INTERVIEW WITH AN AVATAR: A REAL-TIME ENGAGEMENT TRACKING-ENABLED CLOUD-BASED MULTIMODAL DIALOG SYSTEM FOR LEARNING AND ASSESSMENT

V. Ramanarayanan[1], P. Lange[1], D. Pautler[2], Z. Yu[3], M. Mulholland[4], and D. Suendermann-Oeft[1]

[1]Educational Testing Service, San Francisco, CA
[2]Intention Perception, LLC, Dunedin, New Zealand
[3]Carnegie Mellon University, Pittsburgh, PA
[4]Educational Testing Service, Princeton, NJ

vramanarayanan@ets.org

ABSTRACT

Recent advances in immersive computing technology have the potential to accelerate development of engaging intelligent agents that can guide one or multiple phases of learner instruction, learning, and assessment (both formative and summative) [1]. Such technologies are also important since they offer opportunities for personalizing the learning environment to each learner or trainee, providing a natural and engaging interface that can adapt to their individual strengths and weakness in real time so as to increase the efficacy of training [2]. In this spirit, we present a multimodal dialog system equipped with a virtual human avatar interlocutor. The agent, developed in Unity with WebGL support, and leverages the HALEF (Help Assistant–Language-Enabled and Free)[1] open-source cloud-based standard-compliant dialog framework [3]; see Figure 1. The system runs in a web browser on a multitude of devices (including laptops, smartphones, or tablets). It records video and speech data from users and, further, tracks and reacts to users' state, such as engagement, in real time, based on a combination of visual and verbal features. We designed and implemented a conversational job interview task based on the proposed framework. In this scenario, the avatar plays the role of an interviewer and reacts to user disengagement in real-time with feedback strategies designed to re-engage the user in the job interview process, thus providing a more holistic and immersive user experience.

Index Terms— intelligent tutoring, computer assisted language learning, multimodal dialog system, real-time engagement

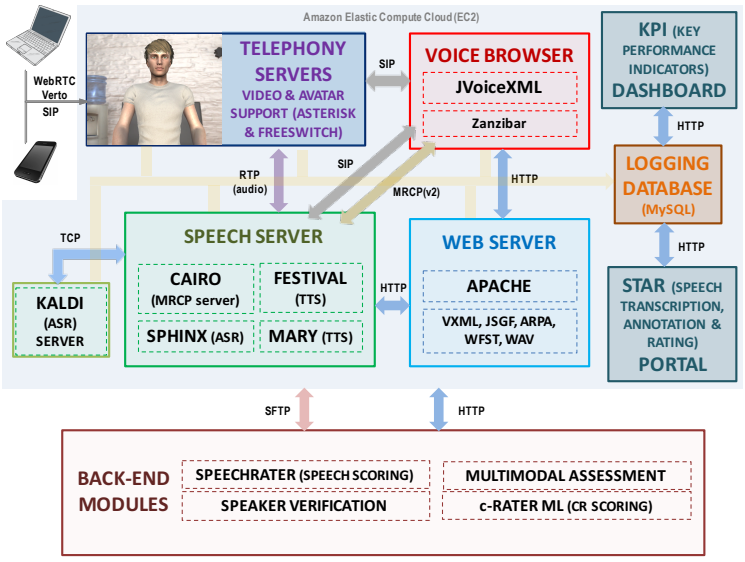

Fig. 1. The HALEF multimodal dialog framework with virtual avatar support for educational learning and assessment applications.

1. REFERENCES

[1] Jean-Claude Martin Bilge Mutlu Mohammed Hoque, Matthieu Courgeon and Rosalind Picard, "Mach: My automated conversation coach," in *Proc. of the 15th International Conference on Ubiquitous Computing (Ubicomp)*, 2013.

[2] Alice Kerry, Richard Ellis, and Susan Bull, "Conversational agents in e-learning," in *Applications and Innovations in Intelligent Systems XVI*, pp. 169–182. Springer, 2009.

[3] Vikram Ramanarayanan, David Suendermann-Oeft, Patrick Lange, Robert Mundkowsky, Aliaksei Ivanou, Zhou Yu, Yao Qian, and Keelan Evanini, "Assembling the jigsaw: How multiple w3c standards are synergistically combined in the halef multimodal dialog system," in *Multimodal Interaction with W3C Standards: Towards Natural User Interfaces to Everything*, p. to appear. Springer, 2016.

[1] http://halef.org