2025 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW) | 979-8-3315-1484-6/25/$31.00 ©2025 IEEE | DOI: 10.1109/VRW66409.2025.00448

Demo: Automated Multiparty Conversations Analysis in VR Massive Multiplayer Online Games Environments

Riccardo Calcagno*[1], Elio Salvadori[1], Oscar Mayora-Ibarra[1], Giovanna Varni[2], and Luca Turchet[2]

[1]Fondazione Bruno Kessler, Italy
[2]University of Trento, Italy

Abstract

The study of social interactions in Massively Multiplayer Online Games (MMOGs) has largely focused on chat communication and speech content analysis. This paper introduces a novel tool designed to analyze social interactions among MMOG players by leveraging Conversation Analysis methodologies to systematically assess multi-party interactions during 6DOF Virtual Reality sessions. The Multiparty Conversation Analysis (MCA) tool can be applied in several MMOG contexts where social analysis is needed. In this demo paper, we present a specific application of the tool focusing on VR MMOG-based therapy for autistic teens. A video demo can be seen in: https://youtu.be/U4A4vkjUrvM

Index Terms: Massive Multiplayer Online Games, VR Environments, Multiparty Interactions, Conversation Analysis.

1 Introduction

Massively Multiplayer Online Games (MMOGs) have become a significant area of interest for researchers studying social interactions in virtual environments. Traditional approaches to analyzing these interactions have predominantly focused on chat communication and speech content analysis. Previous studies have explored the social aspects of digital gaming, emphasizing the importance of various forms of human interaction in virtual worlds or have analyzed linguistic and social interactions in MMOGs, highlighting the role of collaborative dialogue in fostering social cohesion [1]. Other studies have proposed tools and methodologies to analyze player emotional cues, such as the work by [2], which examines the affective aspects of player interactions. Other works leveraged Conversation Analysis (CA) [3], a methodology that systematically assesses the qualities of multi-party interactions such as initiation and response patterns, turn-taking dynamics, gaps and overlaps, e.g. to extract turn-taking cues in single-dialog conditions [4] or to analyze fusions and splits of simultaneous dialogues [5]. However, to the best of our knowledge, there are no previous studies that have been automating CA features extraction respecting the peculiarities of naturalistic-chaotic group interactions ($n > 3$, with multiple simultaneous dialogues) in Virtual Reality-based MMOGs.

This paper describes a prototype of a novel tool designed to analyze social interactions among MMOG players in VR environments by leveraging multi-party CA. It stands out from existing literature by focusing on non-verbal aspects of communication, such as conversation volume and proxemics, rather than solely on text or speech content analysis. The tool provides a comprehensive analysis of player interactions by capturing real-time localized verbal units (such as dialogues), turn-taking data, and navigation cues and then extracting statistical data highlighting individual and group social cues.

*e-mail: rcalcagno@fbk.eu

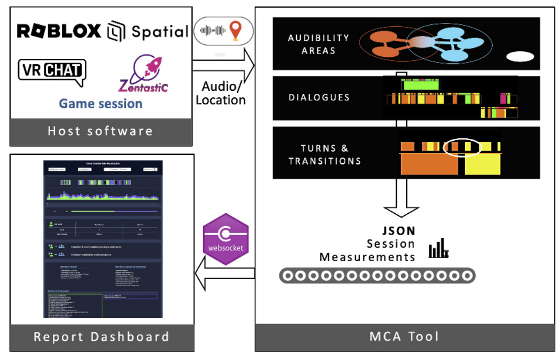

Figure 1: System overview: the MCA Tool as a plug-in of the hosting software and its interface to the Report dashboard

In particular, we present a specific application of the tool focusing on VR MMOG-based therapy for autistic teens. The preliminary evaluation demonstrates its potential as an effective mechanism for therapists to automatically assess and get key insights about social skills improvements while reducing their post-session workload.

2 System Description

Figure 1) describes the main system components:

1. The MMOG software hosting the gameplay sessions and providing the immersive VR context where interactions between avatars' players take place.
2. The **Multiparty Conversation Analysis (MCA) tool** which acts as a plug-in of the host software. The tool is responsible for analyzing the gameplay session in real-time by capturing the datastream about the audio and location of each avatar, and identifying dialogues among players according to their relation of *virtual audibility*[1], their turn-taking features (turns & transitions as defined in [3]) within a session.
3. A **Report dashboard** developed using Javascript technologies and designed to provide a clear and interactive visualization of the metrics and statistics generated by the MCA tool.

MCA Tool

This is a modular software that runs in background during the gameplay session among a group of players interacting through their *avatars*. The tool is a Unity asset written in C# that can be

[1]A is *audible* by B if the closest A's speech unit has sufficient volume (RMS) to be heard by B depending also on the volume attenuation function of B's *AudioSource* and the distance between A and B.

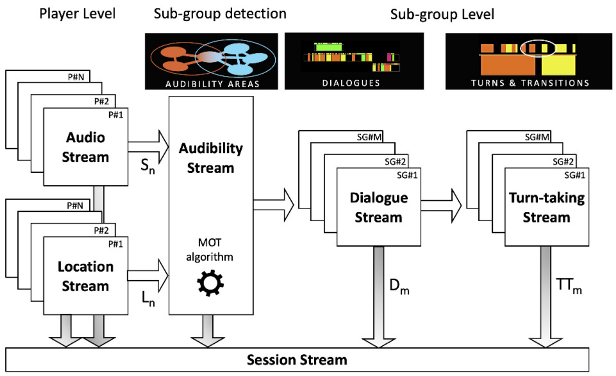

Figure 2: Details of the CA Pipeline Layer

easily integrated within any VR 6DOF application with minimal integration effort. It is based on a Conversation Analysis model divided into two different abstraction layers:
- the **Core** layer defines the abstract types involved in cascading *features*[2] computation such as: 1) the accepted streams in input and their *Unit*, i.e. the base abstract model for each specific feature; 2) the functions and states necessary to build the output *Unit*; 3) the choice of fusion mechanisms of input streams, which includes: temporal synchronization of streams, definition of conditional triggers of computation based on the arrival of a proper batch of *Units*.
- the **CA Pipeline** layer is composed by the following *feature* streams chained in a tree-like dynamic structure (see Figure 2):

1. Multiple *SpeechStream*, one per avatar, delivering *Units* defining speech periods and associated volume;
2. Multiple *LocationStream*, one per avatar, delivering its discrete movements (start & final location);
3. One global *AudibilityStream* which, by synchronizing and computing inputs from all *SpeechStream* & *LocationStream*, is responsible for detecting and tracking all the sub-groups of avatars in which there are the minimal heuristic conditions for a *floor*[3] to exist among them. By design, the members of such sub-groups are the nodes of the largest full-connected graphs of players sharing the perpetually assessable relation of *virtual audibility*. By "tracking the sub-groups" it is intended to track them consistently throughout time with a custom Multiple Object Tracking (MOT) algorithm to ensure that fusion and split of *floors* are handled properly.
4. Multiple *DialogueStream* instantiated when processing a new identifier of all sub-groups identified and transmitted by the *AudibilityStream*. Until the identifier stops coming in as input, the *DialogueStream* inspects the *SpeechStream* of each member of the sub-group to orchestrate the construction and finalization of new "dialogue" *Units*, thus reducing the multi-party problem into a number of multiple single-party problems.
5. Multiple *TurnTakingStream*, each corresponding to a distinct *DialogueStream*. The stream, by inspecting a "dialogue" *Unit* and the speech periods of its members, is responsible for modeling the series of turns and transitions shaping the *floors* within a dialogue.
6. One *SessionStream*, horizontal across all streams, and responsible for the specialized accumulation of real-time data to statistically describe certain group- or individual-coded behaviors. A JSON report is delivered at the end of the session.

[2]Features can be speech, movement, dialogue, turn, etc...
[3]The term refers to the space where one player has the right to speak while others listen [3]. Holding the floor means having control over the conversation.

Report dashboard

At any point during the session, the web dashboard can be accessed to review the accumulated statistics such as:

1. *Dialogues Timeline* presents the session's dialogues as segments arranged chronologically. Each segment includes details on speaker turns, and the voice units it comprises.
2. *Voice Activation Timeline* consist in a timeline highlighting each player's vocal contributions computed over 3" windows.
3. *CA Foundational Statistics* describes the frequency and duration of dialogues and speaking turns for the selected player.
4. *Dialogue-Inducing Spatial Movements* gives insights into the frequency of avatar movements aimed at initiating a dialogue in relation to the selected subject.

3 Preliminary Evaluation

A preliminary assessment of the MCA tool has been performed to explore its impact in supporting therapists to monitor VR MMOG-based therapeutic sessions with autistic teens. The tool was integrated into Zentastic [6], a VR serious game developed by MEEVA Srl (an FBK spinoff) aimed at delivering Social Skills Training (SST) from remote. The performed tests proved the MCA tool can successfully operate as a background task without altering the game's dynamics or performance. After each game session, therapists could access the Report dashboard to review post-session statistics. A preliminary evaluation study involving three therapists with sound experience in VR-based group treatments reported that the MCA tool was deemed beneficial to assess the session intervention efficacy by unveiling quantitative insights previously inaccessible using traditional qualitative-based strategies (gold standard).

4 Conclusions

This paper describes a novel tool leveraging a multi-party Conversation Analysis model designed to analyze interactions among MMOG players and extract statistical data highlighting social cues. While the preliminary evaluation has been focused in the area of VR-based group therapies for autism, further application domains can be envisioned in perspective, e.g. cyber-bullying detection; user engagement boosting via group-based preferences prediction.

Acknowledgments

This work was funded by Hub Life Science – Digital Health (LSH-DH) PNC-E3-2022-23683267 - Project DHEAL-COM - CUP C63C22001970001, Ministry of Health (Italy) under the Piano Nazionale Complementare al PNRR Ecosistema Innovativo della Salute (Code PNC-E.3).

References

[1] T. Quandt et al. *Multiplayer – The Social Aspects of Digital Gaming*. Routledge Editor, 1 edition, 2014. 1

[2] H. Chen et al. Tell me how you play: Exploring ways to enhance the gaming experience in asymmetric multiplayer VR games through affective state visualization. In *ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 2024. 1

[3] H. Sacks et al. A simplest systematics for the organization of turn taking for conversation. In *Studies in the Organization of Conversational Interaction*, pages 7–55. Academic Press, 1978. 1, 2

[4] P. Anna et al. Conan: A usable tool for multimodal conversation analysis. *ACM ICMI '21: International Conference on Multimodal Interaction*, 2021. 1

[5] M. A. Paul et al. Where's the "party" in "multi-party"? Analyzing the structure of small-group sociable talk. *ACM Conf. on Computer Supported Cooperative Work*, 2006. 1

[6] S. Gabrielli et al. Co-design of a Virtual Reality multiplayer adventure game for adolescents with Autism Spectrum Disorder: Mixed Methods Study. *JMIR Serious Games*, 2023. 2