
Polat, Alper

# A micro-analytic investigation of gaming interactions: displays of understanding in Overwatch

Master's Thesis

FACULTY OF EDUCATION

Master's Degree Program in Learning & Educational Technology

2020

University of Oulu
Faculty of Education
A micro-analytic investigation of gaming interactions: displays of understanding in Overwatch (Alper Polat)
Master's Thesis, 62 pages, 1 appendix
July 2020

This thesis examines the gaming interactions of players in Overwatch, a multiplayer, fast-paced, first-person shooter video game. The data contains players' practice matches against other teams to improve their team synergy, and to try out new additions to their roster. Two hours of their video-recorded, naturally-occuring gaming interactions in English are analyzed using Conversation Analysis to uncover insights on how players achieve intersubjectivity (i.e. mutual understanding). As this phenomenon has not been subject to research in gaming contexts yet, the analytical focus is on displays of understanding in second turn (i.e. a turn that is addressed to its prior turn) during both talk and action sequences. Four types of displays of understanding are identified and discussed in detail: non-understanding, verbal display, embodied display, as well as embodied and verbal display. The study also discusses players' engagement in pre-sequential planning actions, their use of in-game jargon and shorthand expressions. The findings contribute to the growing body of research on gaming interactions, and to intersubjectivity research within social interaction. The analyses bring evidence to players attaining more in-game success when mutual understanding is achieved, and players being more prone to in-game failures when it is not. Furthermore, discussions regarding players' speaker selection and repair practises to overcome interactional troubles are made.

Keywords: displays of understanding, intersubjectivity, multiplayer, gaming, social interaction, conversation analysis

# Contents

# 1 Introduction

For a couple of decades, video games have been a significant source of entertainment for many people. Game developers have introduced many different types of games (e.g. first-person shooters, multiplayer online battle arenas, real-time strategy games, sports, and so on). Each different type has various in-game objectives and they thus provide rich opportunities for social interaction to take place, as well as problem solving and collaboration skills to be used (Bluemink & Järvelä, 2011; Gee, 2006). More importantly, as many games are mainly in English, this means that these skills are manifested in either a second or foreign language by many players (Sylvén & Sundqvist, 2012). Given this case, the research into gaming contexts would reveal significant insights on how English is used for problem solving and collaboration outside of classrooms (Chen & Huang, 2010; Hung, 2007; Sylvén & Sundqvist, 2012). These insights then can be used in supporting extra-curricular activities in schools, lesson planning, and curriculum development (Caldwell, Osterweil, Urbano, Tan, & Eberhardt, 2017; Olsson, 2012; 2016; Sundqvist, 2009).

For successful problem solving and collaboration, it is important that group members within teams achieve mutual understanding (Järvenoja & Järvelä, 2009; Paans, Onan, Molenaar, Verhoeven, & Segers, 2019; van de Sande & Greeno, 2012). However, group members do not always reach a mutual understanding easily (van de Sande & Greeno, 2012). This problem gains a new shape in multiplayer video games that require mobile actions in continuously changing contexts (Mondada, 2013). In addition, the increased amount of in-game tasks with a time constraint may make it more challenging to reach mutual understanding. To overcome this challenge, players of multiplayer video games timely coordinate their in-game actions (Mondada, 2013; Piirainen-Marsh & Tainio, 2014; Reeves, Laurier, & Brown, 2009). These timely coordinated in-game actions "not only concern the moment-by-moment unfolding of actions but also the participants' anticipations, based on their skilled interpretation in real time of projected trajectories of actions, both in talk and in the game, as well as the participants' planning of game strategies within more extensive lines of conduct." (Mondada, 2013. p. 302). Some examples to this phenomenon entail actions such as movement, grouping, attacking, defending, using abilities, and many more, all of which are carried out by player controlled in-game 'avatars' or bodies (Bennerstedt, 2008; Reeves et al., 2009). Other than these bodily actions, players also use voice chat software to verbally communicate with each other real-time,

during which they typically engage in game-related, goal-directed interactions such as communicating about the objectives of video games (Moore, Ducheneaut, & Nickell, 2007). These verbal and bodily interactions together constitute significant interactional resources, through which players achieve 'intersubjectivity', a term that refers to mutual understanding between speakers within a given context (see e.g., Schegloff, 1992).

However, even though these timely coordinated interactions are important for achieving mutual understanding, they have not been subject to detailed research (Mondada, 2013). To fill this gap, this study uses Conversation Analysis (CA) as the research methodology. CA is a research methodology that aims to "describe, analyze, and understand talk as a basic and constitutive feature of human social life" (Sidnell, 2010b. p.1). Additionally, embodied actions are included within its framework (Mondada, 2014, 2019; Rauniomaa & Keisanen, 2014). CA provides a detailed, moment-by-moment analysis that focuses on the sequentiality aspect of social interaction, which makes it an effective methodology to inspect how mutual understanding is achieved in social interaction (Schegloff, 1992; Sidnell, 2010). To these ends, this thesis will aim to shed light on the timely coordinated interactions through which players attempt to arrive at mutual understanding as they play a multiplayer first-person shooter game called Overwatch. Detailed observations will be made of players' verbal interactions intertwined with their in-game avatars' bodily interactions. This will be done because in the game world players are represented by their in-game avatars and they orient themselves to their avatars' bodily interactions as if they are real bodies.

In order to understand how players arrive at mutual understanding in gaming contexts, the present study focuses on 'displays of understanding' in the second turn (i.e. a turn that is addressing a prior turn) within in-game talk and action sequences (e.g. talk-centered planning situations, fast-paced combat situations). Furthermore, the cases in which understanding is not displayed are also discussed. Displays of understanding refer to 'relevant' verbal turns and/or embodied actions deployed by participants to demonstrate that they have understood what is going on in the interaction (see e.g. Sidnell, 2010. pp. 59-76). More detailed explanations of this phenomenon will be provided in the next chapter alongside the gaming behavior of players. The research questions guiding the analyses are:

*1- What types of display of understanding do participants use in Overwatch?*

*2- What are the interactional resources used by the participants during display of understanding moments?*

*3- How does displaying understanding affect players' in-game success?*

In the third chapter, the research methodology, its methodological principles, as well as the data, and transcription procedures will be further elaborated on. After that, in chapter 4 the analyses and findings will be presented. Lastly, in chapter 5, the importance of these findings, limitations of the present study, and suggestions for future research will be discussed.

# 2 Literature Review

In this chapter, I will set the grounding of this thesis based on discussions of the literature in two sub-chapters. With the first sub-chapter, I will discuss the CA work on understanding in interaction. In the second sub-chapter I will discuss research on gaming interactions.

## 2.1 Understanding in interaction

In CA, "understanding" can be understood by looking at the reciprocal relationship between two turns-at-talk. For instance, when the person A asks "Would you like to drink something?" a potential answer by the second person B can be a range of things like accepting/declining "Yes/no", or specifying what B would like to drink "Water, soda" etc. instead of what might sound odd to A such as "I am fine, how about you", which could be more relevant if A was asking "How are you?". So, here it is seen that in the first turn that A produced, s/he has made relevant a set of responses for B to choose from, and respond. The second speaker B, while responding to A's question with a relevant response, will, therefore, show his/her own understanding of A's turn simultaneously. Furthermore, if in the second turn there is an unexpected response, A might extend the talk by another turn and try to fix the issue by doing a range of things such as rephrasing, clarifying, asking another question, and so on. In other words, understanding is concerned with "... what a second participant *does* in response to a first participant" (Mondada, 2011. p. 543). Participants, in the second turn address the first, continuing with actions that show their understanding (Heritage, 2009). Consequently, actions can be interpreted as what participants do with their turns-at-talk (Drew, 2013; Sidnell, 2010). For instance, when someone asks "are you thirsty?" the purpose of the turn could be merely an inquiry, or to offer something to drink, when the ensuing interaction is observed the purpose can be uncovered.

Within CA, this aforementioned relation of two turns-at-talk is referred to as "adjacency pairs" (Schegloff & Sacks, 1973). An adjacency pair consists of two pair parts: a first pair part, and a second pair part. Examples of such paired actions are questions-answers, greetings-greetings, offer-accept/decline, and alike (Schegloff, 2007). When a first pair part has been produced, it makes the second pair part relevant, as Schegloff (1968) notes: "given the first, the second is expectable; upon its occurrence it can be seen to be a second item to the first; upon its nonoccurrence it can be seen to be officially absent - all this provided by the occurrence of the

first item." (p. 1083). For example, these are actions that can be questions-answers, greetings-greetings, offer-accept/decline, and alike (Schegloff, 2007). After the second pair part has been completed, the first speaker can do one of two things: accept the response and continue, or point to a trouble source. This occurs in the next turn after the second pair part, which is called "third position" (Heritage, 1984). If a response in the second pair part is not in line with the first pair part, it will likely be treated as a problem source by the first speaker and be subject to a repair attempt (Schegloff, Jefferson & Sacks, 1977; Schegloff, 1992) in the third position (Heritage, 1984) so that understanding can be restored. Repairs refer to interactional mechanisms deployed by participants to overcome interactional problems that may stem from speaking, hearing, and understanding (Hayashi, Raymond & Sidnell, 2013). For instance, repair can be in the form of asking a clarification question, asking for repetition and so on. Furthermore, participants hold each other accountable for putting the relevant actions turn-by-turn, which is called sanctioning. As Sidnell (2010) asserts: "Adjacency pairs allow then for a framework of understanding that is constructed and sustained on a turn-by-turn basis" (p. 66).

In social interaction, speakers carry out actions not *just* by speaking, while they speak they simultaneously use various resources of the body (Goodwin, 2000; Hindmarsh, Reynolds, & Dunne, 2011; Lindwall, 2014), such as gestures (Mori & Hayashi, 2006; Belhiah, 2012), and gaze (Korkiakangas & Rae, 2014). Prior studies have shown how embodied resources are consequential for achieving understanding in interaction (Goodwin, 2000; Hindmarsh, et al., 2011; Lindwall, 2014; Mondada, 2014b; Mondada, 2019). For instance, a second pair part in an adjacency pair may include embodied actions to demonstrate that the second speaker has understood the first pair part, and provided a relevant response (e.g. nodding, see Stivers, 2008; 2010; Stivers, Mondada & Steensig, 2011 for more instances). Furthermore, a second pair part may be entirely made up of embodied actions (Rauniomaa & Keisanen, 2012). Therefore, this 'adjacency' should not be necessarily limited to turns-at-talk, but include embodiment as well (Lindwall & Ekström, 2012).

Building on these earlier research, we can claim that the understandings shown in the second pair part can have characteristics such as (1) verbal, (2) embodied, (3) both verbal and embodied. Furthermore, there can be (4) non-understandings or misunderstandings that can lead to troubles in interaction (e.g. delays, silences, etc.). More specifically, non-understandings are characterized by a missing second pair part or a relevant second action whereas, in misunderstandings the second pair part exists, but it does not provide the expected response to

the first pair part due to reasons such as misinterpretation of the first pair part (see Schegloff, 1987 for more examples).

In sum, we know that understandings are made visible in the second pair part using a variety of interactional resources, and are treated as sufficient or insufficient (e.g accepted, declined, repaired, etc.) in the third position. But how is understanding displayed? While displaying their understanding, the speakers tend to either merely claim that they have understood (e.g. "oh-prefaced turns" see Heritage, 1984), or demonstrate their understanding (e.g. make an interpretation). Consider the following example from Sacks for more clarification on this matter:

(Sacks, 1992:II:141)

| | | |
|---|---|---|
| 1 | A: | where are you staying |
| 2 | B: | Pacific Palisades |
| 3a | A: | oh at the west side of town |
| vs | | |
| 3b | A: | oh Pacific Palisades |

We see that in 3a, A makes an interpretation of what B said in line 2 by re-phrasing the location. By doing so, A demonstrates understanding and this enables B and the analyst to infer that A knows where Pacific Palisades is. Whereas in 3b, A simply claims understanding by merely repeating the said location. Hence, from 3b we cannot make analytical claims that A knows where the said place is located. I will establish my observations based on this explanation of displaying understanding. My analysis will be concerned with demonstrations of understanding as "displays of understanding" rather than claims, for claims arguably do not provide a fully relevant action in the second pair part when in-game actions are considered.

In normal conversation, speakers treat understandings as taken-for-granted, which means that they do not explicitly orient to understandings unless there is a problem with them (e.g. a person would not ask "do you understand?" unless they pick up a set of clues that point to problems in the second speaker's turn). In other words, it can be said that even though understandings are almost always there in the conversation, they have an invisible status (Schegloff, 1992; Seedhouse, 2004; Mondada, 2011). However, the analyst can access and analyze understanding by looking at the sequences in which turns and actions are produced (Seedhouse, 2004;

Mondada, 2011), or by looking at the preceding turn that made the second turn relevant (Sidnell, 2010). In the analysis section, I will be documenting the gaming interactions in a similar way.

To summarize, I have reviewed how "understanding" is perceived and studied in CA research. The emphasis has been on adjacency pairs framework. Within this framework, understanding is displayed in the second pair part to a relevant first pair part (e.g. question, request, summons, directive, instruction, etc.) by using different interactional resources (e.g. verbal, embodied). Moreover, I have shown the difference between claims and demonstrations of understanding, to which I will further refer when considering displays of understanding in my analysis. Lastly, I briefly elaborated on the status of understanding in conversation (i.e. taken-for-granted), and explained how my analysis is going to treat this phenomenon while exploring understanding in gaming interactions. Next, I review the literature on gaming interactions and point out the research gap I will try to address.

## 2.2 Interaction in multiplayer video games

This section will address studies of interaction in multiplayer video games. That said, first it is relevant to first consider what players generally do in games, then consider particular properties of physical location of players, for it will have an impact on how they interact with each other. I will start by noting the players' general in-game behavior regardless of setting. Then, I will move on to describe the setting-specific affordances for interactions (i.e. what the players can do while interacting). For that purpose, I will discuss co-present as well as geographically dispersed settings. At the end of this chapter, I will provide a brief summary of this chapter and point out the research gap.

In multiplayer video games, players come together in virtual worlds (e.g. a map or game world) (Reeves, et al., 2009; Berger, Jucker, & Locher, 2016) to constantly carry out coordinated actions such as "grouping", "fighting", "moving", and "waiting" (Bennerstedt & Ivarsson, 2010), and many more to fulfill in-game tasks such as defeating an enemy team or a group of in-game monsters. While doing so, each player takes on different roles (e.g. tank, damage, healer) (Bennerstedt & Linderoth, 2009), and attempt to competently execute their role-specific actions to assist their teammates (Ducheneaut & Moore, 2005). During their coordination, the players refer to each other in various ways. Players either address (1) the player directly, (2) their in-game avatar, (3) the player and their avatar together (Mondada, 2013; Baldauf-

Quilliatre & Colón de Carvajal, 2015). They use in-game jargon to guide their in-game actions more concretely and smoothly (Wright, Boria & Bradenbach, 2002; Bennerstedt, 2008). Lastly, after long hours of playing with each other, the players are shown to develop an "interactional synchrony", a rhythm that allows them to coordinate their actions timely and effectively (Piirainen-Marsh & Tainio, 2014).

When players are co-present with each other, they tend to play games on either the same screen (see; Aarsand & Aronsson, 2009; Piirainen-Marsh & Tainio, 2009; Mondada, 2011b, Mondada; 2013), or they set up multiple screens in the same room (see Sjöblom, 2008; Keating & Sunakawa, 2010). This means that the way they interact with each other does not get limited to the on-screen, in-game actions. In other words, they still use their real bodies (i.e. embodied actions such as gestures, gaze, etc.) and talk to each other (i.e. verbal actions) when interacting with each other. While playing, players use a range of verbal resources such as prosody, deixis (e.g. pronouns), repetition, as well as embodied resources such as gestures, gaze, pointing (Keating & Sunakawa, 2010). These interactional resources are used in activities such as questioning, analyzing, observing, or commenting on the co-present players' gameplay (Sjöblom, 2008; Tekin & Reeves, 2017). When the game is suspended, players retract to rest or do other real-life activities (Mondada, 2011b; Mondada, 2013).

To summarize, we can infer that when co-present, players are able to share the same screen or have access to their co-present players' screens. Hence, they can use their real-life bodies more effectively while interacting with each other, which can eliminate many obstacles in communication. For instance, while referring to something on-screen, they can point at it to clarify any misunderstandings, or they can make sure that their co-present player is looking at the same thing/same direction as they are, for "looking" can be significant for in-game collaboration (see Reeves et al., 2009). Moreover, because they can talk to each other in real-time and convey their message coupled with these embodied actions, we can claim that the players have more interactional resources to work with than geographically dispersed settings.

While players are geographically dispersed, they do not have access to each others' screens physically. Yet, they can still see each others' avatars and where these avatars are looking, or which embodied actions these avatars are carrying out as far as the game in question allows (see Manninen & Kujanpää, 2005; Bennerstedt, 2008; Bennerstedt & Linderoth, 2009). For instance, a player can interact with other players via their in-game avatars' embodied actions

(e.g. waving, greeting, etc.). The player can also infer the direction to which another player is looking from their in-game avatars, or whether their in-game avatar is fighting, resting, dancing, jumping, or doing any other embodied action that the game supports. Moreover, for games that do not necessitate fast-paced gameplay, players tend to use in-game text chat (see Bennerstedt, 2007 for an example), and for the fast-paced games, they prefer using built-in voice chats that are in the game or use other voice chat software (e.g. TeamSpeak or Discord) (Moore, et al., 2007).

In other words, while geographically dispersed, for verbal interactions players rely on voice chat software, and for embodied actions players rely on the in-game avatars. Even though the voice chat can be a sufficient substitute for verbal interactions, in-game avatars are still far from accurately and fully representing real-life embodied interactions, or conveying the actual state of their controlling players (Kohonen-Aho & Vatanen, 2020; Manninen & Kujanpää, 2005; Moore et al., 2007). Furthermore, the players do not have access to each others' screens, which creates an information gap (see Balaman, 2015). In other words, a player does not have access to see what another player sees on their screen, this information can be, for instance, that player's abilities (e.g. their ability cooldowns - the time required for an ability to recharge after being used). Limitations such as these will have a significant impact on the interaction between players, even more so if the games are fast-paced. This is because that the players will have to communicate with each other many things (e.g. abilities, enemy locations, game plan, etc.) simultaneously as they are trying to complete the in-game objectives (see Reeves et al., 2009, Mondada, 2011b, Mondada, 2013).

To summarize, gaming research successfully shows the following: players engage in a diverse set of in and out of game interactions while playing games, they make use of game jargons, they value in-game competence, they attempt to collaborate with each other to fulfill in-game tasks, and they do not limit their communication to in-game text chat, but use other software to talk and interact with each other. However, detailed studies remain scarce on how players timely coordinate their activities in gaming (see also Mondada, 2013). As Reeves, Greiffenhagen & Laurier (2017) state, the in-game actions "are sequentially organized and environmentally positioned by players in a purposeful, concerted way" (p. 22). This paves the way for a detailed analysis of these interactions. Therefore, the scope of this study is to examine these "timely coordinated" (Mondada, 2013), sequentially organized and environmentally positioned purposeful in-game actions. To my knowledge, there is currently no research addressing

players' displays of understanding from this perspective. So, I will focus on how players display their understanding of each others' actions, which is key to establishing mutual understanding between themselves. I also aim to contribute to the literature with explorations of voice chat as a medium of talk, as most of the previous research has focused on in-game text chat in geographically dispersed settings and real-life talk in co-present settings.

# 3 Methodology

In this section, I will elaborate on the research method used for the study by pointing out its basic principles, as well as analytical procedures. I will also provide detailed information about participants and data. I will explain what the data contains and how it was transcribed with minute detail. Lastly, I will talk about the transcription conventions before transitioning to the analysis.

## 3.1 Research Method

In this thesis, Conversation Analysis was used as the research method. It is a process-oriented, data-driven, qualitative research methodology that studies naturally occurring interactions with the aim of uncovering patterns through which people interact and behave (Sidnell, 2010b), in contexts ranging from everyday life to institutional (ten Have, 2007). To uncover these patterns of interaction, the analyst's duty is to investigate how participants enact interactional practises. These interactional practises are: (1) turn-taking, (2) sequencing, (3) overall structuring, (4) repair (Wong & Waring, 2010). Turn-taking is the fundamental practise upon which conversation is built, it is concerned with how participants construct and allocate turns (Sacks, Schegloff & Jefferson, 1974).

Sequencing practises are more general practises than turn-taking, they are ways through which participants initiate talk, or respond to talk, while carrying out actions like greetings, story-telling, making requests, invitations (Sidnell, 2010; Wong & Waring, 2010). Overall structuring practises are concerned with how participants open or close sequences of talks (Sacks, 1995). For instance, saying "hi" to initiate a greeting sequence, or "don't I know you from somewhere?" as a part of so-called "pick-up lines" (see Sacks, 1995). Repair practises refer to actions enacted for solving interactional problems or "infelicities" (p. 1) that may stem from speaking, hearing, understanding (Hayashi, Raymond & Sidnell, 2013). Repairs are initiated to point out a problem source, they are finalized when there is a "solution or abandonment of the problem" (Schegloff, 2000. p. 207).

Conversation Analysis employs an emic approach to the analysis, in other words participant's perspective. Therefore, the analyst's aim is to uncover the aforementioned interactional patterns strictly from the talk-in-interaction (i.e. the data itself) rather than consulting external theories

or explanations (Seedhouse, 2005). To gain access to participant's perspective, it is significant to consider the adjacency pairs, context of the talk, and the sequentiality. Observing adjacency pairs allows the analyst to uncover the participant's interpretation of a prior turn within the context it was made available to all of the participants of that interaction. Observing sequentiality grants the analyst participant's perspective on what the participant deemed relevant in any point of the interaction. Hence, all of the claims that are made from the data are strictly based on the participants own interpretations of the interaction and their contributions to it rather than a pre-developed set of coding categories or external theories. Furthermore, the data is transcribed in minute detail in order to accommodate the need to gain access to participant's perspective, as well as to make the data available for scrutiny for readers (Seedhouse, 2004).

At the beginning of a conversation analytic research, the analyst makes an 'unmotivated looking' into the data to catch any phenomenon of interest (ten Have, 1997; ten Have 2007). In other words, the analyst approaches the data without any prior thoughts with regards to what might come out from the data, so that no potentially interesting phenomenon is missed. After something of interest has been found, the analyst starts looking for other similar instances of that particular interactional practice so that a sufficient collection of instances can be made. While making these collections, the analyst can see how that interactional practice is enacted by the participants. Consequently, the analyst is able to describe in detail the enactment process of this interactional practice within that particular context as well as out of that context. More specifically, as Sidnell (2013) notes, the analyst can describe "generic, context-independent properties" of the interactional practice (p. 78).

In my thesis I carried out the analysis in following steps:
1- Unmotivated looking
2- Noting possible interesting phenomenon for research
3- Making basic transcriptions
4- Choosing an interesting phenomenon for research
5- Making a collection
6- Making detailed transcriptions for the collections
7- Making detailed analyses of each example

After I collected the data, I watched the entire data repeatedly to notice an interesting phenomenon to study. My initial personal interest was to try to find a phenomenon relating to the development of second language interactional competence. It is important to note that unmotivated looking refers to a certain kind of "open-minded" approach to the data. In other words, even though the researcher has some broad ideas or interests in mind, they do not only focus on finding a phenomenon related to those. Hence, my initial notes of interesting phenomenon were mostly related to this framework, but were not limited to it. Next, I made basic transcriptions. Then, after formulating more ideas I finally came to the conclusion that I should research how players understand each other during the fast-paced, chaotic in-game moments instead of second language interactional competence matters as the data seemed not suitable for it. So I started focusing on 'displays of understanding' and started making collections where these took place. To find a diverse set of examples to capture this basic phenomenon of interaction, I focused on the most basic and salient features of displays of understanding. These formed the categories that I included in my analysis section. Next, I made detailed transcriptions of the instances that I collected within each category so that I could make minute-detail, sequential, robust analyses. Making detailed transcripts also aims at making the data available to other researchers for more scrutiny and transparency. Lastly, I made detailed analyses of every example within all categories.

### 3.2 Participants

There are eight participants from various European countries (i.e. United Kingdom, France, Germany, Norway, Slovakia). All of the players play the game in English. Except three of them, all are non native speakers of English. Their ages range from 15-27. Every participant except the oldest one is studying in high school at the time of data collection. All of the participants are competitive players with high ratings in Overwatch.

### 3.3 Data

As the nature of this research suggests, the collected data is in digital video format. Recordings of multiple online gaming sessions were acquired from one of the participants (TIM) over two weeks. These reached a total of 14 hours video data, 2 hours out of which were transcribed and analyzed. The participants gathered up online multiple times in two weeks during which they played Overwatch as a team and practiced against other teams, which is called "scrimming" in

the gaming society. In other words, two teams arrange matches during which they try to utilize various tactics against each other to improve their understanding of the game as well as their team synergy. Therefore, the matches are only for practising purposes and have no rewards or punishments at the stake. Scrimming is also done in order to find new members to a team that is recruiting new players to their ranks. This means that in some of the matches one or two participants were swapped out with other participants. All the recordings are from the perspective of TIM, who is the support player of his team.

### 3.4 Overwatch

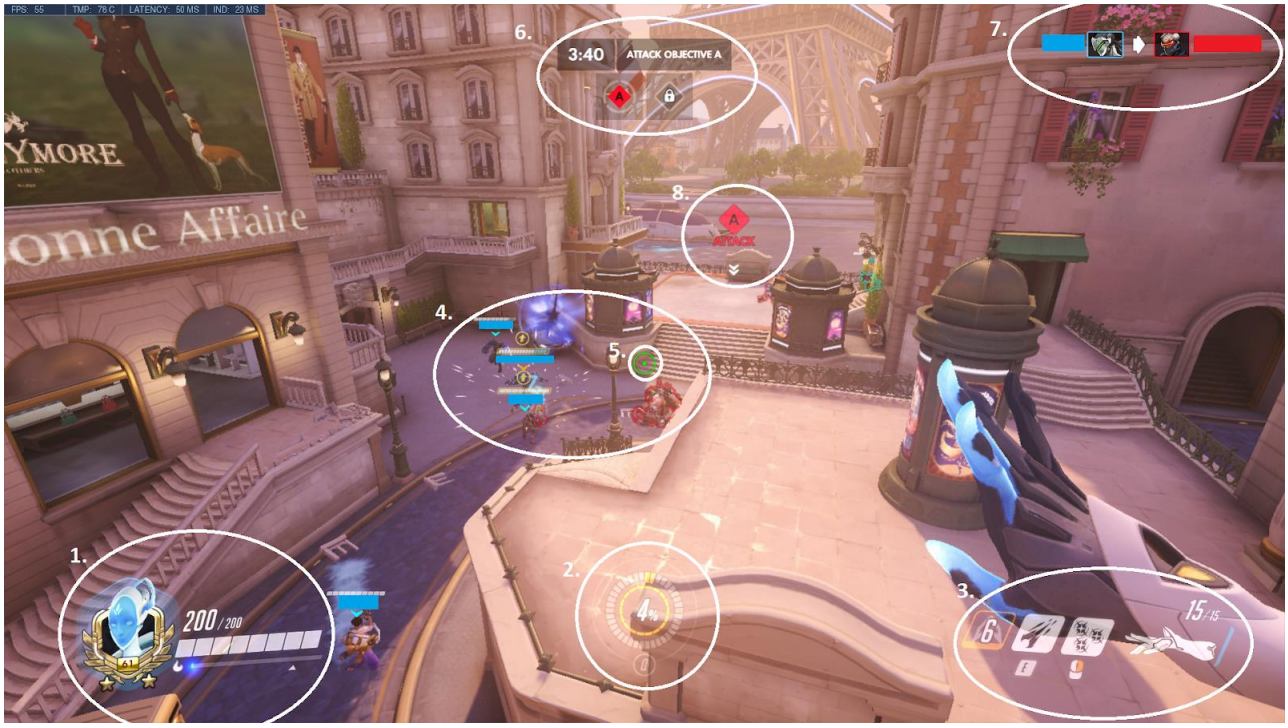
Overwatch is a multiplayer, team-based, first-player shooter (FPS) game. The game is centered on weapon-based combat in a first-person perspective, which means that the players experience the game take action from the eyes of the protagonist/hero. These heroes are in-game avatars that represent the geographically dispersed players. Each hero has a unique name (e.g. Mei, Reaper, Lucio) and its own unique set of abilities (e.g. Ice Wall, Death Blossom, Speed Boost) to be used in the battle. There are three main roles in the game: *tank, damage*, and *support*. Tank role mainly focuses on employing the front lines in battle, to take damage and attention from enemies, create space for their own team to deal damage and kill the enemy heroes to secure the objective, while damage role as its name suggests, focuses on dealing damage and removing enemies from the battle, and lastly support role makes use of various skills to keep their teammates alive, assist them to be more efficient in getting rid of their enemies. Each of the aforementioned roles is crucial for a team's ultimate goal: to secure the victory. Hence, the game's audience is those players who wish to compete with others on various settings through diverse game mechanics and according to a set of collaborative objectives (see below for a detailed discussion of objectives). Action based nature of Overwatch obliges players to communicate with their teammates quickly real-time, make prompt decisions to overcome their enemies and carry out the aforementioned missions various game modes contain. To carry out these missions, players tightly coordinate their actions with other players while they are grouping, fighting, defending, hiding and such. Moreover, there is a built-in voice chat feature in the game that allows players to speak with each other. Hence, the players develop novel ways of communicating through the uses of "shorthand" or "rapid" expressions. For instance, to talk about the low health status, players may say "he is one" meaning that "he is at one health point", which indicates that the enemy health point is dramatically low and that particular enemy can

be removed from the battlefield without hassle. These expressions will be explicated further in the analysis section.

As mentioned earlier, there are various in-game objectives depending on the chosen game mode. These game modes are: *assault, control, escort, hybrid.* Each game mode can be played on its own unique set of in-game locations (i.e. 'maps') (Reeves, et al., 2009). Moreover, game modes are named after the in-game objectives they contain, hence the names for game modes are interchangeably used with the in-game objectives. The game modes are further elaborated below:

1. Assault: In assault, the attacking team must make two capture points. The defending team tries to protect the points and prevent the attackers from making these capture points until the attacking team's time runs out. When a point has been captured by the attacking team, their time gets extended. At the end of a round, teams get swapped and the attackers start defending, while defenders start attacking.
2. Control: In control, both of the teams are attempting to capture a point and hold the control of it until they reach 100% completion. When the point is captured by a team, they start defending it from the opposing team so that they can make it to full completion before their enemies. Control game modes are in best-of-three format, in order to win, a team must make two full completions. Unlike the other game modes, this mode has no time limit.
3. Escort: In escort, the attacking team has to deliver the payload to a final checkpoint. The defending team attempts to prevent this by blocking the path of the payload. Each escort map has two to three checkpoints that attackers must make through, the time gets extended once a checkpoint has been reached by the attackers. The team that delivers the payload further than the other gets the victory.
4. Hybrid: Hybrid modes are a mixture between assault and escort. The attacker team must capture a point, then deliver the payload to a final checkpoint while defenders try to prevent this from happening.

In-game interface involves multiple features that players constantly orient to while playing. These are described with the picture below:




1. Hero's status bar - It contains a small picture of the hero, their health points (200/200 in this case), and their on-fire bar, which indicates how well the player is performing at that particular moment.
2. Hero's ultimate ability status - It shows the percentage of power accumulated so that the hero's unique ultimate ability can be used. Players accumulate this power by dealing damage to the enemy team, or healing their own teammates. Ultimate abilities can only be used when this circle fills up and 100% has been reached. Right below the circle, the key assigned for the ultimate ability is shown.
3. Hero's normal abilities & ammo - Besides ultimate abilities, each hero has its own unique set of normal abilities. Each ability is displayed with a small picture, with assigned keys beneath them. The abilities need to re-charge for a while once they are used. While they are re-charging the picture shows a number that represents the time (in seconds) required for the ability to be available. On the right side of the abilities, the hero's ammo is shown. Heroes can attack their enemies with a limited amount of ammo (15/15 in this case) before they have to reload. Some heroes have unlimited ammo and do not need to reload, if that is the case, it is represented with an infinity symbol (∞).

4. Allied and Enemy heroes - The avatars of allied and enemy heroes are shown. Above an avatar, it's health status is displayed with a small white bar. Below it, there are the nicknames of players (censored with blue for allies, red for enemies). For players to distinguish an ally from an enemy, enemy heroes have a small red lining surrounding them.
5. Player's crosshair - This small green circle indicates where the player aims with their weapon. In FPS games, the crosshair is always fixated right in the center of the screen.
6. Objective status - Here, the objectives are shown as well as the time remaining for the completion of them.
7. Kill feed - Information related to which hero has been killed by whom, using which abilities are shown here. Each player in the game is notified once someone dies, or when their ultimate ability does something (e.g. killing another hero, getting destroyed).
8. Objective location - The in-game location of the objectives are shown with letters representing them. Additionally, borders of the locations are highlighted with colored lines around them (blue for ally controlled, red for enemy controlled). For escort maps, the direction that a payload follows is highlighted.

Overwatch was chosen as the context to analyze interaction for multiple reasons. First of all, the game design creates room for interaction by requiring players to take on and fulfill different roles, which promotes interaction to occur. This interaction, however, is elusive to examine due to the fast-paced nature of the game. As they try to get rid of their enemies, players have to keep in mind many things such as their own abilities, their teammates' abilities, enemy abilities, their locations, and many more. Hence, players find themselves in a chaotic in-game battle environment where they need to carry out their role specific tasks while simultaneously communicating the information they have with their teammates and achieve a synchrony. Therefore, for a team to be successful, each player must timely coordinate their actions with their teammates (Mondada, 2013). Because the research on these timely coordinated interactions is scarce, the present study aims to contribute to fill that gap. Finally, the current popularity of Overwatch made it a relevant source to examine. Even though the game's publisher Blizzard has not revealed the actual number of players in 2019, it was announced to be around 40 million back in 2018. In the next section I will explain how the transcriptions were made, and illustrate the details they contain.

### 3.5 Transcribing Gaming Interactions

Conversation Analysis as a methodology requires the transcription of naturally occurring video data for detailed moment-by-moment analysis. For the data to be presented as complete and accurate as possible, two transcription conventions were used together. Namely, Jefferson (2004) method for transcription of talk, Mondada's (2014) approach for transcription of embodied actions. Both of these transcription conventions were developed for face to face interaction, however, in the current study they are adapted to transcribe the talk of the players and embodied actions of video game avatars which are designed to be somewhat representative of real life interactions (Berger et al., 2016). See Appendix 1 for the full list of markers and symbols used throughout the data.

Transcript example below will be used to illustrate the transcripts presented for analysis in this thesis. Even though their in-game nicknames serve as a means to anonymize them, pseudonyms (i.e. ATI, DEN and so on) will be used in all transcripts in order to protect the privacy of the participants, as well as to ease the reading process. Before the numbered lines, all the heroes that exist during that particular transcript, map and the game mode are noted. Even though each team consists of six heroes, not all of the heroes are listed to preserve clarity and readability. The numbered lines present the talk and pauses as they occur in the data. The first three letters after a line number indicate the speaker. If there is an embodied action taking place within a turn, those will be marked using various symbols (i.e. +, %, &, etc.), and the corresponding embodied action will be demonstrated below the spoken turn. According to the Mondada method (2014), an embodied action begins where a symbol is placed in the corresponding turn (see line 1). The duration of the embodied action is represented by a continuation arrow (--->), until the same symbol is reached, which indicates that the embodied action has ended (see line 3). Figures indicate the screenshots taken at the exact moment in the turn, marked with square (#). All of the lines where embodied actions are indicated begin with the speaker's pseudonym, in order to identify who the action was carried by. Lines where figures are attached begin with a short form of figure (fig), and the figures are numbered.

Consequently, embodied actions indicate the actions of the in-game avatars rather than the players themselves. These actions involve a variety of in-game actions such as gaze (i.e. where the in-game avatars look), ability usage (i.e. special abilities of these avatars), movement (i.e. avatar movements), and so on. To preserve the readability, movements and gaze were only

transcribed wherever it affected the analysis. Moreover, since the current study only had access to TIM's point of view in the games transcribed, the access to other players' gazes and embodied actions is limited.

Excerpt0: Example transcript

Map: King's Row (Hybrid)

Heroes: TIM (Lucio), DEN (Orisa), ATI (Mei) …

Enemy heroes: None

```
1   ATI: oh they're #+%up there up there
    ati                +shoots upwards-->

    den                 %shoots upwards-->

    fig               #fig.1

2         (0.2)
```

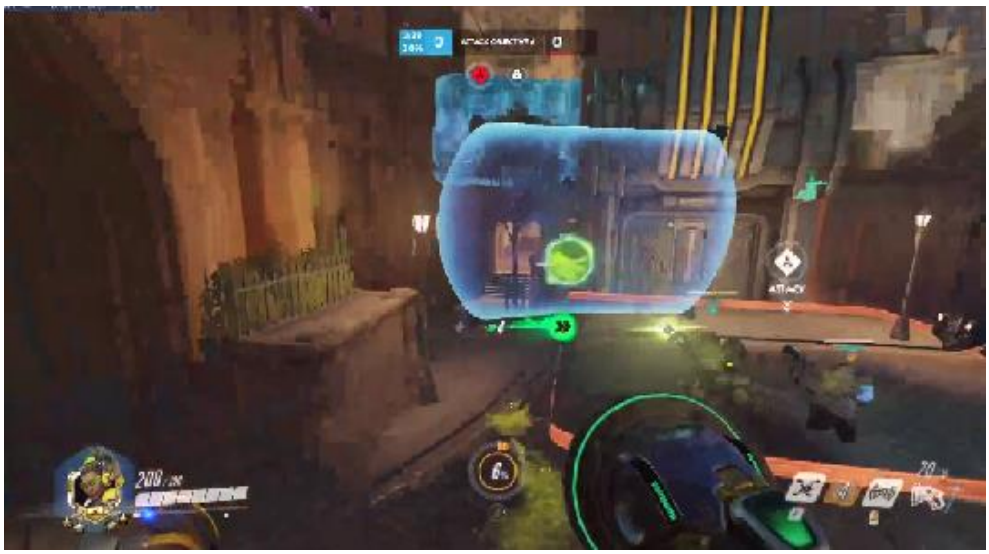


figure 1

```
3   DEN:    +%*they're coming left *they're coming left*
    ati -->+

    den  -->%

    tim       *looks upwards-------*looks left---------*
```

# 4 Data Analysis

This section will illustrate the analysis carried out for this thesis. As emerged from the data, various types of display of understanding were identified. These were first grouped into two main categories: *Displays of Non-Understandings*, and *Successful Displays*. There were two reasons for this choice. First, while the first category encompasses cases of interactional troubles and how these are dealt with, Successful Displays focus on cases where mutual understanding is achieved by players. Second, these main categories have different interactional and game-related consequences as will be emphasized in the analyses. Then, the second category was further seperated into three sub-categories. These are namely: *Verbal Displays of Understanding, Embodied Displays of Understanding, Embodied and Verbal Displays of Understanding*. In sum, display of non-understanding indicates that there has been a demonstration of non-understanding either by a mere claim of understanding rather than a demonstration of it, or by the entire absence of a second pair part. This second pair part refers to potentially relevant actions either with verbal or embodied resources. Second, verbal display indicates that the understanding has been demonstrated during a verbal turn. Third, embodied display indicates that the understanding has been demonstrated using an embodied resource by the player controlled avatars (i.e. gaze, ability use, movement, etc.). Lastly, embodied and verbal display indicates that the understanding has been demonstrated using both verbal and embodied resources simultaneously, or in contingence (i.e. in close proximity time-wise).

The analysis will be presented in the following order: (1) introducing the background for the excerpt, (2) presenting the excerpt, (3) turn-by-turn analysis of the excerpt, (4) the analytical significance of the excerpt. Within the same category, the excerpts presented after the first one will be used to describe the variance in that category. In other words, the excerpts within the same category will differ from each other in certain aspects, these will be elaborated on while introducing the background for the following excerpts after the first one was presented.

## 4.1 Displays of Non-Understanding

This section focuses on excerpts where interactional troubles occur and lead players to not achieve mutual understanding, making them more prone to in-game failures. Both non-understanding and misunderstanding moments are analyzed. Particular attention is paid to the events leading up to the potential cause of this phenomenon (overlaps, no-hearing etc.), as well

as to the events following the display of non-understanding moments (i.e. sanctioning, repair). Players usually ignore these moments arguably because there are no negative consequences for the team, or when there are other immediate focuses in-game. However, sometimes, when these hurdles occur, the in-game consequences for the team are grim. For instance, a teammate might die because of it, or an attack/defense might fail. Furthermore, these moments may trigger upcoming interactional sequences during which remedies are occasionally attempted to solve the non-understanding case.

In the first excerpt, I show a case where a display of non-understanding gets ignored. The team is escorting the payload to the second checkpoint, during which they encounter enemy pharah shooting from above.

Excerpt1: Save ults
Map: King's Row (Hybrid)
Heroes: ALP (Sombra), MIT (Lucio), KLD (Orisa)
Enemy heroes: Pharah

| | | |
|---|---|---|
| 1 | DEN: | we might lose so lets save ults. |
| 2 | ELO: | yeah don't barrier [here.] |
| 3 | DEN: | [yeah ] △dont. |
| | **tim** | **△uses ultimate** |
| 4 | DEN: | oh why- |
| 5 | ATI: | >we can win ⌂we can win.< |
| | **col** | **⌂kills enemy doomfist** |
| 6 | COL: | >i killed doom< keep going. |
| 7 | TIM: | yeah i didn't hear. |
| 8 | DEN: | [no problem.] |
| 9 | COL: | [it's ok we ] won. |

In line 1, DEN suggests to save 'ults', which refers to saving the remaining ultimate abilities of his teammates for an upcoming teamfight. For this, he gives the reason that they might lose the ongoing teamfight. ELO in line 2 follows this decision by saying "yeah don't barrier", which is TIM's ultimate ability. As he is finalizing his utterance, he overlaps with DEN saying "yeah",

as DEN continues with his turn, TIM uses his ultimate. DEN in line 4 questions this decision by saying “oh why-”, ATI cuts him off by quickly saying that they can win the teamfight deemed as potentially lost by DEN earlier. During ATI’s quick utterance, the team is informed of COL removing enemy doomfist from the teamfight. COL also utters this fact in line 6, and reinforces his team’s morale by telling them to “keep going”. In line 7, TIM gives an account of why he used his ultimate “yeah i didn’t hear”. This is the first time in this interaction where an explanation is made regarding not having listened to an earlier callout. Given TIM’s explanation, we can claim that TIM displayed a non-understanding with his ultimate usage. In other words, TIM has not demonstrated the relevant action of keeping his ultimate for the next teamfight, and he gives an account of this. However, the team seems to not treat this non-understanding as a negative case. DEN and COL give a ‘by-pass’ to this non-understanding by saying “no problem”, and “it’s ok we won”. Here, players arguably ignore the display of non-understanding by TIM and they do this because the in-game consequences for the team were not grim. In other words, despite the earlier callout by DEN, the team has won the teamfight.

This excerpt is significant in demonstrating how players ignore some of the non-understanding moments. The main reason for this seems to be the players’ assessment of in-game consequences. DEN and COL assess the situation as non-problematic, and because it is deemed as not negatively consequential for the team, the non-understanding is ignored.

The following excerpt will demonstrate an example in which a display of non-understanding by one player (ELO) takes place, but this time it is not ignored. The fact that non-understanding has been displayed is going to be sanctioned by another team member.

Here, the team is trying to push the payload to the second checkpoint. They are almost there but there is one final stretch, at which their enemies are attempting to stop them.
Excerpt2: Said that

Map: Blizzard World (Hybrid)

Heroes: TIM (Lucio), ELO (Moira), SIN (Sigma), DEN (Orisa)

Enemy heroes: Reaper, Orisa

```
1   TIM: #they are stacking high ground.
    fig  #fig.2
```

2 *#(0.2) *#(1.0)*#

**tim *turns behind*turns in front then quickly turns behind***

**fig #fig.3 #fig.4 #fig.5**

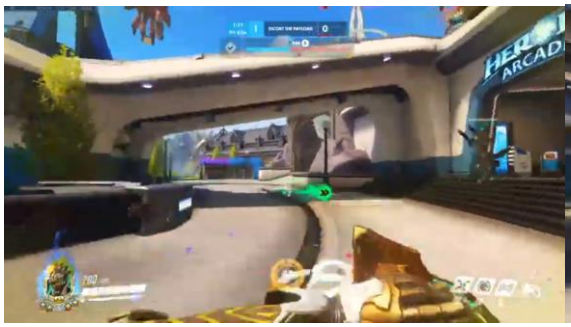


figure 2 (looking in front)

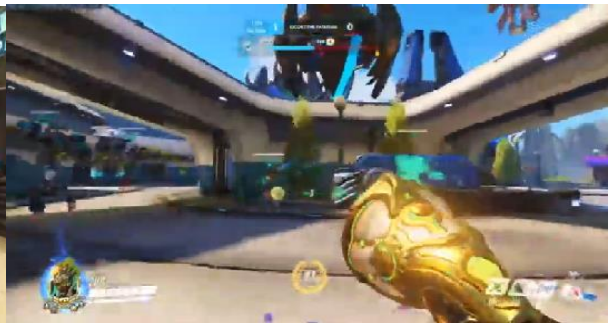


figure 3 (looking behind)

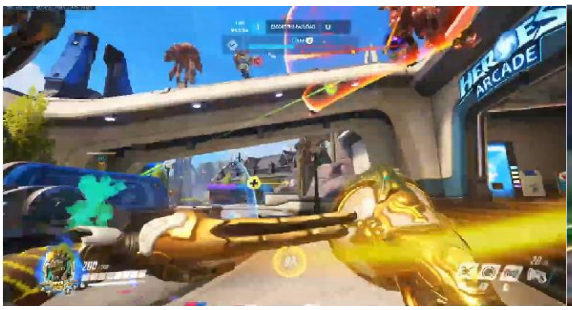


figure 4 (looking in front)

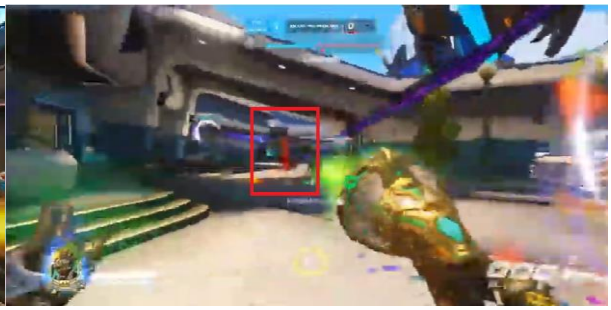


figure 5 (looking behind)

3 TIM: >REAPER BEHIND [+↑TP- R- &BEHIND TP BEHIND↑<]

4 DEN: [+>yeah &i'm bongoing i'm

5 bongoing.<]

**tim +shoots at reaper--->**

**den &uses ultimate**

6 (1.5)

7 DEN: [>orisa one*+ orisa one.<]

8 SIN: [orisa orisa.]

9 (0.5)

**tim *turns to orisa**

**tim -->+shoots at orisa--->**

10 ●(0.2)

**elo ●dies to reaper**

11 TIM: +△ [beating].

12 ELO: +△ [oh reap]er behind.

**tim -->+△uses ultimate**

13 (2.0)

14 DEN: yeah TIM said that.

In line 1, TIM signals to his team that the enemies are gathering up on the high ground. On his screen, we see that he is referring to the balcony above his team (#fig.2). At the end of his turn, he looks behind to assure there is nobody (#fig.3). After this, he looks in front of himself to check his enemies (#fig.4), then looks behind himself again. As he looks back, he spots the enemy Reaper teleporting behind his team. Reaper's teleportation looks similar to a red shadow as can be seen (#fig.5). In line 3, TIM produces a first pair part with his warning action; he shouts and warns his team that one enemy player is right behind them, effectively making relevant a second action (i.e. acting carefully according to the warning). Players use warnings like these so that their entire team knows the danger in their backlines, and reacts to it appropriately either by taking the enemy out, or by running away from that enemy. The physical location of players' in-game avatars make it even more relevant for them to do something about a threat if they are closer to that threat (e.g. Reaper). Other players except ELO are further away from Reaper during this exact moment, so they do not immediately show a reaction or give a response. However, there is an expectation that ELO does something so he would not die. This will also be seen from DEN's utterance in the ensuing interaction.. TIM later listens to DEN's callout at line 7 and starts shooting at Orisa. From this moment onwards, TIM continues shooting at Orisa until line 10. He then starts casting his ultimate ability then notifies his team (beating). Overlapping with this, ELO dies in line 10 and reacts to the fact that Reaper was behind him all along (oh reaper behind). His turn is marked by a change of state token "oh" (Heritage, 1984), which means that this knowledge that Reaper was behind him has just been received. Hence, ELO has not shown a relevant second action that would display his understanding (i.e. running away from Reaper), for he died to the Reaper behind him despite TIM's warning in the first pair part back in line 3. After two seconds of silence, which can be interpreted as an impending trouble (Jefferson, 1989) in line 13 DEN sanctions ELO's failure to act in response to the previous warning by TIM as he points out to the earlier turn where TIM stated that Reaper was behind them (yeah TIM said that). In other words, this turn can be

interpreted as DEN's possible expectation of ELO listening to the warning by TIM and running away from where he was to avoid dying to the enemy Reaper. Why ELO has not displayed understanding to TIM's callout can be explained by looking at TIM's turn. More than half of his turn in line 3 is overlapping with DEN's turn in line 4, thus it might have led to a mishearing or no hearing by ELO, which can be confirmed by the change of state token "oh" he used in his turn. Another explanation why this might have happened could be that ELO focusing on Orisa just like TIM was doing, so he could have missed the callout. Yet, there is no conclusive evidence for these in the transcript as there was no access to his screen. Furthermore, there was no claim nor a demonstration of understanding in this transcript. There was however, only a case of display of non-understanding that led to in-game as well as interactional troubles.

In my data, while usually ignored, such displays of non-understanding are occasionally followed by sequences in which the player who failed to display understanding is held accountable for their failure. In other words, the player who does not display their understanding with a relevant second turn or action gets called out in the talk. These sanctions are carried out usually at the end of the sequence, and are occasionally remedied with repairs. Next, I will analyze an excerpt where repair follows the sanctioning. Also, within the same example I will show a claim of understanding taking place that does not turn into a sufficient display.

In this excerpt, the team is attempting to go to point A in Lunar Horizon in order to deliver an attack and capture the point. However, DEN and TIM got stranded on the way due to the enemy team blocking the path while they were attempting to cross earlier. Here, they are re-attempting to cross the path and unite with their team before attacking. An important aspect of this phenomenon is "fake go", which means that the players fake their running action, then return back to their starting point. They do so to make the enemy team use some of their abilities and waste them, so they can run back to the spot they have decided without giving any casualties to the enemy abilities. This can be interpreted similarly to the use of decoys in a real battlefield so an enemy fires their bullets at the decoy - essentially wasting their bullets.

Excerpt3: Fake go
Map: Horizon Lunar Colony (Assault)
Heroes: TIM (Lucio), ELO (Moira), ATI (Mei), LUC (Reaper), DEN (Orisa)
Enemy heroes: Mei,

```
1   DEN: er we'll fake and go here TIM okay?=
```

2 TIM: =yeah.
3 [three (.) two (.) one.
4 DEN: [three (.) two (.) one.
5 ■%✦(3.0) % ✦(0.5)✦
**tim ■uses speedboost**
**tim %runs to gate--%**
**den ✦runs to gate-->✦runs back and stops**
**den ✦runs to gate-->**
6 DEN: okay they just. ♦ (0.3) ✦not-
**mei ♦uses ice wall**
**den -->✦**
7 (0.5)
((DEN's path is blocked))
8 DEN: #●oh no.-
**den ●dies**
**fig #fig.6**
9 (0.2)

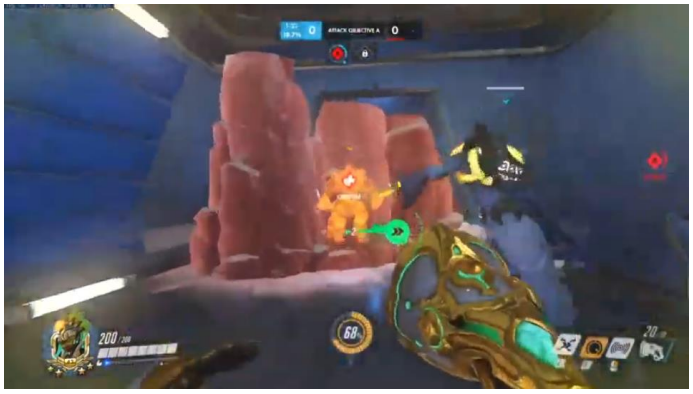


figure 6

10 [okay.
11 ELO: [no.
12 LUC: that's unlucky.
13 (1.0)
14 DEN: i ↑said fake go TIM.
15 and you just $↑left$
16 TIM: ↑hi:h (1.0) $sorry$ (0.5) u:m,
17 (3.0)

```
18   SIN:          just go speed him back.
19                 (0.5)
20   TIM:          yeah (0.8) guess i have to.
```

DEN opens up by briefing the tactic to TIM that they will “fake and go” in line 1, which constitutes a first pair part that requires a relevant second action (i.e. TIM to provide an agreement/disagreement). During the latching turn in line 2, TIM affirms this by saying (yeah), which can be interpreted as a claim of understanding as the running action has not started yet. In the overlapping lines 3 and 4, both DEN and TIM do a countdown meaning that they will go together when the countdown ends. At the end of this countdown, TIM uses his speed boost ability and they start running towards the gate they want to reach. Until after the 3 seconds silence, TIM continues running and makes it through the gate, he then stops there. However, DEN has not done the same as TIM, he had started running briefly and went back to where both TIM and DEN started running from, which was his intention of fake go. After TIM stops, there is a 0.5 secs silence, after which DEN says (okay they just), which could be arguably the first indication of a trouble within this action sequence. He tries to cross the path and reach the gate alive, then enemy Mei uses her ice wall to block his path (marked with ♦, also see #fig.6). Immediately after this, DEN cuts off most likely due to surprise that mei blocked his path (△%not-). In what follows, DEN gets killed by the enemy team and says “oh no” before cutting off again. After a very brief silence (0.2 secs) he says “okay” in line 10, which overlaps with ELO’s “no”. In line 12, LUC reacts to this by saying “that’s unlucky”. Similarly to the previous excerpt, in the third position (lines 14-15) DEN addresses TIM’s failure of providing the relevant second action (i ↑said fake go TIM and you just $↑left$) with an accusation that TIM just left him alone before he could also cross the gate himself. He is doing so by referring to his earlier turn in line 1, during which he briefed the plan to TIM. DEN marks the verb “said” with a rising pitch to emphasize (Walker, 2012) what he had told TIM earlier, and he repeats parts of that specific turn “fake go” so that he is pointing out the trouble source. He then finishes off with another rising intonation ($↑left$). In line 16, TIM responds to this first by laughing (↑hi:h), then saying ($sorry$) and produces a continuation marker “um” (Clark & Tree, 2002). It is also seen that in his turn there are intra-turn silences in between “hih” and ”sorry” as well as between “sorry” and ”um” (1 sec and 0.5 secs, respectively), which is considered an indication for hesitating (Jefferson, 1989). After this, a potentially troublesome long silence (3 secs) (Jefferson, 1989) occurs before SIN

proposes a candidate solution to the situation (`just go speed him back`). Meaning that TIM should run back to the starting location and use his speed boost ability for him so that DEN can return back to his teammates without spending too much time. Half a second of silence follows this and TIM affirms this in line 18 by saying "yeah, guess I have to", and by accepting the candidate solution, also closes the sequence.

Previously in excerpt 2, we have seen that ELO produced an "oh" prefaced turn that indicated his late noticing of Reaper and DEN sanctioned this display of non-understanding similarly to this case. However, there was no response by the sanctioned player (ELO) afterwards, or no repair attempt took place likely because ELO had died and could not temporarily perform any in-game actions. This example differs from that in the way that the sanctioned player (TIM) responded by apologizing, confirming that the understanding did not take place regardless of the affirmation he had shown in line 2 (`yeah`). In other words, TIM only claimed understanding rather than demonstrating it. Moreover, when the sequence was about to end in silence, another player stepped in to fill in with a candidate repair to handle the situation. TIM carried out the repair to close the sequence. The excerpts also differed in the way that while in excerpt 2 the sanctioned player himself (ELO) suffered the in-game consequences of his action (i.e. dying) whereas in excerpt 3 the sanctioned player (TIM) put another player's (DEN) avatar at risk, and eventually it died, which also could arguably explain why more interactional work was needed to resolve the issue.

From the next sub-chapter and onwards, I am going to present cases where understanding is actually displayed, rather than merely claimed. I will provide evidence to this from sequential analyses of players' in-game actions and their talk.

### 4.2 Successful Displays

In this section, I will present excerpts where players achieve mutual understanding and overcome game-related troubles successfully. Three sub-categories will be presented: Verbal Displays of Understanding, Embodied Displays of Understanding, Embodied and Verbal Displays of Understanding.

### 4.2.1 Verbal Displays of Understanding

This sub-category focuses on the display of understanding carried out using verbal resources (i.e. in spoken turns). Display of understanding in verbal turns is a common occurrence in the data, similar to everyday conversations. How players do this in-game is demonstrated with the next excerpt. At this moment, the team is attempting to capture the second objective in Temple of Anubis and end the game by doing so.

Excerpt4: Pull
Map: Temple of Anubis (Assault)
Heroes: TIM (Lucio), ATI (Doomfist), COL (Reaper), DEN (Orisa)
Enemy heroes: Junkrat, Orisa

```
 1  COL: =>do you have pull [do you have pull?<
 2  TIM:                    [amping heal.
 3       (1.5)
 4  DEN: i've in two:
 5       (0.5)
 6       >i have it now i have it now.<
 7       (0.3)
 8  COL: >okay okay just tell me just tell me<- [( )    ]
 9  TIM:                                       [speeding].=
10  COL: =when you pull them.
```

As the fight ensues, in line 1 COL quickly asks "do you have pull" in repetition. This is directed at DEN as COL already knows that DEN's hero (Orisa) has the pull ability, yet COL does not know whether the ability is ready or not. Hence, COL's turn serves as an availability check for the pull ability. TIM's turn in line 2 overlaps with COL's as he is signalling to his team that he is using healing for his team (`amping heal`). After a silence in line 3 (1.5 secs), in line 4 it is seen that DEN responds to COL (`i've in two:`). In this case, DEN is trying to convey that "I have my ability in two seconds", this is another case of a shortened expression used by players commonly. In line 6, he signals that his ability is ready with his quick repetition (`>i have it now i have it now<`). COL responds to this in lines 8 and 10 by asking DEN to let him know when he is going to use his pull ability, in between these lines TIM overlaps with COL by interjecting his usage of speed boost ability.

This excerpt brings evidence to the fact that even in chaotic in-game battle situations, players are able to display understanding of each other's turns during their own turns. This took place between DEN and COL during lines 1-10. COL first asked DEN if he has pull (line 1), which was attended to by DEN in lines 4 and 6. Then in lines 8 and 10, COL expanded on it by telling him "okay okay just tell me just tell me when you pull them". Simultaneously, with this turn, COL has treated DEN's verbal display of understanding as accepted. In the end, they created a successful joint play that led to the capture of the point, with multiple verbal displays of understanding taking place in the process. Another important note is that COL's turns explicitly demand a verbal response. Expressions such as "do you have pull" are information requests that could be answered with a yes/no, or as DEN did by giving the timing of the ability to recharge. Furthermore, it creates an increased response relevance in the particular way COL formed it. First, it is designed as a question that demands an answer. Second, it is repeated twice quickly to "mobilize" a response from the recipient DEN (Stivers & Rossano, 2010). Hence, it can be seen that DEN provided the relevant verbal display of understanding. Conversely, if COL was asking DEN to use an ability that would demand an embodied response (i.e. the usage of that specific ability), which could be considered as an immediate next step after the check COL did for the pull ability.

Whereas the previous one was during combat, the next excerpt varies from the previous one in that it is taken from an out of combat planning situation. Here, the team is organizing their attack on the first point in King's Row. They begin by scouting the enemy team from their starting location, during which they discuss from which angle they can carry out their attack.

Excerpt5: then we go
Map: King's Row (Hybrid)
Heroes: TIM (Lucio), DEN (Orisa), COL (Genji)
Enemy heroes: Doomfist

```
1   TIM: they're holding top right er we should
2        we could rotate behind them or.
3        (0.5)
4   TIM: [through hotel].
5   DEN: [let's check  ]which second dps they have first yeah.
6        (0.5)
```

```
 7   TIM: doom.=
 8   DEN: =oh it's doom yeah yeah we go hotel okay?
 9   TIM: yeah.
10   COL: cool cool cool.
```

TIM tells his team where their enemies are and where his team should attack from in line 1. He pauses for half a second after he says his turn-final "or", which may indicate that he is looking for another candidate location to add to his turn. After the silence, he adds the location in line 4 (`[through hotel]`), which partially overlaps with DEN's turn. DEN is making a contribution to TIM's plan, which also displays his understanding verbally at the same time (`[let's check ]which second dps they have first yeah`). After a half second silence, TIM displays another verbal understanding by pointing (`doom=`) (i.e. doomfist) as the second dps (i.e. damage) hero of the enemy team. DEN, in line 8, marks his turn with the news receipt "oh" and confirms that TIM's hotel plan should be carried out (`yeah yeah we go hotel okay?`). In lines 9 and 10, TIM first gives an affirmative response "yeah", followed by an assessment by COL "cool cool cool", which closes the sequence. After this, the team then proceeds to attack from the hotel (lines omitted).

In this excerpt, the players were making an attack plan out of combat so no immediate ability use or movement was required of them. Hence, it is seen that during this planning moment they treat the use of verbal resources sufficient and do not make use of embodied resources. The reason for including this excerpt was to demonstrate the variety of sequences during which players verbally display understanding. Regularly, the sequences are taken from in-combat action sequences however the players also talk and discuss out of combat as seen from this excerpt.

### 4.2.2 Embodied Displays of Understanding

During this sub-category, the analytical focus is on the embodied resources used to display understanding of each other's turns (i.e. gaze, ability use, movement). Close attention will be paid to the sequential unfolding of the display of understanding moments, similar to earlier sections.

In this excerpt, the team members are standing on top of the balcony in the objective they are defending, waiting for the enemy team to gather up and attack. Simultaneously, they are trying to spot where the enemies will approach from.

Excerpt6: Corridor

Map: Oasis (Control)

Heroes: TIM (Lucio), COL (McCree), DEN (Orisa)

```
1   COL: #they might go corridor↑
    tim *>>looks at the main gate-->
    fig  #fig.7
2         (1.0)
3         or main.
4         (2.0)
5         yea.
6         (1.0)
7         corridor [corridor].
8   DEN:            [corridor].
9           *(1.0)*+#
    tim -->*.....*looks at the corridor---
    tim             +shoots-->>
    fig              #fig.8
```

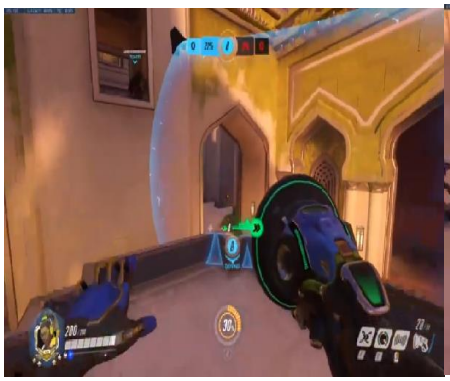


figure 7

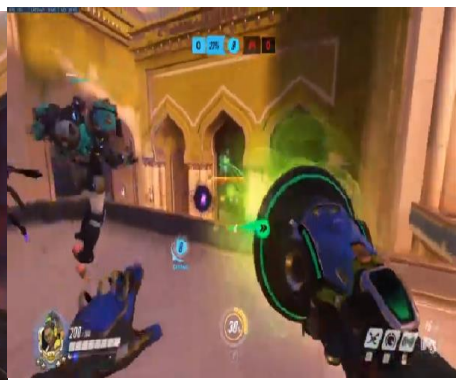


figure 8

The excerpt starts with COL giving a candidate location in line 1, which is followed by a short silence (1 sec) and another candidate location in line 3 (`or main`). These receive no response during the long silence (2 secs) and the team members continue looking around, walking back and forth and so on. However, after the line 5 (`yea`) and a short silence (1 sec), COL spots the

enemies in line 7 (corridor [corridor]). DEN also spots the enemies and we see this due to his turn in line 8 ([corridor]), which overlaps with COL's repetition of corridor. Both of these turns can be considered as a first pair part of an adjacency pair, which creates space for a relevant responsive action. More specifically, because the turn is built as an announcement of an in-game location, it makes actions related to that in-game location (i.e. corridor) relevant. These actions can be for instance saying something about the location, shifting the gaze towards there, shooting there, using abilities targeting that location and so on. It is seen here that TIM carries out this relevant responsive action. He has been looking at the main gate since before the excerpt started (#fig.7), which continued until the contributions by COL and DEN at lines 7 and 8 (see * marker for TIM's gaze). After TIM receives the information that enemies are in the corridor, he walks backwards and starts staring and shooting at the corridor (#fig.8). This shift of gaze and shooting also was done by TIM in a timely way. He only shifted his gaze and started shooting when his teammates were absolutely sure that the enemies were approaching from the corridor. The evidence to this can be found by looking at the multimodal lines between line 1-9. Fundamentally, this gaze shift and the subsequent shooting action by TIM pictures how players timely use their embodied actions in Overwatch to display understanding of other players' turns, as well as their actions. Lastly, the team was able to successfully defend the point and drive the attacking team away.

In the previous excerpt, TIM displayed his understanding using embodied resources. The following excerpt will demonstrate variance in the way that all of the team members will collectively and simultaneously display their understanding using embodied resources.

During this excerpt, the team is preparing to attack point A in Temple of Anubis. They are preparing to cross the defense enemy team has established at the gates of the point. Their aim is to make TIM use his speed boost ability, after the use of which they all plan to run together to bypass the enemy defense.

Excerpt7: Three two one go
Map: Temple of Anubis (Assault)
Heroes: TIM (Lucio), ELO (Moira), COL (Reaper), SIN (Sigma), DEN (Orisa)
Enemy Heroes: Not relevant
Evr - everyone in the team

```
 1   SIN: are you ready?
 2        (1.0)
 3   TIM: i got stunned.
 4        (1.0)
 5   SIN: [ready?
 6   COL: [okay.
 7   ELO: yeah yeah.=
 8   TIM: =okay three two: one.
 9        %■speeding.
     evr  %runs--->
     tim   ■uses speed boost
10        (2.0)%#
     evr   --->%
     fig        #fig.9
11   DEN: nice.
```

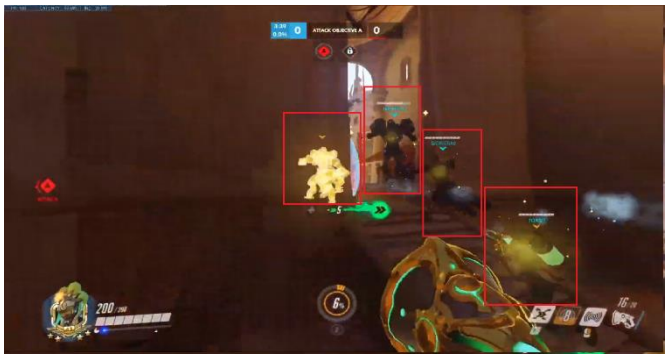


figure 9

Notice how lines 1-8 serve as a pre-sequence (i.e. lines of talk that prepare a main sequence of talk/actions to take place) to prepare as a team for the coordinated play (i.e. the joint run), which we will observe as a synchronized, embodied display from everyone. SIN opens up in line 1 by asking his team if they are ready. There is a 1 sec silence following this, after which TIM in line 3 says "i got stunned" to indicate that he is not ready yet. Another 1 sec silence follows this and in line 5 SIN repeats his question partially (`ready?`), this overlaps with COL's "okay", which is a minimal contribution to the context. However, it is unclear whether he is just confirming the situation of TIM, or trying to state that everyone in the team is okay to go. In line 7, ELO gives the go-ahead response (`yeah yeah`) to signal that the team is ready. This latches with TIM's turn in line 8, during which he counts down (`okay three two: one`),

right before using his speed boost ability and signalling to his team in line 9 “speeding”. The team starts running slightly before TIM notifies them that he is speeding, nonetheless, they all are under his speed boost ability’s benefits during the time period. During the silence in line 10 (2 secs), everyone in the team looks at the same direction and runs past the enemy defense as can be seen (#fig.9), with the exception of ELO, who is not in this particular frame but very close to TIM around the time where the screenshot was taken. DEN then proceeds to close the sequence with a third turn (Schegloff, 2007a) with a positive assessment “nice” in line 11.

This excerpt focuses on a recurrent theme in this data where the embodied displays of understanding demonstrated by every teammate simultaneously. To do this, players frequently start out with these pre-sequence (lines 1-8) where they plan their upcoming actions, and they execute this plan afterwards (line 9 onwards), then finalize by confirming the successful display of understandings (line 10). Players make clear their readiness to make a joint play, they do so by sequentially placing their talk and actions in relevance to what took place previously. In other words, before they try to make a timely coordinated play, they ask each other of their in-game status “are you ready?”, or “are you stunned?” so that they can check if all of the teammates are ready for it. Their in-game avatars looking in the same direction and moving towards it in harmony indicates that mutual understanding has taken place and the relevant actions have been provided by everyone, thus, display of understanding has been done. With the help of these resources, the players engage in successful joint plays that lead to in-game success (e.g. capturing the point).

In excerpt 6, we have seen that players make location callouts and look at those directions to display their embodied understanding, whereas in excerpt 7 it is shown that players are also carrying out actions such as grouping and moving together while simultaneously looking at the same direction. Furthermore, in excerpt 6 there was one player who performed the responsive embodied action while in excerpt 7 there was a pre-sequence which served as a preparation for the group movement.

### 4.2.3 Embodied and Verbal Displays of Understanding

Displays of understanding within this sub-category indicate the displays in which verbal resources are used simultaneously or in contingence with the embodied resources. The analytical focus will be on the sequential unfolding of these moments as well as the speakers’

selection of the upcoming interactant, who will provide a relevant embodied action, and display their understanding verbally while doing so.

Next excerpt demonstrates how players display understanding towards each other's turns and actions using both embodied and verbal resources in contingence. The team is fighting to capture the first point in King's Row. They have just confronted their enemies and the battle ensues.

Excerpt8: Booped
Map: King's Row (Hybrid)
Heroes: TIM (Lucio), ATI (Mei), COL (Reaper), DEN (Orisa)
Enemy heroes: Lucio, Orisa

```
 1   TIM: △beatin¤g.
     tim  △uses ultimate
     ori          ¤uses ultimate
 2        they bongo'ed.
 3   DEN: okay >♠[i'm pulling] i'm pulling<
 4   ATI:       ♠[ready]
     den        ♠uses pull ability
 5   DEN: i'm i'm ♣i pulled em TOAste.
     ati          ♣uses ultimate
 6        (1.0)
 7   COL: bongo's one ¤⌂bongo's one.
     luc              ¤uses ultimate
     col               ⌂destroys orisa's ultimate
 8        (0.3)
 9   COL: ▼no more bongo.=
     col  ▼uses wraith
10   TIM: =they [beated].#
11   DEN:       [nice  ].#
     fig                  #fig.10
```

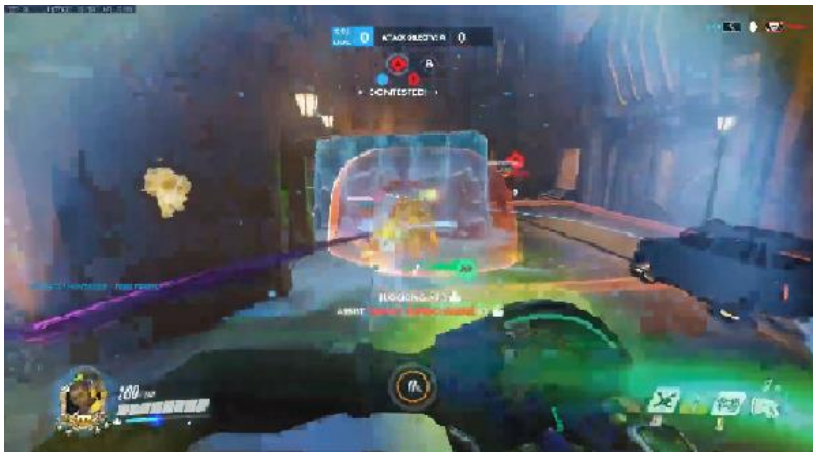


figure 10

12 DEN: reaper behind >REAPER BEHIND*+

13 #REAPER BEHIND↑< (0.2)+▲ ( )

**tim *turns behind**

**tim +shoots-->**

**tim -->+▲boops**

**fig #fig.11**

14 (0.5)

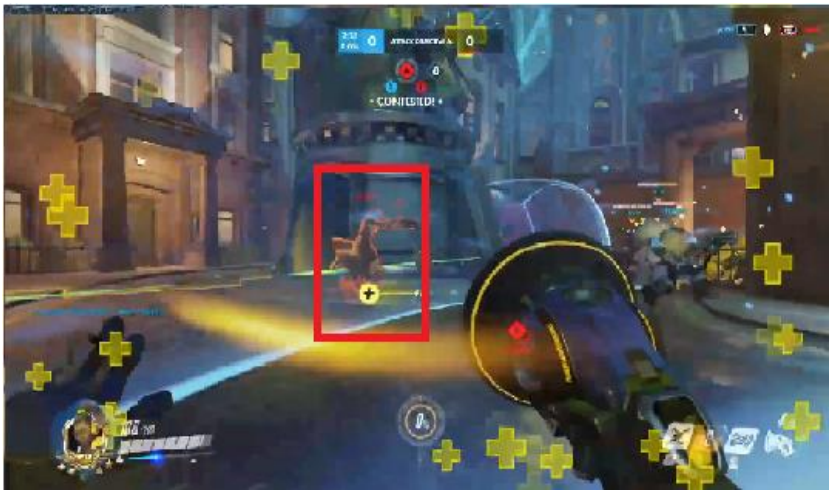


figure 11

15 TIM: booped him away.

In lines 1-10 the team can be seen to destroy enemy Orisa's ultimate ability "bongo" and announce that the enemy Lucio has used his ultimate ability. DEN then closes this action sequence by saying "nice" in line 11. He then signals to his team that Reaper is behind them, which is repeated three times. Two of these are louder and quicker than the first, (reaper behind >REAPER BEHIND*▲ REAPER BEHIND<). He manages to draw TIM's attention

to the incoming danger. We see this when the second repetition of "Reaper behind" has been finished, TIM has turned behind to shoot at Reaper (see `#fig.10` and `#fig.11`), which shows embodied display by TIM. Until the end of third repetition as well as a very brief silence (0.2 secs), TIM continues shooting at Reaper. After this, he uses his boop ability to push Reaper away (see line 12), then reports this in line 14 after the 0.5 secs of silence (`booped him away`), with which he is displaying his understanding verbally. By doing this, TIM ensures that his team is safer as there will be no Reaper threat behind his team's ranks. Consequently, TIM sets up his team for in-game success (e.g. capturing the first point of King's Row).

In this excerpt, it is evident that players are competent in conveying information via verbal and prosodic resources (lines 11-12) and quickly orienting to those using embodied resources (i.e. TIM's avatar turning back, shooting, and booping see *, ▲ and + markers in line 12). These embodied actions show that the information DEN conveyed has been received by TIM, hence understanding has been displayed with embodied resources initially. However, there is still space to provide relevant upcoming talk (see the silence in line 13). Line 14 further provides evidence to this. TIM verbally reports his action after having received the information and showed his embodied display. To summarize, players are able to make relevant upcoming actions for each other, conceive these and fill in with the actions required by them in sequentially relevant positions. While doing so, they effectively display their understanding using both embodied and verbal resources (lines 11-14).

In the previous example, we have seen that TIM reacted to the calls by DEN (i.e. doing something to take care of the enemy Reaper) and performed a relevant action (i.e. attacking, and pushing the enemy Reaper away from the backlines of his team). However, DEN did not specifically select TIM to do the action there. He only signalled to his whole team, and TIM picked up the responsibility to fill in with a relevant action himself. This is due to two reasons: first, as the support player of his team he has to keep his teammates safe. Second, his hero Lucio has abilities tailored for this purpose (e.g. speed boost/boop). This is similar to a next speaker self-selection (Sacks, 2004). in a way that TIM self-selected himself to carry on the task.

In the next excerpt, TIM will be selected by DEN to perform a certain action, of which TIM displays verbal and embodied understanding in contingence. During this excerpt, the enemy team has won the previous teamfight and TIM's whole team is waiting to respawn (i.e. avatars'

coming back to life) at the next checkpoint. TIM is stranded at the previous checkpoint and is trying to run back to the next checkpoint where he aims to unite with his team. However, TIM is not the only person that is stranded. DEN shares his fate, hence he wants to go back to his team as soon as possible, too.

Excerpt9: Don’t leave me

Map: Blizzard World (Hybrid)

Heroes: TIM (Lucio), DEN (Orisa)

```
1   TIM: #%okay (0.5) they used [reaper sigma].
    tim   %runs towards the exit--->
    fig  #fig.12
```

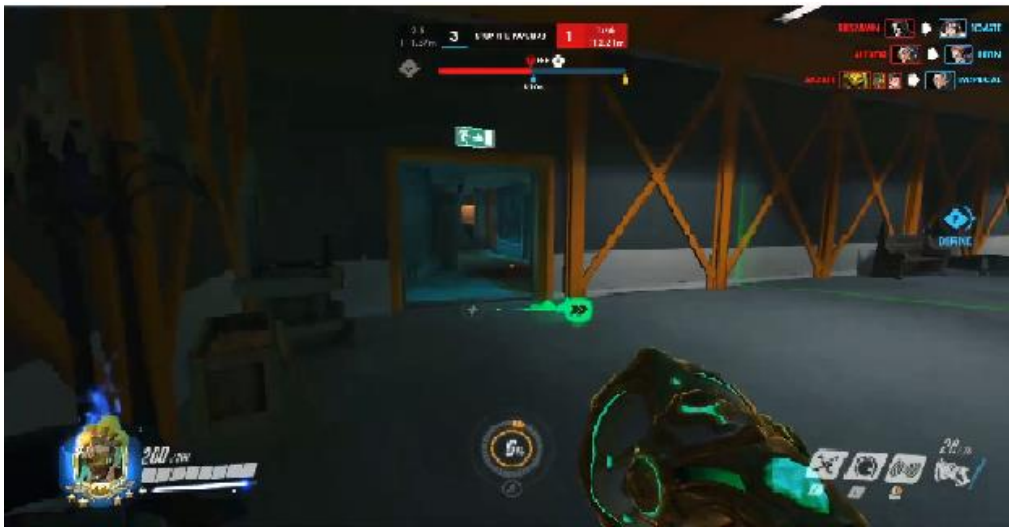


figure 12

```
2   DEN: [don't leave me mitsu].
3        don't leave me mitsu.%
    tim                  --->%stops
    tim                      %runs back towards DEN
4        (0.5)
5   TIM: #got you. ■speeding.
    tim            ■uses speed boost
    fig  #fig.13
```

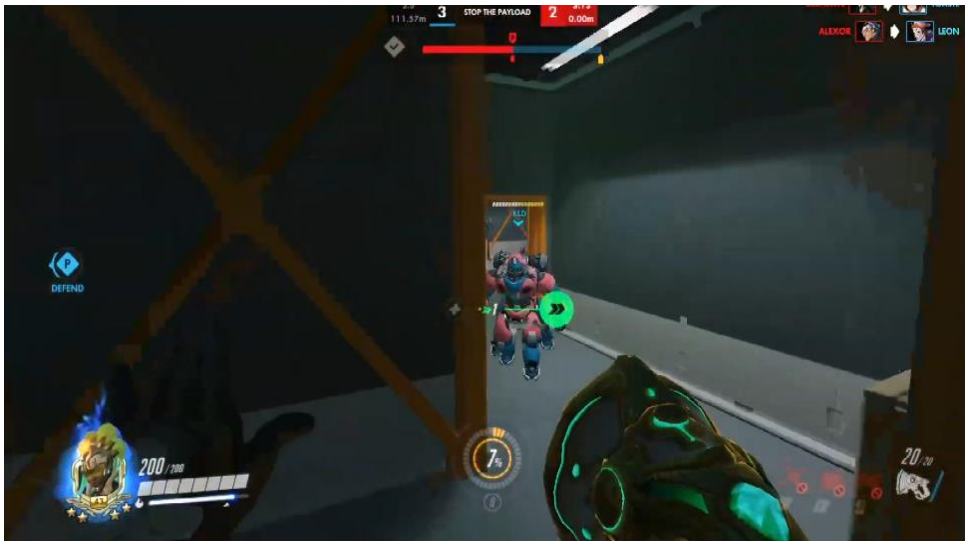


figure 13

In line 1, as regularly seen in my data, TIM starts talking about the enemy team's ultimate abilities right after the previous fight has ended. As he is producing his turn, he is already running towards the exit as can be seen (#fig.12). Before TIM finishes his turn however, DEN overlaps with him ([dont leave me mitsu]). TIM does not react until after the second repetition from DEN in line 3. At the end of DEN's turn in line 3, TIM stops and runs back towards DEN. Thus, TIM displays through embodied means his understanding of DEN's turn. In line 5, he is already united with DEN (#fig.13), and he uses his speed boost while signalling to DEN that he is going to be safe with TIM (#got you +speeding). By doing this, he also places a verbal display of understanding because he shows that he has understood DEN's turn as a request for help of running back to his teammates. Even though within this excerpt there is no immediate in-game success, all of the teammates uniting as fast as possible helps them to engage in successful joint plays. Ensuring teammates grouping up serves as one way to promote the joint plays.

This excerpt explicates the variance of how interactants fill in with the relevant actions while displaying understanding both verbally and embodily. Here, DEN selected the next interactant (TIM) to provide a relevant upcoming action. TIM mainly gets selected due to the hero he plays (i.e. Lucio), who has the ability to speed boost his teammates so that they run faster. When compared with the earlier example where TIM turned around and booped Reaper away from his own team, it is seen here that players use a diverse set of ways to distribute tasks, or pick them up without explicit assignments.

In the next excerpt, the display of understanding is rather reversed to verbal and embodied rather than embodied and verbal. It is taken from the moment before excerpt 4. The team is trying to capture the second point in Temple of Anubis. They have just lost AYA, yet they still want to continue with the attack.

Excerpt10: oh yeah i'm on him

Map: Temple of Anubis (Assault)

Heroes: TIM (Lucio), ATI (Doomfist), COL (Reaper), DEN (Orisa)

Enemy heroes: Junkrat, Orisa

```
 1   DEN: i'm not pulling not yet.=
 2   COL: =i think we can still [win this.
 3   TIM:                       [still winnable.
 4        (0.5)
 5   DEN: sigma l[ow.
 6   COL:        [orisa low.
 7   TIM: orisa no fortify.=
 8   DEN: =pulling junk behind us junk behind us.
 9        (0.2)
10   ATI: oh yeah i'm on him (0.5) he's ⌂#one.
     ati                                ⌂kills Junkrat
     fig                                 #fig.14
```

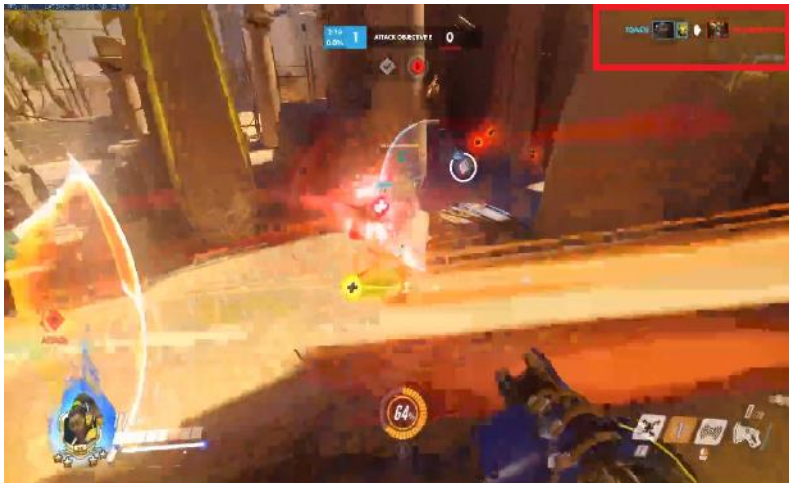


figure 14

```
11   DEN:         nice.=
```

The excerpt starts with DEN signalling to his team that he won't use his pull ability just yet (`i'm not pulling not yet`). This is immediately followed by COL's latching turn in

line 2. He orients to the fact that AYA died by saying “still”, while showing confidence that they will succeed in their attack (`we can still win this`). Overlapping with COL, TIM does the same in line 3 by telling his team that victory is indeed possible (`still winnable`). A negligible short silence (0.5 secs) follows this. Then, in lines 5 and 6, DEN and COL note that enemy Sigma and Orisa are low, respectively. TIM in line 7 provides more information regarding the enemy Orisa’s situation by saying she has no fortify ability (`orisa no fortify`). DEN’s turn in line 8 starts immediately after TIM’s turn, he is signalling to his team that he is pulling the enemy hero Junkrat, which is shortened to “junk” to save time. By producing this first pair part he is not just telling that he is pulling Junkrat here, he is using this turn to tell his teammates to help him kill Junkrat, hence this turn can be interpreted as a request for assistance. This is evident from the fact that he is notifying his teammates of Junkrat’s location (`junk behind us`) and his action of pulling him. Another, perhaps more important evidence can be found in ATI’s treatment of DEN’s turn as a request for assistance to kill Junkrat. After a very short silence (0.2) secs, in line 10, it is seen that ATI is providing the required assistance to eliminate Junkrat. He receives the new knowledge that junkrat is being pulled and we see this as he marks his turn with a reactive token “oh yeah” (Young & Lee, 2004). He then signals to his team, and especially to DEN, that he is on Junkrat. A short silence follows this (0.5 secs) and he conveys the information that Junkrat is close to death (`he's one`). However, even before he finishes his turn, he kills Junkrat. By doing this, he helps his team capture the point successfully. As with every other killing blow in Overwatch, this information is conveyed to every player in the game located on the top right corner (`#fig.14`). DEN reacts to this in line 11 by saying “nice”, by doing this he treats ATI’s display of understanding as sufficient and closes the previous sequence where he requested assistance with Junkrat.

This excerpt demonstrates that during combat situations, displays of understanding can also occur in a reversed way (i.e. verbal can be first and be followed by an embodied display). It could be argued that this occurred because in line 8 DEN announced that he required assistance with Junkrat and that ATI provided verbal display of understanding to this before actually assisting. Another possible reason for this could be the amount of time the player has had at his disposal. For instance, when excerpts 8 and 10 are compared, it could be possible that ATI had more time to react and put a verbal response first to the Junkrat danger in comparison to TIM

reacting to Reaper being behind his team, essentially having to use his boop ability first and foremost, then announcing that he did it.

### 4.3 Summary of main findings

The data analysis carried out in the current study aimed to explicate the different types of displays of understanding in Overwatch, and the interactional resources used during these displays. Two main (i.e. displays of non-understanding, successful displays) categories, three sub-categories were identified (i.e. verbal display of understanding, embodied display of understanding, embodied and verbal display of understanding). Making four categories in total, with each indicating variances in different aspects of social interaction (e.g. repair, speaker/interactant selection, sequence organization), and they were taken from different types of in-game interaction (e.g. in-combat, out of combat). In addition, the moments in which understanding has been displayed successfully lead to more coordinated play within the team and affected their success positively.

In the display of non-understanding category, within excerpt 1 “Save ults” we have seen that players can tend to ignore some displays of non-understanding depending on the situation. The reason for this was that there was no negative in-game consequence for the team. During excerpt 2 “Said that” we have seen that DEN sanctioned ELO’s display of non-understanding as ELO’s avatar got killed by the enemy Reaper. However, the sanctioning did not lead to further interactional efforts such as repair to resolve the case, the reason of which could be that the sanctioned player had already died and could no longer revert the situation. Furthermore, in excerpt 3 “fake go” DEN sanctioned TIM’s not carrying out the action made relevant by himself earlier. This, unlike the previous example, consequently triggered a repair attempt by another team member (SIN). After this, TIM finished the sequence by delivering the relevant interactional effort (i.e. repair) to solve the problem. From this, it can be seen that during the first display of non-understanding moment, it was not deemed as a trouble source, hence no other interactional efforts were made. During the second example, it was indeed treated as a trouble source by DEN. Nonetheless, repair did not take place. Similarly, in the third example the display of non-understanding was treated as a trouble source. Yet, repair did take place and the sequence was thus closed with a problem solution. Moreover, in this category it was also seen that when players fail to display understanding of each other’s turns and actions, they can

be more prone to in-game failures such as the death of their character or failure to capture objectives.

In the verbal display category, it was seen that players deem it sufficient to put verbal responses during in and out of combat situations when the verbal response is enough by itself to solve the situation at hand. Excerpt 4 “pull” brought evidence to this as COL and DEN made a joint play in combat. Furthermore, the turns designed by COL made it increasingly relevant for DEN to give a verbal response, during which he simultaneously displayed his understanding of COL’s turns. Another finding was that TIM’s contributions did not draw any turns addressing them because they were designed as mere announcements, and were placed in overlaps. In Excerpt 5 “then we go” the players were shown to interact out of combat, during which they deemed it sufficient to use only verbal resources to display understanding. This was arguably due to that when out of combat, it is not required of the players to make quick plays with their avatars as they are not confronted with enemies. Hence, in out of combat situations embodied actions were rather limited and verbal contributions were more prevalent.

Embodied displays were significant in that the players controlled their in-game avatars timely coordinated in order to demonstrate that they are in sync with their teammates. TIM’s shift of gaze and shooting in excerpt 6 “corridor” displayed his embodied understanding of the earlier turns, where his teammates by their announcements of the said location made actions related to it relevant. It was also observed that this embodied display was timely placed in order to make use of in-game time and resources more effectively. While in excerpt 6 it was only one player that carried out the responsive action, in excerpt 7 “three two one go” it was shown that the whole team timely and collaboratively carried out actions such as moving. This was evidenced by their countdown and other pre-sequential availability checks (e.g. “are you ready?”), all of which served to prepare the team for the timely coordinated play. Another observation was that players’ avatars looked in the same direction while they carried out this play.

Embodied and verbal displays varied from embodied displays category in the way that verbal resources were used in contingence with the embodied actions in order to display understanding of another player’s turn. Within the category, another variance was identified in the selection of next interactant. During excerpt 8 “booped” we have seen that TIM self-selected himself to provide a relevant action (e.g. booping Reaper away from his team), whereas during excerpt 9 “don’t leave me” DEN selected the next interactant (TIM) to provide the relevant action. Both

of these speaker selections occurred because of TIM's hero (i.e. Lucio), and its abilities to boop and speed boost. In the last excerpt of this category "oh yeah i'm on him" it was observed that the sequential unfolding of displays of understanding was reversed (i.e. to verbal + embodied). This was arguably due to the amount of time the players had at their disposal. While in excerpt 8, TIM had to immediately react to the Reaper threat behind his team, in excerpt 10 ATI arguably had more time to first provide a verbal display of understanding to DEN's assistance with Junkrat, then deliver the killing blow (i.e. embodied display).

Apart from display of understanding moments, an overarching finding across all of the sub-categories where understanding was successfully displayed was that players found more in-game success despite the chaotic, fast-paced in-game battle situations. It could be argued that this was due to their timely coordinated plays and interactional synchrony. More evidence supporting this was found as players also carried out actions such as countdowns before timely coordinated plays, and they engaged in pre-sequential interactions during which they talked about their upcoming plays before executing them. Players were also shown to make use of in-game jargon as well as shorthand expressions likely due to the shortage of time and having quite many tasks at hand to carry out. Another conversation analytic finding was identified in third position (Schegloff, 2007a). Even though it was not the main scope of this research, it is worth mentioning because it may contribute to the potential upcoming research in gaming contexts. Minimal contributions in the third position such as "yeah", "okay", and "nice" were commonly found throughout my data, which are also quite typical in ordinary conversational settings. However, they could have different implications in game settings.

# 5 Conclusion

The aim of this study was to examine the timely coordinated interaction of players in online multiplayer gaming settings as the conversation analytic research on this area is still relatively scarce. Hence, the present study aimed at filling that gap as well. Amongst other reasons such as popularity of the game, the most important reason Overwatch was chosen as the game to analyze was that it naturally facilitates timely coordinated interaction with its fast-paced, team-based qualities. The focus was on displays of understanding within the interaction, for examining how humans establish mutual understanding is one of the core aims of Conversation Analysis. Research questions of this study were:

*1- What types of display of understanding do participants use in Overwatch?*
*2- What are the interactional resources used by the participants during display of understanding moments?*
*3- How does displaying understanding affect players' in-game success?*

The analysis of the data yielded two main categories (i.e. Displays of Non-Understandings, and Successful Displays), in total making four types of display of understanding as elaborated on in the analysis section: Non-Understanding, Verbal, Embodied, Embodied and Verbal. These categories were based on the most salient features of players' displaying understanding of each other's turns and actions.

Within the first category, examinations were made of moments where players failed to arrive at a mutual understanding. This was made clear in the interaction through their misunderstandings, or non-understandings. The players were shown to hold each other accountable for these moments, and try to overcome these hurdles in interaction by using interactional resources such as repair. Another observation was that players were more prone to in-game failures such as the inability to capture an objective when they could not achieve mutual understanding.

In the second category, moments during which players displayed their understanding through spoken turns were examined. It was shown that despite the chaotic in-game battle situations, players were able to arrive at mutual understanding and overcome their enemies. The verbal turns had prominent features such as quick repetition, or shorthand expressions. It is perhaps

due to the fast-paced nature of Overwatch that players interact in this way, which also opens up the possibility for future research as conversation analytic studies of these interactions remain largely unexplored in the literature.

In the third category, the examinations were on moments where players displayed understanding via embodied channels. These embodied channels were mainly based on the player controlled avatars' gaze and bodily orientation. It was observed that the players were able to act in sync and made this available to their teammates through the use of their avatars for instance by looking at the same place as their teammates or moving together with them. Moreover, these actions in sync led to more in-game success.

Within the last category, observations were made of instances where both embodied and verbal channels were used by the players. The players were able to display understanding with their avatars' gaze or body movement and support it with spoken turns. This was due to one possible reason. The in-game roles of the players and the tasks these roles entail make it more relevant for them to provide announcements after they have made plays (e.g. support character Lucio booping away Reaper). A reversed sequential unfolding was also observed in this category. In other words, a player first provided verbal response and supported it with embodied resources. This was possibly because of the shortage of time and amount of in-game actions creating a challenge. Another observation was made on the speaker selection, there were instances of both speaker self-selection and other-selection. This, too, related to the roles of in-game characters players controlled. Support characters were shown to be more relevant in terms of providing verbal responses and embodied displays because of their protective and other-enhancing abilities. As similar to the verbal and embodied categories, the players were able to achieve better in-game results through their establishment of mutual understanding.

The study context was different from many of the existing research on gaming contexts (see e.g., Bennerstedt, 2007; 2008; Bennerstedt & Ivarsson, 2010; Mondada, 2013). These research focused on other types of games (e.g. *FIFA, World of Warcraft*) in which actions required from players are arguably slower in pace than a first-person shooter such as Overwatch. Furthermore, players experience the game take action from a more 'distanced perspective', which allows them to predict upcoming actions with more ease and make their own actions in more calculated, recognizable ways (Mondada, 2013). Mondada (2013) characterizes this perspective as a "zenithal view ..., which captures mobile trajectories in a distanced way, maximizing the

possibility of projecting and anticipating them" (p. 337). Despite this difference in context, many of these studies established a common finding that players show their competence of the game through a set of situated in-game practises (Bennerstedt, 2008; Mondada, 2013; Piirainen-Marsh & Tainio, 2009; 2014). Players seem to assume that their teammates are competent in the game, and they constantly engage in competent actions such as displays of attention (Bennerstedt & Ivarsson, 2010), coordinated teamplay, and visual alignment (Bennerstedt, 2008). The present study contributed to the literature by evidencing in minute-detail players' competent, collaborative in-game actions, as well as how these are manifested in talk and action.

As a rather rare study in terms of its context and methodology combined, Rusk & Ståhl (2020) focused on a first-person shooter game *Counter-Strike: Global Offensive* (CS:GO) using CA. Even though they put emphasis on the sequences of Kill (K-) and Death (D-) events, many of their findings are in line with the current study. First and most central to the present study, they found that players engage in callouts as part of game strategy (Rusk & Ståhl, 2020). They state that "Callouts are employed to co-construct a shared knowledge and understanding of the game environment through sharing game relevant information such as the locations and intentions of teammates and opponents." (Rusk & Ståhl, 2020, pp. 22-23). The present study demonstrated in detail how players make, and orient to these callouts during talk and action sequences in gameplay. Second, they illustrated that players in CS:GO engage in efficient communication (i.e. short and clear interaction) (Rusk & Ståhl, 2020), which resonates with the current findings about "shorthand/rapid expressions". Third, players treat it okay to sometimes not respond to other players when they are concerned with other immediate in-game actions (Rusk & Ståhl, 2020). Even though this was not highly focused on during the analysis, a similar case was found in excerpt 2.

This thesis was written in the program of Learning & Educational Technology, and in my study it was my initial interest to try to understand language learning better. Hence, I find it my responsibility to elaborate on the language learning aspect of the study. Principally, Conversation Analysis aims to describe the machinery of human interaction (Sidnell, 2010b). This is done primarily through spoken and embodied interaction and with a specific interest on what is made available to all the participants rather than the researcher. Theories are not consulted and the analyses are based strictly on data and principles of conversation (Seedhouse, 2004). As this is the case, coining an example as "learning" is rather far fetched. Learning itself

being understood as an internal, mental, elusive phenomenon, we may need a case to be repeated a lot of times, or after seeing a learner's first encounter with a word or phrase (i.e. a potentially problematic issue for the learner, so we can claim that the learning is possible) we need to see that word or phrase to be used again in the future to overcome the situation (Kasper, 2004; Wagner, 2004). For this purpose, the research should be designed in a way to acquire thick longitudinal data. However, even though this can be done, it would still not be certain whether the future utterances of the same word or phrase is incidental learning or not. In sum, if learning is understood as an internal mental state, CA might not be able to provide evidence for other than incidental learning (Wagner, 2004).

On the other hand, if learning is considered from a social viewpoint, CA can provide fruitful evidence in terms of describing the participation of a learner in the environment (Pekarek Doehler, 2010). Through the tracking of longitudinal data, CA can show an increase or decrease in participation, the quality of the participation and so on (Pekarek Doehler, 2010). Building on this perspective, situated learning can indeed be evidenced and the explanation of inner mental states can be left for another methodology (Wagner, 2004).

The context for learning can be everywhere, so limiting the analyses to the classroom would be an outdated fashion (Sundqvist, 2009). Instead, CA's powerful tools that describe the structure of language use should be used in everyday contexts in order to understand language learning better (Wagner, 2004). Following this trend, I was able to describe the language use of a group of Overwatch players. However, all of the players were already very fluent in their English and my data was not longitudinal. Because of this, the study does not focus on the initial aim of shedding light on language learning. To do this, future research should aim to collect much more data than I could. Future research should also look to recruit participants that are mixed in their English language proficiency, so that more potential opportunities for language learning can be created. Another suggestion for data collection is the interface of the participants. In the present study, the data is from only one participant so accessing all of the participants' in-game interfaces was not possible. If the future studies could collect all of the participants' in-game interfaces, more detailed and accurate analyses can be made, or different avenues for research can open up.

In the end, video games form a significant part of everyday life for many people, and with them human behavior continues to change and evolve. Thus, opportunities for research are definitely

rich in the exploration of these new contexts. Not only with CA but also via the use of other methodologies must we aim to continue developing our understanding of humans in change.

## References


Aarsand, P. A., & Aronsson, K. (2009). Response cries and other gaming moves—Building intersubjectivity in gaming. *Journal of Pragmatics*, *41*(8), 1557-1575.

Balaman, U. (2015). A conversation analytic investigation into the impact of task design on the emergence of information gaps. In *Proceedings of 17th International CALL Conference, Task Design and CALL, Tarragona, Spain* (pp. 6-8).

Baldauf-Quilliatre, H., & de Carvajal, I. C. (2015). Is the avatar considered as a participant by the players? A conversational analysis of multi-player videogames interactions. *PsychNology Journal*, *13*(2).

Belhiah, H. (2012). Gesture as a resource for intersubjectivity in second-language learning situations. *Classroom Discourse, 4* DOI:10.1080/19463014.2012.671273

Bluemink, J., & Järvelä, S. (2011). Elements of collaborative discussion and shared problem solving in a voice-enhanced multiplayer game. *Journal of Interactive Learning Research*, *22*(1), 23-50.

Bennerstedt, U. (2007). Avatars & interaction in gaming: Dysfunctional Interaction or a Practice of Players. In *Game in'Action Conference*.

Bennerstedt, U. (2008). Sheeping, sapping and avatars-in-action: An in-screen perspective on online gameplay. In *Proceedings of the [player] conference* (pp. 28-52).

Bennerstedt, U., & Ivarsson, J. (2010). Knowing the way. managing epistemic topologies in virtual game worlds. *Computer Supported Cooperative Work, 19*, 201-230. doi:10.1007/s10606-010-9109-8

Bennerstedt, U., & Linderoth, J. (2009). The spellbound ones: illuminating everyday collaborative gaming practices in a MMORPG. In *CSCL (1)* (pp. 404-413).

Berger, M., Jucker, A., & Locher, M. (2016). Interaction and space in the virtual world of second life. *Journal of Pragmatics, 101*, 83-100. doi:10.1016/j.pragma.2016.05.009

Caldwell, K. E. H., Osterweil, S., Urbano, C., Tan, P., & Eberhardt, R. (2017). “I Just Don’t Know Where to Begin”: Designing to Facilitate the Educational Use of Commercial, Off-the-Shelf Video Games. In *Serious games and edutainment applications* (pp. 625-648). Springer, Cham.

Chen, H. H. J., & Huang, W. Y. C. (2010, April). Examining the potentials of computer games for English learning. In *2010 Third IEEE International Conference on Digital Game and Intelligent Toy Enhanced Learning* (pp. 134-138). IEEE.

Clark, H. H., & Tree, J. E. F. (2002). Using uh and um in spontaneous speaking. *Cognition*, *84*(1), 73-111.

Drew, P. (2013). Conversation analysis and social action. *Journal of Foreign Languages*, *37*(3), 2-19.

Ducheneaut, N., & Moore, R. J. (2005). More than just ‘XP’: learning social skills in massively multiplayer online games. *Interactive technology and smart education*.

Gee, J. P. (2006). Are video games good for learning?. *Nordic Journal of Digital Literacy*, *1*(03), 172-183.

Goodwin, C. (2000). Action and embodiment within situated human interaction. *Journal of pragmatics*, *32*(10), 1489-1522.

Hayashi, M., Raymond, G., & Sidnell, J. (2013). Conversational repair and human understanding: An introduction. In M. Hayashi, G. Raymond & J. Sidnell (Eds.), *Conversational repair and human understanding* (pp. 1-40). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511757464.001 Retrieved from https://www.cambridge.org/core/books/conversational-repair-and-human-understanding/conversational-repair-and-human-understanding-an-introduction/E7ED58DA2A3FA0884CF0F18BD9BDE68D

Heritage, J. (1984). A change-of-state token and aspects of its sequential placement. *Structures of social action: Studies in conversation analysis*, 299-345.

Heritage, J. (2009). Conversation analysis as social theory. *The new Blackwell companion to social theory*, 300-320.

Hindmarsh, J., Reynolds, P., & Dunne, S. (2011). Exhibiting understanding: The body in apprenticeship. *Journal of Pragmatics*, *43*(2), 489-503.

Hung, A. C. Y. (2007). Video games in context: An ethnographic study of situated meaning-making practices of Asian immigrant adolescents in New York City. In *DiGRA Conference*.

Jefferson, G. (1989). A standard Maximum Silence in Conversation. In D. Roger, & P. Bull (Eds.), Conversation: An Interdisciplinary Perspective. Philadelphia: Multilingual Matters LTD.

Jefferson, G. (2004). Glossary of transcript symbols with an introduction. *Pragmatics and Beyond New Series*, *125*, 13-34.

Järvenoja, H., & Järvelä, S. (2009). Emotion control in collaborative learning situations: Do students regulate emotions evoked by social challenges. *British Journal of Educational Psychology*, *79*(3), 463-481.

Kasper, G. (2004). Participant orientations in German conversation-for-learning. *The Modern Language Journal*, *88*(4), 551-567.

Keating, E., & Sunakawa, C. (2010). Participation cues: Coordinating activity and collaboration in complex online gaming worlds. *Language in Society, 39*(3), 331-356. doi:10.1017/S0047404510000217

Kohonen-Aho, L., & Vatanen, A. (2020). Participation and co-presence in the virtual world of Second Life: Transitioning from a gathering to an encounter [Discussion Paper]. Journal for Media Linguistics. Retrieved from: http://dp.jfml.org/wp-

content/uploads/DP002_2020_Kohonen-Ato_Vatanen_Participation-and-co-presence-in-the-virtual-world-of-Second-Life.pdf

Korkiakangas, T., & Rae, J. (2014). The interactional use of eye-gaze in children with autism spectrum disorders. *Interaction Studies, 15*(2), 233-259. doi:10.1075/is.15.2.12kor

Lindwall, O. (2014). The body in medical work and medical training: An introduction. *Discourse Studies, 16*(2), 125-129. doi:10.1177/1461445613514671

Lindwall, O., & Ekström, A. (2012). Instruction-in-interaction: The teaching and learning of a manual skill. *Human Studies*, *35*(1), 27-49.

Manninen, T., & Kujanpää, T. (2005). The hunt for collaborative war gaming-Case: Battlefield 1942. *Game Studies*, *5*(1), 538-540.

Mondada, L. (2011). Understanding as an embodied, situated and sequential achievement in interaction. *Journal of Pragmatics*, *43*(2), 542-552.

Mondada, L. (2011b). The situated organisation of directives in French: Imperatives and action coordination in video games. *Nottingham French Studies*, *50*(2), 19-50.

Mondada, L. (2013). Coordinating mobile action in real time: The timely organisation of directives in video games. In *Interaction and Mobility*. Berlin, Boston: De Gruyter. doi: https://doi.org/10.1515/9783110291278.300

Mondada, L. (2014). Conventions for multimodal transcription. Retrieved from https://344cc026-c96f-49aa-b4bc-071b454d3061.filesusr.com/ugd/ba0dbb_986ddd4993a04a57acf20ea06e2b9a34.pdf (last accessed 22nd April 2020)

Mondada, L. (2014b). Pointing, talk, and the bodies: Reference and joint attention as embodied interactional achievements. *From Gesture in Conversation to Visible Action as Utterance: Essays in Honor of Adam Kendon,* , 95-124.

Mondada, L. (2019). Contemporary issues in conversation analysis: Embodiment and materiality, multimodality and multisensoriality in social interaction. *Journal of Pragmatics*, *145*, 47-62.

Moore, R., Ducheneaut, N., & Nickell, E. (2007). Doing virtually nothing: Awareness and accountability in massively multiplayer online worlds. *Computer Supported Cooperative Work, 16*, 265-305. doi:10.1007/s10606-006-9021-4

Mori, J., & Hayashi, M. (2006). The achievement of intersubjectivity through embodied completions: A study of interactions between first and second language speakers. *Applied Linguistics, 27*(2), 195-219. doi:10.1093/applin/aml014

Olsson, E. (2012). " Everything I read on the Internet is in English". On the impact of extramural English on Swedish 16-year-old pupils' writing proficiency.

Olsson, E. (2016). *On the impact of extramural English and CLIL on productive vocabulary*.

Paans, C., Onan, E., Molenaar, I., Verhoeven, L., & Segers, E. (2019). How social challenges affect children's regulation and assignment quality in hypermedia: a process mining study. *Metacognition and Learning*, *14*(2), 189-213.

Pekarek Doehler, S. (2010). Conceptual changes and methodological challenges: On language and learning from a conversation analytic perspective on SLA. In *Conceptualising 'learning'in applied linguistics* (pp. 105-126). Palgrave Macmillan, London.

Piirainen-Marsh, A., & Tainio, L. (2009). Other-repetition as a resource for participation in the activity of playing a video game. *The Modern Language Journal*, *93*(2), 153-169.

Piirainen-Marsh, A., & Tainio, L. (2014). Asymmetries of knowledge and epistemic change in social gaming interaction. *The Modern Language Journal, 98*(4), 1022-1038. doi:10.1111/modl.12153

Rauniomaa, M., & Keisanen, T. (2012). Two multimodal formats for responding to requests. *Journal of Pragmatics*, *44*(6-7), 829-842. doi:10.1016/j.pragma.2012.03.003

Reeves, S., Brown, B., & Laurier, E. (2009). Experts at play: Understanding skilled expertise. *Games and Culture, 4*(3), 205-227. doi:10.1177/1555412009339730

Reeves, S., Greiffenhagen, C., & Laurier, E. (2017). Video gaming as practical accomplishment: Ethnomethodology, conversation analysis, and play. *Topics in Cognitive Science*, *9*(2), 308-342.

Rusk, F., & Ståhl, M. (2020). A CA perspective on kills and deaths in Counter-Strike: Global Offensive video game play. *Social Interaction. Video-Based Studies of Human Sociality*, *3*(2).

Sacks, H., Schegloff, E.A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language 50*(4), 696-735. doi:10.1353/lan.1974.0010.

Sacks, H., 1992. Lectures on Conversation [1964–72], 2 Vols.. Basil Blackwell, Oxford.

Sacks, H. (1995). *Lectures on conversation*. Oxford: Blackwell.

Schegloff, E.A. (1968). Sequencing in Conversational Openings1. *American Anthropologist, 70*: 1075-1095. doi:10.1525/aa.1968.70.6.02a00030

Sacks, H. (2004). An initial characterization of the organization of speaker turn-taking in conversation. *Pragmatics and beyond new series*, *125*, 35-42.

Schegloff, E. A. (1987). Some sources of misunderstanding in talk-in-interaction. *Linguistics*, *25*(1), 201-218. doi:10.1515/ling.1987.25.1.201

Schegloff, E. A. (1992). Repair after next turn: The last structurally provided defense of intersubjectivity in conversation. *American Journal of Sociology, 97*(5), 1295-1345. Retrieved from http://www.jstor.org/stable/2781417

Schegloff, E. A. (2000). When 'others' initiate repair. *Applied Linguistics, 21*(2), 205-243. doi:10.1093/applin/21.2.205

Schegloff, E. (2007). The adjacency pair as the unit for sequence construction. In *Sequence Organization in Interaction: A Primer in Conversation Analysis* (pp. 13-21). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511791208.003

Schegloff, E. (2007a). *Sequence Organization in Interaction: A Primer in Conversation Analysis*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511791208

Schegloff, E. A., & SACKS, H. (1973). Opening up closings. *Semiotica, 8*(4), 289-327. doi:https://doi.org/10.1515/semi.1973.8.4.289

Schegloff, E. A., Jefferson, G., & Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language*, *53*(2), 361-382. doi:10.2307/413107

Seedhouse, P. (2004). Conversation analysis methodology. *Language Learning, 54*, 1-54. doi:10.1111/j.1467-9922.2004.00268.x

Seedhouse, P. (2005). Conversation analysis and language learning. *Language teaching*, *38*(4), 165. doi: https://doi.org/10.1017/S0261444805003010

Sidnell, J. (2010). Action and Understanding. In *Conversation analysis: An introduction* (pp. 59-76). Chichester, U.K.: Wiley-Blackwell. Retrieved from http://pc124152.oulu.fi:8080/login?url=

Sidnell, J. (2010b). *Conversation analysis: An introduction*. Chichester, U.K.: Wiley-Blackwell. Retrieved from http://pc124152.oulu.fi:8080/login?url=

Sidnell, J. (2013). Basic Conversation Analytic Methods. In J. Sidnell & T. Stivers (eds.), The Handbook of Conversation Analysis (pp. 77-99). Chichester: Wiley-Blackwell.

Sjöblom, B. (2008). Gaming as a situated collaborative practice. *Human IT, 9*(3), 128-165.

Stivers, T. (2008). Stance, alignment, and affiliation during storytelling: When nodding is a token of affiliation. *Research on Language and Social Interaction, 41*(1), 31-57. doi:10.1080/08351810701691123

Stivers, T. (2010). An overview of the question–response system in American English conversation. *Journal of Pragmatics*, *42*(10), 2772-2781. doi:10.1016/j.pragma.2010.04.011

Stivers, T., Mondada, L., & Steensig, J. (2011). Knowledge, morality and affiliation in social interaction. In T. Stivers, L. Mondada & J. Steensig (Eds.), *The morality of knowledge in conversation* (pp. 3-24). Cambridge: Cambridge University Press. doi:10.1017/cbo9780511921674.002

Stivers, T., & Rossano, F. (2010). Mobilizing response. *Research on Language and social interaction*, *43*(1), 3-31.

Sundqvist, P. (2009). *Extramural English matters: Out-of-school English and its impact on Swedish ninth graders' oral proficiency and vocabulary* (Doctoral dissertation, Karlstad University).

Sylvén, L. K., & Sundqvist, P. (2012). Gaming as extramural English L2 learning and L2 proficiency among young learners. *ReCALL*, *24*(3), 302-321.

Tekin, B. S., & Reeves, S. (2017). Ways of spectating: Unravelling spectator participation in Kinect play. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 1558-1570). doi:10.1145/3025453.3025813

ten Have, P. (1997). In the presence of data: conversation-analysis as 'empirical philosophy'. In *conference on Ethnomethodology, an improbable sociology, Cerisy-la-Salles, France*.

ten Have, P. (2007). Doing conversation analysis; pages 4-13. (2nd ed., ). London: SAGE Publications, Ltd. doi:10.4135/9781849208895

van de Sande, C. C., & Greeno, J. G. (2012). Achieving alignment of perspectival framings in problem-solving discourse. *Journal of the Learning Sciences*, *21*(1), 1-44.

Wagner, J. (2004). The Classroom and Beyond. *The Modern Language Journal, 88*(4), 612-616. Retrieved June 8, 2020, from www.jstor.org/stable/3588592

Walker, G. (2012) Phonetics and prosody in conversation. In: Sidnell, J. and Stivers, T., (eds.) The Handbook of Conversation Analysis. Blackwell Handbooks in Linguistics . John Wiley and Sons, Ltd , Chichester, West Sussex , pp. 455-474. ISBN 9781444332087

Wong, J., & Waring, H. Z. (2010). *Conversation analysis and second language pedagogy: A guide for ESL/EFL teachers*. Routledge. doi:10.4324/9780203852347

Wright, T., Boria, E., & Breidenbach, P. (2002). Creative player actions in FPS online video games: Playing Counter-Strike. *Game studies*, *2*(2), 103-123.

Young, R. F., & Lee, J. (2004). Identifying units in interaction: Reactive tokens in Korean and English conversations. *Journal of Sociolinguistics, 8(3), 380-407.*

# Appendix 1

## Transcription conventions

1. Transcription of speech (Jefferson, 2004)

| | |
|---|---|
| [ | A left square bracket on two successive lines indicates the beginning of overlapping talk by two or more speakers. |
| ] | A right square bracket on two successive lines indicates the end of overlapping talk by two or more speakers. |
| = | Equal signs indicate no pause between the utterances. |
| (.) | A dot in parentheses indicates a pause, which is shorter than 0.2 seconds. |
| (0.5) | Numbers in parentheses indicate the length of a silence in seconds and tenths of seconds. |
| . | A dot indicates a falling or final indication. |
| , | A comma indicates slightly rising or continuous intonation. |
| ? | A question mark indicates rising intonation. |
| ::: | Colons indicate the stretch of the immediately prior sound. More colons, the longer the stretching. |
| - | A hyphen indicates a cut-off of a word. |
| HI | Capital letters indicate a loud voice |
| ↑ | The upward arrow indicates rise in pitch. |
| >hi< | Right/left carats indicate that the utterance between them is faster than the surrounding talk. |
| ha ha | Indicates laughter. Laughter can be referred to in different ways. |
| ( ) | Empty parentheses indicate that the transcriber was unable to hear what was said. |
| (( )) | Double parentheses indicate the transcriber's comment or descriptions. |

2. Transcription of embodiment (Mondada, 2014)

| | |
|---|---|
| * * | Gestures and descriptions of embodied actions are delimited between |
| + + | two identical symbols (one symbol per participant) |
| Δ Δ | and are synchronised with correspondent stretches of talk. |
| *--> | The action described continues across subsequent lines |

| | |
|---|---|
| -->* | until the same symbol is reached. |
| >> | The action described begins before the excerpt's beginning. |
| -->> | The action described continues until the excerpt's end. |
| …. | Action's preparation. |
| ---- | Action's apex is reached and maintained. |
| ric | Participant doing the embodied action is identified when they are not the speaker. |
| fig | The exact moment at which a screenshot has been taken |
| # | is indicated with a specific sign showing its position within turn at talk. |

3. Other abbreviations used in the transcriptions

| | |
|---|---|
| evr | everyone |