THE DISCOURSE OF ONLINE LIVE STREAMING ON TWITCH: COMMUNICATION BETWEEN CONVERSATION AND COMMENTARY

DANIEL RECKTENWALD

PhD

The Hong Kong Polytechnic University

2018

The Hong Kong Polytechnic University

Department of English

The Discourse of Online Live Streaming on Twitch: Communication between Conversation and Commentary

Daniel Recktenwald

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

March 2018

Certificate of originality

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text

_____

Daniel Recktenwald

Abstract of dissertation

Live streaming on *Twitch* is a new form of online broadcasting of video games. Aside from the visual game, it features computer-mediated communication between a speaking streamer and a large chatting audience. This communicative setting is quickly expanding to other types of broadcasting. So far, there has been a limited number of sociological and ethnographic studies that argue that live streaming is a mixture of conversation and gameplay commentary. This conversation is said to break down with an increase in audience size and devolve into a chaotic waterfall of text. These widely shared assessments and arguments are based on common-sense and ad-hoc definitions of terms such as conversation and commentary, but they have not been tested or studied trough discourse analysis of natural occurring data. Therefore, the study sets out to describe the organization of discourse of live streaming on *Twitch*.

The second chapter is a detailed introduction to the phenomenon of video game live streaming, the existing prior research as well as the communicative environment of *Twitch*'s website.

Chapter three is dedicated to the linguistic literature review and will present the analytical framework of the Birmingham school and its current descriptions of spoken conversation. To apply the framework to live streaming of video games, the review will also introduce contemporary research on chat communication and online video chat. This introduction will provide a suitable starting point to discuss the discourse of *Twitch*.

The methodology chapter explains the transformation of the visual gameplay, the streamer's spoken language and the written chat into a coherent transcript, which is open to the discourse analysis. This transcript is studied in the three original research chapters five, six and seven. Each chapter describes the organization of discourse of live streaming from a different perspective, which will allow an overall synthesis of the study in the conclusion.

Chapter five begins with the distribution of discursive moves between the participants and describes who is talking to whom, for which purposes and how often. There are different directions of communication such as from viewer to streamer, from streamer to viewer or from viewer to viewer. Often, these directions of talk and chat correspond to different social purposes within the communication. For example, viewers tend to write questions in the chat, when they address the streamer. In turn, streamers tend to give spoken responses to these questions as they address their viewers. It leads to a very consistent pattern in the dialogue. However, the chapter will also show that the largest amount of communication is not strictly dialogical. Streamers and viewers produce monological moves that talk about the game but have no direct recipient. Dialogical communication between the parties and more monological communication about the game co-occur. Therefore, the chapter will conclude that a direct equation of live streaming with conversation or commentary is too limited and needs further refinement to accurately account for the discourse of live streaming.

In the sixth chapter, the dissertation shows how dialogical and monological moves combine to larger units of discourse and how they relate to the simultaneously unfolding gameplay. This will include the most common dialogical exchange patterns between streamer and viewers, which will be explained

in reference to the streamers' status as players, the chat participants' status as viewers, and their different modes of communication. Their discussion reveals that research on cross-modal discourse must distinguish spoken-to-written communication from written-to-spoken communication. They have a different form and spoken-to-written communication features a new discursive move, the topicalizer, which is absent from written-to-spoken communication. The topicalizer is a repetition of a written chat message in order to topicalize it in the spoken mode. The analysis also shows that the streamer's monological commentary consists of two discursive practices with distinguishable content and a different temporal relationship to the unfolding gameplay. Lastly, the chapter demonstrates that the organization of discourse is highly cyclical and re-occurs in patterns that are aligned to the structure of the game and the organization of the broadcast.

Chapter seven studies 'donation alert messages', which are a unique type of cross-modal exchange that receives privileged interactional treatment during live streaming because they are elicited by the payment of a 'donation'. The chapter classifies the most common types of these messages based on their form and content and discusses their role in the overall organization of discourse on *Twitch*.

Chapter eight summarizes the findings of the three original research chapters and synthesises them in a new descriptive model. The model is an adaptation of Birmingham school's original description of discourse applied to live streaming. It is able to explain which discursive practices are likely to occur in different situations of the broadcast. Finally, the chapter also contains critical reflection on the research and suggest potential future research projects that may challenge established notions of casual conversation as talk that is unaffected by instrumental or monetary motivations.

Acknowledgements

"No man is an island"

A PhD has a single author, but there are always several people involved in the process. There is a list of people without whom none of this would have been possible and it goes as far back as to my final year of undergraduate studies with classes in sociolinguistic and language & gender. My teachers Dr. Bettina Kraft and Dr. Dorothea Halbe taught in an engaging style that encouraged students to critically reflect on their preconceptions and at times even question the reading list. As supervisors for my Bachelor's, they have been patient and supportive and helped me to write my thesis.

After graduation, I was applying for my Masters and I still remember the info evening at Saarland University, where Dr. Nele Gerhardt promised to help me in the application process. Afterwards, she would play a vital role in convincing Prof. Neal Norrick to hire me as a research assistant and I cannot overstress how important it was to be part of his research team and look behind the scenes of academia and university. During that time, I also took academic writing with Dr. Carrie Ankerstein and she taught me the skills of the trade. For the first time I was writing coherent arguments and was not just putting thoughts into words. Without a doubt, my studies at Saarland University have been the best time of my life, professionally as well as personally and it convinced me to further pursue higher education.

It may sound strange to some, but good friends are essential for a successful completion of a dissertation and on several occasions, I had lengthy conversations with my dearest friends Kat & Peggy, who helped me to put things into perspective over a beer or hot pot. For a long time, I was unable to appreciate the tremendous opportunities that I have had access to and spent too much thinking about the negative aspects of a PhD.

In our department, Hans, Stephen and Bernadette consistently fought on the behalf of the students and tried to address some of the issues that were arising and helped many students in the department in a period of great transition. Part of this transition was also my own extension for one semester at which point Hanna and Aditi became my new supervisors and assisted me in the last steps of the journey. Their feedback brought in new perspectives and insights, which greatly improved upon the quality of my work. The last people I have to thank are also the people I have to thank the most. Three times my parents took me back in. In the safe and familiar environment of our home, I was able to fully focus on my work and write my B.A., M.A. and PhD. This has been the hardest thing I've ever done, and I am thankful for everyone who played a role in it. I am eternally grateful,

Daniel

Table of contents

List of figures

List of tables

Chapter 1- Introduction

1.0 Initial remarks

Live streaming is a digital media broadcast over the internet, as well as a social experience between its media users. As a broadcast, it is a live transmission by a host or 'streamer' about his or her hobbies to a spectating audience. Live streaming occurs on websites such as *Twitch*, *Facebook Live*, *YouTube Live, Twitter's Periscope, NicoNico* (Japan)*, Afreeca* (Korea), *Douyu* (China) and there is an increasing diversity in activities that are performed including video gaming, mukbang (Korean 'eating broadcasting'), fishing, v-logging, singing or composing music, drawing and fitness workouts. In China, the world's largest and most innovative live streaming market, the revenue was estimated to be 20.6 billion Yuan for 2016 and is forecasted to have doubled to 43.6 billion Yuan in 2017[1].

The most popular content in terms of viewer numbers is the live streaming of video games and its global market leader is *Twitch*, which describes itself as "the world's leading social video platform and community for gamers (Twitch, 2015)." The social experience is realized through the communication between the participants. On *Twitch*, the streamer is playing a video game while talking to his or her viewers. In turn, the viewers can send written chat messages to him or her. So far, academic research dedicated to live streaming is limited and is spearheaded by researchers in ethnography, sociology, game studies as well as human-computer interaction (HCI). These studies conventionally label live streaming as a form of "conversation" or "commentary" but often omit detailed descriptions of its language (Smith et al., 2013; Hamilton et al, 2014). Linguistic studies that analyze the discourse of live streaming are notably rare. This is a problem as well as an opportunity. It is difficult to approach, handle or even define a research phenomenon that is not well established. There is no pre-approved way to study live streaming, which makes it harder to zero-in on the relevant research. At the same time, a study on a new topic is an opportunity because there are few path dependencies found in research. Research into live streaming is 'wide open' with respect to unique and study-worthy phenomena as well as the number of approaches that can lead to insights.

This study is designed to describe the discourse of live streaming on *Twitch* and provide a holistic understanding of its form, content and overall organization. Its description of the discourse is the result of a gradual process that is data-driven and theory-guided. In this dissertation, live streaming is understood as a form of computer-mediated communication (CMC), and the study will be the first micro-level description of the interaction between the

[1] http://www.scmp.com/tech/enterprises/article/2124662/chinas-booming-live-streaming-industry-may-have-reached-its-peak

speaking streamer and the chatting audience under consideration of the video game. The study adjusts prior conventions of data collection, language transcription and uses the analytical framework of the Birmingham School of Discourse Analysis (Sinclair & Coulthard, 1975). This puts the study in the area of discourse analysis, specifically the discourse analysis of CMC, but it also maintains the importance of the video game for the broadcast and communication.

Such an undertaking raises the question as to why one should study live streaming and the broadcast of gaming, with the next subsections presenting live streaming as a study-worthy and under-researched phenomenon (1.1), the broadcast of games and gaming as a specifically interesting area of investigation (1.2) and why I am qualified to undertake this research project due to my own background (1.3). This background consists of my formal training in several linguistic approaches, as well as my steady personal interest in gaming and social media. There is a synergetic relationship between the two aspects that benefits the depth and breadth of the study. Moreover, this has informed the study's perspective on live streaming as a form of computer-mediated communication that is describable through the means of discourse analysis.

Section 1.4 deconstructs live streaming into its integral parts and participants. The purpose of this is to point out its novel features and how they are relevant for linguistic descriptions of computer-mediated communication. It also serves to justify the research question of the study and explain the dissertation's structure. This dissertation consists of eight chapters that consecutively built on one another. Each part of the dissertation is influenced by the former and reciprocally influenced by the latter, with every chapter having an important role in answering the overall research question. This step-wise analysis of live streaming results in a coherent and holistic descriptive model for the new form of computer-mediated communication

1.1 Live streaming as broadcasting and media convergence

Live streaming has a literal meaning as broadcasting technology and a more figurative meaning as a set of media practices. As a technology, live streaming is a form of digital distribution of information. Digital information can be distributed as 'data' and 'codata'. Data is information, which is fixed in a tangible form and usually anchored on the storage medium such as CD-ROMs, DVDs or on USB sticks. Stored information is persistent until it is removed from the carrier. This persistence makes them re-readable or re-playable.

Codata is information that is continuously broadcasted and is not stored on persistent carrier medium. The physical carrier is the digital signal that is transmitted from an information source to a designated receiver of the information.

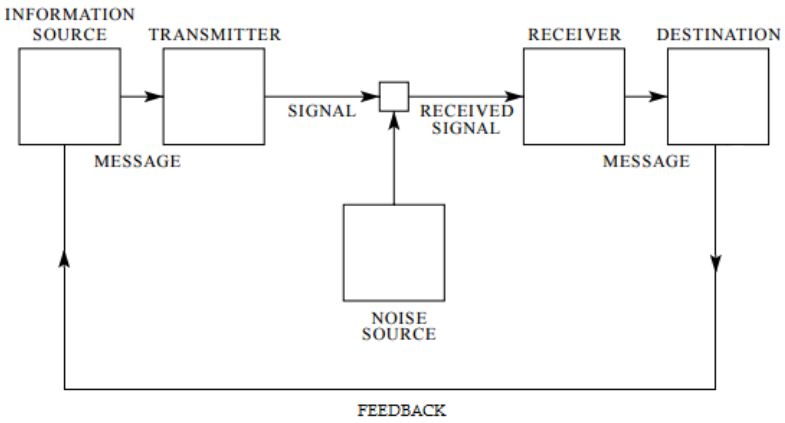


Figure 1 General communication system

This transmission process is described in Shannon-Weaver's (1948, p.3) 'model of communication'. It concisely represents the flow of information during transmission that begins with the information source. In live streaming, this message is sent as a digital signal over the internet to the receiving computer, which in turn decodes the signal and presents it to the 'destination' or viewer. The digital signal is continuous, volatile and once the transmitter stops sending, the message ceases to exist. Analogue examples of such media transmission systems are radio stations or commercial television. On TV or radio, the feedback loop was optional and rarely used. If it was desired, people relied on external media such as telephone call-ins. In this sense, the traditional model was mostly one directional from sender to receiver, privileging the sender because they controlled the content of the message as well as its distribution channel. However, this transmission model also imposed cost on the sender in the form of media content production and the maintenance of infrastructure for their distribution.

As a media practice, these costs were refinanced though the sale of advertisement or through the broadcast of premium / pay-per view channels. During the 20th century, economic pressures and fierce competition lead to several mergers across the globe. In the US, five companies *Comcast*, *Walt Disney*, *21st Century Fox*, *Time Warner* and *CBS/Viacom* own 90% of their domestic media landscape and the dominant process of traditional broadcasting is media concentration (Noam, 2016). It funnels the editorial discretion of what should & should not be produced into the hands of only a few actors. Programming is decided from the top-down and as these actors try to capture large portions of the mainstream market, they homogenize the content to appeal to as many viewers as possible (cf. Noam, 2016; Jenkins, 2005; 2006). Therefore, the 'offline' media concentration has led to less diversity in broadcasting in the number of media producers and the types of media content that are produced.

This development saw a major disruption in the early 2000s with the emergence of web 2.0 (O'Reilly, 2004). Many new websites do not produce their own content but instead

specialize in the transmission of content and its distribution to the users[2]. They serve as a market or middleman that connects producers with consumers. Often, these websites are highly specialized and have a single dominant overarching service. For example, social network sites such as *Facebook* connect friends and acquaintances, or *Twitter* is used by musicians, politicians, companies or activists to reach out to their followers (Boyd & Ellison, 2007; Page, 2012). Other examples are the Hong Kong based *9GAG*, an image-sharing website for humorous pictures and the video-sharing platform *YouTube*.

Live streaming is the newest trend of digital media distribution and mixes the features of an ephemeral live broadcast with the advantages of the decentralized and user-driven web 2.0. There is little editorial oversight and few interventions so long as the the content does not conflict with the website's terms of service. Streamers can use computers, gaming consoles or even smartphones to broadcast globally, which significantly decreases barriers of entry and allows more individuals to start producing media content. There is a renewed diversification of media formats that go beyond the mainstream of pop music or television dramas**.**

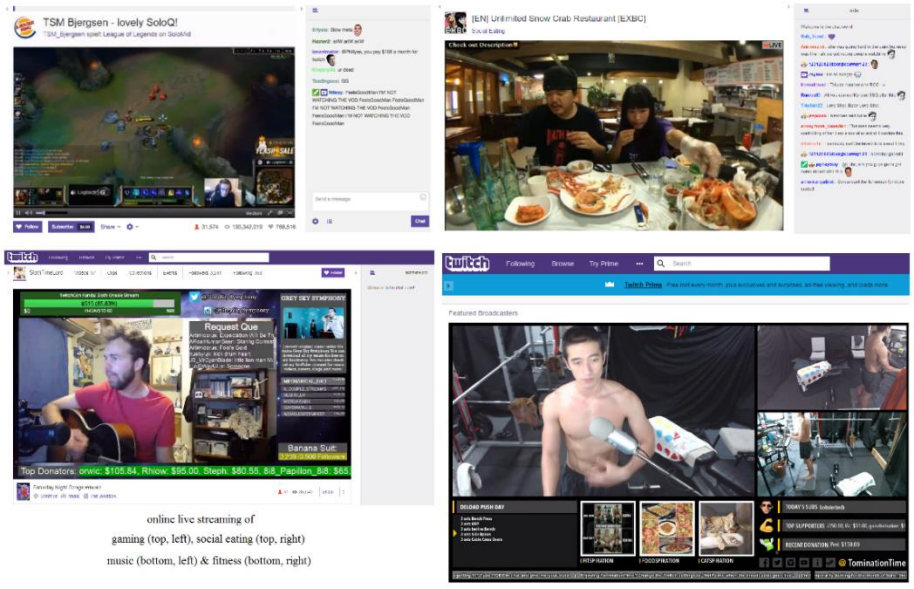


Figure 2 Different types of livestreams

Online broadcasting enables the scaling of niche programs that have difficulties in finding a distribution through traditional channels (cf. Taylor, 2011). Although there may not be enough viewers for gaming, fitness or eating broadcast in a single country, there is enough of an audience spread across the globe. In its most positive formulation, live streaming is part of the

[2] Internet distributed television such as *Netflix* and *Amazon Prime video* are 'old' media that use the internet solely as a transmission channel, while maintaining a premium channel subscription model

wider democratization of media from the bottom-up, for example viewers can choose a program that fits their individual interest or even start broadcasting themselves (Jenkins, 2005; 2006).

Historically, such innovative processes have begun at the margins and mostly appealed to younger early adopters that are open to change and new developments (cf. Rogers, 1962; Jenkins, 2005; 2006). Over time, the previously marginalized content becomes more widely accepted and ordinary types of media content start to migrate to the new distribution platform. This process has already been observed for *Facebook* and *YouTube*. *YouTube* started out with niche programs such as anime music videos or self-recorded prank videos and has since branched out to distribution of pop music, internet-based TV stations and films. The same development towards diversification and growth is also visible in the sector of live streaming. *Twitch* sees increasing competition from the offerings of *YouTube*, *Twitter*, *Facebook* and *Instagram* in the West and *Douyu.com* in China. In 2014, *Amazon* purchased *Twitch* for $970 million dollars and *Twitch*'s annual report show a consistent growth in viewers and hours watched (Twitch, 2014; 2015; 2016; 2017). This suggests that live streaming might reach levels of mainstream significance that are comparable to social network sites or video on-demand services. The rise of these type of sites and technologies have brought about many changes to people's social behavior and new media practices.

The new media practices have been very interesting for linguistics because they usually include a communicative dimension. New media websites provide unique and novel affordances that have been utilized in unforeseeable and creative ways (Hutchby, 2001). Linguistic research has consistently described these practices and studied their specificity but also generalized their implication for the linguistic models (Bolander & Locher 2014; Giles et al., 2015, 2017). In live streaming, the streamer is using the spoken mode to viewer groups, who chat in the written mode. In linguistics, such discourse is characterized as 'cross-modal' because it crosses over from one mode of communication to another (Rosenbaun et al., 2016b). However, cross-modal communication has been subject to only a limited number of studies (cf. Sindoni, 2014; Rosenbaun et al., 2016a, b). The studies cover online chat programs, such as *Skype* or *Google Hangouts*, where conversation is the only activity and cross-modal communication is an exception.

Live streaming is different because streamer and viewers are distinguishable parties with different access points to the communication and cross-modal communication is the norm in the interaction. Additionally, live streaming features a simultaneously performed activity that influences the discourse between the participants. Although live streaming has been hypothesized as a conversation, no prior studies have analyzed the actual discourse under

consideration of the broadcasted activity. Therefore, it is unclear how the participants negotiate the dialogue between streamer and audience, nor how the video game impacts the patterns of communication.

Moreover, live streaming is closely linked to a new method of monetization through so called 'alert messages'. These messages are special because the viewer pays for them to appear saliently in the visual center of the broadcast. Alert messages are a type of promoted message that first appeared in January 2014 and have since become ubiquitous across almost all live streaming platforms. It is a genuinely novel communicative practice that has no obvious 'offline' precursor and appears to significantly affect the patterns of interaction during live streaming. The impact of the game and the existence of alert messages challenges an overly simplistic categorization of live streaming as a conversation and invites a closer look at its linguistic practices and overall organization.

Once it is established that live streaming is a subject worthy of linguistic investigation, one might hear a critical follow-up question.

1.2 Why study the broadcast of gaming?

Every research project needs strong reasoning to justify the undertaking. However, in some areas this justification is almost automatically assumed due to the merit that is given to the discipline. Within the humanities and social sciences, there is a reversed position and studies may face skepticism and allegations of being frivolous, such as being only for the pleasure of the researcher and without an intrinsic value (Sandelowski, 1997; Glass, 2009). This can be particularly problematic, if there is stigma surrounding the research topic. Games and gamers often remain stigmatized in dominant discourses about the supposed violence-inducing effects of video games or claims over video game addiction (Bax, 2016; Brus, 2012; Shaw, 2012).

Amidst these strong negative currents, there is a slow but steady change in public opinions on video games (McKernan, 2013; Szablewicz, 2015). A main reason is their increasing relevance as an economic product. In the US in 2016, video games generated a revenue of $30.4 billion[3], which is comparable to the 38 billion of the film industry[4]. The number of gamers is also increasing, and more than half of all US households own dedicated gaming hardware (Entertainment Software Association, 2015). These developments in the gamer landscape also influence live streaming of video games, which is projected only to grow more and remain the most significant form of live streaming (Warman, 2015).

[3] http://www.theesa.com/article/u-s-video-game-industry-generates-30-4-billion-revenue-2016/
[4] https://www.statista.com/topics/964/film/

Academia has seen a strong increase in studies about games around a variety of new research topics (Mäyrä et al., 2013; Quandt et. al; 2015, Melcer et al., 2015). Historically, the first perspective that emerged was 'ontological games research' (Carter et al., 2014), which is concerned with the study of games as digital objects. Games are a uniquely human invention that consist of complex rules, goals and in many cases a central narrative (Murray, 1997; Juul, 2003; Salen & Zimmerman, 2003). Therefore, games are inherently study-worthy like other forms of human expression and this research describes the integral parts of games and how they interact. The central question of ontological research could be described as 'what are games'?

Secondly, games have become a vital part of contemporary popular culture on the internet and mediate the experiences of everyday life (Couldry & Hepp, 2013; Livingstone, 200)). For example, on internet forums such as *Reddit* or the social media website *9GAG*, users routinely comment critically and humorously on global events through game references and game memes (cf. Knobel & Lankshear 2007). Games are part of humanity's cultural repertoire through which they frame the interpretation and production of meanings. Their study reveals how people make sense of the world and what they consider (extra-)ordinary. As cultural artefacts, they are ideological vehicles that can convey dominant narratives such as 'the war on terror' in games like *America's Army* or *Call of Duty: Modern Warfare* (cf. Nichols, 2009); or they might reinforce values of individualism, capitalism, and neoliberalism (Frasca, 2001; Perez Latorre, 2015; Möring & Leino, 2016). From this perspective, games-in-culture are influential 'texts' of the late 20th and early 21st century that are connected to gamer culture, popular culture and thereby all human culture (cf. Miller, 2006). The study of games and gamer culture increase our understanding of contemporary life. The central question of such research could be summarized as 'what & how do games mean'?

In a third perspective, video games are a site for informal learning. This includes second language learning, situated learning and collaboration as well as the general acquisition of computer-related skills (Piirainen-Marsh & Taino, 2009; Lange 2011; Hung, 2009; Bennerstedt & Ivarsson, 2010). The players' engagement with and in video games prepares them for other tasks that involve computer technologies and digital interfaces. The central question of this research strand is 'how could games be used for learning or training'?

These three perspectives produce very valuable insights and inform several aspects of the study. For example, the conceptualization of games as dynamic systems influences the development of the transcription scheme for online live stream and the research on learning in games assists in the explanation of examples.

Yet, the main reasons for studying the broadcast of gaming for this dissertation fall within practice and epistemological research (cf. Carter, 2014). Practice research is the study of games-in-use or during play and it explores the game-related practices that they enable. Since this study is situated within linguistic research, it is particularly interested in how gaming is part of the organization of discourse. Thereby, the study connects to previous ethnographic and sociological research about the "player as commentator" or the "meaningful conversation" between streamer and viewers (Smith et al., 2013; Hamilton et al., 2014). Importantly, these previous discussions about the communication of live streaming have remained rather superficial and lack actual analysis of discourse. This study can fill this research gap by looking at the communication on *Twitch* at the micro-level and describe the discursive practices based on natural occurring data. The micro-level analysis will confirm, refine and challenge previous conceptions about live streaming of video games.

The research will also allow more general conclusions and descriptions of live streaming that are not specific to the broadcast of video games. In this sense, live streaming of video games is an archetype for a new form of computer-mediated communication, whose linguistic features are typical for many varieties of live streaming.

1.3 About the author

The origins and trajectory of any research is influenced by the researcher and their relationship to the topic. In most cases, researchers gravitate towards areas where their expertise overlaps with their interest. The same is true for this study that formally began in 2014, but in practice is tied to my tertiary education, formal training as a linguist, and my history as a gamer. In some disciplines, it would be frowned upon to talk about oneself because it would threaten the myth of complete objectivity (Mittrof, 1972; Ratner, 2002), but especially the social sciences and humanities acknowledge that a researcher's relationship to the research topic is not necessarily a problem and often a potential resource. I follow Aarseth (2003) and Witkowski (2012), who argue that having lived experience as a gamer can benefit the interpretation of a data-driven and theory-guided analysis. It combines the knowledgeable gamer and analytical researcher, which is together more than the sum of its parts.

My past as a gamer began with single player role-playing games such as *Pokémon* in 1996 on the *Nintendo Game Boy* and *Final Fantasy VII* in 1997 on the *Sony PlayStation*. As an 11-year-old, I was unable to critically deconstruct or even explain why the games appealed to me, but I know that they allowed me to make friends over a shared hobby. Classmates played the same games and every morning we could talk and compare our adventures.

In 2002, my parents finally decided to purchase internet access and I was able to switch from gaming consoles and handheld devices to desktop computers with online games. This transition was very important to me and the study. I started chat online on *Internet Relay Chat* (*IRC*) and began playing the online soccer manager *Hattrick.org*. In *Hattrick*, users were able to manage their own clubs, but the matches were calculated by the game engine based on the qualities of the teams. Saturday 6 PM was match day and matches played out in real time and were watchable on the website. As a result, many users watch their games in real time and simultaneously joined the *IRC* chat channels. Participants watched each other's matches together and communicated within the chat. It was a very rudimentary version of game-spectatorship and through the chat it became a shared social experience among users. It fostered virtual communities that translated into local and national meetups (cf. Rheingold 1993). Online acquaintances became offline friendships that are lasting to this day (cf. Antheunis et al., 2012). Games and social media on the internet develop and change at a rapid pace, but these friendships survived the decline in popularity of *Hattrick.org* and *IRC*. We migrated together to *World of Warcraft* in 2006, which was the breakout title of the massively multiplayer online role-playing game (MMORPG) genre. *World of Warcraft* spawned a decade of research in the social sciences and humanities about social groupings, new forms of virtual communities and reconsiderations of the relationship between play and work (Steinkühler & Williams, 2006; Ducheneaut et al. 2006; Taylor, 2009; Turkle, 2012). For me, it was the home of my guild 'RISE'.

World of Warcraft had an integrated chat with several channels for different purposes. The most important channels were the guild chat and raid chat, which allowed social communication and talk about the collaborative play. Most players of our guild greatly enjoyed this time but there was a major difference to the experience in *Hattrick*. *World of Warcraft* never had dedicated meet-ups on the regional or national level and all friendships remained exclusively online. This may be one of the reasons why none of my friendships and most guilds in *World of Warcraft* do not last. Online chat, co-play and game spectatorship brought the people together and produced greatly rewarding social experience, but the social relationships remained one-dimensional and limited to the game. As the game lost its appeal, so did the connections to the co-players.

The last notable game I started playing was the Multiplayer Online Battle Arena (MOBA) game *League of Legend* and it also has been the first title I started watching on *Twitch* in 2012. *League of Legends* is part of new generation of games that have replaced persistent player-run communities of clans or guilds with algorithmic and AI driven matchmaking. Instead of manually having to form groups with co-players and opponents, the game automatically

provides teams and groups for the duration of a round (Zagal & Mateas, 2007; Zagal et al. 2008). After the round, the players are mixed again with a completely new set of opponents and allies. This significantly hinders or at least changes the type of community formation that occurs in newer games (cf. Kou & Gui, 2014). The new format has clear advantages as there is much less organizational and social skills required to play games. The game system forms the teams and players only need to focus on the match at hand. For my personal enjoyment of play, it was a negative change because it diluted my social experience and I stopped playing video games altogether in 2014 until my graduation in late 2018.

However, during the time as an active player, I experienced virtual communities of online chat and gaming and I acquired a game literacy that allows me to read and understand the events of a gaming broadcast (Buckingham & Burn, 2007; Zagal, 2008). I am familiar with the language of gaming and its highly specialized jargon (Ensslin, 2012; Consalvo, 2009), which is beneficial for this study.

At the same time, there is no risk of overly techno-optimistic interpretations because my greater distance to games and gaming occurred simultaneously with my professional transition from undergraduate to post-graduate. In my bachelor's thesis, I worked with Critical Discourse Analysis (Fairclough, 1989), an analytical framework designed to question dominant modes of representation through systematic analysis of semantics and grammar. Although it is not my main analytical framework anymore, its core idea of critical deconstruction has remained an important part of my thinking and analytical practice. The semantic and grammatical research has been replaced by discourse analysis, which studies the meaning of communication through its unfolding in the interaction. Throughout my Master studies, I have applied such discourse analysis to different types of media texts including TV shows, documentary films and most importantly videos on *YouTube* (cf. Recktenwald, 2014). My Masters thesis was a study on the dialogue of *YouTube* Let's Play videos. It discussed the talk of the YouTuber to his or her imagined target audience and the avatars in the game. The study was an indirect pre-cursor to this research project on a smaller scale and it allowed me to do research on interaction in computer-mediated environment that includes a talking gamer and the unfolding of gameplay.

The difference between *YouTube* Let's Play videos and broadcasting on *Twitch* is the live element. It has a significant impact on the communicative practices that I was observing. When I started watching in 2012, neither the research in linguistics, nor game studies about *Twitch* were addressing these new practices. Therefore, this dissertation originates at the intersection of my professional training as a linguist and my personal interest in the subject matter.

1.4 Research questions and structure of dissertation

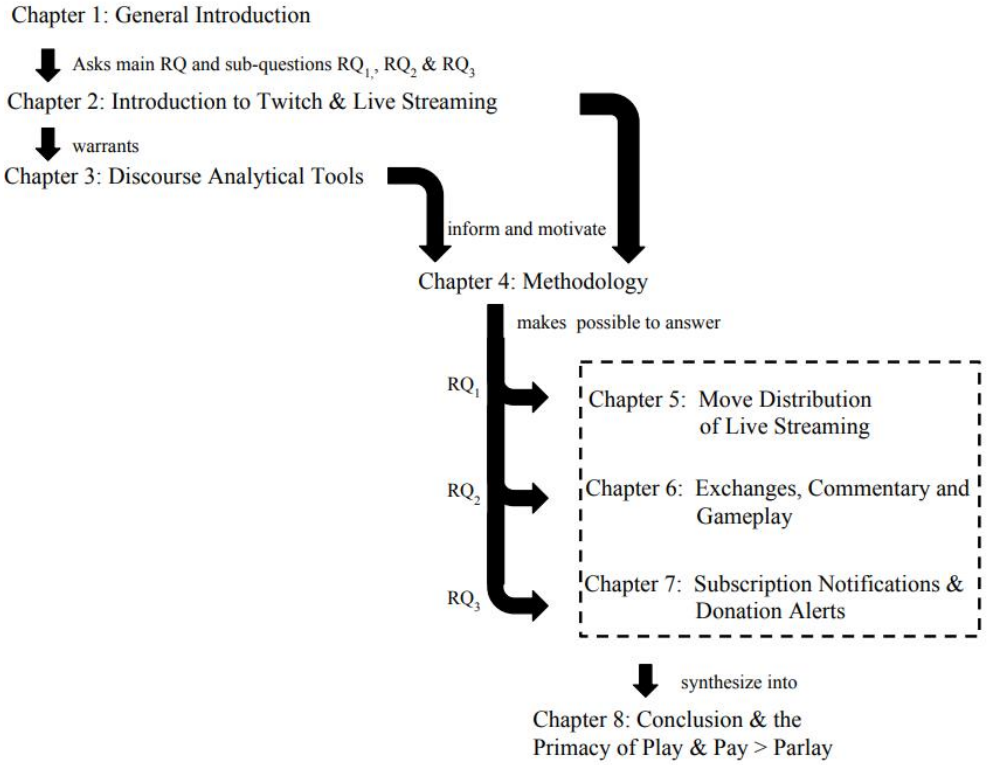


Figure 3 Structure of dissertation

The figure shows the structure of the dissertation and the relationship between the chapters. This study treats live streaming as a form of computer mediated communication that co-occurs with the broadcast of video gaming. The gaming and communication are interconnected interactional phenomena that influence each other. The streamer is talking, while s/he is playing the game and the audience is chatting as they spectate the gameplay. This leads to a cross-modal discourse that is impacted by the configuration of the speaking and writing parties and the simultaneous developments within the game. Considering the underexplored nature of this communication, the study asks the research question:

RQ: How is discourse organized during live streaming of video games?

To answer this broad research question, some qualifiers are necessary, and the research question has to be broken down into three more manageable parts. Each sub-question (RQ_1, RQ_2, RQ_3) will correspond to an original research chapter. Together, they will enable a synthesis in the dissertation's final chapter. In the conclusion, the dissertation will propose a descriptive model that can explain the organization of discourse of live streaming from its smallest units of discursive moves and exchanges to the larger units of organization such as the

daily broadcast session and the weekly streaming schedule. This model has been developed over time as different aspect of the discourse of live streaming have been studied.

A starting point has been the previous research on live streaming on *Twitch* and video games. Chapter two will offer a comprehensive summary of the most important studies that discuss *Twitch*'s interaction and will point out commonly agreed ideas about Twitch's communication. These descriptions usually distinguish broadcasts based on the size of the spectating audience. Small or medium-sized channels are labelled 'conversation', 'commentary' or narration, whereas large-sized broadcasts are named as 'cacophony' or 'waterfall of text' (Hamilton et al., 2014; Smith et al., 2013; Karhulahti, 2016; Nematzadeh et al., 2016; Gandolfi, 2016). These labels have been applied in very categorical terms and were often seen as being mutually exclusive. Each of them suggests vastly different patterns in the organization of discourse between participants. Conversation implies many dialogical exchanges and communication between participants, whereas narration or commentary hint towards more monological stretches of talk by the streamer (cf. Eggins & Slade, 1997). Lastly, cacophony and waterfall of text would suggest the absence of any order in the communication. A proper evaluation of their appropriateness is difficult because prior studies on *Twitch* remain very vague in their use of the terms and do not test them on actual data. Instead, they rely on ad-hoc and common sensical definitions and auto-ethnographic observations. The studies lack an analytical entry point that would allow them a more detailed study of the data.

As part of a solution to this problem, the second chapter will give a detailed description of *Twitch*'s website and introduce video games as dynamic systems. The website is the communicative environment within which communication unfolds. Its description will establish the streamer and their viewers as two distinct parties with different sets of affordances, i.e. streamers speak and play, while the audience is chatting and spectating. The streamers must negotiate the demands of the communication with the demands of the interaction in the game. The game limits the actions that are possible in play and imposes interactional demands on the 'streamer-as-player'. The type of demand depends on the game and its game pace because fast-paced game requires more direct attention than slow paced games.

Moreover, the introduction to the website and video games are necessary for the development of a transcription scheme for live streaming in chapter four. This scheme includes the spoken language of the streamer, the written chat of the viewers as well as the unfolding video game. It transforms the interaction into a written record that is accessible to linguistic analysis. The features and details of the transcript depend on goals of the study as well as the analytical framework and its unit of analysis.

Therefore, the third chapter is the linguistic literature review and it explains how the discourse of live streaming will be studied. The main tools for language description and discourse analysis are taken from the Birmingham school of Discourse Analysis. This approach positions discursive moves as the basic unit of communication and exchanges as the central unit of dialogue (Sinclair & Coulthard, 1975). Through the study of moves and exchanges, discourse is analyzable in a systematic way. The chapter will present the model as it was first developed to describe discourse as a system of ranks, where lower level ranks, such as moves and exchanges, combine into larger level units, such as 'transaction' and 'the interaction' (Sinclair & Coulthard, 1975; Francis & Hunston, 1992). In its original formulation, it described class room discourse in reference to the unfolding lesson. More generally, the model is suitable to describe the organization of discourse in relationship to an activity, for example the broadcast of video games.

The second and the third chapter lead to a state-of-the-art overview of studies that are relevant for live streaming and the description of its discourse. Chapter two presents the previous interpretations of *Twitch*'s discourse as dialogical conversation, monological commentary or 'cacophony' and the third chapter introduces the linguistic tools that can reliably assess discourse as a dialogue or monologue. This will make it possible to describe *Twitch*'s interaction by comparing the prevalence of dialogical communication between participants with the prevalence of more monological discourse such as commentary or 'cacophony'. Therefore, the first sub-question RQ_1 will be**;**

> How are the discursive moves distributed between streamer and audience and does this distribution change with audience size and the pace of the game?

This question can be answered by counting the number of discursive moves by the streamer and the viewer and comparing them to the total number of moves that were produced. Each discursive move in the data set has been coded for its origin and its recipient. The results of the coding are discussed in chapter 5 and they show that previous categorical conception of conversation, commentary or cacophony are problematic. Live streaming has a distribution of moves that is partially dialogical, i.e. between streamer and audience member, as well as moves that are more monological and without an addressee. This pattern varies very little across broadcasts regardless of pace of the game or the size of the audience. The study interprets these findings to indicate that the organization of discourse is rather consistent across different video games and very similar for channels with large or medium-sized audiences. Monological and dialogical moves occur in channels of all types and sizes.

On its own, chapter 5 cannot become more detailed because it is limited to the study of individual moves and their distribution between participants. To continue the description of live streaming, the study has to look beyond isolated moves and ask $RQ_{2:}$

> How are the discursive moves combined into larger units of discourse and how is this process influenced by the unfolding gameplay?

This research question will be the topic of chapter 6. The chapter will describe how discursive moves by streamer and viewer are combined into dialogical exchanges (6.1) or monological stretches of commenting and reporting (6.2). The sections will give detailed accounts of their respective phenomena and demonstrate that the emergence of game events often changes the trajectory of discourse. For example, an exchange between streamer and viewer can become interrupted by a sudden game event, or a series of fast-paced game events which may prevent the streamer from continuing that dialogue with the audience.

This influence of the game on the organization of discourse also persists on higher ranks (6.3). For example, its organization in rounds and matches leads to cyclical repetition in the discourse. As the different stages of the game repeat, the patterns of the communication repeat as well (see section 4.1.2 & 6.3). Streamers play many matches in their broadcast sessions. They can last from 6 to 12 hours and during that time, the same discourse topics re-emerge routinely as similar game events or stages of the game repeat. The daily broadcast structures the interaction between the participants, which is itself part of a larger weekly streaming schedule. This schedule determines on which days of the week the streamer is broadcasting and interacting with the audience.

Based on these findings, the chapter will demonstrate that the discourse of live streaming is very structured, as it adheres to the organization of the video game in rounds and levels as well as the organization of the broadcast. This is comparable to a daily and weekly work shift. This finding makes it possible to explain when certain discourse practices such as cross-modal exchanges or monological commenting occur, and how they are connected to the broadcast and the unfolding gameplay. However, the occurrence of ‘alert messages’ and their responses by the streamer cannot be explained via a reference to game events or the stage of the broadcast. They require an additional inquiry and the last research question RQ_3 will ask:

> What is the social purpose of alerts messages and what is their role in the organization of discourse?

This question is addressed in chapter 7, where the study discusses the two different types of alert messages, namely subscription notifications and donation alert messages. The chapter will show that they work similarly to cross-modal exchanges but require a financial payment for the dialogical initiation. The chapter will explain the exact mechanism that is unique to live streaming and then provide a typology of donation alert types and their social purposes. This typology will suggest that alerts message have an interpersonal purpose for the audience, giving them the opportunity to interact with the streamer on a more intimate level. They receive almost immediate and very direct feedback and have privileged communicative access to the streamer.

Chapter 8 is the conclusion and discussion chapter that brings together the findings of the analyses in chapters 5, 6 & 7. It will synthesize them and give an overall assessment of the organization of discourse and argue that there is an underlying principle that can explain the priorities that are given to some types of interaction over others. In live streaming, there is a primacy of play & pay over parley, where the payed communication and the unfolding gameplay receive preferential interactional treatment over the more conversational discourse (parley). This principle can be embedded into a descriptive model that accounts for the organization of live streaming on *Twitch*.

Non-linguistic organization	*Discourse*
Streaming Schedule	
Broadcast Session	**Interaction**
Pay & play > parlay	**Transaction**
	Exchange / Commenting & Reporting
	Dialogical Moves / Monological Moves

Table 1 Organization of discourse in live streaming

The table above is a representation of the model and it is an appropriation of the original description by the Birmingham school (see chapter 3; Sinclair & Coulthart, 1975). It argues that moves, exchanges and transactions form a rank scale and that they are tied to the non-linguistic organization of the broadcast. For example, if there is an important game event, the participants transactions are very likely to be commenting or reporting on it. In the absence of game events or between rounds, the discourse consists of more dialogical exchanges between streamer and viewer. The overall pattern of the interaction in discourse is cyclical and bound to the cyclical organization of the broadcasting session and the weekly streaming schedule

The final section 8.3, initiates a discussion about a suitable avenue for future research. Traditionally, casual conversation has been understood as “talk for the sake of talk (Eggins &

Slade, 1997, p.27)" and free from monetary considerations. People socialize voluntarily because it is part of human nature. However, internet celebrities on streaming platforms, including *Twitch*, foster groups of followers that they call 'their community'. These communities are branded around the streamer's persona. The larger these communities become, the more difficult it is for the streamer to have meaningful interactions with their members. At the same time, the growth of the community is an important goal for many streamers (TwitchCon, 2015). Hamilton et al. (2014) consider this a paradox because meaningful interaction stands in conflict with success as a streamer. However, this paradox quickly disappears if one assumes a utilitarian and profit-driven motivation for interpersonal talk. Streaming platforms promote new forms of social interaction, where new media celebrities engage in social communication with the intent of extracting monetary value from their viewers. These developments already occurred as part of cam girl culture in the late 1990s (Senft, 2008), but have reached a new level of institutionalization and professionalization particularly in Mainland China. This may habituate vulnerable individuals to seek out interpersonal contact with financial means, rather than building lasting friendships with people who are genuinely interested in them.

Chapter 2 – Live streaming and game spectatorship

2.0 Initial remarks

This study discusses live streaming as a form of computer-mediated communication that is embedded within the broadcast of a video game.

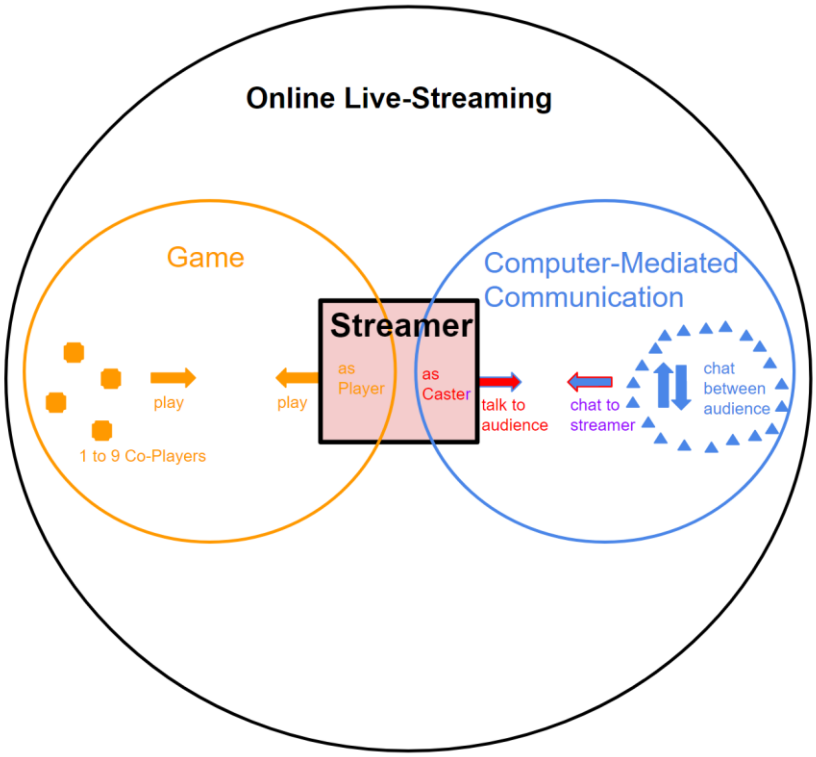


Figure 4 Conceptual visualization of live streaming

The figure is the visualization of the study's analytical lens as it is applied to live streaming. Live streaming consists of two connected interactional phenomena. The streamer-as-player is playing the video game with his or her opponents. The interaction occurs within the confinement of the digital video game. The number of co-players varies by game, but often ranges from one to nine. In figure 4, this interaction is color-coded in orange[5].

Then, there is the computer-mediated communication between the streamer-as-caster (short for broadcaster), who is talking to his or her chatting spectators. Usually, the audience are not co-players of the streamer and only watch the game. They participate through spectatorship and written computer-mediated communication. Audience members can chat with each other or write to the streamer, while streamers mostly talk to their audience. This

[5] The color coding of orange, red and blue will be used consistently as visual cues to support the analysis in chapter 6 and 7

means there is mono-modal, written communication among viewers and cross-modal communication between streamer and audience (cf. Rosenbaun et al., 2016a).

The visualization also signals that the streamer must negotiate the computer-mediated communication with the gaming, while the viewers' communication is not hindered by the demands of the game. In order to describe the organization of the discourse, the dissertation will have to address these interconnected features of live streaming.

Section 2.1 will historically situate the phenomenon of video game spectatorship. Game spectatorship, as it is practiced today on *Twitch*, is part of a development that began 'offline' in the video gaming arcades during the 1970s and received a significant boost in the late 90s and early 2000s with the proliferation of the internet (cf. Borrowy, 2012). The section will also outline several developments of 'online' game spectatorship including game replay files, machinima, *YouTube* videos up until 'emulated' livestreaming. It demonstrates that game spectatorship is not a completely new phenomenon and has existed in different forms that depended on the technological capabilities of its time. Game spectatorship and the communication that it affords has co-evolved and part of this process was the transformation of experimental streaming website *Justin.TV* to modern day *Twitch*. Live streaming on *Twitch* is the result of media convergence that brings together players and game spectators in an environment that facilitates communication between the two parties (cf. Jenkins, 2005; 2006).

Section 2.2 is dedicated to the growing body of research about the interaction on *Twitch.* The review will present common themes in these discussion about *Twitch's* communication and will evaluate them from a discourse analytical perspective. It will be argued that prior studies suggest valuable concepts such as 'player as commentator' or streaming as an informal conversation, but often lack actual description of their proposed ideas (cf. Smith et al., 2013, Hamilton et al., 2014). Often, these studies discuss *Twitch* based on secondary data sources such as interviews and questionnaires, but they rarely analyze naturally occurring communication. This dissertation will take their arguments as starting points, which invite further micro-level discourse analysis to improve our understanding of live streaming and to fully account for the organization of its discourse.

The communication on *Twitch* is an archetype for a new form of computer-mediated communication, where the discourse is embedded into the broadcast of an activity. To make this computer-mediated communication analytically accessible, section 2.3 will introduce the layout and features of the website and its different communicative affordances for streamers and viewers. Novel to live streaming is that streamer and audience are two distinct parties, which have different sets of affordances. It is an interesting site for discourse analysis because

its configuration of participants and communicative setting are underexplored in linguistics. Particularly important to the communication is also the involvement of the streamer-as-player with the video game.

Therefore, the study requires a basic understanding of games and play. 2.4 provides the necessary metalanguage to discuss gaming and games in analytical terms. It will position games as the framing structure of play, which consists of game components, rules & procedures, challenges & rewards and a temporal segmentation in rounds (Salen & Zimmerman, 2003; Hitchens & Tychsen, 2009). These properties of games combine into an interactive system that makes gaming possible (cf. Juul, 2002).

The introduction to play and games will serve two analytical purposes. First, it will facilitate the analyses of the game-related discourse on *Twitch*. Research about live streaming has identified talk about the game as an important element of the communication, but provided very few explanations and examples (cf. Smith et al., 2013; Karhulahti, 2016; Gandolfi, 2016). The proper vocabulary will enable a more detailed description that goes beyond the intuitive labels discussed in 2.2. Secondly, the introduction to the temporal organization of video games will inform the data selection in 4.1 and will become a key aspect in the description of the organization of discourse in section 6.3.

In its four subsections, the second chapter will introduce and reflect on the current knowledge about live streaming and inform the steps in the analysis. However, before doing all this, it will be important to begin outlining the historical development of video game spectatorship and player-audience interaction.

2.1. A history of game spectatorship: from the arcade to *Twitch*

Nowadays, there is a glut of games that are available to play, whereas in previous decades, playing video games was more difficult. Similarly, the proliferation of video games has changed the possibilities for game spectatorship. One can distinguish four phases of video game spectatorship that corresponded to major shifts in the patterns of gaming (Borrowy, 2012; Taylor, 2012; Jin, 2010; Jin & Chee, 2008). The first two were ‘offline’ and cover the golden age of the arcades in the late 70s and early 80s prior and the mainstream popularization of the home console market through the *Atari 2600* and the *Nintendo Entertainment System* (Taylor, 2012). The third stage in the mid to late 90s is a period of transition that featured local network play in LAN cafes and large-scale LAN events but also the beginnings of internet play (Jin, 2010; Hutchins, 2008). It paved the way for the various forms of internet-based game spectatorship that began with replay files, machinima, *YouTube* let’s play videos (Lowood, 2006), up to early forms of ‘emulated’ live streaming and the transformation of *Justin.TV* to

Twitch. This development is multi-facetted and covers issues such as technological inventions, economics, politics and culturally significant video game titles and genres (Loftus & Loftus, 1983; Hemphill, 2005; Lowood, 2006; Jin & Chee 2008; Jin, 2010; Taylor, 2009; Lin & Sun 2011; Borrowy, 2012; Taylor, 2011, 2012; Witkowski, 2012). This section will focus on the connection between the co-evolution of gaming technologies and how they are tied to the practices of game spectatorship and player-viewer interactions.

2.1.1 In the gaming arcade

The period between 1978 to 1982 is often referred to as the golden age of the arcades because of their importance as a social meeting point for American youths (Loftus & Loftus, 1983; Lin & Sun 2011; Taylor, 2012). Video games required bulky and expensive arcade video game machines and most machines were built to run only a single type of game. Due to this technical and economical limitation, the video game arcade developed as a physical, social space and business. The video game arcade machines were played by players but there were also onlookers that watched the games or roamed the arcade (Lin & Sun, 2011). The appeal of the gaming arcade was the opportunity to play games while interacting with other young people in absence of parents and adults (Loftus & Loftus, 1983). The gameplay and communicative interaction were strongly influenced by the physical properties of arcade machines.

Figure 5 Video game arcade machine

Figure 5 shows[6] two video game arcade machines of the games *Pacman* (1980) and *Street Fighter 2* (1992). The machines were designed for one or two players that stand close to one another and close to the machine. Due to the size and orientation of the CRT-monitor, spectators needed to stand very close and were often looking over the shoulder of the other player. The shared physical space allowed for direct spoken interaction between players, and spectators (Lin & Sun, 2011). These interactions have the potential of being facilitative as well as disruptive for the gameplay (cf. Bowman et al., 2013). Participants may give verbal assistance

[6] Source: arcade-museum.com

to one another, taunt or heckle each other (Piirainen-Marsh & Tainio, 2009, 2014; Kappen et al., 2014). The negative or disruptive result of taunting has been described as "choking", where the player underperforms in the game under the social pressure (Lin & Sun, 2011; Bowman et al., 2013). However, if players beat the game or their opponent, the viewers join the celebration, which is perceived as very pleasurable and socially rewarding (Kappen et al., 2014).

In conclusion, the technological limitations of the arcades lead to co-located play in a public setting, but in turn this enabled social interaction in a shared environment. In the wake of the success of the gaming arcades, the TV industry attempted to adapt arcade gaming to a television format. The show *Starcade* was a hybrid between a quiz show and arcade video gaming competition. The broadcast enabled game spectatorship from the living room, yet it missed all the previously mentioned social elements. Lacking this appeal, the show was cancelled after its two-year run from 1982 to 1984 and it was also the last attempt by a commercial broadcaster to bring game competitions to American TV screens.

However, this failure of game broadcasting over cable television did not mean the end for all types of game spectatorship and interaction in the living room.

2.1.2 Game consoles and local co-play

Parallel to the developments of the gaming arcades, the home console market saw its first breakthrough with the *Atari 2600*. It sold over 30 million devices over its five-year live span from 1977-1982 and despite a temporary gaming market recession (1983 – 1985), the release of the *Nintendo Entertainment System* led to a popularization of affordable and consumer friendly video gaming systems (cf. Taylor, 2012). Consoles brought gaming inside the homes and moved it from the public spaces of the arcades to a more private setting. This changed the way players, spectators and technology interact. For example, players can co-operate in character planning, combat decision making or may take turns while trying to defeat a difficult game challenge (Gajadhar et al., 2009; Piirainen-Marsh, 2012). There is also a growing body of research dedicated to the study of play among co-located participants that share a gaming console (Maurer et al., 2015; Vellosso & Carter, 2016; Gomez Maureira & Verbeek, 2016; Tekin & Reeves, 2017).

Their works problematize the distinction between the prototypical player and a prototypical spectator and raise the question of the demarcation line between play, co-play and spectatorship. Gajadhar et al. (2009) argues that all participants taking part in the activity have the status of co-players. This would also include every spectator, who is talking about the game and in their sense, spectators 'play' the game.

This study agrees that play, co-play and spectatorship can transition gradually but it maintains that the control of the game (via the input device) is a necessary condition for play (cf. Eskelinen, 2001). Genuine (co-)play requires the player to give commands over the gaming hardware, wheras the discourse by the spectator does not rise to this level. Therefore, the study will not treat the viewers of live streaming as players and their discourse is not conceptualized as a form of play.

With this interpretation, play on consoles has historically been limited to two to four players that share a single television screen in a private home. This setup promoted games that were suitable for this context such as fighting games (*Street Fighter* in 1987), jump & runs (*Super Mario Bros* in 1985), racing games (*F-zero* in 1990) and soccer simulations (*FIFA International Soccer* in 1993). The games allow for quick transitions between player, co-player and spectator roles and they have a low barrier of entry for new players.

2.1.3 Networked gaming

The situation was different for desktop computers because they were always designed for a single user, who has persistent control of the gaming input devices. Instead of sharing a screen and controllers, multiplayer on computers evolved by connecting computers into networks. This made playing together significantly costlier and required more physical space. Although private 'LAN parties' existed, the cost and space requirements of networked gaming brought them back into the public spaces of internet cafés 1990's.

Figure 6 Starcraft in a Korean 'PC bang' and Counter-Strike in an American internet cafe

Internet cafes have a booth layout, which gives everyone a space to sit in front of their computer. The computers are the access point to a collectively shared game-space and it facilitates games that are enjoyable in larger groups and require an increased coordination among players such as First-Person Shooters, Real-Time Strategy games and Massively Multiplayer Online Role-Playing Games (cf. Taylor, 2011). There was a co-evolution of technological progress and game design, which lead to new gaming and spectator practices (Seo, 2013; Seo & Jung, 2014). Many gamers enjoyed the chance to compete in multiplayer games and created guilds or clans with like-minded players. In these clans or guilds, they trained their skill in the most popular video

game titles (Witkowski, 2012). Other players formed organization such as the *Cyberathlete Professional League* (1997), where the newly formed clans could compete for prize money. This segment of the gamer population transformed from serious and committed players to full-time, professional gamers that play in e-sport[7] tournaments. There are detailed accounts of this process of professionalization and the emergence of e-sport that cover the status of professional gaming as a sport (Hemphil, 2005; Wagner, 2006; Hamari & Sjöblom, 2017), the training and acquisition of gaming expertise (Rambusch et al., 2007; Witkowski, 2012), and the representation and participation of women in professional gaming (Taylor, 2009; Taylor, 2012).

Most significant for this project is the way competitive online gaming transformed game spectatorship (Hutchins, 2008; Jin, 2010; Taylor, 2012; Taylor, 2016). Gradually, video games were developed with spectatorship in mind and online broadcasting services developed to facilitate viewership.

Year	**Games, Leagues & Tournaments**	**Game Related Broadcast Technologies**
1997	Cyberathlete Professional League	
1998	Release: *Starcraft* (SC)	Offered online play but no form of online spectatorship.
1999	Release: *Counter-Strike* (CS)	
2000	1st Global Tournament: World Cyber Games	
2001		HLTV Streams (Live); Replay Systems for CS and SC
2002	Major League Gaming (MLG)	
2003		
2004	Release: *World of Warcraft*	
2005		Warcraftmovies.com & Machinima
2006		*YouTube* & Let's Play Videos (Video on Demand)
2007	Intel Extreme Masters by ESL	**Justin.TV*, precursor to *Twitch* (Live),
2008		
2009	Release: *League of Legends* (LoL)	Gaming.Justin.TV (Live)
2010		
2011	1st LoL World Championship	*Twitch* (Live)
2012	Release: Counter-Strike GO (CS:GO)	*Azubu.TV* (Live)
2013		*Hitbox* (Live)
2014	*The 4th International* tournament with a cash prize of over 10 Million USD	
2015		*YouTube Gaming* (Live)
2016	CS:GO "E-League"	
2017	Over 1 million concurrent viewers watch E-League finale of *CS:GO* on *Twitch*.	

Table 2 Milestones in e-sport and game broadcasting

[7] short for electronic sport

Table 2 shows a timeline of important game releases, gaming leagues & tournaments on the left; and online distribution platforms on the right. Important technologies that enabled game spectatorship often followed major releases of online games or the creation of gaming tournaments and organizations. Online game spectatorship evolved as a complex media ecology and increasingly facilitated the means of communication between players and spectators (cf. Taylor, 2012).

Replay files

Initially, the very successful games *Starcraft* and *Counter-Strike* did not allow any form of broadcasting and spectatorship. To watch games the spectators had to be physically present at the internet cafe or LAN tournament. Only in 2001, the first replay systems were developed. These replays were recordings of the game via the game software. Technically, they are not in video files such as *.avi or *.mkv files and were saved in proprietary formats. Their advantage is their relatively small file size (80 to 200 KBs), which allowed them to be shared despite the very low internet bandwidth of the time. As a downside, they had to be download from dedicated fan sites and could only be watched within the video game client of *Starcraft* or *Counter-Strike*. They did not contain any audio commentary or other ways of interaction between player and spectators, but they were the first form of internet-distributed and easily shareable game spectatorship.

'Machinima' and 'let's play videos'

Communication between player and viewer began with Machinima videos of *World of Warcraft* and later *YouTube* let's play videos. Released in 2004, *World of Warcraft* is a Massively Multiplayer Online Role-Playing Game, whose player base grew to 12 million people in late 2010, making it the most popular video game in the world at that time. The game created large guilds and communities that played together in PVP combat or in PVE[8] 'raids' versus very strong computer-controlled bosses (cf. Taylor, 2009; Chen 2009).

Machinima videos were edited montages that showed very skilled players in these PVP & PVE encounters. Players began to stylize their videos using cinematic techniques and music and successful video makers became celebrities in the World of Warcraft community (cf. Lowood, 2006). The distribution of these videos occurred mostly via the website *Warcraftmovies.com*, which was the most popular hosting site for *World of Warcraft* videos.
An important social aspect of *Warcraftmovies* was its commentary system. It allowed viewers

[8] PVP is short for player versus player and PVE stands for player versus environment

to write messages and comments to the video producer and player.

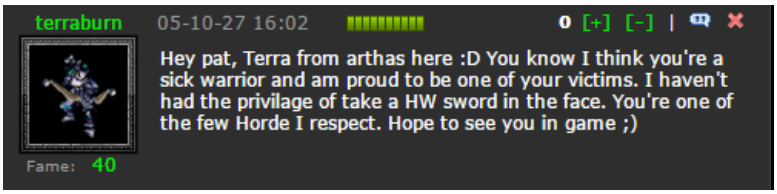


Figure 7 Warcraftmovies comment

The screenshot[9] in Figure 7 is an original comment from the year 2005 taken from one of the most watched machinima, called "High Warlord Tauren Warrior (insane crits)". It shows a comment from the user 'Terra' to the video producer 'Pat'. The viewer writes to the video maker and expresses his gratitude, his respect and that he enjoys playing against him. *Warcraftmovies* enabled the distribution of game videos and its written comment system made it possible to communicate to the player and video producer. Spoken communication is usually absent because machinima videos have a strong emphasis on music and do not feature commentary by the player.

This started to change with the emergence of let's play videos on *YouTube* (cf. Smith, et al. 2013). They reduced the prominence of music soundtracks and foregrounded the spoken communication of players (Nylund, 2015). The spoken discourse of let's play videos is usually recorded while gaming and it captures the experience of the player in the moment of play (cf. Nguyen, 2017). In my M.A thesis, I studied the discourse of let's play videos and I argued that the spoken communication has several directions of talk and social purposes (Recktenwald, 2014). Firstly, there is the narration of the unfolding gameplay and discussions about potential player actions. Secondly, let's players sometimes perform the role of their avatar and talk 'in-character' to the other non-player characters (cf. Gee, 2007; Nguyen, 2017). Lastly, let's players talk to their imagined audience. After the upload, the audience can write a message in the comment section on *YouTube*. Let's play videos offer digital distribution of the video game footage and a bi-directional communication between player and viewers. In this sense, the discourse in *YouTube* 'let's play videos' is very similar to live streaming, as it is a form of cross-modal communication between a speaking player and a spectating audience (cf. Dynel, 2014b; Smith et al., 2013). However, there is an important difference in the unfolding of the communication because *YouTube* videos are asynchronous. The spoken commentary is performed in anticipation of the audience reaction. Only after the editing and upload, can the audience comment – and this might incur a delay of days, weeks or even years.

Live streaming is a much more immediate interaction and requires consistent and

[9] taken from http://www.warcraftmovies.com/movieview.php?id=9199

uninterrupted bandwidth to stream in high resolutions (cf. Aparicio-Pardo et al., 2015; Pires & Simon, 2014), which was rarely available to private households during the mid 2000's.

2.1.4 'Emulated' live streaming

To circumvent this bandwidth restriction, audiences had to appropriate several different technologies to emulate a live streaming experience. In his keynote speech of the inaugural TwitchCon conference, the moderator "djWHEAT" outlined his perspective on the emergence of live streaming of video games and argued that "users moved towards live streaming without realizing it (TwitchCon fieldnotes, 2015)." The rise of live streaming is a step-wise development of actors that use ever-changing technologies. He illustrated this on the example of the spiritual predecessor to live streaming of *Counter-Strike*. It involved three separate technologies, which were used in conjunction. While it is not necessary to have a complete understanding of all the processes of this user-driven media convergence (cf. Jenkins, 2005; 2006), it illustrates that there was a demand for live interaction and live spectatorship of gaming before the existence of *Twitch*.

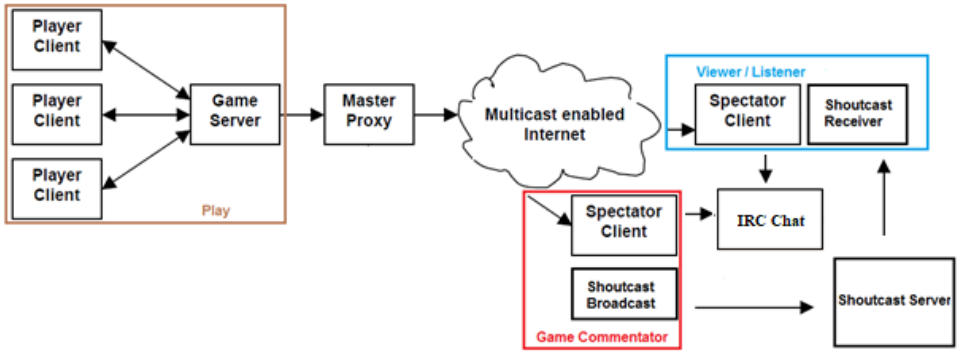


Figure 8 "Emulated" live streaming

Figure 8 shows 'emulated' live streaming and is a modification of Otten's (2001, p.7) '*HLTV* Architecture'. It demonstrates the combination of *Half-Life TV, Shoutcast Internet Radio* and *Internet-Relay Chat* to create an experience that is similar to live streaming. First, there is the game server that hosts the match of *Counter-Strike*. The players connect to the game server with their individual 'player clients' and all the gameplay occurs within this closed environment (green square). Their gameplay is sent to an external master proxy, which broadcasts the game signal over the internet. The broadcast is in a proprietary format and requires a game 'spectator client' of *Half-Life TV*. On the receiving end, the spectator client interprets the signal and shows the gameplay on the screen of the spectator (blue square). This transmission required very low bandwidth but there was no game commentary and neither was it possible for communication between participants (cf. replay files). The second technology was the audio streaming software *Shoutcast*, which allowed the game commentator to distribute his or her audio commentary via a dedicated *Shoutcast* server.

The audience could connect to the server and listen to the broadcast with their *Shoutcast* receiver. At this point the audience could see the game through their *HLTV* client and listen to the game commentator through *Shoutcast* internet radio. The third piece of software was *IRC* chat, where the audience and the game commentator could meet in a designated *IRC* chat channel and communicate with one another. Once *HLTV*, *Shoutcast* and *IRC* were set up and synchronized, the experience was very comparable to modern day live streaming. For the first time, it was possible to chat with other spectators, while watching the broadcast of a video game and listen to audio commentary by a game reporter.

The section highlights creative re-appropriations of technologies to create new media practices of gaming and spectatorship. There was a consistent demand for game spectatorship, even if there were significant barriers of entry. In the case of 'emulated' live streaming, it needed advanced technical knowledge to set up three separate programs for a combined purpose. There was no platform that brought the gaming content, spectatorship and the possibility for communication together in a way that was accessible to a general audience

2.1.5 From *Justin.TV* to *Twitch*

Justin.TV, the actual precursor to *Twitch*, suffered from the opposite problem. It was a live streaming website that offered means of communication and distribution but lacked original media content. *Justin.TV* was founded in 2007 by Justin Kan and Emmet Shear as one of the many new media experiments. Initially, it was a platform for Justin Kan to broadcast his life with a head-mounted webcam. In October of the same year, they opened the service for the general public and broadcasting of gaming was just one of many other categories next to sport, music, socializing or technology. By July 2008, more than one million users registered with the site, yet much of this growth in the early years has been associated with the broadcasting of pirated sport events of the National Football League or the Ultimate Fighting Championship. After a legal dispute with copy right owners, *Justin.TV* became a distribution platform without content and a stagnating audience.

The growth and transformation of *Justin.TV* to *Twitch* was not an immediate revolution. Instead, it started out rather modestly, at the intersection of two struggling parties with shared interests. Since the gaming section of *Justin.TV* was still demonstrating growth, Emmet Shear started to stir the company in this new direction. In 2009, they started a subdomain *gaming.Justin.TV* specifically dedicated to the broadcast of gaming. Its continued success led to the rebranding of the website to *Twitch* in 2011 and it became one of the first live streaming websites specifically dedicated to gaming. Other competitors such as *Azubu.TV*, *Hitbox.TV* and

most notably *YouTube Gaming* tried to enter the market but despite their efforts, *Twitch* was able to translate its first-mover advantage into an ongoing and dominant market position.

The history of video game spectatorship and communication started several decades ago and is still evolving today. The section demonstrated that players and viewers always utilize the technologies of their time. Each of them possessed unique properties that are employed for different play and communicative practices. The next section is specifically dedicated to *Twitch* and the interactional phenomena that have been discussed in prior research.

2.2 The study of *Twitch*

Research on *Twitch* is still very limited, but nevertheless it has already ventured in several directions. For example, computer scientists see live streaming as a technological challenge for data transfer (Kaytoue et al., 2012; Pires & Simon, 2014; Deng et al., 2015). In marketing, it has been studied as a platform for advertisement and promotion of products (Plath, 2015; Raes, 2015; Heuer, 2017). This section will discuss the research that is directly relevant to communication on *Twitch*. This includes *Twitch*'s self-understanding and mission statement, prior research on the appeal or motivations for watching games, and the existing descriptions on *Twitch*'s interaction.

2.2.1 Twitch's mission statement and its appeal

Twitch's mission statement is important because a company will support developments that are in accordance with its self-understanding.

Figure 9 Twitch - social video for gamers

> Twitch is "fun and represents a compelling new social network to connect with friends and fans over a shared love of games. Many broadcasters are making a living on Twitch based solely on how they entertain and interact with their audiences (Twitch, 2015; emphasis mine)."

Twitch sees itself primarily as a social experience, with gaming being an important facilitator for social relationships. These relationships are said to be between 'friends and fans', often enacted through 'entertainment and interaction' on the website. This positions the social discourse and communication as an important aspect of live streaming, next to gaming and entertainment.

These three themes of gaming, entertainment and social interaction are also dominant in the sociological and ethnographic research about *Twitch*. Within these academic disciplines, studies discussed the appeal of live streaming, the live streaming motivations and live streaming types. Although their work is related in content, most studies have their own focus and propose their own terminology:

- Cheung & Huang (2011) create a taxonomy of audience types
- Smith et. al (2013) suggest a taxonomy of live stream types
- Hamilton et al. (2014) describe key elements of the interaction
- Gandolfi (2016) presents orientations towards live streams
- Sjöblom & Hamari (2016) speak of audience motivations
- Karhulahti (2016) distinguishes streamer roles

Upon close reading of these works, it becomes possible to identify the three overarching categories that motivate the spectatorship of gaming according to *Twitch's* mission statement and the academic literature.

	Live stream types, appeals and motivations								
Author	**Ludic appeal**			**Social contact**			**Entertainment value**		
Cheung & Huang	Commentator	Curious	Pupil	Assistant	Crowd	Bystander	Entertained	Inspired	Unsatisfied
Gandolfi	Challenge-oriented spectacles			the Exchange			the Exhibition		
Smith et al.	E-sport	Speedrun		Let's Play					
Sjöblom & Hamari	Cognitive Motivations			Personal Integrative	Social Integrative		Affective Motivation	Tension Release	
Karhulahti	Impersonal Livestream	Personal Livestream							
Hamilton et al.				Informal Social Interaction			Shared experiences around ephemeral in-game events.		

Table 3 Summary of live stream typologies

Table 3 above groups the individual typologies evident in the literature into the three main aspects of ludic appeal, the social contact and the entertainment value.

Ludic appeal is a broad category that covers the game and gaming-related reasons to watch a broadcast on *Twitch*. Many spectators are very interested in improving their own gameplay. They take the role of 'pupils' that look up to the very skillful streamers as role models and gaming mentors. They have 'cognitive motivation' as they observe the livestream and ask

the streamer for tips and tricks (cf. Georgen et al., 2015; Kow & Young, 2013; Karhulahti, 2016). Another subset of the ludic appeal is the spectatorship of excellent play and greatness in a video game. Spectator types such as the 'curious' want to see innovative and well executed gameplay and appreciate the streamer's high performance. They are not necessarily interested in learning or emulating what they see, but instead they are attracted by watching the best players in the world during e-sport competitions or speed runs (Smith et al., 2013).

The second type of appeal is the social contact among participants. Cheung & Huang's (2011) 'the crowd' and 'bystander' describe spectators that enjoy the feeling of togetherness as they join to support their friends. Their motivation is the social connection with people that have the same hobby. Gandolfi (2016) describes the interaction in these type of livestreams as an exchange between streamer and viewers, which fosters social ties between the participants. Furthermore, he argues that they are "trying to build a bond beyond the game on the screen (2016, p.76)." This bond is usually conceptualized as a virtual community that is maintained through the interpersonal communication in the livestream (cf. Rheingold, 1993; Hamilton et al., 2014).

The third overarching category in the literature is the entertainment value. Watching live streams can relieve tensions from everyday life. Alternatively, viewers may have 'affective motivations', that is they gain enjoyment simply from watching the broadcast (cf. Sjoblöm & Hamari, 2016). They are the 'entertained' spectator, who prefers to watch the potentially stress-inducing gameplay. For this type of audience, live streams are similar to 'let's play' videos due to the appeal they offer as a form of 'vicarious play', where the streamer plays on the behalf of the audience (Smith et al, 2013; Glas, 2015). The streamer supplements the gameplay with humorous or exaggerated commentary and the livestream is a staged performance or exhibition (Nyland, 2015; Gandolfi, 2016; Nguyen, 2017)

The separation into three major appeals is very common across the studies. They present valuable insights that should match the intuitive understanding of researchers who are familiar with live streaming. In addition, they serve as an overall frame of reference that can support the interpretation process of my study. Nevertheless, there is some point of concern because the categories are derived from secondary data sources such as interviews, questionnaires or they are based on auto-ethnography and observations. Survey respondents, interviewees and auto-ethnographers orient towards desired norms and the information given is biased towards expected and accepted patterns (cf. Crown & Marlow, 1960). While the findings of these studies are very balanced and reasonable, they attest to a lack of data-driven studies about the livestream interactions. Furthermore, none of them tested how their proposed categories were

enacted. If there was an analysis, the methodology remained notably vague as in Gandolfi (2016), who writes “play sessions were analyzed by taking notes concerning the performance and building relationships with the audience (p.75)”. One may ask what was noted down, based on which criteria and what analytical frameworks guided the analysis of the data, if any?

Similar questions and concerns also extend to the research on *Twitch*’s communication. The next section will discuss seminal studies that have influenced the design of this research project and it will also point towards potential issues that arise and how they have been addressed in this study.

2.2.2 Communication on *Twitch*

Research on *Twitch* consistently points towards the important role of the streamer, the influence of the game and the effect of the audience size for the patterns of communication. The studies make interesting observation but often remain vague in the details. For example, Smith et al. (2013) propose the term “Player as Commentator (p.135)” to describe the streamer’s talk, but the idea is not fully explored and suffers from a very general description that reads “narrating what is happening, what parts of the game are good and what parts are bad, etc (Smith et al., 2013, p.135).” This description is broad and does not explain how the ‘narrating’ looks like. Similarly, Gandolfi (2016, p.77) describes the narration of the gameplay as “top-down”, without providing criteria or an illustrating example. Karhulahti’s (2016) study is grounded in Goffmanian frame analysis (1974) and he considers the narration of gameplay the essential component of his ‘play frame (competition)’, Karhulahti (2016) argues that the streamer enters a ‘play frame (competition)’ during a match and that this play frame is visible in the streamer’s physical as well as linguistic orientation towards the game. In this phase, streamers’ turn their head towards the monitor and their body torque is oriented towards mouse and keyboard. This signals the streamers’ involvement with the game. The talk consists of the aforementioned narration of the gameplay. What the contributions by Gandolfi (2016), Smith et al. (2013) and Karhulahti (2016) have in common is that they describe game narration as communication, which originates from the streamer and takes the “pertinent game situation (Gandolfi, 2016)” as its topic. However, they do not write about the features of these narrations.

There is disagreement on the effect of the game situations on the overall organization of discourse. For Smith et al. (2013), the player as commentator is in control of the communication and is only slightly affected by the “constraints of the game (p.135)”, which are not further defined. The players alternate relatively freely between ‘narrating’, which is more monological and talk about “tactics & strategies (Smith et al. 2013, p.133)”, which is more dialogical with the audience. For Karhulahti (2016), the type of communication depends on the status of the

game because he strictly aligns the 'play frame (competition)' to the duration of a match or round. If the competition of the match is ongoing, the player remains within the play frame, which he associates with narration. Only after the round does the streamer enter the 'interview frame (conversation)', where they turn to the chat, read the messages and respond to questions. Considering both positions, it remains unclear if the unfolding of the game is just a minor 'constraint' on an otherwise 'free' discourse, as suggested by Smith et al. (2013), or if the game has a significant impact on the organization of talk (Karhulahti, 2016). Both studies present their observations based on auto-ethnography, the researcher's intuitions and background knowledge, which makes their contrasting positions difficult to assess.

This study uses the idea of narration as a type of talk about game and game events and further develops it in the transcription scheme for live streaming on *Twitch* (see 2.4 & 4.2). This will allow a closer look at the narration segments of live streaming and reveal that it is more accurate to distinguish narration into two separate linguistic practices, which will be referred to as commenting and reporting (6.2). They have different linguistic features and a different temporal relationship to the game events. This indicates that the organization of the game is impacting the organization of discourse, but not necessarily in a static and binary fashion, such as in-game or out of game (cf. Karhulathi, 2016).

Aside from the description of the streamer's game 'narrations', previous studies also discussed the dialogical communication between streamer and audience. They argue that the size of the audience may affect the communicative patterns (Hamilton et al., 2014; Nematzadeh et al., 2016). The chat in small (<10) and medium-sized (~150 viewers) streams produces meaningful interaction between participants and the communication is labeled as a 'conversation' or an 'informal social interaction' (Hamilton, et al. 2014, p.1315). In large-sized chats the interaction is said to change and:

> The chat becomes a source of breakdowns. It transforms from a meaningful medium of discussion into an illegible waterfall of text, scrolling up the page so quickly that it cannot be read. Participants can no longer follow the conversation. [..] (it) changes to something like the roar of the stadium (Hamilton et al. 2014, p. 1319; emphasize mine)

Hamilton et al. (2014) hypothesize that information overload is a major reason for this transformation because there are too many messages for a meaningful discussion (cf. Jones et al., 2004). Hamilton et al. (2014) apply a prescriptive and normative position to the *Twitch* chat. They frame orderly conversation as a desirable standard for communication, which they set in opposition to the 'illegible waterfall of text'. In response to Hamilton et al.'s (2014) research,

corpus linguistic studies investigated the written chat of medium-sized and large-sized streams (cf. Olejniczak, 2015, Ford et al., 2017, Nematzadeh et al. 2016). These studies found that messages in the larger chat rooms are:

- shorter with regards to word count
- lower in lexical density
- shorter in message uptime
- higher in the number of emoticons

Based on these findings, the studies draw similar conclusions to Hamilton et al. (2014) and agree in their negative assessment of large-sized streams.

Author	**Large-sized Streams**	**Medium-Size Streams**
Hamilton et al.	Roar of the Crowd	Meaningful medium of discussion; informal social interaction
Olejniczak	'Cheering' crowds	Meaningful exchange
Nematzadeh et al.	Cacophony	Conversation

Table 4 Comparison of large and medium-sized streams

The table summarizes the common descriptions for large-sized and medium-sized online live streams. Nematzadeh et al. (2016) describe *Twitch*'s chat in large channels as a 'cacophony' and Olejniczak (2015) argues that they have "little to no time for meaningful exchange of thoughts (p.332)."

There are three points of criticism against these arguments. First, it is problematic to take isolated linguistic features, such as word count or lexical density, and conclude that the discourse in large-sized channels is less orderly and less meaningful. Hereby, Hamilton et al. (2014), Nematzadeh et al. (2016), and Olejniczak (2015) rely on an ad-hoc and common-sense definition of the term conversation as 'ordered talk' but they do not describe necessary conversational features. Thereby it is difficult to measure or asses *Twitch*'s orderliness and explain why large-sized channels are an 'illegible waterfall of text' that do not meet the conversational standard.

Secondly, they break down Twitch's communication into a binary opposition with the audience size as the single determining factor. The communication is either an orderly conversation or chaotic and this pattern of discourse is assumed to never change throughout the broadcast.

Third, these studies on *Twitch's* chat have systematically excluded the communication of the streamer. By ignoring the most salient party, it becomes difficult to evaluate them as either orderly conversation or a different communication type (cf. Cheung, 2017). Therefore, it

must be said that the early research on *Twitch* treated discourse as a list of separate items and discussed them in isolation of one another.

The dissertation agrees that conversation and commentary are suitable starting points, but any description of *Twitch*'s discourse must look at the audience, the streamer and the unfolding gameplay in conjunction. Moreover, it is necessary to have a proper linguistic foundation, which provides criteria for judging communication as a conversation, commentary or cacophony. These criteria must be applied to and tested on authentic data. The result will be a much more detailed and accurate description of the organization of the discourse of live streaming.

To get to this point, the dissertation takes several steps. In 2.3, it will present the *Twitch*'s website and position it as the communicative environment of the computer-mediated communication. The section will show the different communicative affordances that are available to the streamer and the viewers. The streamer's communication is mostly impacted by the unfolding of the game, whereas the chat communication is influenced by the number of chat participants. This influence of the audience size plays an important role in the selection of the recorded livestreams (4.1) and there will be a distinction between medium-sized and large-sized live streams. To study the game's impact on the communication, section 2.4 will introduce video games as dynamic systems of rules, goals and levels and the interpretation of games as system informs the development of the transcription scheme (see 4.2). Next to the transcription of spoken language and written chat, the annotation of game events will be a core feature of the transcript. This transcript is designed to capture the interaction in an authentic and yet analysable format. To understand some of the transcription choices, it is important to have a solid understanding of the communicative environment of Twitch. Its layout and features enable the interaction between the participants. The next section will introduce this process by discussing *Twitch* from the perspective of its affordances.

2.3 The communicative environment of *Twitch* and its affordances

The term "affordances" is widely used in many academic disciplines such as psychology, design studies, sociology, as well as communication and media studies. In each of these disciplines there is a slightly different application of the concept, but they all derive their interpretation from common point of origin.

Gibson (1979) first coined the term affordances to describe the relationship between animals and their habitat. He argued that animals do not perceive the environment as such, but rather through its affordances. By this Gibson (1979) means the possible actions that an object

provides to its user. An illustrarive example is the case of the bee hive. A bee hive is a man-made structure, which the honey bee uses as its nest and honey production hub. For the bee, they provide protection from harsh weather, an easily defendable entrance against possible intruders and sufficient space as a breeding ground for their colony. Honey bees perceive bee hives as ideal housing for their nests. In other words, the bee hive "affords" this type of use to the honey bee. For people, a bee hive has a very different set of affordances, as it is an easily transportable bee nest that can be driven around to different fields. They have removable frames, which house the honey combs, which are harvestable by the bee keeper. The bee hive is more akin to a form of live stock and a farming technique. From Gibson's (1979) perspective this means that affordances are more than a material property, but also a relationship between the material object and its users. Importantly, Gibson's affordancs were a static concept as they do not change. An object has fixed properties, which enable or afford a fixed number of practical actions.

Norman (1988), a researcher in design studies, built on Gibson's (1979) concept of affordances but rejected its inflexibility. He described affordances as "perceived and actual properties of the thing, primarily those fundamental properties that determine just how the thing could possibly be used' (1988: 9). In his tradition, affordances are variable and they may change over time as the needs of the user changes. Historical examples are knives or axes, which have been tools for cutting meat or chopping lumber but also have been used as weapons during times of conflict.

Norman (1988) saw the concept of affordances as valuable in order to distinguish efficient from poor design. In a well-designed object, the material properties give visual cues for its intended use and they are particularly suitable for this action. The affordances are designed into the object as the designer has a certain use in mind. Hence, the blade of a butter knife is wide and flat to spread a topping, whereas the blade of a butterfly knife is sharp and pointy to penetrate organic tissue. What Gibson's (1979) and Norman's (1988) affordances have in common is that they mostly apply to physical objects, which are used for specific material actions. They have a 'realist' perspective, which means that tools have actual and measurable properties that enable clearly-defined use cases.

Other researchers approach affordances from the perspective of social constructivism (cf. Grint & Woolgar, 1997). In this tradition, affordances come into existence as an object that is used by a person for a particular purpose. The butter knife becomes a butter knife since it is used to spread butter. In court, a butter knife may be "the murder weapon" and thus its status as an object, its uses and affordances are socially (re-)mediated. While this focus on usage, at the

expense of material properties, is useful to prevent too deterministic views, it has also weakened the analytical usefulness of the concept of affordances. Since any object can be used in an almost infinitve number of ways, for an almost infinite number of social purposes, the description of its affordances is an exercise in imagination. The realist and social constructivist position present two extremes in their description of affordances.

The dissertation will take a point of view that is in-between these two opposits, which was proposed by Hutchby (2001; 2014) in his discussion of communicative affordances. He argues that the affordances of a given technology are always both functional and relational (p.448). They are functional in the sense that their material properties are enabling and constraining the use of the technology by an actor (realist position). But they are also relational in as far as they afford different uses to different actors in different situations (constructivist position). This third way can explain why a tool may have i) an obvious primary purpose, ii) is usable in a variety of different occasions and iii) or may never be used for other tasks. Affordances are inside or outside of the "perceiptual range" of actors (2014, p. 87). The less suitable a technology appears for a given purpose, the less likely it is used.

For this dissertation, the range of affordances can be described a priori, by looking at the features of the technology, but this description must consider how the technology is commonly used. Therefore, Twitch's affordances cannot be summarized to a finite list of items or features, but rather the description must also explain how the participants typically use the platform for their communicative goals.

Streamers on *Twitch* have their own unique channel website, which is freely accessible at www.twitch.tv/username. This study takes the streamers' individual website as the environment within which the communication takes place. Thereby, it excludes additional layers of communication that might occur through external programs such as *Skype* or *Discord*. The following figure is a schematic representation of *Twitch'*s channel website.

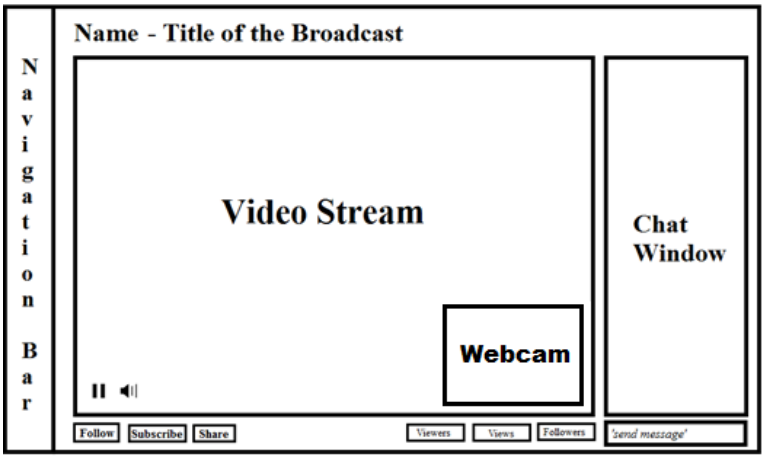


Figure 10 Schematic representation of Twitch's website

It shows the environment of the communication and enables an analytical deconstruction. On the left, there is a navigation bar, which allows users to browse the website and look for different games and channels. Once viewers have selected the livestream of their choice, the navigation bar is of little consequence. The same is also true for the top section of the page, where the streamer's username and the title for the current broadcast are displayed.

On the right is *Twitch's* chat window for the audience. There, the audience can communicate with written chat messages and website specific emoji. There is no upper limit on the number of chat participants and some channels can have thousands of viewers, which leads to a very active and quickly moving chat. In the center of the web page is the video stream. The video stream contains the video game footage as well as the embedded web cam recording of the streamer.

Figure 11 Enlarged screenshot of 'webcam' recording.

The figure shows an enlarged and cropped example taken from the web cam section. The web cam is usually mounted on top of the streamer's primary monitor and records the broadcaster's upper torso, face and his or her spoken communication. Parts of their arms, hands and lower body are rarely visible. For the viewers, *Twitch* has a very low barrier of entry. Within a single browser window, they can simultaneously see the game, hear the streamer's spoken language and write in the chat. For the broadcasters, the interaction is more difficult because of their pre-occupation with the game. They have to control the game input devices with their hands, which prevents them from writing many chat messages. Moreover, they must constantly look at the game and it is difficult for them to see the chat.

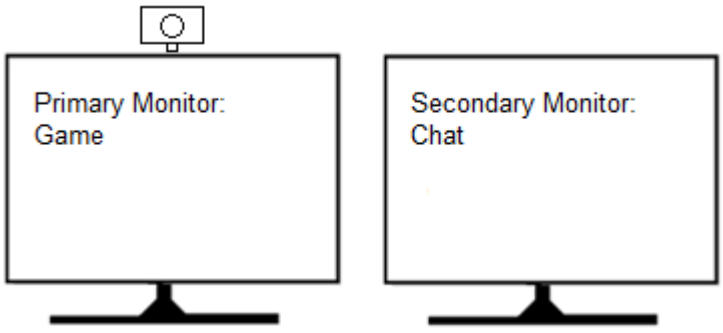


Figure 12 Dual monitor Setup

The figure shows a common solution to these shortcomings in affordances – the dual monitor setup. The primary monitor is positioned in front of the streamer and it displays the game. The

secondary monitor is positioned nearby and requires a slight body or gaze re-orientation.

Figure 13 Interactional re-orientation towards chat.

This behavior becomes visible through the web cam footage. While looking at the game, the streamer's gaze is straight forward with a slight downward angle. When they are reading the chat, they turn their gaze away from the primary screen towards the secondary monitor. On the second monitor, they are watching their own broadcast and reading the incoming chat messages by the audience.

This physical and technological configuration has significant interactional consequences. There is a delay between the streamer uploading the video signal from his or her computer to the *Twitch* server; and there is another delay from the *Twitch* server before it sends the broadcast to the viewers. It takes around 8 to 12 seconds for the complete transmission, which means that the audience and streamer are always slightly out of sync in their interaction. As the streamer is looking to the chat monitor, s/he sees the broadcast and chat from several seconds ago. As a result, they can never respond to the most current message and are restricted to 'older chat'. This split in timelines will become very visible in the transcript of the interaction (see chapter 4).

Another interactional affordance of the dual monitor setup is that the streamer cannot look at the game and the chat monitor at the same time. They must decide where to prioritize their gaze. As they look to the game, the chat messages are outside of their field of vision and cannot be read. If streamers decide to look at the chat, it is difficult to play succesfully. Therefore, they have to negotiate the two roles of streamer-as-player and streamer-as-caster.

In turn, this means it can be very difficult for audience members to be perceived by the streamer and it has also led to the development of third-party software that enables paid messages that are displayed directly within the video livestream.

Figure 14 Schematic representation of on-screen alert message

Figure 14 above shows the typical location of such so called 'alert messages'. These messages are triggered by a paid channel subscription or direct 'donations' from a viewer to the streamer. Thereby, they circumvent the problem of low visibility for the audience. So far, very few studies mention donation alert messages (cf. Raes, 2014; Sjoblöm & Hamari, 2016; Heuer, 2017), although the phenomenon has become wide-spread and is also common on *YouTube Live* and *Twitter*'s *Periscope*. However, as a communicative practice, they have not yet received any academic attention. Chapter 7 is specifically dedicated to these alert messages and will argue that they are a privileged communicative access to the streamer enabled by a monetary transaction. This makes them highly relevant for the overall organization of the discourse.

The following table is a succinct summary of the discussion in this section, showing the main differences in affordances between the broadcaster and their audience within the communicative environment of live streaming.

Participant	Visual Field	Interaction with the Game	Communicative Resources	Number
Broadcaster	Game or Chat	Play	Mostly Spoken Language	Individual
Audience	Game and Chat	Spectatorship	In Chat: Messages and Emoji On Stream: Donation Alert Messages	Unlimited

Table 5 Streamer and viewer asymmetries

Most of the points described in Table 5 are generalizable to a variety of live streaming types or other live streaming websites and are not unique to broadcasting on *Twitch*.

The next section will explain the relevant dimensions of games and play that are crucial for studying the discourse on *Twitch*.

2.4 Games and play

Video games are a very salient part of the broadcast and although many people have an intuitive knowledge about games, even students enrolled in game studies programs are often unable to

analyze them in a critical manner. They confuse being able to play a game with their comprehension of that game (Zagal, 2008). The study of video games has notable and early precursors such as Avedon & Sutton-Smith's (1971) edited volume *The Study of Games,* Huizinga's (1938) *Homo Ludens* and Caillois's (1961) *Man, Play and Games.* Huizinga (1938, p.13), for example, defines play as:

> free activity standing quite consciously outside "ordinary" life as being "not serious", but at the same time absorbing the player intensely and utterly. It is an activity connected with no material interest, and no profit can be gained by it. It proceeds within its own proper boundaries of time and space according to fixed rules and in an orderly manner (own emphasis).

From Huizinga's (1938) perspective, play occurs within in a separate sphere outside of ordinary life, which he calls 'the magic circle (1938, p.10)'. It separates what is inside (play) and outside (ordinary life) and it structures the organization of play 'according to fixed rules and in an orderly manner'. Huizinga (1938) describes the magic circle as the defining feature that structures the play. The magic circle is comparable to a willing adherence to games and a promise to play according to its rules and limitations.

This early perspective is reasonable to start a first discussion about games, but it is not specific enough for modern day video games and their broadcast on *Twitch.* This requires a more fine-grained explanation of the individual parts of video games, which has been provided by formalist approaches to game studies and their analyses of games. A formalist approach studies "a game independent of context, that is, without regarding which specific people are playing a specific instance of the game (Lankoski & Björg, 2015, p.23)." Within this paradigm, games are complex and dynamic structures that frame the play of the players. Games are understood as "game-systems" or "state machines" that operate on an input-output logic (cf. Järvinen, 2003; Juul, 2004; Björk & Holopainen, 2003). They have an initial state, which receives a play-input by the player and produces a new output based on internal calculations. This output then becomes the new game state for the next steps of the 'game machine'.

Figure 15 Games and play

Figure 15 is a reprint of Salen & Zimmerman (2003, p.85) and describes play as an integral part of games and games as the frame within which the play operates. Games limit the possible play-inputs and at the same time, the play-inputs are the engine of the interactive game system (cf. Calleja, 2011). This relationship between play-actions and game reactions 'engages', 'involves' or 'immerses' the player with the game (McMahan, 2003; Calleja, 2007; Lankoski & Björg, 2011; Henricks, 2015). For discourse analytical purposes, it is useful to further dissect this game system into its relevant properties that interact with one another and influence the actions of the player. They include the games components[10], rules & procedures, goals, challenges & rewards and the segmentation of games in time and levels. This section will explain these categories with examples taken from soccer, chess and poker because these games are commonly known and ideal to illustrate the concepts. In the analysis chapters, the study will demonstrate how the game properties influence the discourse of live streaming.

Components

Components (Eskelinen, 2001) or units (Bogost 2010) are the parts of the game that players can manipulate via play. The most central component of many games is the avatar, i.e. the virtual body that the players control to interact within the game world.

Rules and procedures

Rules and procedures are the core of any game system. Huizinga (1938, p.11) says that "all play has its rules" and Parlett (1999) reformulates this position to the extreme and argues, "every game is its rules, for they are what define it (Parlett 1999, p.3)." Admittedly, these definitions are very narrow, but they highlight the importance of rules for games. Rules license certain actions and prohibit others and thereby structure the play. Carter et al., (2015), Consalvo (2005),

[10] There is a significant variation in terminology for very similar concepts. Alternative overviews are provided by Carter (2015), Juul (2003) and Salen and Zimmerman (2006)

Järvinen (2003), Juul (2003) discussed different types of rules which can be summarized in a typology of game rules.

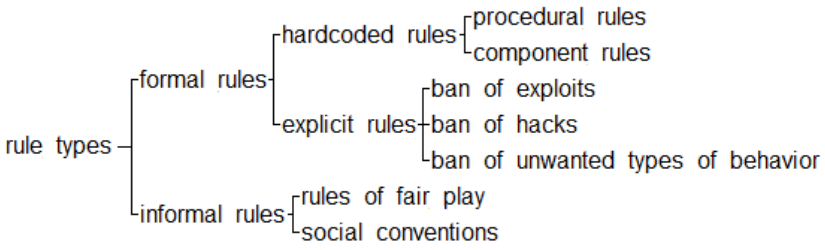


Figure 16 Typology of rules

A typology of relationships between different types of rules in games can be represented as in Figure 16 above. Carter (2015, p.1) begins with a basic distinction between formal rules and informal rules. The formal rules are produced by the game developer and come in two types. One type are the 'hardcoded rules' that are programmed in the video game software and these are further divided into 'procedural rules' and 'component rules' (Juul, 2003; Järvinen, 2003). Procedural rules define the actions that are possible in each situation – an example of this would be the mandatory turn-taking between players during chess. It is a procedural rule that each player has to move before it is the opponents turn. Component rules, by contrast, attribute values to components that are controlled by the player. So, for example the pawn in chess has a much lower value than the king, since it is the loss of the king that ends the game.

Often, component rules and procedural rules interact or can be combined in advantageous ways as 'game mechanics' (Järvinen, 2003). A pawn sacrifice is a game mechanic that relies on the compulsion to move. A player uses a low valued pawn (component rule) to force the opponent to make a disadvantageous move based on the procedural rule. The goal-oriented application of such game mechanics is a strategy (Juul, 2002).

Component Rule & Procedural Rule => Game Mechanic
Game Mechanic & Goal-Oriented Application => Strategy.

Therefore, game strategies are the means to pursue the game goals within the limits of the hardcoded rules.

The second type of formal rules are the 'explicit rules' which are set up to ward against unwanted player practices such as misbehavior, exploits or hacks. Misbehavior are all forms of tricking, lying or otherwise abusing other players to gain advantages or to cheat them. Exploits are abusive manipulations of the game in ways that lead to "glitches". A glitch is where the game does not behave according to the design of the procedural rules and produces outcomes

that are unintended. Hacks are comparable to exploits, but additionally involve a third-party program, which directly interferes with the game software.

Then there are the unwritten informal rules. They comprise the ideas of fair play, which are often less clear, much more context dependent and open to negotiation between players and game designers (Carter, 2015; Consalvo, 2005; Doherty et. al., 2014; Kücklich 2007; Moeller et al., 2009). Examples are 'ganking', which is short for gang killing a single victim with a large group of players; 'cheesing', the use of unconventional but effective tricks; and camping, the repetitive killing of a player from an advantageous position. Informal rules are community norms and are less binding than the official rules. Breaching them is often seen as unsportsmanlike and unfair but may not be punished by the game developers. Being a fair player means following the official as well as informal rules of a game (Wirman, 2007).

Goals, challenges and rewards

The goal of a (fair) player is to defeat the game's challenge within the confinement of the rules and receive the reward of the game. There are optional goals, which are possible but not necessary; in contrast to mandatory goals that are required in order to progress in the game (Juul, 2007, 2010). Most games contain a mixture of optional and mandatory goals and after a challenge is defeated, a new challenge arises in the next level or round (Zagal et al., 2008). It demonstrates that rules, goals, challenges and rewards are closely tied to one another and to the segmentation of games.

Segmentation of games, their time and pacing

There are several competing models that try to explain the segmentation of games and game time (Elverdam & Aarseth 2007; Eskelinen, 2001; Hitchen & Tychsen, 2009; Juul, 2004; Zagal et al., 2008; Zagal & Mateas, 2007, 2015). This leads to considerable variety in time categories and terminology. For this project, the most important categories are:

1) Cycles, durations and rounds

2) Game pace, haste and tempo

Temporal cycles, durations and rounds are an interaction of procedural rules and the game's goal. In poker or in chess, the procedural rule of turn taking leads to an alternation and cycling between the players until the goal of the games is reached and the round is over. Soccer has a different configuration of rules and goals, because a match has a fixed time limit (procedural rule) and the team that scores more goals during the duration of the match is the

winner. Cycles, durations and rounds are the temporal segmentation of games and they structure the temporal organization of play.

Within each round, the unfolding of game time can be described in terms of game pace. "Game pace is a subjective measure of how fast the game feels to the player (Hitchens & Tychsen 2009, p.12)." This feeling of pace can be deliberately designed and balanced (Hitchens & Tychsen 2009; Davies 2009; Yildirim, 2016). For example, a game can have 'haste', which means that with any passing of real-world time the game state changes (Elverdam & Aarseth, 2007). Many popular online games cannot be paused and even if a player does not input any commands, it remains possible for his or her opponents. This imposes time pressure on the players to act rather than to wait.

Tempo describes the number of simultaneous game events or player actions (Davies, 2009). A game can have haste but still possess a manageable tempo. An example is blitz chess, which has more haste and time pressure than normal chess. However, it still has the same tempo since players can only move one piece per turn. Other games operate in real-time and have no restriction on the number of moves. They allow the input of several commands in quick succession and increase a game's tempo. High tempo and haste lead to the perception of a high game pace, which has significant consequences on the approach players can take and the skills they need (Lewis et al., 2011; Caplar et al., 2013; Castaño et al., 2015). While slow-paced games allow for in-depth planning, fast paced games require quick decision-making and rapid inputs of commands (Kow & Young, 2013). Depending on the game's pace, this can be a significant interactional demand on the streamer-as-player, which is likely to influence the communicative behavior of the streamer-as-caster. Therefore, game pace is an important criterion for the selection of the games and the study will describe the temporal organization of *World of Warcraft's* Arena PVP, *League of Legends* and *FIFA* in section 4.1. The three games have a discernably different pace which affects the cross-modal exchanges between streamer and audience (see 6.1) as well as instances of commenting & reporting (see 6.2).

2.5 Chapter summary

The second chapter began with the historical developments that have led to the emergence of live streaming of video games. In 2.2, it summarized the existing literature on the motivations and reasons for game spectatorship and presented its three overarching appeals. It also argued that these appeals correspond to different patterns of communicative interaction between participants. For example, social contact can be achieved through an informal social conversation between streamer and viewers.

The section also presented other conceptualizations of the discourse of *Twitch* as narration, commentary, cacophony and an illegible waterfall of text. It pointed out common and shared description among previous authors, disagreements among their opinions as well as my criticism towards their arguments. Their ideas are very useful but often remain too vague or rely on secondary data sources. To describe the organization of discourse, the dissertation will connect to their contributions and improve upon the outlined issues. This began with the introduction to the communicative environment of *Twitch* in 2.3. and continued with the more detailed discussion of video games in 2.4. This will enable a new transcription process in the fourth chapter and will make live streaming accessible to micro-level linguistic research. Once transcribed, this data can be analyzed with a robust and yet adjustable discourse analytical framework.

Chapter 3- Linguistic literature review

3.0 Initial remarks

The role of this chapter is to provide a comprehensive overview of the relevant linguistic literature and explain how it informed the methodology and analysis of the thesis. This will begin in 3.1 with the introduction of the Birmingham school of Discourse Analysis (Sinclair & Coulthard, 1975) as the analytical framework. So far, this framework has only been applied to spoken conversation. This means that there is a significant gap between its traditional use and the current online live streaming data that I intend to apply it to. Therefore, the section will also present contemporary studies in online chat and video-mediated communication. These studies relied on the Conversation Analysis (Sacks et al., 1974) and Goffman's (1981) work on participation frameworks. The section will point out the strengths of these two approaches and demonstrate that despite their differences to the Birmingham School, they have an overlap in the units of analysis. Thereby, they inform my reading and application of the Birmingham School.

The unit of analysis and language description of the Birmingham School are the discursive move and the exchange between speakers. Moves and exchanges can be studied sequentially on the micro-level to reveal how participants interact with one another but also how discourse is produced in reference to the non-linguistic context. This makes the framework and the units of analysis ideal for the study of online live streaming with its dialogue between streamer and viewers and more monological segments of narration or commentary about the video game.

Section 3.2 and 3.3 will summarize prior research on the properties of spoken face-to-face conversation and written chat in *IRC* and the sections will discuss their organization as well as their usual content. This review serves as a point of reference to compare the discourse of live streaming. The two sections will begin to highlight the importance of the linguistic mode of communication for the patterns of interaction within a conversation. This theme will be further explored in 3.4.

It will cover video-mediated communication, for example *Skype* or *Google Hangouts* and communication on a video platform such as *YouTube*. Research in this area is a rather recent phenomenon (Rosenbaun et al, 2016b). 'Video chat' allows for cross-modal communication between a speaking and a writing party, which makes it in many respects comparable to live streaming. However, the section will also present important differences between video chat and live streaming as participants in online video chat can "mode-switsch" from speech to writing and vice-versa (Sindoni, 2014). In live streaming, the access to the communicative mode is

fixed according to the role of the participant as streamer or viewer. While viewers are free to chat, streamers have to negotiate their talk with video game play.

Aside from the sequential study of discourse, research on video-mediated communication has studied how speech and writing are used to address different participants for different purposes. For example, in online video chat, participants use spoken communication to address an interlocutor as part of the main discussion and turn to written communication for less significant topics with another person (Sindoni, 2014). This indicates that the choice of mode and the direction of communication between participants are important features in the organization of discourse.

This aspect can be studied by taking a 'synoptic view' on the communication (Eggins & Slade, 1997). The synoptic view assesses the distribution of discursive moves between participants and their usual functions (see 3.1, 4.3 & chapter 5). It complements the sequential analysis because it can evaluate the prevalence of different orientations of communication from, for example, streamer to viewer, viewer to streamer and among viewers. It is particularly useful in the case of live streaming because it is unknown how common dialogical exchanges or monological narration are. The chapter will continue with the explanation of how Eggins & Slade's (1997) synoptic view can be used to create a coding of discursive moves for *Twitch* live streams. This coding has been applied to the transcript of the data, which consists of live streams of varying sizes (large / medium) and different game paces (slow / medium / fast). In chapter 5, the distribution of moves between participants will be analyzed to evaluate if audience-size and game pace influence the organization of discourse.

The sixth chapter will build on this analysis of the discursive moves and describe how moves are combined into larger units of discourse during live streaming. Moves and exchanges are the main unit of analysis and they are described through the discourse analytical toolkit of the study.

3.1 Discourse analytical tools

The analytical framework for this dissertation is the Birmingham School of Discourse Analysis (Sinclair & Coulthard 1975) but its application is also informed by recent research that has used Conversation Analysis (Sacks et al., 1974) and the work of Goffman (1981). These research traditions share as a major premise that communication is an interactional achievement between participants. This idea influences all other aspects such as data sources, unit and level of analysis and the considerations of the context of situation (Malinowski, 1923)

This research is committed to the study of naturally occurring discourse based on recordings of authentic interactions. It is qualitative research that objects to the exclusive study

of invented sentences (cf. Chomsky, 1965; Searle, 1969). Constructed and idealized scenarios are representations of the researcher's judgments about language and do not account for the linguistic reality (cf. Goodwin & Heritage, 1990). Sentences never occur in isolation and are always part of wider communication or discourse. Discourse is a process, where meanings are produced and exchanged by participants with communicative goals. Participants orient towards one another and apply practical reasoning to the communication and treat utterances as meaningful in a given situation (Garfinkel, 1967; Grice, 1968). The context of this language production and its interpretation is similarly emergent and cannot be reduced to a fixed list of extra-linguistic features (Goodwin & Heritage, 1990). Context goes beyond the social and spatial setting and includes other elements that participants can make relevant for the ongoing interaction (Goodwin & Duranti, 1992; Auer 1996). This remains true for live streaming with its rapid changes in the non-linguistic context. It means that the selected linguistic tools must be able to describe the emergent discourse between participants under consideration of the unfolding game. This section will introduce the Birmingham school and explains why it is suitable for this task once one considers the additional insights provided by Conversation Analysis and the recent applications of Goffman (1981).

These traditions were originally designed to describe and analyze the organization of spoken discourse in the class room (Sinclair & Coulthard, 1975), on the telephone (Sacks et al., 1974) or in ordinary conversations (Goffman, 1981). Their shared observations are that spoken discourse is produced in distinct units of talk that unfold over time and involve the alternation of speaker and listener roles. Goffman's work), Conversation Analysis and the Birmingham School have separately developed analytical categories to describe these patterns. The dissertation will describe the organization of discourse through the terminology of the Birmingham school because it provides a very explicit theorization of the relationship between different 'ranks' of discourse and because it suggests a clear link to the organization of the non-linguistic context.

The next subsection 3.1.1 will illustrate the Birmingham School theory and explain the categories using simple examples. Afterwards, section 3.1.2 will demonstrate the near-equivalence of their terminology with concepts in Conversation Analysis and in the work of Goffman (1981).

3.1.1 The Birmingham School of discourse analysis

The Birmingham school describes the organization of discourse across three horizontally connected levels and vertically integrated ranks.

Levels and ranks

Non-linguistic organization	*Discourse*	*Grammar*
course		
period	LESSON	
topic	TRANSACTION	
	EXCHANGE	
	MOVE	sentence
	ACT	clause
		group
		word
		morpheme

Figure 17 Levels & ranks of discourse

The figure shows the relationship of ranks and levels for the example of classroom discourse. The three levels are grammar, discourse and the non-linguistic organization of the activity. In order to provide a proper introduction, this section will explain all three, but the thesis will only use the levels of discourse and non-linguistic organization. Questions of grammar are outside of the research focus because they are less relevant for the organization of larger units of interaction.

A complete class consists of the spoken discourse of a 'lesson' and its non-linguistic organization in a 'period'. The lesson consists of the content of the class room discussion and is embedded in the temporal organization of a 45- or 60-minute period. This means that there is a correspondence or connection between the levels of discourse and the level of non-linguistic organization. Similarly, there is a connection between the organization of discourse and the organization of grammar.

Example (1)

```
Beijing is the capital of the People's Republic of China
and its second largest city.
```

The utterance can be described on the level of grammar as a complete sentence and on the level of discourse as a move with the speech function of a statement (cf. Eggins & Slade, 1997). Moving vertically, each level has a rank scale and the higher ranks are constituted by lower level units. The example sentence consists of two clauses and when looking at it as a move it consists of two acts.

Moves are the basic unit of communication (cf. Halliday, 1984), but not all communication

is a dialogue or equally dialogic. Communication can also be more monologic and consist of a series of discursive moves by a single speaker. Monologues can address a listening audience and appear like a dialogue (Bell, 1984; Frobenius, 2014), but a requisite for proper conversation is an exchange between speakers. The exchange is the "basic unit of interaction (Sinclair & Coulthard, 1975)" between participants and a minimal dialogue consists of at least one exchange. In their original formulation, Sinclair & Coulthard (1975) argued that every exchange consists of exactly three slots named initiation, response and follow-up, which are filled by exactly three moves.

Example (2)

I: What is the capital of China?

R: Beijing.

F: Right.

In the example, an eliciting move serves as the initiation for the exchange. After the initiation, there is a speaker shift to the interlocuter, who produces an elliptical statement "Beijing ~~is the capital of China~~" that fills the response slot and is the answer to the question. After the response, there is a second speaker shift back to the initiator, who produces an affirmation "'right" into the follow-up F slot.

An exchange is a sequential structure of slots, whose elements are defined by their position in the exchange. Importantly, the slots of an exchange do not possess their own speech functions. The speech functions are realized by the moves that fill the slot. An exchange can be initiated with an elicitation as in example 2 but also with an attending move such as "hey John" (cf. Eggins & Slade, 1997).

Depending on the type of move that is put in a slot, the exchange has a certain 'prospection' (Sinclair, 1992). Prospection is a form of trajectory for a dialogical exchange and it limits the number of appropriate reactions. An initiation filled by an elicitation prospects a responding statement, whereas an initiation that is filled by an attending move prospects a responding greeting. Therefore, it is possible to speak of an exchange structure E = I → R→ F, which describes a sequential order of slots and a pattern of expectation.

The model was developed for classroom data and proofed to be very useful for this orderly type of spoken discourse between teacher and student. However, as it was applied to other settings, it proved to be too narrow and could not account for the diverse forms of communication. The prescriptive stance of the model that each exchange slot with exactly one move has been replaced by more descriptive and open-ended revisions (Coulthard & Montgomery, 1981). Exchange slots can be filled by several succeeding moves as shown in the

next example:

Example (3):

I: `What is the capital of China`

R: `I am not sure.`
`It could be Beijing`

F: `Good guess.`

It demonstrates that the distinction between exchange slots and its move is not trivial. In this example, the response is filled by two moves to form a statement. Only afterwards, there is a speaker-shift and follow-up. Moreover, there are exchanges without a follow-up, if the response sufficiently reacts to the initiation. It means that the follow-up slot is an optional element, which is indicated by rounded brackets in the exchange formula E = I → R → (F).

The last revision of the original framework accounted for the fact that there can be resets in the exchange sequence (Coulthard & Brazil, 1992; Sinclair, 1992).

I: `Do you want something to drink?`

R/I: `Do you have lemonade?`

R: `Yes, I have a few bottles.`

F: `Then I'll take one.`

The elicitation "`do you have lemonade`" fills the response slot but is interpreted as a second initiation. Effectively, the element re-initiates the I → R → (F) sequence[11] and therefore receives a double labeling R/I. In spoken dialogue, exchanges can be described as E = I→ (R/I) → R→ (F), whereby the I and R slots are mandatory elements and R/I[12] and F are optional (Coulthard & Brazil, 1992). The more flexible approach to exchange structures allowed for a wider application of the model to other settings of everyday conversation or spoken interaction (Francis & Hunston, 1992; Stenström, 1994).

The concepts of moves, exchanges and prospection will be the most important tools for the description of discourse of live streaming but at this point it is very important to highlight a difference in understanding between Sinclair & Coulthard (1975) and my application of their work. They consider the formula E = I→ (R/I) → R→ (F) as an abstraction that can explain all spoken exchanges. An improvement of the formula would mean to find exceptions to their rule and propose a new formula that can explain all patterns of exchanges. This study will not attempt to propose a 'catch-all' formula, which is at the risk of becoming too general and lose its analytical grip and descriptive force. Instead, it uses the exchange structures as descriptive

[11] Within Conversation Analysis, such restarts are known as side-sequences (Jefferson, 1972)

[12] Response / Initiations have also been called challenge moves (C) in Burton (1980) and Warren (2006). This study will prefer R/I because the term challenge is closer to a function of moves than a description of a slot.

tools that can highlight the form of specific discursive practices. In turn, I will look at the variations of the exchanges and the reasons for these variations in order to describe the organization of discourse.

Above the level of exchanges are transactions. Transactions are a complete discussion of a topic and consist of several exchanges. Very little is known about their internal structure and attempts to formalize them have had limited success (cf. Sinclair & Coulthard, 1975). The flexible and emergent nature of spoken discourse is likely to make it impossible to ever produce an adequate description of all possible variations. The only certainty is that they are a "topic-unit" that is tied together by the shared theme across its series of exchanges (Francis & Hunston, 1992). Despite their decreasing analytical grip, transactions remain a valuable concept because they can be used to describe the content of communication or patterns of communication that are related to the content (see 6.3)

The highest unit of discourse organizations in the class room is the 'lesson', which is a series of transactions limited in time by the 45- or 60-minute 'period'. Similar to transactions, the highest ranks of discourse have a lot of variation and their exact description is not possible. However, it can be said that the organization of the lesson is attuned to the non-linguistic organization of the period and that the sum of all periods constitutes the course (Coulthard & Montgomery, 1981).

Most studies within the Birmingham School were dedicated to exclusively spoken, face-to-face encounters. Other forms of conversation, especially those in a computer-mediated setting, used the research paradigms of Conversation Analysis or built upon the linguistic contributions by Goffman (1981). The next section will introduce these works and demonstrate that their basic premises are very similar to the Birmingham school. Therefore, this research on computer-mediated communication has many implications for the Birmingham school and its application to the live streaming on *Twitch*.

3.1.2 Conversation Analysis and participation frameworks

Conversation Analysis and Goffman (1981) have categories that are very similar to the Birmingham School's notions of move and exchange. CA's turns and adjacency pair are related concepts.

Utterance	Conversation Analysis	Birmingham School DA
What is the capital of China?	Turn Position: First-pair part	Move (Elicitation) Position: Initiation
	Floor Transfer	Speaker shift
Beijing	Turn Position: Second-pair part.	Move (Statement of Fact) Position: Response
	Adjacency pair	Exchange

Table 6 Comparison between CA and Birmingham School categories

Speakers take turns in a dialogue and the current speaker holds the conversational floor, while other participants are listening (Sacks et al., 1974). For Conversation Analysis, listening is an active process and interlocutors observe the speech to find points for a possible floor transfer.

In the example above, the production of a first-pair part clearly marks the point of floor transfer to the listener. The first-pair part creates 'conditional relevance' for the second speaker to produce a turn that relates to the question (Schegloff, 1968; 2006). Second-pair parts are predicted or demanded by the first-pair part. In the terminology of the Birmingham school, first-pair parts have prospection and an adjacency pair consists of a predicting first-pair part and a predicted second-pair part. This makes them conceptual equivalents to exchanges with initiations and responses. The differences between turns, moves and exchanges arises from the ethnomethodological roots of Conversation Analysis. It sees communication as a procedural 'speech-exchange system' (Sacks et al., 1974, p.719) and it has a more mechanistic view of language, which is less interested in the study of grammar, speech functions and the role of contexts. Turns and adjacency pairs have no functional description, which is a central quality of moves and exchanges (cf. Kerbrat-Orrechioni, 2004).

However, Conversation Analysis provides additional contributions in other aspects of discourse description as it has been applied to many communicative settings and has been more rigorous in its descriptions of the conversational floor. My understanding of the Birmingham School has been refined by the contributions of Conversation Analysis (see 3.2), which will enable an in-depth analysis of the sequential unfolding of the discourse of live streaming.

The third relevant research strand for this study is the work of Goffman (1981). Goffman (1981) criticized the dyadic bias of early conversational models and divided talk into a 'production format' of speakers and a 'participation framework' of listeners. Levinson (1988) suggested a revision of the terms into 'production roles' and 'reception roles' because they are more symmetrical, and he used 'participation framework' to discuss the intersection of

production roles with reception roles. The study will adopt this change in terminology as it is more intuitive and does not interfere with Goffman's (1981) original descriptions. The production roles are the animator, author and principal. The principal is the origin of the opinions that are expressed. The author puts the opinions into a message and the animator gives the message to the recipients. In many situations, these three roles are realized by the same person, but exceptions are also very common as for example in quotations or reported speech (cf. Holt, 1996). The following figure is a reprint from Kerbrat-Orecchioni (1990) and illustrates Goffman's reception roles.

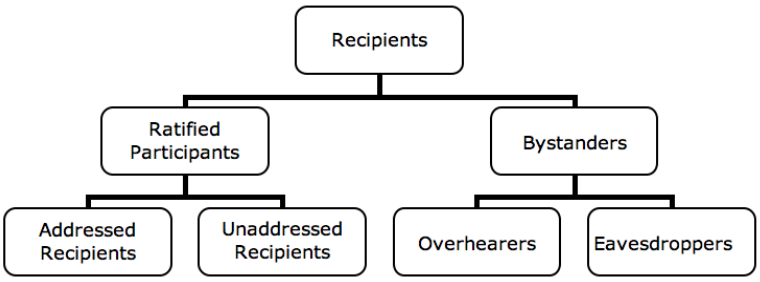


Figure 18 Goffman's 'reception roles'

The recipients consist of ratified participants, which includes the directly addressed recipients and the unaddressed recipients. The directly addressed recipients are the primary target of the message and they might be addressed via their name, a pronoun or a clearly marked physical orientation (Goffman, 1976). Unaddressed recipients also influence the communication because the producer is acutely aware of their presence (Clark & Carlson, 1982; Bell, 1984). Bystanders or unratified participants are not known to the speaker and have no effect on the wording of the message.

The participation framework of production and reception roles has seen meticulous refinement by Clark & Carlson (1982) and Levinson (1988) and while their work is widely acknowledged and cited, many studies tend to go back to Goffman (1981) for practical reasons because Levinson's (1988) categories are difficult to measure or operationalize. For example, Levinson (1988, p.173) introduces a 'ghostor', which is a ghost writer who is co-present with the animator and he contrasts it with the 'formulator', who is a non-present ghost writer. While such distinctions are often interesting for linguistic theory, they often remain impractical or overly detailed.

Participation Frameworks are congruent with a description of discursive moves within the Birmingham School. As discursive moves are produced, they also assign speaker and listener roles. The main differences between participation frameworks and moves are in the terminology that is applied, and the descriptive angle taken. Goffman (1981) focuses on the participants,

who situate themselves in speaker and listener roles via dialogue. The Birmingham school focuses on moves in a dialogue that situate the speaker and listener. The implication for my application of the Birmingham school is that discursive moves are not just defined by their position in an exchange and their speech function, but also by the interlocutors of the message. Depending on the configuration of participants, the same message may realize different speech functions and elicit different types of reactions. At the end of section 3.4, the dissertation will describe how this insight can be used to study talk and chat between streamer and viewers and how it will be helpful in the description of the organization of discourse.

Beforehand, it is necessary to introduce the features of spoken (3.2), written (3.3) and cross-modal conversation since conversation is the most common description for the discourse of live streaming (Hamilton et al., 2014; Olejniczak, 2015; Gandolfi, 2016; Nematzadeh et al., 2016).

3.2 Spoken conversation

Conversation is a well-researched and established term in linguistics and can serve as a frame of reference for the analysis and discussion of live streaming. However, there are also many competing terms such as chat, talk or dialogue that are often treated interchangeably and there is a general lack in consistency in wordings (Kerbart-Orrechioni, 2004). This is problematic because micro-level linguistic study requires academic accuracy in their description of communication. Therefore, the section will start with a review of the most common terminology and explain how they are used in this research project. The most general category is the interaction, which can be linguistic and non-linguistic. Interaction consists of an action that is performed and an orientation towards a person or an object (cf. Sacks et al., 1974). For example, streamers interact with their audience via communication, but they also interact with the game through play (cf. Newman, 2002; Lankoski, 2011)

'Communication' refers to any use of the semiotic system of language in its spoken or written mode. If the communication occurs in the spoken mode, then the dissertation will use the terms 'talk' and 'speech'. Non-verbal communication such as gestures lie outside of the study, because they are unavailable to the audience, difficult to perform for the streamer and complicated to combine with the unit of analysis, i.e. moves and exchanges.

Dialogue has a common association with talk between exactly two participants (cf. Halliday, 1984). The Greek prefix *dia-* is often wrongly interpreted as 'two' although it means 'through' (Dynel, 2014a). Dialogue is first and foremost a discourse pattern and for communication to be dialogical, it must consist of proper exchanges and floor transfers (cf.

Sinclair & Coulthard, 1975; Matthiessen & Slade, 2011). Dialogue is not restricted to the spoken mode and can happen between any number of participants. Dialogue between exactly two participants is described as 'dyadic' (Gumperz & Hymes, 1972) and if there are more than two participants the phrase multi-party dialogue will be used. Kerbrat-Orrechioni (2004) suggest the terms duologue and polylogue, which this dissertation considers an unnecessary coinage (cf. Dynel, 2014a).

Finally, conversation is a spontaneous dialogue that is typically in the spoken mode (cf. Halliday, 1984). It focuses on the expression of interpersonal meanings between participants (Eggins & Slade, 1997; Warren 2006; Matthiessen & Slade, 2011). Its written equivalent is chat, which is colloquially also used to refer to spoken conversation, but in academic contexts mostly refers to written online chat (Rintel & Pittam, 1997). The following table provides an overview of the common terminology and how it will be used in this dissertation.

<table>
<tr><th>Label</th><th>Linguistic Mode</th><th>No. of Participants</th><th>Interaction via</th><th>Content</th></tr>
<tr><td>Interaction</td><td rowspan="6">Spoken
or
Written</td><td rowspan="3">Unspecified</td><td>Language & Actions</td><td rowspan="8">Unspecified</td></tr>
<tr><td>Communication</td><td>Language</td></tr>
<tr><td>Dialogue</td><td>Language</td></tr>
<tr><td>Dyad / dyadic
<s>Duologue</s></td><td>2</td><td>Language</td></tr>
<tr><td>Multi-party dialogue
<s>Polylogue</s></td><td>>2</td><td>Language</td></tr>
<tr><td>Talk / Speech</td><td>Spoken</td><td>Unspecified</td><td>Language</td></tr>
<tr><td>Conversation</td><td>Mostly Spoken</td><td>Unspecified</td><td>Language</td><td>Socially-oriented</td></tr>
<tr><td>Chat</td><td>Mostly Written</td><td>Unspecified</td><td>Language</td><td>Socially-oriented</td></tr>
</table>

Table 7 Forms of interaction and conversation

In linguistic tradition, conversation has a been described as "the primordial site of human interaction (Goodwin, 2000)." Its main purpose is the social contact and it has no pragmatic goals or motivations and is "talk for the sake of talk (Eggins & Slade, 1997, p.21)." Conversation has a free organization, which emerges dynamically and is not tied to the level of 'non-linguistic organization'. Studies that apply the principles of the Birmingham school to the study of spoken conversation tend to exclude the non-linguistic context from their descriptions and analysis (Francis & Hunston, 1992; Stenström, 1994; Eggins & Slade, 1997; Warren, 2006; Matthiessen & Slade, 2011). This dissertation considers this a missed opportunity. By excluding the systematic study of the context, it becomes impossible to relate stages of discourse to stages of non-linguistic organization (cf. Winn, 2015). It raises risks because the assumed free organization of conversation might be the result of a selective transcription or a lack of

consideration for the context. Therefore, the dissertation concludes from this point forward that even the study of conversation must have a consistent annotation of context in order to fully utilize the Birmingham School and evaluate the impact of contextual factors.

Returning to the discoursal organization of conversation, most studies identify two dominant patterns. The first pattern is 'narrative chunks' or narrations (Tannen, 1987; Eggins & Slade, 1997; Norrik, 2000). These chunks are stretches of talk are produced by a single speaker, who is consistently holding the floor. They are a series of moves without prospection and there is little to no turn-taking with other parties. Therefore, narration can be categorized as more monological chains of moves with an overarching topic (see transaction) and no turn taking with other parties. (cf. Matthiessen & Slade, 2011). This formal definition of narration is still congruent with the description of gameplay 'narrations' discussed in chapter 2 (Karhulahti, 2016). Therefore, it is possible to study the narration of gameplay by looking at more monological stretches of talk or chat among the participants. A major difference between conversational narration and gameplay narration is the content and social purpose. Conversational stories share parts of the teller's inner life to create a better mutual understanding, social bonding and a decrease in social distance (cf. Coupland & Jaworski, 2003; Matthiessen, 2014). This cannot be assumed for gameplay narrations and their content and purpose have to be analyzed from the data.

The second pattern typical of conversation are segments of talk with very frequent speaker shifts. This frequent and less orderly floor transfer co-occurs with rapid shifts in topics (Tannen, 1987; Stenström, 1994; Matthiessen & Slade, 2011). Such segments contain dialogical exchanges with a clear prospection between moves. The content of the exchanges attends to social roles, is often humorous and contains markers of friendship and solidarity among participants. Therefore, the study will use the term conversational only in reference to exchanges that express interpersonal meanings, yet will fall back to the term dialogical to describe exchanges about other contents such as tactics or strategies.

Within the Birmingham School, there are three major typologies of moves that discuss 'casual conversation' (Eggins & Sladen, 1997), 'everyday conversation' (Francis & Hunston, 1992) or 'spoken interaction' (Stenström, 1994). This research sees them as mutually complementary because none of them claim to be the definitive descriptions of spoken discourse. Their systems are flexible and;

> "new acts, moves and exchanges can be added as the need arises, so long as such new additions are sufficiently generalizable and so long as the basic theoretical principles of Sinclair and Coulthard (1975) are adhered to (Francis & Hunston 1992, p.156)."

There is a lot of agreement and overlap in their typologies and an utterance such as "`Hi John`" would be considered an 'attending move' in Eggins & Slade (1997) or a 'greeting' in Francis & Hunston (1992). Each of them also describes unique moves that are not mentioned in the other typologies. Examples are 'apologies' in Stenström (1994), concluding moves that summarize a topic in Francis & Hunston (1992) or detaching moves that serve as an exit to the conversation (Eggins & Slade, 1997). Instead of inventing a completely new arsenal of moves it is much more reasonable to build on existing descriptions. The combination of their typologies is sufficient to describe most cases of face-to-face conversation[13]. The study will only coin new moves for the description of novel phenomena (see 'topicalizer' in 6.1).

The discursive organization of live streaming can be studied with the analytical tools of the Birmingham School by comparing live streaming to the well-established descriptions of spoken conversation. Studies on spoken conversation provide suitable reference points for dialogical talk and narration and will be used in 6.1 for the discussion of exchanges and in 6.2 for the analysis of commenting and reporting about gameplay. The dissertation maintains that it is necessary to explain the organization of discourse in reference to the 'non-linguistic organization' of live streaming. This will occur throughout chapter 6 but also specifically in 6.3. The section will demonstrate that stages of live streaming lead to re-occurring developments in the discourse. Whenever possible, the description of spoken moves by Francis & Hunston (1992), Stenström (1994) and Eggins & Slade (1997) will be applied.

However, these main works have been developed for exclusively spoken conversation in face-to-face settings. Live streaming is greatly different (see section 2.3), and its study has to consider the features of internet chat conversation (3.3) as well as cross-modal, video-mediated communication (3.4). Therefore, the following two sections will introduce internet chat and video-mediated communication and elaborate on their implications for the study.

3.3 Internet chat

The term chat colloquially refers to informal speech and it has been used to describe the loosely ordered segments of talk in conversation (Eggins & Slade, 1997). In research on computer-mediated communication, chat refers to written conversation that emerged with internet relay chat (*IRC*) and instant messaging software in the mid to late 1990s (Rintel & Pittam, 1997).

[13] The complete list of moves and speech functions is listed in the appendix.

The written chat is the main mode of communication for the audience on *Twitch* and interestingly *Twitch*'s chat system still relies on *IRC* technology. *IRC* chat is a technological predecessor, which makes its prior study highly relevant for the discussion of live streaming. Prior research covered the aspects of lexis, grammar, the social purpose of chat and the organization of its turn-taking. (Cherny, 1995; Crystal, 2005; Werry 1996; Paollilo, 2001). This section will only cover findings that are directly relevant to this research project and suggests Rheingold (1993), Walther (1996), Werry (1996), Garcia & Baker Jacobs (2010), Rintel & Pittam (1997), Crystal (2005; 2008; 2011); Herring (1999); Simpson, (2005), Paolillio (2001), Paolillio & Zelenkauskaite (2013) for a comprehensive discussion of all major phenomena.

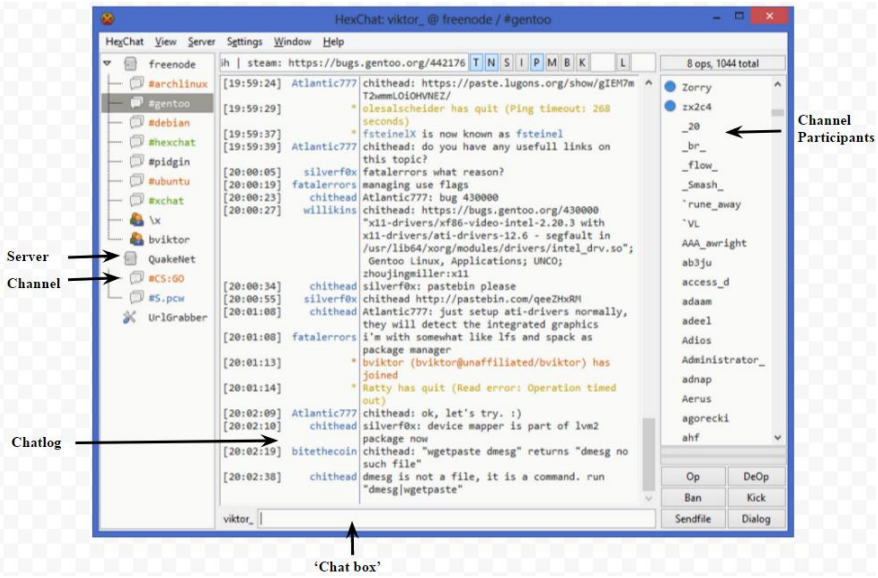


Figure 19 Annotated screenshot of internet relay chat

The figure shows an annotated screenshot of internet relay chat and is a reprint from the official website https://hexchat.github.io/screenshots.html. *IRC* chat was released in 1988 and uses the client-server principle. The users connect their clients to one of the many publicly open servers and then join a 'chat channel' with an overarching theme, which brings the participants together. In the figure the server is 'Quakenet' and the channel is 'CS:GO', a sequel to *Counter-Strike*, the popular online game already mentioned in the second chapter.

Twitch hosts its own server with the address *irc.chat.twitch.tv*. When audience members visit the website Twitch.tv with their web browser and start watching the stream, they automatically log in to *Twitch*'s *IRC* server and join the streamer's chat channel. The chat is accessed through the browser but handled in the background by *IRC* servers and technology.

Therefore, *Twitch*'s chat shares many of its affordances. In contrast to face-to-face communication, chat is not limited to its immediate spatial surrounding. It enables a global gathering of participants in a shared chat channel. Chat produces 'virtual communities' and facilitates communication about niche activities with the most notable example being the live

streaming of video games (cf. Rheingold, 1993; Paollilo, 2001). *IRC* research considered chat communication as self-contained within the channel (cf. Jones, 2004). All communication occurs in writing, between the participants of the chat channel and free from an outside context. On *Twitch*, it is unclear how much communication occurs among channel participants and how much of its communication is addressed at the live streamer or in reference to the game. The description of the organization of discourse of live streaming has to consider the prevalence of intra-chat communication in comparison to cross-modal communication. This is another reason to study the distribution of moves between streamer and viewers, because it can account for their preferred directions of communication from streamer to viewer, viewer to streamer, oramong viewers (see 4.3 & chapter 5).

Regarding the content of chat communication, research on *IRC* has shown that the exchange of socially-oriented meanings dominate, with participants writing about their lives and hobbies in humorous or creative ways. Chat communication has a 'ludic vein' and is full of word play (Simpson, 2005). This finding is echoed in the literature on live streaming (cf. Hamilton et al, 2014; Gandolfi, 2016).

Differences between spoken and chat conversation are predominantly attributed to the modal affordances of writing vis-à-vis speech. These differences are understood in terms of inequalities and inadequacies. Depending on the author and the aspect that is discussed, speech is considered superior or inferior to writing (cf. Kress, 2005; 2009; 2010). For example, Crystal (2001) and Walther (1996) see online chat as an impoverished form of communication and argue that it is inferior for expressing social cues because it lacks visual or auditory signals such as facial expressions, intonation and pronunciation. Such negative evaluations persist even in contemporary discussions on *Twitch*'s chat. Hamilton et al. (2014) describe the chat as a 'cold medium' in reference to McLuhan (1964). This research takes the position that chat communicates social meanings with different practices, which are not intrinsically inferior to their embodied or face-to-face counterparts. In chat, linguistic resources such as lexis, spelling, grammar and punctuation are used to emulate the features and meanings of spoken conversation (Crystal, 2008). The most notable feature of chat communication is emoticons[14], a portmonnaie term for emotion + icon. They are punctuation marks that signify facial expressions and convey the emotional undertone of a message (Thompson & Filik, 2016). Emoticons and Emoji are very common in *Twitch* chats and it is important to represent them accurately in the transcript

[14] Emoticons such as :=) are not to be confused with emoji, which are pictograms that cannot be subdivided into smaller punctuation marks.

to highlight the audience's reactions to game events or the streamer's talk. Emoji make *Twitch* chats extremely lively and the opposite of a cold medium. It will be demonstrated that they are very important for analyzing 'commenting' (see 6.2) and that they, furthermore, play an important role in the distribution of communication between participants (see chapter 5).

The last relevant topic from the discussion of web chat are its mechanisms for turn-taking and how they are different between spoken conversation and chat. Speech is vulnerable to parallel production of talk. If there are several simultaneous speakers, it becomes indiscernible noise. Speech production is easily observable in face-to-face communication because it is produced sound by sound. This leads to mitigating communicative practices that minimize disturbances and ensure a smooth conversation, i.e. exchanges with pauses and speaker shifts.

Writing in chat rooms is produced letter by letter but it is sent in the form of messages consisting of only a few words or sentences. On the upside, this means that chat systems support a higher number of participants because they do not impact each other's message production. Only as the message is sent, does it become visible to other readers. As a downside, the readers of *IRC* chat are mostly unaware of the message production process[15]. This makes turn negotiation very difficult and leads to "disrupted-turn adjacency (Herring, 1999)". Initiations and responses that belong together as part of an exchange become separated by messages from third parties. In larger chat rooms, this is the norm and exchanges are intertwined with one another. They develop very rarely in a clearly distinguishable series. Moreover, new chat messages take the position of older messages. Older chat messages are moved outside of the chat window and become invisible. As a result, there is less facilitative coordination and more competition between participants.

This has severe implications for the concept of floor in CMC environments. Taking a turn, in other words, producing moves, is insufficient for floor control (Edelsky, 1981; Herring, 2010). The floor is an interactional and collaborative achievement and a message requires ratification by a reader or interlocutor to become the floor (Simpson, 2005; Herring, 2010). This ratification of moves is often taken for granted in spoken conversation, but it cannot be assumed for multi-party chat or live streaming on *Twitch*. The exchange structure E = I→ (R/I) → R→ (F), which is typical for dyadic spoken exchanges, is very likely to be different in live streaming. In order to describe the organization of discourse, one has to study the exchange structures that are actually realized, because these will demonstrate how the discourse is successfully tied together. Moreover, it becomes possible to properly define the 'breakdown' of communication

[15] Newer chat software such as WhatsApp or Facebook Messenger indicate message production, for example via oscillating ' … '. This feature is not present in IRC chat or *Twitch* Chat.

that was suggested by Hamilton et al (2014). A breakdown describes a move that serves as an initiation and which possesses prospection but never receives an appropriate response by another party. They are failed attempts at gaining the conversational floor, and so exemplify unsuccessful communication. The study will discuss their prevalence in chapter 5 and will discuss their form in terms of "incomplete exchanges" in section 6.1. Thereby, the study will challenge the negative label of 'waterfall of text' (Hamilton et al, 2014) and provide a much more accurate description of the phenomenon.

The streamer's spoken discourse and the viewer's written chat are the main forms of communication. So far, talk and chat have been introduced in their separate environments of face-to-face conversation and chat channels. The communicative setting of live streaming brings both modes of communication together (see section 2.3). This leads to interactional practices that cannot be explained by studying speech and writing in isolation. Therefore, the next section introduces research on 'video-mediated communication' (VMC)[16].

3.4 Video-mediated communication

Studies on video-mediated communication are still rare and are a very recent phenomenon in linguistics (Frobenius, 2014; Dynel, 2014b; Sindoni, 2014; Rintel, 2015; Rosenbaun et al. 2016a, b; Rosenbaun & Licoppe, 2017). To the best of the author's knowledge, all of these are focused on 'conversational' types of video-mediated communication such as *Skype* calls, *Google Hangout* video chats, *Camfrog* video chat or *YouTube* video blogs. The conversation is the only form of interaction and the participants are not engaged in another activity.

Research in video-mediated communication is concerned with the simultaneous and overlapping use of written and spoken conversation and its novel communicative practices. Video chat software allows participants to choose their mode of communication. This leads to the development of a written and a spoken conversational floor. Participants actively navigate both floors as they participate in the conversation. Sindoni (2014) describes "mode-switching" as transitions from writing to speech and vice versa. Participants switch to speech to make more salient contributions to the conversation and switch back to writing when they type less significant information. Speech is an upgrade in participation status and participants seek the spoken conversational floor to gain more attention (Sindoni, 2014; Rosenbaun et al., 2016a). In

[16] The dissertation uses the term 'video-mediated communication' because the reviewed studies have used it in self-reference and it is becoming its own research genre. However, the term is neither superior nor more accurate than CMC, especially since the chat during VMC is not actually video-mediated. Both terms, CMC and VMC, suffer from similar shortcomings for describing online communication (cf. Jucker & Dürscheid, 2012). The coinage of a new term is often a political decision and an artificial demarcation in academia.

turn, interlocuters treat speaking participants as the center of the interaction and give more attention to their talk. Writing is used to downgrade the participation status and participants switch from speech to writing if they want to be less involved in the discussion (Rosenbaun et al., 2016b). Written messages are also more likely to be ignored, or they are less likely to receive overt responses. Similarly, in *YouTube* video blogs, the spoken utterances of the video maker are very likely to elicit several written responses in the comment section, but most written comments will not receive any follow-up (Bou-Franch et al., 2014).

I interpret these results in terms of Kendon's (1992) distinction between the front stage and backstage of a conversation and I want to argue that the spoken floor is treated as the front stage and the written floor has the role of the backstage of the conversation in online video chat. Central participants are far more likely to use the spoken mode because it is much more salient, and participants attribute it with having more 'authority' in the conversation (cf. Sindoni, 2014).

In contrast to video chat, live streaming has a fixed configuration of modes and the streamer has a monopoly on the spoken floor, with the audience competing in the written chat. This means the findings from VMC research cannot be transferred directly but they provide important implications. The description of the organization of discourse should consider the interaction between the written and spoken floors. It requires a transcript that distinguishes the spoken communication from the written chat and it motivates the multi-column format of the transcript in this study (see 4.2). This will make it possible to observe whether there is a preference for communication in a certain direction, that is, do viewers prefer to communicate with the streamer or with one another. Interactional preferences and successful exchanges indicate a dominant position in the discourse, which are very relevant to its overall organization. Moreover, it will be important to describe how the discourse transitions between the spoken and written floors.

Exchanges between participants that communicate in different modes will be referred to as cross-modal exchanges (cf. Rosenbaun et al., 2016a). For example, a written message might serve as an initiation and receive a spoken response, or a streamer's utterance can initiate a discussion and receive written responses. Importantly, prior research treated cross-modal communication as a single phenomenon and did not distinguish between cross-modal communication from speech-to-writing or from writing-to-speech. This study highlights the recurring need to distinguish between these two directions of communication. They occur in a different frequency (see chapter 5) and are realized in distinguishable exchange patterns (see 6.1, 6.2). In spoken-to-written communication, the audience can take a turn and briefly hold the floor. In written-to-spoken communication, streamers 'give' the spoken conversational floor to

the chatting audience as they attend to their questions. This 'turn-giving' will become visible in the organization of the cross-modal exchanges (see 6.1). Without a spoken response, it is difficult for written initiations to hold the conversational floor. Similarly, spoken responses depend on the interaction of the streamer with the game. The descriptions of the exchange structures will present new findings such as the most common patterns of interaction between streamer and audience, as well as a new discursive move. This discussion will be able to explain most of the cross-modal exchanges that occur on *Twitch*.

The only exception are alert messages. Alert messages are a form of cross-modal communication through a dedicated and unique communicative pathway (see chapter 7). These messages require the payment of money, they appear in the center of the live stream and receive preferential interactional treatment by the streamer. The chapter will position them as a form of privileged cross-modal communication, whose occurrence significantly alters the unfolding of the discourse.

The analysis in each chapter presents a different perspective on the organization of discourse. Communication on *Twitch* is a complex system and its holistic description requires the study of its cross-modal exchanges between the streamer and their viewers, as well as the study of its segments of narration or commentary. Both descriptions need to consider the emergent unfolding of the video game, as well as the overall prevalence of the dialogical or monological forms of discourse.

However, the prevalence of a linguistic practice cannot be assessed merely through the detailed and qualitative study of examples of exchanges or narrations. This requires a complementary approach that studies discourse from a "synoptic perspective (Eggins & Slade, 1997)". Eggins & Slade (1997) applied this perspective to their study of spoken conversation. They coded the discursive moves of spoken discourse for its speaker and according to their particular speech functions. Their stidy revealed an uneven distribution of moves and move types between the participants. For example, some members of the conversation produced more moves and their moves were more assertive. The interlocutors produced fewer moves and they tended to have a more facilitative purpose. These patterns were interpreted as indicators for linguistic solidarity and discursive dominance in the casual conversation. Importantly, Eggins & Slade (1997) did not consider the recipients of discursive moves.

In video-mediated communication, studies have taken a similar synoptic view and used Goffman's (1981) participation framework, to describe possible configurations of production roles (speakers & writers) and recipients (listeners & readers) on platforms such as *YouTube* or *Hangouts* (Boyd, 2014; Dynel, 2014b; Rosenbaun et al., 2016b). As argued in 3.1, participation

frameworks describe discursive moves that are sent between participants. This research showed that the choice of linguistic mode was related to conversational practices such as upgrading one's participation status or discussing less pressing background information (cf. Sindoni, 2014; Rosenbaun et al., 2016b). However, their work did not consider how common such practices were, because participation frameworks are not countable.

This study will combine the strength of Eggins & Slade's (1997) description of discursive moves and enhance it through the additional insights taken from the applications of Goffman's (1981) participation frameworks to video-mediated communication. The study has developed a coding scheme for the discursive moves of live streaming (see 4.3) and coded each move of the transcribed data for its sender (whether the streamer or a chat participant), and its intended recipient via the program *UAM corpus tool*. After the complete annotation of the data, the program calculates the distribution of moves between participants. This distribution will show the relative amount of talk and chat as well as their direction of communication. For example, it will be possible to describe the prevalence of the streamer's talk to an individual audience member in comparison to talk that has no direct addressee (see chapter 5). A closer look at these two types of moves shows that the former tends to be part of dialogical exchanges, whereas the latter cover instances of narration or commentary. This makes it possible to discuss the frequency of cross-modal exchanges and the more monological discourse. In turn, the description of the frequency of moves will support the qualitative analysis in chapter 6 because it will contextualize patterns that are unusual for ordinary conversation such as 'incomplete exchanges' or 'exchange chaining'.

The next chapter will provide a detailed explanation of the methodology. It will describe the criteria-driven data selection as well as the transcription of the ephemeral live stream to a persistent and analyzable transcript.

Chapter 4 – Methodology

4.0 Initial remarks

Every live stream on *Twitch* is a unique and ephemeral interaction between streamer, viewer and game, which leads to two consequences for this study. First, there must be principles to the selection of live streaming broadcasts. Secondly, these broadcasts must be captured and put into a format that is accessible to discourse analysis. This chapter will explain this step-wise process and how it has been informed by the literature reviews in chapter 2 and 3.

Section 4.1 is dedicated to the data collection and explains the criterion-driven recording of twelve hours of naturally occurring live streams. The first criterion has been the video game and the study recorded broadcasts of three different games of varying game pace and goals. The selected games *World of Warcraft*, *League of Legends* and *FIFA 2015* are typical for live streaming and yet diverse enough to form a representative sample of different game types (4.1.1). The second criterion is channel size. Section 4.1.2 will explain the selection of a medium-sized channel and large-sized channel for each of the three games. This makes it possible to assess the impact of audience size on the organization of discourse in chapter 5 and 6. The selection of three games and two channels per game leads six broadcasts. Each broadcast was recorded for two hours, for a total of twelve hours.

The recordings included 54 'alert messages', which is deemed insufficient for their dedicated study. Therefore, 4.1.3 will explain how I collected an additional 96 alert messages to arrive at 150 alerts messages. Together, the 150 alert messages and twelve hours of video recording constitute the collected data. However, in this (recorded but untranscribed) form, it is not possible for linguists to analyse the discourse of live streaming. To make the data accessible to linguistic analysis the video recordings have to be transformed into a written transcript.

As there is no previous set of suitable transcription conventions, section 4.2 outlines the development of the novel transcription system, which is part of the original research contribution of the dissertation. The section is a revised version of Recktenwald (2017) and has been reprinted with permission of the *Journal of Pragmatics* and in compliance with the copyright agreement of Elsevier[17]. The section will review existing conventions in conversation and discourse analysis and point out problems for their direct application to live streaming data. The review covers studies on video game play and talk in the living room (Mondada, 2012; Piierainen-Marsh, 2012), written chat in video games (Bennerstedt, 2008) as well as research on video-mediated communication on *Skype*, *Google Hangouts* or *Camfrog* (Sindoni, 2014,

[17] https://www.elsevier.com/about/our-business/policies/copyright/personal-use

Rosenbaun et al, 2016a; Licoppe, 2017).

The transcription system developed in this study will systematically account for the communication in the spoken and written mode as well as the continuous progress in the game. When appropriate, the transcription scheme follows established principles of discourse analysis, but it also argues for necessary adjustments, which include a multi-column layout and the annotation of the video game in the form of game-events. The multi-column layout will provide a visible distinction between the spoken discourse by the streamer and written chat by audience. The annotation of the gameplay is based on my understanding of video games as dynamic systems (see 2.3) and the specific properties of the selected video games (see 4.1.1). The systematic transcription of discourse and the annotation of game events generates a textual rendition of the twelve hours of data. This textual rendition is subject to the research questions.

Section 4.3 covers the operationalization of RQ_1, RQ_2, and RQ_3 and it will explain how the Birmingham School is applied to the data. The discussion of the results occurs in the original research chapters 5, 6, and 7. The 8th and last chapter brings together the separate sets of findings and synthesizes them to a general description of the organization of live streaming.

4.1 Data collection

To describe the organization of discourse, the data collection has to build a corpus that is an accurate representation of live streaming on *Twitch*. The review of live streaming in 2.2 identified the size of the audience and the unfolding game as relevant factors for *Twitch*'s discourse (cf. Hamilton et al., 2014; Gandolfi, 2016). For example, streamer's have to negotiation their communication with the input-demands of the game. Additionally, the audience is said to behave differently in medium and large-sized chat channels. Therefore, the data collection must consider both aspects to arrive at a balanced sample. The next two sections explain the implementation of this considartion.

4.1.1 Game selection

First, it was necessary to get an overall picture of the most significant games that are live streamed on *Twitch*. Since 2012[18], *Twitch* releases monthly or bi-monthly lists of the 20 most watched games. Up to February 2015, all published lists were collected with the last one being "Top Twitch Games January 2015"[19]. Out of these individual lists, an alphabetically ordered complete list of all popular titles was generated (see appendix). The list showed that many

[18] https://blog.twitch.tv/top-twitch-games-for-december-2012-aa592b0a3bc5
[19] https://blog.twitch.tv/top-twitch-games-for-january-2015-44be9c5b332f

games have been very popular for brief periods of time and only a small number of games are played and streamed consistently. From this smaller subset, three games were selected.

Game	**Developer**	**Genre**	**Pace**
League of Legends	Riot Games	MOBA	Slow
FIFA 15	EA Sports	Sport Simulation	Medium
World of Warcraft	Blizzard Entertainment	MMORPG	Fast

Table 8 Selected video games

As of 2018, all three games are still popular on *Twitch*[20]and each title is the most popular of its genre. Moreover, they have been the subject of significant prior studies due to their cultural significance (Crawford, 2005; Ducheneaut et al. 2006; Jarret, 2016). The selection of games that are popular for extended periods ensures that they are typical for live streaming and the diversity of the titles with their different rules, goals and game pace ensures the representativeness of the research project.

League of Legends

With 27 million daily and 67 million monthly players, *League of Legends* is the most-played computer game in the world[21] as well as the most broadcasted game on *Twitch* from 2014 to 2017 (Twitch, 2014; 2015; 2016; 2017). Its success vividly highlights the synergetic relationship between live streaming, online games and e-sport (see 2.1.). *League of Legends* is a MOBA, which is short for Multiplayer Online Battle Arena. It is a game genre that developed out of strategy games in 2003 (Jarret, 2016) and it shares many of its conventions regarding pacing, goals and game structure. It is a multiplayer game between two teams of five players. Each player controls an avatar, which is called ‘champion’ and the match begins with a preparation phase called ‘champion select‘.

[20] The *FIFA* franchise is played in ist current version *FIFA 18*, which is very similar to previous versions.
[21] https://blogs.wsj.com/digits/2014/01/27/player-tally-for-league-of-legends-surges/

Figure 20 Avatar selection in League of Legends

The figure is a screenshot taken from *League of Legends* during the champion selection phase. The players of each team select their 'champion', which they want to play for the duration of the next round. Available champions are represented via thumbnails in the middle of the screen and selected champions are shown on the left and right columns. The selection takes place turn-by-turn and takes around five minutes to complete. In this phase, the game is very slow and requires very little input from the streamer. After this preparation period, the proper match starts, and the players are teleported onto the playing field 'Summoner's Rift'. On Summoner's Rift, the two teams play against one another and their goal is the destruction of the base of the opposing team. The following figure shows an annotated version, which will assist in the brief explanation[22].

Figure 21 Annotated summoners rift

The home base of each teams is on opposite ends of "summoner's rift" and they are connected via the three lanes called 'top', 'mid' (short for Middle) and 'bot' (short for bottom). As the

[22] Source: http://www.businessinsider.com/league-of-legends-pentakills-2015-8

match begins, players move their avatars towards one of the three lanes and each team attempts to push their enemies towards the opposing team's home base. During a match, players also distinguish between the early 'laning phase' and the later 'team-fight phase'. Both phases feature combat between avatars and attacks on defensive structures ('towers'). Each match is a very slow siege towards the main base and the match ends when either base 1 or 2 is destroyed. After the match, players may inspect game statistics on the 'post-match screen', stop playing or play another round.

Stages	**Usual Duration**	**Game Goals**
1) Champion Select	~5 minutes	Selection of Avatar
2) Match on Summoner's Rift: - Laning Phase - Team-fight Phase	~ 30 – 45 Minutes (10 – 15 minutes) (20 – 30 minutes)	 Defeat players & towers (optional) Destroy enemy base (mandatory)
3) Post-Match Screen	----	----
4) Stop playing or new round	----	----

Table 9 Organization of League of Legends

The table summarizes the organization of *League of Legends*. Every match takes between 30 to 45 minutes[23] and is a slow progression from champion select towards the main goal and the conclusion of the match. There is no time pressure and in practice it is very difficult to win quickly. The game's pace is further limited by the number of actions a player can perform. Every avatar attack has a 'cooldown period', which prevents them from being re-used too often. The game's goals and the restrictions on avatars lead to an overall slow game pace. The game revolves around strategic decision making instead of fast executions of commands.

For this game, the annotation of game events includes the transitions between the stages of the match as well as the achievement of optional and mandatory game goals.

FIFA 2015

The second selected game is the sport simulation *FIFA 2015*. The *FIFA* franchise is the most popular sport simulation and there is an annual release of a new version (Crawford, 2005, 2015; Baerg, 2006; Consalvo et al., 2013). These versions are iterations of one another and are overall very similar in gameplay. The gameplay consists of team management and soccer matches. Most streamers play in online leagues or online tournaments against other players.

[23] See http://www.leaguemath.com/match-duration-analysis/

Figure 22 Team management and soccer match in FIFA 2015

The figure consists of two screenshots[24] that show both aspects of the game. On the left, the players manage the team by changing the roster and formation. On the right, the player controls the avatars in the match. The success on and off the field are closely connected because avatars have the same strengths and weaknesses as the real-life counterparts. A strong squad is a significant advantage in a match. There is no time limit for the team management and streamers can invest significant amounts of time to trade virtual soccer players in preparation for a match. In turn, matches reward the player with in-game currency that can be reinvested in improvements of the team. The matches of *FIFA 2015* are played on a virtual pitch and have similar rules as traditional soccer. For example, there are restrictions on tackles & fouls that are sanctioned with yellow & red cards. *FIFA 2015* and ordinary soccer have in common that ball possession is a key indicator for success. A team in possession of the ball has more influence on the direction of the game and ball possession has been shown to positively correlate with win-rates in real-world soccer (Göral, 2015) and in *FIFA 2015*[25]. *FIFA* matches are won by the team that scores more goals over the fixed period of time and if there is no immediate winner, the match may continue with overtime or even a penalty shootout.

[24] Source: http://www.ea.com/fifa15ut-and/images/build-and-manage-your-ultimate-team

[25] http://www.fifaencyclopedia.com/fifa-15-passing-possession

Stages	Usual Duration	Game Goals
1)Team management	----	Trading of virtual soccer players
2) First Half-Time	6 Minutes	Ball Possession (optional) Scoring Soccer Goals (mandatory)
3) Half-Time Break	~1 Minute	
4) Second Half-Time	6 Minutes	
5) Overtime (optional)	2 x 3 Minutes	
6) Penalty Shootout (optional)	~1 Minute	
7) Stop playing or new match	----	----

Table 10 Organization of FIFA 2015

The game pace of *FIFA 2015* can be categorized as medium because there is a fixed time limit, which creates time pressure for the team that is behind in goals. Compared to *League of Legends*, *FIFA 2015* has an increased tempo and players can perform many actions in a few seconds. The streamer is constantly engaged and is trying to secure the ball and score a soccer goal.

This is accounted for in the transcription by including the scoring of soccer goals, transitions between the game's stages, changes in ball possession and breaks in the gameplay due to fouls and free kicks.

World of Warcraft

The last selected game is *World of Warcraft*. It was initially released in 2004 and is the most popular title of the massively multiplayer online role-playing game (MMORPG) genre. Despite a recent decline in player numbers, it still reports an active player base of 5.6 million players and it is one of the few digital games that continues to be played for over a decade[26]. Its lasting success made it one of the most researched games. Among many other aspects, studies discussed its social systems of guilds & groups, player motivations, play styles as well as player cooperation & coordination in combat (Steinkühler & Williams 2006; Ducheneaut et al., 2006; Bennerstedt & Ivarsson, 2010; Williams & Kirschner, 2012; Bean & Groth-Marnat, 2014; Witkowski, 2012).

World of Warcraft allows for several types of play (cf. Suznjevic & Matijasevic, 2009). For example, there is 'questing', where a single player performs tasks in the game world such as collecting objects. Then there is 'raiding', which features a group of players fighting together versus computer-controlled boss monster. This study will focus on the game mode 'arena' because it is very popular, and many streamers exclusively broadcasts ii on *Twitch*.

[26] http://www.mmo-champion.com/content/5063-WoW-Down-to-5-6-Million-Subscribers

Figure 23 WoW arena

The annotated figure is taken from *World of Warcraft* and it shows a player-controlled avatar in the 'Circle of Blood Arena'. Arenas in *World of Warcraft* are virtual fighting-pits, where players fight against one another in small teams of two versus two or three versus three. The goal of a round is to defeat or 'kill' all members of the opposing team and win the round. After victory or defeat, players leave the arena and may quickly start the next match.

Stages	**Usual Duration**	**Game Goals**
1) Arena Match Preparation	~1 minute	Discuss strategy against opponent (optional)
2) Arena Match	~1-3 Minute	Defeat opposing team (mandatory)
3) Stop playing or new round	----	----

Table 11 Organization of World of Warcraft arena

The rules and organization of the game mode are very simple, and its appeal is the fast-paced combat between players. The average match lasts only between 1 to 3 minutes and within this period, the player is constant engaged and fighting without pauses or slowdowns in the action. *World of Warcraft* arena requires quick reflexes and a lot of attention for short bursts of time. Player-actions must be coordinated, teamwork is rewarded and there is very little tolerance for mistakes (Bennerstedt & Ivarsson, 2010; Witkowski, 2012). The transcription of *Word of Warcraft*'s gameplay will include any avatar 'death' or 'kill', the different stages of *World of Warcraft* arena and the outcomes of matches.

This brief introduction showed that *World of Warcraft, FIFA* and *League of Legends* can be distinguished in terms of their goals, game pace and temporal organization. They build a diverse set of games that is described in a standardized format. This makes it possible to develop a consistent annotation for all three games with game events as their basic unit. As this annotation is applied to the broadcast of the three different games, it becomes possible to draw generalizable conclusions about the impact of the video game on the organization of

communication (see chapter 6) and the transcript reveals how discursive moves or exchanges relate to the annotated game events. This link between game and communication is an important step for describing the organization of discourse.

However, other steps are also necessary because prior research on computer-mediated communication and live streaming has argued that the number of chat participants changes the patterns of communication (Herring, 1999; Jones et al, 2004; Hamilton et al., 2014; Ford et al., 2017). Medium-sized live streams are understood as a "meaningful medium of discussion (Hamilton et al., 2014)" that contain "meaningful exchanges (Olejniczak, 2015)". Larger audiences of over 1000 viewers are said to turn into a spectacle (cf. Debord, 1967; Szablewicz, 2015; Gandolfi, 2016), where the communication of the audience mimics the "roar of the crowd", a "waterfall of text" or "cheering" (Hamilton et al., 2014; Olejniczak 2015). Therefore, the data selection must also consider different audience sizes, which is achieved through the process of channel selection.

4.1.2 Channel selection and data recording

Channel selection refers to the selection of a livestream broadcast based on the number of its viewers. It does not consider the specific broadcaster/streamer/host in the decision making. This is an important point because Gandolfi (2016) and Hamilton et al. (2014) suggest that the 'persona' of a streamer might also influence the communication. However, they do not discuss which type of 'persona' exist and it is unclear how they could be identified. A streamer's persona may express itself in his or her communicative style with the audience (cf. Tannen, 1987), but this can only be studied through the analysis of the data. For example, section 5.4 will discuss an outlier in the communicative patterns and will suggest that this might be due to streamer's less talkative style. However, his style was not knowable beforehand, which makes the communicative style of the streamer an unsuitable criterion for data selection. The broadcasted game and the size of the audience are visible and unambiguous, which allows the data selection to make informed decisions based on solid criteria.

The literature review suggests making a distinction between medium-sized channels of ~100 – 150 viewers and large-sized channels of over 1000 (Hamilton et al., 2014; Olejniczak, 2015; Nematzadeh et al. 2016; Ford et al., 2017). Briefly, I also considered to add the category of small channels with less than 20 viewers. However, this idea was quickly abandoned because there was not enough communication in these channels. Streamers barely talked, and the audience wrote few messages. It might indicate that live streaming requires a critical mass to start and maintain a conversation (Rogers, 1962) but beyond this general interpretation the study did not pursue the issue further and settled for channels of two distinct sizes.

For each game there is one medium-sized and one large-sized channel, which leads to a total of six broadcasts. The next decision that had to be made was about the necessary number of hours to come towards a reliable analysis. There is no established standard or golden rule for to establish an appropriate size. Many studies do not explicitly discuss the number of hours that have been used (Reeves et al, 2015; Rintel et al. 2015). On the lower end, there are studies that quote 90 minutes to three hours of original data (Mondada, 2012; Rosenbaun et al., 2016a). On the upper end there are claims of over 800 hours (Sindoni, 2014). The middle ground consists of studies that use six to twenty hours of data (Gerhard, 2008; Piierainen-Marsh, 2012; Rosenbaun & Liccoppe, 2017). Based on these considerations and practical feasibility, the study decided for twelve hours of data. Twelve hours are comparable to previous research projects and it is possible to divide them evenly across six broadcasts. This configuration allows for two types of comparison. First, it will be possible to compare the communication of online live streaming across the game types by keeping the audience sizes constant.

Game	**Channel Size**
League of Legends	Medium
FIFA 2015	
World of Warcraft	
League of Legends	Large
FIFA	
World of Warcraft	

Table 12 Comparison across games

The table shows the six broadcasts as they group in two triplets with a constant channel size. This makes it possible to study the discourse of live streaming across the different games and it can reveal the influence of the game's properties such as pace or goals. For example, it will be argued that the quicker games *FIFA 2015* and *World of Warcraft* feature more 'commenting' than 'reporting', but in the slow-paced *League of Legends*, the opposite is true (see 6.2).

The second type of comparison keeps the game as a constant and compares the discourse across channel size.

Game	Channel Size
League of Legends	Medium
	Large
FIFA 2015	Medium
	Large
World of Warcraft	Medium
	Large

Table 13 Comparison across channel size

The table shows the grouping into three duplets of medium and large-sized channels. This makes it possible to study the discourse under consideration of the audience size. For example, this distinction makes it possible to argue that in large-sized streams the audience prefers to communicate with the streamer and is less interested in chat with fellow viewers. In medium-sized channels, this tendency is reversed and there is still a substantial amount of dialogue between viewers (see 5.4).

In the next step, the criterion-based data selection must be realized into practice. It began by initially observing the 'game directory' of *World of Warcraft*, *League of Legends* and *FIFA 2015* in March 2015. On Twitch, the game directory lists all active channels that currently broadcast the particular game[27]. Consistently streaming channels that had the appropriate audience size and were broadcasting in English were collected in a list (see appendix). The focus on English is motivated by its status as the lingua franca in computer-mediated communication (Seidlhofer, 2005) and it is a practical necessity due to my own language proficiency. Streamers that were non-native speakers were included, as long as they broadcasted in English. This is not deemed problematic because the study is not concerned with issues of correctness or linguistic practices that distinguish native from non-native speakers (cf. Long, 1983). From the list of suitable streamers, six channels were randomly selected and afterwards each channel was recorded for two hours. The broadcasts were screen-recorded directly from the channel page with the freeware program *Open Broadcaster Software*. This produces a video file in the *.avi format, which contains the streamer's spoken communication, the gameplay and the written chat by the audience. The recording includes six rounds of *League of Legends*, six matches of *FIFA 2015,* and thirty-four rounds of *World of Warcraft* arena. This was deemed a sufficiently large data size that also adequately fulfills the criterions for the game and channel selection.

[27] For example, the directory of *League of Legends* is available at www.twitch.tv/directory/game/LeagueofLegends

It is important to note that the written chat is part of this video. However, in this form the chat is more comparable to a photo of a text rather than a proper text file. Therefore, the chat messages are not accessible for traditional word processing or linguistic tools. To solve this issue, *Twitch*'s chat was simulatenously captured or 'logged'. *Twitch*'s chat runs on its own *IRC* server and it is possible to connect to this server with an ordinary *IRC* client such as *gIRC*, *mIRC* or *Chatty* (see 2.3; 3.3). The study used *Chatty* to live record all incoming chat messages of the selected broadcasts into a separate 'plain text file' (*.txt). This made the manual transcription of several thousand written chat messages unnecessary.

The recording also captured 54 instances of on-screen alert messages (see fig. 11). 45 messages were 'subscription notifications' and 9 were 'donation alerts' (see 2.3 & chapter 7). Their occurance quickly influences the interaction between streamer and audience, which means they have a significant impact on the organization of discourse. Moreover, a preliminary observation suggested that they have very different purposes in comparison to other types of talk and chat. Their number was decided to be too small of a sample size to produce a valid analysis and it was desirable to achieve a wider spectrum of messages to assure data saturation (Mason, 2010).

4.1.3 Collection of alert messages

This study treats alert messages as a special type of linguistic exchange between viewer and streamer. The viewer is the initiator of the exchange and pays to produce a discursive move in the center of the video. The streamer is the recipient of the payment and in most cases also the addressee of the message. Therefore, the collection of additional examples focused on the interaction between these two parties and actively searched them them for the sake of efficiency.

Many streamers upload their past broadcasts on *YouTube*, where they become archived and the result is is an abundance of recordings that can be searched through quickly.

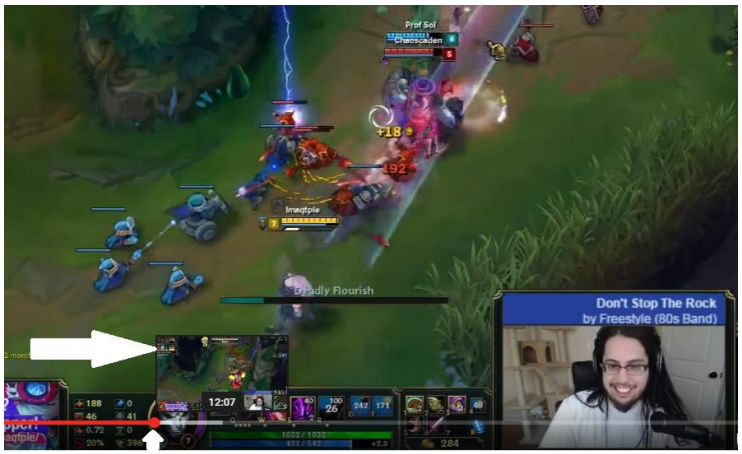


Figure 24 Annotated screenshot of a YouTube recording

The annotated screenshot shows the process of searching for alert messages on *YouTube*. There

is a video timeline at the bottom of the recording. The lower arrow indicates the position of the cursor as it is moving through timeline of the video. The larger upper arrow shows a picture-in-picture preview of the upcoming video. In this case, the alert message is visible within the preview at 12:07 minutes. With this information it is possible to fast-forward to the exact position and screen capture the alert message. 96 additional alert messages have been collected with this method and increased the total number of alert messages to 150 with 44 being 'donation alerts' and 106 being 'subscription notifications'. At this sample size, the data reached a saturation point and every type of alert message that was later identified in the analysis has been represented with several examples (cf. Mason, 2010; see chapter 7).

Taken together, the 150 alert messages, twelve hours of chat log and the recorded video are the 'raw data' used for the study. The next section has to explain how this 'raw data' can be made accessible to discourse analysis and the study's linguistic tool kit. This requires an appropriate analytical entry point to the 'raw data', which will be the transcription of the interaction into a textual format.

4.2 Data transcription

Transcripts are very useful in academic practice as they are accessible to data coding and subsequent analysis (cf. Meredith, 2015). They are more than a simple change from the audio-visual to the written mode, but the product of a transcription process that transforms the original interaction into a static written representation (Bauman & Briggs, 1990). There are many consequences to this process that have been addressed in comprehensive discussions about the transcription practices of spoken language, the format and representational choices in transcripts, as well as historical trends and future challenges (Ashmore, 2004; Bucholtz, 2000, 2007; Duranti, 2006; Dressler & Kreuz, 2000; Du Bois et. al., 1993; Ochs, 1979; Edwards & Lambert, 1993; Goodwin, 1994; Mondada, 2007, 2014; Neville, 2015).

This study reflects on this literature and its relevant information for the transcription of live streaming of video games. It informed several decisions in the development of the transcription system, but it was also necessary to pursue completely new directions in order to adequately represent the data and study the discursive moves and their interaction with the gameplay. Therefore, the section discusses general concerns about transcripts, the choice for a basic layout, the spoken & embodied communication of the streamer, the representation of the chat communication, and the annotation of the video game material.

General concerns about transcripts

Transcripts are symbolic representations of actual interactions. Interactions between participants occur in unique and passing moment of time. Through the process of recording, they become fossilized as ‘raw data’ (see 4.1). This raw data is then transformed into a transcript. This step always leads to a loss of information because it condenses the recorded audio-video data into a textual rendition. The ability of this textual rendition to capture the qualities of interaction is its ‘fidelity’ (Ayaß, 2015) or ‘aptness’ (Bezemer & Marves, 2011). Transcripts of high fidelity allow the reader to imagine the unfolding of the communication and are generally very desirable. Fidelity makes the transcript an intersubjective document that accurately represents the ‘raw data’ and allows the reader to follow the presented analysis (Psathas & Anderson, 1990).

At the same time, transcripts must posses the necessary ‘analytic utility’ to answer the research question (cf. Ashmore & Reed, 2000). Often, researchers must strike a balance between ‘basic’ or ‘fine’ transcripts (Ayaß, 2015). ‘Fine’ transcripts contain more information, but there is a risk of them becoming too convoluted and full of unneeded and complicated details that only distract from the important points. Given the many intricacies of live streaming with its cross-modal communication, game-related jargon and number of participants, this project decided to begin with a ‘basic’ transcription that only covers the linguistic features that are directly relevant for this research project. This will help to understand the fundamental interactional processes between streamer, chat and the game.

Many of the transcription choices have been adopted from Dressler & Kreuz (2000). Their transcription system is based on an extensive meta review and their design principles ‘consensus’, ‘parsimony’ and ‘extensibility’ are particularly suitable for the undertaking of this study. Consensus refers to an agreement among researchers in terms of symbol use. Whenever possible, a transcript should adopt established conventions to describe a well-researched feature because it facilitates intersubjectivity and lowers the barrier of entry for other readers. Similarly, parsimony describes the attempt to minimize the number of conventions used in a transcript. Discourse phenomena should be accounted for with as few symbols as possible to avoid overly difficult transcripts. Extensibility acknowledges that no transcription scheme is ever final or complete and it should be possible to amend a transcription scheme for different purposes by introducing new symbols or conventions.

The basic layout of the transcript

For any transcript the first concern is the basic layout and the unit of transcription. In Conversation Analysis, the Jeffersonian 'play-script format' (Meredith, 2015) has been widely adopted since its initial use by Sacks et al. (1974). All participants are presented in a single column and speaker shifts are indicated prior to the utterance. The format has proven to be very useful in many different situations such as co-present dyads, small group conversation and video-mediated communication (Kerbrat-Orecchioni, 2004; Rosenbaun et al., 2016a, b).

An alternative is the use of distinct 'participant columns' (Ochs, 1979). In this format, every speaker receives his or her own designated column and the columns are positioned next to one another. For live streaming, there are several reasons to favor this format because it provides a separate column for the streamer's communication, the audience's chat and the unfolding game. The streamer and audience perform different activities and communicate in different linguistic modes. Distinct columns increase the fidelity of the transcript because they emphasize that streamer's speech is audible, whereas the written messages are visual. Moreover, in a single-column layout the high volume of chat messages would make the streamer's speech appear disjointed in the transcript. A multi-column layout offers some remedy to this problem.

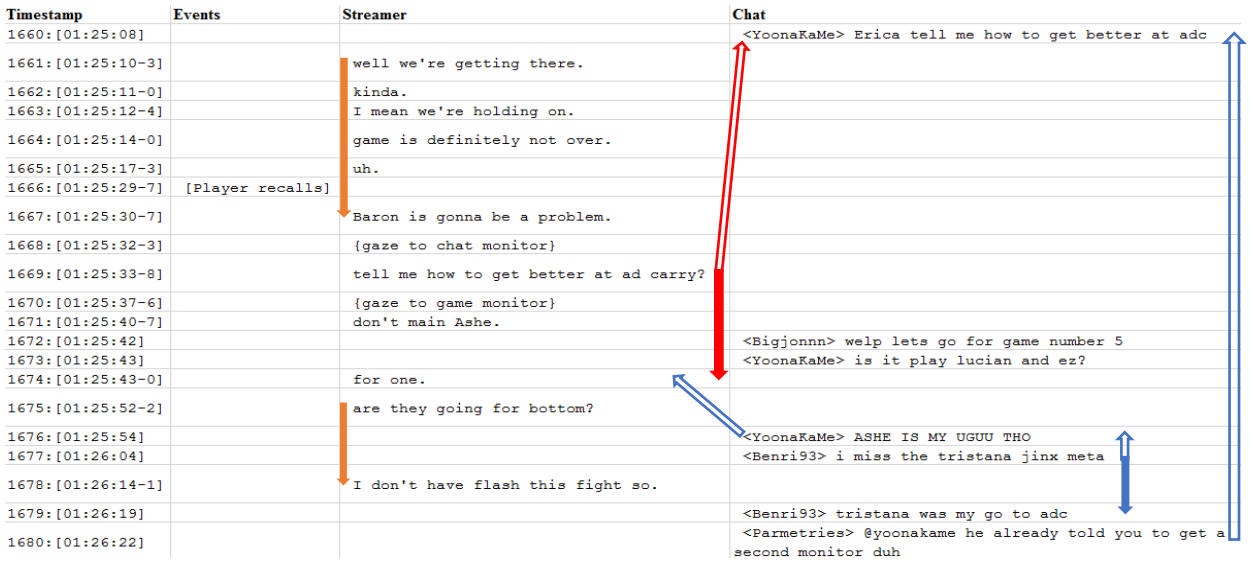


Figure 25 The basic layout of the transcript

For now, the figure is only an example to illustrate the basic layout. The first column contains the time stamps and line numbers needed to demonstrate temporal relationships between utterances, messages or the video game. At times, it was necessary to abridge segments of the transcript because *Twitch*'s chat is very active and page space can be limited. In such cases, the colon ':' symbol is placed in the respective row to signal the omission of less relevant data in the excerpt. The three main columns are dedicated to the annotation of the game, transcription

of the streamer and the written chat by the audience. It will enable the first systematic study of the interaction on the micro-level and there is an intrinsic iconicity in the format, which is supported by the colored arrows. Mono-modal communication is a coherent stretch of text within a single column. Cross-modal communication is as text that is cross-referencing another column.

Arrows are another visual aide because they highlight stretches of talk or chat that belong together. For example, the arrow from line 1661 to 1667 shows a stretch of talk by the streamer. The arrow is to the left of his discourse, closer to the game 'event' column and colored in orange[28]. The position and color signal that the talk is about the game and not directly dialogical with the audience. Blue arrows indicate stretches of writing by an audience member, and an arrow in red indicates cross-modal talk by the streamer to a viewer. In this example, the streamer is responding to an initiation from the chat. This initiation is identified by the upward pointing arrow (l.1660). Downward pointing arrows with color filling signal chunks of discourse, whereas upwards pointing arrows without color filling indicate references to prior discourse.

In contrast to previous research on video-mediated communication (cf. Rosenbaun et al. 2016b; Sindoni, 2014), this project distinguishes cross-modal exchanges from speech-to-writing and from those that communicate from writing-to-speech. Both types of cross-modal communication are visible as explicit discourse references across columns and they can be distinguished based on the direction and color of arrows. This format is useful to demonstrate that spoken-to-written exchanges have a different structure from written-to-spoken exchanges (see 6.1). Similarly, the game's influence on the discourse becomes clearly visible as a relationship between the 'events' column and the streamer's talk or viewer chat (see 6.2).

These advantages of a multi-column format are not unique to the live streaming of video games. In other broadcast genres such as social eating or dancing, it will remains useful to separate the activity, the talk by the broadcaster and the chat by the audience for analytical reasons.

Transcribing the streamer

The streamers' most prevalent communicative resource is speech. It is transcribed in intonation units and follows many conventions outlined in Dressler & Kreuz (2000). However, there are also necessary adjustments, which require an explicit explanation. For example, quotation

[28] The color scheme is analogue to the colors in figure 4 on page 19. In figure 4, orange indicates the interaction of the streamer with the video game red indicates the streamer's talk to the audience and blue indicated the viewer's chat.

marks “ “ signal a sudden and temporary change in the voice quality of an utterance. Pitch modulation is often used to indicate ‘reported speech’ and other ‘external voices’ (Holt, 1996). This has been identified as a common practice in the streamer’s cross-modal exchanges. They use an ‘external voice’ to highlight the written messages by the viewers (see 6.1).

Another issue is the transcription of pauses. Short pauses of less than 0.5s have been transcribed as ‘ (..) ‘. Longer pauses are not transcribed directly and must be inferred from the time stamps and the colon symbol. Online live streaming is in an ‘open state of talk’ (Goffman, 1981), which invites but does not require constant communication. Therefore, it is common have long segments without a spoken utterance by the streamer.

Lastly, the streamer must negotiate his or her embodied communication with the demands of the video game. Their use of gestures for communicative purposes is extremely limited as their hands are occupied with the control of the mouse and keyboard. Non-verbal communication was not systematically transcribed because of its low prevalence and the study’s focus on the spoken and written communication in moves and exchanges. The only embodied behavior that has been transcribed consistently is gaze shifts. They have been annotated in ‘{ }’ curly brackets in the streamer’s column. It is very difficult for the broadcaster to look at the game and chat at the same time (see 2.3). Their gaze signals their primary involvement with game or chat. Gaze shifts are points of transition between these two types of interaction. They are a visible cue that assists in the study of exchanges between streamer and audience.

Audience messages

The recorded chat logs are used as a starting point for the chat column of the audience. In these chat logs, the default structure of every message is [*time stamp*] <*username*> *chat message*. The time stamp is used to coordinate the chat messages with the transcription of the streamer’s talk and the annotation of the game. A particular feature of *Twitch’s* chat is its wide-spread use of platform specific emoji. These emoji often depict the faces of prominent streamers and staff member and their facial expressions often represent an emotion. The ‘Kappa’ emoji  stands for sarcasm and the emoji ‘FailFish’  shows embarrassment. Emoji are important pragmatic markers and signal speaker’s intent for a written message (Thompson & Filik, 2016). Since the chat logs are in plain text format *.txt , they include the emoji as written text commands. For example, the consecutive use of the emoji  in the chat would be represented in the chatlog as ‘KappaKappa’. In the final excerpts of the examples, the ‘emoji text’ has been manually replaced with the corresponding ‘emoji picture’ to increase the accuracy and fidelity

of the transcript.

A high volume of emoji has been described as one of the disruptive features of *Twitch*'s chat, which turns the conversation into a 'waterfall of text' or 'cacophony' (Hamilton et al, 2014; Nematzadeh et al., 2016; Ford et al., 2017). In 6.2, the study will show that the distribution of emoji is not random or chaotic. They are systematically tied to the unfolding of the gameplay and the occurrence of game events.

The annotation of game events

The representation of the game is very important for the study of live streaming, but an adequate approach is not intuitively obvious. Most studies on video games rely on screenshots as a means of representation. Ethnographers and researchers from the symbolic interactionist paradigm tend to give holistic accounts of player interactions and screenshots generally only set the stage for the researcher's thick descriptions (Steinkühler & Williams, 2006; Taylor, 2009; Williams & Kirschner, 2012). In Conversation Analysis, screenshots of games are used to support specific points of the analysis (Mondada, 2012; Piirainen-Marsh, 2012). However, these studies also show the limitation of screenshots as a format for the accurate representation of emergent gameplay. When studying a soccer goal in *FIFA 2010*, Mondada (2012, p. 241) must include the explicit description "*luc scores the goal for Real Madrid*" into the written transcript. She uses a static screenshot to illustrate a soccer attack, but the screenshot cannot convey the outcome of the gameplay because the screenshot is a frozen representation of a unique moment. In ethnomethodology, 'vignettes' are used to approximate the sequential developments in games (Bennerstedt, 2008; Bennerstedt & Ivarsson, 2010; Reeves et al., 2009). Vignettes are a series of still-images that form a 'visual narrative' (Ayaß, 2015). They promise greater fidelity in the illustration of the gameplay and can represent the progression of the gameplay. However, vignettes always remain game-specific because they rely on the surface representations of the game, i.e. its graphics, to describe its status. Therefore, they offer only limited analytical utility for a study dedicated to the broadcast of several different games.

A transcript designed to study live streaming of games more generally must look at the common core of all games. An adequate solution is the annotation of game events based on the fundamental properties of video games and gaming (see 2.4). Video games are interactive systems that operates like 'state machines' (Järvinen, 2003). When players play games, they provide the necessary input to the current game state and the game calculates the output based on its procedural rules and internal organization (Juul, 2004; see chapter 2). Game events are meaningful and salient transition from one game state to another.

There is a growing body of evidence that players perceive game events as distinct categories. In several studies, gamers were monitored during play and psychophysiological measurements were taken. The measurements were pupil dilation, electrocardiograms and skin conductance. They reveal that states of mental arousal correlate with unfolding gameplay (Ravaja et al., 2006; Kivikangas et al., 2011). Most notably, Weber et al. (2009) measured peaks and shifts in heart rate and were able to map them with particularly intense game events. The body of gamers subconsciously reacts to the perception of game events on screen.

Game events are part of the game system and are realized through the interaction of the players with the system. They include processes that are countable, clearly observable or related to the game's goals or structure. The annotated game events in this study are derived from the goals and rules of *World of Warcraft*, *FIFA 2015* and *League of Legends* such as scoring a soccer goal or scoring a 'player kill'. Therefore, the principle of annotation is consistent across all three titles and also applicable to other games. The description of the game events is decidedly neutral, aims for objectivity and is using very general phrases such as "`opponent team scores a soccer goals`" or "`end of first half-time`". Other features of game events are not transcribed and are only assumed to be important, if they are made relevant by the participants in discourse (cf. Sacks, 1984). If a game event has a specific significance for the broadcast, it will emerge from the spoken and written discourse, instead of being prescriptively encoded in the annotation.

The game events are annotated in the dedicated 'events' column. They share this column with the on-screen alert messages, which also appear on the primary gaming monitor of the streamer (see 2.4). The chat occurs on the secondary monitor and requires a gaze shift. In this sense, the order of the columns 1) events, 2) streamer and 3) 'chat' is also a visual metaphor. The discourse of the streamer is positioned in the middle and between the two. Arrows and discourse references "to the left" of the transcript correspond to an interactional orientation to game events and alert messages. Arrows and discourse references "to the right" of the transcript are instances of cross-modal communication and an orientation to the written chat. This makes the three columns of the transcript a very accurate textual representation of the natural occurring interaction.

This transcription method has been applied to the six broadcasts and each transcript has been saved in a separate *Microsoft Excel* sheet. These six sheets and the 150 alert messages are the data of the study. A digital copy of the complete data set has been provided to the library of PolyU and is available upon request via the Hong Kong interlibrary loan system. The next section will explain how this data was studied and the research questions were operationalized.

4.3 Operationalization of research questions

The operationalization of the research question is the cornerstone of the methodology. It builds on the literature review of live streaming and linguistics, influences the data analysis, and the interpretation of the results. The study argues that live streaming on *Twitch* is novel form of computer-mediated communication similar to conversation and video-mediated communication. However there are also substantial differences with regard to the number of participants, their access to the spoken and written mode, and the influence of the unfolding gameplay. Descriptions of live streaming are vague and often ignore the role of the streamer in the interaction (Smith et al., 2013; Hamilton et. al, 2014; Gandolfi, 2016; Cheung, 2017). So far, there is no comprehensive linguistic description of the discourse of live streaming and therefore this thesis pursues a very broad research question.

RQ: How is discourse organized during live streaming of video games?

To make this question answerable, the overarching research question will be broken down into three sub-questions.

RQ_1: How are the discursive moves distributed between streamer and audience and does this distribution change with audience size and game's pace?

RQ_2: How are the discursive moves combined into larger units of discourse and how is this process influenced by the unfolding gameplay?

RQ_3: What is the purpose of alert messages and how do they influence the organization of discourse?

Each of the sub-questions will receive a dedicated research chapter to describe the discourse of live streaming from a different perspective. Afterwards, their individual findings will be synthesized into a conclusion to answer the main question.

RQ1 is motivated by the argument that a live stream is a discourse of a certain type. It is said to be either a conversation, a 'top-down' narration or cacophony (Hamilton et al, 2014; Karhulahti, 2016; Nematzadeh et al., 2016; Gandolfi, 2016). Previous studies suggested the important role of the number of chat participants and the broadcasted game. However, they remain vague in their description of conversational or narrational features and do not test their

proposed categories on actual data. They lacked the linguistic criteria and the analytical entry point to describe the discourse. To bridge this gap, the third chapter introduced the discourse analytical frameworks of the Birmingham school, which describes interactions through moves and exchanges. A conversation consists of several dialogical exchanges between two or more parties (see 3.2), whereas narration and commentary are more monological and form a coherent series of moves by a single speaker.

Therefore, it is possible to assess the prevalence of monologue and dialogue by studying the distribution of discursive moves between the streamer and audience (see RQ1). A discursive move can be described through its sender and recipient configuration. For example, a move can be produced by the streamer to a chat participant. In this sense, each move has a direction of communication. For a single move, this information is not very meaningful, but if this description is applied to each move of s representative data set, it shows who is communicating to whom during the broadcast of live streaming on *Twitch*.

This idea has been operationalized with the free software *UAM Corpus Tool*. In *UAM Corpus Tool*, it is possible to design a coding scheme and then manually apply the scheme to discourse units within a corpus. In this study, the unit of coding is the discursive move and each move has been coded for its sender and intended recipient. The coding started out with two primary assumptions. First, there is the mandatory distinction between the speech of the streamer and the chat of the audience. Secondly, due to the public nature of online live streaming, all audience members are considered ratified recipients (cf. Dynel, 2014a).

Any type of manual coding is a subjective process and to assure consistency the coding was improved over time. This meant that names of categories have been changed to be more

appropriate[29] and that the coding of moves saw constant revisions (cf. Saldana 2009). The coding of each move has been decided based on its linguistic features such as cohesive devices, terms of address or its prospection in the ongoing communication. In contentious cases, the original recording has been re-examined.

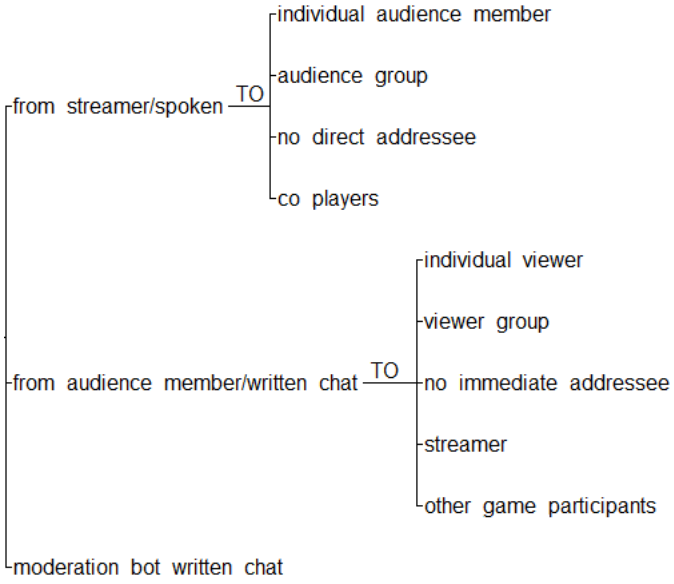


Figure 26 UAM coding scheme template

The figure shows the template of the final *UAM* coding scheme as it was applied to the transcripts of the data. The template shows all possible configurations between senders and recipients that have been identified during the coding process. Yet, any given move is always coded as a specific and signle category. For example, a move may be classified as 'from streamer/spoken to individual audience member' if the streamer uttered the move towards an audience member. If an audience member addressed the streamer, the move has been coded as 'from audience member/written chat to streamer'.

It is also important to highlight some consequences of the *UAM* software. *UAM* demands that every label for a category is only used once. This means that similar or equivalent labels must be renamed. If a viewer is addressed by the streamer, it is coded as 'to individual audience member'. If the viewer is addressed by a fellow chat participant, then the label 'to individual viewer' was used. There is no conceptual difference between the two, and it is only due to a software limitation. The same reason explains the differences between 'to collective audience' and 'to collective viewer' as well as 'to no direct addressee' and 'to no immediate addressee'.

In chapter 5, each of these categories will be explained via representative examples that illustrate the content of the categories and assist in the interpretation of the coding results. Once

[29] For example, the viewers of the live stream have been mislabeled as "chat participants". This turned out to be a logical error because all chat participants are viewers, but not every viewer is a chat participant.

every move has been manually coded, *UAM* automatically calculates the distribution of moves in each category for a given transcript. All six transcripts were coded separately and *UAM* produced six tables of move distributions. These six tables can be systematically studied and compared due to the criteria-driven game and channel selection. Chapter 5 will analyze the distribution of moves for each broadcast to see if the patterns of communication change for broadcasts with a different audience size or games pace. In answering RQ_1, the chapter will provide the first of three perspectives on the organization of discourse and its results will contextualize the sequential micro-level study in chapter 6.

Chapter 6 is dedicated to RQ_2 and will study the discourse of live streaming through the description moves and their combination in larger units of discourse. This has been operationalized by applying the discourse analytical tools of the Birmingham school to the transcripts (see 3.1.1). There were several close readings of the data and I examined the linguistic interaction close to the annotated game events and gaze shifts. Dialogical segments between participants have been analyzed as exchanges and recurring patterns emerged. The recurring exchange structures and their formulas are understood as common communicative practices during live streaming. The analysis reveals that dialogical exchanges differ depending on the initiating party, i.e. the streamer or the chat. After their initial description, the analysis provides potential reasons for these differences in the exchange structure (see 6.1).

A similar multi-step process took place for the study of the more monological discourse, which has previously been classified as narration of gameplay (Karhulahti, 2016). The transcript has been searched for segments of discourse that are without prospection. The study looked for discursive moves that do not form exchanges with different participants. The close reading of these segments suggested that narration of gameplay is not a single phenomenon and is better understood as two distinct communicative practices, which will be called ‘commenting’ and ‘reporting’. The analysis continued to explore the two categories and concluded that commenting and reporting have a significantly different forms and purposes in discourse (see section 6.2).

The study of exchanges, commenting and reporting also suggested that these communicative practices alternate with one another in an orderly manner. This indicates that there is a higher level of organization of discourse. Therefore, the transcripts have been studied for transactions, i.e. complete discussions of a topic. It became apparent that there is a finite set of topics that are discussed and that transactions re-appear periodically (see 6.3). Certain types of transactions consistently re-appeared at the same point in the game. For example, discussions about the selection of an avatar began during ‘champion select’ in *League of Legends* or during

the team management phase in *FIFA 2015* (see 4.1.1). Often, the interaction has been very context bound and closely connected to the 'non-linguistic organization' of the broadcast. The non-linguistic organization of the broadcast is cyclical because streamers play many rounds of their game for durations of 4 to 12 hours. These long broadcasts are repeated in daily streaming sessions and according to a weekly streaming schedule. The study will demonstrate this cyclical repetition of the broadcast and its discourse with examples and evidence taken from the recorded data. The fifth and sixth Chaptercan explain most patterns of the interaction and the organization of discourse.

The exception are alert messages, which are not tied to game stages or game events but almost always have a significant impact on the surrounding discourse. They are directly addressed with their own research question in Chapter 7. RQ_3 will discuss the communicative purpose of alert messages and their influence on the organization of discourse. Alert messages must be distinguished in ''subscription notifications' and 'donation alert messages' (see 2.3). Both types will be analyzed as exchanges that are initiated by the audience and responded to by the streamer. They appear on the gaming monitor and involve a monetary transaction from the viewer to the streamer. The study of these exchanges will show that alert messages receive preferential interactional treatment. The analysis will demonstrate that streamers quickly orient towards alert messages, even if there are other game events that demand their attention. The chapter will argue that alert messages allow the audience to purchase discourse participation and interaction with the streamer.

It raises the question what the privileged discursive access is used for, because the purpose of alert message must justify the payment. The study of the 150 alert messages and their responses suggests that they are primarily used for interpersonal exchanges and section 7.3 will present the five dominant categories. On the surface, alert messages are very similar to casual conversation or phatic discourse because they discuss topics associated with the development of friendships (Baym, 2010). However, they are also very asymmetrical because the audiences are very interested in the streamers, whereas the streamers mostly react to or acknowledge the message but then quickly proceed to talk about the game. The viewer is much more invested in the streamer than vice versa and streamers appear to react to alert messages mostly because they are paid to do so. This finding is the third important perspective on the organization discourse of live streaming.

The three perspectives will be synthesized in chapter 8 and the study will bring them together under a unifying umbrella. The overall organization of discourse will be expressed through a re-appropriation of the discourse model of ranks and levels by Sinclair & Coulthard

(1975; see fig. 17 in 3.1.1). The new formulation of the model adheres to the principles of the Birmingham school and can account for the discourse of online live streaming. Thereby it answers the original research question and is the most accurate description of the discourse of live streaming to date.

Chapter 5- Move distribution of live streaming

5.0 Initial remarks

Whats the point of the chapter

RQ_1: How are the discursive moves distributed between streamer and audience and does this distribution change with audience size and game's pace?

First, section 5.1 will show the individual distribution in each of the six live streams. Each broadcast has been manually coded with *UAM Corpus Tool* and the software generated its specific distribution of communication between participants. These six figures have been annotated with additional information that are relevant for analysis in the later steps.

The second subsection will present the discourse that is typical for each direction of the discourse flow and how the participants indicate the desired recipient. For example, it will show that the discourse 'from the streamer to individual audience member' has different forms and function from the discourse that operates in the opposite direction 'from the audience member to the streamer'. This will fill out the abstract distribution of numbers within in a live stream with more concrete meanings and it will be understandable what a high share of a certain direction of discourse means for the overall communication of this live stream. A higher share of dialogical discourse between participants will make the overall communication of a broadcast more dialogical, whereas a higher percentage of more monological segments lead to a more monological broadcast.

In a third step, the discourse will be compared across the six broadcasts. The share of each direction of discourse will be compared along the dimension of channel size (large-sized vis-à-vis medium-sized) as well as the dimension of game type. This will allow to interpret if a type of communication such as 'audience member to streamer' is affected by the viewer number or the game's pace. Certain directions of communication increase or decrease in frequency depending, whereas others remain consistent. This means the step-by-step comparison will uncover, which distributions of discourse flow are generally typical for all online live streaming and which patterns are dependent on specific aspects of the broadcast.

Section 5.4 will take this series of findings and synthesize them into a multi-faceted synoptic view that provides a detailed and differentiated look on the analyzed discourse. This description of the dominant distributions of communication is the first of its kind for the video-mediated communication of online live streaming, and it will also support and contextualize the study of exchanges and gameplay commentary in chapter 6.

5.1 The coding results

After the manual coding, *UAM corpus tool* automatically produces a figure, which shows the distribution of move types.

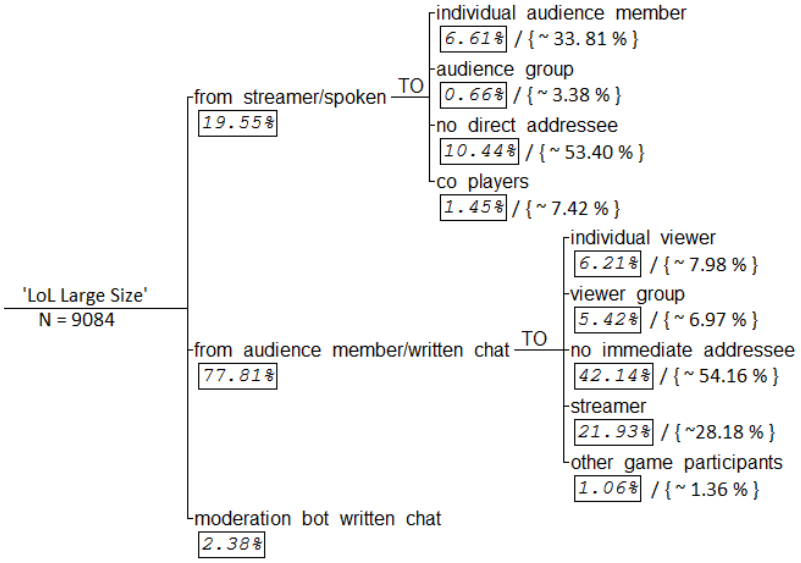


Figure 27 Distribution of moves in large-sized League of Legends broadcast

On the left of the figure, there is the abbreviated name of the game, the size of the channel, and the number of annotated moves of the broadcast. In this case, the figure represents the distribution of moves for the large-sized *League of Legends* (LoL) live stream and a total number of 9084 moves were coded. *UAM Corpus tool* gives you the 'absolute share' in percentages for each category of coding. For example, 19.55 % of all moves in this broadcast were produced by the streamer in the spoken mode, whereas the audience wrote 77.81 % of all moves and 2.38 % were produced by as moderation bot. The second tier of each category further distinguishes the recipients of the move. For example, 6.61% all 9084 moves were coded as 'from streamer/spoken to individual audience member'. This absolute share is unsuitable for meaningful comparisons because it may lead to wrongful conclusions. The absolute share of 'audience member/written chat to individual viewer' is 6.21%, which appears to be very similar to the 6.61% of 'streamer/spoken to individual audience member'. In absolute numbers, both categories are almost equally common, and one could conclude that the streamer is talking to the audience about as much as individual audience member may talk among themselves. However, there are vast discrepancies in the number of spoken and written moves because there are many more audience members and as a result there is a lot more communication by chat participants in general. This will automatically lead to high absolute share of their discursive moves. The streamer utters significantly fewer moves, but his or her talk 'from streamer/spoken to individual audience member' still reaches a total share of 6.6.1%. Therefore, the study manually calculated the relative share for each category and put them into the curly brackets { }. This is a simple task via the mathematical rule of three:

19.55 % of all moves = 100 % of streamer's communication

6.61 % of all moves = Y (the relative share of 'to individual audience member')

The two equations can be solved for Y with $Y = \frac{\mathbf{6.61\% \times 100\ \%}}{\mathbf{19.55\%}} = \sim 33.81\%$

33.81% of the streamer's talk addresses an individual audience member and it is the 2nd most frequent category of spoken discourse behind 'from streamer/spoken to individual audience member' with [10.44%] {~53.40%}.

For 'audience member/written chat to individual viewer', the absolute share is 6.21% but this only accounts for a relative share of $\frac{\mathbf{6.21\% \times 100\%}}{\mathbf{77.81\ \%}}$ = 7.98 % because of the very high number of written chat messages. The relative share is much small, which means that the viewers spend relatively few moves to communicate with one another.

Relative distributions and their frequency are a better tool for comparison because they compare the streamer's spoken discourse and direction of talk to his or her conversational alternatives, which indicates their preferred communicative partners. The same logic applies to the moves by the written chat. Therefore, the analysis in section 5.3 will study and discuss the relative distribution and the relative frequency found in the curly brackets rather than the absolute numbers that are printed inside the round box.

After this explanation of the general layout, the section will continue to show the distribution of moves of the five remaining broadcast before giving examples for each category of coding in 5.2

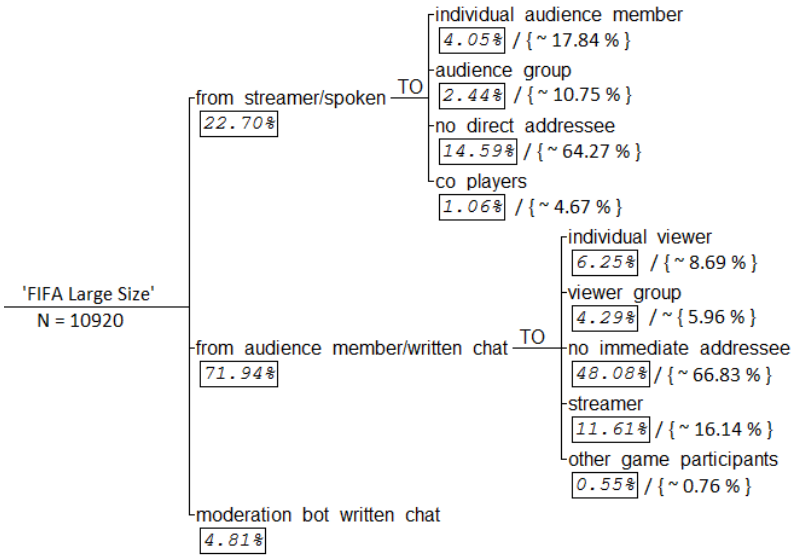


Figure 28 Distribution of moves in large-sized FIFA 2015 broadcast

The figure shows the distribution of moves in the large-sized *FIFA 2015* broadcast, whose, share of talk and chat are very similar to the large-sized *League of Legends* stream.

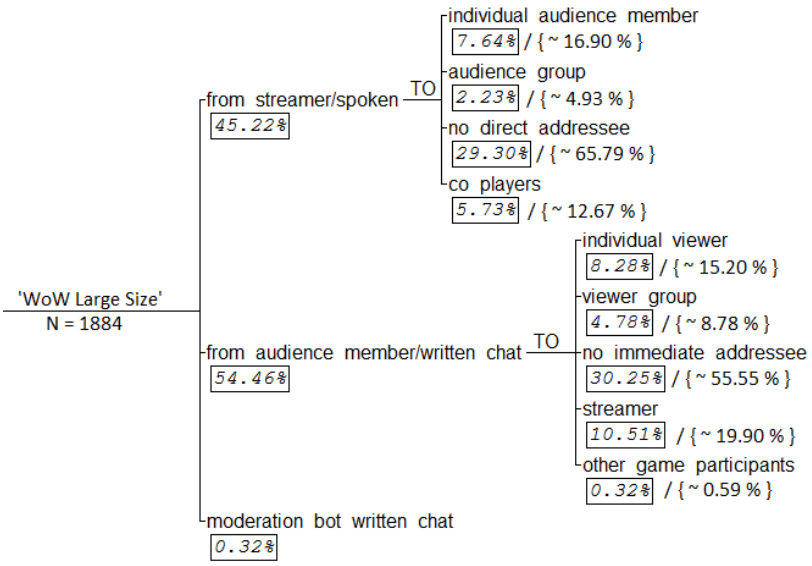


Figure 29 Distribution of moves in large-sized World of Warcraft broadcast

The large-sized *World of Warcraft* broadcast features the lowest number of moves (N=1884) and has patterns that are comparable to medium-sized broadcasts. This finding will be discussed in section 5.4 and it will be speculated that the streamer talked very little, which may impact the number of chat messages.

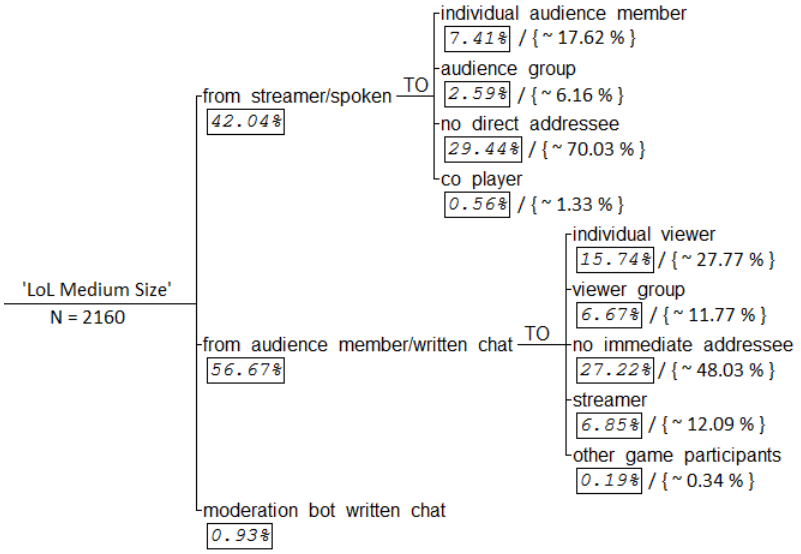


Figure 30 Distribution of moves in medium-sized League of Legends broadcast

For medium-sized live streams there was no comparable outlier and all three broadcasts have a very similar distribution of moves. Their number of total moves varied from 2160 as in the case of the medium-sized *League of Legends* broadcast to 3146 moves for the medium-sized *FIFA*

2015 stream. The number of total moves is lower in medium-sized broadcasts due to the significantly smaller number of chat participants.

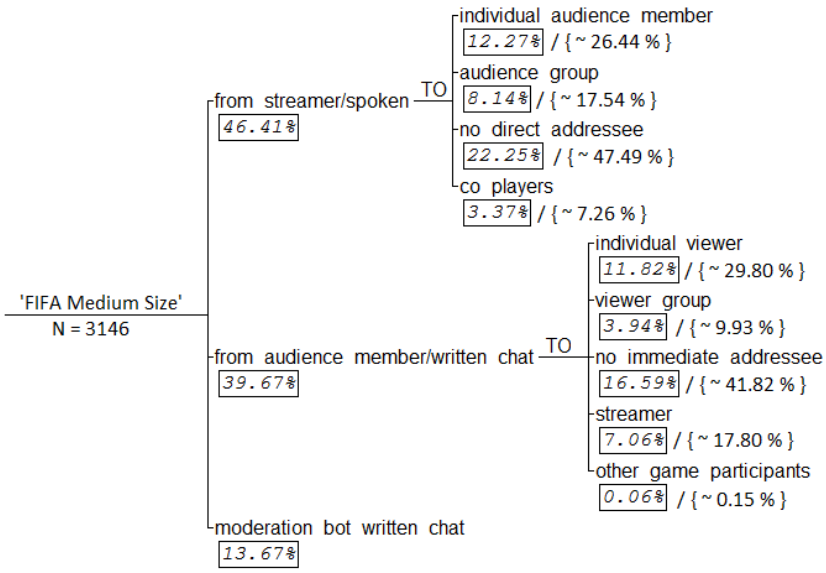


Figure 31 Distribution of moves in medium-sized FIFA 2015 broadcast

The next figure shows the distribution of moves for the medium-sized *FIFA 2015* broadcast. It is the most active live stream of its size. However, this will be put into perspective in the discussion sections 5.2 and 5.3, when the chapter addresses moderation bot messages because this broadcast has by far the largest number of moderation bot messages.

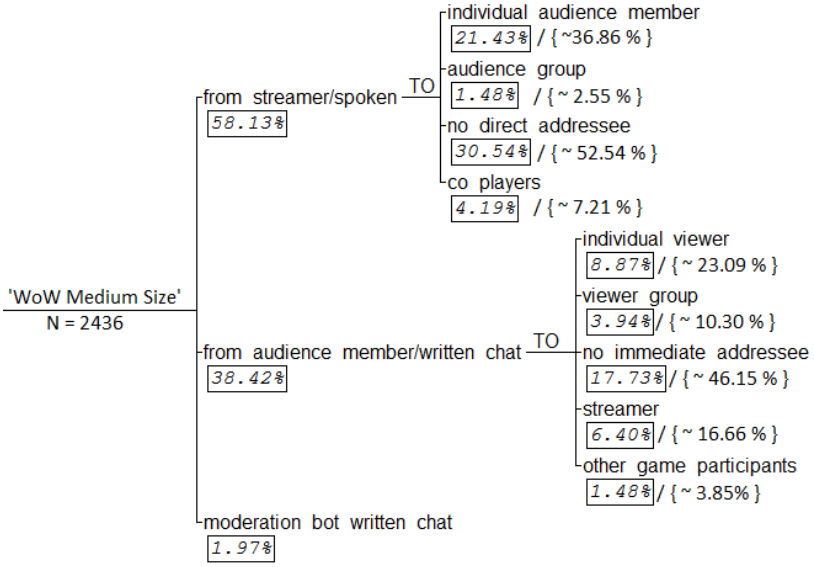


Figure 32 Distribution of moves in medium-sized World of Warcraft broadcast

The last figure is of the medium-sized *World of Warcraft broadcast*. It features the largest share of talk by a streamer in relation to the total number of moves N = 2436 and it is the only broadcast, where there was more talk than there were chat messages.

The six figures show the patterns of communication in each of the selected channels. They show who is communicating to whom and how often. However, they do not indicate what the communication is about or explain the form that it takes. Therefore, the next section will

contextualize the numbers by providing examples for each of the categories.

5.2 Examples for each coding category

It is important to remember that the criterion for the coding of moves was their sender and recipient configuration. The coding categories emerged from the data and each coding decision is based on linguistic evidence found in the transcript. The most common and explicit indicators have been terms of address and textual cohesion between moves (Halliday & Hasan, 1976). During the coding process, it appeared that the directions of communication may correspond with certain communicative functions. This means participants may talk to different people for different purposes. This point was not statisitically tested and there was no further research into this issue. Therefore, I do not want not argue that this pairing of direction of communication and its purpose could withstand the scrutiny of quantitative research but it was nevertheless a very salient observation, which will be demonstrated via the upcoming examples.

Streamer/spoken to audience group

Messages by the streamer to the 'audience as a group' were identified through terms of address in the plural such as 'you', 'you all', 'guys', 'you guys', 'chat' or an inclusive 'we'. Notably, the terms of address often have a gender bias and label audience members as males. There were no instances of 'hey girls / gals', even though there were chat participants with female names. It supports prior arguments that live streaming on Twitch is often a male-gendered domain (cf. Anykey, 2015, 2016).

(1) Example – 'you guys'

```
you guys are asking me to play all kinds of fucking.
sort of troll champs.
```

The example is taken from *League of Legends* and consists of a single move by the streamer and serves as a response to an audience initiation. Such responses to a collective audience only occur if the audience is talking about similar topics or is asking similar requests. They have a low prevalence (see 5.3), which indicates that it is rare for the chat to speak with a unified voice.

(2) Example – 'another question for you guys?'

```
another question for you guys.
I was thinking about doing the player road to glory around pink Reus.
But I did wanne see what you guys thought.
so we may do a strawpoll.
```

The example is taken from *FIFA 2015* and the streamer utters a series of four moves. Together, they form an initiation by the streamer to the audience viewers. He invites them to give their opinion on an upcoming decision in the live stream.

Utterances coded as 'streamer/spoken to audience group' are almost always part of a larger exchange between streamer and viewers. They are an explicit one-to-many dialogue, but they are also very rare as it appears difficult to have a genuine group conversation.

Streamer/spoken to individual audience member

The coding for this category has been very surprising because streamers rarely use terms of address when they talk to an individual audience member. They could mention the username of the chat participant, but this rarely occurred. In ordinary streamer-to-viewer dialogue, streamers tend to utter a repetition of the viewers message.

(3) Example - game-related responses

```
"do I like Aubameyang"?
I've never I never bought his tool card.
but his loan card.
was great.
his loan card was great for like a couple of games.
and then he fell off.
```

The example is taken from *FIFA 2015* and begins with such a repetition. Repetitions of this type will be discussed as 'topicalizers' in section 6.1. They are produced at the start of a spoken response before the streamer is answering the question by the viewer. This pattern of exchanges is very robust and almost all moves that were coded as 'streamer/spoken to individual audience member' are part of an exchange that has been initiated by a viewer. The topics under discussion may vary. In this example, the streamer is asked to evaluate the virtual game character '`Aubameyang`' and provides a detailed answer that relies on insider knowledge about game properties such as '`tool cards`' or '`loan cards`'. Game-related talk is the content of most exchange, but it is also appropriate for the audience to ask questions about certain aspects of the streamer's private life.

(4) Example – conversational responses

```
camero are you fucking serious?
yeah got a 1967 camero.
have a little money saved up for it so.
I think it's gonna be like 8 to 10 grand.
```

The fourth example is taken from *World of Warcraft* and the topic is completely unrelated to the live streamed game. The streamer is asked about his recent car purchase and gives a very casual and colloquial response. The important aspect is the structural similarity to the previous example. The response has no term of address and starts with a spoken repetition of the question before providing the answer. It means that the streamer is very much reacting to the audience but it is extremely rare that they take initiative to talk to people. Nevertheless 'streamer to individual audience member' is the second most common coding for the spoken discourse, which means they respond to many questions in quick succession. This point will be further explored in chapter 6 as 'exchange chaining'.

Streamer/spoken to no direct addressee

Moves coded as 'streamer to no direct addressee' have no explicit referent and the discourse does not interact with prior messages by the audience. There is no prospections and turn-taking with a discourse partner. This category covers the two distinct discourse practices of reporting and commenting, which are discussed in detail in section 6.2.

(5) Example of reporting – 'their sidelaners were roaming'

```
we got both side lane turrets.
because their sidelaners were roaming.
like Blitz and Hecarim were roaming mid.
they BACKED UP.
because they were scared as shit of my ult.
```

The example is taken from *League of Legends*. Reporting is elicited by a single game event but usually goes beyond a simple narration of the gameplay. It discusses the event through a series of coherent and connected moves that give important background information, explain the decision making of the players and the consequence in the game. The content of these moves is game-specific and filled with jargon.

(6) Example for commenting - ‘oh shit’

```
nice.
:
the fuck kinda pass was that?
:
oh shit.
:
come on.
:
nice.
alright whatever.
```

This excerpt is taken from *FIFA 2015* and shows a segment of commenting. It consists of a series of moves that do not combine into a larger unit of communication. Commenting is a communicative practice that can only be understood in reference to the transcribed game event. The moves are response cries to the game event that give an emotional commentary and contain almost no jargon (Goffman, 1978).

Commenting and reporting have in common that they are a monological type of discourse. Moves that have been coded as ‘streamer/spoken to no direct addressee’ do not engage with chat messages explicitly and are produced when the streamer is primarily oriented towards the game. ‘Streamer/spoken to no direct addressee’ is consistently the most frequent coding across all six broadcasts, which means that significant parts of the streamer’s discourse are not conversational.

Streamer/spoken to co-players

Streamers sometimes produce moves that are addressing co-players or even game avatars (cf. Baldauf-Quilliatre & de Carjaval, 2016). The term of address is usual the nickname of the co-player or the avatar name. This practice is interesting because the nomimal addressee cannot hear the message. On the literal level, the discourse appears to talk to a person, but in practice it is talking about a player or game avatar.

(7) Example – ‘Come on team mate’

```
Come on Esubar do it.
:
nice.
good job dude.
```

The example is taken from *World of Warcraft* arena and the streamer is cheering for his team mate ‘Esubar’, who is still in combat with the opponents. After the victory, he congratulates him for his performance in the game.

(8) Example – ‘please don’t kill me’

```
"please don't kill me".
"please don't kill me".
OH^ FUCK^ OFF^.
"ODDONE".
```

In example 8, the streamer is begging an opposing player ‘Oddone’ not to kill him in *League of Legends*. Despite the serious wording, the utterances are clearly marked as playful by the shift in voice quality and the laughter during the exaggerated ‘OH FUCK OFF’.

(9) Example – ‘What the fuck’

```
NO:
WHAT THE FUCK LEO MESSI.
TEAM OF THE YEAR MESSI.
WHAT THE FUCK ARE YOU DOING?
HOW DO YOU MISS THAT?
MESSI.
```

The streamer misses a goal opportunity in *FIFA 2015* with the game avatar ‘Messi’. He loudly and vividly complains that the avatar made the mistake and is at fault, although the streamer is controlling the virtual character. The discourse in all three examples is dialogical on the surface.

‘Come on Esubar’ and ‘please don’t kill me’ are commands and ‘how do you miss that?’ is a question that addresses a nominal recipient. However, these utterances are more similar to commenting because they are always linked to a game event, where the addressed co-player carries a lot of responsibility. The discourse is not a dialog with the co-players but an exaggerated comment about them. They are the least common coding for the spoken discourse, which could be due to their staged nature.

Audience member/written chat to viewer group

Discourse coded ‘from audience member/written chat to viewer group’ contain explicit terms of address towards the chat such as ‘guys’, ‘boys’ or ‘chat’. Alternatively, the message mentions the streamer in third person and is talking about him or her. Usually, the messages invite other chat participants to respond to the initiation.

(10) Example ‘how old is the streamer’

```
<hakkesho> how old is bjersen [streamer]
:
<jujumanji> 19
<bandicosplay> bjerg is 19 :)
<sooted> bjerg is 19
```

The chat participant is asking about the age of the streamer and despite a minor spelling error other readers understand the request for information because such questions are very common. They interpret the question as being directed at everyone in the chat and quickly respond by giving the correct answer.

(11) Example - ‘his record guys?’

```
<danb1990> whats his record guys?
```

In this example from *FIFA 2015*, the chat participant is addressing the audience members as ‘guys’ and is asking about ‘his record’, which in this context refers to the number of wins and losses of the broadcaster.

Messages ‘from audience member/written chat to viewer group’ are almost exclusively initiations that ask a question and initiate an exchange with other viewers. Their counterpart are messages that have been coded as ‘audience member/written chat to individual viewer’.

Audience member/written chat to individual viewer

Moves coded from ‘audience member/written chat to individual viewer’ are responses to the previous category of initiations. The appropriate coding can be identified through the very common chat convention ‘@-username’ or through explicit linguistic references in the response.

(12) Example ‘how does nick do that’

```
<H2hfifa> Guys how does nick do that celebration with holding his hands up
<Lewez24> @h2hfifa by scoring
<Marzui19> @h2hfifa right stick up
```

In example 12, an audience member is initiating a question to the viewer group. The question is about an event in the live stream, where the streamer ‘nick’ has performed a special celebration after scoring a soccer goal. The question receives two direct responses and ‘@-

username' is fronted at the beginning of the message. It clearly marks him or her as the addressee of the message. The written responses are very short and provide all the necessary information in a single message in a very elliptical format, which is typical for speech like chat communication (cf. Werry, 1996).

Most dialogical exchanges in the chat consist of an initiation 'to the viewer group' and responses 'to the individual viewer'. Most questions are asked to the general public in the chat room aRather than a specific participant. Arguably, this indicates that chat participants are not very familiar with one another. They tend to ask the group as a whole because there is no suitable chat participant that could be asked directly. In turn, the response can address the initiator of the question because s/he is singled out as a potential conversational partner. The intra-chat communication is very anonymous and there is very little phatic communication between participants. It may indicate that chat participants are barely interested in one another and are oriented towards the streamer or the game

Audience member/written chat to streamer

The strong interest in the streamer is also visible in the high frequency of discourse 'from audience member/written chat to streamer'. The coding was the second or third most common category in the chat communication (see 5.3), and mostly consists of initiations that ask questions to the streamer or request him or her to perform a certain action. The terms of address can very between the username of the streamer or even their real name.

(13) Example – '@username'

```
<Mayxlol> @spb_89 Why no deadly poison?
```

Example 13 is taken from *World of Warcraft*. The move is an initiation by a chat participant that asks the streamer about a specific choice in the game. Notably, such initiations front the term of address and follow the @-username convention to mark the streamer as the intended addressee.

(14) Example – 'Eric play alone'

```
<Nvidia807> Eric play alone or meteos duo
```

In the 14th example, the streamer's first name is used as the term of address and it is again fronted at the beginning of the message.

The discourse on *Twitch* is very direct and "bald on record" (Levinson et al., 1987), which indicates a low social distance between streamer and viewers. The request is framed as

a command although the initiator has no inherent authority to tell the streamer what to do. Nevertheless, it is acceptable to give direct commands such as '`play alone or meteos duo`' without hedging. If the streamer is not interested, s/he will usually just ignore the message and will not treat it as a breach of etiquette (see 6.1).

Audience member/written chat to no immediate addressee

Discourse coded as from 'audience member/written chat to no immediate addresse' has no prospection and does not combine into exchanges with the streamer or audience members.

(15) Example - 'rekt'

```
<thecookieshow> rekt <abby315> WHAT
lol
<thelasttheory> rekt
<twizzler76> nice
<rinsbac> LOL
<johnrcrv> R E K T
<sam14146> rekt
<dtkaze> so good lol..
<painjames> GOD
```

In some respects, these moves are comparable to the streamer's spoken commenting. The excerpt is taken for *League of Legends* but is just as likely to occur in *World of Warcraft.* These moves are exophoric, refer to a game event and evaluate it in an emotional tone that is often exaggerated or playful. Across all six broadcasts, it is by far the most common type of discourse in the chat and it is presumably this kind of communication that several authors have in mind when they label *Twitch*'s chat as cacophony, noise or illegible waterfall of text (cf. Hamilton et al., 2014; Olejniczak, 2015; Nematzadeh et al. 2016.).

This study objects to these negative labels and the associated argument that these messages are disruption for an otherwise orderly conversation (cf. Hamilton, 2014). Instead, the study will show that the occurrence of these messages is very regular and that they are an essential part of the overall organization of discourse (see 6.2).

Audience member/written chat to other game participant

The category 'audience member/written chat to other game participant' was applied to audience messages that explicitly address other players in the game with the streamer. Usually the chat participants will refer to them via the username or the name of the avatar that is controlled

(16) Example – ‘Yasuo stfu’

<metalfrider> Yasuo [game participant] stfu Bjerg [streamer] is 1000 times better than you kid !!!

<caliraised93> YASUO [game participant] WE KNOW WHO YOU ARE

The example is taken from *League of Legends* and shows two chat participants that write messages to another game participant. The term of address ‘Yasuo’ is fronted before the main body of the very emotive and disparaging messages.

These messages are not part of a genuine dialogue because *Twitch’s* chat is not visible to the co-player and they are completely unaware of the message. The purpose of such messages is to signal support for the streamer by criticizing the opponent.

Moderation bot/written chat

The last category is produced by moderation bots. Their messages are triggered by pre-defined parameters. Since they are not sentient, it cannot be said that they address someone in the sense that human communication addresses interlocutors. They give background or promotional information, police misbehavior and are used for chat games.

(17) Example – ‘promotional and background information’

<Moobot> this is nicks twitter! https://twitter.com/Nick28T

<Moobot> For XBOX or PSN coins visit http://justfifacoins.com/ and use NICK at checkout

The computer script can periodically repeat a fixed message that has been designed by the streamer. In this example, it features advertisement of the streamer’s *Twitter* as well as an affiliate code to a third-party website. Affiliate codes are a commission system that is very common in online marketing. If viewers buy ‘PSN coins‘ from the website with the code, the streamer will receive a commission

(18) Example – ‘chat moderation’

<pdealla> https://www.twitch.tv/pdealla

BAN: pdealla

Chat bots produce a short message when they ban users for misbehavior. In this case, a viewer posted his or her own live stream address. This practice is frowned upon because it is seen as an attempt to poach viewers to a different channel.

(19) Example – ‘chat games’

```
<RevloBot> "Win" was picked as the winning option! 13 winners split 4995 pugbucks in winnings.
<RevloBot> steyn_king has 121 pugbucks
<RevloBot> danb1990 has 31 pugbucks
<RevloBot> rushxk has 695 pugbucks
```

Chat games are rare and only one broadcast featured them. They have a significant impact on the number of chat messages produced by the moderation bot. In this case, the chat game is a type of gambling system and users can bet a fictional currency on outcome of the broadcasted games. If they bet on the correct outcome of a match, the bot produces a ‘payout’. Chat games are a very idiosyncratic feature of a very small number of channels and the study will not address them much further.

5.3 Comparison across broadcasts

The study will compare the share of each category across broadcasts and the format of the tables is designed to facilitate the comparison.

From moderation bot/written chat

Size / Game	Large-sized		Medium-sized	
	Share in %	Frequency	Share in %	Frequency
World of Warcraft	0.32	3rd out of 3	1.97	3rd out of 3
FIFA	4.81	3rd out of 3	13.67	3rd out of 3
League of Legends	2.38	3rd out of 3	0.93	3rd out of 3

Table 14 Moves from Moderation Bot/Written Chat

In this case, the table presents the share of all moves coded as ‘from moderation bot/written chat’. It is possible to compare these shares across games by going from top to bottom, or to compare the shares across audience sizes by going from left to right. Frequency describes how common a coding is in comparison to its alternatives. For example, ‘moderation bot/written chat’ has been compared to ‘streamer/spoken’ and ‘audience member/written chat’. ‘Moderation bot/written chat’ is the least common coding and is therefore consistently ranked as 3rd out of 3 options. In some cases, color coding is used to visually highlight sections in the table. In table 14, it is the medium-sized *FIFA 2015* live stream, which shows a significant increase in moderation bot messages due to the lottery-like chat game. For the other broadcast, the share of moderation bot messages was in the low single digits. The main function of bot messages is to support the main interaction between streamers and viewers and aside from the policing misbehavior (see 6.3.2), moderation bots have little to no discernable effect on the communication.

From streamer/spoken

Size / Game	Large-sized		Medium-sized	
	Share in %	Frequency	Share in %	Frequency
World of Warcraft	45.22	2nd out of 3	58.13	1st out of 3
FIFA	22.70	2nd out of 3	46.41	1st out of 3
League of Legends	19.55	2nd out of 3	42.04	2st out of 3

Table 15 Moves from Streamer/Spoken

The next table shows the share of the streamer's spoken moves in each of the six broadcasts. It demonstrates that in every channel the streamer produces remarkable amounts of talk, even if they are outnumbered by their viewers. The average streamer talks much more than the average chat participant is writing. In the medium-sized broadcasts with 100 – 150 viewers, there is almost a parity between the number of spoken and written moves. In the large-sized broadcasts of *FIFA* and *League of Legends*, the higher number of chat participants (>1000) leads to a shift and the streamers' share decreases to 19.55% and 22.70%.

The exception is the large-sized *World of Warcraft* broadcast, which has a share of spoken discourse of 45.22 % much closer to the medium-sized channels. Intuitively, one might assume that this *World of Warcraft* streamer is extremely active and talking very much. However, figure 29 on page 94 shows that the opposite is true. The large-sized *World of Warcraft* stream features the lowest number of moves in the data set (N=1884). This number is more comparable to that of the medium sized channels because the participants produce very little discourse. The study cannot explain this outlier with certainty, but the finding suggests that the liveliness of the streamer is another mediating factor for the communicative activity in the chat (cf. Hamilton et al, 2014). Even a large-sized broadcast can have a rather inactive chat if the streamer is not facilitating the interaction.

From audience member/written chat

Size / Game	Large-sized		Medium-sized	
	Share in %	Frequency	Share in %	Frequency
World of Warcraft	54.46	1st out of 3	38.42	2nd out of 3
FIFA	71.94	1st out of 3	39.67	2nd out of 3
League of Legends	77.81	1st out of 3	56.67	1st out of 3

Table 16 Moves from Audience Member/Written Chat

This table shows the distribution numbers for the "audience member/written chat messages". In medium-sized broadcasts, the spoken and written communication have an almost equal share, whereas the chat messages in the large-sized broadcasts substantially outnumber the talk. Given the numerical asymmetry between streamer and audience, this finding is not surprising, but it will be important to understand the patterns of exchanges between participants in chapter

6. For now, the study will increase the delicacy and take a closer look at the spoken discourse.

From streamer/spoken to…

The coding revealed that the streamer produces language in four different directions. They speak to 1) an individual audience member, 2) the viewer group, 3) no direct addressee, or 4) the co-player.

From streamer/spoken to individual audience member

Size / Game	Large-sized		Medium-sized	
	Share in %	Frequency	Share in %	Frequency
World of Warcraft	16.90	2nd out of 4	36.86	2nd out of 4
FIFA	17.84	2nd out of 4	26.44	2nd out of 4
League of Legends	33.81	2rd out of 4	17.62	2nd out of 4

Table 17 Moves from streamer/spoken to individual audience member

The table shows the share of discourse that is produced 'from streamer/spoken to an individual audience member'. This type of talk mostly describes spoken responses by the streamer to the written initiations by the audience. It is consistently the second most frequent direction of talk, regardless of the type of game or the size of the audience. Even if the number of viewers is very large, streamers do not become more likely to talk to them as a collective group. There is a preference for a dyadic dialogue with the audience. The viewers produce many initiations (questions or request) and streamers select from these questions and respond to them one by one. In turn, this leads to a high share of discourse 'from streamer/spoken to an individual audience member' in medium and large-sized channels.

From streamer/spoken to audience group

Size / Game	Large-sized		Medium-sized	
	Share in %	Frequency	Share in %	Frequency
World of Warcraft	4.93	4th out of 4	2.55	4th out of 4
FIFA	10.75	3rd out of 4	17.54	3rd out of 4
League of Legends	3.38	4th out of 4	6.16	3rd out of 4

Table 18 Moves from streamer/spoken to audience group

Table 18 shows the discourse from 'streamer/spoken to the audience group'. Its low share re-affirms the previous argument and talk to the audience group is rare and often much lower than responses to an individual viewer. Overall, it is a dispreferred option and only used in exceptional cases, for example, when the streamer is requesting information or actions from the audience. This low prevalence of talk to the viewers as a group is very different from the computer-mediated communication on *YouTube*. On *YouTube*, video hosts tend to address

their collective audience of future listeners (cf. Frobenius, 2014). Arguably, this is because they have to initate the conversation and the viewers can only respond after the upload. On *Twitch*, the viewers can go first, initiate written questions and the streamers adjust their communicative behavior and provide more individual responses. The coding from 'streamer/spoken to the audience group' also has two outliers with the medium-sized and large-sized broadcast of *FIFA 2015.* The study revisited the transcripts of the broadcasts and found that both streamers talk more about their community of viewers and their shared undertaking on *Twitch*. This talk about community is reflected in more talk towards the audience as a group. Unfortunately, the study does not have sufficient evidence to argue whether this is due to the selected streamers or if it is property of the game. It is possible that the streamer's soccer club in *FIFA 2015* provides a shared identity that the streamer and the viewers can rally behind.

From streamer/spoken to no direct addressee

Size / Game	Large-sized		Medium-sized	
	Share in %	Frequency	Share in %	Frequency
World of Warcraft	65.79	1st out of 4	52.54	1st out of 4
FIFA	64.27	1st out of 4	47.49	1st out of 4
League of Legends	54.16	1st out of 4	70.03	1st out of 4

Table 19 Moves from streamer/spoken to no direct addressee

Moves coded as 'from streamer/spoken to no direct addressee' consists of game-related commenting and reporting (see 5.2; 6.2). It is the most frequent talk across all six channels regardless of game type or channel size. There is a considerable margin between the share of talk to 'no direct addressee' and the second most frequent type of talk to an 'individual audience member". Spoken discourse without a direct addressee makes up the largest share of the streamer's talk and more than half of their speech is monological and about the game. This means that live streaming is not a conversation between streamer and viewers in the linguistic sense. It is a form of computer-mediated communication that is, at times, conversational and at other times behaves differently. This organization in dialogical and monological talk is further explored in 6.2, when the study will argue that the presence or absence of game events are an important factor for the organization of discourse in stretches of monologue and dialogue.

From streamer/spoken to co-players

Size / Game	Large-sized		Medium-sized	
	Share in %	Frequency	Share in %	Frequency
World of Warcraft	12.67	3rd out of 4	7.21	3rd out of 4
FIFA	4.67	4th out of 4	7.26	4th out of 4
League of Legends	7.42	3rd out of 4	1.34	4th out of 4

Table 20 Moves from streamer/spoken to co-players

The talk 'to the co-players' is an entertaining performance by the streamer (cf. Fernández-Vara, 2009; Nyland, 2015) and puts praise or blame on fellow team mates or opponents. Its share on the streamer's talk is very low and it is rather infrequent. The study could not identify any consistent patterns for this type of talk even as it revisited the transcripts of the two outliers, the large-sized *World of Warcraft* broadcast and the medium-sized *League of Legends* stream. The most likely explanation for their deviance is individual differences in the personality of the live streamer. The *League of Legends* streamer often commented on negative outcomes as 'unlucky' and put no blame or praise on other participants. The streamer of *World of Warcraft* describes his co-players as 'noobs[32]' or monkeys and these insults were often framed as a form of direct address as in '`you noob`' or '`you are playing like a monkey`'. These differences in individual behavior possibly explain the different amount of talk to co-players.

Other than that, the distributions of moves are very similar across the six broadcasts. This indicates that the streamer's patterns of talk mostly stay constant even if the game or the audience size changes. When they engage in dialogue, there is a dyadic preference and they tend to respond to written initiations by the viewers. Talk to the viewers as a group is much rarer even in large-sized live streams. The most frequent type of talk has no addressee at all, which means that the largest share of streamer communication is monological and about the game.

In the next part, the study will compare the distribution figures in the chat.

<u>From audience member/written chat to…</u>

Chat messages are produced in five directions. There is communication from an audience member to 1) another game participant, 2) no immediate addressee, 3) the viewer group, 4) an individual viewer, or 5) the streamer.

From audience member/written chat to other game participant

Game \ Size	Large-sized		Medium-sized	
	Share in %	Frequency	Share in %	Frequency
World of Warcraft	0.59	5th out of 5	3.85	5th out of 5
FIFA	0.76	5th out of 5	0.15	5th out of 5
League of Legends	1.36	5th out of 5	0.34	5th out of 5

Table 21 Moves from audience member/written chat to other game participant

Throughout all six broadcasts, the chat's communication to another game participant was the least frequent. It is not intended as genuine dialogue and usually is an indirect signal of support for the broadcaster. Since there are also other practices to signal one's appreciation of the streamer such as compliments, the chat messages 'to other game participants' offers little benefit and may be perceived as too artificial. This could explain their low prevalence in live streaming

From audience member/written chat to no immediate addressee

Game \ Size	Large-sized		Medium-sized	
	Share in %	Frequency	Share in %	Frequency
World of Warcraft	55.55	1st out of 5	46.15	1st out of 5
FIFA	66.83	1st out of 5	41.82	1st out of 5
League of Legends	54.16	1st out of 5	48.03	1st out of 5

Table 22 Moves from audience member/written chat to no immediate addressee

Communication that was coded 'from audience member/written chat to no immediate addressee' was the most frequent type of written discourse across all broadcasts. Just like the streamer's talk, the largest share of the audience's discourse is also not dialogical. Instead they produce many emphatic moves in reaction to the game events on screen. It is another indication that the linguistic model of conversation is not a perfect description and needs further adjustment to accurately account for the organization of discourse. In 6.2, the analysis will show that segments of more monological and more dialogical discourse alternate in a systematic manner

From audience member/written chat to viewer group

Game \ Size	Large-sized		Medium-sized	
	Share in %	Frequency	Share in %	Frequency
World of Warcraft	8.78	4th out of 5	10.30	4th out of 5
FIFA	5.96	4th out of 5	9.93	4th out of 5
League of Legends	6.97	4th out of 5	11.77	4th out of 5

Table 23 Moves from audience member/written chat to viewer group

The discourse from 'audience member/written chat to viewer group' is consistently the second least frequent direction of chat communication, which can be explained with its usual purpose. The discourse tends to contain initiations to other chat participants that ask streamer-related or game-related background information. They are a small set topics that viewers discuss with one another and most participants treat these questions as an interactional 'backstage' (cf. Kendon, 1992; Rosenbaun et al, 2016b). They cover less important information such as the streamer's age, or what s/he is doing.

From audience member/written chat to individual viewer

Size / Game	Large-sized		Medium-sized	
	Share in %	Frequency	Share in %	Frequency
World of Warcraft	15.20	3rd out of 5	23.09	2nd out of 5
FIFA	8.69	3rd out of 5	29.80	2nd out of 5
League of Legends	7.98	3rd out of 5	27.77	2nd out of 5

Table 24 Move from audience member/written chat to individual viewer

Chat messages 'from audience member to an individual viewer' usually serve as responses to initiations of the category 'written chat to viewer group'. Notably, in every broadcast, there is more 'written chat to individual viewer' than there are 'written messages to the viewer group'. In turn, this means that there are more responses than there are initiations to respond to. A written initiation is likely to receive several responses by chat participants. This finding will support the description of mono-modal exchanges in the chat in 6.1. They are a relatively rare phenomenon because there are few initiations, but if there is an initiating first turn, then there will be several responses.

The table also suggests that discourse from 'written chat to individual viewers' is more common in medium-sized channels than it is in large-sized chats. In medium-sized channels, it is the second most frequent direction of chat communication, whereas it is the third most frequent for large-sized chats. A possible explanation is that in the smaller channels, the chat participants are more likely to continue a transaction about a given topic. After a first initial exchange, other participants may produce follow-ups or continue the ongoing discussion with extended exchanges. In the larger-sized channels, it is very likely that the transaction finishes after a single exchange and the communication in the broadcast switches to another topic.

From audience member/written chat to streamer

Size / Game	Large-sized		Medium-sized	
	Share in %	Frequency	Share in %	Frequency
World of Warcraft	19.90	2nd out of 5	16.66	3rd out of 5
FIFA	16.14	2nd out of 5	17.80	3rd out of 5
League of Legends	28.18	2nd out of 5	12.09	3rd out of 5

Table 25 Move from audience member/written chat to streamer

The communication from 'audience member/written chat to streamer' are often initiations in the form of questions or request. They combine with 'streamer/spoken to individual audience member' to form cross-modal exchanges (see 6.1). Within the large-sized channels, the audience initiations towards the streamer is the second most frequent direction of chat, whereas it is the third most frequent direction of chat for the medium-sized broadcasts. This suggests that in the large-sized streams the communication to the streamer is more desirable than communication with fellow chat participants. In large-sized channels, the streamers are already new media celebrities and well-known on *Twitch.* Therefore, chat participants are very eager to communicate with them and are also willing to pay for brief moments of interaction (see chapter 7; Sjoblöm & Hamari, 2015). Medium-sized channels have more interaction between viewers and the streamer is still perceived as an ordinary gamer. In this sense, the largesized channels are not necessarily less dialogical, but their dialogue has a different preferential orientation. The more viewers there are, the more they funnel the discourse towards the single streamer.

5.4 Concluding thoughts on the discursive moves

This chapter set out to describe the discourse of online live streaming by looking at the distribution of moves between participants. It provided a more nuanced perspective that replaces binary and essentialist labels such as 'waterfall of text' or 'cacophony' with a balanced description and analysis. The literature review in chapter 2 suggested that the audience-size and type of game could have a significant effect on the patterns of communication. Slower-paced games may allow for more dialogical discourse between streamer and viewer because the streamer has time to re-orient towards the second monitor, read the chat and respond. Fast-paced games could discourage such cross-modal communication because streamers may find it difficulty to turn away from the game (see 2.3). Similarly, the literature review also suggested that medium-sized live streams could contain more elements of dialogue because there is less 'disruptive' communication and more 'orderly conversation' (cf. Hamilton et al., 2014).

Within this study, the game type and game pace had no immediately measurable effect for the distribution of moves. The faster-paced games did not feature more monologue (or less dialogue) than slower paced games. This observation is strengthened by the qualitative study of exchanges, reporting and commenting in the following chapter. Across different games, the patterns of communication are very comparable, and the discourse patterns appear to be very has very consistent regardless of the game that is played. This means that the description of *Twitch's* discourse is generalizable for games of different types and it is not necessary to make further distinctions between them.

As for the impact of the audience size, the study indicates that it has no effect on the communication of the streamer. Their talk to an individual audience viewer is always more frequent than their talk to the audience group. They have a consistent dyadic preference that addresses audience members one-by-one even if the viewer number is increasing and it becomes very difficult to talk to everyone. Instead of talking to the viewers as a group, they continue dyadic exchanges, which they chain in a series. This exchange chaining leads to minimal dialogical pairs between the streamer and his or her viewers.

For the written chat, there are nuanced differences between medium-sized and large-sized channels. In medium-sized channels, there is more communication from one chat participant to another. A qualitative look suggests that this is mostly due longer exchanges between the chat members (see 6.1). In medium-sized broadcasts, the production of new messages is slower, and messages have a longer uptime in the visual field of the chat window. There is sufficient time to read the text, think of a suitable response and to formulate it in the dedicated chat box. The difference in chat pacing makes it possible chat participants to have a coherent discussion within the chat.

In the larger-sized broadcasts, there are many more written chat messages than in the medium-sized channels. Naturally, this is due to the increase in the number of participants. However, there are also changes in the communicative patterns. Communication towards the streamer becomes more frequent and there is less chat among viewers. The streamer is the preferred discourse partner for the audience. S/he is more famous than his or her counterpart in the medium-sized broadcasts. Viewers use the chat to produce written initiations, often questions, towards the streamer. Many of these questions fail to receive a response because it is impossible for the streamer to answer them all. The result is many ignored exchanges (see 6.1), which are still dialogical in their intent. Therefore, they should not be evaluated as less meaningful or defunct forms of communication comparable to a 'waterfall of text' or cacophony. The distribution of moves is mostly similar in medium-sized and large-sized

broadcasts, which indicates that the organization of discourse of live streaming is nearly identical in the two categories.

They both contain conversational segments but also substantial amounts of monological talk or chat. The monological moves combine to chunks of reporting or commenting on the game (see 6.2). Therefore, it is inaccurate to describe the discourse of live streaming solely as a conversation or narration. However, to go beyond this initial observation, the study had to ask the second research question RQ_2 and describe in detail how discursive moves are combined into larger units of discourse and how is this process is influenced by the unfolding gameplay. The second research question will be addressed in chapter six and during its discussion, the findings of this chapter will help to contextualize and understand the structure of exchange.

Chapter 6 - Exchanges, commentary and gameplay

6.0 Introduction to the Chapter

Discourse is constituted from smaller units that build into larger ones. This chapter will describe how the different sets of moves that were discussed in the 5th chapter combine to larger units of discourse such as dialogical exchanges, segments monological reporting or commenting and finally to complete transactions (Eggins & Slade, 1997; Francis & Hunston, 1992). The discussion will begin with the smaller units and gradually move to towards the description of the larger patterns of the interaction. A critical question for such an argument is the representativeness of the selected examples for the overall data set. As a mostly qualitative researcher, I began the study with several close readings of the transcripts. In each reading, I highlighted notable exchanges and game-related communication in the transcript. To some degree, this was a subjective process, but it was informed by the prior research on *Twitch* with its distinction in conversation and narration as well as the analytical framework of the Birmingham school. Segments of the data were manually annotated, and patterns became increasingly salient over time. A total of 59 excerpts were copied into separate excel tables for a more fine-grained study (see digital appendix – 'Corpus of selected exampels'). Admittingly, this method and sample does not rise to the level of a perfectly exhaustive and definite description of all potential patterns. Nevertheless, it suggested several distinct discourse practices, which re-occurred periodically and have a consistent set of features.

Section 6.1 will begin with the patterns in the formation of complete and incomplete exchanges. Complete exchanges bring the participants together in dialogue. There are three participant configurations that correspond to three vastly different patterns of complete exchanges. There are 1) written intra-chat exchanges between viewers, 2) cross-modal exchanges from viewers to streamer and 3) cross-modal exchanges from streamer to viewers. Of these three patterns, the cross-modal exchanges from viewer to streamers is the most common and most interesting one. Their discussion will reveal a new discursive move, 'the topicalizer', which has not been described in previous computer-mediated settings. The topicalizer is a spoken repetition that topicalizes the written discourse by the chat. It has a fixed position in the exchange sequence and is for the organization of complete cross-modal exchanges.

Then there are incomplete exchanges that begin to form an exchange but do not fully complete. There is break in the conventional sequence of dialogue. In 'ignored exchanges' the initiation is unable to pick up a response and in 'interrupted exchanges' the response is temporarily paused and resumed later. Ignored exchanges are typical for the chatting viewers, whereas interruptions only affect the speaking streamer. The study will explain these different

exchange patterns and relate them to the unfolding gameplay. Complete dialogical exchanges occur in absence of game events. If there are game events, exchanges might be ignored or interrupted because the participants orient towards the game and produce monological discourse that is not directly addressing a specific recipient.

Section 6.2 is dedicated to the study of this game-related talk & chat. There are two distinct communicative practices, namely commenting and reporting. The analysis will reveal that they possess very different forms and purposes for the broadcast. Reporting is reflective, educational and only produced by the streamer. Commenting is spontaneous, emotional, and is produced by streamers and viewers. Commenting and reporting depend on the presence of game events, which means that the alternation between dialogue and monologue is tied to the unfolding of the game. There is a pairing of the organization of discourse and the non-linguistic organization of live streaming.

Section 6.3 will continue this argument by moving to the next higher ranks, i.e. the transactions and the complete linguistic interaction. There is a fixed set of topics that re-emerge continuously throughout a live stream and these transactions are connected to the cyclical organization of the live streamed video game in rounds and levels. As different stages of the game repeat, the same transactions re-occur. This pattern is repeated throughout the interaction between streamer and viewers for the duration of the daily broadcast. Lastly, these daily broadcasts are chained together into a weekly streaming schedule that the streamer adheres to.

The organization of discourse between streamers and viewers is highly organized from the smallest to the largest ranks. In absence of game events, there are dialogical exchanges between participants, whereas the sudden appearance of game events leads to commenting and reporting. Transactions are tied to the game's cyclical organization and the interaction takes place during scheduled times that follow the rhythm of a work week. This explanation will be able to account for most of the ordinary discourse of live streaming on *Twitch.*

6.1 Dialogue segments and exchanges

Exchanges are the basic unit of dialogue that contain an obligatory initiation and an obligatory response to form at least one adjacency pair I→R. Optional elements were response-initiations (R/I) that reinitiate the talk as well as a follow-up (F) (cf. Francis & Hunston, 1992). This structure has been summarized in the formula E = I→(R/I)→R→(F) for complete mono-modal spoken dialogue. In online live streaming, exchanges often have different features, which are due to participant configuration and their available linguistic modes. This section will highlight common patterns and will begin with complete exchanges in 6.1.1. In 6.1.2, the dissertation

will show incomplete exchanges, which are segments of communication that do not develop into a proper exchange.

6.1.1 Complete exchanges

Complete exchanges in online live streaming come in three types. There are the mono-modal exchanges that occur exclusively between audience members and there are two types of cross-modal exchanges between the streamer and the audience. This subsection will begin with rather simple examples and move towards the more complex exchanges.

Mono-modal exchanges

(1) Example – 'How old is the streamer'

Timestamp	Events	Streamer	Chat
2216:[00:19:07]			<hakkesho> how old is bjersen
2219:[00:19:10]			<costlymilk> 12
2225:[00:19:15]			<jujumanji> 19
2229:[00:19:17]			<bandicosplay> bjerg is 19 :)
2233:[00:19:18]			<sooted> bjerg is 19

The example was already presented in the fifth chapter to show moves that are directed towards the viewer group. In the context of this chapter, it constitutes a complete exchange with a single initiation that is met with several responses. The responding moves are produced independently of one another and there is no indication for coordination between participants. There is no follow-up by the initiator, which means the simplest complete exchange in a multi-party dialogue has the exchange structure $E_{mono\text{-}modal} = I_{written} \rightrightarrows R_{written}$, where $\rightrightarrows$ indicates the possibility of several responses for a single written initiation. This exchange structure is very common for the discussion of background information in the chat. Such background information is unrelated to the current events in the livestream and discusses information about the streamer or general questions about the video game.

(2) Example – 'Butland or Courtouis'

Timestamp	Events	Streamer	Chat
345:[00:08:40]			<ddavila101010> Butland or courtouis?
353:[00:08:53]			<ProfitBirdAus> @ddavila101010 Butland imo
355:[00:08:55]			<caleb_oltmans> @ddavila101010 Butland
357:[00:09:08]			<da_closer6> Butland
:			
559:[00:15:01]			<ddavila101010> Thanks guys

The second example is taken from *FIFA 2015*, where Butland and Courtouis are playable avatars. The initiator is talking about a theoretical choice within the management of the game. S/he can buy a digital version of the players 'Butland' or 'Courtouis' and is asking the viewer group for opinions. To assist in this decision, a responder has to know about the avatar's strengths and weaknesses. The verdict is anonymous, and three responses suggest the avatar

'Butland'. The follow-up by ddavila1010101 is in the plural '`guys` (l.559)' and implicitly assumes that it is an all-male chat and the messages thanks all three chat members simultaneously. Separate follow-ups to each response would be a possibility but are absent from the data. This indicates that there is a preference for a single follow-up and an exchange structure of $E_{mono\text{-}modal} = I_{written} \rightleftarrows R_{written} \rightarrow F_{written}$ for multi-party chat. Despite their simple form, they already require basic game knowledge. The next two examples are more elaborate discussions in the chat and they are tied more directly to the gameplay on the screen. The game is an ever-present frame of reference for the communication and exchanges can be very jargon-heavy and require a lot of tacit knowledge.

(3) Example – 'Jayce Buff'

Timestamp	Events	Streamer	Chat
225:[00:05:01-7]	[Jayce(e) kills Gragas]		
:			
230:[00:05:19-8]	[Jayce (e) kills Victor]		
234:[00:05:23]			<%NexAfire> They buffed Jayce? I've seenso many lately
235:[00:05:30]			<loingirdler> yeah they buffed
236:[00:05:36]			<@%Zegers93> They did
237:[00:05:50]			<Slash_Evan> they're always buffing jayce
238:[00:05:50-2]		hey man these bottom lane.	
239:[00:05:52-9]		uargh it is such a bad spot to be in right now.	
240:[00:05:53]			<Slash_Evan> cus no one plays him
241:[00:05:55]			<Slash_Evan> xD
:			
249:[00:06:10]			<%NexAfire> Yet Yorick remains the same
250:[00:06:13]			<loingirdler> Better damage and scaling on his melee q and mana cost reduced

The example 'Jayce Buff' is taken from *League of Legends* and begins with the enemy Jayce player defeating two team mates of the streamer (l.225; l. 230). The streamer reports on the bad outlook of the game (l.238 – 239), but does not engage with the chat because he is too busy with the gameplay.

There are two separate floors and while the streamer is focused on the game, the viewers have a background discussion about the game avatar 'Jayce'. The discourse in the chat is occasioned by the game events but it is not commenting or reporting on it. The initiation of the exchange is a question "`they buffed Jayce`?" (l.234)", followed by an 'elaborating move' "`I've seen so many lately`" (l.234), which provides additional information that is relevant to the discourse. In this case, the elaborating move warrants the question as reasonable. "`I've seen so many lately` (l.234)" indicates that the avatar Jayce might have been 'buffed', or strengthened.

There are four responses to the question (l.235; l.236; l.237- 241; l.250). The first two responses R_1 (l.235) and R_2 (l.236) are affirming moves that provide positive confirmations. R_3 (l.237-241) and R_4 (l.250) are elaborating moves that rely on game-specific expertise. The strengthening (buff) of Jayce indeed happened "`cuz no one plays him` (l.240)". To

understand R_3, it is necessary to know that there is a correlation between avatar's strength and how often the character is played. Since gamers want to win their matches, strong characters are played more often than weak characters. The viewer is saying that the game developers strengthened the character Jayce to increase his popularity in the game.

The follow-up (l.249) refers back to this response R_3 and the message "`cuz no one plays him` (l.240)" and builds on this tacit knowledge. 'Yorik' is another avatar that is also played very rarely because he is weak and "yet Yorik remains the same  "[30], which means he is not buffed / strengthened. There is a mix of colloquial language with oral features, game-specific jargon and unsaid presuppositions (Hornyak, 2016). This informal expert discourse allows the exchange of very complex meanings in a concise manner.

On the surface, the discussion is casual but without the insider knowledge, the meaning of the exchange remains opaque. This phenomenon can also be demonstrated in the fourth example, where different degrees of game-expertise lead to a disruption in an exchange.

(4) Example – 'raid style party frames'

Timestamp	Events	Streamer	Chat
1603:[01:25:29]			<sravenous> What is that ui showing party member's hp in square box
1607:[01:25:48]			<Cherrybombx> Sravenous i was wondering the same
1608:[01:26:00]			<Laydeeqt> @sravenous, you mean raid style party frames?
1611:[01:26:16]			<sravenous> Yea white color for priest and yellow for rogue
:			
1613:[01:26:27]			<sravenous> What is that ui name?
1614:[01:26:30]			<Cherrybombx> @Laydeeqt I was wondering aobut the druids healthbar in orange how do u get that
1618:[01:26:45]			<Laydeeqt> @sravenous, you have that in interface raid style party frames is the name
1619:[01:26:55]			<cdewsdiploma> It's default frames. Click arrow, click yellow i, click use class colors, click use raid style party frames.
1620 :[01:26:57]			<Laydeeqt> @Cherrybombx, that's default for me
1623:[01:27:16]			<Cherrybombx> hmm wtf for me it isnt
1624:[01:27:32]			<sravenous> Oh so i that is basic ui from blizzard? I didn't know
1627:[01:27:44]			<Laydeeqt> @sravenous, yes it is
1630:[01:28:08]			<sravenous> Thank you!!

The excerpt is taken from *World of Warcraft* and shows several exchanges that form a transaction. The topic is the user interface of the game and the topic are its 'raid style party frames'. The exchanges occur during an arena match of the streamer, playing a 'rogue' together with a 'priest' team mate against an opposing team of 'druid' and 'rogue'. During the exchanges, the streamer is pre-occupied with the match and does not join in the discussion. Therefore, his utterances have been abridged.

[30] The emoji is named 'painlessFeels ' and depicts a sad and anthropomorphic frog face

I 1603 <sravenous> What is that ui showing party member's hp in square box

R_1 1607 <Cherrybombx> Sravenous i was wondering the same

R_2 1608 <Laydeeqt> @sravenous, you mean raid style party frames?

In the first exchange, the initiation is a question about the UI element (l.1603) and it receives two responses R_1 and R_2. Both responses mark the intended recipient with a fronted 'attending move' and indicate that the following information is designed for 'sravenous'. R_1 is a statement that shows 'Cherrybombx' has a similar question. R_2 is a proper response to the initiation and contains the complete answer to the original question. The UI element is called "`raid style party frames` (l.1607)." However, the response R_2 is formulated as an interrogative question and not as declarative statement. This syntactical choice and the fact that 'sravenous' and 'Cherrybombx' appear unfamiliar with the term 'raid style party frames' lead to a misunderstanding and disruption in the exchange. Presumably, 'sravenous' reads 'raid style party frame' as a general description and understands R_2 as a request for clarification. R_2 "`You mean raid style party frames?` (l.1608)" is treated as a response-initiation and 'sravenous' produces an affirmation as a response "`Yea white color for priest and yellow for rogue` (l.1611)." The response expects a final follow-up, which answers the question about "`that ui`" element. This would end the exchange with the pattern $E_{mono} = I_{written} \rightarrow R/I_{written} \rightarrow R_{written} \rightarrow F_{written}$. For Laydeeqt, the message R_2 (l.1608) was meant as an answer to the question. Therefore, s/he treats the affirming move "`Yea white color for priest and yellow for rogue`" (l.1611) as a follow-up that ends the sequence $E_{mono} = I_{written} \rightarrow R_{written} \rightarrow F_{written}$. Both participants have access to the visual information on the screen and discuss them via exophoric references. However, there is an asymmetry in their game knowledge and use of jargon, which leads to a misalignment of expectations for the exchange. Despite its informal features, participation in the discourse of live streaming requires significant expertise, where viewers must be able to 'read' the game and discuss it with appropriate terminology. Most chatting participants possess enough game knowledge and such misunderstandings are surprisingly rare. In this case the misunderstanding is not immediately resolved, 'sravenous' (l.1613) and 'Cherrybombx' (l.1614) simultaneously start new overlapping exchanges about the same topic.

I 1613 <sravenous> What is that ui name?

R 1618 <Laydeeqt> @sravenous, you have that in interface raid style party frames is the name

Laydeeqt responds more explicit and incorporates elements of the question into the response

(cf. Du Bois & Giora, 2014). However, due to the lack of punctuation, poor use of conjunctions and a conflation of terms, his contribution remains difficult to understand. A fully-worded answer could be "@sravenous, you have that user interface (UI) setting in the interface folder and raid style party frames is its name". There is an unmentioned distinction between a specific user interface setting, i.e. raid style party frames, and the interface folder, which houses the interface setting.

'Cherrybombx' is also unable to identify 'raid style party frames' as a proper name and falls back to an exophoric description that uses the visually present game as a point of reference. In the live steam the "`druids healthbar in orange`" (l.1614) is visible to every chat member.

I 1614 <Cherrybombx> @Laydeeqt I was wondering aobut the druids healthbar in orange how do u get that

R_1 1619 <cdewsdiploma> It's default frames. Click arrow, click yellow i, click use class colors, click use raid style party frames.

R_2 1620 <Laydeeqt> @Cherrybombx, that's default for me

F 1623 <Cherrybombx> hmm wtf for me it isnt

The Initiation and request for information is met with two responses that reinterpret 'druid healthbar in orange' as a shorthand description for 'raid style party frames'. However only 'cdewsdiploma' is complying with the request and provides a step by step guide as a series of commands in his Response R_1.

The fourth exchange serves as closure to the transaction.

I 1624 <sravenous> Oh so i that is basic ui from blizzard? I didn't know

R 1627 <Laydeeqt> @sravenous, yes it is

F 1630 <sravenous> Thank you!!

'sravenous' begins to understand that the 'raid style party frames' are a "`basic ui` (l.1624)" setting in the game but tries to conform with a final exchange. The initiation contains the phrase "`from blizzard` (l.1624)" as a synecdoche, where the software developer of the game, *Blizzard Entertainment*, stands in for the game *World of Warcraft*. Laydeeqt infers this meaning correctly and produces an affirming move in the response, which ultimately leads to an appreciative follow-up that ends the discussion about the interface (l.1630).

The four examples demonstrate that even the intra-chat communication is usually grounded in some aspect of the live stream. Chat participants discuss background information about the streamer or the video game. Their language is often very informal and contains many

spoken features. Therefore, the discourse appears similar to a casual conversation (cf. Hamilton et al., 2014). However, messages are also rich in jargon and contain many unsaid presuppositions (Hornyak, 2016). In most cases, participants are experienced enough and exchange very complex meanings in very few turns. Misunderstandings are due to a lack of game knowledge or asymmetries in game knowledge. Chat participants mitigate this issue by producing exophoric references that depend on the mutually shared live stream (cf. Halliday & Hasan, 1976).

Cross-modal exchanges

In cross-modal exchanges, the floor transfer from initiation to response or from response to follow-up coincides with a shift in the mode of communication. Therefore, one can distinguish exchanges that originate in the speech of the streamer and those that start with the writing of the audience.

(5) Example – 'another question for you guys'

Timestamp	Events	Streamer	Chat
346:[00:01:42-0]		another question for you guys.	
352:[00:01:44-9]		I was thinking about doing the player road to glory around	
358:[00:01:45-9]		pink Reus.	
361:[00:01:47-6]		making it a little more interesting.	
367:[00:01:49-2]		{Gaze to Chat Monitor}	
368:[00:01:49-3]		instead of using just a striker.	
373:[00:01:50-9]		But I did wanne see what you guys thought.	
381:[00:01:54]			<Haroon_786123> Yesssssss
382:[00:01:54]			<Quinn_connolly> Yes
383:[00:01:54]			<Donster13> Bale
385:[00:01:56]			<Billyspencerzz> Nick use tots muller
386:[00:01:56]			<%Mathias_chnk> reus
388:[00:01:57]			<Angel123mer> best player in the world messi
390:[00:01:57]			<Jannik1607> nooo
391:[00:01:57]			<%AlexB> yesss
394:[00:01:58]			<Carmelo2221> Make it Around TOTS SANCHEZ
395:[00:01:58]			<%lewis_loves_banter> yeah
396:[00:01:58]			<Kamisafan45> Yess
397:[00:01:58]			<Dakram1> Pink bale

The fifth example shows a streamer-initiated exchange. The topic of the exchange is the selection of the game avatar in *FIFA 2015* (l. 352). The streamer wants to incorporate his audience in the decision-making and outlines his idea[31]. He has a monopoly on the spoken floor and can utter several moves to make his point. The audience keeps their responses very t short and provides suggestions without further explanation. Effectively, their messages are

[31] Reus, Bale, Müller and Sanchez are soccer players and potential avatars in the soccer simulation

condensed commands or requests that are disambiguated by the streamer's question and the context of avatar selection in the game.

Notably, the streamer's initiation receives significantly more responses than the audience initiations in the previous examples. Although there are many viewers to chat with, a disproportionate number of moves is directed towards streamers (see chapter 5). It suggests a preference for communication with the streamer and the audience-streamer dialogue is treated as the main interaction.

The form of the streamer-initiated exchanges shows very little variation from $E_{cross\text{-}modal} = I_{spoken} \rightleftarrows R_{written}$ and the close reading of the data did not identify a clear case of a spoken follow-up F_{spoken}. In streamer-initiated exchanges, F_{spoken} is impractical because of the delay of the broadcast. As streamers end their initiation, they would have to wait the 8-12 second for the video to reach their viewers and the additional time for the audience to produce written responses. Although this is theoretically possible, it was not recorded and rarely practiced. In general, streamer-initiated exchanges are very rare and they are vastly outnumbered by the audience-initiated exchanges (see chapter 5).

(6) Example – 'Spell Vamp on Kennen?'

Timestamp	Events	Streamer	Chat
10105:[01:32:51-2]			<generalchiefy> @TSM_Bjergsen Spell vamp on kennen?
10128:[01:33:06]		{gaze to chat monitor}	
10130:[01:33:09]		{gaze to game monitor}	
10133:[01:33:10]		Spell vamp on Kennen?	
10134:[01:33:10-1]		no it just delays his core items too much.	
10138:[01:33:15]		like it delays his Abbysal Zhonia's.	
10144:[01:33:17]		combo.	
10147:[01:33:19]		or whatever AP item you wanne built for.	
10151:[01:33:22]		and it's pretty much only good in lane.	
10152:[01:33:22-4]		it doesn't do anything in teamfights.	
10162:[01:33:29]		people used to built it.	
10166:[01:33:30]		cuz spellvamp used to be so much better.	
10167:[01:33:30-3]		used to be.	
10171:[01:33:33]		the items just were a lot better.	

The sixth example is taken from the large-sized *League of Legends* stream and is abridged to highlight the most common features of audience-initiated cross-modal exchanges. They usually have the form $E_{cross\text{-}modal} = I_{written} \rightarrow R_{spoken}$ and the written initiations almost always consist of two or more moves produced in a single message. The audience's initiation begins with an attending move, '`@TSM_Bjergsen`' to address the streamer followed by the actual request or question. If the audience member wants to write several moves, they produce a long message

and do not split their text across short messages[32]. There is a competition for chat space and viewers need to make sure that their initiation remains coherent in the very active chat channels.

Streamers do not compete with other speakers, but they must deal with their own set of limitations. To see the initiation (l.1105), streamers must turn to the chat monitor and read the chat message. Often, there will be several competing initiations and the streamer must choose among them, indicate his or her choice and then produce the actual response. The linguistic practice that evolved to solve this issue is a special form of repetition. In example six, the streamer utters "`spell vamp on Kennen?` (l. 10133)" to refer to the question in line 10105. Such repetitions are almost universal to streamers' responses in cross-modal exchanges. They are a novel linguistic feature that has not been discussed in prior descriptions of oral repetitions (Tannen, 1989), or spoken conversation (Eggins & Slade, 1997) or video-mediated conversation (Rosenbaun et al., 2016). The function of these repetitions is to index a written message from the chat and turn it to the topic of the spoken floor. Therefore, this study will refer to such moves as topicalizers. The topicalizer in example six "`spell vamp on kennen` (l.10133)" is an incomplete repetition of a previous chat message (l,10105). It topicalizes the central proposition of the message but leaves out the attending move '`@TSM_Bjergsen`' and the name of the initiator. Topicalizers are always in the first position of the response slot.

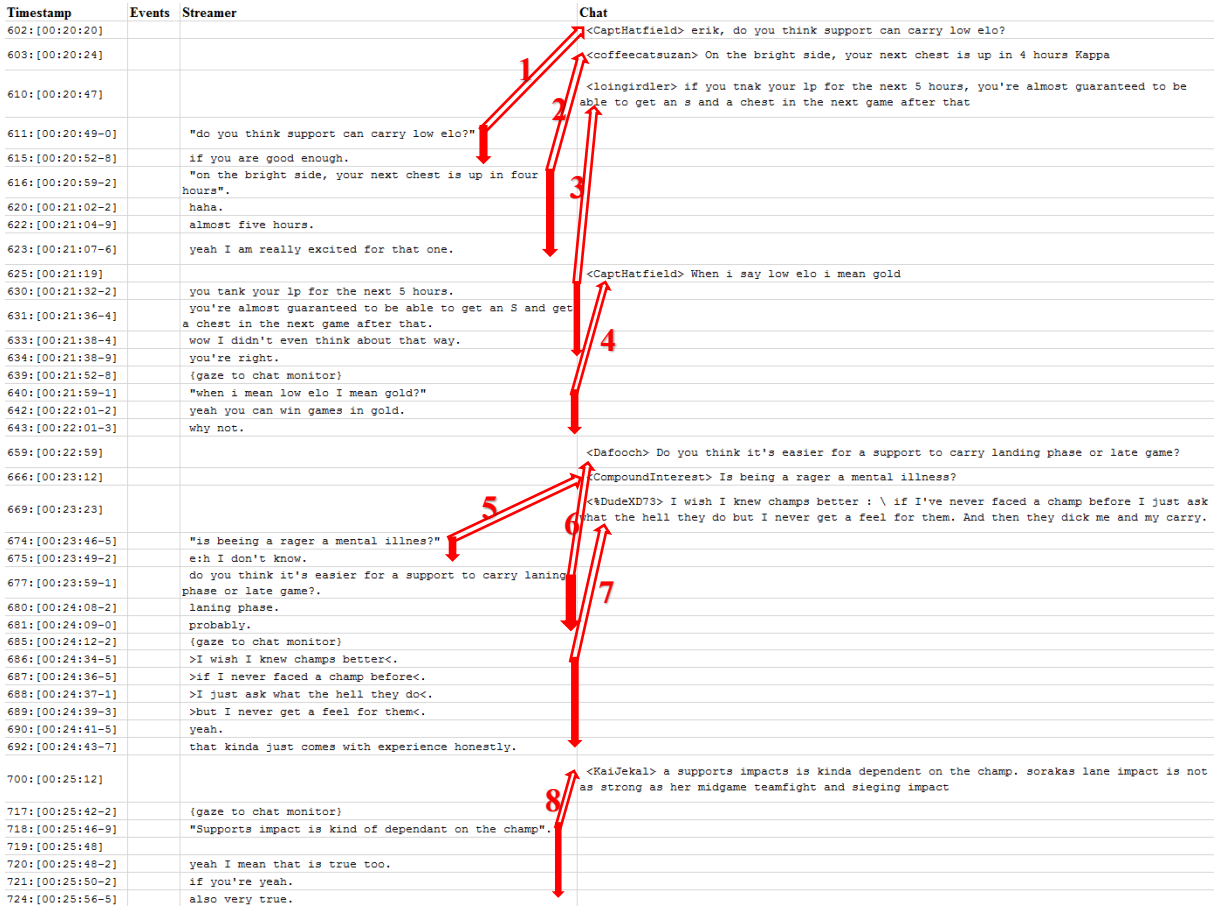
After 'spell vamp on Kennen' is established as the topic, the streamer proceeds to answers the question in a series of elaborating moves (10134 – 10171). The streamer argues that 'spell vamp on Kennen' is reasonable but not optimal because there are other game components that are better. He provides alternative choices and also gives a concrete argument against the strategy 'spell vamp on Kennen' as it is "`good in lane` (l.10151)" but very weak "`in teamfights` (l.10152)". It is a utilitarian argument, where a game strategy is evaluated according to its usefulness. The streamer provides an insightful and extended explanation to argue at length against the proposition 'spell vamp on Kennen'. The argument is full of jargon and interconnected concepts. A listener must know the meaning of 'spell vamp', 'lane' or 'teamfight' and s/he must be able to interpret them in relationship to one another. This requires extensive game-specific expertise and such examples are prevalent in all three games of the study. Most exchanges discuss the broadcasted game in an expert discourse and gamers that are unfamiliar with a specific title may not be able to follow the discussion.

Another notable property of online live streaming is how streamers respond audience-

[32] This is different from texting or traditional IRC chat, which emulated the production of spoken intonation units (Crystal, 2005).

initiated cross-modal exchanges in a series. In written or spoken mono-modal conversation, exchanges have free turn-taking allocation, which leads to a rapid shift between different speakers or overlapping exchanges (Herring, 1999; Matthiessen & Slade, 2011). There is a constant back and forth between the participants (Tannen, 1987). In cross-modal dialogue, the parties cannot overlap or interrupt because they operate on two distinct floors and on slightly alternated timelines. This leads to cross-modal exchange chaining.

(7) Example – 'can support carry?'



Timestamp	Events	Streamer	Chat
602:[00:20:20]			<CaptHatfield> erik, do you think support can carry low elo?
603:[00:20:24]			<coffeecatsuzan> On the bright side, your next chest is up in 4 hours Kappa
610:[00:20:47]			<loingirdler> if you tnak your lp for the next 5 hours, you're almost guaranteed to be able to get an s and a chest in the next game after that
611:[00:20:49-0]		"do you think support can carry low elo?"	
615:[00:20:52-8]		if you are good enough.	
616:[00:20:59-2]		"on the bright side, your next chest is up in four hours".	
620:[00:21:02-2]		haha.	
622:[00:21:04-9]		almost five hours.	
623:[00:21:07-6]		yeah I am really excited for that one.	
625:[00:21:19]			<CaptHatfield> When i say low elo i mean gold
630:[00:21:32-2]		you tank your lp for the next 5 hours.	
631:[00:21:36-4]		you're almost guaranteed to be able to get an S and get a chest in the next game after that.	
633:[00:21:38-4]		wow I didn't even think about that way.	
634:[00:21:38-9]		you're right.	
639:[00:21:52-8]		{gaze to chat monitor}	
640:[00:21:59-1]		"when i mean low elo I mean gold?"	
642:[00:22:01-2]		yeah you can win games in gold.	
643:[00:22:01-3]		why not.	
659:[00:22:59]			<Dafooch> Do you think it's easier for a support to carry landing phase or late game?
666:[00:23:12]			<CompoundInterest> Is being a rager a mental illness?
669:[00:23:23]			<%DudeXD73> I wish I knew champs better : \ if I've never faced a champ before I just ask what the hell they do but I never get a feel for them. And then they dick me and my carry.
674:[00:23:46-5]		"is beeing a rager a mental illnes?"	
675:[00:23:49-2]		e:h I don't know.	
677:[00:23:59-1]		do you think it's easier for a support to carry laning phase or late game?.	
680:[00:24:08-2]		laning phase.	
681:[00:24:09-0]		probably.	
685:[00:24:12-2]		{gaze to chat monitor}	
686:[00:24:34-5]		>I wish I knew champs better<.	
687:[00:24:36-5]		>if I never faced a champ before<.	
688:[00:24:37-1]		>I just ask what the hell they do<.	
689:[00:24:39-3]		>but I never get a feel for them<.	
690:[00:24:41-5]		yeah.	
692:[00:24:43-7]		that kinda just comes with experience honestly.	
700:[00:25:12]			<KaiJekal> a supports impacts is kinda dependent on the champ. sorakas lane impact is not as strong as her midgame teamfight and sieging impact
717:[00:25:42-2]		{gaze to chat monitor}	
718:[00:25:46-9]		"Supports impact is kind of dependant on the champ".	
719:[00:25:48]			
720:[00:25:48-2]		yeah I mean that is true too.	
721:[00:25:50-2]		if you're yeah.	
724:[00:25:56-5]		also very true.	

The excerpt is taken from *League of Legends* and occurred between two rounds of the game. It shows a series of eight cross-modal exchanges of the form $E_{cross\text{-}modal} = I_{written} \rightarrow R_{spoken}$, which demonstrates the consistency of the pattern. The overall topic of the exchanges is 'can support carry?'. It asks if the role of 'support' is strong enough to impact the outcome of matches in *League of Legends*.

$I_{written}$ `602<CaptHatfield> erik, do you think support can carry low elo?`

R_{spoken} `611"do you think support can carry low elo?"`
`614 yeah.`
`615 if you are good enough.`

In all 8 exchanges, R_{spoken} begins with a topicalizer that repeats the main proposition of the

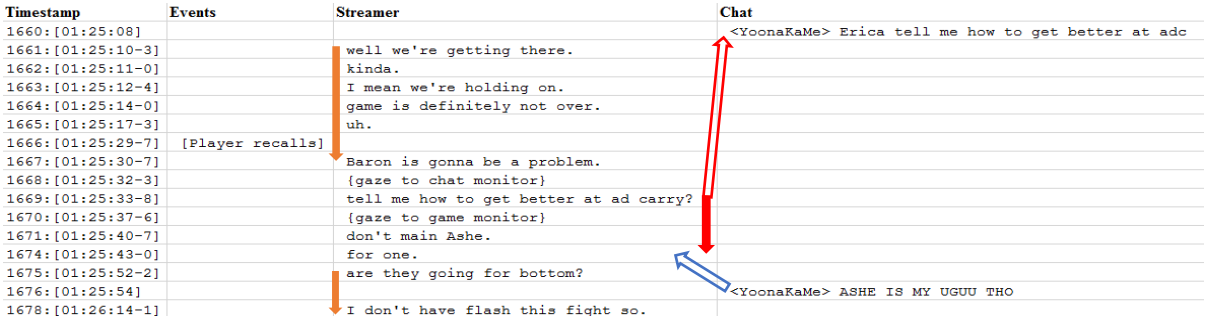
initiation. Inverted commas highlight an audible shift in voice pitch. This is a common practice to signal reported speech of non-present parties in face-to-face conversation (Holt, 1996; Couper-Kuhlen, 1999). In cross-modal exchanges, the difference in voicing signals that the message comes from a chat member and raises the content of the message to the spoken floor. The topicalizers identify who the streamer is talking to and what they are talking about. Due to the delay, the streamer always refers backwards in time and toplicalizes 'older' discourse. In example seven, the temporal gap between an initiation and its topicalizer is between 11 seconds (l.669 to l.686) and up to 1 minute (l.659 - l.677). In a sense, the streamer is always behind the current discussion and cannot participate in fast paced and developments of a topic in the chat.

Instead, they practice exchange chaining, where they participate in the discussion by responding to audience-initiations in order of their production in the chat. Exchange chaining has a consistent pattern of topicalizing and responding to initiations. Streamers alternate their dialogical partner in every exchange. This behavior is similar to a public Q & A, where audience members take turns and pose questions to a main speaker, who will then answer them one-by-one. The turn-taking by the initiators is replaced by topicalizer uttered by the streamer. This type of cross-modal communication may be described as turn-giving rather than turn taking because viewers rely on the streamer to acknowledge their question.

Exchange chaining is less fluid than ordinary conversation and makes the discourse disjointed and 'bouncy'. Each exchange is more insular, stands on its own for extended periods and there may be significant many transitions between topics. In the excerpt above, the 1st, 4th, 6th, 7th and 8h cross-modal exchange roughly share the same the topic of 'can support carry', whereas the 2nd, 3rd and 5th exchange are completely unrelated to it. Importantly, this is not an argument for technological determinism, where a 'deficiency' in the technology leads to a 'deficient discourse' (cf. Walther, 1996; Crystal, 2011). Streamers could ignore interfering or unrelated messages, but as a practice they decide to topicalize messages in order of their appearance and decide to talk to as many viewers as they reasonably can.

Limits to dialogical exchanges arise from the number of written messages or the occurrence of game events. If a streamer wants to respond during a round, they must negotiate the demands of the communication with the demands of the gaming activity. This leads to quick transitions between dialogical exchanges and the more monological reporting.

(8) Example – ‘Don’t main Ashe’



Timestamp	Events	Streamer	Chat
1660:[01:25:08]			<YoonaKaMe> Erica tell me how to get better at adc
1661:[01:25:10-3]		well we're getting there.	
1662:[01:25:11-0]		kinda.	
1663:[01:25:12-4]		I mean we're holding on.	
1664:[01:25:14-0]		game is definitely not over.	
1665:[01:25:17-3]		uh.	
1666:[01:25:29-7]	[Player recalls]		
1667:[01:25:30-7]		Baron is gonna be a problem.	
1668:[01:25:32-3]		{gaze to chat monitor}	
1669:[01:25:33-8]		tell me how to get better at ad carry?	
1670:[01:25:37-6]		{gaze to game monitor}	
1671:[01:25:40-7]		don't main Ashe.	
1674:[01:25:43-0]		for one.	
1675:[01:25:52-2]		are they going for bottom?	
1676:[01:25:54]			<YoonaKaMe> ASHE IS MY UGUU THO
1678:[01:26:14-1]		I don't have flash this fight so.	

The excerpt is again taken from *League of Legends* in a moment where the streamer is likely to lose the match. He is pre-occupied with the game and reports on the negative outlook and a possibility for a comeback (l.1661 – 1667). The streamer is fully engaged with the game and his discourse reflects on the match.

As experienced gamers, streamers know when a game will demand their attention and when there is a brief window that requires less game input. As the streamer uses the ‘recall’ ability (l.1666), he teleports to a safe area and gameplay temporarily becomes less urgent. This freedom from the constraints of the game is used as a window of opportunity for dialogue. He turns his gaze to the chat monitor (l.1668) and notices the previous initiation (l.1660). It contains an attending move, Erica, followed by a command “`tell me how to get better at adc` (l.1660)”. The streamer utters the corresponding topicalizer (l.1669) and verbalizes the abbreviation adc into ad carry. Reading requires the streamer to look at the chat monitor, but afterwards he can immediately re-orient back to the game monitor (l.1670). The practice of gaze shift → topicalizer → gaze shift is very common for cross-modal exchanges during a match and minimizes the time a streamer must look away from the game monitor. The verbal production of the response does not require vision of the chat.

In this case, the response consists of a brief statement “`don’t main Ashe for one` (l.1671; l.1674)”, which answers the original question and tells the viewer not to play the character ‘Ashe’. The streamer does not qualify his opinion via elaborating moves because he is quickly re-immersed in the unfolding gameplay as the enemy team is continuing their attack (l.1675). Streamers dynamically transition their interactional orientation between audience communication and game. They engage in dialogical exchanges when it is possible but also focus on the game when it becomes necessary. Karhulathi’s (2016) postulated ‘play frame (competition)’ and ‘interview frame (conversation)’ are useful descriptive categories to understand this behavior, but they are realized on much smaller scales then previously theorized. Streamers juggle their roles as player and caster and look for opportunities of communication. In absence of immediate game events, streamers can temporary alternate between the game-focused and audience-focused interaction. This can be seen in their gaze

shifts between chat and game monitor and their alternation between dialogical exchanges and the more monological reporting and commenting. In a slow-paced game such as *League of Legends,* there can be many such transitions during a single match, whereas the faster paced *FIFA* and *World of Warcraft* allow for fewer transitions. In these two games, there are more pressing game events that require the streamer's immediate attention, which is detrimental to the production of complete exchanges. The results are longer segments of reporting and commenting as well as instances of incomplete exchanges.

6.1.2. Incomplete exchanges

Incomplete exchanges are exchanges that start out with a proper initiation but then have breakdown in their exchange structure. They are an unsuccessful attempt at dialogue and come in the type of ignored exchanges and interrupted exchanges.

In the fifth chapter, the dissertation showed that the audience in large-sized streams produces many discursive moves that are addressing the streamer (~ 16% to 28%). This substantial share of moves is directed at a single individual. There is an imbalance between the number of chat participants that attempt to initiate an exchange with the streamer and the streamer's ability to respond to them. Streamers mitigate this issue by chaining exchanges and responding to as many chat messages as possible (see example 7). However, in very active channels this is not enough, and many initiations will remain incomplete.

(10) Example – '@Streamer'

Timestamp	Events	Streamer	Chat
822:[00:06:27]			<danielson108> @TSM_Bjergsen hey is Orianna a good champ?
823:[00:06:27]			<xexclusity> @TSM_Bjergsen, is their a console to the virtual reality, or is it all done in the part you wear?
825:[00:06:29]			<jdarocha210> @tsm_bjergsen can you play Yasuo
826:[00:06:29-4]		oh end the game.	
828:[00:06:30-7]		{gaze to chat monitor}	
830:[00:06:32]			<t0xicragnar> @Tsm_bjergsen wghat u think about ziggs?
835:[00:06:33]	[player wins round]		
838:[00:06:34]			<frozenliquidz> @tsm_bjergsen do you still play any nidalee?
842:[00:06:37]			<beaglejuce> bjerg what tips do you have for climbing elo?
843:[00:06:37]			<tritosblast> @Tsm_bjergsen Why dont u Duo with oddone?? he would get so happy!
846:[00:06:38]			<qelbiv> the question is, what item were you trying to do ?
848:[00:06:39]			<kaimiloaisbomb> @Tsm_bjergsen why or don't get play anymore?
849:[00:06:41]			<tyedyecloudlol> @Tsm_bjergsen why is diana weak in the current meta?
857:[00:06:44-4]		is there a console to the virtual reality?	
858:[00:06:44-9]		or is it done.	
859:[00:06:45]			<kungen29> How is the life of being an pro/lcs gamer? please answer :)
860:[00:06:45]			<sooted> @Tsm_bjergsen When do you get your HTC M9?
862:[00:06:45]			<the_tacos_doctor> @Tsm_bjergsen what do u think of the new champion bard and do u think he is op?@Tsm_bjergsen what do u think of the new champion bard and do u think he is op?
863:[00:06:45-3]		done in the part you wear.	
868:[00:06:46-9]		it is all done in the part you wear.	

The excerpt is taken from the large-sized *League of Legends* broadcast and demonstrates the phenomenon of ignored exchanges. As the streamer is about to win the round (l.835) his attention is freed from the demands of the game. He orients towards the chat in preparation of cross-modal exchanges (l.828) and selects from the many potential initiations.

$I_{written}$	823	<xexclusity> @TSM_Bjergsen, is their a console to the virtual reality, or is it all done in the part you wear?
R_{spoken}	857	is there a console to the virtual reality?
	858	or is it done.
	863	done in the part you wear.
	868	it is all done in the part you wear.

The streamer's Response R_{spoken} begins with the topicalizer (l.857, l.858, l.863) and provides a statement of fact as the answer (l.868). During this brief period, the excerpts features thirteen initiations but only can be picked up by the streamer. The other twelve remain unsuccessful in requesting a response. Chat participants have only limited recourse to this problem of visibility and compete for spoken responses. One solution is self-repetitions in order to increase their visibility.

$I_{written}$	862	@Tsm_bjergsen what do u think of the new champion bard and do u think he is op?@Tsm_bjergsen what do u think of the new champion bard and do u think he is op?

In line 862, the message consists of an attending move '@TSM_Bjergsen' followed by and two game-related questions. Afterwards, the moves are repeated verbatim, which suggests that it is an instance of copy & paste. The message becomes bigger and more noticeable, which may increase the user's chance to be seen.

Other types of self-repetition can be seen in example 11 taken from *FIFA*

(11) Example – 'Hey shoutout pls'

Timestamp	Events	Streamer	Chat
9030:[00:36:12]			<Visca_karloo> Hey Nick <3 Is this the PMRTG?
:			
9272:[00:38:13]			<Visca_karloo> @nick_28t Heyy :D
:			
9417:[00:39:32]			<Visca_karloo> Hey Nick :D <3
:			
10156:[00:45:31]			<kushisweed> @Nick_28T Shoutout plz
:			
10510:[00:47:31]			<kushisweed> @Nick_28T Can i get a shoutout me mother is dying
:			
10686:[00:48:23]			<kushisweed> @Nick_28T Shout my mother is dying plz
:			
10939:[00:49:40]			<kushisweed> @Nick_28T Shout my mother is dying plz she wants a proper goodbye

In this case, the audience members seek acknowledgement by the streamer in the form of a greeting or 'shoutout'. They remain unnoticed due to the high-volume messages (abridged). As a result, they repeat a similar message with slight variations in a brief interval. The messages are designed to be noticeable with a fronted attending move, emoticons or an invented story of a terminally-ill mother (l.10939).

These repetitions are a rational behavior for the individual to get noticed, but in turn

they increase the overall chat activity and number of incomplete exchanges. In large quantities incomplete exchanges might appear as a waterfall of text, but it is an inaccurate description because they are intended as initiations for an orderly dialogue. Ignored exchanges can be formalized as $E_{ignored} = I_{written} \not\Rightarrow R$ with a crossed arrow indicating the breakdown of the exchange sequence. The initiations have prospection and demand responses, but they occur in a communicative environment where a response is not always given. Ignored exchanges appear to be uncommon in the slow-moving chat of the medium-sized broadcast but, they may even be the standard for large-sized live streams.

The other type of incomplete exchanges are interrupted exchanges, which only happen to streamer.

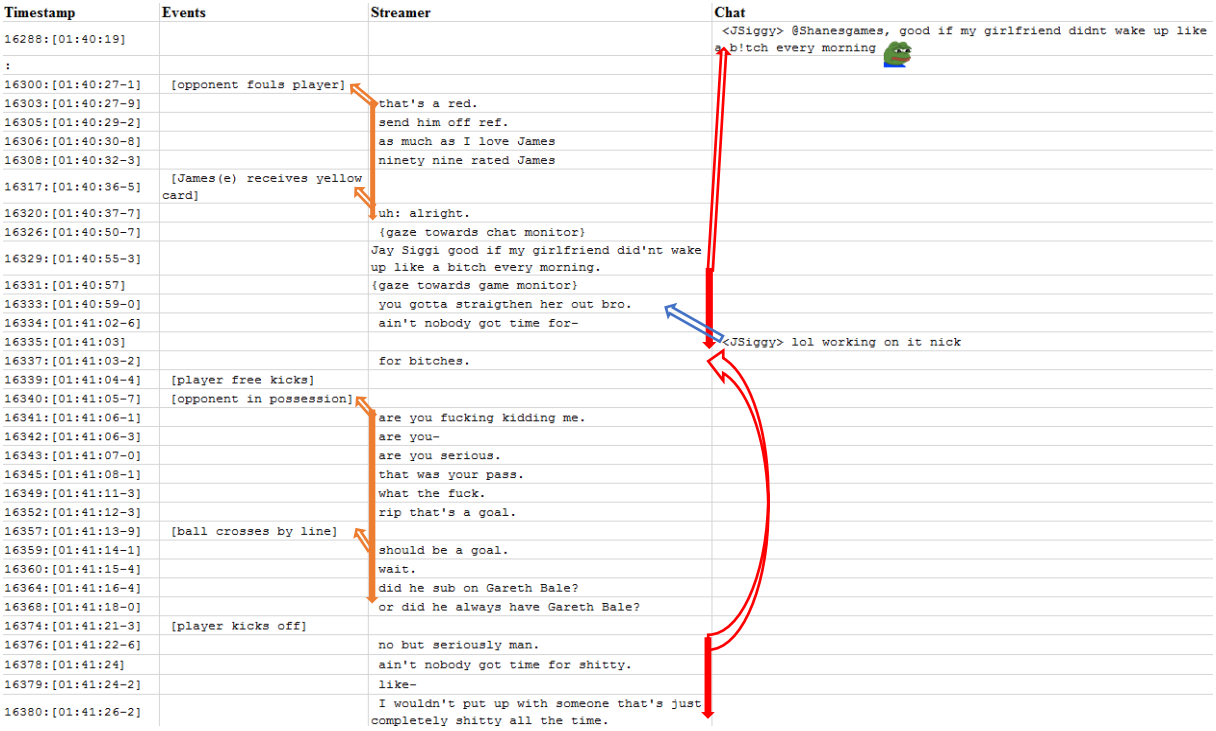
(12) Example – 'Dude watch the fuck out' from *World of Warcraft*

Timestamp	Events	Streamer	Chat
1037 :[00:41:41]			<dukuzz> what are your top 3 favorite characters in dbs rn?
:			
1054:[00:41:59-9]		what are your top 3 characters in dbs.	
1055:[00:42:00-9]		right now?	
:			
1091 :[00:42:44-2]	[Arena beginns]		
1093:[00:42:46-5]		{gaze to game monitor}	
1094:[00:42:48-6]		I don't really like berus.	
1095:[00:42:50-6]		he is kinda meh.	
1096:[00:42:51-3]		(wis) is cool.	
1097:[00:42:56-2]		I just go fucking Goku Wis Hit.	
1098:[00:42:57-3]		those are my favorite.	
1099:[00:42:58-3]		I like Vegeta too though.	
1101:[00:42:59-2]		I wanna see fucking brolly.	
1102:[00:43:00-3]		that's what I wanna see.	
1104:[00:43:02-6]		that make me happy dude.	
1105:[00:43:04-2]		fucking super sayin god brolly.	
1106:[00:43:05-0]	[opponent attacks team mate]		
1107:[00:43:06-0]		dude watch the fuck out.	
1108:[00:43:09-5]		I am a nerd dude.	
1109:[00:43:10]	[player is stunned]		
1111:[00:43:12-1]		I am actually dead too.	
1112:[00:43:13-3]		I am a dead nerd.	

Interrupted exchanges begin with a written initiation in the chat. In this example, the chat participant asks the streamer about his three favorite anime characters from the series "`dbs`(l. 1037, short for Dragon Ball Super)." The streamer is a fan of Japanese anime and after the topicalizer, he begins a very long response that discusses several characters of the show. He continues even as the arena match begins in line 1091. His team mate is suddenly attacked (l.1106) and he self-interrupts his response. The game event elicits an interactional re-orientation towards the gameplay. The game-related commenting consists of attending move "`dude`" to his co-player as well as a command "`watch the fuck out` (l.1107)." After this self-interruption, the streamer revisits the previous response with "`I am a nerd dude` (l.1108)". Importantly, the addressed "`dude` (l.1108)" is again dukuzz, the original initiator of the question from line 1037 and not the co-player, who has been addressed as 'dude' in line 1107. The statement thematically reconnects the detailed discussion of anime characters and

restarts the exchange with the audience member. The demands of the game have temporarily paused the streamer's response but as the game event passes, he can re-orient back to the audience. In this case, the self-interruption was brief and there was only one game-related utterance before the streamer quickly resumed the previous discussion. However, the time frames of self-interruption and resumption of responses can also be significantly longer, if there are several competing game events.

(14) Example – 'Ain't nobody got time for'



Timestamp	Events	Streamer	Chat
16288:[01:40:19]			<JSiggy> @Shanesgames, good if my girlfriend didnt wake up like a b!tch every morning
:			
16300:[01:40:27-1]	[opponent fouls player]		
16303:[01:40:27-9]		that's a red.	
16305:[01:40:29-2]		send him off ref.	
16306:[01:40:30-8]		as much as I love James	
16308:[01:40:32-3]		ninety nine rated James	
16317:[01:40:36-5]	[James(e) receives yellow card]		
16320:[01:40:37-7]		uh: alright.	
16326:[01:40:50-7]		{gaze towards chat monitor}	
16329:[01:40:55-3]		Jay Siggi good if my girlfriend did'nt wake up like a bitch every morning.	
16331:[01:40:57]		{gaze towards game monitor}	
16333:[01:40:59-0]		you gotta straigthen her out bro.	
16334:[01:41:02-6]		ain't nobody got time for-	
16335:[01:41:03]			<JSiggy> lol working on it nick
16337:[01:41:03-2]		for bitches.	
16339:[01:41:04-4]	[player free kicks]		
16340:[01:41:05-7]	[opponent in possession]		
16341:[01:41:06-1]		are you fucking kidding me.	
16342:[01:41:06-3]		are you-	
16343:[01:41:07-0]		are you serious.	
16345:[01:41:08-1]		that was your pass.	
16349:[01:41:11-3]		what the fuck.	
16352:[01:41:12-3]		rip that's a goal.	
16357:[01:41:13-9]	[ball crosses by line]		
16359:[01:41:14-1]		should be a goal.	
16360:[01:41:15-4]		wait.	
16364:[01:41:16-4]		did he sub on Gareth Bale?	
16368:[01:41:18-0]		or did he always have Gareth Bale?	
16374:[01:41:21-3]	[player kicks off]		
16376:[01:41:22-6]		no but seriously man.	
16378:[01:41:24]		ain't nobody got time for shitty.	
16379:[01:41:24-2]		like-	
16380:[01:41:26-2]		I wouldn't put up with someone that's just completely shitty all the time.	

Example 13 is taken from *FIFA 2015* and shows an exchange that is interrupted for 18 seconds by two game events (16337 – 16374). Chat messages that did not contribute to the discussion have been abridged. Prior to the exchange, the streamer is engaging in unrelated commenting and reporting (l.16003 – l.16320). There is a slowdown in the gameplay because of a foul (l.16300). The streamer uses this opportunity to transition from the game-related talk to the dialogical exchange. The sequence of the response consists of a gaze shift towards the chat monitor (l.16326), the topicalizer (l.16329) and a series of responding moves[33] that advise the viewer to take actions and sanction the behavior of his girlfriend (l.16333 – l.16337). The initiator positively receives the response and produces an appreciative follow-up (l.16335).

[33] The example was selected because of its suitable structure and not due to its content. Casual sexism against women is a very prevalent issue in gamer culture and *Twitch*'s chat (cf. Taylor, 2009; Anykey, 2016).

Before the streamer utters the epithet (l.16337), there is a slight hesitation in his utterance and he pauses after "`for-`", before he fully commits to the phrase "`for bitches` (l.16337)". This might suggest that, for a moment, he is considering the appropriateness of the term but is immediately consumed by the demands of the game (l.16339). His response is self-interrupted due to the bad pass (l.16345), which almost leads to a goal for the opponent (l.16352). The notable and sudden game events required an interactional re-orientation to the game. A second slowdown in line 16374 allows the streamer to restart the initial response. With the statement "`no but seriously man` (l.16376)", he tries to frame the previous discourse about 'problematic' girlfriends as playful and non-serious. There is an initial attempt at a reformulation "`aint noboady got time for shitty` (l.16378)", which is a partial repetition of line 16334. The last item that is missing is the noun that serves as the head of the noun phrase '`shitty +` (female)'. The streamer cornered himself syntactically and is looking for the right word (l.16379) before opting for a complete reformulation. This reformulation is less categorical and slightly less offensive because it replaces the offensive head of the noun phrase 'bitches' with a neutral had and an elaborating move, i.e. "`someone that's just completely shitty all the time` (l.16480)." With this reformulation, the streamer ends the exchange and there was no follow-up or any further reference.

Interrupted exchanges consist of a written initiation followed by a first segment of a spoken response that is self-interrupted due to an immediate game event. Afterwards, there is a restart of the response. It highlights that the streamer was not finished with the exchange and wants to resume his or her discussion with the previous audience member. Therefore, interrupted exchanges can be formalized as $E_{interrupted} = I_{written} \longrightarrow R_{spoken\ (interrupted\ \&\ restarted)}$.

Summary of exchanges

The study identified five types of exchanges, whose relationship can be visualized in the following typology.

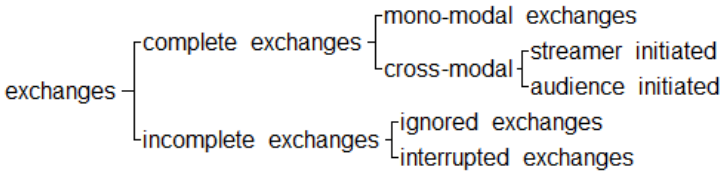


Figure 33 Typology of cross-modal exchanges

Complete and incomplete exchanges can be divided according to the criterion of success, i.e. is there a proper initiation that is met with a sufficient and complete response. A complete exchange is a successful dialogical unit, whereas incomplete exchanges features a disruption in the exchange structure. Exchanges are further divided according to their modal

properties. Mono-modal exchanges start and end in the chat, whereas cross-modal exchanges have a floor transfer across the communicative modes. This floor transfer can happen from writing to speech as in the case of audience-initiated exchanges or from speech to writing as in the case of streamer-initiated exchanges.

The mono-modal, intra-chat communication usually has the exchange structure:

$E_{mono\text{-}modal} = I_{written} \rightleftarrows R_{written} \rightarrow F_{written}$

Streamer-initiated cross-modal exchanges have the form:

$E_{cross\text{-}modal} = I_{spoken} \rightleftarrows R_{written}$

In both cases, a single initiation receives several written responses, which are easy to identify due to their textual cohesion and sequential implicature. This pattern of many Rwritten is possible because of the number of participants and the affordances of chat. Even if the chat messages are produced simultaneously, they appear as discrete units in the chat window and do not overlap (cf. Herring, 1999). Due to the higher number of participants, large-sized live streams will have an increased number of $R_{written}$ in comparison to the medium-sized broadcasts.

Audience-initiated cross-modal exchanges have the form:

$E_{cross\text{-}modal} = I_{written} \rightarrow R_{spoken} \rightarrow F_{written}$

They are different from streamer-initiated exchanges in form, number and purpose. Audience-initiated cross-modal exchanges are the most common type of complete exchange, because they are more frequent than streamer-initiated exchanges and initiations towards other audience members (cf. chapter 5; 'written chat to audience group'). This highlights the necessary distinction of cross-modal exchanges that originate in the spoken or the written mode, which has been absent in previous studies on video-mediated communication (Sindoni, 2014; Rosenbaun et al., 2016b).

A second reason for the prevalence of audience-initiated exchanges is the streamer's practice of exchange chaining, where they rotate between interlocuters and topics. This makes exchange chaining dialogical but also very different from ordinary turn-taking in face-to-face conversation. The chat participants are not taking the floor but rather it is given to them by the streamer. This turn-giving is realized by a new practice of spoken repetitions. There is a competition among audience-initiations and the streamer must highlight who s/he is responding to and what s/he is talking about. Both goals are achieved by the topicalizers that occur as the first move in a spoken response. Topicalizers raise a written message to the spoken conversational floor[34] and give the contribution much more salience in the live stream. There is

[34] This is makes them different "mode-switching (Sindoni, 2011)" or "upward participation (Rosenbaun et al., 2016) , where the video-chat participant can themselves shift into the spoken mode to gain more recognition (cf. section 3.1 video-mediated communication and participation frameworks.

very little variation in the form of the topicalizer an it is close to a translation of the written message to a spoken utterance. Variations between the wording of the initiation and topicalization are 1) corrections of spelling errors, 2) a full lexicalization of abbreviations, 3) omission or repetition of the attentive move, and the 4) reversal of the origo of the initiation.

1) $I_{written}$ "is their a" → R_{spoken} "is there a"

2) $I_{written}$ "better at adc" → R_{spoken} "better at ad carry"

3) $I_{written}$ "Erika, tell me how" → R_{spoken} "tell me how"

$I_{written}$ "nick can you" → R_{spoken} "Nick can you"

4) $I_{written}$ "do yousee yourself" → R_{spoken} "Do I see myself"

After the topicalizer, the streamer uses his or her oral monopoly to produce a response. Useful responses are more than a bare minimum answer to the proposition. They provide a thorough explanations to the game-related questions. Their elaborating moves are grounded in the game-expertise of the streamer and include significant amounts of gamer jargon (cf. Eggins & Slade, 1997; Enslinn, 2011; Gandolfi, 2016). The game-related questions are concerned with general scenarios and rarely relate to the moment-by-moment unfolding of the game. To the contrary, the immediate gameplay has detrimental effects on dialogical exchanges.

Many incomplete exchanges occur as the streamer has to focus on the match. In the faster-paced games, there is constant interactional demand by the game. As streamers play the match, they cannot turn their gaze towards the chat and many initiations are overlooked. This leads to the pattern of ignored Exchanges $E_{ignored} = I_{written} \not\Rightarrow R$, where written messages fail to elicit a spoken response. Alternatively, the streamer might begin his or her response but then practices an interrupted exchange $E_{interrupted} = I_{written} \longrightarrow R_{spoken\ (interrupted\ \&\ restarted)}$. The self-interruption and game-related discourse is licensed by the notable game event and temporarily takes precedent over the dialogical exchange (cf. Gerhardt, 2012). The restart of the response only occurs as it becomes interactionally feasible, usually after a slow down in the game. The different orientations between game and audience are constantly negotiated in the embodied gaze shifts and the patterns of discourse.

Dialogical exchanges are best understood as the type of discourse that occurs between participants and around game events**.** This 'around-ness' has a temporal dimension as well as topical one. The game frames the discourse, and the exchanges are occasioned by the game's components, rules or strategies. However, the exchanges do not directly relate to the sequential unfolding of the gameplay and they ccur in the absence of game events. This insight can explain

why there are more dialogical exchanges outside of a match, but it can also account for moment-by-moment transitions around a game event during a match (cf. Karhulahti, 2016).

Moreover, this perspective remains analytically useful for live streams that do not have clear separations into matches or rounds. For example, the communication of live streaming of eating is influenced by its own type of events such as the of unwrapping food items or chewing. Different activities have different interactional demands but the dialogical communication between participants will always revolve around stream-related events. It distinguishes exchanges from the more monological discourse that is elicited by the activity and is about the events, namely commenting and reporting.

6.2 Commenting and Reporting

This study identified commenting and reporting as two discursive practices that are about the immediate game events during a match. They have distinct set of features regarding their temporal proximity to the game event, the types of moves that are produced and the context dependence of discourse. In pointing out their differences, this section provides a clearer picture of the commentator function of the streamer, whose talk has been loosely described as "narrating what is happening, what parts of the game are good and what parts are bad, etc. (Smith et al., 2013, p.133)." This and similar worded descriptions are correct in their assessment but they are also too vague and benefit from more analytical precision.

(14) Example – go in on Thresh

Timestamp	Events	Streamer	Chat
298:[00:08:12-7]	[Aurelion Sol(e) kills Gragas]		
299:[00:08:22-2]		I really didn't wanna (go in on Thresh).	
303:[00:08:42-5]		damn they're so low.	
304:[00:08:43-5]		I didn't wanna go in on Thresh.	
306 :[00:08:45-1]		that's the problem	
307:[00:08:45-9]		I felt like if I went in on Thresh.	
308:[00:08:48-4]		we have a hundred percent gonna lose that fight.	
309:[00:08:50-4]		that's exactly what happened last time.	
310:[00:08:51-9]		and they're ahead.	
311:[00:08:53-5]		I don't think I can go for Thresh ever.	
312:[00:08:55-9]		if I use my combo on Thresh.	
313:[00:09:00-6]		I have nothing for Aurelion or Ez.	
314:[00:09:02-0]		and I think we just lose then.	
315:[00:09:13-9]		man they were so low.	

Example 14 is taken from *League of Legends* and is a typical instance of reporting. It features a stretch of talk that is lasting over one minute and it is about a single game event. After a fight with a negative outcome, the streamer reports on his decision not to 'go in on Thresh'. Twice, He raises it as the topic of the segment (l.299; 304), and then justifies his actions with his assumptions and game knowledge.

304 I didn't wanna go in on Thresh
307 I found like if I went in on Thresh.
308 we have a hundred percent gonna lose that fight.
309 that's exactly what happened last time

In line 304 and 307, the underlined verbs phrases are explicit references to his 'epistemic status' of what he knows or believes to be true (Heritage, 2012). The streamer calculated the risks and consequences of 'going in on Thresh' and concluded that it would be a certain loss. This inner thought process is verbalized via elaborating moves that explain the most likely outcome based on past experiences (l.307 - 309). Reporting gives the audience insights in the thought process of the streamer-as-player so that they can understand the reasons for his or her decisions.

The discourse is used as an explanation similar to some of the previous dialogical exchanges. Reporting is also comparable these exchanges with regard to its significant use of in-game jargon and its reliance on tacit knowledge (Hornyak, 2016).

312 if I use my combo on Thresh.
313 I have nothing for Aurelion or Ez.

To understand the statement, it is necessary to know the specific meaning of the term 'combo' in the context of this match as well as the implication of using the 'combo on Thresh'. Abilities in *League of Legends* have long cooldowns that make them unusable for a certain amount of time. Using the 'combo on Thresh' triggers this cooldown and would make the player vulnerable for counter-attacks by other avatars such as 'Aurelion' or 'Ezrael'. By inference, the better game strategy is not to use 'combo on Thresh' and keep the combo for more important game events. The spectatorship of gameplay paired with such an informative discourse leads to informal learning about the game (Kow & Young, 2013; Georgen et al.,, 2015). This study generally agrees but must qualify this assessment by highlighting an important difference between explanatory exchanges and reporting. Exchanges are initiated by questions of an audience member and receive responses that are tailored to their question. Reporting is grounded in unique game events and its circumstances. As a result, the information might be less generalizable to other matches. The spectator must apply inductive reasoning to draw conclusions from the specific instances of reporting. Exchanges and reporting both use game-specific jargon to expresses complex game-specific meanings, but the monological reporting is less instructive, more implicit and more requires more inferences by the spectator.

While reporting was relatively common in the slow-paced game *League of Legends*, there were fewer such instances during a match of the faster-paced games *FIFA 2015* and *World of Warcraft*. Reporting consist of a series of reflective discursive moves that requires longer pauses between game events. The transcript of the next example makes differences between reporting and commenting immediately visible.

(15) Example – pass it up the field

Timestamp	Events	Streamer	Chat
12094:[01:02:02-9]		wake up boyz.	
12100:[01:02:07-6]	[player in possession]		
12105:[01:02:12-4]		okay.	
12106:[01:02:13-7]	[opponent in possession]		
12111:[01:02:15-0]		fuck dude.	
12112:[01:02:16-0]		>WHAT THE FUCK<	
:			
12369 :[01:05:01-8]	[ball crosses touch line, opponent in possession]		
12370:[01:05:01-9]		NO:.	
12374:[01:05:03-3]		you fucking morons.	
12376:[01:05:05-3]		pass it up the field	
12379:[01:05:07-0]		what the fuck.	
12383:[01:05:09-8]	[player in possession]		
12384:[01:05:10]			<%FCSPJoern>
12385:[01:05:10-7]	[opponent in possession]		
12387:[01:05:12]			<MATtheTANK> lol
12389:[01:05:12-0]		o:h my god.	
12391:[01:05:13-3]		my brain.	
12392:[01:05:14]			<Tjenks7>
12394:[01:05:15-9]		oh lord.	
12401:[01:05:17-3]	[player in possession]		
12402 :[01:05:19]			<Fifa_ashton>
12403:[01:05:19-2]	[opponent in possession]		
12408:[01:05:20]			<%ALinkzYT> LOL nice pass
12409:[01:05:21]			<Ra2wu>
12412:[01:05:22-1]		oh for fucks sake dude.	
12413:[01:05:24]			
12414:[01:05:25]			<%Luisito91987>
12415:[01:05:26]			<DClover03>
12416:[01:05:27-9]		DEFENDer.	
12417:[01:05:29-5]		OH GOD.	
12418:[01:05:30-7]	[player in possession]		
12420 :[01:05:32-5]		yes okay go go go.	

The excerpt is taken from a match of *FIFA 2015*. The streamer is fully engaged with the play and there is a constant back and forth between the two teams. The ball possession is constantly changing, which is visible in the transcript through a high number of game events. Each game events receives a brief comment by the streamer. The typical purpose of comments are reacting moves or response cries (Goffman, 1981; Stenström, 1994). They display a general emotional stance of the speaker towards the commented game event. Positive response cries are elicited by game events that lead to positive outcomes for the streamer;

```
12418 [player in possession]
12420 yes okay go go go
```

while negative game events elicit a more negative reaction.

```
12106 [opponent in possession]
12111   fuck dude.
12112   >WHAT THE FUCK<
```

For many streamers, this reaction is almost instantaneous, and a given series of game events will lead to a corresponding series of response cries. The impact of the game event for the streamer-as-player predicts the linguistic reaction of the streamer-as-caster and the discourse reflects the developments in the fast-paced game. The stretches of commenting are more fragmented because every response cry is only relevant to its most immediate game event and response cries do not feature elaborating moves that construct an overarching topic.

Commenting has a primarily social purpose. It contains colorful language and is rife with expletives and epithets. The response cries are an entertaining performance designed to amuse the audience via exaggerated and emotive discourse (cf. Fernández-Vara, 2009; Nylund, 2015; Pellicone, 2016, 2017). During commenting, streamers react like an ordinary player and their emotional outbursts are relatable for other gamers and are a vicarious experience (cf. Aarsand & Aronson, 2009; Conway 2013; Glas, 2015).

The audience is also commenting and is producing reacting moves with a focus on interpersonal meanings. They rely on emotive written messages such as "`lol` (l. 12387)" or use *Twitch's* emoji. *Twitch*'s emojis have a fixed meaning across all games. Often, they are iconic images that display emotional states such as embarrassment  (l.12384) or amusement  (l.12414). Other emoji are visual metaphors and use inside-jokes, for example the salt shaker  (l.12392, 12415) stands for being angry or crying. People produce tears when they are sad or upset and these tears are salty. The salt shaker is a playful mocking of the streamer's exaggerated response cries to the negative game events. The audience is performing their spectator role and signals involvement with the streamer and the unfolding game events (cf. Tannen 1987; Cheung and Huang, 2011).

This type of discourse does not amount to a full conversation or dialoge, since commenting has no prospection and does not expect a proper response. Nevertheless, it is also not random or disruptive to the broadcast (cf. Hamilton et al., 2014). It only occurs in close temporal proximity to game events and both parties produce intersubjective reacting moves that foreground emotional and interpersonal meanings.

The micro-level study also reveals segments of commenting that contain discursive moves that behave differently from response cries. It is a discourse that is even more anchored around game events. This study will refer to them as pivoting moves and will explain them via example 16.

(16) Example – Disgusting roam

Timestamp	Events	Streamer	Chat
2099:[00:17:43-7]		°holy shit our supports here°.	
2105:[00:17:53-7]	[Player kills Ahri(e)]		
2109:[00:17:58]			<lyrute> lol ez
2110:[00:17:58-0]		holy shit.	
2111:[00:17:58-8]		what a disgusting roam.	
2113:[00:18:02]			<iwiggs> shadow op
2114:[00:18:02]			<thankyouflyingpotato> rekt
2116:[00:18:03]			<john95x> god morg
2118:[00:18:04]			<fataldigits> level 2 roam LOL.
2125:[00:18:07]			<xkah06x> lol
2126:[00:18:07]			<junglecleave> dat roam
2129:[00:18:08]			<ichorrr> dat gank doe
2130:[00:18:08]			<metalfrider> LOL BM AHRI
2131:[00:18:09]			<lordgecco> rip
2133:[00:18:10]			<1breadstick> that roam
2140:[00:18:12]			<paulostein> 2EZ
2141:[00:18:12]			<chocolate_metaphor> omfg that morg

The excerpt covers a segment of commenting that is taken from *League of Legends*. It features a brief combat between the streamer, an ally team member 'Morgana' and an opposing player with the avatar 'Ahri'. Throughout the segment the streamer is physically oriented towards the game and is not directly communicating with the audience.

In line 2100, the streamer performs a "noticing (Schegloff, 1988)" or alerting move (Stenström, 1994) that highlights the presence of his fellow team mate. The use of profanity and softness of his utterance `°holy shit our support is here°`(l. 2100) suggest surprise (cf. Wilkinson & Kitzinger, 2006). Together, the streamer and the friendly player defeat 'Ahri' (l.2105) and afterwards, many audience members are commennting with reacting moves / response cries:

```
2109 <lyrute> lol ez
2110 Streamer: holy shit.
2114 <thankyouflyingpotato> rekt
2125 <xkah06x> lol
2131 <lordgecco> rip
2140 <paulostein> 2EZ
```

They are casual & colloquial messages that contain very little jargon. For example, '`rekt`' (short for wrecked), '`rip`' (short for rest in peace) and '`EZ`' (short of easy) are non-literal and playful ways to comment on a player's victory. The messages are very unspecific and explain very little about the game event. This distinguishes them from pivoting moves:

```
2112 What a disgusting roam
2119 <fataldigits> Level 2 roam LOL
2127 <junglecleave> dat roam
2130 <ichom>dat gank doe
2134 <Ibreadstick>that roam
```

Pivoting moves discuss a local meaning for the game event. This process of sense making involves significant amounts of gamer jargon and game knowledge. In the example above, the most used nouns are 'roam' and 'gank'. The participants produce a spontaneous interpretation

of the game based on their prior experience. They classify the unique encounter as a recognizable game strategy. 'Gank' (short for Gang Killing) and 'roam' are near synonyms and describe a series of actions, where a player assists a team mate in combat. To do this, the player must 'roam' from his or her lane (see figure 21, p.70) to assiste the team mate. Together, they 'gank' the opponent and defeat him or her.

Through the pivoting moves, the game event on-screen becomes a 'gank' or a 'roam' in the discourse. Pivoting moves attribute game events with specific meanings. This can be a general classification as 'gank' or 'roam' as well as more particular and notable qualities. For example, the streamer signals that the roam was surprising (l.2099) and "`disgusting` (l. 2112)". Another chat participant describes it as a "`Level 2 roam LOL` (l. 2119)". The discourse marks the game event as extraordinary because it happened very early into the round before the start of the team fight phase (see table 9; p.70). This inference requires extensive knowledge of *League of Legends* and it shows that pivoting moves are more than emotive response cries. Aside from the emotional component, they also possess a proposition or idea that is conveyed, and they classify a unique game encounter as a recognizable type of game event.

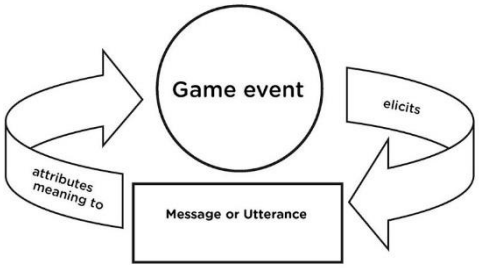


Figure 34 Pivoting moves

The figure demonstrates the conceptual schema for pivoting moves. Salient and notable game events elicit responses from the participants in the form of written messages or spoken utterances. The moves attribute the game event with a local meaning. This local meaning depends on the participant's interpretation of the game event at the backdrop of their game knowledge and experience. There does not seem to be a finite list of game events that can be highlighted in this process. For example, in *World of Warcraft* chat members commented on a game event as "`amazing kiting to win`". 'Kiting' describes a technique of running away from one's opponent, where it becomes difficult for them to attack the player.

In *FIFA 2015*, participants commented on "`finesse`" free kicks and goals. Pivoting moves are different in every video game, because each video game is a different frame of reference. The only constant is the interactional unfolding and its general function. It is elicited by the game event on screen and comments about it in discourse.

Summary of reporting and commenting

Section 6.2 presented reporting and commenting as the two main discursive practices that discuss the moment-by-moment unfolding of the game. Reporting and commenting are the types of discourse that occurs in the immediate temporal proximity of game events and is directly talking or chatting about them.

Reporting features a stretch of talk of several moves that is elicited by a single game event. It is exclusive to streamers and is reflective and explanatory. It verbalizes their thought process and allows the spectators to comprehend the streamers decision making in the gaming situation. It is more common in slow-paced games, because they provide the streamer with enough time to formulate his or her thoughts.

Commenting is a more fragmented discourse and occurs in fast-paced games. Each move is referencing a different game event and is written or uttered as a quick reaction quickly after it. Its purpose is to amuse the participants with exaggerated performances of player and spectator roles. Commenting is realized by response cries / reacting moves and pivoting moves. Reacting moves are mostly emotional & interpersonal and their meaning is not game specific. Pivoting moves are specific to games and game events and categorize the game event as a strategy or gameplay practice. Commenting is very prevalent in the medium to fast-paced games. The quick succession of game events puts on a continuous cognitive and interactional demand. It is very difficult for the streamer to play competitively on an extremely high level and to formulate their thoughts and decisions in a clear manner.

The analysis did not reveal any differences in the streamer's discourse that could be attributed to audience size. Neither streamer-reporting nor streamer-commenting appear more likely in large-sized broadcasts vis-à-vis medium-sized ones. There is however, a notable higher volume of audience-commenting, which can be explained with the higher number of chat participants.

6.3 Transaction and interaction

Above the level of exchanges, reporting and commenting the description of discourse becomes increasingly difficult because there can be considerable variation. Nevertheless, there are patterns in the discourse of live streaming, which can be described by utilizing the concept of transaction and interaction. Transactions have a temporal and content dimension because they are a series of exchanges that are bound together by a coherent topic. The interaction consists of a series of transactions.

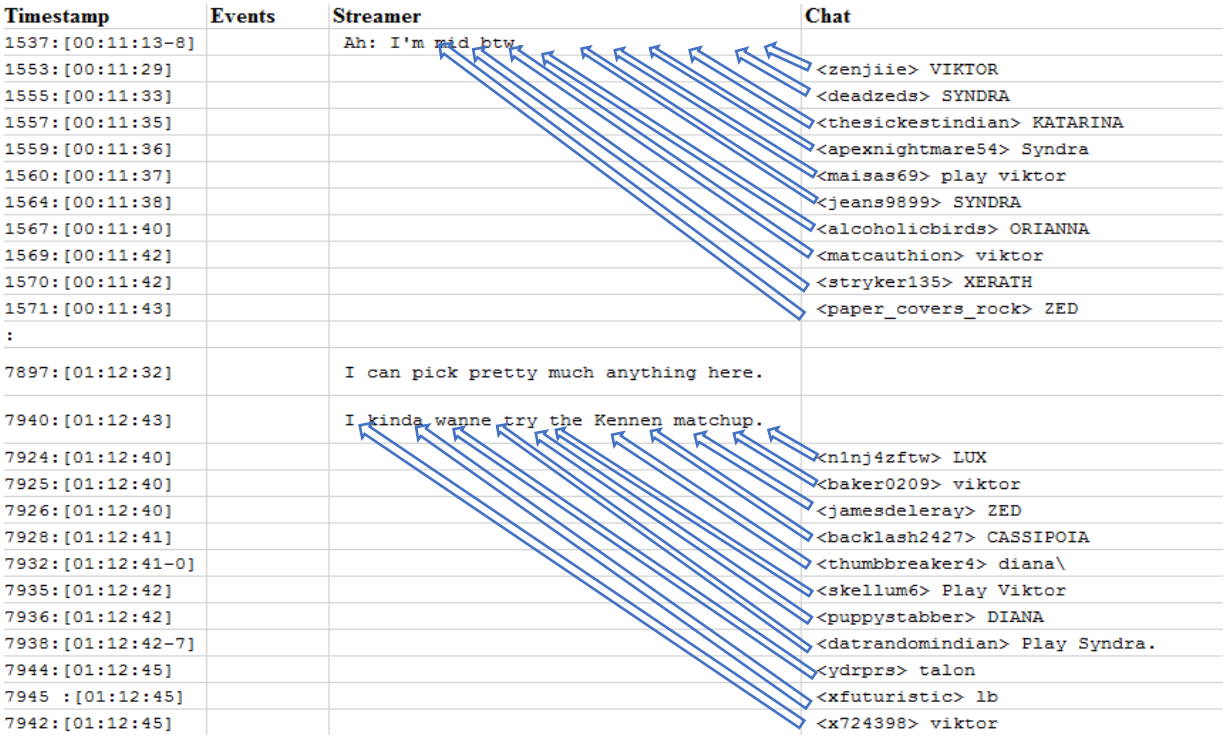
Transactions and interactions on *Twitch* repeat regularly in a systematic manner and 6.3

will argue that they have a cyclical organization. The smallest cycle of topic recurrence is tied to the game round.

The cycle of the round

Rounds are an important organizational principle of video games (Zagal et al., 2008). Section 4.1.1 (p.70ff.) described the temporal organization of the three games used in this study. The introduction showed that these games have a period of team preparation that precedes every round. During this time, the streamer decides which avatars s/he wants to play. This avatar can be a fighter in *League of Legends* or a soccer player in *FIFA 2015*. The transactions in the discourse orient towards the game stages.

(17) Example – Avatar Selection



The abridged example is taken from *League of Legends* and shows the discourse during two avatar selections. The two selections are separated by one round of the game, which lasted slightly less than 1 hour. Despite this temporal gap between line 1571 and line 7891, the transactions have the same form and functions. Streamer and audience discuss the upcoming avatar selection. This stage is one of the few opportunities for the audience to directly influence the gameplay and they vividly request the avatar that they would like to see. The requests mostly consist of the fighter's name, often written in capital letters. The game's stage contextualizes the communication so tightly that no further elaboration by the audience is necessary. These truncated requests only make sense at the 'character selection' stage of the game and do not

occur in other stages[35]. Other stages have their own typical discourse such as commenting & reporting, which only occurs during the match. As different game stages repeat periodically throughout the rounds, the transactions repeat as well.

(18) Example - Cross-modal exchanges and commenting across two rounds

Timestamp	Events	Streamer	Chat
444:[00:16:05-7]	[arena beginns]		
445:[00:16:05-7]		{gaze to game monitor}	
446:[00:16:11-6]		Paladin.	
449:[00:16:15-1]		and monk	
492:[00:18:43]			<Dingodrunk> is outlaw fun? i never see anyone play it
499:[00:19:08-1]		fucking bitch.	
501:[00:19:11-1]	[Player dies & loses arena]		
502:[00:19:12-6]		{gaze to chat monitor}	
504:[00:19:18-7]		is outlaw fun?	
505:[00:19:19-3]		it just doesn't do damage	
507:[00:19:21-6]		that's the downside to it.	
510:[00:19:28]			<RelapseGalore55> is ret strong atm?
512:[00:19:30]			<theSGfire> This is like the worst comp world
519:[00:19:40-4]		is ret strong?	
520:[00:19:41-4]		oh fuck yeah.	
521:[00:19:42-0]		ret is really strong dude.	
523:[00:19:45-8]		this is like the worst combo in the world {laughs}.	
524:[00:19:48-2]		eh you know it is not that good.	
526:[00:19:51-3]		that is for sure.	
531:[00:20:03-6]		{gaze to game monitor}	
536:[00:20:10-3]		ret rogue.	
551:[00:20:32-3]	[arena beginns]		
554:[00:20:45-3]		I got rogue.	
559:[00:21:01-5]		kick the pally.	
560:[00:21:04-1]	[player kills opponent]		
561:[00:21:04-4]		nice.	
563:[00:21:11]			<StefanSWAG> you are a crazy mf
564:[00:21:24-0]		rets are no joke.	
565:[00:21:24-3]		So I am not gonna fool around with him.	
566 :[00:21:32-5	[Player wins arena]		
567:[00:21:34-3]		{gaze to chat monitor}	
568:[00:21:37-0]		you are a crazy motherfucker.	
571:[00:21:39-5]		thank you.	

The example features two consecutive rounds of *World of Warcraft arena*. In the first round (l.444 - 501), the streamer's team loses versus a team of a paladin and monk. In the second fight (l.551-566), he wins versus a 'ret' and a 'rogue' player. The discourse patterns of cross-modal exchanges and commenting repeat across the two rounds.

	First Round	Second Round
Start of arena	446 Paladin	536 ret rogue
	449 and monk	
	:	:
End of arena	499 fucking bitch	561 nice

At the start of each round of arena, the streamer names the classes of the two opponents. Afterwards, the combat of the match starts, and the streamer is commenting on the unfolding

[35] Example 5 from FIFA 2015 (section 6.1.1, p.123) showed the same linguistic behavior at the same type of game stage.

gameplay[36]. At the end of the round, the streamer is commenting on victory as '`nice` (l.561)', whereas he insults his opponents after his loss (l.499).

Since *World of Warcraft arena* is very fast-paced, the dialogical exchanges occur as exchange chaining between rounds.

I_1 492 <Dingodrunk> is outlaw fun? i never see anyone play it?

R_1 504 is outlaw fun?
505 it just doesn't do damage
507 that's the downside to it.

I_2 510 <RelapseGalore55> is ret strong atm?

R_2 519 is ret strong?
520 oh fuck yeah.
521 ret is really strong dude.

I_3 512 <theSGfire> This is like the worst comp world

R_3 523 this is like the worst combo in the world?
524 eh you know it is not that good.
525 I'll give you that one dude.
526 that is for sure.
528 definitely not that good.

Every spoken response begins with a topicalizer that repeats the previous written initiation (l.504, l. 519, l.523), followed by the answer to the question. The streamer is chaining the exchanges in consecutive order and this pattern is repeated immediately after the victory in the second round.

I_4 563 <StefanSWAG> you are a crazy mf

R_4 568 you are a crazy motherfucker.
571 thank you

The content of the discourse of *FIFA 2015*, *League of Legends* and *World of Warcraft* is different, but the underlying cyclical organization remains the same. The difference lies in the length of a cycle, which is due to the duration of the rounds. The fast-paced game *World of Warcraft* has short rounds of one to three minutes (see example 18), whereas the slow-paced League of Legends has cycles that take around one hour (see example 17). Therefore, the topics and patterns of transactions in *World of Warcraft* repeat every few minutes, whereas the discourse in *League of Legends* repeats every hour.

[36] The commenting segments are abridged due to their length

Above the level of the game round is the daily broadcast. Each broadcast consists of many rounds that are chained in a series

The cycle of the daily broadcast

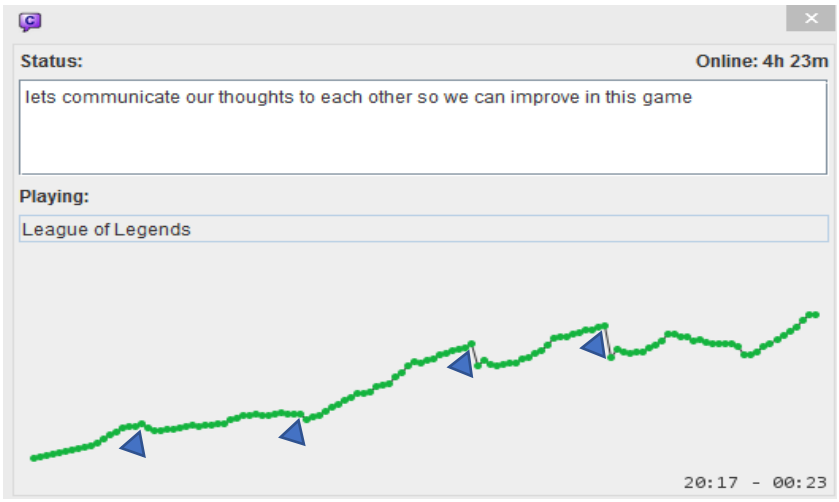


Figure 35 Development of viewer numbers

The figure is an annotated screenshot from the chat program *Chatty*, which was used to log the chat data. The screenshot shows the typical development of viewer numbers of a broadcast. The x-axis represents the passing of time and the y-axis symbolizes the increase or decrease in viewers in the channel. The blue triangles highlight the peaks in viewership, which occur at the end of each round. This peak is followed by a sharp decline because many viewers temporarily stop watching, while the streamer is preparing for the next round. Over the course of the next round, they return and new viewers join the broadcast. This leads to the typical pattern of growth in viewer numbers throughout the daily broadcasting. Usually every peak is higher than the previous one.

Since streamers are interested in a high number of viewers, the audience's behavior incentivizes long broadcasts with short pauses. For professional and aspiring full-time streamers, this conflicts with their basic physical needs such as food, sleep or going to the bathroom. These needs must be negotiated with the game rounds and the overall broadcast.

(19) Example – ‘make it through the whole thing’

Timestamp	Events	Streamer
614:[00:17:00-2]	[player wins match]	
:		
650:[00:18:09-7]		you know I may need to put in some food honestly.
651:[00:18:11-0]		I am getting kinda hungry.
652:[00:18:12-9]		I only had a what's it called.
653:[00:18:13-9]		I had like an Arby's sandwich.
655:[00:18:15-7]		{gaze to game monitor}
656:[00:18:15-7]		right before I streamed.
657:[00:18:18-2]		but that was.
658:[00:18:18-7]		I mean I am not too far into the stream.
660:[00:18:19-8]		{gaze to chat monitor}
661:[00:18:19-8]		but I gotta keep like.
662:[00:18:21-5]		at a normal state you know.
663:[00:18:22-8]		to make it through the whole thing.

In this excerpt, the streamer has just won a round of *FIFA 2015* and informs his audience that he is becoming hungry because he ate too little before the live stream[37]. This puts him in a problematic position because he cannot cook, play and talk at the same time. Cooking requires prolonged pauses and even eating a snack takes several minutes. In the context of live streaming, pauses are disruptions of the broadcast session and are usually explicitly justified by the streamer.

The argument by the streamer is notable because it frames eating in relationship to the live stream. He consumed food in preparation for the broadcast (l.656) but it was not enough for him (l.657). A pause for a meal would be too early because he is ‘`not too far into the stream` (l.658)’ yet. There is an expectation of how often and how early a streamer can pause during a broadcast and the pause must be justified to the audience. The streamer must eat ‘`to make it through the whole thing` (l.663).’ Eating is a tool to sustain the daily broadcast sessions and meal times are adjusted to the demands of the game round and the ongoing broadcast. It must occur in the pauses between rounds and these pauses must be brief. The example demonstrates that there is a pressure to play many games in a row and continuously interact with the audience.

Streamers reportedly aim for a broadcast duration of at least 3-4 hours[38] and the professional streamers in this study play for six to ten hours per day. Once enough rounds have been played, streamers end their daily broadcast.

[37] Arby’s is an American fast food chain.

[38] https://www.reddit.com/r/Twitch/comments/2vg3ey/streaming_regulars_how_long_is_your_average/

(20) Example – ‘tired & got to go to the bathroom, I’ll see you guys tomorrow’

Timestamp	Events	Streamer	Chat
1933:[01:39:11-3]	[Enemy team wins]		
1948:[01:39:41]			<loingirdler> Tomorrow we have a diamond to masters stream?
1956:[01:39:55-5]		I think the tiredness and the fact that I really got to go the bathroom	
1957:[01:39:57-7]		{gaze to chat monitor}	
1959:[01:40:00-4]		Diamond to Master's stream tomorrow.	
1962:[01:40:02-2]		Sure looks like it.	
1968:[01:40:15-1]		alright guys I'm gonna call it there.	
1969:[01:40:16-1]		I don't really feel like playing.	
1971:[01:40:18-0]		obviously.	
1972:[01:40:19-6]		I am tired.	
1974:[01:40:20-3]		I gotta go to the bathroom	
1978:[01:40:24-9]		I'll be on tomorrow.	
1979:[01:40:26-1]		what is today anyways.	
1980:[01:40:27-8]		Wednesday now?	
2009:[01:41:37-1]		I'll see you guys tomorrow.	
2010:[01:41:38]		I guess Diamond to Master stream tomorrow.	
2011:[01:41:40-3]	[Stream Ends]		

The example begins after the streamer lost a round of *League of Legends*. It is possible to queue up for another round or stop playing. In this example, the streamer prefers to end his broadcast. Notably, is the justification as to why no further rounds can be played. Twice in quick succession, the streamer states that he is physically unwell because he is tired and needs to go to the bathroom (l.1956; l. 1972; l. 1974). Streamers almost always give reasons for ending their broadcast, which indicates that there is a sense of obligation to continue.

The continuation is provided in the form of consecutive live stream on the following day. This daily cycle is expected by the participants. The chat member initiates a cross-modal exchange and asks if the streamer will broadcast the next day and continue his progress within the game’s internal ranking system, ‘`tomorrow we have a diamond to masters stream` (l.1948)’. The streamer topicalizes the question (L.1959) and affirms it with ‘`sure looks like it` (l. 1962)’. A new broadcast the next day is normal and expected. The progression to ‘master’ was not reached today and there must be a new broadcast. This interpretation is further supported by the streamer’s final statement ‘`I guess diamond to master stream tomorrow` (l.2010)’. The verb phrase *I guess* is often used to hedge complying moves (cf. Eggins & Slade, 1997), when speakers are not fully agreeing but feel obliged to do so. Tomorrow’s broadcast is unavoidable and almost mandatory and there is no happiness or enjoyment in the announcement of the streamer that he is going to play the video game the next day.

Importantly, this organization of day-after-day is not a consequence of the structural organization of the video game. In theory, the streamer could turn off the broadcast, go to the bathroom, rest for a while and then resume his gaming session. However, this is not practiced and instead the broadcast has a logic that is similar to work life. Every day, there is a single broadcast session that spans many hours with few and short breaks. This pattern optimizes the viewer numbers over the duration of the broadcast (see fig. 35).

Once a broadcast ends, it is finished for the whole day. This organization of daily

broadcasts combines into the highest level of non-discursive organization, i.e. the weekly streaming schedule.

The weekly streaming schedules

Streamers organize their broadcasts according a weekly streaming schedule

Figure 36 Streaming schedule

The figure shows the weekly streaming schedule of two streamers used in this study. On the left is an US American streamer, who plays *League of Legends* and on the right is a Canadian broadcaster, who is playing *FIFA 2015*. In their profile, they mention the weekdays they broadcast as well as their usual starting time of '3:00PM CST' on the left or '4-6PM UK' on the right. The starting times are set in the afternoon or early evening of the target audience. Since soccer is more popular in Europe than in North America, the Canadian *FIFA 2015* streamer begins to broadcast at 8 or 9AM of his local time, which corresponds to 4-6PM in the United Kingdom. The viewership of each region is highest in the late afternoon or early evening when audience members have left school, University or work (Kaytoue et al., 2012). Streamers plan around the leisure time of their viewers as it increases the number of their spectators.

Many streamers will broadcast six or seven days each week and if they take a day off, it is usually on Friday's, Saturday's or Sundays. Their weekly streaming schedule has the cycle of a work week. This idea has also been expressed by one the live streamers of the study.

(21) Example – play all day dude, seven days a week

Timestamp	Events	Streamer	Chat
7402:[01:09:32]			<symphee> How is the life of a pro player ? Would change being a pro player ?
7432:[01:09:47]		life of a pro player.	
7434:[01:09:48]		it's alright.	
7455:[01:09:52]		we just fucking play all day dude.	
7460:[01:09:53]		I wake up.	
7474:[01:09:56]		and I eat.	
7479:[01:09:57]		and go to the office.	
7484:[01:09:58]		and then we practice from the office.	
7491:[01:10:01]		for like 8 hours.	
7494:[01:10:02]		9 hours.	
7495:[01:10:02-2]		and then we go back home.	
7500:[01:10:04]		I play solo queue stream.	
7522:[01:10:14]		I don't really get to have a lot of free time	
7528:[01:10:15]		we practice five days a week.	
7532:[01:10:17]		and we have the LCS.	
7537:[01:10:19]		two days a week.	
7552:[01:10:23]		so use all.	
7558:[01:10:24]		seven days a week.	

A chat participant initiates the cross-modal exchange and asks about the life as a professional video gamer and streamer (l.7402). The streamer topicalizes the question (l. 7432) and his response provides a detailed account of his daily and weekly routine as professional gamer and live streamer. The first move of the response characterizes the life of a live streamer as '`alright` (l.7434)'. He is neither enthusiastic nor endorsing. His daily activity consists of over 8 to 9 hours of continuous play, only interrupted by the necessary sleep and food (l.7460; l.7464). This daily pattern repeats during the whole work week and is only alternated on the weekend, when training and streaming are replaced by professional online tournaments (l.7528, l.7537). All seven days of the week are '`used` (l.7552)' for optimal efficiency as a player and streamer. His response succinctly summarizes the different cycles of repetition from individual game rounds, over training and broadcasting days to the complete organization of a gaming week.

This non-linguistic organization of play also corresponds to the linguistic organization of discourse. On the smallest level, the individual round of play has stages with corresponding stages of discourse. For example, as game events occur during a match, the streamer's and audience's moves combine to segments of commenting and reporting. In absence of game events, participants engage in exchanges and viewers will initiate questions about the game. If there are no pressing game events or if there is a pause between rounds, the streamer may chain several exchanges in a series to talk to as many viewers as possible. Game rounds are repeated throughout the daily broadcast session, which can last anything from 6 to 10 hours. During this period, pauses for food or the bathroom are minimized to maintain a constant interaction with the audience. Once the streamer turns off the gaming broadcast, the communicative interaction ends. It starts again the next day with the beginning of the new broadcast. Lastly, the daily broadcasting sessions are repeated in a weekly streaming schedule that has the organization of a work week. All this makes live streaming an extremely repetetive activity and most of its

discourse follows this predictable cyclical structure. The cyclical organization is a universal feature and it is independent from the type of game or the size of the spectating audience.

The only discourse that operates differently are paid alert messages. They do not follow a cyclical structure and are not tied to game events. Their special position and role in the discourse makes them an important topic that requires a separate discussion in the seventh chapter.

Chapter 7- Paid alert messages

7.0 Initial remarks

Alert messages are a very recent phenomenon. They were first invited by the company 'streamlabs' in 2014 and at that time, alert messages were exclusive to *Twitch* and there was only one program that could display them. Initially, they were not planned as part of my research. However, throughout the last four years, alert messages have become ubiquitous on *Twitch* and they are also emerging on other live streaming services such as *YouTube Live* or *Twitter*'s *Periscope*. As the study progressed it became clear that alert messages cannot be explained in reference to game events or the organization in rounds or the daily broadcast. At same time, it was obvious that they had an immediate impact on the organization of discourse and that it was necessary to account for them. Therefore, the study formulated RQ_3 which asks about the purpose of alert messages and their position in the organization of discourse. This issue will be addressed in chapter seven.

Alerts messages appear in middle of the live stream window (see fig. 37, p.154). To produce such a message, a viewer must visit a third-party website and pay a certain amount of money, which usually fluctuates between $3-5USD per message. They are an interactional practice of their own kind that cannot easily be compared to other patterns of communication. Previous research on communication on *Twitch* often treats them as a side note ând describes hem only a few sentences (Hamilton, et al., 2014; Plath, 2015; Karhulahti, 2016). On *Twitch*, alert messages are usually distinguished in 'subscription notifications' and 'donation alerts', and these terms are increasingly adopted in academia. Research tends to treat them as genuine donations and designs questionnaires that ask for 'donations' motivations (Raes, 2015; Sjoblöm & Hamari, 2015; Groß et. al, 2017). These studies are closed-question online surveys that may have an affirmative bias towards the conventional narrative surrounding 'donation alerts'. The studies consistently report that viewers see the alert messages as 'donations' that support streamer (Raes, 2015; Sjoblöm & Hamari, 2015; Groß et. al, 2017).

So far, there has been no research that uses authentic alerts messages as primary data and there are no studies that discusses them as a form of viewer-streamer interaction. This chapter will address this gap by looking at 150 collected alerts messages through the lens of the discourse analytical framework. The chapter will begin with an explanation of alert messages as discursive moves. They are initiations sent by the viewers to the streamer. However, their process of text production is more complicated than ordinary chat messages. 'Subscription notifications' and 'donation alert' messages are produced through separate mechanism.

Section 7.1 will thoroughly explain the required steps for sending an alert message and how this process impacts the textual format of the message on the screen. In turn, the format of

‘subscription notifications’ and ‘donation alert’ messages influence their use in discourse as well as their treatment by the streamer.

Section 7.2 will discuss subscription notifications as a tool to discursively construct a community within the live stream. The notification is celebrated as rite of passage of a viewer from ‘pleb’ to subscriber and community member. The section will present the typical features of this performance and discuss the streamer’s work to maintain their ‘community’.

7.3 is dedicated to alert messages that ask the streamer for information or request them to do certain actions. These initiations receive long and extended responses that go beyond the usual practice of cross-modal exchanges. The subsection will argue that these types of messages are more comparable to a paid service than a pro-social donation that ’support’ streamers (Ariely et al., 2009; Bekkers & Wiepking, 2011; Penner et al., 2005).

Aside from these utilitarian ‘donations’, there are many ‘donation’ alert messages that are primarily phatic or interpersonal in content. Section 7.4 will discuss these messages and argue that they are used by the viewers to joke with the streamer, negotiate music taste, display appreciation and share personal narratives. With regard to their content, phatic ‘donation alert’ messages are the most conversational discursive practice of live streaming. They lead to dialogical exchanges that focus interpersonal meanings, which are traditionally associated with the formation of friendships (Eggins & Slade, 1997; Boyd, 2010). The payment temporarily enables a privileged and personalized communicative access to the streamer (cf. Goldhaber, 1997).

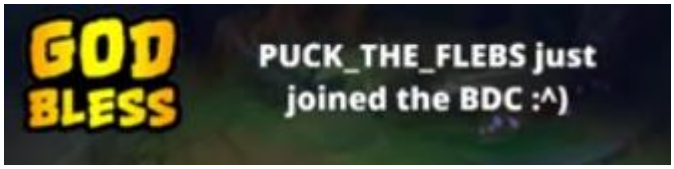
7.1 Alert messages as paid moves in exchanges

Alert messages elicit a communicative exchange between viewer and streamer. The viewer takes the initiating role and the streamer is the responding party. This initiation can occur through a ‘subscription notification’ or a ‘donation alert’ message and both have separate mechanism for text production outside of *Twitch’s* website. The process begins with the streamers, who must install an ‘alert software’ on their computer. This software is connected to

Twitch's servers and produces an on-screen alarm if it is triggered. Figure 37 shows the usual position and size an alert message

Figure 37 Schematic representation of alert message on screen

If a viewer presses the 'Subscribe' button at the bottom of the page, *Twitch's* server will communicate this button-press to the alert software on the streamer's computer. In turn, the alert software will produce a 'subscription notification' on the screen of the streamer. In a third step, this subscription notification is then broadcasted over *Twitch*'s website to other viewers.



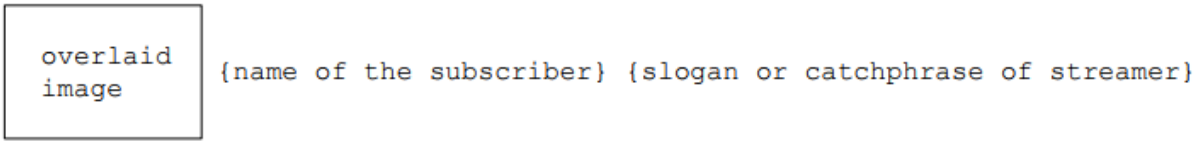


Figure 38 Schema of a subscription notification

Figure 38 shows a screen captured subscription notification and its schematic representation below. The syntax of 'subscription notifications' consists of an overlaid image, the subscriber's name and a slogan or catchphrase. Image and the catch phrase are designed by the streamer. The viewer is only pressing the 'Subscribe' button and the pre-defined subscription notifications will appear. This type of alert message affords few possibilities for individualization and customization, which is why it is used construct community and group identity (see 7.2.1).

The second type of alert message allows more customization. The viewer must go to the

external website of the alert software and visit the ‘donation landing page’ of the streamer.

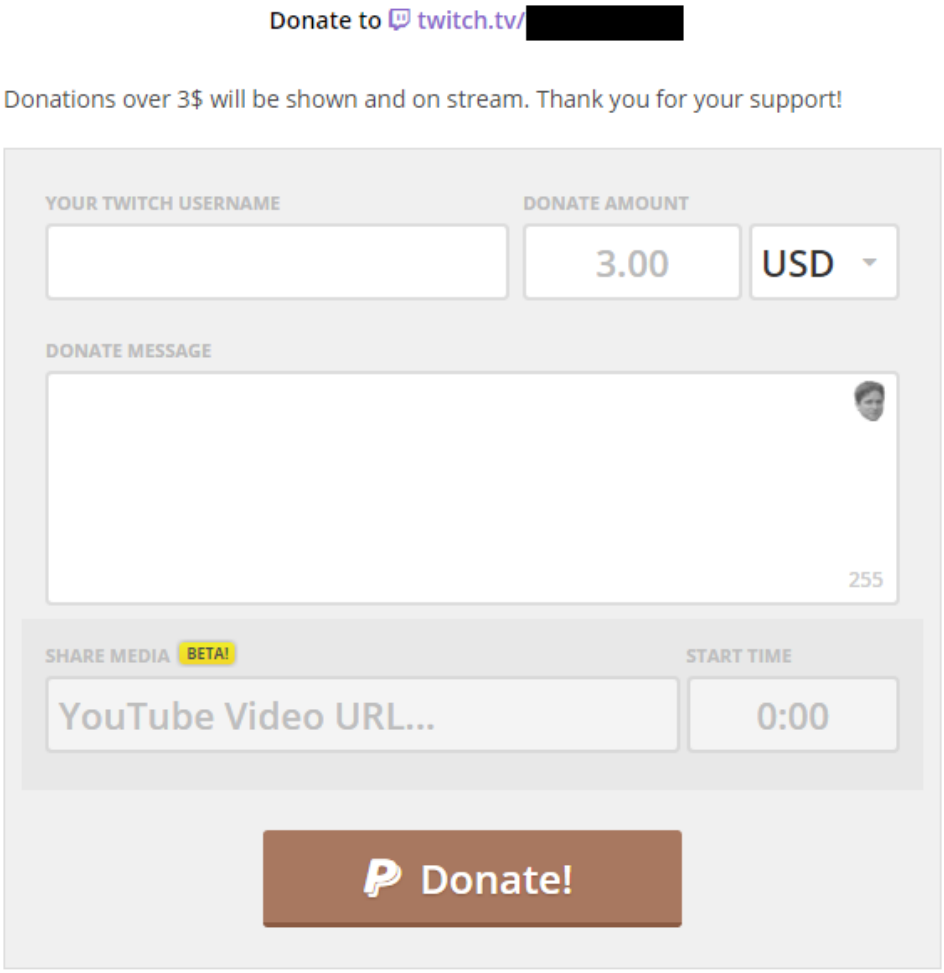


Figure 39 Donation alert landing page

Every streamer has their own ‘donation landing page’. At the top, the page indicates the recipient of the ‘donation’ and there is a field with a message written by the benefactor of the ‘donation’. It mentions the minimum amount for a ‘donation’ to trigger the on-screen message and thanks the patron for the ‘support’. In this study, the recorded payments fluctuated between $3US to $50US but payments of several hundred or even thousands of US Dollars are possible[39].

Below the first row, the patron enters the amount of money s/he wants to send and his or her *Twitch* username. The message is written in the main text field called ‘Donation Message’. This field can contain up to 255 keystrokes, which is about twice as long as the maximum length of a tweet (140 keystrokes). Alert messages provide sufficient space for longer texts. In the bottom row, the patron can add links to external websites such as *YouTube*. In paid messages, this is encouraged but posting a link in the ordinary chat will result in an immediate chat ban (see example 18, p.105).

[39] https://www.youtube.com/watch?v=A4Gux-8N1fc

Once the message is entered by the patron, s/he can press the 'PayPal Donate!' button at the bottom of the page. The term 'donate' is used five times on the landing page and it matches the common conceptualization of alert messages as donations and support for the streamer. Framing alert messages as donations highlights a presumed act of voluntary giving by the viewer. From this point, it is understandable why previous research mostly looked for the motivations in 'donation' behavior (Raes, 2015; Plath, 2015; Groß, 2017).

However, such a perspective backgrounds the content of alerts message and does not consider their interactional treatment by streamer in discourse. In this chapter, alert messages are understood as discursive moves that require an initial payment. The message only appears on the screen after the payment is processed through *PayPal.*

The format of alert messages is similar to subscription notifications.

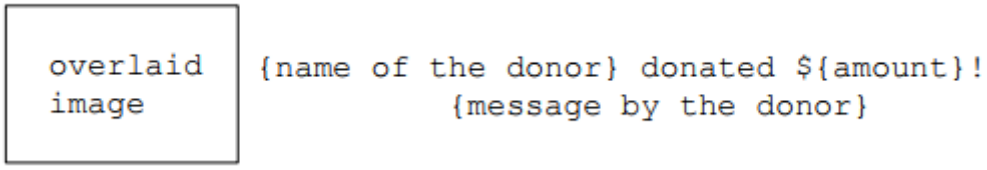


Figure 40 Schema of a donation alert message

It will usually feature an image on the left, next to the name of the 'donor', and the payed amount in the first row. The text message is usually in the second row.

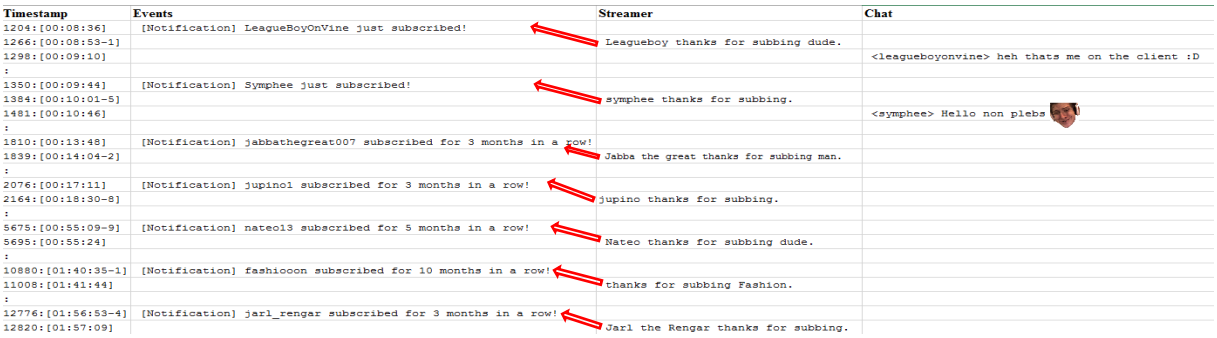
As mentioned in the methodology (4.1.3), this study passively recorded 54 alert messages and decided to collect an additional 96 to arrive at a sample size of 150. At this sample size, the data reached a saturation point and the identified alert type categories were represented in several instances (Mason, 2010). The categories are derived from a bottom-up and data-driven content analysis (Saldana, 2009), which looked at the written message of the alert as well as its interactional treatment by the streamer. The study found three major categories. There are subscriptions notifications that are used to discursively construct a community (7.2), alert messages by viewers that demand information and actions from the streamer (7.3), and alert messages that serve as interpersonal discourse between streamer and viewer (7.4). Similar to the analysis in chapter 5 and chapter 6, I do not claim that this is a comprehensive list of all possible types of alert messages. Yet, they are the ones that have been identified most commonly in the data.

7.2. Subscription notifications and community.

Subscription notifications are a very homogeneous category. On the surface, their purpose is to give notice when an audience member subscribes to the stream. However, this notification is embedded in a wider interactional practice that provides a personalized mark of distinction to

the subscriber and separates him or her from ordinary non-paying viewers. At the very least, they receive a brief word of thanks by the streamer or they may be welcomed to the 'community' of the streamer.

(1) Example – 'thanks for subbing'



Timestamp	Events	Streamer	Chat
1204:[00:08:36]	[Notification] LeagueBoyOnVine just subscribed!		
1266:[00:08:53-1]		Leagueboy thanks for subbing dude.	
1298:[00:09:10]			<leagueboyonvine> heh thats me on the client :D
:			
1350:[00:09:44]	[Notification] Symphee just subscribed!		
1384:[00:10:01-5]		symphee thanks for subbing.	
1481:[00:10:46]			<symphee> Hello non plebs
:			
1810:[00:13:48]	[Notification] jabbathegreat007 subscribed for 3 months in a row!		
1839:[00:14:04-2]		Jabba the great thanks for subbing man.	
:			
2076:[00:17:11]	[Notification] jupino1 subscribed for 3 months in a row!		
2164:[00:18:30-8]		jupino thanks for subbing.	
:			
5675:[00:55:09-9]	[Notification] nateo13 subscribed for 5 months in a row!		
5695:[00:55:24]		Nateo thanks for subbing dude.	
:			
10880:[01:40:35-1]	[Notification] fashiooon subscribed for 10 months in a row!		
11008:[01:41:44]		thanks for subbing Fashion.	
:			
12776:[01:56:53-4]	[Notification] jarl_rengar subscribed for 3 months in a row!		
12820:[01:57:09]		Jarl the Rengar thanks for subbing.	

The first example shows seven subscription notifications that are spread throughout the recorded broadcast of the large-sized *League of Legends* live stream. They have no relationship to the game events or prior discourse. Streamers very consistently react to subscription notifications with standardized spoken responses. This means they treat the notification as an initiation in an exchange.

Subscription notification messages mention the duration of the ongoing subscription. In the first two cases (l.1204; l.1350), the individuals are first-time subscribers, whereas the other viewers are subscribing for several months in a row. Streamers responds swiftly to each of the notifications with their catch phrases or slogan. In this example, the spoken phrase is "`{subscriber name} thanks for subbing`" or "`thanks for subbing {subscriber name}`". Other streamers have more elaborate responses, but this example is ideal to show that there is a standardized format with very little variation. The spoken response is an internalized linguistic practice, which is thanking the subscriber. As a move, 'thanking' is interpersonally oriented and shows gratefulness of the "thanker" towards the "thankee". The thanking moves single out a specific viewer and make his or her action the topic of the streamer's spoken discourse. Subscribing is a positive behavior and will lead to an appreciative response by the streamer. Notably, every thanking move also contains the name of the initiator as an explicit acknowledgement of the person. Paid cross-modal exchanges mention the subscribers name, whereas ordinary cross-modal exchanges only topicalize the content of the message (see 6.1).

This difference in status is also visible in the reactions of the viewers to their respective notification and thank you note

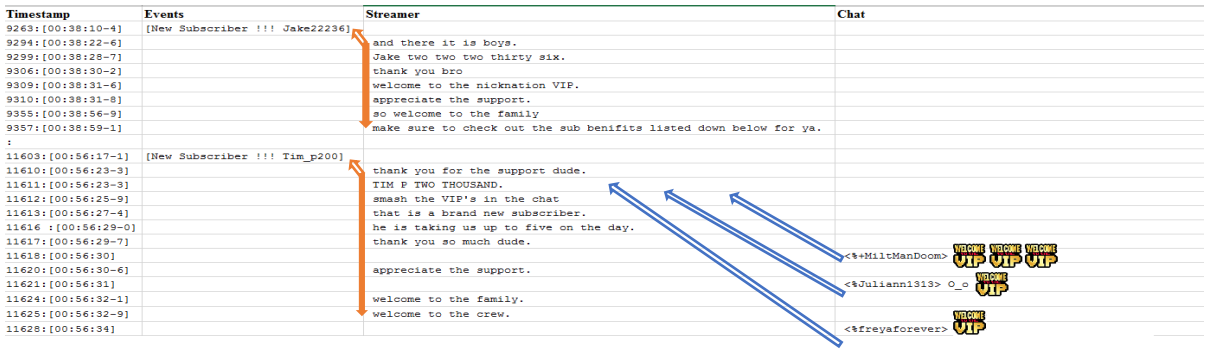
```
1204 [Notification] LeagueBoyOnVine just subscribed!
1298:[00:09:10] <leagueboyonvine> heh thats me on the client :D
:
1481:[00:10:46] <symphee> Hello non plebs
```

In the first case (l.1298), the viewer & donator performs a noticing (Schegloff, 1988) that highlights his recent subscription. For a moment, his position is raised because he is at the center of the broadcast and not at the sidelines in the chat.

Similarly, in the second case (l.1481), the audience member is writing about his upgrade in status. On Twitch, the term 'pleb' (Latin for common folk / ordinary citizen) is a playful metaphor for non-paying users. 'Plebs' are lesser members than the paying subscribers. On the surface, the user is greeting his or her fellow paying subscribers.

However, in naming the subscribers 'non-plebs' and using the subscriber-specific emote 'bjergtroll', s/he is ridiculing ordinary viewers and produces a minor taunt. As a paying viewer, s/he is not a pleb anymore and can elevate him or herself above others. In many channels, this distinction between subscribers and 'plebs' goes even further, and streamer's actively try to produce a sense of community through their discourse.

(2) Example – 'welcome to the family'



The second example features two subscription notifications and reactions by streamer and audience. The subscription notification serves as the initiation and the streamer gives a spoken response. First, he thanks the subscriber and signals his appreciation. This includes mentioning the name of the subscribers (l. 9299; l. 11611) and thanking moves that frame the subscription as an act of support (l. 9306 & l.9310; l.11620). Thereby the streamer establishes a personal connection between himself and the viewer.

In a second step, the responses are also establishing a connection between the audience members as part of an overarching community. The streamer begins with an attending move '`there it is` (l.9294)', which gathers the collective attention of the viewers and the new subscriber is welcomed in the '`nicknation VIP` (l.9309)'. It is a play on words and compound of the streamer's name 'nick_28t' and the concepts of nation & VIP. Nation and VIP are non-literal metaphors for the privileged or special group of paying subscribers.

The subscription notification is the moment of transition from 'pleb' to member of the 'nicknation VIP'. Often, this transition is linguistically celebrated by the streamer as well as

other audience members. In line 1162, the streamer requests from his subscribers to '`smash the VIP's in the chat`'. 'The VIP' refers to a channel specific emoji that is only usable by other subscribers. The streamer elicits a welcoming message from the old subscribers to the newest member. Therefore, the subscriber is not only acknowledged by the streamer but also by his or her new peers of 'VIP' viewers. The subscription notification and the responses by streamer and subscribers are linguistic performances of a ritualistic rite of passage. The subscriber enters the streamer's community.

The unifying element of this community is their connection to the streamer, i.e. they are part of the 'nicknation' and are all subscribers of his channel. This is an appeal and construction of a streamer-centric group identity, which is common to most broadcasters. Other recorded examples are the '*Voyscouts*' of the streamer 'Voyboy', which is a pun on his name and the youth group boy scouts, or the '*Weeb gang*' by the streamer Nightblue3. The streamer serves as the figurehead and the viewers are grouped around him or her.

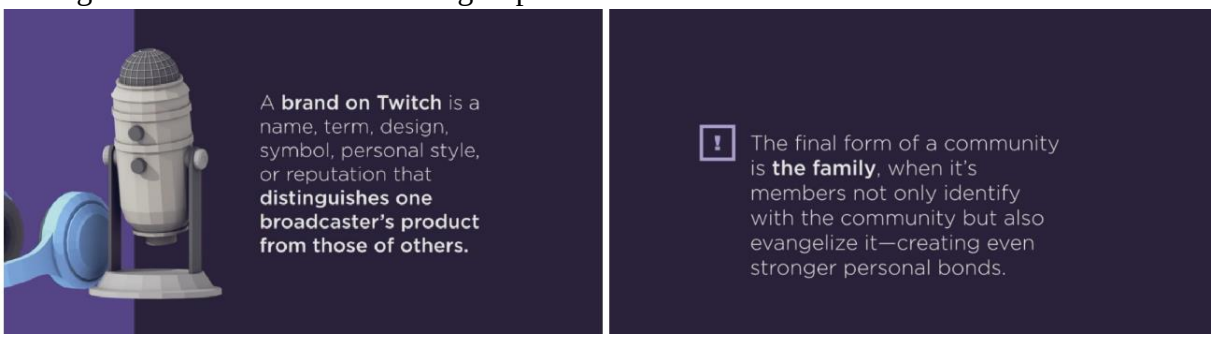


Figure 41 'Your brand on Twitch' - slides from TwitchCon 2015

Among streamers, this is often discussed as building a brand identity around oneself. Figure 41 shows two slides from the official 'Your Brand on *Twitch*' panel at *TwitchCon* 2015. On the left slide, it mentions the features of the brand such as 'names, terms, design, symbol, personal style or reputation'. In the example above, the streamer coined the 'nicknation VIP' around his personal image and created the 'VIP emoji'. It is an extension of his persona that is available for the other viewers to use in the chat. Subscribers become part of this streamer-centric community through the payment of a monthly fee. Such branded communities have an assumed natural progression towards 'a family', where viewers welcome new members into their midst. The streamer maintains his or her special position, but it is hoped that viewers develop strong interpersonal feelings for one another. This idea of a family-like community is widely shared among streamers on *Twitch* and they are often proclaimed in micro-level linguistic practices. In example 2, both subscribers are welcomed to 'the family' as a response to their subscription (l. 9355; 11624). The streamer discursively constructs a network of social relationship between viewers. It is a naturalized and dominant discourse (cf. Fairclough, 1989), and it frames

subscribers as persistent social groupings, which begins with the subscription notification and the entry to family.

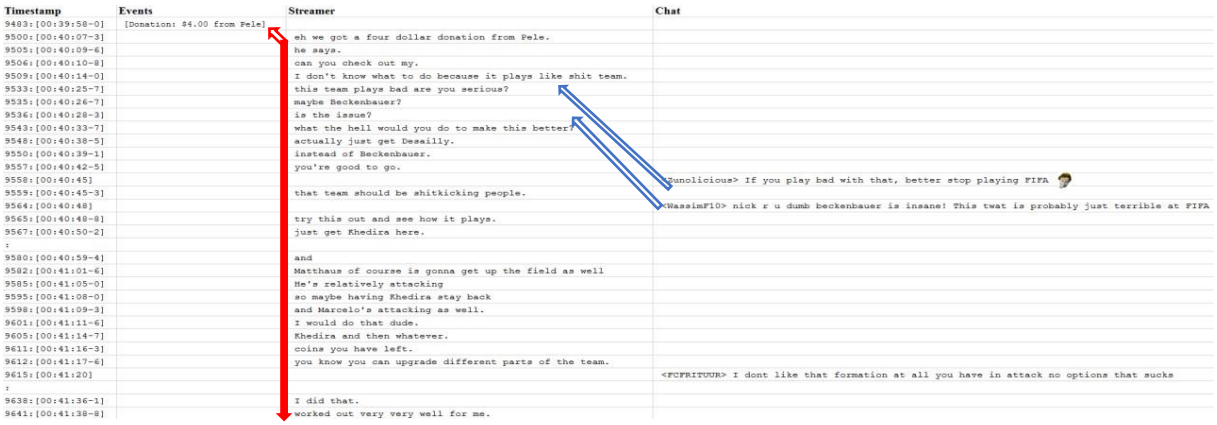
It is difficult to assess to what extent the participants truly understand or believe these relationships to be close-knit, but it is possible to study and critique their linguistic behavior. Aside from streamer-driven subscription rituals, there is very little communication on *Twitch* that indicates a sense of community or personal bonds. Only a minority of discursive moves is intended as communication among audience members and its content is mostly related to the game or background information about the streamer (see chapter 5 & 6). Personal questions are rarely discussed between chat participants and isolated instances of social talk or chat hardly constitute a community. This study did not identify communication that is designed to construct a persistent group identity. For the most part, the discourse is informal and friendly, but it indicates lose or weak interpersonal ties (Granovetter, 1983). Chat members rarely talk about themselves or one another and their only shared common ground is the streamer or the video game. As the streamer ends his or her daily broadcast, viewers quickly leave the channel (see 6.3). There is little interest in each other and explicit references to a shared group identity are almost exclusive to subscription notifications. Nevertheless, streamers and *Twitch* are monetary invested in the practice and the appearance of a close and personal community. A larger number of subscribers leads to more income for the streamer and the website. Interestingly, the chat participants are habituated to interpret a subscription payment to the streamer and *Twitch* as "growth" of their own personal network and community (cf. Sjoblöm & Hamari, 2016). This is a paradox because the study of large-sized broadcasts demonstrates that there is a decrease in audience to audience communication (see 5.3.), more exchange chaining with the streamer (see 6.1) and an increase in ignored exchanges (see 6.1.2).

However, these downsides for the viewers are masked behind the dominant reading of positive "community growth", which is mostly beneficial to *Twitch* and the streamer. As the viewer numbers are growing, 'donation alert' messages become increasingly helpful or necessary for communicating with the streamer because they appear in the visible center of the stream, are much more noticeable, and have additional force because they come attached with a financial payment.

7.3 Alert messages as request and service

Donation alerts may ask for detailed advice or specific actions, where the imposition on the streamer is larger than in ordinary cross-modal exchanges. These 'donationy' are more comparable to a paid service requested by the viewer and provided by the streamer.

(3) example – I don't know what to do



Timestamp	Events	Streamer	Chat
9483:[00:39:58-0]	[Donation: $4.00 from Pele]		
9500:[00:40:07-3]		eh we got a four dollar donation from Pele.	
9505:[00:40:09-6]		he says.	
9506:[00:40:10-8]		can you check out my.	
9509:[00:40:14-0]		I don't know what to do because it plays like shit team.	
9533:[00:40:25-7]		this team plays bad are you serious?	
9535:[00:40:26-7]		maybe Beckenbauer?	
9536:[00:40:28-3]		is the issue?	
9543:[00:40:33-7]		what the hell would you do to make this better?	
9548:[00:40:38-5]		actually just get Desailly.	
9550:[00:40:39-1]		instead of Beckenbauer.	
9557:[00:40:42-5]		you're good to go.	
9558:[00:40:45]			<Gunolicious> If you play bad with that, better stop playing FIFA
9559:[00:40:45-3]		that team should be shitkicking people.	
9564:[00:40:48]			<WassimF10> nick r u dumb beckenbauer is insane! This twat is probably just terrible at FIFA
9565:[00:40:48-8]		try this out and see how it plays.	
9567:[00:40:50-2]		just get Khedira here.	
:			
9580:[00:40:59-4]		and	
9582:[00:41:01-6]		Matthaus of course is gonna get up the field as well	
9585:[00:41:05-0]		He's relatively attacking	
9595:[00:41:08-0]		so maybe having Khedira stay back	
9598:[00:41:09-3]		and Marcelo's attacking as well.	
9601:[00:41:11-6]		I would do that dude.	
9605:[00:41:14-7]		Khedira and then whatever.	
9611:[00:41:16-3]		coins you have left.	
9612:[00:41:17-6]		you know you can upgrade different parts of the team.	
9615:[00:41:20]			<FCFRITUUR> I dont like that formation at all you have in attack no options that sucks
:			
9638:[00:41:36-1]		I did that.	
9641:[00:41:38-8]		worked out very very well for me.	

Example 3 is a 'donation' alert that was recorded in the *FIFA 2015* broadcast and it features the patron ('Pele'), who is seeking advice for his team because his squad "`plays like shit team` (l.9509)". Upon seeing the message, the streamer immediately acknowledges the patron by name (l.500) and topicalizes the donation alert by verbally repeating the on-screen message (9505 – 9509). The patron is asking for individual advise that is helpful for his personal situation. The encounter lasts 1 minute and 30 seconds, which is the longest cross-modal response in the data set. During this time, the streamer does not interact with another participant.

The paid exchange is treated very different from the quick and brief exchange chaining. In a first step, the streamer must inspect the team of the patron on an external website, which takes him around 10 seconds to open. He then produces two checking moves "`this team plays bad are you serious?` (l.9533)", which signal his surprise about the negative assessment of the team. In the streamer's own words, the team is strong and "`should be shitkicking people` (l.9559)".

Nevertheless, he keeps his composure and provides an analysis of potential problems. Throughout the excerpt, the streamer gives several suggestions for improving the team and changing some players. For example, he states that "`maybe Beckenbauer?` (l.9535)" could be an issue and suggests "`Desailly` (l.9548)" or later "`Khedira` (l.9567)" as suitable alternatives. His analysis accepts the patron's premise that the problem lies with the team and the avatars. The streamer remains very diplomatically and supportive of the patron even though he indicated his surprise earlier. The same cannot be said about the other audience members and their follow-

ups to the streamer's spoken response.

R_{spoken} `this team play bad are you serious?`

$F_{1written}$ `If you play bad with that, better stop playing FIFA` 

:

R_{spoken} `maybe Beckenbauer?`
`is the issue?`

$F_{2written}$ `nick r u dumb Beckenbauer is insane!`
`This twat is probably just terrible at FIFA.`

Both follow-ups are very confrontational and criticize the streamer and the patron for claiming that the issues are with the team. The squad is evaluated as very capable, consisting of strong avatars and is not to blame for the lack of success. For the chat participants, this warrants the comments "`better stop playing FIFA`", "`nick r u dumb`" and "`this twat is probably just terrible at FIFA`". They are direct attacks at the patron and the streamer. The patron is open to insults because he is, presumably, a bad player and bad players should quit the game. The streamer is insulted as 'dumb' for playing along with this pretense.

Viewers have no financial incentive and they are less likely to accommodate the patron. While their remarks are unnecessarily offensive, they may also be more genuine. A third message by another audience member supports this argument. The team has "`no options in attack that sucks` (l. 9615)". The formation of the team is too defensive and there are signs for mismanagement by the team owner. The patron has a strong squad but is making poor decisions, which are detrimental to his success in the game. The audience sees the problems of the team with the player & manager and is very willing to confront the patron by expressing negative opinions about him and his actions.

The streamer must be much more diplomatic and has to be careful with his words. He is trying to help the donor without criticizing him. Streamers must consider the current alert message, but also future messages. Potentially offending one of the patrons can alienate parts of their viewership that otherwise might have asked them for advice via a payed message. From this perspective, the alert message is comparable to a paid service and the streamer tries to make his customer happy, which will help him to attract new patrons.

In the data set, alert messages have been used for a variety of requests and they included playing a certain fighter in *League of Legends* or playing a few more rounds of *World of Warcraft* arena. In comparison to the ordinary exchanges, such alert messages have an increased imposition and there is a high pressure on the streamer to comply with the request because of the financial payment.

Alert messages as a service may be previously unknown to a reader of this work, but they are intuitively understandable. An initiating party is paying a responding party to perform

a certain action that s/he may desire. The mechanism of the exchange is new, but its purpose is very familiar. The third main category of alert messages is very different. Often, alert messages are used to initiate interpersonal exchanges with the streamer and such paid messages are not covered in previous research.

7.4 Alert messages as interpersonal exchange.

In these alert messages the viewer is seeking personal contact with the broadcaster. Broadcasters routinely respond to such messages with appreciation and provide positive social feedback. The social contact is more explicit and individualized than in ordinary cross-modal conversation on *Twitch*. The study has classified the paid interpersonal exchanges into four sub-categories. Alert messages are used to 1) send humorous remarks or jokes, 2) negotiate and discuss music taste, 3) encourage and admire the streamer and 4) for self-disclosure and brief personal narratives. Importantly, interpersonal communication about such topics has consistently been identified as part of friendship formation and friendship maintenance (Norrick, 1993, 2000; Baym & Ledbetter, 2009; Baym 2012; Selfhout, 2009; Hall, 2011; Policarpo, 2014).

Therefore, the four types of interpersonal exchanges are in many respects similar to befriending, but it is doubtful that they will lead to genuine friendship or any persistent social ties between viewer and streamer. Alert messages are very illusive and soon after the paid exchange is concluded, the viewer is again an ordinary spectator and the broadcast quickly continues.

Humorous messages

Humor is an interpersonal resource that can be used to signal a speaker's stance towards other people or groups. Joking with someone is a form of alignment and laughing at someone may indicate underlying resentments (Zillman & Stocking, 1976; Du Bois, 2007; Knight, 2013). In spoken conversation, humor is regularly employed to improve the mood of participants and create solidarity (Eggins & Slade, 1997; Norrick, 1993). On *Twitch*, these basic functions of humor are similar, but the exchange pattern of joking is adapted to fit the technological affordances of alert messages.

(4) Example – Chat Offender

The user produces a donation alert message that is initiating a joke in an exchange and the streamer produces a spoken response almost immediately

$I_{written}$:

R_{spoken}: 01 i1014i thanks for the four twenty.
02 I am an official chat offender
03 I need to warn every streamer
04 that I am in their chat room
05 and Pam stole three dollars from your wallet
06 so who is the real threat here?
07 {smiles & laughter}
08 wait Pam stole money from me?
09 holy shit.
10 "alright".
11 "we're done it's over".
12 "I'm not gonna verify this claim".
13 "I'm gonna trust you dude".
14 thanks for dropping the knowledge on me dude
15 I appreciate that.

The first utterance of the response is a thanking move, which mentions the username of the patron and acknowledges the payment of $4.20USD. This acknowledgement is always among the first moves and usually positioned before the topicalizer. The signal of appreciation for the payment precedes any reaction to the content of the message and is a central aspect of the response. Only after the personal acknowledgement, the streamer utters the topicalizer (l. 2 - 7) and is reading out the joke.

The initiator describes him or herself as a "chat offender", who must inform the streamer of his or her presence. The message is a humorous and creative adaptation of the American punishment for sex offenders, who must inform their neighbors about their criminal history. The second part of the joke accuses the streamer's girlfriend 'Pam' to have stolen $3 USD out of the streamer's wallet, when he was in the bathroom earlier. The two parts of the joke are barely connected and are only tied together by the word 'threat', which applies to "chat offenders" as well as thieves.

The first part of the joke is positively received by the streamer via visible smiles and loud laughter in line 07. It validates the message as a successful joke and the patron as a good

jokester. In line 8, the streamer performs a registering move and signals the reception of thé new and surprising information. Arguably, the second part is less funny and even face threatening, but the streamer's response continues with the playful theme of the message. There is a notable and hyperbolic shift in his voicing (l.10 – 13), which indexes the non-seriousness of his utterances. He constructs a fictional scenario, where he must separate from his girlfriend over the reported theft (l.11). Within this joke, he aligns with the patron and expresses his trust (l.13) and appreciation (l.15) for the information (cf. Du Bois, 2007; Englebretson, 2007). The streamer co-constructs the humorous encounter by working with the material provided by the patron. This behavior is similar to friendly banter commonly found in ordinary spoken conversation. However, such banter often features verbal jousting or one-upmanship, where both speakers continue the joke and try to outwit one another with increasingly absurd scenarios (Norrick, 1993; 2000). In face-to-face conversation, there could be another exchange or at least a brief follow-up to the response. This is almost never the case on *Twitch* because every 'one-up' or new turn would require the viewer to send another donation alert message. Even a short transaction with only a few exchanges would quickly cost a significant sum of money. Therefore, viewers put their complete contribution in a single turn, even if it leads to disjointed messages that express two or more seemingly unconnected sentences.

Conversational joking is adapted to the cost of alert messages and a similar process also explains the format of the riddle in the next example

(5) Example – 'riddle me this'

$I_{written}$:

During the period of data collection, there were several high-profile shootings of unarmed African Americans, which received significant media attention. One of these incidents was the lethal encounter between Trevor Martin and George Zimmerman. In court, the defense attorney argued that Trevor Martin's sweater with a hood made him look suspicious and dangerous. George Zimmerman pursued Trevor Martin and it came to an altercation during which Zimmerman shot Martin. The riddle plays on this theme and the primary butt of the joke are 'black males', who are laughed at for getting shot. This basic idea is masked by a series of

conceptual metaphors that link the removal of the foreskin (circumcision) to the removal of sweater with ‘a hood’. Hood has a double meaning and this resolution may be perceived as humorous.

Outright racism is banned but humor circumvents this taboo and is hiding the aggression (Zillman & Stocking, 1976). Discriminatory and misogynistic messages against women and minorities are common on *Twitch* (see example 13 in 6.1.2, p.133; Anykey, 2016). Even if the streamer is not verbally responding, they still notice the message and receive the payment. They do not have to actively endorse the content and there may be some appreciation of the joke by the audience.

From an interactional perspective, the performance of riddles or knock-knock jokes is a dyadic exchange with three to five phases (cf. Harries, 1971; Glazier & Glazier, 1976). They are an oral tradition and that expects a back and forth between the individual that is asking the riddle and the listener. The following table separates the riddle in its expected form in spoken conversation.

Exchange Slot	**Realization**
Initiation	`Ok Sneaks, riddle me this.` `Why are all black males circumcised?`
Response	Ø
Follow-up	`Because George Zimmerman will shoot anything that’s black and wears a hood`

Table 26 Alert message riddle

A riddle should consist of three phases. The initiation begins with an attending move ‘`ok Sneaks`’ and the command ‘`riddle me this`’. It starts the interaction with the participant and frames the following exchange as a non-serious riddle. It mentally prepares for a non-literal question that is not solvable with common logic. Afterwards, the main part of the riddle is presented to the listener. Riddles often contain conceptual metaphors that play with word meanings in different domains (Glazier & Glazier, 1976; Lakoff & Johnson, 1980). In an oral riddle, the listener will have time to think and will be asked to respond with a possible solution. This phase is usually very enjoyable to the participants because it poses a mental challenge and a test of wit. The phase may also be dragged out and involve teasing until the responding party gives in and admits defeat. This important segment is skipped in donation alerts. The message already contains the punchline, gives away the solution to the riddle, and exposes the underlying metaphor. The three phases that would be typical for a spoken performance are moved together in a single donation alert message.

The ‘donor’ can never be certain that the streamer will respond to a riddle initiation. Although donation alerts receive more responses than ordinary chat, there are no guarantees. The production of an initiation and a follow-up would require a second payment. Instead of following the traditional oral pattern, the viewer re-appropriates and adapts the message to fit the requirements of the situation with its financial constraints and its uncertain participation by the streamer.

Such constraints are not specific to riddles or conversational joking because they are a basic feature of alert messages. Viewers anticipate their own future turns and pre-emptively produce them in the initiation of the exchange. It minimizes the required payment and serves as a safeguard against being ignored. The riddle is shown in its entirety for $4 USD instead of $8 USD and is not reliant on the streamer’s guesswork and assistance

Negotiation of Music Taste

Music preferences are a strong determiner of friendship formation, especially among adolescents and young adults (Selfhout et al., 2009). They serve as a shared starting point that can branch out to other areas and lead to lasting friendships (Baym & Ledbetter, 2009).

In many donation alert messages, the viewers present their music taste and will link a URL of a *YouTube* video as part of the alert message. They hope for a reaction by the streamer and in some cases, streamer’s play the audio track in the background.

(6) Example – for all the old school 90’s kids

$I_{written}$:

In example 6, the donor is addressing the ‘`the old school 90’s kids`’ via an attending move. The phrase positively marks a certain segment of the population that was a teenager during the 90s. The linked song is the intro music to an American TV show that ran on the TV channel *TeenNick*, a subsidiary of *Nickelodeon* for teenagers.

Presumably, the donor is him or herself a member of this generation and is sharing a song s/he listened to when s/he was growing up in these formative years. S/he constructs a group identity of ‘`old school 90’s kids`’ that are assumed to have a shared cultural repertoire

that is familiar with this particular TV show and its theme song. Therefore, the donation message is not just about the music but also about the patron's identity and other people like him or her, who '`know this song`' and can relate to it on a personal level. Donation alerts are used to share songs that are of special importance to the patron and carry deeply emotional meanings and nostalgic memories. Thereby, the viewers share a part of themselves on screen and invite the streamer and other viewers to engage with them and their past.

(7) Example - 'When we were young'

I$_{written}$: 

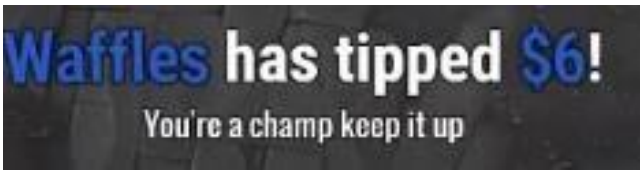
In example seven, the donation alert is directly addressing the streamer and begins with words of encouragement that are very typical for donation alerts. For now, the analysis will focus on the second part of the message that reads "`Also I would love to hear your opinion of this song me and some friends did when we were young (15 years ago)`". It is a careful request to play and evaluate the song. The song is very personal to the patron because it was produced by his band in his teenage years. Artists often have very intimate relationships to their songs because they express thoughts and emotions that they had at the time of writing and composing. The donor is seeking acknowledgement and validation for the musical creation of his younger self. It is very likely that the song is still relevant for him, since he is willing to pay four dollars for the opportunity to share the video with the streamer.

Neither the sixth nor the seventh example received a spoken response and it is impossible possible to know the reason why. It can be speculated that sharing *YouTube* videos and external URLs is prone to an interactional failure because the messages need the active participation by the streamer. Streamers must go out of their way to open the link in the web browser. This is problematic if the streamer is busy with the game, or if they are not in the mood to listen to a certain song. Donation alert messages come with a sense of obligation and its paid initiations are more likely to receive a response, but the desired reaction is not mandatory.

<u>Encouragements and Displays of Admiration</u>

Messages of encouragement and displays of admiration have various degrees of intensity. They range from friendly and supportive to the intrusive, which crosses personal boundaries between viewer and streamer.

(8) Example – Keep it up

$I_{written}$: 

Example eight is typical for a supportive donation alert message. The user pays $6 USD to send a brief message that displays his or her admiration and encourages the streamer in his broadcast. The first move "`you're a champ`" refers to the streamer and is a display of admiration. Other examples from the data read '`the best`', '`number one`', '`the best streamer`', '`favorite streamer`', '`hero`' or '`the god`'. Viewers talk to the streamer in very positive terms and aggrandize them with exaggerated and superlative descriptions. These descriptions are always subjective and different patron's will refer to different streamers as 'the best'.

The second move '`keep it up`' is encouraging the streamer to continue the broadcast and it also serves as the closing move of the message. As the viewers produces the alert message, there is already an indication that they will not produce a concluding follow-up. Viewer's initiate and end their part of the exchange within a single turn.

(9) Example – 'my favorite streamer'

Iwritten: 

In example nine, the alert message begins with an attending move '`ey voy`' that attempts to capture the attention of the streamer. 'Voy' is an abbreviation of the streamer's username Voyboy. It is a diminutive that makes the address even more informal and this informal theme continues throughout the message. The patron makes him or herself relevant to the livestream by establishing him or herself as a long-time viewer, who has '`been watching ~~you~~ him for a while`'. It gives credibility to the display of admiration that calls the streamer his or her '`favorite streamer by far`'. In the third line, the patron expresses his intention of '`spreading some cheer`' because of the birthday.

On the literal level, the paid message is casual and a small celebration. The study cannot evaluate if this is truly the case, but it can highlight the context of the exchange. On his or her birthday, the patron is sitting in front of a computer to watch a broadcast on *Twitch* that

cyclically repeats every day. On this occasion, the person is paying $5.96 US to tell the streamer that it is his or her birthday, which may or may not lead to a spoken response. From this perspective, the cheerful celebration appears to be a payment for social acknowledgement. In such a scenario, the casual wording would be more comparable to a defense or coping mechanism that masks the call for attention.

This interpretation also fits with the observations from other examples, where viewers very explicitly try to reach out to the streamer and attempt to create a personal connection. In many donation messages, the viewers report on personal hardships, their struggles and how the social interaction during the live stream is an outlet to their problems (see example 11). However, there is also a limit to bringing your personal life in the broadcast and some audience members become overly invested and cross personal boundaries.

(10) Example – "accept me as your friend"

Example 10 features two adjacent donation alert messages and the spoken response by the streamer that occurred in between them. The text indicates that there was one preceding alert message by the patron, but it was not possible to recover the alert from *YouTube*.

$I_{written}$:

R_{spoken}: "eh".

"No"?

$F_{written}$:

In the first donation alert, the viewer lays out his personal feelings towards the streamer. First, the patron asks the streamer not to hate him. He speaks under the assumption that the streamer harbors strong negative emotions towards him. Secondly, the patron expresses that he genuinly considers the streamer as his or her friend although this feeling is not reciprocated. Therefore,

the viewer asks the streamer to a change his behavior and befriend him. Lastly, the viewer writes '`love you`' as a closing move. It is a farewell message, which reinforces the donor's statements and finishes his or her part of the exchange.

The streamer's response is a brief and almost confused rejection. Friendship or love cannot be demanded from one party to another and has to develop naturally over time and in both directions (Hall, 2011). This is difficult in live streaming because there is an asymmetry in the interaction. Streamers must interact with many viewers at a time and each viewer is almost anonymous to them. They know them only by their screen name and their chat interactions. However, viewers watch the streamers and listen to them for extended periods of time. Therefore, they may develop strong feelings of affection and assume or hope that their feelings are shared by the streamer. At TwitchCon, streamers reported that in some cases viewers could become "too needy" or "too clingy" and overinterpret the interpersonal relationship (TwitchCon field notes, 2015). The example shows many signs of an overattachment towards the streamer. There is a mixture of strong and conflicting emotions, which are very likely inaccurate. The most likely scenario is that the streamer does not know the viewer and has neither a negative nor a positive opinion about him or her.

Due to the dispreferred response '`eh no`', the patron pays for another message, where he re-states that he does not want to buy the streamer's friendship that his feelings are genuine "`I do care` [about] `you`" and "`you have friends`". The request to work on the friendship is softened with "`You do not need to add me`". The viewer is talking about the streamer 'adding' the donor on the so called 'friends list' in the video game *League of Legends*. Being on the streamer's friends list, would allow the donor to send private messages within the video game and he would not need to send donation alerts on *Twitch*. It is a more direct level of communicative access to the streamer, which is rarely given to viewers.

The exchange ends with a promise by the viewer to contact the streamer again and the closing move '`love you`'. The patron still desires the streamer to be his friend and considers it a possibility in the future.

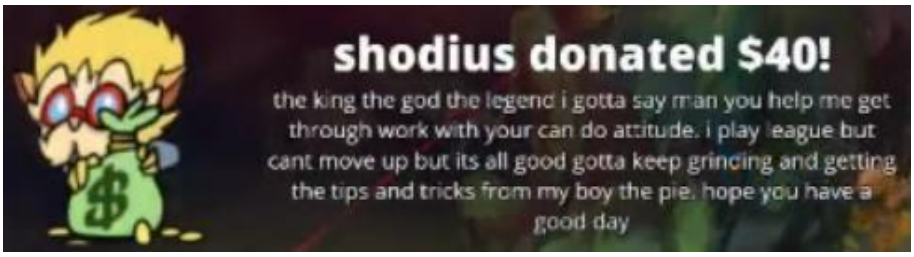
Self-disclosure and brief personal narrative

In ordinary chat, it is very difficult for viewers to talk about themselves because they cannot hold the conversational floor for a prolonged period. Investing much effort in a message is futile, if the chat is moving quickly and it is uncertain if the streamer will read or respond to the message. Donation alerts are a remedy that comes with a financial incentive for the streamer to react to the messages. This makes them more suitable for longer and personalized

messages by the audience.

(11) Example – “you help me get through work”

$I_{written}$:



The alert message begins with an attending move “`the king the god the legend`”, which is simultaneously a display of admiration by the viewer. The donor is sharing his or her negative experiences and frames the broadcast as a relief and help to “`get through work`” due to the positive attitude of the streamer. There is a theme of perseverance despite hardships in the professional life and video game.

In the video game, the donor is not as successful as s/he wants to be and “`can't move up`”. He refers to the game’s internal ranking system and ‘moving up’ means progressing towards higher and more prestigeous ranks. In the viewer’s story of perseverance against the odds, the streamer takes the double role of mentor that helps with “`tips and tricks`” and is a friend-like figure, who is referred to as “`my boy the pie`”. ‘The pie’ is diminutive form of the streamer’s nickname Imaqtpie. It constructs familiarity and social closeness between streamer and viewer. In the viewer-constructed narrative, there are two protagonists, namely the donor and the streamer, who have a shared connection through the live stream. The message creates a social relationship between them that one-directional from viewer to streamer. The story is told by the viewer and expresses how s/he relates to the streamer. The streamer is put in the story and has the roles that are given to him. The streamer is ‘the boy’ or friend of the viewer, because the viewer says he is. The production of the story required the payment of money, and in this case the viewer payed $40 USD. The high amount may indicate the gratitude that the viewer feels towards the streamer.

If there was no such payment, the message would not appear on screen. In this sense, the users pay for alert messages to produce more salient texts, where they can inform the streamer what the broadcast means to them and how they personally relate to the streamer. If the streamer responds to such messages and acknowledges the viewers, they may co-tell a coherent narrative via donation alert messages.

(12) Example – the history of you and me

The title of the 12th example is an allusion to the co-telling between two viewers and the streamer. The narrative begins with the donation alert by an audience member and the first spoken response by the streamer.

$I_{written}$: 

R_{spoken}:

01 Psycho donated 20 bucks thanks dude
02 hey I am new to the stream.
03 I just subbed and I wonder why you don't play competitive.
04 eh thanks dude.
05 uff long story {laughter}.
06 I had a really long career
07 I played competitively for about four five years
08 I was pretty much part of the NA scene when it first started
09 a:nd part of LCS when it first started
10 I went to season two world championships
11 I almost qualified for season one world championships
12 and pretty much have been playing for the longest time man
13 ehm
14 got to the point where I kinda needed to
15 take a break from competitive play
16 and just do my thing
17 and about a year and a half ago I left the team that I was on
18 at that time team curse
19 now it is team liquid
20 and I started full time streaming
21 I started doing my own thing
22 stopped living in gaming houses
23 yeah I've been really happy man
24 ever since
25 not that I've been unhappy playing competitively
26 but I've done it for so long
27 I had to take a step away from the grind man

In the story, the viewer positions him or herself as new player of *League of Legends*, who 'always use to hear about' the streamer. Thereby, the viewer ties his personal experience with the game to the streamer's career. The streamer is special to the viewer and the viewer wants to hear the streamer's story from him directly. By mentioning his subscriber status, he signals that he is already paying $5 US per month on top of $20 US in the donation alert and the monetary reinforcement creates a sense of obligation to answer

The spoken response starts with a thanking move that mentions the name of the donor

and the payed amount. Afterwards, the streamer utters the topicalizer (l. 2 – 3) before telling the "`long story`" of his professional career. The main themes of this story are similar to example 21 (p.151). It is a very sober and unenthusiastic account of the life as a professional gamer (l. 5 – 27). The story begins four years in the past and mentions individual events such as the start of the "`NA scene` (l.8)", the "`season one world championship` (l. 11)" or '`team liquid` (l. 19)'. It covers several years in very few sentences and focuses on the most important steps of his past. There is only one element in the story that is repeated. The competitive career is explained as a tiresome "`grind` (l.27)" that he had to "`take a break from` (l.15)" or "`take a step away from` (l. 27)". The repetition vividly highlights the negative evaluation of the professional gaming career as a very difficult activity. Additionally, the evaluation also serves as the coda to the story and connects the past events to the present (Labov & Waletzky, 1967). The streamer is a full-time broadcaster on *Twitch* because his e-sport career was not very desirable. He shares about himself and his negative experiences. However, and importantly, his story is not connected to the viewer or anyone else. It is his individual and personal account of the last several years.

Around 30 seconds later, the viewer sends another donation alert message and starts a second exchange that picks up at the later part of the story, when the streamer has joined '`team liquid` (l. 19)'

$I_{written}$:

R_{spoken}:

```
01 Psycoos thanks for the four twenty and he said
02 I chose team liquid to follow for lcs
03 and my friends informed me about their past
04 and how you where their prime
05 hey thanks man
06 yeah I had some great times on team curse
07 wasn't the highlight of my career
08 but we were really successful
09 we just never made it to worlds or anything
10 we were like good in random periods
11 where it did not matter
12 the first season of LCS
13 like the first split of LCS
14 we were the best team that existed until like playoff
```

In the initiation, the donor positions himself as supporter of the streamer's new team. This support is framed as an active choice due to information he received from his or her friends

about the streamer's past and his important role for the team "`you where their prime`". The reasoning is streamer-centric, and the donor connects his present decisions to the streamer's past actions. Streamer and donor are the main protagonists of the story and the streamer is aggrandized for his gamer career.

The streamer's response lacks a clear trajectory. The initiation by the viewer does not ask a question and it provides little specific information that the streamer could talk about. It is very difficult for him to connect to an anonymous viewer. Therefore, he can only thank the patron twice (l. 1; l. 5) and reiterate some aspects of his career in a rather disjointed story, which does not involve the viewers.Around 1 minute later, there is an additional message by a different viewer.

$I_{written}$: 

The donation alert also has the streamer's gaming career as a topic. It characterizes the previous two exchanges as a "`brief history`" and uses them as a point of transition to his or her own story. The message uses jargon and mentions details such as "`the own3d days with IRC`" to give a temporal context and to claim authenticity. In the new story, lozzamaniac was there with the streamer from the beginning ("`only get to watch your stream`"), and is still there to celebrate the current successes ("`how far you've come man`"). In this version of the story, 'lozzamaniac' is the co-protagonist with the streamer.

On the surface, the three donation alerts and the two spoken responses have the hallmarks of a well-told personal story. They discursively create a shared narrative thread, which begins in the past and "connects moments and their men (Goffman, 1967, p.3)." Beneath this surface, the narrations are streamer-centric and the social ties between participants appear very weak. In each version, the donor and streamer are the sole protagonist that experience the successes and failures of the streamer together. The streamer's responses barely connect to the viewers and it also needs repeating that this co-constructed story was only enabled by combined payment of $34.20USD. The donors have to pay for taking their turn and giving their account of the streamer's history. Yet, this story only remains relevant for a short moment and the discourse of the live stream quickly moves on.

7.5 Chapter summary

The chapter set out to answer the third aspect of the research question and inquire about the form and social function of alert messages and their role in the organization of discourse. This question was motivated from the observation that alert messages work differently from ordinary cross-modal exchanges. Therefore, the chapter began with an explanation of their text production and showed the communicative features of alert messages such as their salience on screen, their financial requirement, and their default format. Their set of affordances lend themselves for messages that the participants consider very important and that are not very time sensitive. This explains why alert messages are not used for cheering and are generally unconnected to the unfolding gameplay. Moreover, they do not contain simple game questions that can be answered easily with a web search. If they are used to ask for information, they elicit individual advise or ask special requests to the streamer (see 7.3). In these cases, the requests have a higher imposition than the ordinary conversation and may even interrupt or delay other types of interaction. The discussion also showed that streamers will be careful with their reactions to these requests and avoid face threatening answers. It has been argued that this is, to some extent, motivated by the financial incentive to react quickly and politely to alert messages.

This pattern of quick and polite reactions has been demonstrated very clearly for subscription notifications. Streamers develop a routine response to the announcement of a new subscriber. It becomes a standardized ritual, where the viewer transitions from a 'pleb' to community member. The transition is stylized by the streamer through his or her catch phrases and emoji. The catch phrases and emoji are designed in accordance with the streamer's persona. The paying community members are understood as an extension of the streamer's identity. *Twitch* and streamers argue that this fosters interpersonal bonds between streamer and viewers but also among community members. However, outside of subscriber notifications, the study did not identify many instances of community building discourse. Most communication is about immediate game events or discussion about the game (see chapter 6). There is also a fast turn-over between communicative participants, which leads to a general anonymity between users. At best, the relationship between viewers could be described as a weak tie (Granovetter, 1983), which is the result of the shared interest in gaming and the streamer. Therefore, it is unsurprising that in large-sized channels, the discourse funnels towards the streamer and s/he is the recipient of most of the attention. It is more important for viewers to interact with the streamer than it is to interact with other viewers. In this context, donation alert messages become particularly useful to the most committed spectators, who feel invested in the broadcast (cf Sjoblöm & Hamari, 2016). The alert messages allow them to circumvent the competition in the chat and

initiate an exchange with the streamer on the main screen. Based on the content analysis, the study argues that the interpersonal exchanges come in four major categories. There are humorous exchanges, displays of admiration, negotiations of music taste and segments of narrative self-disclosure. Although the categories are different in the details, there is overlap in their form and social purpose. They are streamer-centric practices that spectators use to relate to the broadcaster on a personal level. Viewers attempt to humor the streamer, show him or her their favorite music or express how important the streamer is to them and their life. This is mostly appreciated by the streamer but may also escalate, when the viewers become overly attached and emotional.

There is an asymmetry in familiarity, which was very visible in examples 10, 11 and 12. Viewers are very knowledgeable about their favorite streamers and can create extended narratives, where they co-occur with the streamer in a shared timeline. These narratives are one-sided because streamers do not perceive the viewers as members of their past. They produce interpersonally-oriented responses that react and relate to the contribution by the viewer as good as possible. Exception are extreme cases, when viewer completely misjudge the status of their social relationship to the streamer.

The form of donation alert messages is adapted to the payment mechanism. Messages are longer than ordinary chat messages and they are pre-planned to express all the intended points in a single turn. Closing moves or other follow-ups are written directly in the first initiation, which avoids unnecessary turn-taking and makes the exchange more affordable and less risky. The payments attached to alert messages give the initiations considerable pragmatic force and create an incentive for the streamer to react. This does not mean that streamers respond to all alert messages but there is a sense of obligation. Outside of this study, I observerd many instances of viewers complaining that their message to the streamer has not been read or answered. There is an expectation that alert messages receive a response in a timely manner. Streamers even apologize to the viewer and admit that they were at fault. Such an expectation does not exist for ordinary chat messages, where it is expected that most initiations by the audience will be ignored. The streamer and viewers treat alert messages as a privileged form of communication that should receive preferential treatment in the interaction on *Twitch.*

Chapter 8- Conclusion and future research

8.0 Initial remarks

The aim of the dissertation was the description of the organization of discourse of live streaming, which was motivated by two major factors. First, the study wanted to improve our knowledge of the cultural phenomenon of live streaming of video games. Language and communication are one of many windows into this sub-culture and research on interaction can reveal what is valued by participants in an activity. The study began by historically situating live streaming of gaming and presented previous research from sociology and ethnography. Their descriptions about live streaming on *Twitch* (see section 2.1) are very valuable, but they are distributed across many papers and are often not connected to one another. Section 2.2 ordered the many pieces of information and presented the first coherent overview of the status quo on research on *Twitch*.

The overview compared earlier contributions, pointed out shared themes, and issues of contention. There was agreement in the description of live streaming as form of conversation, commentary or narration. Disagreement or unclear description existed in the details of these descriptions. For example, 1) how does conversation and commentary alternate, 2) what is the role of game events, and 3) what makes the discourse of live streaming conversational. Karhulahti (2016) proposes a model of frame shifts that is tied to the status of the round. Within the round, there is narration about the game. Outside of the round, the 'interview frame (conversation)' takes place and streamers chat with their viewers. Smith et al. (2013) give the 'player as commentator' much more communicative freedom in the interaction and argues that streamers are in control. Their two academic positions are irreconcilable, and yet it has been previously impossible to judge or compare them. Both of their arguments were based on ethnographic studies and their observations, which are inaccessible to critical inquery.

Another issue in the previous literature has been the binary distinction between medium-sized and large-sized channels. The former have been characterized as orderly conversations whereas the later are considered as a chaotic cacophonies (Hamilton et al, 2014; Nematzadeh et al., 2016). My review critizised that the literature on this topic has not considered the dynamic character of discourse or the role of the streamer in the communication (cf. Ford et al., 2017).

All these research gaps or conflicting information exist because prior research has not studied the interaction of streamer, game and viewers with natural occuring data. Without a proper theoretical foundation and an analytical access point, descriptions of the discourse of live streaming must remain superficial and limited to common sense explanations. It is at this point, where the study intervenes and describes the discourse of live streaming via the framework of the Birmingham school. The framework is uniquely suitable for this endeavor

because it is designed to study the organization of discourse in reference to its non-linguistic context. Moves and the exchange are its units of analysis and the building blocks of dialogical conversation and more monological narration. So far, the Birmingham school has mostly been applied to spoken dialogue. Cross-modal conversation has only been discussed within adjacent research paradigms such as Conversation Analysis (Rosenbaun et al, 2016a) and it produced micro-level descriptions of online video chat. However, online video chat is different from live streaming because it lacks a simultaneous performed activity.

Therefore, the second motivation has been the value of the study for linguistics. There are no descriptions of live streaming with its emergent discourse, its simultaneously unfolding game events, and unique configuration of participants. For linguistics, the study generated new insights about micro-level practices in a novel and unique setting of computer-mediated communication.

Section 8.1 will outline how the findings of the study in its three original research chapters lead to a new descriptive model that accounts for the organization of live streaming. The model is a re-appropriation of the Birmingham school's description of discourse via levels and ranks (see figure 17, p.49) and it will be explained from the bottom-up. Each section will explicitly refer back to the findings in the chapters to discuss how they have influenced the development of the data-driven and theory-guided model.

In 8.2, the study will reflect on its own limitations and how they could be improved upon. The section will critically discuss issues with the transcription, data selection and my interpretation of the social discourse. The transcription format was decidedly simple to lay a foundation and start basic research about live streaming. There remain many phenomena that are open to a more detailed study. To make them available to research, the transcript system needs further refinement. By giving a review of the most recent developments in multi-modal transcription, the section can suggest potential improvements, which would open up new avenues for future research.

The section will also address concerns with the data selection, which were difficult to foresee in the planning process. Unfortunately, some of the studies limitations only became appearant during or after the actual analysis of the data. These deficits are part of the learning experience of the PhD and are an important lesson.

The third potential concern arises from my interpretation of the discourse and its ability to construct a community. As indicated in chapter seven, I am skeptical that *Twitch*'s communication is fostering persistent communities. This stance is open to criticism because the collected data was a snapshot of only two hours per broadcast. One cannot exclude the

possibility that there is social discourse that would become visible in a longitudinal study. Regular members may form a discourse community, whose conventions and practices were not properly registered. Secondly, the research cannot account for the lived experience of viewers. They may perceive the discourse on *Twitch* as socially enriching, even if there is little discourse that is explicitly interpersonally-oriented.

The study concedes these two points but remains doubtful about the construction of genuine social communities. Social media become increasingly professional and this professionalization goes hand in hand with new monetization strategies. Therefore, the dissertation will end with a brief outlook on the monetization of social interaction and its potential implications in section 8.3.

8.1 Synthesis of study

The discourse of live streaming features two parties, hundreds or even thousands of participants and an unfolding video game. All of this is brought together in the communicative environment of the website. Although this makes discourse appear chaotic, the study argues that the communication of live streaming has an emergent order. Temporarily, the discourse may develop into a certain direction, but it always returns or resets to its central topics and patterns of interaction.

Non-linguistic organization	***Discourse***
Streaming Schedule	
Broadcast Session	**Interaction**
Pay and play > parley	**Transaction**
Game Events	**Exchange / Commenting & Reporting**
	Dialogical Moves / Monological Moves

Figure 42 Organization of discourse of live streaming

The figure is the descriptive model for the organization of live streaming. It is not a fixed set of rules and instead describes strong interactional tendencies that are grounded in the preferences of the participants. The model adheres to the central tenets of the Birmingham school and argues that lower level units of discourse combine to larger ones and that the organization of discourse is oriented to the non-linguistic organization of the live stream.

The smallest unit of communication is the move. Each move in the dataset has been coded for its producer and its recipient. Chapter five demonstrated that the communication can be grouped into monological moves that have no prospection with other discourse, and

dialogical moves that initiate or respond to a participant. In all six broadcasts, monological moves were the most common category for streamers and viewers. This finding suggests that the discourse of live streaming is not an ordinary conversation and needs additional layers of description. The coding of dialogical moves revealed many peculiar properties. For example, streamers tend to produce more responses and react to individual initiations by their viewers. The streamer is more reactionary, and this pattern is consistent across all broadcasts regardless of the game type and size of the audience.

The discourse by the audience is streamer and game-centric. Dialogical moves are written to initiate questions to the streamer and the viewers prefer to interact with the streamer over a dialogue with other viewers (see chapter 5). If viewers are talking among themselves, the topics are treated as background information. They discuss the streamer or the current state of play (see example 10 – 12, p.101f.). Discursive moves about the game and discursive moves with the streamer are the interactional frontstage and the discourse between viewers is the interactional backstage (cf. Kendon, 1992).

At this point, the model moves up the rank scale from the dialogical move to the exchange. In the sixth chapter, the study discussed the five common exchange structures (see figure 33, p.134). Their formulas describe the default patterns of interaction between participants. For example, a streamer-initiated request to the audience can be described as $E_{cross\text{-}modal} = I_{spoken} \rightleftarrows R_{written}$, where the streamer's spoken initiation leads to many responses by the viewers. However, such streamer-initiated exchanges are very unpopular because of *Twitch's* technical limitations and 8-12 second delay. It leads to user practices that mitigate the issue and streamers tend to react to audience initiations, rather than initiating a discussion themselves. The most common exchange structure is $E_{cross\text{-}modal} = I_{written} \rightarrow R_{spoken}$. If the game permits it, the streamer will chain exchanges of this type together in a series and respond to as many as possible.

A unique feature of audience-initiated cross-modal exchanges occurs in the beginning of the R_{spoken} slot. Streamers have to negotiate the demands of the game with the demands of the dialogue. They must decide whether they focus on game or chat. If there is no ongoing and pressing game event, they turn their gaze to the secondary monitor, read the chat and select one of the many audience initiations. Afterwards, they produce the 'topicalizer' usually in the form of a repetition. The topicalizer mentions the main proposition of the initiation. There is minimal variation in this pattern (see p.136) and the streamer uses the topicalizer to reference the question and raise it on the spoken floor. After the topicalizer, streamers re-orient towards the game, which minimizes the time that they are distracted from the gameplay.

Similarly, interrupted exchanges $E_{interrupted} = I_{written} \rightarrow R_{spoken\ (interrupted\ \&\ restarted)}$ are disrupted in the spoken part of the response due to the sudden emergence of a game event. The streamer quickly orients to the gameplay and is commenting on it. This stops their previous response segment. The re-orientation to the dialogue only occurs as the action in the game is slowing down and the streamer has sufficient time.

If there is a series of pressing game events, the streamer will not turn to the chat for extended periods and it leads to many ignored exchanges $E_{ignored} = I_{written} \not\Rightarrow R$. It demonstrates that the dialogical talk is subordinate to the engagement with the game events. Dialogical exchanges are produced in absence of game events and game events can even disrupt exchanges. Quite the opposite holds true for commenting and reporting. They occur in direct temportal proximity and discuss the game-related content. Therefore, it can be said that game events are an important element for the organization of discourse in dialogical exchanges or more monological segments of commenting and reporting.

The next higher rank in the description of discourse are transactions. Their defining feature is the coherence about a single topic. In live streaming on *Twitch*, transactions can consist of a series of dialogical exchanges or a series of commenting and reporting. If there is a game event, the next transaction is very likely to be about the play. The game is a dynamically emerging context and the presence or absence of game events changes the trajectory of transactions.

Paid alert messages are similarly influential and as they occur on screen, streamers quickly tend towards them. Paid alert messages are much more likely to receive a response and they have been classified as a privileged form of communication. They are treated as more important for the interaction and this is an accepted norm in the discourse of live streaming.
The paid alert messages are the most conversational discourse segments of discourse and may be part of a wider development of commodification of social interaction on the web (see 8.3). Every alert message is purchasing a single transaction but as the broadcast continues, the discourse also continues in a new direction. This moment-by-moment development of the discourse's transaction is tied to the non-linguistic organization of the broadcast.

The study describes this rank of the non-linguistic organization as the primacy of pay and play over parley to accounts for the interactional preferences of the participants. For example, streamer and viewers will engage in ordinary dialogical exchanges, i.e. parley[40], if there are no relevant events in the play and if there are no paid alert messages. In each situation,

[40] The study uses the term parley, because the dialogue is at least partly motivated by the goal of 'growing the community' of viewers, who are then expected to provide income to the streamer.

streamer and viewers make choices in the communication and decide to discuss a certain topic at the expense of another. On the level of transactions, the discourse of live streaming is not pre-determined but the primacy of pay and play over parley is the principal of its dynamic emergence and continuous development.

The organization of play is tied to the structure of games in rounds or matches. As similar stages of the game re-occur, their prototypical transactions tend to follow. At the stage of character selection, the transaction is likely to be about the upcoming character choices. Similarly, after a victory or loss in a round, streamer and viewers tend to comment on the outcome of the match or report what could have been done differently. This pairing of game rounds and transactions occurs for the full duration of the daily broadcast.

The daily broadcast session has the length of a work day and as the broadcast ends, the interaction between streamer and viewer ceases. Importantly, the interaction among the viewers end as well. Within minutes after the broadcaster goes offline, the hundreds or even thousands of chat participants leave the chat channel. The streamer and the gameplay are the occasion that brings the viewers to the interaction. Without these two aspects, viewers have no interest in further communication and depart from the website. This point highlights the streamer-orientation of the discourse and challenges claims of community between viewers. The viewers only return to the chat for the next broadcast on the next day. The interaction has precisely defined temporal limits. They occur according to a weekly streaming schedule, which has a set time frame that usually ranges from afternoon to evening.

In summary, the organization of discourse of live streaming is dynamically emergent but always dependent on the non-linguistic organization of the broadcast. On every rank, there is a preference to communicate with the streamer or to comment about the unfolding gameplay in the live stream. If the context permits it, the smallest units of discourse successfully combine into larger units such as cross-modal exchanges or reporting. Participants orient their individual contributions to what they perceive as the most important aspect in the current situation of the broadcasts. The descriptive model illustrates their usual choices from the bottom-up. It describes the overall organization of discourse of live streaming and can explain its constituting parts. Thereby, the study has reached its goal and provided a holistic account of *Twitch*‘s discourse, which includes the streamer, the viewers and the unfolding game play. In the next sections, it is important to look ahead and critically reflect on the limitations of the study as well as future directions

8.2 Critical reflection, limitations and potential improvements

Admittingly, the discussion of limitations is often difficult because research is designed to

address or avoid glaring issues. Therefore, limitations arise from the different choices that are made during the research process. This section will reflect on potential concerns about the transcript, the data selection and the interpretation process. None of these issues are too severe to discredit the contributions of the thesis, but academic integrity requires that certain points are discussed under consideration for future research.

The transcript

A central issue for transcription is capturing the unfolding the interaction. The devised transcription scheme was a first step to begin the study live streaming. It used the chat log of the written chat, transcribed the streamer's spoken language and annotated the gameplay as game events. This made it possible to study the coordination of communication and game on the micro-level up to the largest patterns of discourse. For this purpose, the transcript has been designed as rather broad because it was intended to facilitate the basic research of the study.

Thereby, it left out many aspects that are common to contemporary and multi-modal transcripts in Conversation Analysis, interactional sociolinguistics and other branches of discourse analysis. This section will introduce key aspects of modern multi-modal transcription and how their ideas and principles could be used for a refinement of the transcription scheme or future research. The addressed aspects are 1) a more detailed transcription of chat communication, 2) the annotation of streamer prosody and voicing 3) the transcription of gestures, embodied actions and spatial configurations.

1) One of the earliest criticisms of chat logs and their use in research arose from the fact that chat logs are textual 'products', which are the result of an intricate writing 'process'. In his detailed study Beißwenger (2008) demonstrated that chatting parties in *IRC* alternate between monitoring the chat and looking at the keyboard to type messages. As chat participants notice the posting of a new message, they may revise their initial text and address the new topic. This is usually done by deleting parts or sections in the 'non-sent' message. In academic studies (Beißwenger, 2003; Meredith, 2015) such revisions have been transcribed via strike through text:

I_1 `We could watch the game at home.`

R_0 ~~`hmm I am not sure`~~

I_1 `or in a bar.`

R_1 `yeah let's go out.`

In this invented example, the respondee begins the response to the first part of the initiation but deletes the message as a new suggestion appears on the screen. The latter suggestion 'to go out'

is expressed as the prefererred option and the hesitant rejection R_0 is never sent. Such transcriptions highlight that there is much more conversational coordination in written chats than previously thought. Chat participants constantly negotiate their own messages and turn-taking with the flow of conversation. This insight could be particularly useful to refine the description of commenting by the audience. In my study, commenting has been conceptualized as an individual's textual reaction to a game event. An alternative framing is suggested in interactional sociolinguistics and reception studies. Gerhardt (2012) analysed sport spectatorship on TV in the living room. In her study, the spectating participants negotiated their own commenting with the game events on-screen, the sport announcer talk by the TV commentator, and the discourse of the other co-present spectators in the living room. Gerhardt's (2012) finding of an ongoing turn-negotiation may have a similar implication for the reception and discourse production during online live streaming.

It is possible and even likely that game spectators monitor not just the unfolding of the video game but also the simultaneously developments in the chat. As game events occur, they may react to them, but also to the commenting of other audience members. Such a more audience-focused design might reveal deeper insights in the construction of (digital) sociability (Simmel, 1949) and mediated social presence (Sivunen & Nordbäck, 2014)

2) Voice quality is another aspect that received a very rudamentary transcription in this dissertation. The three main features were sentence-final intonation to distinguish questions and statement, shifts in speech volume to identify response cries and exaggerated performances, and notable shifts in voice quality to highlight the reported speech in topicalizers. Voicing is a pragmatic feature that clarifies the meaning of utterances. It has been a constant research interest in Conversation Analysis (Jefferson, 2004), interactional sociolinguistics (Couper-kuhlen, 1999,) and multimodal interaction analysis (Norris, 2011; Norris & Mayer 2014). Studies in these areas pointed out the complexities of voicing and discussed:

- voice onset
- intonation contours
- word stress via pitch or amplitude
- voiced or voiceless production of consonants
- audible breathing sounds
- laughter

A more delicate transcription of these features would enable the fine-grained study of the streamers' play performance, and how they facilitate the viewers' enjoyment of the vicarious experience (Glas, 2015). Performance studies have been applied to online live streaming and

YouTube 'Let's Play' videos to describe the staged enactment of the player role and the exaggerated displays of enthusiasm (Lowood, 2006; Nyland, 2015; Postigo, 2016). The emotional outbursts during game broadcasting are deliberate performances for the spectating audiences. This expression of emotions has to be well balanced. If the streamer displays too few of them, s/he may be perceived as boring but if the reactions are too frequent or too forceful, the streamer may be seen as 'fake' or inauthentic.

However, the features of 'good' and 'succesful' performances are rather unknown because performance studies often lack actual descriptions of performances. Given a proper transcription of the prosodic features of the streamer's talk, a future study could identify the micro-linguistic realization of play performances. This research would connect to existing literature on streamer persona's and styles (Gandolfi, 2016; see also next section, data selection). For example, the streamer Tyler "Ninja" Blevin became the most popular streamer on *Twitch* in 2017 & 2018 and regularly has over 100 000 viewers because of his strong appeal to younger viewers[41]. Similarly, the success of "PewDiePie" as the world's most popular YouTuber and 'Let's Player' has also been attributed to his ability to hold the attention and interest of very young audiences. The study of the prosodic features of "PewDiePie" and "Ninja" could lead to recommendations of 'best practices' for aspiring live streamers as well as theoretical descriptions of humorous voicing in play-performances.

3) The third limitation and area of potential improvement lies in the transcription of gestures, embodied actions and spatial configurations. Historically, transcripts and the research that they facilitated have been tied to the available recording technologies. For example, Sacks (1984) focused on speech and talk-in-interaction because he had access to tape recorders. Later, Conversation Analysis began to use video, when recording devices became more affordable (Heath et al., 2010). It allowed the inclusion of physical actions in the study of communication. Nevile (2015) describes this new interest in multi-modal communicative ressources as the "embodied turn" and there is a wide range of embodied features that are discussed in contemporary research (Nevile, 2015; Mondada, 2014; Sindoni, 2014; Norris, 2014) including:

- gaze
- eye movement
- posture
- facial expressions
- gestures

[41] http://www.espn.com/esports/story/_/id/24159881/ninja-talks-fortnite-chicago-being-role-model-younger-audiences-red-bull-rise-dawn-event

- pointing
- body torque
- proxemics, i.e. the distance to the camera or between participants
- kinesics, i.e. other types of body movement

Compared to my own research and the design of my transcript, studies on multi-modal interaction demonstrate a minute attention to detail. The focus is on an accurate description of a small-scale phenomenon in all of its interactionally relevant facets.

A key concern has been the representation of the simultaneous usage of several communicative ressources, e.g. speech and pointing (Goodwin, 2000), or the representation of a sequence of embodied actions and their relationship to the ongoing discourse (Mondada, 2011, 2014). Although Nevile (2015) asserts that "no commonly shared format for representing and transcribing embodiment has yet emerged", there is a growing number of studies in Conversation Analysis that apply the multimodal transcription conventions by Mondada (2014). Her transcription conventions have two major strenghts. First, it is an "open" system and new conventions for the annotation of embodied practices can easily be added. This is especially important for online live streaming and its communication with and through information technologies, such as computers or cell phones. Live streaming usually includes the physical handling of an input device, which is essential to the interaction. Secondly, her conventions allow a textual description of ongoing physical actions and a temporal alignment of these actions with the ongoing talk[42]. In her transcript, gestures and descriptions of embodied actions are delimited between two identical symbols and are synchronized with the correspondent stretches of talk. The result is a very precise, multi-modal description of the participants interaction that contains the talk, bodily conduct and the interaction with physical objects.

Such a more detailed transcript is very valuable to the analysis of online live streaming because very little is known about the way streamers negotiate the demands of the game, i.e. game-oriented gaze, posture and control of the gaming hardware, with the competing needs of the interaction with the audience. My own research project only covers the surface of this complex interaction but there are two ongoing dissertation projects by Neise (forthcoming) and Meyer (forthcoming) that are significantly more detailed and have a special emphazise on the role of physical actions and their embeddedness in the spatial environment of the interaction.

[42] The ability to account for the progression of actions has been seen as one of the major shortcomings of screen shots (see section 4.2).

The data selection

The game and channel selection have been informed by the literature review, the game's success on *Twitch* and to some extend the authors knowledge about video game. This criteria-driven selection of broadcasts and channels occurred prior to any linguistic analysis. Therefore, the study remained unaware of other mediating factors that were underrepresented in the literature review or difficult to assess prior to the transcription. Gandolfi (2016) and Hamilton et al. (2014) mention a streamer's conversational style and argue that it might influence the interaction with the audience. This hypothesis has not been tested in the research at this point and this study did not consider streamer types as a major factor during data selection. There was little previous literature about them and it is difficult to identify streamer types before the actual study of the data. Therefore, the research operated under the assumption that the streamer's personality would not significantly alter the discourse of live streaming. For the most part, this initial assumption was supported by the findings in the analysis. All streamers engage in the same discursive practices as they transition between audience-related and game-related talk. However, the fifth chapter demonstrates that there are differences in the rate of talk. The large-sized WoW streamer talked significantly less than all other participant and his audience also produced fewer chat messages, which suggests that the two behaviors are interconnected. A vocal and active streamer may encourage a more active audience. This raises the question if there are also other differences and factors that may be relevant, but have not been accounted for in the data seleciton

One candidate for such variation is the gender and gengered language of *Twitch*. In many examples of the study, participants communicated under the assumption that they are all men or mostly men. Streamers and viewers referred to each other as 'guys' or 'boys' and there is a significant amount of swearing and cursing. Female terms of address were absent. One may ask if this is the result of the sample of selected streamers, who were all male, or a feature of *Twitch's* discourse more generally. There is an argument to be made for both cases and neither of them can be dismissed without further inquiries. On the one hand, gamer culture and the culture of gaming is often male-gendered and male-dominated (Taylor, 2012). In this scenario, female streamers orient towards the dominant cultural norms and adopt male-gendered discursive practices in their interaction with the audience. Linguistically, they might position themselves as 'one of the guys' and treat everyone 'buddies'. It would imply that *Twitch's* discourse is overall or in tendency a male-gendered domain.

On the other hand, it is possible that the sample does not represent the full spectrum of linguistic diversity. The data selection of this study excludes the criterion of gender and only

selects for games and channel sizes. Thereby, the sampling unwillingly reproduces existing gender biases on *Twitch*. Since there are significantly more male streamers, the selected channels also had a male bias. The only recourse to this issue is an active intervention at the point of data selection. One might record and study male and female streamers in an equal proportion. However, such an intervention would require a justification that is connected to the research question and literature review. Partially, the research must become about language and gender. This was not deemed necessary or desirable for this study because there was no prior literature that indicated that the organization of discourse may vary by gender. This initial assumption has been supported by the results of the study and there is nothing to suggest that male and female streamers might act differently when they are commenting, reporting or talking to the audience. If there are gender differences to be found, they operate on a different linguistic level that is not part of the research focus of the study.

It is an interesting research question if there are nuanced live streaming styles that correspond to gender, or if there is a prototypical discourse that is emulated and adhered to by nearly all participants.

Interpretation of data

The third potential limitation arises at the intersection of the data with the researcher. Although the interpretation was guided by the discourse analytical framework, it still involves a process of reading and sense making. My own position is that of an informed observer, who is very familiar with video games and *Twitch* but not part of a specific live stream community. Similarly, the data covered a diverse set of channels of various sizes and game types. From this vantage point, I also argued that there is very little discourse that is primarily interpersonal. Although the concepts of family and friends are part of Twitch's self-characterization (see 2.2.1), genuinely befriending discourse appears superficial and rare. There is informal and friendly communication that indicates a low social distance, but detailed and personalized discourse tends to be streamer-centric and is often tied to paid alert messages. With such linguistic evidence, the study is very critical of notions if family or friends and considers them inadequate even in a more metaphorical sense.

However, there are three possible arguments that this scepticism might not be waranted. First, the data covered a two-hour period in each of the broadcasting channels. In such a brief window of time, it is difficult to uncover channel-specific discourse. Hamilton et al. (2014) argue that many streamer's and streamer communities have their own set of inside jokes that are known to regular viewers and core members. A spectator that is unfamiliar with the channel may overlook its specific discursive practices and cannot evaluate their interpersonal function

for the participants. A solution to this concern could be longitudinal research, which focuses on a smaller number of channels and their unique elements. This type of research has already been undertaken in other CMC environments including *Twitter* (Dayter, 2016) or 'camgirl' video chat (Senft, 2008). These studies demonstrate that sub-cultures develop distinct interpersonally-oriented linguistic practices that are rooted in the community's particular experiences and beliefs. These practices promote a shared group identity among members without the need of direct and close interpersonal exchange between individuals. It is possible that *Twitch's* channels have similar qualities and form their own sub-communities within the larger community on *Twitch*.

A second potential argument in the defense of meaningful social ties could consider the interpersonal function of *Twitch*'s game-related discourse. Participants have a well-established and shared linguistic repertoire, which they apply for their specific purposes on *Twitch.* For example, streamers engage in reporting that is full of gamer jargon to express complex game events and their decision making. A shared linguistic repertoire, common linguistic goals and a specific linguistic context are the main features of a discourse community (Swales, 1990). Discourse communities do not necessarily feature interpersonal communication between members, but they may lead to a sense of belonging centered around the topic of interest. The shared interest is expressed in an expert discourse that conveys specialized and domain specific meanings. In the context of *Twitch*, the spectators communicate with like-minded individuals, who are very interested in video games. The perception of *Twitch*'s discourse depends on the position of its reader. An informed but detached researcher may interpret the communication as less social than it appears to its core participants. Regular viewers of a channel may see themselves as part of a tight-knit community, even if they are not directly talking to each other about personal issues (Hamilton et al., 2014).

Linguistic behavior in an interaction must be distinguished from its perception and the lived experience of the participants. Throughout the thesis, the study remains critical of reported behavior or reported intentions (cf. Hamilton et al., 2014; see section 2.2). Nevertheless, my position is not denying or delegitimizing the audiences' and streamers' subjective experiences during live streaming. The research format cannot evaluate the emotional affect that the communicative interaction may elicit in the viewers or streamers. Even superficial social contact may feel deeply rewarding to viewers because their communication is combined with a hobby that they feel very passionate about. Due to the study's design, there is a persistent epistemological blind spot that emphasizes linguistic data at the expense of the participants' perspectives. To address this concern, the study concludes with suggestions for future research

that build and elaborate on the findings of the dissertation, while also taking the participants point of view into consideration.

8.3 Future directions – the commodificaiton of social interaction

One of the most interesting aspects of this study has been the important role of paid alert messages for the discourse of live streaming. They greatly influence the organization of communication and they are very important to streamers and viewers. They are a linguistic and sociological phenomenon that is interesting in itself, but also in its broader significance for concepts such as interpersonal interaction, casual conversation as well as labor and capital.

Traditionally, the study of conversation and other forms of interpersonal talk excluded economic considerations and often defined this talk in contrast to utilitarian thinking. Eggins & Slade (1997) define casual conversation as 'talk for the sake of talk' and Goodwin (2000) names conversation as the 'primordial site of human interaction'. Conversation and social interaction are framed as a natural and pure phenomenon that is part of the human condition. Interpersonal contact is understood as an intrinsic desire of all people and they engage in conversations to meet their interpersonal needs. From this perspective, interpersonal interaction is thoroughly positive, beneficial to everyone involved and should be encouraged in all its diverse forms. Most of the initial research on *Twitch* shares this enthusiastic outlook and lauds *Twitch*'s communication as a meaningful interaction that enriches the experience for all parties (cf. Olejniczak, 2015, Ford et al., 2017). In this line of thinking, paid alert messages are tokens of appreciation by the viewers to the streamers for their commendable live stream. The live stream is an entertaining social good that is supported through well-meant 'donations' by patrons, who mostly report personal and social motivations (cf. Sjoblöm & Hamari, 2016). I shared this thinking and its reasoning at the beginning of my studies in 2014. However, my attendance at the inaugural TwitchCon convention, reading-suggestions by my supervisors and detailed analysis of my data began to challenge my early assumptions and one-dimensional perspective.

In the data, aspiring and professional streamers continuously stress the notion of work (see 6.3), and at the convention many participants had a very professional perspective on their play and live streaming. Such aspects are outside of the core area of linguistics and are often discussed within cultural studies. Cultural studies encourages the critical and reflective reading and extends notions of labor and capital to new phenomena. Within this research tradition, there are several relevant works that discuss gaming and let's play videos as a new type of labor (Jin, 2008; Postigo, 2016). The studies demonstrate that labor occurs in many different forms and is not exclusive to traditional factory of office work (Pearce, 2006). Play becomes labor as players

professionalize their gaming activity and their motivation shifts from enjoyment of the game towards monetary considerations.

Through continuous play-labor, players can accrue gaming capital (Walsh & Apperley, 2009; Consalvo, 2005). This gaming capital has an embodied and institutional dimension (cf. Bourdieu, 1986). Embodied gaming capital refers to the players knowledge about the game and his or her physical habitus. Experienced gamers know the right strategies and they internalized the necessary controls of mouse and keyboard to perform them. Their play is far superior to ordinary people and the very best players compete in global tournaments (Jin, 2008; Taylor, 2012). Through their successes, they build institutional gaming capital, which are marks of distinction such as tournament titles, in-game achievements but also followers on social media platforms such as *YouTube* (Walsh, 2008; Postigo, 2016). Institutional gaming capital leads to the recognition of a player by his or her peers as an exceptional gamer. Embodied and institutional gaming capital can be transformed to economical capital (Bourdieu 1986). For example, excellence in play (embodied capital) may translate into the winning of tournaments and a payout of prize money (Walsh, 2008). Institutional gaming capital may lead to sponsorships or large viewer numbers of one's *YouTube* channel. In turn, the viewer numbers on *YouTube* generate advertisement revenue for the player. There is a monetization of play through established channels with methods that are very similar to professional sports. Teams, tournaments and *YouTube* are institutions that pay the players for their play-labor (Postigo, 2016).

Play-labor is also very important on *Twitch,* but it is not the primary mechanism for generating the streamer's income. The play of the video game is necessary to get viewers interested at the beginning, but their long-term commitment to the channel depends on their enjoyment of the interaction (Hamilton et al., 2014; Sjoblöm & Hamari, 2016). Viewers must appreciate the streamer or otherwise they are unlikely to continue watching. As demonstrated in chapter 7, the paid alert messages revolve around the exchange of interpersonal meanings between streamer and viewers. So far, there are no studies that discuss alert messages or the interpersonal discourse on *Twitch* in light of this process of commodification.

However, there is prior research in cultural studies that may be a suitable starting point to combine the linguistic research on conversation with research on play and labor. Hochschild (1983) discusses the commercialization of human feelings in her study of flight attendants in an American airline. The work of flight attendants consists of several tasks such as maintaining order on board, preparing the cabin for start and landing as well as catering the passengers with food and refreshments. Outside of these manual tasks, flight attendants are also expected to be

very polite, friendly and attentive. It is an essential requirement of their job and they engage in 'emotional labor' (Hochschild, 1983). Their emotional labor is a performance that may not correspond to their true inner states or mood and their displays of emotion are in service of their company and profession. The polite and friendly demeanor has an underlying utilitarian motivation. This brief explanation demonstrates that there could be some parallels between Hochschild's (1983) ideas and my potential future research.

Arguably, live streamers on *Twitch* perform two kinds of labor. They perform play-labor in the video game and their interpersonal communication may be conceptualized as a form of emotional labor. Commenting consists of exaggerated emotional outbursts and is a performance to entertain the viewer. The response to subscription notification is a ritualistic inclusion of the viewer into the streamer-centric community. Both linguistic practices cater to the emotional and social needs of the viewers. They have characteristics of labor, if they are motivated by the financial incentives that are provided in and through the paid alert messages. The potential earnings of professional live streamers have been noticed and are actively discussed on *Twitch's* online forum, newspaper articles[43] and at TwitchCon. Live streaming is seen as a viable career path and many gamers try to emulate the successes of high-profile streamers. Since sociable and pleasant interactions are necessary for maintaining and growing an audience, the friendly talk blurs the lines between interpersonal and functionally-orientation. Traditionally, these two positions have been studied from perspectives that take very contradictory starting points and it would be a valuable inquiry because it could improve our knowledge about live streaming and refine the understanding or definitions of casual conversation in linguistics.

The design of the study would continue the trajectory of the thesis, while improving on some of its shortcomings. The research would move towards a longitudinal perspective and include secondary data sources that provide the perspective of the live stream participants, i.e. the streamers and viewers. As the primary data, the screen recordings would maintain their central position. My previous experiences suggest that successful streamers are very skeptical of academic research and do not consider it worthwhile or even a risk for their careers. However, medium-sized channels remain more approachable and could be more cooperative.

These channels would be recorded periodically during the time frame of one to three months. It would allow the researcher to identify peripheral and core members and look for signs of continuation in their discourse (cf. Hamilton et al., 2014). The study would look for common themes or topics that consistently re-occur such as channel-specific inside jokes or references to past events. They could indicate the existence of a community repertoire that is

[43] https://www.nytimes.com/2017/04/12/technology/personaltech/paying-for-live-stream-chat.html

shared among participants for social purposes. Additionally, I would look for signs of persistent social ties between participants. One may study if participants are talking or writing with one another and about themselves or if the discourse indicates familiarity. The findings of the thesis suggest that streamers maytalk about their private life but it has been very difficult to assess the prevalence of self-disclosure by viewers.

The future study's longitudinal design would provide additional opportunities to discover such discourse. Its distinction in core and marginal members would make it possible to evaluate if there are more interpersonal exchanges between established members. Additionally, it would be very interesting to see if subscribers and donators behave differently or if they are treated differently by the streamer outside of the context of their paid alert messages. Identifying interpersonal discourse practices that only occur between streamer and subscribers but are absent from the streamer's interaction with his or her other core members would make strong case for a selective commodification of interpersonal communication.

The analysis of the interactional data can be supplemented with secondary data sources. Semi-structured interviews with the streamer, subscribers and core members of the channel would provide the perspective of the participants and contextualize the interpretation of the researcher. One might ask audience members to explain some of the channel-specific discourse practices or if they are familiar with other viewers in the channel. If the channels are more than very loose associations of spectators, then the subscribers and core members should be able to talk about the channel community, their defining characteristics, and why they appreciate the channel. A similar set of question can also be asked towards the streamers to see if their reported reception of the interpersonal communication matches the discursive practices in the channel.

As an additional topic, streamers could be asked to comment on their role in the maintenance of the channel community. This part is very sensitive because the researcher must be careful not to pre-frame the possible responses. Ideally, the streamer should talk freely about his or her efforts to foster communication without being primed for an interpretation in terms of labor. It may be achieved through varies questions that are related to the issue. For example, one may ask about the performance element of live streaming, the constant need to portray positive emotions and engage the audience. Questions about frictions with audience members and feelings of constraints in the streamer-viewer interaction are interesting as well. They indicate that streamers must censor their wording and expressions to fulfil their duties as a streamer.

Panel discussions at TwitchCon and other conferences such as 'Pax' demonstrate that streamers are often very willing to report honestly and openly about their relationships on

Twitch, even if it addresses unpleasant or negative topics. Their feedback is a form triangulation of the primary data, which will result in a realistic assessment of the quality of social interaction and community construction. It will make it possible to refine the binary distinction between interpersonal discourse and utilitarian or goal-oriented talk towards a more nuanced position. This makes it a sensible extension of the dissertation and a valuable avenue for future research.

Appendix

Transcription Conventions

Symbol	Meaning
TEXT	Emphasis or higher volume
°text°	Softer speech.
.	Falling final intonation
?	Rising final intonation
:	Abridged columns & elongated vowel sounds
-	Word cut off
“	Audible shift in voice quality such as a rise in pitch
{ }	Physical actions by the streamer
(..)	Brief pauses within an intonation unit
=	Latching
[ ]	Overlap in speech
<text>	Slower speech
>text<	Faster speech
	Reference to Prior Discourse
	Coherent Stretch of Discourse

Top 20 Games December 2012- February 2015

Name	Genre
Alien: Isolation	First Person Shooter
Assassin's Creed 4	Action Game
*ArmA III	First Person Shooter
Battlefield 4	First Person Shooter
Bioshock: Infinite	First Person Shooter
Call of Duty Modern Warfare	First Person Shooter
*Counter-Strike: Global Offensive	First Person Shooter
Cube World	Open World Game
Dark Souls I	Action Game
DayZ	First Person Shooter
*Destiny	First Person Shooter
*Diablo III	Massively Multiplayer Online Roleplaying Game
*Dota 2	Multiplayer Online Battle Arena
Dying Light	First Person Shooter
*EVE Online	Massively Multiplayer Online Roleplaying Game
***FIFA Soccer**	**Sport Simulation**
*Final Fantasy XIV	Massively Multiplayer Online Roleplaying Game
Game Dev Tycoon	Economic Simulation
*GTA V	Action Game
*Hearthstone	Trading Card Game
Heroes of Newerth	Multiplayer Online Battle Arena
*H1Z1	First Person Shooter
Injustice: Gods Among Us	Fighting Game
***League of Legends**	**Multiplayer Online Battle Arena**
*Magic: The Gathering	Trading Card Game
*Minecraft	Open World Game
Monster Hunter 3	Massively Multiplayer Online Roleplaying Game
Neverwinter	Massively Multiplayer Online Roleplaying Game
Outlast	Role Playing Game
*Path of Exile	Massively Multiplayer Online Roleplaying Game
Payday 2	First Person Shooter
Pokemon Red/Blue	Role Playing Game
Pokemon X/Y	Role Playing Game
*RuneScape	Massively Multiplayer Online Roleplaying Game
*Smite	Multiplayer Online Battle Arena
*Starcraft 2	Real Time Strategy Game
Spelunky	Side Scroller
Super Street Fighter IV	Fighting Game
Super Smash Bros.	Fighting Game
The Binding of Isaac	Action Game
The Evil Within	Action Game
The Legend of Zelda	Role Playing Game
The Sims 4	Life Simulation
The War Z	First Person Shooter
Titanfall	First Person Shooter
Ultimate Marvel vs. Capcom 3	Fighting Game
Watch Dogs	Action Game
Wildstar	Massively Multiplayer Online Roleplaying Game
***World of Warcraft**	**Massively Multiplayer Online Roleplaying Game**

* indicates game that are still live-streamed as of 2018

Bold font indicates the selected games

List of Channels

Game	Size	URL
League of Legends	Large-sized	https://www.twitch.tv/tsm_dyrus
League of Legends	Large-sized	https://www.twitch.tv/c9sneaky
League of Legends	Large-sized	https://www.twitch.tv/aphromoo
League of Legends	Large-sized	https://www.twitch.tv/voyboy
League of Legends	**Large-sized**	**https://www.twitch.tv/tsm_bjergsen**
League of Legends	Large-sized	https://www.twitch.tv/imaqtpie
League of Legends	Medium-sized	https://www.twitch.tv/karasmai
League of Legends	**Medium-sized**	**https://www.twitch.tv/painlessdeath22**
League of Legends	Medium-sized	https://www.twitch.tv/ipav999
League of Legends	Medium-sized	https://www.twitch.tv/rflegendary
League of Legends	Medium-sized	https://www.twitch.tv/pantsaredragon
FIFA	Large-sized	https://www.twitch.tv/castro_1021
FIFA	**Large-sized**	**https://www.twitch.tv/nick28t**
FIFA	Large-sized	https://www.twitch.tv/drjarba
FIFA	Medium-sized	https://www.twitch.tv/spenfc
FIFA	**Medium-sized**	**https://www.twitch.tv/therealpugzilla**
FIFA	Medium-sized	https://www.twitch.tv/buckarmy
FIFA	Medium-sized	https://www.twitch.tv/brendo_yo
World of Warcraft	Large-sized	https://www.twitch.tv/xaryulol
World of Warcraft	Large-sized	https://www.twitch.tv/venruki
World of Warcraft	Large-sized	https://www.twitch.tv/cdewx
World of Warcraft	**Large-sized**	**https://www.twitch.tv/psherotv**
World of Warcraft	Large-sized	https://www.twitch.tv/snutzy
World of Warcraft	Large-sized	https://www.twitch.tv/mitchjones
World of Warcraft	Medium-sized	https://www.twitch.tv/tosantribe
World of Warcraft	Medium-sized	https://www.twitch.tv/hotted89
World of Warcraft	Medium-sized	https://www.twitch.tv/smexxin
World of Warcraft	Medium-sized	https://www.twitch.tv/maldiva
World of Warcraft	**Medium-sized**	**https://www.twitch.tv/woundman**

Bold font indicates selected channels

Moves in Casual Conversation (Eggins and Slade, 1997)

Speech function labels for OPENING moves

Speech function	Discourse purpose	Congruent mood	Example
attending	attention seeking	minor; formulaic	Hey David!
offer	give goods and services	modulated interrogative	Would you like some more wine?
command	demand goods and services	imperative	Look.
statement:fact	give factual information	full declarative; no modality; no appraisal	You met his sister
statement:opinion	give attitudinal /evaluative information	full declarative; modality and/or appraising lexis	This conversation needs Allenby.
question:open:fact	demand factual information	wh-interrogative; no modality; no appraisal	What's Allenby doing these days?
question:closed:fact	demand confirmation/ agreement with factual information	polar interrogative; no modality; no appraisal	Is Allenby living in London now?
question:open:opinion	demand opinion information	wh-interrogative; modality/appraisal	What do we need here?
question:closed:opinion	demand agreement with opinion information	polar interrogative; modality/appraisal	Do we need Allenby in this conversation?

Figure 42 Opening Moves and Speech Functions

Summary of continuing speech functions

Speech function	**Discourse purpose**	**Congruent mood**	**Example**
Continue:monitor	check that audience is still engaged	elliptical major clause or minor clause with interrogative intonation	You know? Right?
Prolong:elaborate	clarify, exemplify or restate	full declarative, linked (or linkable) by: for example, I mean, like	see moves: 4b, 10b, 17b, 17c, 20b, 21b, 46c, 56c in 5.1
Prolong:extend	offer additional or contrasting information	full declarative, linked (or linkable) by: and, but, except, on the other hand	see moves: 42b, 45b, 55c, 58b, 59b in 5.1
Prolong:enhance	qualify previous move by giving details of time, place, cause, condition etc.	full declarative, linked (or linkable) by: then, so, because.	see moves: 3b, 30b, 43b in 5.1
Append: elaborate	clarify, exemplify or restate previous move after intervention by another speaker	elaborating nominal group	see moves: 14, 20a and 62 in text 5.1
Append:extend	offer additional or contrasting information to previous move after intervention by another speaker	extending nominal group	see example from text 3.1 presented above
Append:enhance	qualify previous move after intervention by another speaker	enhancing prepositional/adverbial phrase	see example from text 3.1 presented above

Figure 43 Continuing Moves and Speech Functions

Summary of sustaining responding speech functions

Speech function	Discourse purpose	Congruent mood	Example
engage	show willingness to interact by responding to salutation, etc.	minor: typically "yea" or matched response	Hi-*Hi*; Nick? – *Yea?*
register	display attention to the speaker	repetition of speaker's word(s); paralinguistic expressions such as Mmm, Uh huh; ritual exclamations; minor clauses	see moves: 29, 38a, 41/a, 55/b in text 5.1
comply	to carry out demand for goods and services	non-verbal; expressions of undertaking (e.g. "OK")	Can you pass the salt, please? – *Here/ [passes it]*
accept	to accept proffered goods and services	non-verbal; expressions of thanking	Have another? – *Thanks/ [takes one]*
agree	to indicate support of information given	Yes; positive polarity;	see move 16 in text 5.1
acknowledge	to indicate knowledge of information given	expressions of knowing	see moves 10/a, 11, 13, 41/b in text 5.1
answer	to provide information demanded	complete missing structural elements	Where's Allenby? – *In London*
affirm	to provide positive response to question	Yes; positive polarity	Have you heard from him lately? – *Yes, I have/only yesterday*
disagree	to provide negative response to question	negation of proposition	Is he in London now? – *No*
non-comply	to indicate inability to comply with prior command	non-verbal; no expressions of undertaking; negation of verbal command	Could you pass me the salt, please? – *sorry/can't reach/got my hands full*
withhold	to indicate inability to provide demanded information	negative elliptical declarative	When is he due back? – *I've no idea*
disavow	to deny acknowledgement of information	expressions of disclaiming knowledge	see move 24/b in text 5.1
contradict	to negate prior information	No; switched polarity	see moves: 3/a, 7/a, 43/a, 46/a in text 5.1

Figure 44 Responding Moves and Speech Functions

Summary of sustaining rejoinder speech functions

Speech function	Discourse purpose	Congruent mood	Example
check	to elicit repetition of a misheard element or move	elliptical polar interrogative	see moves: 18, 33, 46b in text 5.1
confirm	to verify information heard	elliptical wh-interrogative; wh/element from prior move	see move 24a in text 5.1
clarify	to get additional information needed to understand prior move	elliptical interrogative; wh/new element (not in prior move)	see move 24c in text 5.1
probe	to volunteer further details/implications for confirmation	full clause, new subject, etc. but in logico-semantic relation with the moves it's tracking or tagged declarative	see move 27 in text 5.1
resolve	to provide clarification, acquiesce with information	elliptical declarative; mood adjunct of polarity or modality	see moves: 28, 35, 37, 39 in text 5.1
detach	to terminate interaction	silence; expression of termination	see move 52 in text 5.1
rebound	to question relevance, legitimacy, veracity of prior move	wh-interrogative, elliptical	see moves: 2a, 48 in text 5.1
counter	to dismiss addressee's right to his/her position	non-elliptical declarative; negation of understanding/rightness	see move 46b in text 5.1
refute	to contradict import of a challenge	elliptical declarative; negation	see move 47 in text 5.1
re-challenge	to offer alternative position	elliptical interrogative	see move 23 in text 5.1

Figure 45 Rejoinder Moves and Speech Functions

Acts of Everyday Conversation (Francis & Hunston, 1992)

The acts of everyday conversation

Label	*Symbol*	*Realization and function*
framer	fr	Realized by a closed class of items: (i) 'OK', '(all) right', 'anyway' and their variants, where the item precedes an exchange-initial move head ('anyway' may also be embedded in a move head); (ii) 'well', 'now', 'good' and their variants, where the item precedes an exchange-initial move head and is said with high key falling intonation followed by silent stress. Its function is to mark boundaries in the conversation, where such an interpretation is consistent with considerations of topic.
marker	m	Realized by the same closed class of items as *fr*: (i) 'OK' etc. where the item precedes a non-exchange-initial move head; (ii) 'well' etc. (also 'oh', 'er(m)' and 'look') where not said with high key falling intonation. Its function is to mark the onset of a move.
greeting	gr	Realized by a closed class of items which form the first-pair parts of the adjacency pairs used in the rituals of greeting and leave-taking: 'hello', 'hi' 'good morning', '(good)bye(-bye)', 'have a nice/ good day', 'be seeing you' and their variants. Its function is self-explanatory.
reply-greeting	re-gr	Realized by a closed class of items which form the second-pair parts of the adjacency pairs used in the rituals of greeting and leave-taking: 'hello', 'hi', 'good morning', '(good)bye(-bye)', 'fine thanks (and you?)', 'thank you', 'same to you', 'yeah see you', and their variants. Its function is self-explanatory.
summons	sum	Realized by the ringing of the telephone, a knock at the door, etc., or the calling of somebody's name. Its function is to engage another participant in a conversation or to attract his/her attention.

starter	s	Realized by statement, question, command or moodless item. Its function is to provide information about or direct attention towards the act realizing the move head.
meta-statement	ms	Realized by statement, question or command. Its function is to structure the conversation prospectively in some way, and to obtain a warrant for doing so.
conclusion	con	Realized by a statement or question often with anaphoric reference. Its function is to 'tie up' a particular topic, and to obtain a warrant for doing so.
acquiesce	acq	Realized by 'yes' and other items indicating assent, both verbal and non-verbal. May also be realized by silence, interpreted as a default mechanism whereby failure to protest (*rej*) is an indication of acquiescence. Its function is to provide a warrant for a suggestion as to prospective (*ms*) or retrospective (*con*) structuring made by the other participant in a two-party conversation.
reply-summons	re-sum	Realized by the items used to answer a telephone ('hello', the giving of one's number, etc.) or the door (opening it, calling 'come in', etc.) or by 'yes', 'what?' and other indications of attention (both verbal and non-verbal) given upon hearing one's name called. Its function is to indicate willingness to participate in a conversation, or that one is giving one's attention.
inquire	inq	Realized by questions which seek information as opposed to a 'yes' or 'no' answer, i.e. wh-questions and ellipted forms of these. Its function is to elicit information.
neutral proposal	n.pr	Realized by questions which seek a 'yes' or 'no' answer, i.e. questions beginning 'Do you', 'Are you', etc, and ellipted forms of these. Its function is to elicit a decision between 'yes' and 'no'.
marked proposal	m.pr	Realized by questions which seek a 'yes' or 'no' answer, where the form of the question indicates the polarity of the expected answer, i.e. questions beginning 'Don't you', 'Aren't you', etc. It is also realized by declaratives said with 'questioning' intonation and declaratives followed by tag questions. Its function is to elicit agreement.
return	ret	Realized by question, often ellipted. Its function is to seek clarification of a preceding utterance.
engage	eng	Realized by 'mm', 'yeah' and low or mid key 'echoes'. Its function is to provide minimal feedback while not interrupting the flow of the other participant's utterance.

loop	l	Realized by a closed class of items: 'pardon', 'what', 'eh', 'again', and their variants, said with rising intonation. Its function is to elicit the repetition of a preceding utterance which was not clearly heard.
prompt	p	Realized by a closed class of items: 'hah' (with rising intonation), 'come on', 'go on give me an answer', 'guess' and their variants. Its function is to reinforce the point of a preceding utterance, whether this was to elicit an *i*, a *conc* (etc.) or a *be*. When it realizes a move-head, it follows a silence on the part of 'B'.
observation	obs	Realized by statement. Its function is to offer 'information' which is already part of the shared knowledge of the participants in the conversation. In other words it has a predominantly phatic function.
informative	i	Realized by statement or by 'yes' and 'no' items and their variants, both verbal (e.g. 'I (don't) think so') and non-verbal (e.g. nods and shakes of the head). Its function is to supply information or to give a decision between 'yes' and 'no'.
concur	conc	Realized by low or mid key 'yes' and 'no' items and their variants, both verbal and non-verbal; or by repetition or paraphrase. Its function is to give agreement.
confirm	conf	Realized by high key 'yes' and 'no' items and their variants, both verbal and non-verbal; or by repetition or paraphrase. Its function is to give or assert agreement.
terminate	ter	Realized by low key 'yes' and 'no' items, and their variants, both verbal and non-verbal; or by low key repetition. Its function is to acknowledge a preceding utterance and to terminate an exchange (although it may be followed by further acknowledging moves).
receive	rec	Realized by mid key 'yes' and 'no' items and their variants, both verbal and non-verbal; or by mid key repetition. Its function is to acknowledge a preceding utterance or (as pre-head) to indicate that the appropriate *i*, *be*, etc. is forthcoming.
react	rea	Realized by high key 'yes' and 'no' items and their variants, both verbal and non-verbal; or by high key repetition. Its function is to indicate positive endorsement of a preceding utterance.

qualify	qu	Realized by 'qualified' statement or by tentative 'yes' and 'no' items (where tentativeness is intonationally signalled) and their variants, both verbal ('to some extent yes', 'no not really', 'well I suppose so (not)', etc.) and non-verbal (e.g. shrugging the shoulders). Its function is to qualify a decision or an agreement by indicating that its polarity is not unconditional, or to detail conditions and exceptions.
reject	rej	Realized by statement or by 'yes' and 'no' items and their variants, both verbal and non-verbal. May also be realized by silence, interpreted as a default mechanism whereby failure to supply a *re-gr*, *re-sum*, *i*, *conc*, *conf*, *qu* or appropriate *be* is an indication of rejection. Its function is to refuse to acquiesce to a suggestion as to the structuring of the conversation; or to refuse to give an appropriate answer to a *gr* or a *sum*, or to reject the underlying presuppositions of an *inq*, *n.pr* or *m.pr*; or to indicate unwillingness to comply with a *d*.
reformulate	ref	Realized by statement which paraphrases a preceding utterance. Its function is to acknowledge a preceding utterance or offer a revised version of it.
endorse	end	Realized by statement or moodless item. Its function is to offer positive endorsement of, sympathy with, etc., a preceding utterance ('good idea', 'you poor thing', 'well I never', 'very interesting', etc.).
protest	prot	Realized by statement or by 'yes' and 'no' items and their variants. Its function is to raise an objection to a preceding utterance; it acknowledges the utterance while disputing its correctness, relevance, appropriateness, the participant's right to have uttered it, or anything else.
directive	d	Realized by command. Its function is to request a non-verbal response, i.e. an action.
behave	be	Realized by action. Its function is to provide a non-verbal response to a preceding *d*, whether this involves compliance non-compliance, or defiance.
comment	com	Realized by statement. Its function is to exemplify, expand, explain, justify, provide additional information, or evaluate one's own utterance.

Acts in spoken interaction (Stenström, 1994)

Primary acts	
<accept>	agrees to a <request>, <suggest>, etc
<acknowledge>	signals receipt of information
<agree>	signals agreement with what was just said
<alert>	calls the addressee's attention
<answer>	responds to a <question>/<request>
<apology>	expresses regret
<call-off>	prompts a conversational closing
<check>	asks for clarification
<closer>	ends a conversational closing
<confirm>	responds to a request for confirmation
<disagree>	expresses disagreement
<evaluate>	judges the value of what the previous speaker said
<greeting>	greets somebody or bids farewell
<inform>	provides information
<invite>	asks if somebody 'would like to do X'
<object>	signals a different opinion
<offer>	presents something for acceptance/rejection
<opine>	gives one's personal opinion
<query>	expresses doubt or strong surprise
<question>	asks for information, confirmation, clarification
<react>	expresses attitude and strong feelings
<reject>	disagrees to a <request>, <suggest>, etc
<reply>	responds to a <statement>
<request>	asks somebody to do something
<smoother>	responds to an <apology>
<statement>	informs or expresses opinion
<suggest>	puts forward an idea or a plan
<thanks>	expresses gratitude

References

Aarsand, P. A., & Aronsson, K. (2009). Response cries and other gaming moves—Building intersubjectivity in gaming. *Journal of Pragmatics*, *41*(8), 1557–1575. https://doi.org/10.1016/j.pragma.2007.05.014

Aarseth, E. (2003). Playing Research: Methodological approaches to game analysis. In *The MelbourneDAC - the 5th International Digital Arts and Culture Conference.* Retrieved from http://hypertext.rmit.edu.au/dac/papers/Aarseth.pdf

Aparicio-Pardo, R., Pires, K., Blanc, A., & Simon, G. (2015). Transcoding live adaptive video streams at a massive scale in the cloud. In *MMSys '15 - the 6th ACM Multimedia Systems Conference* (pp. 49–60). New York: ACM. https://doi.org/10.1145/2713168.2713177

Antheunis, M. L., Valkenburg, P. M., & Peter, J. (2012). The quality of online, offline, and mixed-mode friendships among users of a social networking site. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, *6*(3). https://doi.org/10.5817/CP2012-3-6

AnyKey. (2015). *Workshop #1 White Paper: Women in Esports*. San Francisco.

AnyKey. (2016). *Workshop #2 White Paper: Barriers to inclusion and retention: The role of community management and moderation.*

Ariely, D., Bracha, A., & Meier, S. (2009). Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially. *American Economic Review*, *99*(1), 544–555. https://doi.org/10.1257/aer.99.1.544

Ashmore, M., MacMillan, K., & Brown, S. D. (2004). It's a scream: professional hearing and tape fetishism. *Journal of Pragmatics*, *36*(2), 349–374. https://doi.org/10.1016/S0378-2166(03)00005-5

Ashmore, M., & Reed, D. (2000). Innocence and Nostalgia in Conversation Analysis: The Dynamic Relations of Tape and Transcript. *Forum: Qualitative Social Research*, *1*(3). Retrieved from http://www.qualitative-research.net/index.php/fqs/article/view/1020

Auer, P. (1996). From Context to Contextualization. *Links & Letters*, *3*, 11–28.

Ayaß, R. (2015). Doing data: The status of transcripts in Conversation Analysis. *Discourse Studies.* Advance online publication. https://doi.org/10.1177/1461445615590717

Baerg, A. P. (2006). The Role Of The Digital Sports Game In The Sports Media Complex (PhD Dissertation). University of Iowa.

Baldauf-Quilliatre, H., & Carvajal, I. C. de. (2016). Is the avatar considered as a participant by the players? A conversational analysis of multi-player videogames interactions. *PsychNology Journal*, *13*(2-3), 127–148.

Bauman, R., & Briggs, C. L. (1990). Poetics and Performances as Critical Perspectives on Language and Social Life. *Annual Review of Anthropology*, *19*(1), 59–88. https://doi.org/10.1146/annurev.an.19.100190.000423

Bax, T. (2016). "Internet Gaming Disorder" in China: Biomedical Sickness or Sociological Badness? *Games and Culture*, *11*(3), 233–255. https://doi.org/10.1177/1555412014568188

Baym, N. K. (2010). *Personal connections in the digital age*. *Digital media and society*. Cambridge, UK: Polity.

Baym, N. K. (2012). Fans or friends? Seeing social media audiences as musicians do. *Participations - Journal of Audience & Reception Studies*, *9*(2), 286–316.

Baym, N. K., & Ledbetter, A. (2009). TUNES THAT BIND? *Information, Communication & Society*, *12*(3), 408–427. https://doi.org/10.1080/13691180802635430

Bean, A., & Groth-Marnat, G. (2014). Video Gamers and Personality: A Five-Factor Model to Understand Game Playing Style. *Psychology of Popular Media Culture.* Advance online publication. https://doi.org/10.1037/ppm0000025

Beißwenger, M. (2003). Sprachhandlungskoordination im Chat. *Zeitschrift Für Germanistische Linguistik*, *31*(2), 198–231.

Beißwenger, M. (2008). Situated Chat Analysis as a Window to the User's Perspective. *Language@Internet*, *5*(6). Retrieved from http://nbn-resolving.de/urn:nbn:de:0009-7-15329

Bekkers, R., & Wiepking, P. (2011). A Literature Review of Empirical Studies of Philanthropy: Eight Mechanisms That Drive Charitable Giving. *Nonprofit and Voluntary Sector Quarterly*, *40*(5), 924–973. https://doi.org/10.1177/0899764010380927

Bell, A. (1984). Language style as audience design. *Language in Society*, *13*(02), 145. https://doi.org/10.1017/S004740450001037X

Bennerstedt, U. (2008). Sheeping, sapping and avatars-in-action. An in-screen perspective on online gameplay. In *Proceedings - the [Player] Conference* (pp. 28–52). Copenhagen: IT University of Copenhagen.

Bennerstedt, U., & Ivarsson, J. (2010). Knowing the Way. Managing Epistemic Topologies in Virtual Game Worlds. *Computer Supported Cooperative Work*, *19*(2), 201–230. https://doi.org/10.1007/s10606-010-9109-8

Bezemer, J., & Mavers, D. (2011). Multimodal transcription as academic practice: a social semiotic perspective. *International Journal of Social Research Methodology*, *14*(3), 191–206. https://doi.org/10.1080/13645579.2011.563616

Björk, S., & Holopainen, J. (2003). Game Design Patterns. In *DiGRA '03 - Proceedings of the 2003 DiGRA International Conference.* Digital Games Research Association.

Bogost, I. (2010). *Persuasive games*. Cambridge, Mass.: MIT Press Ltd.

Bolander, B., & Locher, M. A. (2014). Doing sociolinguistic research on computer-mediated data: A review of four methodological issues. *Discourse, Context & Media*, *3*, 14–26. https://doi.org/10.1016/j.dcm.2013.10.004

Borrowy, M. (2012). Public Gaming: eSport and Event Marketing in the Experience Economy (Bachelor Thesis). University of British Columbia.

Bou-Franch, P., & Garcés-Conejos Blitvich, P. (2014). Conflict management in massive polylogues: A case study from YouTube. *Journal of Pragmatics*, *73*, 19–36. https://doi.org/10.1016/j.pragma.2014.05.001

Bourdieu, P. (1986). Forms of Capital. In J. G. Richardson (Ed.), *Handbook of theory and research for the sociology of education* (pp. 241–258). New York: Greenwood Press.

Bowman, N. D., Weber, R., Tamborini, R., & Sherry, J. (2013). Facilitating Game Play: How Others Affect Performance at and Enjoyment of Video Games. *Media Psychology*, *16*(1), 39–64. https://doi.org/10.1080/15213269.2012.742360

boyd, d. m., & Ellison, N. B. (2007). Social Network Sites: Definition, History, and Scholarship. *Journal of Computer-Mediated Communication*, *13*(1), 210–230. https://doi.org/10.1111/j.1083-6101.2007.00393.x

Boyd, M. S. (2014). (New) participatory framework on YouTube? Commenter interaction in US political speeches. *Journal of Pragmatics*, *72*, 46–58. https://doi.org/10.1016/j.pragma.2014.03.002

Brus, A. (2012). A young people's perspective on computer game addiction. *Addiction Research & Theory*, *21*(5), 365–375. https://doi.org/10.3109/16066359.2012.733466

Bucholtz, M. (2000). The politics of transcription. *Journal of Pragmatics*, *32*, 1439–1465.

Bucholtz, M. (2007). Variation in transcription. *Discourse Studies*, *9*(6), 784–808. https://doi.org/10.1177/1461445607082580

Buckingham, D., & Burn, A. (2007). Game Literacy in Theory and Practice. *Journal of Educational Multimedia and Hypermedia*, *16*(3), 323–349.

Burton, D. (1980). *Dialogue and Discourse: A sociolinguistic approach to modern drama dialogue and naturally occurring conversation*: Routledge.

Caillois, R. (1961). *Man, play and games*. Urbana, Ill.: University of Illinois Press.

Calleja, G. (2007). Digital Game Involvement: A Conceptual Model. *Games and Culture*, *2*(3), 236–260. https://doi.org/10.1177/1555412007306206

Calleja, G. (2011). *In-game: From immersion to incorporation*. Cambridge, Mass.: MIT Press.

Caplar, N., Suznjevic, M., & Matijasevic, M. (2013). Analysis of player's in-game performance vs rating: Case study of Heroes of Newerth. *arXiv preprint arXiv:1305.5189*.

Carter, M., Gibbs, M., & Arnold, M. (2015). The Demarcation Problem in Multiplayer games: Boundary-Work in EVE online's esport. *Game Studies - the International Journal of Computer Game Research*, *15*(1). Retrieved from http://gamestudies.org/1501/articles/carter

Carter, M., Downs, J., Nansen, B., Harrop, M., & Gibbs, M. (2014). Paradigms of games research in HCI. In *CHI PLAY '14 Proceedings of the first ACM SIGCHI annual symposium on Computer-human interaction in play* (pp. 27–36). New York: ACM. https://doi.org/10.1145/2658537.2658708

Castaño Díaz, C. M., Dorner, B., Hussmann, H., & Strijbos, J.-W. (2015). Scientific Heroes. In *CHI PLAY '15 - The ACM SIGCHI Annual Symposium on Computer-Human Interaction in Play* (pp. 481–485). New York: ACM. https://doi.org/10.1145/2793107.2810313

Cherny, L. (1995). The situated behavior of MUD back channels. In *Proceedings of the AAAI Spring Symposium.*

Cheung, G. (2017). Commentary in: Chat Speed OP PogChamp: Practices of Coherence in Massive Twitch Chat. In *CHI '17: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, (p. 971). New York, NY: ACM.

Cheung, G., & Huang, J. (2011). Starcraft from the stands. In *CHI '11 - Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 763–772). New York: ACM. https://doi.org/10.1145/1978942.1979053

Chomsky, N. (1965). *Aspects of the Theory of Syntax*: MIT Press.

Clark, H. H., & Carlson, T. B. (1982). Hearers and speech acts. *Language*, 332–373.

Consalvo, Mia (2005). Gaining Advantage: How Videogame Players Define and Negotiate Cheating. In *DiGRA '05 - Proceedings of the 2005 DiGRA International Conference.* Digital Games Research Association.

Consalvo, Mia. (2009). Lag, Language, and Lingo: Theorizing Noise in Online Game Spaces. In Wolf, Mark J. P & B. Perron (Eds.), *The video game theory reader 2* (pp. 295–312). New York, NY: Routledge.

Consalvo, Mia, Mitgutsch, K., & Stein, A. (Eds.). (2013). *Sports videogames*. New York: Routledge.

Conway, S. (2013). Argh! An exploration of the response cries of digital game players. *Journal of Gaming & Virtual Worlds*, *5*(2), 131–146.

Couldry, N., & Hepp, A. (2013). Conceptualizing Mediatization: Contexts, Traditions, Arguments. *Communication Theory*, *23*(3), 191–202. https://doi.org/10.1111/comt.12019

Coulthard, M., & Montgomery, M. (1981). *Studies in discourse analysis*. London: Routledge & Kegan Paul.

Coulthard, M. (Ed.). (1992). *Advances in spoken discourse analysis*. London: Routledge.

Coulthard, M., & Brazil, D. (1992). Exchange Structure. In M. Coulthard (Ed.), *Advances in spoken discourse analysis* (pp. 50–78). London: Routledge.

Couper-Kuhlen, E. (1999). Coherent Voicing: On Prosody in Conversational Reported Speech. In W. Bublitz, U. Lenk, & E. Ventola (Eds.), *Coherence in spoken and written discourse: How to create it and how to describe it* (pp. 12–33). Philadelphia: John Benjamins.

Coupland, J., & Jaworski, A. (2003). Transgression and Intimacy in Recreational Talk Narratives. *Research on Language and Social Interaction*, *36*(1), 85–106. https://doi.org/10.1207/S15327973RLSI3601_5

Crawford, G. (2015). Is it in the Game? Reconsidering Play Spaces, Game Definitions, Theming, and Sports Videogames. *Games and Culture*, *10*(6), 571–592. https://doi.org/10.1177/1555412014566235

Crawford, G. (2005). Digital Gaming, Sport and Gender. *Leisure Studies*, *24*(3), 259–270. https://doi.org/10.1080/0261436042000290317

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, *24*(4), 349–354. https://doi.org/10.1037/h0047358

Crystal, D. (2005). The scope of Internet linguistics. Retrieved from www.davidcrystal.com/?fileid=-4113

Crystal, D. (2008). Texting. *ELT Journal*, *62*(1), 77–83. https://doi.org/10.1093/elt/ccm080

Crystal, D. (2011). *Internet linguistics: A student guide*. New York: Routledge.

Davies, M. (2009). Examining Game Pace: How Single-Player Levels Tick. *Gamasutra*. Retrieved from http://www.gamasutra.com/view/feature/132415/examining_game_pace_how_.php

Dayter, D. (2016). *Discursive self in microblogging: Speech acts, stories and self-praise*: John Benjamins.

Debord, G. (1967). *Society of the spectacle*. Detroit: Black & Red.

Deng, J., Cuadrado, F., Tyson, G., & Uhlig, S. Behind the game: Exploring the twitch streaming platform. In *2015 International Workshop on Network and Systems Support for Games (NetGames)* (pp. 1–6). https://doi.org/10.1109/NetGames.2015.7382994

Doherty, S. M., Liskey, D., Via, C. M., Frederick, C., Kring, J., & Liu, D. (2014). An Analysis of Expressed Cheating Behaviors in Video Games. In *Proceedings of the Human Factors and Ergonomics Society 58th Annual Meeting*.

Dressler, R. A., & Kreuz, R. J. (2000). Transcribing Oral Discourse: A Survey and a Model System. *Discourse Processes*, *29*(1), 25–36. https://doi.org/10.1207/S15326950dp2901_2

Du Bois, J. W. (2007). The stance triangle. In R. Englebretson (Ed.), *Stancetaking in Discourse: Subjectivity, evaluation, interaction* (pp. 139–182). Amsterdam: John Benjamins. https://doi.org/10.1075/pbns.164.07du

Du Bois, J. W., & Giora, R. (2014). From cognitive-functional linguistics to dialogic syntax. *Cognitive Linguistics*, *25*(3). https://doi.org/10.1515/cog-2014-0023

Du Bois, J. W., Schuetze-Coburn, S., Cumming, S., & Paolino, D. (1993). Outline of Discourse Transcription. In J. A. Edwards & M. D. Lampert (Eds.), *Talking data: Transcription and coding in discourse research* (pp. 45–90). Hillsdale: Lawrence Erlbaum.

Ducheneaut, N., Yee, N., Nickell, E., & Moore, R. J. (2006). "Alone together?". In *CHI '06 - Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 407–416). New York: ACM. https://doi.org/10.1145/1124772.1124834

Duranti, A. (2006). Transcripts, like shadows on a wall. *Mind, Culture, and Activity*, *13*(4), 301–310.

Dynel, M. (2014a). On the part of ratified participants: ratified listeners in multi-party interactions. *Brno Studies in English*, *40*(1), 27–44. https://doi.org/10.5817/BSE2014-1-2

Dynel, M. (2014b). Participation framework underlying YouTube interaction. *Journal of Pragmatics*, *73*, 37–52. https://doi.org/10.1016/j.pragma.2014.04.001

Edelsky, C. (1981). Who's got the floor? *Language in Society*, *10*(03), 383. https://doi.org/10.1017/S004740450000885X

Edwards, J. A., & Lampert, M. D. (Eds.). (1993). *Talking data: Transcription and coding in discourse research*. Hillsdale: Lawrence Erlbaum.

Eggins, S., & Slade, D. (1997). *Analysing casual conversation*. Oakville: Equinox.

Elverdam, C., & Aarseth, E. (2007). Game Classification and Game Design: Construction Through Critical Analysis. *Games and Culture*, *2*(1), 3–22. https://doi.org/10.1177/1555412006286892

Englebretson, R. (Ed.). (2007). *Stancetaking in Discourse: Subjectivity, evaluation, interaction*. Amsterdam: John Benjamins.

Ensslin, A. (2012). *The language of gaming*. Houndmills, Basingstoke, Hampshire, New York: Palgrave Macmillan.

Eskelinen, M. (2001). The Gaming Situation. *Game Studies - The International Journal of Computer Game Research*. (1). Retrieved from http://www.gamestudies.org/0101/eskelinen/

Entertainment software association. (2015). *Essential facts: about the computer and video game industry*. Retrieved from http://www.theesa.com/wp-content/uploads/2015/04/ESA-Essential-Facts-2015.pdf

Fairclough, N. (1989). *Language and Power*. London: Longman.

Fernández-Vara, C. (2009). Play's the Thing: A Framework to Study Videogames as Performance. In *DiGRA '09 - Proceedings of the 2009 DiGRA International Conference* (pp. 1–9). Digital Games Research Association.

Ford, C., Gardner, D., Horgan, L. E., Liu, C., tsaasan, a. m., Nardi, B., & Rickman, J. (2017). Chat Speed OP PogChamp: Practices of Coherence in Massive Twitch Chat. In *CHI '17: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, (pp. 858–871). New York, NY: ACM. https://doi.org/10.1145/3027063.3052765

Francis, G., & Hunston, S. (1992). Analysing everyday conversation. In M. Coulthard (Ed.), *Advances in spoken discourse analysis* (pp. 123–162). London: Routledge.

Frasca, G. (2001). The Sims: Grandmothers are cooler than trolls. *Game Studies - the International Journal of Computer Game Research*, *1*(1). Retrieved from http://www.gamestudies.org/0101/frasca/

Frobenius, M. (2014). Audience design in monologues: How vloggers involve their viewers. *Journal of Pragmatics*, *72*, 59–72. https://doi.org/10.1016/j.pragma.2014.02.008

Gajadhar, B. J., de Kort, Y. A.W., & IJsselsteijn, W. A. (2009). Rules of Engagement. *International Journal of Gaming and Computer-Mediated Simulations*, *1*(3), 14–27. https://doi.org/10.4018/jgcms.2009070102

Gandolfi, E. (2016). To watch or to play, it is in the game: The game culture on Twitch.tv among performers, plays and audiences. *Journal of Gaming & Virtual Worlds*, *8*(1).

Garcia, A. C., & Baker Jacobs, J. (2010). The Eyes of the Beholder: Understanding the Turn-Taking System in Quasi-Synchronous Computer-Mediated Communication. *Research on Language and Social Interaction*, *32*(4), 337–367. https://doi.org/10.1207/S15327973rls3204_2

Garfinkel, H. (1967). *Studies in ethnomethodology*. Englewood Cliffs: Prentice-Hall.

Gee, J. P. (Ed.). (2007). *Good video games + good learning: Collected essays on video games, learning, and literacy*. New York: Peter Lang.

Georgen, C., Duncon, S. C., & Cook, L. (2015). From lurking to participatory spectatorship: Understanding Affordances of the Dota 2 Noob Stream. In *CSCL 2015 - The 11th International Conference on Computer Supported Collaborative Learning.* Gothenburg: The International Society of the Learning Sciences.

Gerhardt, C. (2008). Turn-by-turn and move-by-move: A multimodal analysis of English live television football commentary. In E. Lavric, G. Pisek, A. C. Skinner, & W. Stadler (Eds.), *The linguistics of football* (pp. 288–299). Tübingen: Narr.

Gerhardt, C. (2012). Notability: The construction of current events in talk-in-interaction. In R. Ayaß & C. Gerhardt (Eds.), *The appropriation of media in everyday life* (pp. 47–77). Amsterdam, Philadelphia: John Benjamins.

Gibson, J. J. (1979). *The ecological approach to visual perception*. New York: Psychology Press.

Giles, D., Stommel, W., Paulus, T., Lester, J., & Reed, D. (2015). Microanalysis Of Online Data: The methodological development of "digital CA". *Discourse, Context & Media*, *7*, 45–51. https://doi.org/10.1016/j.dcm.2014.12.002

Giles, D., Stommel, W., & Paulus, T. M. (2017). The Microanalysis of Online Data: The next stage. *Journal of Pragmatics*, *115*, 37–41. https://doi.org/10.1016/j.pragma.2017.02.007

Glas, R. (2015). Vicarious play: Engaging the viewer in Let's Play videos. *Empedocles: European Journal for the Philosophy of Communication*, *5*(1), 81–86. https://doi.org/10.1386/ejpc.5.1-2.81_1

Glass, R. L. (2009). Frivolous Research. *Information Systems Management*, *26*(2), 209–210. https://doi.org/10.1080/10580530902797631

Glazier, J., & Glazier, P. G. (1976). Ambiguity and Exchange: The Double Dimension of Mbeere Riddles. *The Journal of American Folklore*, *89*(352), 189. https://doi.org/10.2307/539689

Goffman, E. (1967). *Interaction ritual: Essays in face-to-face behavior*. New Brunswick, N.J.: Aldine Transaction.

Goffman, E. (1974). *Frame analysis: An essay on the organization of experience*. Boston: Northeastern Univ. Press.

Goffman, E. (1978). Response Cries. *Language*, *54*(4), 787–815.

Goffman, E. (1981). *Forms of talk*. *University of Pennsylvania publications in conduct and communication*. Philadelphia: University of Pennsylvania Press.

Goldhaber, M. H. (1997). The Attention and the Net. *First Monday*, *2*(7). Retrieved from http://firstmonday.org/article/view/519/440

Gómez Maureira, M. A., & Verbeek, F. (2016). The Impact of Co-Located Play on Social Presence and Game Experience in a VR Game. In *Proceedings of the 1st International Joint Conference of DiGRA and FDG.* ACM.

Goodwin, C., & Duranti, A. (1992). Rethinking Context - An Introduction. In A. Duranti & C. Goodwin (Eds.), *Rethinking context: Language as an interactive phenomenon* (pp. 1–42). Cambridge: Cambridge University Press.

Goodwin, C. (1994). Professional Vision. *American Anthropologist*, *96*(3), 606–633.

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

Goodwin, C., & Heritage, J. (1990). Conversation Analysis. *Annual Review of Anthropology*, *19*(1), 283–307. https://di.org/10.1146/annurev.an.19.100190.001435

GÖRAL, K. (2015). Passing Success Percentages and Ball Possession Rates of Successful Teams in 2014 FIFA World Cup. *International journal of Science Culture and Sport*, *3*(9), 86. https://doi.org/10.14486/IJSCS239

Granovetter, M. (1983). The Strength of Weak Ties: A Network Theory Revisited. *Sociological Theory*, *1*, 201. https://doi.org/10.2307/202051

Grint, K., & Woolgar, S. (1997). *The machine at work: Technology, work and organization / Keith Grint and Steve Woolgar*. Cambridge: Polity Press.

Gros, D., Wanner, B., Hackenholt, A., Zawadzki, P., & Knautz, K. (2017). World of Streaming. Motivation and Gratification on Twitch. In *SCSM '17 - Proceedings of the 9th International Conference on Social Computing and Social Media* (pp. 44–57). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-58559-8_5

Gumperz, J. J., & Hymes, D. (Eds.). (1972). *Directions in sociolinguistics: The ethnography of communication*. Oxford: Blackwell.

Hall, J. A. (2011). Sex differences in friendship expectations: A meta-analysis. *Journal of Social and Personal Relationships*, *28*(6), 723–747. https://doi.org/10.1177/0265407510386192

Halliday, M. (1984). Language as code and language as behaviour: a systemic-functional interpretation of the nature and ontogenesis of dialogue. In R. P. Fawcett (Ed.), *Open linguistics series. The semiotics of culture and language* (pp. 3–35). London: Pinter.

Halliday, M. A. K, & Hasan, R. (1976). *Cohesion in English*. Hoboken: Taylor and Francis.

Hamari, J., & Sjöblom, M. (2017). What is eSports and why do people watch it. *Internet Research*, *27*(2). Retrieved from http://ssrn.com/abstract=2686182

Hamilton, W. A., Garretson, O., & Kerne, A. (2014). Streaming on Twitch. In *CHI' 14 - Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1315–1324). New York: ACM. https://doi.org/10.1145/2556288.2557048

Harries, L. (1971). The Riddle in Africa. *The Journal of American Folklore*, *84*(334), 377. https://doi.org/10.2307/539632

Heath, C., Hindmarsh, J., & Luff, P. (2010). *Video in qualitative research: Analysing social interaction in everyday life*. *Introducing qualitative methods*. Los Angeles: Sage.

Hemphill, D. (2005). Cybersport. *Journal of the Philosophy of Sport*, *32*(2), 195–207. https://doi.org/10.1080/00948705.2005.9714682

Henricks, T. S. (2011). Caillois's Man, Play, and Games: An Appreciation and Evaluation. *American Journal of Play*, *3*(2). Retrieved from http://www.journalofplay.org/sites/www.journalofplay.org/files/pdf-articles/3-2-article-cailloiss-man-play-and-games.pdf

Heritage, J. (2012). Epistemics in Action: Action Formation and Territories of Knowledge. *Research on Language & Social Interaction*, *45*(1), 1–29. https://doi.org/10.1080/08351813.2012.646684

Herring, C. S. (2010). Who's got the floor in computer-mediated conversation? Edelsky's gender patterns revisited. *Language@Internet*, *7*(8). Retrieved from http://nbn-resolving.de/urn:nbn:de:0009-7-28579

Herring, S. C. (1999). Interactional Coherence in CMC. *Journal of Computer-Mediated Communication*, *4*(4), 0. https://doi.org/10.1111/j.1083-6101.1999.tb00106.x

Heuer, C. (2017). Der Einfluss von Echtzeitinteraktivität auf die Kundenbindung: Eine empirische Analyse der Online Plattform Twitch (Bachelor's thesis). Hochschule für Wirtschaft, Technik und Kultur, Stuttgart.

Hitchens, M., & Tychsen, A. (2009). Game Time: Modeling and Analyzing Time in Multiplayer and Massively Multiplayer Games. *Games and Culture*, *4*(2), 170–201. https://doi.org/10.1177/1555412008325479

Holt, E. (1996). Reporting on Talk: The Use of Direct Reported Speech in Conversation. *Research on Language & Social Interaction*, *29*(3), 219–245. https://doi.org/10.1207/s15327973rlsi2903_2

Hornyak, C. (2016). NIGHT EYES, JARGON, AND YOU: THE LANGUAGE OF POWER WITHIN WORLD OF WARCRAFT. *Well Played*, *5*(1), 46–63.

Huizinga, J. (1938). *Homo ludens: A study of the play-element in culture*. Kettering, OH: Angelico Press.

Hung, A. C.-Y. (2009). The Order of Play: Seeing, Teaching, and Learning Meaning in Video Games. In *DiGRA '09 - Proceedings of the 2009 DiGRA International Conference.* Digital Games Research Association.

Hutchby, I. (2014). Communicative affordances and participation frameworks in mediated interaction. *Journal of Pragmatics*, *72*, 86–89. https://doi.org/10.1016/j.pragma.2014.08.012

Hutchby, I. (2001). Technologies, Texts and Affordances. *Sociology*, *35*(2), 441–456.

Hutchins, B. (2008). Signs of meta-change in second modernity: the growth of e-sport and the World Cyber Games. *New Media & Society*, *10*(6), 851–869. https://doi.org/10.1177/1461444808096248

Jarret, J. (2016). Critically Approaching the Playful and Participatory Genealogy of MOBAs. In *1st International Joint Conference of DiGRA and FDG.* ACM.

Järvinen, A. (2003). Making and Breaking a Game: A typology of Rules. In *DiGRA '03 - Proceedings of the 2003 DiGRA International Conference* (pp. 68–79). Digital Games Research Association.

Jefferson, G. (1972). Side Sequences. In D. N. Sudnow (Ed.), *Studies in Social Interaction* (pp. 294–333). New York: New York Free Press.

Jefferson, G. (2004). Glossary of transcript symbols with an introduction. In G. H. Lerner (Ed.), *Conversation Analysis: Studies from the First Generation* (pp. 13–31). Amsterdam: John Benjamins.

Jenkins, H. (2005). Welcome to convergence culture. *Receiver*, *12*, 1–6.

Jenkins, H. (2006). *Convergence culture: Where old and new media collide*. New York: New York University Press.

Jin, D. Y. (2010). *Korea's online gaming empire*. Cambridge, Mass.: MIT Press.

Jin, D. Y., & Chee, F. (2008). Age of New Media Empires: A Critical Interpretation of the Korean Online Game Industry. *Games and Culture*, *3*(1), 38–58. https://doi.org/10.1177/1555412007309528

Jones, R. H. (2004). The Problem of Context in Computer Mediated Communication. In P. LeVine & R. Scollon (Eds.), *Discourse and technology: Multimodal discourse analysis* (pp. 20–33). Washington, D.C.: Georgetown University Press.

Jones, Q., Ravid, G., & Rafaeli, S. (2004). Information Overload and the Message Dynamics of Online Interaction Spaces: A Theoretical Model and Empirical Exploration. *Information Systems Research*, *15*(2), 194–210. https://doi.org/10.1287/isre.1040.0023

Jucker, A. H., & Dürscheid, C. (2012). The Linguistics of Keyboard to screen Communication. *Linguistik Online*, *56*.

Juul, J. (2002). The Open and the Closed: Games of Emergence and Games of Progression. In *Proceedings of the Computer Games and Digital Cultures Conference* (pp. 323–329). Tampere: Tampere University Press.

Juul, J. (2003). The Game, The Player, The World: Looking for a Heart of Gameness. In *DiGRA '03 - Proceedings of the 2003 DiGRA International Conference.* Digital Games Research Association. Retrieved from http://www.jesperjuul.net/text/gameplayerworld/

Juul, J. (2004). Introduction to Game Time / Time to Play: an examination of Game temporality. In N. Wardrip-Fruin & P. Harrigan (Eds.), *First person: New media as story, performance, and game* (pp. 131–140). Cambridge, Mass.: MIT Press.

Juul, J. (2007). Without a goal. In B. Atkins & T. Krzywinska (Eds.), *Videogame, player, text* (pp. 510–515). New York: Palgrave.

Juul, J. (2010). In search of lost time. In *FDG '10 - Proceedings of the Fifth International Conference on the Foundations of Digital Games* (pp. 86–91). New York: ACM. https://doi.org/10.1145/1822348.1822360

Kappen, D. L., Mirza-Babaei, P., Johannsmeier, J., Buckstein, D., Robb, J., & Nacke, L. E. (2014). Engaged by boos and cheers: the effect of co-located game audiences on social player experience. In *CHI PLAY '14 Proceedings of the first ACM SIGCHI annual symposium on Computer-human interaction in play* (pp. 151–160). New York: ACM. https://doi.org/10.1145/2658537.2658687

Karhulahti, V.-M. (2016). Prank, Troll, Gross and Gore: Performance Issues in Esport Live-Streaming. In *1st International Joint Conference of DiGRA and FDG.* ACM.

Kaytoue, M., Arlei, S., Cerf, L., Meira Jr., W., & Raissi, C. (2012). Watch me playing, I am a Professional: a First Study on Video Game Live Streaming. In *WWW 2012 - Proceedings of the 21st international conference on World Wide Web*. New York: ACM.

Kendon, A. (1992). the negotiation of context in face to face interaction. In A. Duranti & C. Goodwin (Eds.), *Rethinking context: Language as an interactive phenomenon* (pp. 326–332). Cambridge: Cambridge University Press.

Kerbrat-Orecchioni, C. (1990). *Les interactions verbales*. Paris: Armand Colin.

Kerbrat-Orecchioni, C. (2004). Introducing polylogue. *Journal of Pragmatics*, *36*(1), 1–24. https://doi.org/10.1016/S0378-2166(03)00034-1

Kivikangas, J. M., Chanel, G., Cowley, B., Ekman, I., Salminen, M., Järvelä, S., & Ravaja, N. (2011). A review of the use of psychophysiological methods in game research. *Journal of Gaming & Virtual Worlds*, *3*(3), 181–199. https://doi.org/10.1386/jgvw.3.3.181_1

Knight, N. K. (2013). Evaluating experience in funny ways: How friends bond through conversational hum. *Text & Talk*, *33*(4-5). https://doi.org/10.1515/text-2013-0025

Kou, Y., & Gui, X. (2014). Playing with strangers: understanding temporary teams in league of legends. In *CHI PLAY '14 Proceedings of the first ACM SIGCHI annual symposium on Computer-human interaction in play* (pp. 161–169). New York: ACM. https://doi.org/10.1145/2658537.2658538

Kow, Y. M., & Young, T. (2013). Media technologies and learning in the starcraft esport community. In *CSCW 2013 - Proceedings of the 2013 conference on Computer supported cooperative work* (pp. 387–399). New York: ACM. https://doi.org/10.1145/2441776.2441821

Kress, G. (2005). Gains and losses: New forms of texts, knowledge, and learning. *Computers and Composition*, *22*(1), 5–22. https://doi.org/10.1016/j.compcom.2004.12.004

Kress, G. (2009). What is Mode. In C. Jewitt (Ed.), *The Routledge handbook of multimodal analysis* (pp. 54–67). New York: Routledge.

Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication*. London: Routledge.

Labov, W., & Waletzky, J. (1967). Narrative Analysis: Oral Versions of Personal Experience. In J. Helm (Ed.), *Essays on the Verbal and Visual Arts: Proceedings of the 1966 annual spring meeting* (pp. 12–44). Seattle: University of Washington Press.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.

Lange, P. G. (2011). Learning Real-Life Lessons From Online Games. *Games and Culture*, *6*(1), 17–37. https://doi.org/10.1177/1555412010377320

Lankoski, P. (2011). Player Character Engagement in Computer Games. *Games and Culture*, *6*(4), 291–311. https://doi.org/10.1177/1555412010391088

Lankoski, P., & Björk, S. (2015). Formal Analysis of Gameplay. In P. Lankoski & S. Björk (Eds.), *Game research methods: An overview* (pp. 23–34). Pittsburgh: ETC Press.

Levinson, S. (1988). Putting linguistics on a proper footing: Explorations in Goffman's participation framework. In P. Drew & A. J. Wootton (Eds.), *Erving Goffman: Exploring the interaction order* (pp. 161–227). Cambridge: Polity.

Levinson, P., Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage*: Cambridge University Press.

Lewis, J. M., Trinh, P., & Kirsch, D. (2011). A Corpus Analysis of Strategy Video Game Play in Starcraft: Brood War. In *COGSCI 2011 - Expanding the Space of Cognitive Science. Proceedings of the 33rd Annual Meeting of the Cognitive Science Society* (pp. 687–693). Austin: Cognitive Science Society.

Licoppe, C. (2017). Showing objects in Skype video-mediated conversations: From showing gestures to showing sequences. *Journal of Pragmatics*, *110*, 63–82. https://doi.org/10.1016/j.pragma.2017.01.007

Lin, H., & Sun, C.-T. (2011). The Role of Onlookers in Arcade Gaming: Frame Analysis of Public Behaviours. *Convergence: the International Journal of Research into New Media Technologies*, *17*(2), 125–137. https://doi.org/10.1177/1354856510397111

Livingstone, S. (2009). On the Mediation of Everything: ICA Presidential Address 2008. *Journal of Communication*, *59*(1), 1–18. https://doi.org/10.1111/j.1460-2466.2008.01401.x

Loftus, G. R., & Loftus, E. F. (1983). *Mind at play: The psychology of video games*. New York: Basic Books.

Long, M. H. (1983). Native speaker/non-native speaker conversation and the negotiation of comprehensible input1. *Applied Linguistics*, *4*(2), 126–141. https://doi.org/10.1093/applin/4.2.126

Lowood, H. (2006). High-performance play: The making of machinima. *Journal of Media Practice*, *7*(1), 25–42. https://doi.org/10.1386/jmpr.7.1.25/1

Knobel, M., & Lankshear, C. (2007). Online memes, affinities, and cultural production. *A New Literacies Sampler*, *29*, 199–227.

Kücklich, J. (2007). Homo Deludens. *Convergence: the International Journal of Research into New Media Technologies*, *13*(4), 355–367. https://doi.org/10.1177/1354856507081951

Malinowski, B. (1923). The Problem of Meaning in Primitive Languages. In C. K. Ogden & I. A. Richards (Eds.), *The Meaning of Meaning* (pp. 296–336). London: Paul, Trend Trubner.

Mason, M. (2010). Sample Size and Saturation in PhD Studies Using Qualitative Interviews. *Forum: Qualitative Social Research*, *11*(3).

Matthiessen, C.M. (2014). Language use in a social semiotic perspective. In A. Barron, P. Grundy, & Y. Gu (Eds.), *The Routledge Handbook of Pragmatics.* London: Routledge.

Matthiessen, C. M., & Slade, D. (2011). Analysing Conversation. In R. Wodak (Ed.), *The SAGE Handbook of Sociolinguistics* (pp. 375–395). London: Sage. https://doi.org/10.4135/9781446200957.n27

Maurer, B., Aslan, I., Wuchse, M., Neureiter, K., & Tscheligi, M. (2015). Gaze-Based Onlooker Integration: Exploring the In-Between of Active Player and Passive Spectator in Co-Located Gaming. In *CHI PLAY '15 - The ACM SIGCHI Annual Symposium on Computer-Human Interaction in Play* (pp. 163–173). New York: ACM.

Mäyrä, F., Van Loy, J., & Quandt, T. (2013). Disciplinary Identity of Game Scholars: An Outline. In *DiGRA '13 - Proceedings of the 2013 DiGRA International Conference.* Digital Games Research Association.

McLuhan, M. (1964). *Understanding media: The extensions of man*. Cambridge, Mass.: MIT Press.

McMahan, A. (2003). Immersion, Engagement, and Presence: A Method for Analyzing 3-D Video Games. In M. J. P. Wolf & B. Perron (Eds.), *The video game theory reader* (pp. 67–87). New York: Routledge.

McKernan, B. (2013). The Morality of Play: Video Game Coverage in The New York Times From 1980 to 2010. *Games and Culture*, *8*(5), 307–329. https://doi.org/10.1177/1555412013493133

Melcer, E., Nguyen, T.-H. D., Chen, Z., Canossa, A., El-Nasr, M. S., & Isbister, K. (2015). Games research today: Analyzing the academic landscape 2000-2014. *Network*, *17*, 20.

Meredith, J. (2015). Transcribing screen-capture data: The process of developing a transcription system for multi-modal text-based data. *International Journal of Social Research Methodology*, 1–14. https://doi.org/10.1080/13645579.2015.1082291

Meyer, N. (forthcoming). Massively Multimodal Communication and Space: A Case Study of Video Game Livestreaming (Dissertation). University of Zurich, Zurich.

Miller, T. (2006). Gaming for Beginners. *Games and Culture*, *1*(1), 5–12. https://doi.org/10.1177/1555412005281403

Mitroff, I. I. (1972). The Myth of Objectivity OR Why Science Needs a New Psychology of Science. *Management Science*, *18*(10), B-613-B-618. https://doi.org/10.1287/mnsc.18.10.B613

Moeller, R. M., Esplin, B., & Conway, S. (2009). Cheesers, Pullers and Glitchers: The Rhetoric of Sportsmanship and the Discourse of Online Sports Gamers. *Game Studies - The International Journal of Computer Game Research*, *9*(2). Retrieved from http://gamestudies.org/0902/articles/moeller_esplin_conway

Mondada, L. (2007). Commentary: Transcript variations and the indexicality of transcribing practices. *Discourse Studies*, *9*(6), 809–821. https://doi.org/10.1177/1461445607082581

Mondada, L. (2012). Coordinating action and talk-in-interaction in and out of video games. In R. Ayaß & C. Gerhardt (Eds.), *The appropriation of media in everyday life* (pp. 231–270). Amsterdam, Philadelphia: John Benjamins.

Mondada, L. (2014). Conventions for multimodal transcription. Retrieved from https://franz.unibas.ch/fileadmin/franz/user_upload/redaktion/Mondada_conv_multimodality.pdf

Möring, S., & Leino, O. (2016). Beyond games as political education – neo-liberalism in the contemporary computer game form. *Journal of Gaming & Virtual Worlds*, *8*(2), 145–161. https://doi.org/10.1386/jgvw.8.2.145_1

Murray, J. H. (1998). *Hamlet on the holodeck: The future of narrative in cyberspace*. Cambridge, Mass.: MIT Press.

Neise, I. (forthcoming). Multimodale Analysen audio-visueller Formate in den neuen Medien. Institut für Deutsche Sprache, Mannheim.

Nematzadeh, A., Ciampaglia, G. L., Ahn, Y.-Y., & Flammini, A. (2016). Information Overload in Group Communication: From Conversation to Cacophony in the Twitch Chat. Retrieved from https://arxiv.org/abs/1610.06497

Nevile, M. (2015). The Embodied Turn in Research on Language and Social Interaction. *Research on Language and Social Interaction*, *48*(2), 121–151. https://doi.org/10.1080/08351813.2015.1025499

Newman, J. (2002). The Myth of the Ergodic Videogame: Some thoughts on player character relationships in videogames. *Game Studies - the International Journal of Computer Game Research*, *2*(1).

Nguyen, J. (2016). Performing as video game players in Let's Plays. *Transformative Works and Cultures*, *22.* https://doi.org/10.3983/twc.2016.0698

Nichols, R. (2009). Target Acquired: America's Army and the Video Game Industry. In N. Huntemann & M. T. Payne (Eds.), *Joystick soldiers: The politics of play in military video games* (pp. 39–52). New York: Routledge.

Noam, E. M. (2016). *Who owns the world's media? Media concentration and ownership around the world.*

Norman, D. A. (1988). *The psychology of everyday things*. New York: Basic Books.

Norrick, N. R. (1993). *Conversational joking: Humor in everyday talk*. Bloomington: Indiana University Press.

Norrick, N. R. (2000). *Conversational narrative: Storytelling in everyday talk*. Amsterdam: John Benjamins.

Norris, S. (2012). Multimodal Interaction Analysis. In C. A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics.* Oxford: Blackwell. https://doi.org/10.1002/9781405198431.wbeal0814

Norris, S., & Maier, C. D. (Eds.). (2014). *Texts, Images and Interactions: A Reader in Multimodality*. Berlin: De Gruyter Mouton.

Nylund, N. (2015). Walkthrough and let's play. In M. Turunen (Ed.), *Proceedings of the 19th International Academic Mindtrek Conference on - AcademicMindTrek '15* (pp. 55–62). New York, New York, USA: ACM Press. https://doi.org/10.1145/2818187.2818283

Ochs, E. (1979). Transcription as theory. In *Developmental pragmatics* (pp. 43–72). New York: Academic Press. *Developmental pragmatics.* (1979). New York: Academic Press.

Olejniczak, J. (2015). A Linguistic Study of Language Variety Used on Twitch: Descriptive and Corpus Based Approaches. In *RCIC'15 - 4th International Conference Redefining Community in Intercultural Context.*

O'Reilly, T. (2005). What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software. Retrieved from http://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html

Otten, M. (2001). Broadcasting Virtual Games in the Internet. Retrieved from www.slipgate.de/download/BroadcastingVirtualGames.pdf

Page, R. (2012). The linguistics of self-branding and micro-celebrity in Twitter: The role of hashtags. *Discourse & Communication*, *6*(2), 181–201. https://doi.org/10.1177/1750481312437441

Paolillo, J. C. (2001). Language Variation on Internet Relay Chat: A social network approach. *Journal of Sociolinguistics*, *5*(2), 180–213.

Paolillo, J. C., & Zelenkauskaite, A. (2013). Real-time chat. In S. C. Herring, D. Stein, & T. Virtanen (Eds.), *Pragmatics of computer-mediated communication* (pp. 109–133). Berlin: De Gruyter Mouton.

Parlett, D. (1999). *The Oxford history of board games*. Oxford: Oxford University Press.

Pearce, C. (2006). Productive Play: Game Culture From the Bottom Up. *Games and Culture*, *1*(1), 17–24. https://doi.org/10.1177/1555412005281418

Pellicone, A. (2016). Performing Play: Cultural Production on Twitch.tv. In *CHI '16: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 244–248). New York, NY: ACM. https://doi.org/10.1145/2851581.2859022

Pellicone, A., & Ahn, J. (2017). The Game of Performing Play: Understanding Streaming as Cultural Production. In *CHI '17: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, (pp. 4863–4874). New York, NY: ACM. https://doi.org/10.1145/3025453.3025854

Penner, L. A., Dovidio, J. F., Piliavin, J. A., & Schroeder, D. A. (2005). Prosocial behavior: multilevel perspectives. *Annual review of psychology*, *56*, 365–392. https://doi.org/10.1146/annurev.psych.56.091103.070141

Perez Latorre, O. (2015). The Social Discourse of Video Games Analysis Model and Case Study: GTA IV. *Games and Culture*, *10*(5), 415–

437. https://doi.org/10.1177/1555412014565639

Piirainen-Marsh, A. (2012). Organising participation in video gaming activities. In R. Ayaß & C. Gerhardt (Eds.), *The appropriation of media in everyday life* (pp. 197–230). Amsterdam, Philadelphia: John Benjamins.

Piirainen-Marsh, A., & Tainio, L. (2009). Collaborative Gameplay as a Site for Participation and Situated Learning of a Second Language. *Scandinavian Journal of Educational Research*, *53*(2), 167–183. https://doi.org/10.1080/00313830902757584

Piirainen-Marsh, A., & Tainio, L. (2014). Asymmetries of Knowledge and Epistemic Change in Social Gaming Interaction. *The Modern Language Journal*, *98*(4), 1022–1038. https://doi.org/10.1111/modl.12153

Plath, H.-F. (2015). Marketing in the Field of Gaming Related Online Live-Streaming (Bachelor's Thesis). Europa-Universität Flensburg, Flensburg.

Postigo, H. (2016). The socio-technical architecture of digital labor: Converting play into YouTube money. *New Media & Society*, *18*. https://doi.org/10.1177/1461444814541527

Psathas, G., & ANDERSON, T. (1990). The 'practices' of transcription in Conversation Analysis. *Semiotica*, *78*(1-2). https://doi.org/10.1515/semi.1990.78.1-2.75

Quandt, T., van Looy, J., Vogelgesang, J., Elson, M., Ivory, J. D., Consalvo, Mia, & Mäyrä, F. (2015). Digital Games Research: A Survey Study on an Emerging Field and Its Prevalent Debates. *Journal of Communication.* Advance online publication. https://doi.org/10.1111/jcom.12182

Rambush, J., Jakobsson, P., & Pargman, D. (2007). Exploring E-Sport: A Case Study of Gameplay in Counter-Strike. In *DiGRA '07 - Proceedings of the 2007 DiGRA International Conference.* Digital Games Research Association.

Raes, T. C. M. (2015). Twitch TV: Motives and interaction, a consumer perspective (Master Thesis). Aalborg University, Aalborg. Retrieved from http://projekter.aau.dk/projekter/en/studentthesis/twitch-tv-motives-and-interaction-a-consumer-perspective(d8c55eb0-3aaf-4ee9-aefd-b74e505e8248).html

Ratner, C. (2002). Subjectivity and Objectivity in Qualitative Methodology. *Forum: Qualitative Social Research*, *3*(3). Retrieved from http://nbn-resolving.de/urn:nbn:de:0114-fqs0203160

Ravaja, N., Saari, T., Salminen, M., Laarni, J., & Kallinen, K. (2006). Phasic Emotional Reactions to Video Game Events: A Psychophysiological Investigation. *Media Psychology*, *8*(4), 343–367. https://doi.org/10.1207/s1532785xmep0804_2

Recktenwald, D. (2014). Interactional Practices in Let's Play Videos (Master Thesis). Saarland University, Saarbruecken. Retrieved from https://www.academia.edu/23345153/M.A._Dissertation_Interactional_Practices_in_Lets_Play_Videos

Recktenwald, D. (2017). Toward a transcription and analysis of live streaming on Twitch. *Journal of Pragmatics*, *115*, 68–81. https://doi.org/10.1016/j.pragma.2017.01.013

Recktenwald, D. (2015). *TwitchCon Field Notes*. San Francisco.

Reeves, S., Greiffenhagen, C., Flintham, M., Benford, S., Adams, M., Row Farr, J., & Tandavantij, N. (2015). I'd Hide You. In *CHI '15 - Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 2573–2582). New York: ACM. https://doi.org/10.1145/2702123.2702257

Rintel, S., O'Hara, K., Rostami Yeganeh, B., & Rädle, R. (2015). Ad hoc adaptability in video-calling. In *ITS 2015 - Nature Meets Interactive Surfaces.* New York: ACM.

Rintel, S. E., & Pittam, J. (1997). Strangers in a Strange Land Interaction Management on Internet Relay Chat. *Human Communication Research*, *23*(4), 507–534. https://doi.org/10.1111/j.1468-2958.1997.tb00408.x

Rheingold, H. (1993). *The virtual community: Homesteading on the electronic frontier*. Cambridge: MIT Press.

Rogers, E. M. (1962). *Diffusion of innovations*. New York: Free Press.

Rosenbaun, L., & Licoppe, C. (2017). Showing 'digital' objects in web-based video chats as a collaborative achievement. *Pragmatics*, *27*(3), 419–446. https://doi.org/10.1075/prag.27.3.05ros

Rosenbaun, L., Rafaeli, S., & Kurzon, D. (2016a). Blurring the Boundaries between Domestic and Digital Spheres: Competing Engagements in Public Google Hangouts. *Pragmatics*, *26*(2), 291–314.

Rosenbaun, L., Rafaeli, S., & Kurzon, D. (2016b). Participation frameworks in multiparty video chats cross-modal exchanges in public Google Hangouts. *Journal of Pragmatics*, *94*, 29–46. https://doi.org/10.1016/j.pragma.2016.01.003

Sacks, H. (1984). Notes on methodology. In J. M. Atkinson (Ed.), *Structures of Social Action* (pp. 21–27). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511665868.005

Sacks, H., Schegloff, E. A., & Jeffersen, G. (1974). A Simplest Systematics For the organization of Turn-Taking for Conversation. *Language*, *50*, 696–735.

Saldaña, J. (2009). *The coding manual for qualitative researchers*. Los Angeles: Sage.

Salen, K., & Zimmerman, E. (2003). *Rules of play: Game design fundamentals*. Cambridge, Mass.: MIT Press.

Sandelowski, M. (1997). "To be of use": Enhancing the utility of qualitative research. *Nursing Outlook*, *45*(3), 125–132. https://doi.org/10.1016/S0029-6554(97)90043-9

Schegloff, E. A. (2006). On possibles. *Discourse Studies*, *8*(1), 141–157. https://doi.org/10.1177/1461445606059563

Schegloff, E. A. (1968). Sequencing in Conversational Openings. *American Anthropologist*, *70*(6), 1075–1095. https://doi.org/10.1525/aa.1968.70.6.02a00030

Schegloff, E. A. (1988). Goffman and the Analysis of Conversation. In P. Drew & A. J. Wootton (Eds.), *Erving Goffman: Exploring the interaction order* (pp. 89–135). Cambridge: Polity.

Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language*. London: Cambridge University Press.

Seidlhofer, B. (2005). English as a lingua franca. *ELT Journal*, *59*(4), 339–341. https://doi.org/10.1093/elt/cci064

Selfhout, M. H., Branje, S. J., ter Bogt, T. F., & Meeus, W. H. (2009). The role of music preferences in early adolescents' friendship formation and stability. *Journal of Adolescence*, *32*(1), 95–107. https://doi.org/10.1016/j.adolescence.2007.11.004

Senft, T. M. (2008). *Camgirls: Celebrity and community in the age of social networks*. *Digital formations: v. 4*. New York: Lang.

Seo, Y. (2013). Electronic sports: A new marketing landscape of the experience economy. *Journal of Marketing Management*, *29*(13-14), 1542–1560. https://doi.org/10.1080/0267257X.2013.822906

Seo, Y., & Jung, S.-U. (2014). Beyond solitary play in computer games: The social practices of eSports. *Journal of Consumer Culture*. Advance online publication. https://doi.org/10.1177/1469540514553711

Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, *27*(3), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Shaw, A. (2012). Do you identify as a gamer? Gender, race, sexuality, and gamer identity. *New Media & Society*, *14*(1), 28–44. https://doi.org/10.1177/1461444811410394

Simmel, G., & Hughes, E. C. (1949). The Sociology of Sociability. *American Journal of Sociology*, *55*(3), 254–261.

Simpson, J. (2005). Conversational floors in synchronous text-based CMC discourse. *Discourse Studies*, *7*(3), 337–361. https://doi.org/10.1177/1461445605052190

Sinclair, J. M. (1992). Priorities in discourse analysis. In M. Coulthard (Ed.), *Advances in spoken discourse analysis* (pp. 78–88). London: Routledge.

Sinclair, J. M., & Coulthard, R. M. (1975). *Towards an analysis of discourse: The English used by teachers and pupils*. London: Oxford Univ. Press.

Sindoni, M. G. (2014). Through the looking glass: a social semiotic and linguistic perspective on the study of video chats. *Text & Talk*, *34*(3). https://doi.org/10.1515/text-2014-0006

Sjöblom, M., & Hamari, J. (2016). Why Do People Watch Others Play Video Games? An Empirical Study on the Motivations of Twitch Users. *SSRN Electronic Journal*. Advance online publication. https://doi.org/10.2139/ssrn.2779543

Smith, T., Obrist, M., & Wright, P. (2013). Live-streaming changes the (video) game. In *EuroITV '13 - Proceedings of the 11th european conference on Interactive TV and video* (pp. 131–138). New York: ACM. https://doi.org/10.1145/2465958.2465971

Steinkuehler, C. A., & Williams, D. (2006). Where Everybody Knows Your (Screen) Name: Online Games as "Third Places". *Journal of Computer-Mediated Communication*, *11*(4), 885–909. https://doi.org/10.1111/j.1083-6101.2006.00300.x

Stenström, A.-B. (1994). *An Introduction to Spoken Interaction*. *Learning about language*. London: Routledge.

Suznjevic, M., Dobrijevic, O., & Matijasevic, M. (2009). MMORPG Player actions: Network performance, session patterns and latency requirements analysis. *Multimedia Tools and Applications*, *45*(1-3), 191–214. https://doi.org/10.1007/s11042-009-0300-1

Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. *Cambridge applied linguistics series*. Cambridge: Cambridge University Press.

Szablewicz, M. (2015). A Realm of Mere Representation? "Live" E-Sports Spectacles and the Crafting of China's Digital Gaming Image. *Games and Culture*, 1–19. https://doi.org/10.1177/1555412015595298

Tannen, D. (1987). Conversational Style. In H. W. Dechert & M. Raupach (Eds.), *Psycholinguistic models of production* (pp. 251–267). Norwood: Ablex Publishing.

Taylor, N. (2009). Power Play:Digital Gaming Goes Pro (Dissertation). York University, Toronto.

Taylor, N. T. (2016). Now you're playing with audience power: The work of watching games. *Critical Studies in Media Communication*, *33*(4), 293–307. https://doi.org/10.1080/15295036.2016.1215481

Taylor, T. L. (2009). *Play between worlds: Exploring online game culture*. London: MIT Press.

Taylor, T. L. (2011). Internet and Games. In Consalvo, Mia & C. Ess (Eds.), *The Handbook of Internet Studies* (pp. 369–383). Oxford: Wiley-Blackwell. https://doi.org/10.1002/9781444314861.ch17

Taylor, T. L. (2012). *Raising the stakes: E-sports and the professionalization of computer gaming*. Cambridge, Mass.: MIT Press.

Taylor, T. J., & Cameron, D. (1989). *Analysing conversation: Rules and units in the structure of talk* ([1st ed., repr.]). *Language & communication library: Vol. 9*. Oxford: Pergamon.

Tekin, B. S., & Reeves, S. (2017). Ways of Spectating: Unravelling Spectator Participation in Kinect Play. In *CHI '17: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems,* (pp. 1558–1570). New York, NY: ACM. https://doi.org/10.1145/3025453.3025813

Thompson, D., & Filik, R. (2016). Sarcasm in Written Communication: Emoticons are Efficient Markers of Intention. *Journal of Computer-Mediated Communication*, *21*(2), 105–120. https://doi.org/10.1111/jcc4.12156

Turkle, S. (2012). *Alone together: Why we expect more from technology and less from each other*. New York: Basic Books.

Twitch. (2014). To Thousand MOREteen. Retrieved from https://www.twitch.tv/year/2014/

Twitch. (2015). Twitch: Social Video for Gamers. Retrieved from https://www.twitch.tv/p/about

Twitch. (2015). The 2015 Retrospective. Retrieved from https://www.twitch.tv/year/2015/

Twitch. (2017). 2017 Year in Review. Retrieved from https://www.twitch.tv/year/2017/

Twitch. (2016). Year in Review 2016. Retrieved from https://www.twitch.tv/year/2016/

Twitch. (2017). 2017 Year in Review. Retrieved from https://www.twitch.tv/year/2017/

Velloso, E., & Carter, M. (2016). The Emergence of EyePlay. In *CHI PLAY '16: Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play* (pp. 171–185). New York: ACM. https://doi.org/10.1145/2967934.2968084

Wall Street Journal (2014, February 6). Twitch ranked fourth in peak internet traffic for US. Retrieved from http://online.wsj.com/news/articles/SB10001424052702304851104579361201655365302

Walsh, C., & Apperley, T. (2008). Gaming Capital: Rethinking Literacy. In *Changing Climates: Education for sustainable futures. Proceedings of the AARE 2008 International Education Research Conference.* Queensland University of Technology.

Wagner, M. G. (2006). On the Scientific Relevance of esport. In *ICOMP '06 - Proceedings of the 2006 International Conference on Internet Computing & Conference on Computer Games Development* (pp. 437–442). [S.l.]: CSREA Press.

Walther, J. B. (1996). Computer-Mediated Communication: Impersonal, Interpersonal, and Hyperpersonal Interaction. *Communication Research*, *23*(1), 3–43. https://doi.org/10.1177/009365096023001001

Warman, P. (2015). *The Global Games Market: Trends, Market Data and Opportunities.*

Warren, M. (2006). *Features of naturalness in conversation*. *Pragmatics & beyond: v. 152*. Amsterdam, Philadelphia: J. Benjamins.

Weber, R., Behr, K.-M., Tamborini, R., Ritterfeld, U., & Mathiak, K. (2009). What Do We Really Know About First-Person-Shooter Games? An Event-Related, High-Resolution Content Analysis. *Journal of Computer-Mediated Communication*, *14*(4), 1016–1037. https://doi.org/10.1111/j.1083-6101.2009.01479.x

Werry, C. C. (1996). Linguistic and Interactional Features of Internet Relay Chat. In S. C. Herring (Ed.), *Computer-mediated communication: Linguistic, social, and cross-cultural perspectives* (new ser. 39, pp. 46–63). Philadelphia: John Benjamins.

Wilkinson, S., & Kitzinger, C. (2006). Surprise As an Interactional Achievement: Reaction Tokens in Conversation. *Social Psychology Quarterly*, *69*(2), 150–182. https://doi.org/10.1177/019027250606900203

Williams, J. P., & Kirschner, D. (2012). Coordinated Action in the Massively Multiplayer Online Game World of Warcraft. *Symbolic Interaction*, *35*(3), 340–367. https://doi.org/10.1002/symb.22

Winn, C. (2015). The Well Played Moba: How Dota 2 and League of Legends use Dramatic Dynamics. In *DiGRA '15 - Proceedings of the 2015 DiGRA International Conference.* New York: ACM.

Wirman, H. (2007). I am not a fan, I just play a lot: If Power Gamers Aren't Fans, Who Are? In *DiGRA '07 - Proceedings of the 2007 DiGRA International Conference.* Digital Games Research Association.

Witkowski, E. (2012). Inside the Huddle: The Phenomenology and Sociology of Team Play in Networked Computer Games (PhD Dissertation). IT University of Copenhagen, Copenhagen.

Yildirim, I. G. (2016). Time Pressure as Video Game Design Element and Basic Need Satisfaction. In *CHI '16: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 2005–2011). New York, NY: ACM. https://doi.org/10.1145/2851581.2892298

Zagal, J. P., Fernandez-Vara, C., & Mateas, M. (2008). Rounds, Levels, and Waves: The Early Evolution of Gameplay Segmentation. *Games and Culture*, *3*(2), 175–198. https://doi.org/10.1177/1555412008314129

Zagal, J. P., & Mateas, M. (2007). Temporal Frames: A Unifying Framework for the Analysis of Game Temporality. In *DiGRA '07 - Proceedings of the 2007 DiGRA International Conference.* Digital Games Research Association.

Zagal, J. P., & Mateas, M. (2015). Analyzing Time in Video Games. In P. Lankoski & S. Björk (Eds.), *Game research methods: An overview* (pp. 37–56). Pittsburgh: ETC Press.

Zillmann, D., & Stocking, S. H. (1976). Putdown Humor. *Journal of Communication*, *26*(3), 154–163. https://doi.org/10.1111/j.1460-2466.1976.tb01919.x