona periodically generates potential questions that could potentially arise from the audience.

The right panel displays the average score across the entire audience, offering users real-time insight into the current effectiveness of their speech. Moreover, a graph illustrates the change over time, accompanied by corresponding speech excerpts. This enables users to pinpoint sections of their speech that resonated convincingly with the audience and those that might require further refinement.

2.3 Implementation Details

AudiLens is implemented as a web-based application utilizing HTML, CSS, and JavaScript. There are two distinct LLM-based pipelines for AudiLens. The first pipeline focuses on generating diverse personas, driven by user input. The second pipeline is built to produce scores and reactions tailored to each persona's attributes. Additionally, the integration of a speech recognition model will enable us to capture user speech and utilize it as input for the system.

3 APPLICATION SCENARIO

3.1 Tutorial

For information-based speeches like a seminar talk or a job talk, the speaker often has to deliver the same speech across multiple settings. We anticipate the AudiLens to be able to support re-structuring or adjusting the speech according to the distribution of the audience.

Imagine a speaker who is giving a talk about the intersection between Human-Computer Interaction (HCI) and Artificial Intelligence (AI). The speaker only has an initial draft prepared, and the goal of the speaker is to motivate the importance of doing Human-AI Interaction (HAI) research. As the speaker is expected to present in an HCI department at a university, they first configure the audience to be HCI researchers but add variations in terms of their AI knowledge and career position. After the talk, the speaker realizes that most audiences found the talk enjoyable and engaging based on the summary from the AudiLens.

Afterwards, the speaker wonders how a more AI-focused audience would perceive the same talk. During audience configuration, they decide to re-configure the audience and add in more audiences who are not familiar with HCI. Seeing the reaction of the audience in different parts, the speaker adjusts the length of each section, to explain the importance of human research in more detail. The speaker continues to use AudiLens with audiences tailored to the institution they are presenting next.

3.2 Debate

We envision AudiLens expanding to speeches involving multiple users, for example for debate where two or more people are arguing with opinions from different sides.

Imagine a scenario where Speakers A and B are debating about implementing universal basic income (UBI). A and B both start the session by creating 50 audiences supporting basic income and 50 audiences going against basic income. The goal for both sides is to bring in as many audiences to their side as possible.

During the debate, Speaker A (Speaking for UBI) notices many questions and confusion arise from the opposing side after stating "UBI can acknowledge the labors of those underpaid." By specifically addressing the questions being asked, Speaker A realizes the need for careful treatment of this argument and includes the question in the main content of the speech to persuade more participants. After the debate, both speakers review the timeline to identify parts with significant opinion shifts. Speaker B realizes one of their claims drew in 20 audiences to their side, and decides to move the particular argument to the end for maximum impact.

REFERENCES

[1] Gati Aher, Rosa I Arriaga, and Adam Tauman Kalai. 2022. Using large language models to simulate multiple humans. *arXiv e-prints* (2022), arXiv–2208.

[2] Gati V Aher, Rosa I Arriaga, and Adam Tauman Kalai. 2023. Using large language models to simulate multiple humans and replicate human subject studies. In *International Conference on Machine Learning*. PMLR, 337–371.

[3] J Maxwell Atkinson. 1984. 15. Public speaking and audience responses: some techniques for inviting applause. (1984).

[4] Steven A Beebe and Susan J Beebe. 2009. e: Public Speaking: An Audience-Centered Approach, 11/e. (2009).

[5] Denis McQuail. 1997. *Audience analysis*. Sage.

[6] Joon Sung Park, Lindsay Popowski, Carrie Cai, Meredith Ringel Morris, Percy Liang, and Michael S Bernstein. 2022. Social simulacra: Creating populated prototypes for social computing systems. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*. 1–18.

[7] Bao Truong, Trung-Nghia Le, Khanh-Duy Le, Minh-Triet Tran, and Tam V Nguyen. 2022. Public Speaking Simulator with Speech and Audience Feedback. In *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. IEEE, 855–858.

[8] Department of Communication University of Pittsburgh. 2015. Audience Analysis. https://www.comm.pitt.edu/oral-comm-lab/audience-analysis.

[9] Eunkyong Lee Yook. 2004. Any Questions? Knowing the Audience through Question Types. *Communication Teacher* 18, 3 (2004), 91–93.

[10] Juan Diego Zapata-Rivera and Irvin R Katz. 2014. Keeping your audience in mind: Applying audience analysis to the design of interactive score reports. *Assessment in Education: Principles, Policy & Practice* 21, 4 (2014), 442–463.

A ETHICS AND BROADER IMPACT STATEMENT

As with any LLM-based system, inherent biases within the LLM may inadvertently introduce biased outcomes, potentially leading to the creation of personas or simulated reactions that reflect these biases. These biases could emerge from the training data, influencing factors like demographic representation or content emphasis. An additional concern is that users might excessively depend on the generated result, potentially inhibiting the development of natural communication skills.

Our strategies for mitigating these concerns are two-fold. Firstly, AudiLens empowers users by offering them control over persona composition and attribute variance, allowing them to customize and counteract any individual biases. Secondly, we will notify users that AudiLens is intended to complement, rather than replace, genuine communication proficiency.

By providing users with a platform to simulate diverse audience reactions, we anticipate that AudiLens provides the social benefit of developing public speaking skills, improving communication effectiveness, and reducing speaking anxiety in various public speeches.