

1 Research Interests

My research focuses on **human-agent multimodal interactions** through virtual environments and **embodied conversational agents (ECAs)**. My research aims to improve naturalness and effectiveness in ECA interfaces by increasing **rapport** between agents and humans. Rapport is the feeling of mutual understanding, or the sensation of being “in sync” with another person. Replicating this sensation with a human and a virtual character is a complicated process, as it does not only involve the **verbal** expression of ideas, but additional **paralinguistic** meaning represented through nonverbal behaviors.

The purpose of these interfaces is to communicate in an unobtrusive, natural way, in this case speech and gesture, to complete nominal and non-nominal tasks. To accomplish this I have worked in a series of studies involving rapport and familiarity behaviors. Previously I did some research on multiparty turn taking behaviors between humans and agents. A summary of those studies in chronological order are described in the following subsections:

1.1 Grounding and Turn-Taking in Multimodal Multiparty Conversation

Before applying our knowledge into conversational agents, we conducted several experiments with natural conversations. One of these studies explored the empirical basis for multimodal conversation control acts. We applied conversation analysis as an exploratory approach in an attempt to illuminate the control functions of paralinguistic behaviors in managing multiparty conversation. We also contrasted our multiparty analysis with an earlier dyadic analysis and, to the extent permitted by our small samples of the corpus, contrast (a) conversations where the conversants did or did not have an artifact, and (b) conversations in English among Americans with conversations in Spanish among Mexicans. Our analysis suggested that speakers tend not to use gaze shifts to cue nodding for grounding and that the presence of an artifact reduced listen-

ers’ gaze at the speaker. These observations remained relatively consistent across languages.

1.2 Multiparty Turn Taking

This is an exploratory study of what happens when a person enters a room where people are conversing, based on an analysis of 61 episodes drawn from the UTEP-ICT cross-cultural multiparty multimodal dialog corpus. We examined the reliability of coding of gaze and stance, and we develop a model for room-entry that accounts for observed differences in the behaviors of conversants, expressed as a state-transition model. Our model includes factors such as conversational task, not considered in existing social-force models, which appear to affect conversants’ behaviors (Novick and Gris, 2013). An ECA implementation of this model was developed shortly after, in which the agents are conversing and they integrate a person to the conversation when someone enters the room they are being projected.

1.3 Familiarity

These agents were developed with additional capabilities and tasks. This is part of a longer-term project to provide ECAs with behaviors that enable them to build and maintain rapport with their human partners. It is both, the complement and prequel of my dissertation work. In this study we focused on paralinguistic behaviors and their role in communicating rapport. This study piloted the investigation of how to signal increased familiarity over repeated interactions as a component of rapport. We studied the effect of differences in the amplitude of several behaviors by an ECA interacting with a human across two conversational sessions (Novick and Gris, 2014). Our main question was whether subjects would perceive more rapport with the agent in the increased-familiarity condition in the second session as having higher rapport. Interestingly, the results were the opposite of those hypothesized. We would later realize that this likely originated from a slight mismatch between verbal and nonverbal actions, as the voice remained constant while the amplitude of the animations increased. This is a great example of the importance of emotion on SDSs, in

particular, for those applied to full-body conversational agents.

1.4 Virtual Rapport

This study is part of my dissertation and thus, my future work. To create the sense of connection, mutual understanding and proper coordination, our agent, Adriana, leads users through a series of activities and conversations. These conversations take place in a survival scenario setting, where you have to collaborate, cooperate, and build a relationship with the ECA to survive in an uninhabited island.

The simulation is being built with the intention to maximize rapport building opportunities, as well as to take advantage of the verbal and non-verbal behaviors in a more immersive environment, where both, the user and the agent can interact with the same objects in virtual space. It provides a flexible storyline that allows decision making, without being an open environment (Gris, 2013).

2 Future of Spoken Dialog Research

Spoken dialog systems will likely be integrated into a wide set of devices in the near future. As we improve speech recognition and tackle practicality issues, dialog systems will hopefully gain popularity.

One important key feature of advanced dialog systems is the use of paralinguistic expression to add additional meaning into the utterances. This would be particularly helpful in distinguishing between identical phrases that can have additional meanings based on the intonation.

In addition, it would be great to see SDSs that are capable of collecting information from previous conversations, and then use it for future interactions. Although I am confident that these systems have several marketing applications, I would prefer to see a more practical implementation. In other words, it would be a great stepping stone towards embodied agents capable of convincingly representing the roles of virtual assistants, trainers, therapists, instructors, or conversational partners.

3 Suggestions for discussion

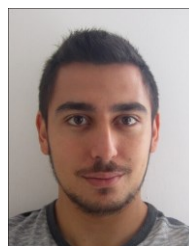
- Semi-scripted dialog systems for storytelling with virtual agents
- Best practices for developing SDSs with graphical components
- Matching of verbal and nonverbal behaviors
- Conversational recall pros and cons

- Social forces and roles applied to SDSs
- Entertainment applications for dialog systems

References

- Gris, Ivan. "Adaptive virtual rapport for embodied conversational agents." *Proceedings of the 15th ACM on International conference on multimodal interaction*. ACM, 2013.
- Novick, David, and Iván Gris. "Grounding and turn-taking in multimodal multiparty conversation." *Human-Computer Interaction. Interaction Modalities and Techniques*. Springer Berlin Heidelberg, 2013. 97-106.
- Novick, D., Adoneth, G., Manuel, D., & Gris, I. (2012). "When the Conversation Starts: An Empirical Analysis". *Joint Proceedings of the IVA 2012 Workshops*. Santa Cruz, USA: Otto von Guericke University Magdeburg (pp. 67-74).

Biographical Sketch



Ivan Gris is a Ph.D. student at the University of Texas at El Paso working under the supervision of Dr. David Novick. He works at the Interactive Systems Group, as part of the Advanced Agent Engagement Team, developing full body embodied conversational agents and immersive, interactive environments.

Previously he worked for Cyber-ShARE, an interdisciplinary research center, where he developed applications for geological scientific visualization, and high performance computing to render complex models in near-real time.

When he is not working in his dissertation, he is working on the research and development of two small companies he created. One of these companies is a tech startup that uses embodied conversational agents, animatronics, scene development, and visual & special effects to develop a unique role playing themed experience.