Exploring the Personality of Virtual Tutors in Conversational Foreign Language Practice

Johanna Dobbriner   **Cathy Ennis**   **Robert Ross**

School of Computer Science
Technological University Dublin
`{johanna.dobbriner,cathy.ennis,robert.ross}@tudublin.ie`

Abstract

Fluid interaction between virtual agents and humans requires the understanding of many issues of conversational pragmatics. One such issue is the interaction between communication strategy and personality. As a step towards developing models of personality driven pragmatics policies, in this paper, we present our initial experiment to explore differences in user interaction with two contrasting avatar personalities. Each user saw a single personality in a video-call setting and gave feedback on the interaction. Our expectations, that a more extroverted outgoing positive personality would be a more successful tutor, were only partially confirmed. While this personality did induce longer conversations in the participants, we found that interactions with both were enjoyed and that user perception of them differed less than intended.

1 Introduction

When learning a foreign language, Computer Assisted Language Learning (CALL) systems and virtual tutors can be a viable alternative for one-to-one conversational practice. Several of these systems have already been tried – embedded in video games, on their own or as part of a larger CALL system (Dalton and Devitt, 2016; Cheng et al., 2017; Wang et al., 2017; Collins et al., 2019; Divekar* et al., 2021). For these CALL applications, the automated tutor is usually embodied, for example as a virtual avatar with a dialogue system for conversing with the student.

There are several possible avenues to increase user engagement with a virtual avatar, one of which could be tailoring specific personality traits of the virtual tutor to the student. We see this tailoring as a manifestation of pragmatics modelling and choice as ultimately the conversational policy must be tuned to and react to the conversational traits of the user. As a first step towards investigating this hypothesis, we designed a Wizard-of-Oz (Kelley, 1984) pilot experiment to see whether there are any observable variations in the interaction and feedback when students are confronted with two opposing tutor personalities.
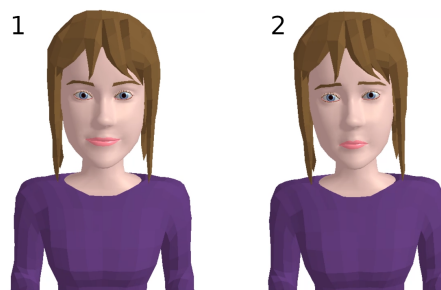

Figure 1: Avatar expressing different personalities – P1 (left) and P2 (right)

2 Experiment Design

For our experiment, each participant provided some demographic information before talking to one of the two virtual avatars for a few minutes in English followed by rating the avatar personality and giving feedback on the conversation itself.

To implement this study, we built upon the expressive avatar from Sloan et al. (2020), animated with an Irish English, female voice [1] to embody our virtual tutor. Our two avatar personalities varied along three of the OCEAN model's (Goldberg, 1990) five dimensions – extroversion, openness and agreeableness and were expressed through dialogue scripts where the avatar exhibited specific personality traits, posture, facial expression and speech characteristics (see Figure 1). Personality 1 (P1) represents the higher end of the 3 dimensions, being open, friendly and sociable, while personality 2 (P2) exhibits low values along the 3 dimensions,
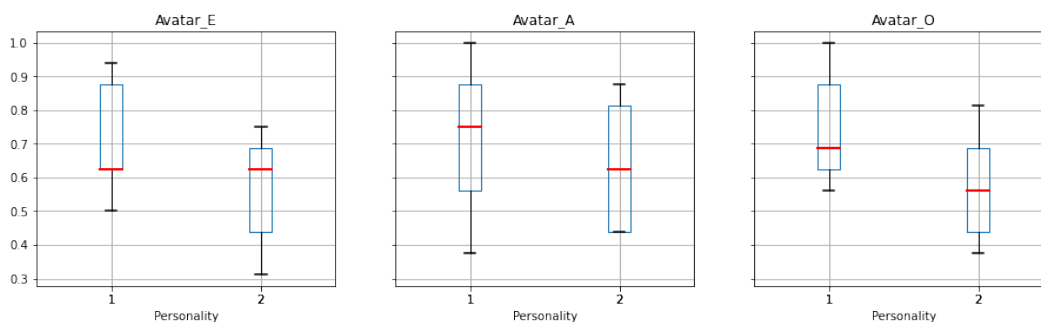
---

[1] `https://www.cereproc.com/en/node/1155`

Figure 2: Boxplots of Avatar personality scores between avatar personalities for Extroversion (Avatar_E), Agreeableness (Avatar_A) and Openness (Avatar_O)

behaving in a more closed off, curt and distant manner.

The participants of this experiment were adult English learners, 18 years or older, who had been learning English between 7 and 22 years ($M = 12.5$, $SD = 4.33$). 18 participants completed the study, 44% male and 56% female, aged 18 - 45 years. At 83%, the majority of participants spoke German as their native language with another 11% Italian and 5.6% (one person) Chinese. The total duration of the participant's interaction with the avatar varied significantly between avatar personalities with P1 interactions lasting an average of 12.15 minutes ($SD = 1.81$) and P2 conversations being markedly shorter at 7.55 minutes on average ($SD = 1.54$). With a sample size as low as this, any statistics computed on the collected data cannot be very robust and all results are to be taken as indicative.

3 Results

We expected P1 to be more pleasant and enjoyable to converse with, which we expected in turn would show itself in positive user feedback and high scores on an avatar personality survey, with low personality scores and fewer participants enjoying the interaction for P2. Additionally, we hypothesised that P1's conversation style would encourage participants to talk more than P2.

Our results as shown in Figure 2, only partly support this hypothesis: While P1 generally achieved higher scores, the box plot for Agreeableness shows much overlap between personalities. For Extroversion, the median is the same in both groups. The clearest distinction is found for Openness. A T-test at a significance threshold of 0.05 confirms these results, with Openness ($T(18) = 2.166$, $p = 0.046$) showing the only significant differ-

ence between P1 and P2, whereas Extroversion ($T(18) = 1.816$, $p = 0.088$) is close to significant and may prove distinct with more participants. Agreeableness ($T(18) = 1.099$, $p = 0.288$) was not perceived as significantly different, likely due to the avatar unintentionally interrupting the participants occasionally due to connectivity issues and human error. The speaking time ratio (human vs. avatar) showed the expected significant difference ($T(18) = 2.525$, $p = 0.022$), so P1 appears to encourage students to talk more than P2. In terms of user feedback, all participants noted they enjoyed the interaction regardless of avatar personality.

4 Conclusions and Future Work

In conclusion, this initial study opens the door to a multitude of interesting research directions that bear further exploration.

In this pilot study we found that only one out of three personality dimensions, Openness, was perceived significantly differently between both groups. Extroversion came close to the significance threshold and Agreeableness was widely distributed with relatively high scores in both groups. However, the participant's speaking time relative to the avatar was significantly higher in the personality designed to be more pleasant overall, thus matching our expectations.

Further research might focus on making the personality differences more apparent, building a chatbot for each personality, adding more personalities exploring measures to automatically adapt to the student.

Acknowledgments

References

Alan Cheng, Lei Yang, and Erik Andersen. 2017. Teaching language and culture with a virtual reality game. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 541–549.

Naoise Collins, Brian Vaughan, Charlie Cullen, and Keith Gardner. 2019. Gaeltechvr: Measuring the impact of an immersive virtual environment to promote situated identity in irish language learning. *Journal For Virtual Worlds Research*, 12(3).

Gene Dalton and Ann Devitt. 2016. Gaeilge gaming: Assessing how games can help children to learn irish. *International Journal of Game-Based Learning (IJGBL)*, 6(4):22–38.

Rahul R. Divekar*, Jaimie Drozdal*, Samuel Chabot*, Yalun Zhou, Hui Su, Yue Chen, Houming Zhu, James A. Hendler, and Jonas Braasch. 2021. Foreign language acquisition via artificial intelligence and extended reality: design and evaluation. *Computer Assisted Language Learning*, 0(0):1–29.

Lewis R Goldberg. 1990. An alternative "description of personality": the big-five factor structure. *Journal of personality and social psychology*, 59(6):1216.

John F Kelley. 1984. An iterative design methodology for user-friendly natural language office information applications. *ACM Transactions on Information Systems (TOIS)*, 2(1):26–41.

John Sloan, Daniel Maguire, and Julie Carson-Berndsen. 2020. Emotional response language education for mobile devices. In *22nd International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '20, New York, NY, USA. Association for Computing Machinery.

Yifei Wang, Stephen Petrina, and Francis Feng. 2017. VILLAGE - virtual immersive language learning and gaming environment: Immersion and presence. *Br. J. Educ. Technol.*, 48(2):431–450.