



Patterns of Communication in Live Streaming

A comparison of China and the United States

SHUQIAN. ZHOU
MEIMEI. WANG

Master of Communication Thesis

Report nr. 2017:088

Acknowledgements

First of all, we would like to express our sincere gratitude to our supervisor, Prof. Jens Allwood, for his weekly meetings during this half year and the constructive suggestions to our thesis. We learned a lot from him, not only his profound knowledge in communication, but also his rigorous academic attitude. We appreciate all of his efforts in our thesis.

We would also like to thank our teachers of Master in Communication. They helped us build our knowledge systems in communication from the very beginning and keep enriching the systems within the two years. They also encouraged us to cooperate with each other and present ourselves a lot. We are grateful for their help in our study.

Our sincere thanks also goes to our classmates, for their help and support for us in the whole two years.

Finally, we must express our gratitude to our parents for providing us with unfailing support and continuous encouragement throughout our two years of study. This thesis would not have been possible without them.

Thank you all again.

Abstract

Live streaming as a social medium provides a multi-functional internet platform for its users to have real-time interaction with others through broadcasting live streaming videos by mobile devices and websites. It brings a new mode of communication. There are previous studies related to its basic usage practices, user's behavior and its applications in specific fields, etc. However, this study is made from a communicative perspective. It aims to describe and analyze the communication patterns in live streaming. The study is a comparative study between China and the United States.

In order to study the communication patterns in live streaming, 106 live streaming videos are observed (the total length of 2251 minutes). Combining qualitative analysis with quantitative analysis, communication patterns in live streaming are analyzed based on relevant communication theories including: Interactive Communication Management, and Multimodal Communication. Cultural differences between China and America are reflected during analyses of communication patterns in live streaming. The findings demonstrate that there are common communication patterns in live streaming in China and the United States. Common communication patterns are mostly influenced or decided by traits of live streaming, the new social medium. Common communication patterns in the two countries inference some general communication patterns in live streaming. But, different communication patterns in live streaming in China and the United States also exist. Indicating the cultural impact of the countries on communication patterns in live streaming.

Key words: Live streaming, Communication patterns, China, the United States, Culture.

CONTENT

1. Introduction.....	6
1.1 What Is Live Streaming and What Are Communication Patterns?.....	6
1.2 Why Studying Communication Patterns in Live Streaming in Both China and America?	7
1.3 Research Question and Purpose.....	8
1.4 The Framework of the Thesis	8
2. Research Background	10
2.1 Development of Live Streaming	10
2.1.1 <i>Development of Live Streaming in the United States</i>	10
2.1.2 <i>Development of Live Streaming in China</i>	12
2.2. Functions and Features of Live Streaming	13
2.3 Related Studies.....	16
3. Theoretical Framework.....	21
3.1 Interactive Communication Management.....	21
3.1.1 <i>Turn Management</i>	21
3.1.2 <i>Sequence</i>	22
3.1.3 <i>Feedback</i>	23
3.2 Multimodal Communication	24
3.2.1 <i>Dimensions of Production and Perception in Multimodal Communication</i>	25
3.2.2 <i>Body Movements</i>	26
4. Methodology	28
4.1 Study Design	28
4.1.1 <i>Observation and Coding Framework</i>	28
4.1.2 <i>Types of Live Streaming to Be Observed</i>	30
4.2 Data Analysis	31
4.3 Ethical Consideration.....	32
5. Results and Analyses.....	33
5.1 Overview of the Observation Data	33
5.1.1 <i>Demographics of Broadcasters</i>	33

5.1.2 Place Broadcasting	34
5.1.3 Number of Viewers	35
5.1.4 Interactivity in China and the USA.....	37
5.2 Interactive Communication Management.....	38
5.2.1 Turn Management.....	39
5.2.2 Feedback.....	41
5.2.3 Sequence.....	49
5.3 Multimodal Communication in Live Streaming	50
5.3.1 Specific Multimodal Communication in Live Streaming	50
5.3.2 Body Movements in Live Streaming	52
5.3.3 Flexibility of Multimodal Communication in Live Streaming	58
5.3.4 Multimodality in Different Contents of Live Streaming	59
6. Discussion	61
6.1 Delayed Answers from Broadcasters	61
6.2 Questions Chosen by Broadcasters	63
6.4 Lack of Understanding.....	65
6.5 Embarrassment Hiding.....	66
7. Conclusion	68
7.1 Patterns of Communication in Live Streaming.....	68
7.2 Differences between Chinese and American Live Streaming.....	69
7.3 Limitations and Future Studies	71
References.....	72
Appendices.....	77
Appendix 1 Observation Schema.....	77
Appendix 2 Cluster Membership in USA	79
Appendix 3 Cluster Membership in China	81
Appendix 4 Division of Work.....	83

1. Introduction

1.1 What Is Live Streaming and What Are Communication Patterns?

In previous studies, the new social medium of live streaming is not given a complete and accurate definition. Powell (2015) says that live streaming essentially allows you to capture and stream, or watch, live video on your mobile device. Hamilton et al. (2014) claim that live streaming enables public broadcast of live audio and video streams alongside a shared chat channel. Pires et al. (2015) mentioned that live video streaming systems are services that allow anybody to broadcast a video stream over the Internet. Juhlin et al. (2010) state that live streaming makes it possible to capture live video on a mobile phone and broadcast it in real time to a web page. Hamilton et al. (2016) claim that live streaming has come to refer to live, streaming, video as well as a set of communication media that enable viewers to interact with each other and the streamer. From these descriptions, there are three factors that need to be considered when defining live streaming, live videos and audios, through mobile devices or internet, interacting or sharing with others. Therefore, based on the definition of Tang et al. (2016), live streaming enables immediate live broadcasting of video and audio from a smartphone, to whomever wants to tune in. We define live streaming as: a social medium that provides a multi-function internet platform for its users to have the real-time interaction with others through broadcasting live streaming videos by mobile devices and websites. Communication patterns refer to specific features of communication that are typical of a certain community or activity, such as typical sequence of events, feedback, turn taking, or spatial arrangements, topics, nonverbal behavior etc. Therefore, the number of such aspects and traits is large and what is at stake is, therefore, to focus on aspects and traits which have turned out to be interesting in a given community or activity (Allwood, 1999). According to Allwood (1999), the concept of "patterns of communication" is fairly general and does not imply very much more than repeated traits of, or aspects of the communication of the members of a certain social or cultural group. He has focused on cultural reasons or influences for communication features

when studying communication patterns. Other studies have analyzed communication patterns from the perspective of management and had claimed communication patterns are structures in which communication flows in an organization (Mishra, 2017). This study also focused on the communication links in work teams according to organizational structures. It is clear that different perspectives focus on different aspects when talking about communication patterns.

This study analyzed communication patterns in live streaming in both China and America, it focuses on features of communication and basic structure of communication patterns in live streaming. Meanwhile, cultural factors are also considered in the analysis, particularly when they are the main reasons for differences between Chinese and American live streaming communication.

1.2 Why Studying Communication Patterns in Live Streaming in Both China and America?

Live streaming not only brings new opportunities for the development of social media, more significantly, it creates new communication modes for social media users. People are not satisfied to post personal status on Facebook or Twitter anymore, they broadcast their lives, e.g. concerts they are enjoying, activities they are participating in, travels they are experiencing, etc. Through live streaming, they share lives and interact with whomever wants to tune in directly and in real time. Moreover, communication is the basis of live streaming. Broadcasters can communicate with viewers through videos, audios or graphics, while the audience can only use text-based comments or tap hearts up to communicate with broadcasters. This interactivity is important in live streaming. As above, this new mode of communication has great potential as data for research. However, through our analyses of previous works, they paid more attention on, for instance, live streaming usage practices; broadcasters' behaviors; audience psychology, etc., but not on patterns of communication. Thus, we study live streaming from a communicative perspective. This study aims to describe and analyze communication

patterns in live streaming.

In general, interactivity in live streaming, and the communication mode brought by this new medium, as well as the analyses of previous studies provide us great motivation to study patterns of communication in live streaming.

In addition, the huge development potential of live streaming attracts a lot of attention in both the USA and China. From 2015, in the USA, 4 live streaming platforms emerged one after another. They belonged to Twitter, Facebook, and Google, which are giants of American social networking or internet companies in America. In China, according to Statistical Report on Internet Development in China (2017), there are more than 200 platforms till 2017. And by the end of December 2016, online broadcast users reached 344 million, accounting for 47.1% of the total Internet users. The very rapid development of live streaming in China and America probably make them the most advanced countries when studying live streaming.

1.3 Research Question and Purpose

This study tries to answers the question: *What are communication patterns in live streaming in China and the United States?*

According to this research questions, purposes of this paper include i) describe and analyze basic features of communication in live streaming, ii) construct the basic structure of communication in live streaming, iii) demonstrate similarities and differences between live streaming communication and face-to-face communication in relevant analyses, (iv) compare specific traits of communication in live streaming in China and America from a cultural perspective. This thesis aims to show a clear and complete picture of communication patterns in live streaming.

1.4 The Framework of the Thesis

This thesis consists of seven sections. The first section, introduction, mainly introduces

research question and purposes of the thesis. The second section, research background, includes developments of live streaming in the United States and China, basic functions and features of live streaming, as well as prior studies in this field. The third section shows the theoretical framework of used in the thesis. After that the method used in this study will be introduced. Then, in the results and analyses section, results from the empirical data are described, analyzed and compared according to the theoretical framework. The next section discusses special aspects focused on during observations. The final section concludes the whole thesis, displaying the main findings in our study.

2. Research Background

In this chapter, we introduce live streaming and describe its development, main functions and features, as well previous related studies. The purpose is giving a clear, general understanding of this new social medium.

2.1 Development of Live Streaming

The emergence of live streaming as a popular medium in recent years provides a new platform for socio-technical interaction in the world (Tang et al., 2017). The popularity of live streaming platforms, Periscope and Twitch, Facebook etc. , have attracted a lot of users, media attention, and capital injection. However, Rome was not built in a day, live streaming has its course of development.

2.1.1 Development of Live Streaming in the United States

In the United States, the earlier live streaming platforms can be traced back to Ustream and Justin.tv. Both of these two platforms were built in 2007. These two social platforms were seen as the founders of live streaming. At the time, their service worked off webcams hooked to computers for the first time (McDermott, 2015). They allowed users to broadcast, and watch, live video streams online.

Ustream (acquired by IBM) was famous for its real-time broadcast of political events, while Justin.tv made the huge success on its gaming channel, which became a separate website called Twitch.tv in 2011. Till 2014 (Twitch was acquired by Amazon with \$970 million), Twitch has become one of the largest live streaming platform in the world and it held the leading position in gaming live streaming market. According to Hamilton et al. (2014), Twitch has over 34 million unique monthly viewers and tens of thousands of streamers.

The success of Twitch is a turning point for the development of live streaming. Huge capital investment by big companies, IBM and Amzon, not only opened new market for live streaming, but also attracted more attention from the public. In 2013, the

implementation of 4G LTE (Long-Term Evolution) mobile networks, a high-speed protocol that transmits data 10 times faster than 3G, made it much easier to send and receive video from cellphones (McDermott, 2015). 4G service broke the technological limitations for live streaming. Therefore, with large capital injection, 4G technology and the maturity of social platform (Twitter, Facebook, etc.), live streaming ushered in the golden age of its development.

In February, 2015, San Francisco entrepreneur Ben Rubin announced the launch of his live streaming' video application, Meerkat (Edelman, 2016). This application allows users to shoot video footage on their smartphones and simultaneously make that footage appear in real time on the internet, and allows watchers to comment live.

Just 2 weeks later, Twitter cut off Meerkat's (shut down in October, 2016) integration with its feed and announced its similar online application, Periscope (McDermott, 2015). Periscope allow users to follow broadcasters and comment on or "heart" the videos. Unlike Meerkat, Periscope broadcasts persist on the app for 24 hours after a filming (McDermott, 2015). Meerkat and Periscope instantly became rivals and have been developing features to set themselves apart from each other.

In August 2015, Facebook introduced its app Facebook Live for only celebrities with the verified Pages. Then on April 6th, 2016, Zuckerberg announced they were launching Facebook Live for all users. Anyone with a phone now has the power to broadcast to anyone in the world when they using Facebook. Besides personal users, there are many news media, such as BBC, starting to broadcast live news by using Facebook Live. After that, on August 26th, 2015, YouTube launched its first live streaming channel, YouTube Gaming—a video game oriented platform. After that, they opened more live streaming channels such as sports, technology, animal etc.

From the above, through the development of representative live streaming platforms in the United States, it is clear to see that, in the year of 2015, there was the boom of live streaming after the success of Twitch. This boom is inseparable from progress of technology, market investment, the participation of social platforms. The continuous emergence and growth of new platforms showed the flourishing development of live streaming.

2.1.2 Development of Live Streaming in China

In China, based on the information we collected from the official websites of live streaming platforms and national Statistical Report on Internet Development (2017), the development of live streaming can be divided into three phases.

Before 2013, 9158, YY Live (changed name to Huya Live in January, 2014) etc. were the main platforms at the time. Early live streaming platforms mainly focused on live show. They provided internet platforms for individual broadcasters to show personal talent (like sing, dance etc.) or chat with viewers through broadcasting on live streaming platforms. Broadcasters could get salaries from platforms. Early live streaming platforms built this underlying business model of live streaming in China.

From 2014 to 2015, gaming live streaming had a huge development in China because of the emergence of two big live streaming platforms, Douyu and Panda.tv.

In January, 2014, Douyu and Zhanqi were launched at same month. Both of them focused on gaming video. However, Douyu defeated Zhanqi after launched and became the leader of gaming live streaming market in China. Douyu's official website showed that it has more than 70% market of gaming live streaming in China. In 2016, its popularity and profitability also won \$100 million investment from Tencent (one of Chinese largest Internet companies).

After that, in October, 2015, Panda.tv was launched by Sicong Wang, son of the richest man in China. Panda.tv also mainly focus on gaming video. It developed very quickly with the support of abundant funding. Till now, it has become the biggest rival to Douyu in gaming live streaming market.

Both Douyu and Panda.tv increase the competitiveness of gaming live streaming in Chinese live streaming market. Moreover, their success also brought the prospect of a better development of live streaming in China.

After 2016, Inke and Yizhibo showed new development direction of live streaming in China. Inke was launched in May of 2015, as a platform of live streaming, not like Douyu or Panda.tv, which focus on gaming video, Inke included so many different contents, such as live show, sports, music, travel etc. It is quite popular as the platform

of live streaming among young people. Yizhibo was launched in March, 2016 by Sina (one of Chinese largest Internet companies). It can be used with Weibo (microblog). Yizhibo developed very fast and got much investments from Sina. It is similar to Inke, users can broadcast many different things in their life, no matter their home dinner or football games they watched.

The popularity of Inke and Yizhibo shows that live streaming in China is changing from specialization to popularization. More and more ordinary people use live streaming to broadcast their daily life. Live streaming does not only focus on live shows or gaming videos anymore. According to Statistical Report on Internet Development in China (2017), there are more than 200 live streaming platforms in China till 2017. Most of the platforms are trying to expand the content of live streaming, especially live streaming about their daily life and live streaming programs in order to attract more types of users. To contrast with live streaming in the United States, Chinese live streaming has its unique business model. On most of American platforms (except Twitch), the broadcasters do not have incomes from live streaming, the audience watch for free. However, in China, broadcasters on most platforms have incomes. Their incomes are depended on the number of the audience and the virtual gifts they get from live streaming. The audience buy virtual gifts by real money on the platforms. After broadcasting, platforms and broadcasters distribute money in certain proportion. In this way, live streaming shows strong profitability in China.

2.2. Functions and Features of Live Streaming

Live streaming provides a multi-function Internet broadcast platform for broadcasters to have real-time interaction with others through immediate live broadcasting of video and audio from a computer or mobile smartphone. Live streaming combines high-fidelity graphics and video with low-fidelity text-based communication channels to create a unique social medium (Hamilton et al. 2014).

The typical live streaming experience consists of a broadcaster broadcasting a video stream accompanied by a dedicated chat channel (Hamilton et al. 2016). Through live streaming's user interface, both broadcasters and audience can see graphics and videos from broadcasters, as well as the comments from all audiences. However, broadcasters cannot use text-based functions to communicate with the audience, in contrast, the audience can only use text-based functions to give comments or feedback to broadcasters.

The essential functions of live streaming are allowing users to broadcast or watch real time live video on relevant platforms. Streaming video in real time means people all around the world can watch whenever/whatever you are broadcasting through relevant platforms (Powell, 2015). Through this function, live streaming provides a new way for its users to share experiences as they happen with no editing or uploading. Real time broadcasting as one of the significant factors of live streaming appeals to the human desire to live out new experiences vicariously through someone else, whether the streams are by a stay-at-home dad cooking dinner or a celebrity taking viewers through a red carpet event, users can broadcast themselves (Brouwer, 2015). Sharing one's own experiences or participating in some things vicariously through others' experiences show the most unique and significant value of live streaming.

Some functions and features of live streaming can be described by introducing two representative live streaming platforms, Periscope (from the United States) and Yizhibo (from China). As two professional and mature live streaming platforms, they have some similar functions.

After users choosing to create new accounts or login by their Twitter/Weibo accounts, they can shoot or watch live streaming videos on the platforms. Both Periscope and Yizhibo allow users to follow broadcasters. When seeing the live streaming videos, viewers can comment on or "heart" the videos. Figure 1 and 2 respectively show two screenshots of viewing a live stream in both Periscope and Yizhibo apps at the time of the observation of this study (April, 2017).

Periscope show the number of viewers in lower right corner as part of the live stream for all to see (Tang et al, 2016). While Yizhibo only show who are watching this video

at the same on the top of the image. Yizhibo shows broadcasters' information at the upper right corner, however, Periscope does not shows broadcasters' information. Periscope broadcasts persist on the app for 24 hours after a filming, while Yizhibo can save the videos more than 24 hours. Both of Periscope and Yizhibo users can set their app to notify them if those whom they follow are broadcasting. In addition, both applications eat cellphone power and bandwidth, so it is recommended that broadcasts be made over a WiFi connection.



Figure. 1 Viewing a live stream in Periscope **Figure. 2 Viewing a live stream in Yizhibo**

From the above, we see that the main functions of live streaming are allowing users to broadcast or watch real time live video. These functions enable remote viewers to engage and participate in shared live experiences (Hamilton et al. 2016). Users can also follow other broadcasters, set notifications, give text-based comments and tap hearts when watching live streaming videos. During live streaming, broadcasters and viewers communicate with each other by different modalities. And their interactions are always real-time and simultaneous.

2.3 Related Studies

The rapid development of live streaming has drawn the attention of some scholars. A number of studies on live streaming have been conducted in different academic fields. McDermott (2015) quoted above and Cui (2016) provided the general introduction of live streaming respectively in the United States and China.

McDermott (2015) introduced important nodes in the development living streaming in the United States: from YouTube, Ustream, to Meerkat, Periscope, and Facebook live. She focused on analyzing the history of development of Meerkat, Periscope, and Facebook live. She also compared advantages and shortcomings of Meerkat and Periscope by analyzing features of their interfaces. Cui (2016) analyzed the categories of living streaming in China. She summarized four live streaming categories: live streaming focusing on publishers' performance, live streaming focusing on audience reaction, live streaming focusing on the content, and live streaming focusing on constructing specific scene. She also claimed that PUGC (professional user-generated content) will replace UGC (User-Generated Content) and PGC (Professionally-generated Content) and becoming the main trend in the future live streaming in China. Tang et al. (2016), Siekkinen et al. (2016), Lim et al.(2012), Juhlin et al. (2010), Weisz et al. (2007) studied the usage practices/patterns of live streaming or mobile live video service.

Tang et al. (2016) made a comparative study about live streaming by comparing Meerkat and Periscope apps for live streaming mobile devices. They described the contents, settings, and other characteristics of live streaming. They found that most of streamers were motivated to broadcast in order to develop their personal brands. Moreover, they studied a range of streamers' responses to their viewers' comments and found that the highest percentage of streamers who actively or sometimes responded occurred in chats, conversely, streaming a professional or amateur event had lower percentages of responsive streamers.

Siekkinen et al. (2016) explored the anatomy of a mobile live streaming service by doing the case study of Periscope. They studied live streaming service to understand its

usage patterns and technical characteristics of the service (e.g. delay and bandwidth) and also performed the adaptation strategies of using Periscope. Lim et al. (2012) studied usage practice by constructing a social media-enhanced real-time streaming video service prototype and conducted a field experiment with actual social media users. Their research results indicate that: inhabited space (the degree of being situated in context and in a meaningful place) and isomorph effects (the degree of preserving the structure of a user's actions) reduce psychological distance between users, and this, in turn, enhances co-experience.

Juhlin et al. (2010) were interested in mobile live video and user-generated content. They provided a qualitative content analysis of four such services (bambuser.com, qik.com, flixwagon.com, and kyte.com.). Their analysis revealed how broadcasters utilize the different affordances (text, photography, audio, video files) of this new medium. Weisz et al. (2007) studied people's experience of watching videos online, while simultaneously chatting with others using a text chat feature. Through experiments, they found that new peer-to-peer video streaming technologies fundamentally change the passive way we experience media, people actively engage with each other as engaging with the video, but active engagement comes at a cost.

Zhang and Liu (2015), Pires et al. (2015), Shamma et al. (2009), Kaytoue et al. (2012) studied the interaction between publishers and viewers in live streaming from the different perspective. Specially, Shamma et al. (2009), Kaytoue et al. (2012) explored the interaction between publishers and viewers when building virtual communities in live streaming.

Zhang and Liu (2015) focused on the multi-sourced live streaming broadcasters' and viewers' interactive experience from the technological perspective. They made a measurement study by taking Twitch as a representative. The results revealed that current delay strategy on the viewer's side substantially impacts the viewers' interactive experience. On the broadcaster's side, the dynamic uploading capacity is a critical challenge. Pires et al. (2015) focused on the User-Generated live video streaming systems. They presented a dataset for three months of traces of both Twitch and

YouTube. They found that both systems generate a significant traffic with frequent peaks at more than 1 Tbps.

Shamma et al. (2009) addressed the ways in which DJs have adopted one webcasting technology, Yahoo Live. They found that three overlapping communities that are important to the DJs, with whom they want to maintain reliable and consistent connection. They also claimed that during the asynchronous interaction, small cues, feedback, music, chat facilities and the potential for eye-gaze with audience members offer a connection between the DJ and their audiences. Kaytoue et al. (2012) studied Electronic-sport (E-Sport) by analyzing Twitch.tv. In their paper, they found that a new Web community for e-sport fans watching live streaming was emerging. Their results also show that tournaments and releases translate into clear growths of the game audience.

Hamilton et al. (2014, 2016) did studies about how live streaming fosters participation and builds community. Hamilton et al. (2014) presented an ethnographic investigation of the live streaming of video games on Twitch. They found that Twitch streams act as virtual third places, in which informal communities emerge, socialize, and participate. The assemblage of hot (video) and cool (text) media enables live streaming to provide an open place for people to go socialize, play, and participate in something larger than themselves. Hamilton et al. (2016) studied how to support communication and participation in multi-stream experiences, they presented the design and evaluation of Rivulet, and found that viewers use all modalities (text chat, Push-To-Talk, and hearts) to engage with the streamers. They also found that multi-stream experiences led to interesting cross-stream interactions. Viewers were able to easily find and participate in streams that addressed their interests and desire for engagement.

House (2016) and Thorburn's (2013) studied live streaming from a political perspective. Thorburn's (2013) article seeks to critically examine the practical application of live streaming video at use in contemporary resistance movements. She claimed that political actors and digital technologies can form unique assemblages, which can both operate as mechanisms of power as surveillance technologies for police forces and open up nodes of counter-power, disrupting state surveillance. House (2016) claimed that

live streaming is a dynamic assemblage of technologies, practices, and participants that destabilizes the boundaries of who is a participant and even what, where, and when is an event. She found that the embodied, immediate, and intimate nature of live-streaming could promote mutual understanding and interaction among protesters, police, and distant viewers.

Ding et al. (2011) explored the behavior of broadcasters in live streaming. Yuan et al. (2016), Li (2015), Jia (2016) studied the audience psychology in live streaming in China. Ding et al. (2011) studied YouTube uploaders and conducted extensive measurement and analysis of them. They found that: among these uploaders, the most active 20% uploaders contribute 72.5% of the videos; there are more than three times male uploaders than female uploaders. Furthermore, they found that much of the content in YouTube is not user generated. Many YouTube uploaders are user copied rather than user generated. Yuan et al. (2016) explored the psychological phenomenon behind live streaming. He claimed that live streaming's turning from the public space to private sector also leads viewers' complex psychological logic. According to Foucault's theory of the panoramic view and Lacan's gaze theory, they claimed that viewer's watching behavior reveal a kind of self-image construction and self-identity reshaping. It also reflected people's needs of peep and sexual drive. Jia (2016) studied audience psychology by using Goffman's Dramaturgical model. He also claimed that live streaming puts back stage in front and it satisfied viewers' need of peep.

Brouwer (2015) and Birkner (2016) studied live streaming from a business perspective. Brouwer (2015) explored how broadcasters could use live streaming for business purposes while delivering high-quality broadcast experiences to their audiences and customers. She claimed that accessibility makes live streaming much more powerful. Birkner (2016) emphasized that the significance of Periscope to organizations is -- allowing organizations to humanize their brand. Periscope makes it simple for brands to connect with their consumers on a personal level.

From the analysis above, previous studies of live streaming are mainly about usage practices; usage patterns; interaction between broadcasters and audiences; building virtual communities; fostering participation; roles of political tools; broadcasters'

behaviors; audience psychology and business benefits. Comparative study, case study and measurement study are used in this field. Twitch, Meerkat, Periscope, YouTube live and Facebook live are usually taken as representatives. In general, prior works focusing on American live streaming platforms give us an early glimpse at a rapidly evolving social communication technology. The studies in China, however, focus more on psychological reasons why the audience watch live streaming videos. All of these show us important studies about live streaming in different fields.

According to the research background, it is clear that studies on live streaming focus on very limited fields. There is no previous study focusing on communication patterns in live streaming.

3. Theoretical Framework

In order to analyze communication patterns in live streaming, we base our analysis on a theory framework. It includes two parts: interactive communication management and multimodal communication. These theories are described one after another.

3.1 Interactive Communication Management

Simultaneous communication in live streaming involves lots of interlocutors. To make this multilateral communication go successfully, interactivity in live streaming is very important. When analyzing interactivity in live streaming, a theory of interactive communication management is used in this study. Interactive communication management (ICM) refers to the features of communication supporting interaction, e.g. mechanisms for management of turns, feedback, sequencing, rhythm and spatial coordination (Allwood, 2008b). Turn management, feedback, and sequence of ICM are used in this study.

3.1.1 Turn Management

A turn is defined as a speaker's right to the floor (Allwood, 1999a). Turn management regulates the interaction flow and minimizes overlapping speech and pauses (Allwood et al., 2007). When we have two-way interactive communication, turn management as mechanisms and processes are essential (Allwood, 2010).

Three general features of turn management are turn gain, turn end and turn hold (Allwood et al., 2007). Turn gain can be divided into turn taking and turn accepting. Turn taking means the speaker takes a turn that wasn't offered, possibly by interrupting, while turn accepting means the speaker accepts a turn that is being offered. Similarly, turn end also can be divided into two types. Turn yielding means the speaker releases the turn under pressure, while turn offer means the speaker offers the turn to the interlocutor, or a turn complete if the speaker signals completion of the turn and end of the dialogue at the same time (Allwood et al., 2007).

3.1.2 Sequence

Activities can be subdivided into different sub-activities. If we take one live stream as an activity, it should consist of several sub-activities. Each sub-activity has its sequence from start to end. A common sequence is the following (Allwood, 1999a).

- (i) Initiation (opening, entering an activity, a sub-activity or a topic)
- (ii) Maintenance (maintaining a sub-activity or topic)
- (iii) Changing (changing a sub-activity or topic)
- (iv) Ending (closing an activity, a sub-activity or a topic)

When it comes to specific patterns of communication, there will be typical sequences. The expression "typical sequences of events" is intended to refer to the fact that what happens in a conversation often happens in a certain sequence (Allwood, 1999b). Different cultures, different activities have their own sequence of communication, varying in initial sequences, medial sequences, and final sequences. Open sequence, continue sequence and close sequence are used to analyze different stages of a meaningful sequence (Allwood et al., 2007).

Contributions often occur in fairly set sequences, such sequences extend from "exchange types" (Allwood et al., 2012). In particular activities, preferred types of responses are used to reply certain comments. For example, in live streaming, broadcasters usually express gratitude verbally and bodily after receiving gifts and hearts. The gestures and utterances of broadcasters are different from gratitude expression in other activities. That is the particular exchange type in live streaming. For example, broadcasters have their own gestures and utterances which only appear in live streaming.

In this study, we noted the number of sequences during observation. One sequence in living streaming consists of one or several interactions of broadcasters and viewers of the same topic. The open sequence refers to the beginning of a new topic. And the continue sequence is the continuing conversation of the same topic as its beginning. Sequence closes before the topic was changed.

3.1.3 Feedback

In order to ensure that communication is going successfully and interlocutors have shared understanding, there is a system of communicative feedback (Allwood, 2010). Communicative feedback can be explained in broad sense or in narrow sense. Communicative feedback (narrow sense) refers to unobtrusive (usually short) vocal or body expressions whereby a recipient of information can inform a contributor of information about whether he/she is able and willing to (i) communicate (have contact), (ii) perceive the information (perception), and (iii) understand the information (understanding). In addition, (iv) feedback information can be given about emotions and attitudes triggered by the information, a special case here being an evaluation of the main evocative function of the current and most recent contributions (Allwood, 2008a). Studies of feedback usually use the concept of narrow sense, because the narrow sense can help scholars focus on ways of giving and perceiving feedback and functions of feedback, etc.

In this study, feedback is used in its broad sense. "Feedback" (broad sense) refers to the fact that speaker as well as listener, in a conversation must know how the other party is reacting (Allwood, 1999b). The interlocutors need to know whether the information is perceived and understood by each other. The speaker also needs to know how the listener reacts to what is being said (Allwood, 1999b). From above, feedback in broad sense contains not only willingness, perception, understanding and emotion, but also the replies to what is being said. In this study, we analyze patterns of communication in living streaming from a broad sense, including analyses from many aspects, such as turn management, feedback, sequence, etc. Using the broad sense of feedback in the study can cover more information and can describe feedback from a broad view. Feedback of broad sense helps us give a comprehensive communication patterns in live streaming.

The main ways of giving feedback linguistically are body feedback, e.g. head movement or smile (Jensen, 2014), and spoken feedback such as "yes". Spoken feedback can be feedback words like "yes, no, m" with various phonological and

morphological operations allowing expansion of these words, repetition of words in a previous utterance to show agreement or to elicit confirmation or more information, and pronominal or other types of reformulation (Allwood, 1999a). In this study, spoken feedback also includes main messages in replies.

Thus, in this study, we are mainly studying "feedback in a broad sense", it contains the main reply to interlocutors. To be more specific, the feedback of the broadcasters was been operationalized as the body and linguistic reply to the comments, questions and gifts, hearts. The feedback of the viewers was been operationalized as the comments, questions, gifts, and hearts.

3.2 Multimodal Communication

According to Allwood (2008c), multimodal communication = co-activation, sharing and co-construction of information simultaneously and sequentially through several modes of perception (and production) (Allwood, 2008c). The basic reason for studying multimodal communication is that it provides data for more complete studies of "interactive face-to-face sharing and construction of meaning and understanding" (Allwood, 2008c). In this study, multimodal communication is also important to help understand communication patterns in live streaming.

Many multimodal communication studies focus on face-to-face communication. That is because face-to-face communication is the chief representative of multimodal communication in our lives. Normal face-to-face communication is multimodal, it employs several modalities of production and perception in order to share information (Allwood, 1998). According to Allwood's analyses, the two primary modes of production are speech and various types of body gestures; the two primary modes of perception are, accordingly, hearing and vision. The spoken message normally predominates, while body gestures provide additional information; gestures are often, in turn, reinforced by prosody. (Allwood, 1998).

A term closely linked to multimodal communication is "modality". The term "modality"

can be used in many ways, but the definition adopted here is that “multimodal information” is information pertaining to more than one “sensory modality” (i.e., sight, hearing, touch, smell or taste) or to more than one “production modality” (i.e., gesture (be used in the sense of any body movement), speech (sound), touch, smell or taste) (Allwood, 2008c). These are the traditional five sense modalities and their production modalities.

In multimodal communication, multimodality give us flexibility in the choice of modality, and also the possibility of being redundant when this is needed, for example, in a complex noisy environment (Allwood, 2013). Especially, flexibility is very characteristic of multimodal communication. In many different contexts, the flexibility of choice can help to improve the efficiency of communication. It reflects one of the advantages of multimodal communication.

3.2.1 Dimensions of Production and Perception in Multimodal Communication

In Allwood’s study (2013), besides the basic five senses and their corresponding production modality, he gave an overview of how dimensions of production and perception can be related in multimodal communication (Table. 1).

PERCEPTION	Hearing Understanding + Attitudinal Reactions	Vision	Touch	Ssmell Smell	Taste
PRODUCTION SPEECH:					
Prosody/Phonology	x	x			
Vocabulary	x				
Grammar	x				
GESTURES:					
Facialgestures		x			
Manual gestures		x			
Body movements		x			
Posture		x			
Touch			x		
Smell				x	
Taste					x
Medium	acoustics	optics	physiology	molecules	molecules

Table. 1 Multimodal face-to-face communication—Perception and production

This table displays how different modalities are produced and perceived through

different media. It is clear that multimodal communication production basically includes two aspects: speech (prosody, vocabulary, grammar) and gestures (body movements, posture, etc.) Therefore, besides the auditory aspects of speech, there are also other types of communicative expression (Allwood, 2013), e.g. facial gestures, posture, etc. In fact, all of 15 types of body movements (Allwood, 2002) may be used/analyzed in multimodal communication.

3.2.2 Body Movements

Allwood (2002) discussed 15 major types of body movements from the perspective of production. They are: Facial gestures, Head movements, Direction of eye gaze and mutual gaze, Pupil size, Lip movements, Movements of arms and hands, Movements of legs and feet, Posture, Spatial orientation, Clothes and adornments, Touch, Smell, Taste, Nonlinguistic sounds. Each type is followed by an account of its functions. These functions given are meant as examples.

He also noted that the majority of the body movements are connected with visual sense, while touch is connected with the sense touch and hear with auditory sense. Smell and taste are in ordinary language more or less neutral with regard to production and perception (Allwood, 2002).

The categories of body movements and analyses of functions provide a significant basis for classifications and analyses of body movements in live streaming. In addition, because Allwood's study was based on face-to-face communication, it gives an opportunity to make comparisons of body movements between face-to-face communication and live streaming communication.

Allwood (2002) also claimed that body movements can be used both together with speech and independently of speech. Communicative expressions over and above auditory aspects of speech can supplement auditory aspects of speech or play an autonomous role in communication (Allwood, 2013).

There are studies focused on this aspect. Allwood and Ahlsén (2009) studied features of gesture types that are produced before or simultaneously with speech. They found

that most of the factual information gestures are produced simultaneously with the target words.

Allwood and Cerrato (2003) studied gestures related to verbal feedback expressions, their results showed that feedback is mostly expressed simultaneously by vocal/verbal and gestural means. Specially, the gestures accompanying verbal and vocal feedback expressions can be broadly categorized according to their function in the given context (Allwood & Cerrato, 2003). In addition, they also claimed 4 ways about how gestural feedback expressions modify the meaning of the vocal/verbal expression: reinforcement (R), adding redundancy by giving the same information as the vocal message; positive (P), indicating a positive reinforcing attitude; negative (N), weakening what has been said vocally; contradicting (C), contradicting what has been said vocally.

These studies show that body movements are a major source of the multimodal and multidimensional nature of face-to-face communication. They offer us a new angle when analyzing body movements in live streaming, especially how body movements are used with vocal/verbal expressions in live streaming.

Multimodal communication studies in face-to-face communication provide us with ideas for observing and analyzing multimodal communication in live streaming. They not only demonstrate dimensions of production and perception in face-to-face communication, but also show the importance of body movements in multimodal communication. All of these can help us build a theoretical framework to analyze multimodal communication in live streaming.

4. Methodology

In this section, the study design, data from analysis and ethical considerations that are relevant to this study are described.

4.1 Study Design

4.1.1 Observation and Coding Framework

Observation and analysis are the main methods of this study. We engaged in a pre-study by watching some live streaming videos and analyzing some features of it before constructing the observational schema and unifying the coding standards.

Observation, as a main method of collecting data in this study, helps us directly see both broadcasters' and viewers' interactions. However, observation sometimes can be subjective because of different observers. Therefore, the data of observation is gathered by same coding standards in order to reduce the subjective impact. So the data is more valid in analysis. Through observation, we can also include some other features which were unpredicted in pre-study.

Based on the pre-study and theoretical framework of this study, observation framework (Appendix. 1) mainly includes three parts:

(i) Basic information of a live streaming video: serial number; category; content; platform; watching time; watching duration; number of viewers; broadcaster's ID, age, gender, and dressing up; broadcasting place; broadcasting place movement. Through these scales, demographic characteristics of the broadcasters and other basic information of live streaming are considered in our analysis.

(ii) Scales to measure the interactivity of live streaming:

- *Number of feedback units*: the feedback broadcasters give to the viewers' comments and questions.
- *Number of (viewer's) questions*: the questions from viewers.
- *Number of (broadcaster's) answers*: the answers given by broadcasters.

- *Number of sequences*: the sequence changing times (according to the description in theoretical framework section).

(iii) Features of multimodal communication in live streaming: the modalities used in live streams. Specially, we focus on body movements in multimodal communication in live streaming. The main categories of body movements (based on 15 types of Allwood's study) and their functions are particularly noticed.

Besides, some other details not belonged to these three parts are recorded in the exception part. The findings are discussed in results and discussion section.

Taking the 18th live streaming videos in China as an example, we coded it as following:

(i) Basic information of a live streaming video:

Serial Number: A18; *Category*: Music; *Content*: Singing and chatting (asking for gifts and other random topics); *Platform*: Yizhibo; *Time (0:00-24:00)*: 2017/4/17 16:25; *Duration (mins)*: 23; *Number of viewers*: 136000;

Broadcaster's ID: ZDJ; *Gender (not sure=0, Female=1, Male=2, multi-gender=3)*: 1; *Age (not sure=0, 0-20=1, 20-30=2, 30-40=3, >40=4)*: 2;

Place (Indoor=1, Outdoor=2): 1; *Place (specific)*: Bedroom; *Place changing (change=1, not change=2)*: 2; *Dress up (exposed=1, unexposed=2, not sure=3)*: 3;

(ii) Scales to measure the interactivity of live streaming:

Number of Feedback units: 68; *Number of Questions*: 22; *Number of Answers*: 8; *Number of Sequences*: 10;

(iii) Features of multimodal communication in live streaming:

Multimodal Communication (sending information): Face, gesture, voice; say and sing "I love you" while wink and smile;

Bodily Movements: Wink, movements of arms and hands (finger point at the camera), facial gestures (laugh);

(iv) *Exceptions*: Broadcaster uses laugh as feedback, but the audience can't know which comment she laughs; the audience interact with each other frequently; broadcaster reads

the comment out but don't know its meaning (several times); someone asked "how is your lips?" the broadcaster doesn't reply, but wipes her lips.

4.1.2 Types of Live Streaming to Be Observed

After constructing the observational framework, we started selecting the live streaming videos by classifying the types of live streaming to be observed in this study.

The types of live streaming are decided by combining types on different live streaming platforms. We summarize the classifications and find that, in general, they include 9 types of live streaming: chat, talk show, music, food, travel, sports, game, news, and activity.

"Chat" refers to talking randomly without a specific topic, while a "talk show" has a specific topic and usually has only one topic. "Music" refers to singing, dancing, or playing instruments by broadcasters. "Game" live streaming is usually showing the skills of playing computer games in real-time. "Travel" live streaming is about attractions or showing sceneries during travels. "Activity" is live streaming broadcasting events or activities like concerts, meetings, press conference etc. "Food" refers to making food or teaching cooking. "Sports" live streaming is usually about broadcasting sport games and fitness.

Our observations cover all these 9 categories in both two countries. During observations, in total, we watched 106 live streaming, 54 of USA (8 chat, 8 talk show, 8 music, 4 food, 6 sports, 6 travel, 6 games, 4 news, 4 activity) and 52 of China (8 chat, 8 talk show, 8 music, 4 food, 5 sports, 5 travel, 6 games, 4 news, 4 activity). Each live streaming videos was watched for 20 to 30 minutes, the total length is 2251 minutes (1088 minutes of American live streaming, 1163 minutes of Chinese live streaming).

Live streaming from China and USA in this study means that the broadcasters are Chinese or Americans. During studies, in order to confirm the nationality of broadcasters, we referred to their Twitter, Facebook or other information showing the nationality of the broadcasters. This provides a better basis for a comparative study of China and the USA.

4.2 Data Analysis

Qualitative and quantitative analysis are both used to analyze the data.

Microsoft Excel 2013 and IBM SPSS Statistics 23 are used in quantitative analysis. By inputting data into Excel, we used IBM SPSS Statistics 23 to process the quantitative data. All of quantitative data form 106 samples in this study are valid to analyze in statistics. Z-score is used to process data as standard scores. Then SPSS is used to check the normality of data in continuous variables of both countries. Some data not according with normal distribution are processed into normal distribution. After normal processing, data in accordance with normal distribution can be analyzed by Cluster analysis and Pearson correlation analysis. Then we examine the reliability and validity. Cronbach's alpha of the 106 samples data of the 6 scales measuring interactivity (Number of feedback units, number of questions, number of answers, number of sequences, feedback per minute, and answer rate) are higher than 0.7 in our study. This means the data in this study has acceptable reliability.

In addition, when doing quantitative analyses, we added two calculated quantitative scales. The calculation of these scales is as following:

- *Feedback per minute*: Divide number of feedback by duration (minutes).
- *Answer rate*: Divide number of answers by number of questions.

This is because each live streaming has its own observational duration. Feedback is very sensitive to observational duration. So we calculate feedback per minute as a new variable. Answer rate are also calculated to measure the interactivity of live streaming. Qualitative analysis in this study is mainly combining the features of live streaming with the relevant theories. ICM theory is used to analyze features of interactivity in live streaming which we recorded in study. Multimodal communication theory is used to analyze characteristics of multimodality in live streaming communication. For instance, we analyze traits of body movements in live streaming by connecting some features we recorded with Allwood's 15 types of body movements.

4.3 Ethical Consideration

During the data collection and analysis processes, there are ethical considerations we have to take note of this study. First, the live streaming we observed is open to all users. We did not observe live streaming only friends can see. Second, to protect the privacy of broadcasters, all personal information in the study is kept confidential. All names of broadcasters are anonymous or mentioned with nicknames during analyses. Third, all the data is just used in academic research and will not be used in business.

5. Results and Analyses

In this part, demographics of broadcasters, broadcasting place and place movement, numbers of viewers, interactivity and multimodal communication in live streaming are studied through analyzing the observational data. Comparisons of China and the USA are also demonstrated in this section.

5.1 Overview of the Observation Data

During April, 2017, 106 live streams are observed in both China (52) and the United States (54) in this study. We observed 20 to 30 minutes per live stream in most live streaming, 2251 minutes (37.52 hours) in total. The live streams we observed are picked randomly but trying to cover as many types of live streaming as possible.

5.1.1 Demographics of Broadcasters

		Age			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Not sure	26	24.5	24.5	24.5
	0-20	3	2.8	2.8	27.4
	20-30	48	45.3	45.3	72.6
	30-40	16	15.1	15.1	87.7
	Over 40	13	12.3	12.3	100.0
	Total	106	100.0	100.0	

Table. 2 Age distribution of broadcasters in all samples

In the 106 samples, there are 62.3% male broadcasters, a little bit more than female broadcasters. We also noted the age of broadcasters by the same way as confirming their nationalities, through their talks in live streaming or their social platforms, such as Facebook and Twitter. As for the age distribution (Table. 2), 45.3% of broadcasters are young people, between 20 to 30 years old. If we remove 26 samples of uncertain age, there are 60% of broadcasters from 20 to 30 years old. Young people show higher

receptivity for live streaming than elder people. Children may not have that much time and freedom for using live streaming. People who are 20 to 30 years old are at age to explore and expand the social network. Also, they usually do not have a family to care of, living alone. So they are free after school or work. Because of these reasons, people from 20 to 30 are the main groups of people to broadcast themselves. This demographic feature appears to both China and America.

5.1.2 Place Broadcasting

			Place	Place movement
Spearman's rho	Place	Correlation Coefficient	1.000	-.726**
		Sig. (2-tailed)	.	.000
		N	106	106
	Place movement	Correlation Coefficient	-.726**	1.000
		Sig. (2-tailed)	.000	.
		N	106	106

** . Correlation is significant at the 0.01 level (2-tailed).

Place (Indoor=1, Outdoor=2) , Place movement (move=1, not move=2)

Table. 3 Correlations between place and place movement

Viewing all 106 samples, 76.4% broadcasters (81) broadcast indoor. And 84% broadcasters (89) just stay in one place, only 16% of them keep moving and changing places during broadcasting. In order to test whether place they broadcasting and place movement are correlated, we use Spearman's rho to test these two variables. As Table. 3 shows, the correlation is significant at the 0.01 level (2-tailed). And the correlation coefficient is -0.726 ($|-0.726| > 0.7$), which means they have strong correlation. Number 1 of place represents “indoor” and 2 represents “outdoor”, but number 1 of place movement represents “move” and 2 represents “not move”, so the correlation coefficient is minus. The correlation means broadcasters who broadcast indoor usually do not move to another place, and broadcasters who broadcast outdoor usually change place during broadcasting. This result is based on the whole 106 samples, so it's valid

in both China and the United States.

Based on observations, several reasons are concluded to explain this phenomenon. First, broadcasters can connect WIFI indoor, but they need to use 4G outdoor. The WIFI signal outside is not that stable and fast as inside WIFI. And 4G sometimes costs a lot of telephone charges. Second, indoor environment is much better than outdoor, less noise, less sunshine reflecting light and less other interference factors. In a better environment, broadcasters can focus more on broadcasting and give more feedback.

For example, Jie (from America) was broadcasting on Brooklyn Bridge on a shining day wearing a pair of sun glasses, she walked slowly because it was crowded on the bridge. She cannot reply immediately and even cannot see all the comments and questions. After a while, she apologized:

“I can’t reply to you sometimes and I can’t see your words clearly because of the sunshine and I’m walking. Oh, my god! It’s too crowded today.”

There are the different categories of live streaming, and we found that the live streams broadcasted outdoor are usually about travel, news and activity. In these three kinds of live streaming, broadcasters need to walk around to broadcast what is happening at the time.

5.1.3 Number of Viewers

The number of viewers in China and the United States is very different. On average, there are 859.74 viewers per live stream in USA, and there are 371281.69 viewers per live stream in China. This is not only explained by the difference in population. The population of China is 4.31 times of America, but the number of live stream viewers is 432 times greater than in the USA.

This is a big difference between the viewers’ in the two counties. The explanation was be given by commercial interests and psychological needs. Yi Jia (2016) described the core of live streaming as interpersonal interaction under the commercial interests. Aiqing Yuan and Qiang Sun (2016) warned that as a commercial force driven product,

live streaming still keeps the purpose of pursuing interests.

According to the Statistical Report on Internet Development in China (2017), there are more than 200 live streaming platforms and 344 million users (47.1% of Internet users of China) of live streaming till January of 2017. As some big American Internet companies joined the live streaming market in the United States, several large Chinese Internet companies also joined the Chinese live streaming market. The companies know the psychological needs of the consumers well.

Different from most of American living streaming platforms, most Chinese live streaming platforms provide high-level users lots of privileges, such as different titles and colors of their usernames. So broadcasters can easily recognize the high-level viewers' comments from the huge number of comments. A quick way to become a high-level user is sending gifts to broadcasters. And all of the gifts are virtual and only sold on the platforms they use. After broadcasting, broadcasters can get some money from platforms according to the gifts they get. In this way, platforms sell gifts to viewers and share incomes with broadcasters, so companies holding these platforms earn a lot. Viewers who send gifts can get the special effects on screens and a "*Thank you dear!*" from broadcasters with inner satisfaction.

Besides commercial and psychological reason, technological improvement also helps the development of live streaming. In 2016, there were 469.2 million users using mobile payment in China. This convenient payment method helps viewers buy gifts whenever they want.

As for broadcasters, broadcasting becomes a new occupation with hundreds of thousands RMB per day in China (Yi Jia, 2016). After it changes to an occupation, broadcasters need to attract viewers and fans because they live by broadcasting. But most of broadcasters in America still broadcast for fun or just killing time.

All of the above are main differences between China and the United States, and cause the huge difference of number of viewers in these two countries.

5.1.4 Interactivity in China and the USA

Descriptive Statistics					
	N	Minimum	Maximum	Mean	Std. Deviation
Number of Feedback units	52	0	138	36.58	28.329
Feedback per minute	52	.00	4.60	1.6420	1.21472
Number of Questions	52	2	166	25.60	28.395
Number of Answers	52	0	37	10.31	8.264
Answer Rate	52	.00	1.00	.5195	.31263
Number of Sequences	52	0	94	20.42	22.346
Valid N (listwise)	52				

Table. 4 Descriptive statistics of interaction data of China

Descriptive Statistics					
	N	Minimum	Maximum	Mean	Std. Deviation
Number of Feedback units	54	0	100	26.37	26.968
Feedback per minute	54	.00	4.15	1.3369	1.31711
Number of Questions	54	0	81	18.67	20.805
Number of Answers	54	0	59	7.81	11.783
Answer Rate	54	.00	1.19	.3792	.35211
Number of Sequences	54	0	56	12.57	14.458
Valid N (listwise)	54				

Table. 5 Descriptive statistics of interaction data of USA

In the last part of data overview, 6 significant factors given below are discussed and compared in China and the USA. The 6 factors are: Number of feedback, feedback per minute, number of answers, answer rate and number of sequences (decided mainly by broadcasters), while the number of questions (mainly decided by the viewers). Combining these two aspects, interactivity can be measured comprehensively.

From Table. 4 and Table. 5, concerning all the 6 factors, China has higher mean than the USA. For example, the mean of "number of feedback units" in China is 36.58, while it is 26.37 in the USA. Standard deviations are similar in both countries except the number of questions and the number of sequences that are different. The standard deviation of "number of questions" in China is 28.395, while it is 20.805 in the USA. Comparing the max value of the two factors in both countries, standard deviations are

influenced by the extreme value. For example, the maximum number of questions in the USA is 81 while it is 166 in China. This extreme value is much higher than its mean value, so the standard deviation of "number of questions" in China is much higher than its in the USA. In the other four factors, they have similar standard deviations, so they have similar dispersion degrees. Assuming that 54 samples from the USA and 52 samples from China are representative, we can say Chinese live streaming was stronger interactivity than live streaming in the USA.

The reasons for the differences in interactivity need to be analyze data. First, in China, there are many more viewers than in the USA. From this aspect, more viewers can interact with broadcasters. They may comment more and ask more questions. But high interactivity of viewers not equals to high interactivity of broadcasters. Broadcasters need motivation to interact with viewers. In USA, broadcasters are the center of live streaming, most of them broadcast for enjoy themselves. However, in China, some of the broadcasters broadcast for commercial benefits, some even make a living by broadcasting. So viewers are the center of broadcasting. Broadcasters need to attract viewers' attention. This makes them more interactive. And if viewers feel they are concerned, they may stay and become fans of the broadcasters. Giving feedback is the only way broadcasters have of making their viewers feel concerned. So from the broadcasters' aspect, Chinese broadcasters have stronger motivation to give feedback. After getting feedback and seeing others getting feedback, viewers may have more willingness to comment or ask questions. All of these reasons can explain why Chinese live streaming seems to have higher interactivity than American live streaming.

5.2 Interactive Communication Management

Communication is based on the interaction of the interlocutors. Different patterns of communication have different ways to manage interactions. Interactive communication management (ICM) provides a framework to analyze the interaction of communication. ICM has three subsystems, turn management, feedback and sequences. In this section,

we go through these subsystems and analyze the patterns of ICM in live streaming.

5.2.1 Turn Management

In live streaming communication, there are usually one broadcaster and many viewers and broadcasters seems have absolute authority in the communication in live streaming. This is different in comparison with face-to-face communication, where all participates may get turn. Here we will study three issues of turn management: who can get turn, how do they get turn, and who really control the turn.

Correlations

		Number of question	Number of answer
Number of question	Pearson Correlation	1	.699**
	Sig. (2-tailed)		.000
	N	54	54
Number of answer	Pearson Correlation	.699**	1
	Sig. (2-tailed)	.000	
	N	54	54

** . Correlation is significant at the 0.01 level (2-tailed).

Table. 6 Correlations between question number and answer number in USA

Taking live streaming in USA as an example, using the method of Pearson correlation (Table. 6), correlations between question number and answer number is significant at the 0.01 level (2-tailed) and the correlation coefficient is 0.699. Processing data of China can get the same results. In both countries, question number and answer number are correlated.

To explain these correlations, we cannot just analyze from one side. Questions and answers promote mutually and co-construct the whole interactive communication. Specifically, viewers raise questions and broadcasters have questions to answer. But if broadcasters do not answer any questions, viewers probably will not like to ask any more. The interactivity may decrease to a low level or stop totally. So answering questions also leads more questions coming up. Therefore, answering question is

necessary and important for broadcasters. For example, when broadcasting activities, there were comments and questions at first. But if the broadcaster did not give any feedback, then “there was silence in the broadcasting world”. Even if some new comers may give comments continuously, they soon keep silent, too.

Therefore, broadcasters and viewers are the co-constructors of the interaction. So viewers can get the turn from raising questions and comments. And broadcasters seem to hold the turn all the time, except when viewers are commentating or asking. Broadcasters need to read at such a moment, so viewers actually get the turn then. Broadcasters can get the turn back from choosing questions and comments, and then give feedback.

Correlations

	Zscore: Feedback per minute	Zscore: Answer Rate
Zscore: Feedback per minute	1	.607**
Pearson Correlation		.000
Sig. (2-tailed)		
N	52	52
Zscore: Answer Rate	.607**	1
Pearson Correlation		
Sig. (2-tailed)	.000	
N	52	52

** . Correlation is significant at the 0.01 level (2-tailed).

Table. 7 Correlations between feedback per minute and answer rate in China

Taking live streaming in China as an example, from Pearson correlation in Table. 7, correlations between feedback per minute and answer rate is significant at the 0.01 level (2-tailed) and the correlation coefficient is 0.607. Processing the data of USA can get the same results and even higher correlation coefficient.

In both two countries, feedback per minute and answer rate are correlated. If broadcasters want to communicate with viewers, the feedback per minute and answer rate result are in high scores. But if broadcasters have less willingness to communicate and do not want to give feedback, the two factors can be low. Broadcasters have the absolute power in live streaming. They can choose whether to give feedback or answer the questions.

To conclude, broadcasters are powerful interlocutors who control the turn. They can get the turn by starting a new topic, giving feedback and answering the questions they want. But they lose turn when they see the comments and questions without talking. Viewers can get the turn by asking questions or giving comments.

5.2.2 Feedback

Feedback processes help interlocutors communicate successfully, making sure that they have contact, perceive and understand each other's emotional-attitudinal reactions and contributed content (Allwood, 2013). Viewers use comments and questions showing whether they can receive information from broadcasters. Sometimes sending gifts also can be interpreted as feedback. The feedback ways are simple for viewers. Broadcasters can only get the information from these ways. Broadcasters give gestures or answers to comments and questions as feedback. Their feedback is multimodal. Gestures like nodding, facial gestures (smiling, frowning etc.), moving close to camera can be interpreted as feedback in body movements. Answering to comments and questions is the verbal way of feedback. Emotional-attitudinal reactions of broadcasters can be found in both gestural and verbal ways.

Broadcaster sometimes may ask for feedback from viewers. During live streaming, broadcasters just face the cameras and see the comments and questions rather than the viewers' face. Broadcasters may feel insecure if viewers do not give feedback, they even ask if the network is broken down. So sometimes broadcasters may ask for feedback. For example, an American policy broadcaster Jos said:

“Can you see me? Can you hear me? Raise your hands if you hear me. Raise your hands! Raise hands!”

Viewers comment *hands emoji* to give feedback.

“Ok! We can keep going on then. What I just said...”

Viewers' feedback in live streaming can be ignored, interrupted, delayed, and vague (lack of understanding). Viewers give lots of comments and questions, broadcasters can

only reply to some of them, especially what they are interested. In this situation, some comments and questions can be ignored. Broadcasters (from China and the USA) often ignore some rude and aggressive comments, they even block the users who give those comments. Blocking (the blocked user cannot comment in the live streaming anymore) can be interpreted as a special feedback.

From the previous study, the share of interrupting feedback shows if the participants interrupt each other frequently, e.g. because the interaction is fast (Ahlsén et al, 2003). The various speed of interaction in different groups of live streaming shows a big difference between China and the USA. We use cluster analysis to discuss the differences in detail. In general, interactions in live streaming are fast. Because usually there is just one broadcaster communicating with lots of viewers simultaneously. Comments and questions are coming all the time during broadcasting, so broadcasters have to read them while talking. Sometimes they are attracted by comments, so their talking can be interrupted. As for delayed answers and vague feedback in live streaming, they are discussed in detail in discussion section.

Our analysis of interactivity in China and USA above shows that interactivity in China is higher than interactivity in USA. If we analyze this more in detail by comparing different types of live streaming in the two countries, there are obvious differences and similarities between China and the USA. In order to compare the interactivity of different categories from China and the USA, we use cluster analysis to divide live streaming video into different categories based on degree of interactivity. The cluster is also based on the 6 factors used in the descriptive comparisons. The results of cluster analysis are introduced respectively. As mentioned, these 6 factors are divided into broadcaster interactivity factors and viewer produced interactivity factor. The broadcaster factors measure feedback from broadcasters' side, while the viewer factor measures feedback from viewers' side. So we can describe feedback more in-depth levels from both sides by analyzing this cluster analysis.

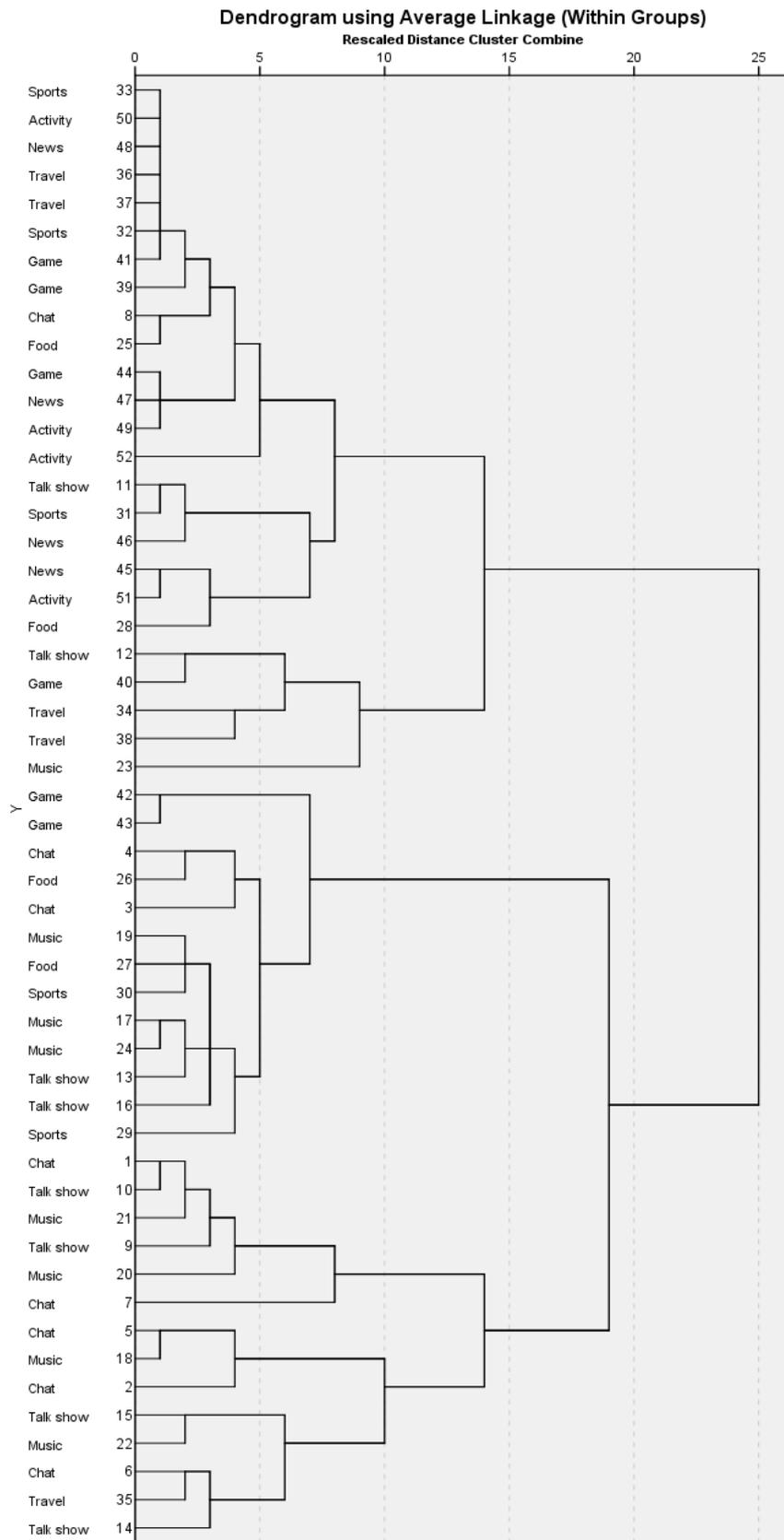


Table. 8 Dendrogram of live streaming in China

According to the dendrogram in Table.8 (also the cluster membership in Appendix 3),

we divide all 52 live streaming videos into three groups. In group 1, chat and talk show are the main categories. Music and food appear much in group 2. And sports, travel, news, game and activity belong to group 3. By analyzing the categories in different groups, we come up with three names of the groups. Chat and talk show are all about talking. Broadcasters keep talking to the viewers. So we call group 1 talking live streaming. Group 2 mainly consists of music and food. Broadcasters display their singing, dancing, instruments playing, and food cooking skills. Besides, they also pay much attention to talk to the viewers. So group 2 is called display-talking live streaming. Group 3 consists of sports, travel, news, game, and activity five categories. In these five categories, broadcasters focus on display rather than talking and there might even be no talking in the whole broadcasting. So group 3 is called display living streaming. 3 groups of live streaming appear different degrees of interactivity. Group 1 has the highest interactivity, group 2 is in medium level, and group 3 interacts less than other groups. The different levels of interactivity relate to the categories of live streaming.

Group	Original categories	Group name
1	Chat, Talk show	Talking live streaming
2	Music, Food	Display-talking live streaming
3	Sports, Travel, News, Game, Activity	Display live streaming

Table.9 Groups of live streaming in China

Talking live streaming is focused on talking. More accurately, it is focused on the talking of broadcasters and comments and questions from the viewers with high level of interaction. Some of the broadcasters even do not have specific topics, just choose the questions and comments which they want to answer. In this situation, viewers actually hold turns and partly control the topic of live streaming. When broadcasters have their own topic for the broadcasting, they keep talking but can easily be interrupted by the feedback from viewers. In talking live streaming, the viewers are almost the only focus of broadcasters. Broadcasters keep interacting with viewers. And because of the high dependence of viewers, feedback in talking live streaming has the highest level in

all three types of live streaming.

In live streaming, broadcasters cannot see the faces of viewers like face-to-face communication, they can just “see” the viewers by their feedback with their usernames. The feedback from viewers represents the viewers by showing their attitudes and thoughts. Feedback from the broadcasters said is the main way for them to know their viewers. Broadcasters give feedback to keep interaction with viewers and keep live streaming going on.

Broadcasters of display-talking live streaming have more foci than talking live streaming. Broadcasters focus on displaying their music and food making procedures to viewers while answering related questions and replying comments. When broadcasters are displaying, they usually do not have time to interact with viewers, even cannot see the feedback from viewers. In order to solve this problem, broadcasters usually stop their display and spend some time replying to the questions and comments from viewers and then come back to display their work. Mixing displaying and replying for every several minutes help broadcasters get feedback and give feedback. Sometimes, broadcasters can interact with viewers while displaying. But this distraction usually slows their work down or causes a little pause during working. Some singers use smiles to give feedback while singing during broadcasting, but viewers cannot know which comments the broadcasters are replying. To conclude, display-talking live streaming has a medium level of interactivity and less feedback than talking live streaming.

Display live streaming is mainly focused on displaying rather than interaction and even no interaction. This group have categories of game, travel, news, sports, and activity. When broadcasters broadcast these things, they show the skills of playing computer games, great sceneries, situations of news, sports games and activities in real-time. Some broadcasters may have interaction with viewers where there are pauses during activities. Some broadcasters do not give any feedback during the whole observation periods. So, display live streaming have the lowest interactivity with no feedback sometimes.

After analyzing the three groups of live streaming in China, we do the same cluster analysis of the data in USA and we get the cluster membership shown in Table.10. The

factors and the methods are same as China, so we have same standards in both countries.

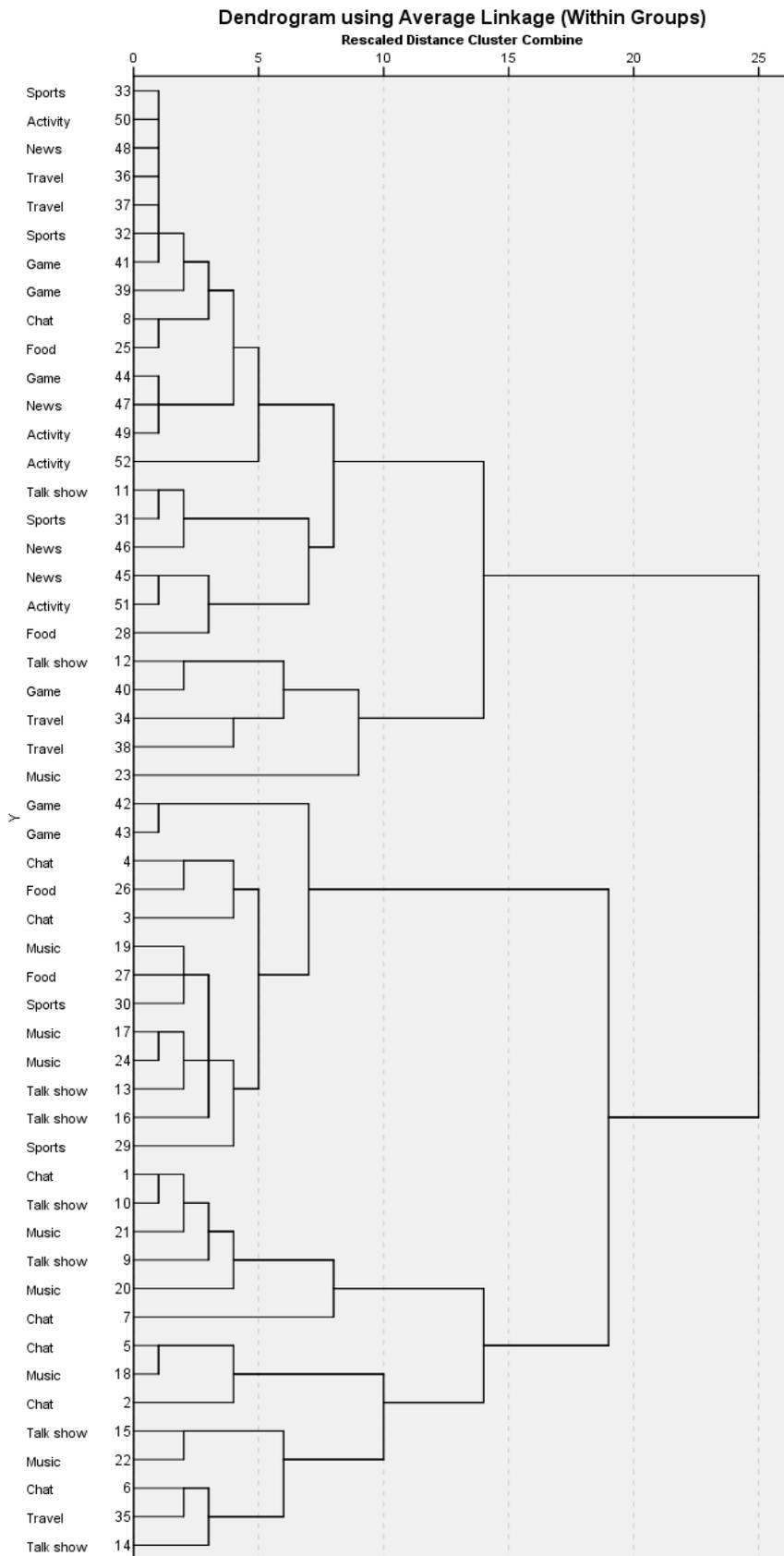


Table. 10 Dendrogram of live streaming in USA

According to the dendrogram (also the cluster membership in Appendix 2) above, we divide live streaming video into groups. If we divide into the same 3 groups as China, there are only a few samples are in group 2. The group 2 has a few representative, and this loses its meaning. So we divide the samples into 2 groups. As Table.11 showing, group 1 has chat and music and group 2 consist of food, talk show, sports, travel, news, game, and activity. Combining features of the data, we name group 1 as two-way interaction live streaming, and group 2 as one-way interaction live streaming.

Group	Original categories	Group name
1	Chat, Music, Interactive travel	Obvious two-way interaction live streaming
2	Food, Talk show, Sports, displaying travel, News, Game, Activity	Obvious one-way interaction live streaming

Table.11 Groups of live streaming in USA

Group 1, as its name, shows obvious two-way interactivity in live streaming. In most of the broadcasting time, broadcasters pay attention to viewers and give feedback all the time. Different from the talking live streaming with the highest interaction in China, music and interactive travel are included in highest interaction group. As mentioned, viewers in China are over 400 times than viewers in USA. Broadcasters in China seldom meet situations of no comments and questions to give feedback. But American broadcasters often face this situation, they cannot keep interact with the viewers without pauses. So during the same time period, Chinese broadcasters have more feedback, more questions to answer and answers, and a higher answer rate. And even if American broadcasters broadcast for chatting, because of the limited comments and questions from the viewers, the number of feedback, questions, answers, and answer rate are limited. So the data of interactivity of chat are similar to the data of music and interactive travel. When broadcasters play the instruments or show the sceneries, sometimes, they are also waiting for the comments and questions. So in the group of obvious two-way interaction, the interactivity is highest in USA but lower than in talking live streaming of China.

Group 2 contains 7 categories, food, talk show, sports, displaying travel, news, game, and activity. These live streams show obvious one-way interactivity characteristics. Broadcasters keep talking or showing viewers about their skills, sceneries, and activities but seldom give feedback or do not give feedback. Broadcasters have a low dependence on their viewers. So in the group of obvious one-way interaction, the interactivity is low in all 7 categories.

Sometimes in news, sports, and activities live streaming, broadcasters cannot see the comments and questions. They may just put cellphones somewhere and letting it broadcast. When the affairs they want to broadcast are finished, they press the button to end the live streaming. Live streams about food and displaying travel focus on displaying the food making procedures or the sceneries. Broadcasters pay more attention to what is displayed than interaction. Broadcasters broadcasting playing computer games have no interaction with viewers because they are just showing their skills to the viewers. Most of the game broadcasters are professional broadcasters, and they make a living by playing games. Broadcasting is one way for them showing their skills and they can also get money by support from the viewers. The games are often very fierce, so they need to focus on the games and have no time to interact with viewers. Viewers always discuss with each other by using the comments function in this situations. The number of viewers of games are extremely large and the interaction between viewers to viewers keeps all through the broadcasting. In China, game broadcasters have similar characteristics as American broadcasters, but they sometimes interact with viewers.

The biggest difference between China and USA in interactivity appears in the talk show. In China, the talk show is in the group with the highest interactivity. But in USA, the talk show is in the group of obvious one-way interaction live streaming, which has low interactivity. American talk show broadcasters share their own opinions to viewers more, so they have limited time to give feedback to viewers. Viewers also express themselves less. But in China, talk show broadcasters prefer to discuss the topics with the viewers. They keep receiving feedback and giving feedback about the topic. So talk shows in China are co-constructed by broadcasters and viewers. Therefore, talk shows

in China have the highest level of interactivity, while talk shows in USA have a low level of interactivity. The difference can be partly attributed to cultural influence. According to Hofstede, the culture of the United States emphasis individualism, while Chinese culture is more collectivism. So American broadcasters may tend to show themselves more. Chinese broadcasters may prefer to discuss with their viewers. This could be the reason of the difference of interactivity in the two countries.

5.2.3 Sequence

The sequence in live streaming usually consists of a comment or a question and one feedback. One sequence in living streaming consists of one or several interactions of broadcasters and viewers of the same topic. Sequences are usually short, but the number of sequences are large, which are the main features of sequences in live streaming. Sequences show strong similarity of live streaming in China and the USA. These features of sequences are valid in both countries. Because the similar sequences or exchange types are mainly depended by the new technology. Culture differences do not have obvious influence in sequence of live streaming. So in this section, we analyze these common features of sequences in both China and the USA.

Broadcasters face tens to hundreds of thousands of viewers during live streaming. When many viewers participate in the live streaming, broadcasters need to reply to the comments and questions. Every feedback to a comment/question and the comment/question consist a whole sequence or a particular exchange type. Because broadcasters seldom wait to communicate with the same viewers to continue their previous sequence. So there are usually many sequence.

In most situations, broadcasters give feedback without following up with a question to the viewers. But when few viewers comment, broadcasters may have several questions-answers as one sequence. If broadcasters ask back to the viewers when they replied, they may wait for viewers' answer. But speaking is usually faster than typing into words, so from viewers receiving questions to finishing typing and sending out, there is still some time. During this time, broadcasters usually wait while answering other questions

or replying to other comments.

Mostly the sequence between broadcasters and viewers is one question-answer or one comment-reply. Because most sequence are short, broadcasters can reply many questions and comments in short duration. And when broadcasters really need to interact with viewers, they may reply briefly and then move to the next comments or questions. Every viewer has his/her own point of view, so the questions and comments can be quite different. Thus, broadcasters needs to pay more attention on understanding and replying to them. Also, viewers come and go quickly, as for new viewers, they know nothing about the topics and the broadcasters, so they may ask some repeated questions. So as we can see, broadcasters have lots of questions and comments to reply to. So the number of sequences in live streaming is large.

5.3 Multimodal Communication in Live Streaming

According to Allwood (2008c), multimodal communication = co-activation, sharing and co-construction of information simultaneously and sequentially through several modes of perception (and production). The basic feature of multimodal communication is employing several modalities of production and perception in order to share information. In this section, we analyze the specific mode of multimodal communication in live streaming, and then focus on body movements in live streaming. Flexibility in the choice of modality is also analyzed. Multimodality used in different contents of live streaming is claimed. Comparisons between China and the United States are also demonstrated.

5.3.1 Specific Multimodal Communication in Live Streaming

During live streaming, broadcasters can use more than one modalities/channels to broadcast and communicate with the audience, while the audience can only use limited ways to communicate with broadcasters. This unique model of communication in live streaming builds its specific multimodal communication patterns. It can be divided into

two parts:

(i) For broadcasters, in live streaming, they are both information senders and receivers. When producing information, they can use more than one channel. Basically, they use auditory aspects of speech, singing or non-linguistic voice, combined with communicative body movements (facial gestures; head movements; hand and arm movements; eyes movements; lips movements, etc.). For instance, one American female broadcaster chatting with viewers about fitness, shook her head saying “No” and then made non-linguistic sounds to show her disagreement with some comments. In addition, internet and live streaming technologies allow them to use more channels to send information, e.g. computer game interface, videos of conference scene, videos of concert scene, videos of football game, music, graphics, etc. For example, most gaming broadcasters only use auditory aspects of speech and game interface to send information.

On the other hand, broadcasters are also information receivers during communication in live streaming. However, compared with sending information, broadcasters perceive information mainly by eyes/visual sense. Because most of the effective information from viewers is text-based comments and heart-shaped images. In China, broadcasters can also get gifts sent by viewers.

(ii) For the audience, as the information receivers, they perceive information mainly by visual sense and auditory sense. All information they get is through screens of computers or smart phones. Therefore, they depend on eyes and ears to receive what broadcasters send to them. Meanwhile, viewers are also information senders. As we mentioned before, they can type text-based comments/questions to have communication with broadcasters and tap hearts, or send gifts (in China) to show love. In general, broadcasters use both the primary modalities produced by human bodies and the secondary channels supported by internet and live streaming technologies to produce information. They perceive information by only visual sense. However, viewers use visual and auditory sense to get information from broadcasters, use text-based comments and heart-shaped images to send effective information. Therefore, it can be described as: broadcasters send information with multimodality, while perceive

information with single modality. However, viewers send information with texts and images, while they perceive information with visual and auditory sense.

The main reasons for these characteristics lie in live streaming technologies. Live streaming technologies do not allow users who are broadcasting to use its typing function, nor allow users who are watching broadcasting videos to use its camera function. This special restriction creates a unique communication mode in live streaming. This mode creates the specific multimodal communication in live streaming. Even if live streaming in China and America have some differences, the basic technologies are the same. Therefore, both China and America live streaming have the features analyzed above.

The influences of this specific multimodal communication are also significant. For broadcasters, sending information through more than one modalities, and seeing text-based comments on the screens can easily distract their attentions from what they are talking about. Text-based comments appear on the screens for a very short time and they are hard to be found. For those broadcasters who want to communicate with viewers, they must read text-based comments/questions very quickly to know what viewers are saying. This process is really challenging for them. It can also cause very short answers or misunderstanding during live streaming (we discussed in the following paper). For viewers, text-based comments are not as direct /fast as speech, and their text-based comments can easily disappear from the screen. To draw attention of the broadcasters to see their comments, they always repeat sending same messages.

5.3.2 Body Movements in Live Streaming

According to Allwood (2002), movements of the body as primary means of expression in communication, they can be used together with speech and independently of speech. The use of body movements in communication is typically connected with simultaneous multidimensionality, both with regard to means of expression and functional content (Allwood, 2002). Body movements are important in multimodal communication. In this study, we focus on features and functions of body movements

in live streaming. The similarities and differences in China and the United States are also analyzed during analyses.

5.3.2.1 Main Types and functions of Body Movements in Live Streaming

In our study, we observed 52 live streaming videos from China and 54 from the United States. Considering that live streaming only allows the audience to use text-based information during the communication, we only focused on body movements from broadcasters. By recording the very high frequency and significant body movements of broadcasters, we found that during the live streaming communication, there are 6 main types of body movements that were used by broadcasters in these two countries. Based on the 15 types body movements described by Allwood (2002), we categorized the main types of body movements used by broadcasters as the following 6 types, each type of body movements is then followed by the functions the movement may have:

(i) Movements of arms and hands: finger points to camera to emphasis; hand moves close to camera to show things in hand; special hand gestures (heart-shaped gesture to show love; OK gesture to show agreement; thumbs up to show something is great; stop gesture to stop something; wave hands to greet).

(ii) Facial gestures: laugh; smile; grimace; happy. To show e.g. emotions and attitudes (Allwood, 2002).

(iii) Head movements: nod to show agreement or admit; shake one's head to disagree or deny.

(iv) Distance: move body close to camera to show closeness (Allwood, 2002) or the willingness to communicate.

(v) Eyes movements: wink to show attraction or happy mood; increased pupil size to show increased interest (Allwood, 2002) or surprise.

(vi) Lip movements: pout to show attraction or love.

All the movements above are the high frequency body movements in live streaming in our observation. Here is the general picture of body movements in live streaming including both these two countries. It is worth noting that functions explained above are based on our observation, they have relevant context when being used during the

communication. For instance, in our analysis, one broadcaster cooked a kind of special mushrooms in his video. He took some mushrooms and moved his hands close to camera several times during broadcasting. The main function of the hand movements in the relevant situation/context is showing things to the audience.

5.3.2.2 Features of Body Movements in Live Streaming

Body movements in live streaming have specific characteristics compared with face-to-face communication. Furthermore, there are also some similarities and differences in these two countries in this field.

(i) Body movements in live streaming mainly occur in the upper body of broadcasters. Both Chinese broadcasters and American broadcasters show this feature.

Except when filmed from distance, broadcasters move their bodies close to cameras, body movements like eyes movements, lip movements, head movements, arms and hands movements all occur in the upper body. This is different to face-to-face communication. According to face-to-face communication, besides the 6 types above, body movements also include movements of legs and feet, posture, spatial orientation, smell, touch, taste, etc. (Allwood, 2002). However, most of the body movements in lower bodies are not significant or with very low frequency in live streaming since they cannot be seen.

One reason for this feature is that those broadcasters who show body movements in live streaming mostly use front-facing camera (as we analyzed in the first feature) and need to keep a closer distance to camera to watch the audience's comments on the screen. When doing this, broadcasters usually show their upper bodies because of the limitation of the screen size. That is why the audiences usually see their upper bodies in live streaming. In order to get better communication results, most broadcasters choose to use body movements in upper bodies.

In addition, because of limitations of the live streaming technology, the existing technology can only allow broadcasters to shoot videos in order to share what they are doing. The help of internet and live streaming platforms cannot decrease the physical distance between broadcasters and viewers, so it cannot make them really touch, smell

or taste like in face to face communication. Therefore, broadcasters use these movements with very low frequency.

(ii) Some body movements are used by both Chinese broadcasters and American broadcasters, while some of them are used by either Chinese broadcasters or American broadcasters.

We summarized that there are 6 main types of body movements in our observation. Each type of them includes specific movements. Some of these body movements are used by broadcasters from both two countries. For instance, finger points to camera for emphasis; hand moves close to camera to show things in hand, are two body movements used with very high frequency in both countries. These features are based on the communication conditions/technology provided by live streaming. The screens between broadcasters and viewers are the most important channel for viewers to see broadcasters (broadcasters can only see text-based comments). When sending information, broadcasters must make sure what they want to show/emphasize can be watched on screens. Therefore, they put fingers or hands close to cameras to draw attentions of the audience. Moving the body close to camera is the same as above situations. Both Chinese broadcasters and American broadcasters move their bodies close to camera, on the one hand to see comments on screens, on the other hand to show willingness to communicate.

Smiles, nods, shaking one's head are also used by both Chinese broadcasters and American broadcasters, while these features could be more influenced by the cultural reasons. Smiles, nods, and shaking one's head have similar meanings in both Chinese and American cultures in most situations. Smile is used to send positive attitudes, a nod is used to show agreement or admitting, shaking one's head is used to show disagree or deny. Therefore they are often used by broadcasters from both two countries with similar meanings.

Besides body movements used by both of Chinese broadcasters and American broadcasters, there are specific hands gestures only used by Chinese or American. For instance, heart-shaped gesture (Figure.3) was only used by Chinese broadcasters. This special gesture is very popular in China among young people. The thumb and forefinger

together meant “money” in past. With time and culture changes, the meaning of this gesture has changed. The shape made up by the thumb and forefinger looks like a little “heart”, at very start, many young people in China use it to show love and appreciation to others. This heart-shaped gesture spread very quickly on the internet. Now, it has become a representative gesture in China to show love used by people of any age. According to Allwood (2002), cultural variation both with regard to means of expression and type of function is considerable for almost all of the mentioned types of body movements. The change in the meaning of this gesture also implies the cultural variation in China.

However, the stop gesture (Figure.4) was only used by American broadcasters in our study. American broadcasters used this gesture to stop viewers’ comments or questions. This hand gesture is a pretty direct signal to stop what is going. We did not found any Chinese broadcasters used this gesture during the live streaming. The reason for this can be cultural.

According to Allwood (1985), culture refers to all the characteristics common to a particular group of people that are learned and are not given by nature (1985). It includes four primary cultural dimensions: patterns of thought (common ways of thinking, includes factual beliefs, values, norms, and emotional attitudes); patterns of behavior (common ways of behaving); patterns of artifacts; imprints in nature. The patterns of thought and behavior play important roles when broadcasters are using body movements in live streaming. For Americans, they are more direct during communication to express their feelings or thoughts, Chinese, however, are more indirect/hesitated to show their attitudes. Chinese prefer to use implicit ways to express, especially when sending negative information. In addition, high context culture allows Chinese to understand each other even through implicit expressions, while Americans usually communicate more directly and explicitly to understand each other.



Figure.3 Heart-shaped gesture



Figure.4 Stop gesture

(iii) Body movements are used together with speech by both Chinese and American broadcasters in live streaming.

From Allwood (2002), body movements can be used both together with speech and independently of speech. Except in several obvious examples, e.g. one male broadcaster from China used his finger to point to his seat belt to answer viewers' questions (if you fastened your seat belts) without saying anything; one female broadcaster from America replied by tidying her hair according to the viewers' comment without speech. In our study, we found that broadcasters used body movements with speech in live streaming. Especially body movements were used with giving verbal feedback during the live streaming communication. From Allwood & Cerrato (2003), feedback is mostly expressed simultaneously by vocal/verbal and gestural means. Our study also displayed this finding.

Broadcasters use many hand gestures, facial gestures and head movements with verbal/vocal feedback. For instance, except heart-shaped gesture is used with saying "Thank you!" or "Love all of you" by Chinese broadcasters; stop gesture is used with saying "Stop it" by American broadcasters, some body movements are used with speech in live streaming by both Chinese and American broadcasters. OK gesture is used with saying "OK, will do that" or "OK, wait a minute"; thumbs up is used with saying "Great" or "Good"; nodding with say "Yes/Yeah" or "Agree"; shaking one's head with stay "No" or "Don't think so"; smile with giving positive verbal feedback; grimace face is used with strange sounds to give negative feedback, etc.

According to the 4 ways concerning how gestural feedback expressions modify the meaning of the vocal/verbal expression (reinforcement; positive; negative; contradicting) proposed by Allwood and Cerrato (2003), body movements used with speech in live streaming mainly reinforce or give positive reinforcing attitudes during the communication. Most above examples give the same information as the vocal message to reinforce. For example, some of Chinese broadcasters use heart-shaped gesture showing enthusiasm. And from our analysis, there were few gestural feedback expressions used with speech in live streaming to indicate the lack of enthusiasm or interest or to denote sarcasm, irony.

Live streaming allows broadcasters to use more than one modality during the communication with viewers. Body movements in live streaming with speech, especially with vocal/verbal feedback, shows that they are one of the major source of the multimodal communication in live streaming.

5.3.3 Flexibility of Multimodal Communication in Live Streaming

In many contexts, flexibility in the choice of modality is needed. Multimodality gives us this flexibility, and also the possibility of being redundant when this is needed, for example, in a complex noisy environment (Allwood, 2013). In face-to-face communication, interlocutors' choices of modality have higher flexibility. Whether they send or receive information, there is no limitation about which kind of modality must be/cannot be used during communication. However, in live streaming, as we described before, broadcasters can use multimodality to produce information, while they can perceive information only by one sense (visual). Viewers use their visual and auditory senses to get information, while they send information only by text-based comments and heart-shaped images. It is clear that live streaming has lower flexibility when users are choosing modalities. This is because the communication mode in live streaming is mainly decided by functions supported by live streaming technologies. The choices of modality are also dependent on the technology. However, in face-to-face communication, interlocutors has more freedom to build their own communication

mode by choosing different modalities.

5.3.4 Multimodality in Different Contents of Live Streaming

From our observations, broadcasters use different modalities/modality combinations when broadcasting different contents.

In general, when the contents of live streaming videos are about chat, talk show, music or food, broadcasters always use auditory aspects (speech, sing or non-linguistic voice) with videos shot by front-facing camera to produce information during live streaming communication. For instance, one of male Chinese broadcaster saying “*love all of you*” with heart-shaped gesture and smile to show appreciation to gifts from viewers. However, when contents are about sports, travel, news, activities or games, broadcasters mostly use auditory aspects with videos shot by rear camera. For instance, when one male broadcaster from America broadcast his travel in New York, he was shooting a street scene with rear camera while introducing it with speech. The biggest difference in using front-facing camera and rear camera is if viewers can see the broadcasters’ body movements of broadcasters.

For broadcasters who chat with viewers, give talk shows, share music or show food/cooking, etc. to viewers mostly use front-facing camera to show their faces in live streaming. Broadcasters using front-facing camera to show more willingness to communicate with viewers. Showing face during the communication displays their sincere attitudes and makes the audience feel more comfortable to give comments. Showing faces are often accompanied with relevant body movements during the communication. Because when broadcasters use front-facing camera, they can also see themselves, to be more natural, they always use some body movements.

However, when live streaming contents are about sports, travel, news or games, etc., broadcasters usually use rear camera to show activities they want to share, landscapes during their travels, sports competitions they are watching, or computer games they are playing, etc. The main purpose of these kinds of broadcasters is showing what they are seeing in real time. Rear camera helps broadcasters to achieve this. But it cannot show

broadcasters' face, so the audience cannot see if they have any body movements. Therefore, when broadcasting chat, talk show, music or food, broadcasters preferred to use this modality combination: auditory aspects (speech, sing or non-linguistic voice) and body movements with front-facing camera. While when broadcasting sports, travel, news, activities or games, they mainly use auditory aspects (speech, sing or non-linguistic voice) with rear camera. From our analyses, Chinese broadcasters and American broadcasters are similar in this part.

In conclusion, in live streaming multimodal communication, broadcasters send information with multimodality, while perceive information with single modality. However, viewers send information with two visual modalities (texts and images), while perceive information with two sensory modalities (visual and auditory sense). There are 6 main types of body movements being used in live streaming by both Chinese and American broadcasters. Flexibility in live streaming multimodal communication is lower than in face-to-face communication. Moreover, broadcasters broadcast different contents of live streaming with different multimodality combinations. Last but not least, from our analyses, we found that differences between Chinese and American broadcasters in using multimodality in live streaming are not obvious. Most of the time, they made the same choices of using multimodality combinations when broadcasting similar contents. The purposes of live streaming, technological limitations as well as some cultural factors can have influences on broadcasters' use of multimodality in live streaming.

6. Discussion

The above analyses reveal significant features of live streaming communication mainly from interactive communication management and multimodal communication. However, as an emerging and fascinating social medium, live streaming also shows some other features. In this section, we discuss these characteristics of live streaming in detail and analyze reasons for them. Differences between China and America are also discussed.

6.1 Delayed Answers from Broadcasters

During live streaming, some broadcasters did not answer some questions until these questions appeared on screens for four or five minutes later. From our analysis, except broadcasters who were chatting with the viewers, the phenomenon of delayed answers from broadcasters is very common in both Chinese and American live streaming. For instance, one male broadcaster from America, who sang while playing the guitar, answered one viewer's question, "Can you play *Hey Jude*?" after almost 5 minutes when this question appeared on the screen. The long-delayed answer is unusual in face-to-face communication, because delayed answers always decrease the willingness of communication between interlocutors, or can even result in the ending of the conversation. Moreover, this usually implies a less polite attitude during face-to-face communication. However, in live streaming, delayed answers still are very common and seems to be a normal situations in communication.

On the one hand, reasons for this phenomenon are relevant with broadcasters. Some of them delay answering questions because they are doing something else when questions appearing on screens. The male broadcaster above did not answer if he could play *Hey Jude* because he was singing and playing the guitar at that time. After finishing that song, he answered with "I can have a try."

However, we also found that some of broadcasters gave long-delayed answers to inspire the communication. In one Chinese female broadcaster's live streaming, for a while

there was no comment/question on screens. So she answered the question “*Where are you from?*” which appeared at the very beginning of the broadcast. Broadcasters did this usually when there were very few questions and comments on screens at some certain moments to inspire viewers’ enthusiasm for interaction.

In addition, we also found that most Chinese broadcasters usually gave apologies or some explanations about why they delayed to answer questions, while most American broadcasters did not do this. Different performances show different cultural backgrounds. In Chinese culture, “和 (he)” is an important factor, it means “harmony”. This traditional cultural factor leads Chinese to keep harmonious relationships with others. According to Hofstede, Chinese culture has a higher score on collective dimension. Chinese broadcasters’ explanations display that they want keep harmonious relationships with viewers. However, according to Hofstede, American culture has a higher score on individual dimension. For Americans, what “I” think or “I” feel is significant to them. They will not give extra explanation if they think there is no need to explain or delayed answers are not caused by them. Furthermore, most Chinese broadcasters’ can get incomes from broadcast, while American broadcasters are not driven by economic interests in the same way. Thus, harmonious relationships with viewers is much more significant to Chinese broadcasters. This can also explain their behaviors.

On the other hand, sometimes, delayed answers are caused by network latency. Smooth live streaming needs the high technical quality of the internet. This is very different from face-to-face communication. When a network latency occurs, viewers cannot see live streaming videos smoothly, meanwhile, broadcasters cannot read text-based comments or questions, either. Delayed answers caused by network latency cannot be attributed to anyone. That is the reason for why delayed or long delayed answers can be accepted by viewers in live streaming.

6.2 Questions Chosen by Broadcasters

In live streaming, when comments and questions are appearing on screens, broadcasters cannot control the numbers of questions or comments. In some broadcasts, which have thousands of viewers, or even tens of thousands of viewers, there are a large number of questions and comments every minute. Although broadcasters cannot limit numbers of questions and comments, they have the freedom to choose. Through our analysis, we found some features when broadcasters were choosing questions to answer.

Both Chinese broadcasters and American broadcasters tried to ignore questions which showing obvious aggressive attitudes. For instance, in a live streaming video of an American female broadcaster, who talked about losing weight by diet, she answered all questions on the screen at certain moment but ignored the question which said “So, *why you are still so fat?* ”; one Chinese male broadcaster talked about affective problem between couples, he ignored the question asked “*Can you broadcast after taking a shower?*”. Both these kind of questions, which implied aggressive attitudes, were ignored by broadcasters.

For American broadcasters, they mostly chose questions which were typed in a language they can understand. For example, one American male broadcaster who played piano said “*Just English Please!*” in his live streaming. This is because most live streaming platforms in the United States are used by people from all over the world. Some of them cannot or are not willing to speak or type English, but they can still see live streaming videos on these platforms. However, their comments or questions may not be understood by some broadcasters. That is why some of American broadcasters only answered questions which were written by languages they know. In China, platforms of live streaming mostly are used by Chinese or someone who know Chinese. Therefore, this phenomenon is rare in China compared with America. This feature demonstrates the significance of language for communication in live streaming.

In general, in terms of answering questions, broadcasters are given a lot of autonomy. They can choose which questions will be answered/ignored and when to answer. The choices of questions are influenced by cultural factors, personal factors, economic

factors, etc. That is one reason for why live streaming attracts more and more people as broadcasters.

6.3 Repetition of Comments from Viewers as Feedback

Repetition of comments from viewers as feedback in live streaming appears extensively both in China and USA. Broadcasters often read out the questions and comments without giving other feedback. For example, one Chinese game broadcaster, Facekaka, read out the comments and questions, then game attracted his attention, so he forgot to answer questions 2 times. This kind of feedback just has one meaning: the broadcasters have seen the comments or questions.

We can analyze reasons of the repeated feedback from the characteristics of live streaming. Broadcasters broadcast in real-time, they need to keep sending out information to viewers. If they stop talking and see comments and questions, the sudden quiet may make broadcasters and viewers feel embarrassed. Broadcasters try to talking while separating part of attention to check comments and questions. It is really hard for our brain to do so many things at the same time, such as receiving and interpreting the information, coming up ideas, organizing the words and speaking them out at the same time.

Broadcasters are often attracted by the comments, questions, and forget what they are talking. This special feedback can be used because of the same reason. Broadcasters read out the questions firstly before answering them, but attracted by another question, so the silence comes and the question are forgotten.

Another reason for the repeated feedback is developed from observational. Sometime when broadcasters finished reading questions and comments, they had obvious pauses with sounds like “eh...”, that may represents broadcasters are thinking. After the little pauses, they turn to another question or comment, and ignore the question or comment they read. In this kind of situation, broadcasters may have no idea to the questions or do not want to give replies.

From all above, there are different reasons for repeated feedback in live streaming.

Broadcasters just repeated the questions and comments without other replies. Both thought interrupted, focus attracted and no idea to answer lead to repetition of comments from viewers as feedback in live streaming.

6.4 Lack of Understanding

Lack of understanding occurs on both sides of live streaming and in both countries. Viewers cannot understand which comment or question is replied by the broadcasters. Broadcasters sometimes cannot understand the viewers' comments and questions. When broadcasters reply to comments and questions without first reading them out, viewers can be confused. Some broadcasters even use facial gestures (smile etc.) as feedback. Viewers cannot know which comment made the broadcaster laugh. Sometimes broadcasters give verbal feedback to the comments or questions, but it's still hard to guess what they are replying to. Fortunately, most of broadcasters are experienced communicators in using live streaming. They first read out the comments and questions and then respond. The reading time gives the broadcasters some time to understand the words and to think about how to reply to the comments and questions. Another kind of lack understanding applies to broadcasters who cannot understand comments and questions. For example, Nuoye, a Chinese chat broadcaster, cannot understand the comment:

“Dong bu dong you ...”

“Dong bu dong you ...”, “eh...”, “What are you talking about?”

In this example, Nuoye tries to read and understand the meaning of the comment, but failed. This comment is not a sentence. The broadcaster cannot get any information from it, so she asked “What are you talking about?”.

People sometimes type fast and make some mistakes. So the comments cannot express the right meanings. Viewers type quickly because of the instantaneity of live streaming. Subtopics may change quickly in live streaming. In order to catch the right time to send

the comments out, viewers have to keep a high speed of typing. So mistakes appear all the time. And in order to catch the turn, viewers try to type as short as possible. In this situation, some comments may lead to lack of understanding or lead to misunderstandings. Lack of understanding maybe caused by a missing punctuation. For example, an American food broadcaster, Cliff, cannot understand his viewers:

“Some people are foodies”

“Is this a comment or a question?”

Cliff does not know whether it is a comment or a question. It influences the feedback a lot, so Cliff decided to ask clearly.

When it comes to live streaming in USA, taking Periscope as an example, many comments on Periscope are in different languages. Broadcasters cannot understand all of them. Some broadcasters just ignore them, some try to understand, some emphasize “English only please”.

Live streaming provides people with a great way to communicate in real-time, but phenomena like lack of understanding and misunderstanding are inevitable.

6.5 Embarrassment Hiding

Broadcasters sometimes feel embarrassed during broadcasting and then they often try to hide the embarrassment. Embarrassment appears more in highly-interactive live streaming. More in detail, when there are no comments and questions from the viewers, the pauses and the quiet usually make broadcasters obviously embarrassed. So many broadcasters use background music while broadcasting. When there is nothing to talk about, they just turn up the sound and sing with the music. Some female broadcasters often play with their hair to hide embarrassment.

The embarrassment is amplified by live streaming. If person A communicates with person B face-to-face, and they have nothing to say. They can see each other’s facial expressions and feel the emotions. But in live streaming, only broadcasters face the cameras, seeing their own faces, knowing nothing about the viewers’ emotions.

In highly-interactive living streaming, broadcasters depend on their viewers a lot. Although the broadcasters control the turn, the viewers can choose whether to see the live streaming and whether interact with the broadcaster. If a broadcaster has no viewer, the interaction will end. Correspondingly, if no broadcasters broadcast, there will no viewers anymore. From this point of view, broadcasters and viewers are important to each other. Successful live streaming makes both broadcasters and viewers feel comfortable rather than embarrassed.

7. Conclusion

In the conclusion section, patterns of communication in live streaming will be summarized based on the theoretical framework in the first part. Differences between China and USA are summarized in the second part. In addition, limitations and future studies are discussed.

7.1 Patterns of Communication in Live Streaming

People from 20 to 30 are the main groups of people who broadcast themselves. 76.4% broadcasters (81) broadcast indoor. And 84% of the broadcasters (89) just stay in one place, only 16% of them keep moving places during broadcasting. Broadcasters who broadcast indoor usually do not change to another place, and broadcasters who broadcast outdoor usually change places during broadcasting. Good Internet and environment indoor help broadcasters to focus their attention and give more feedback. Broadcasters are powerful interlocutors who can control turn management. They can get the turn by starting a new topic, giving feedback and answering questions they like. But they lose turn when they see comments and questions and forget talking. Viewers can also get the turn by asking questions or giving comments. The turn changes frequently in live streaming, therefore the sequences in live streaming are usually short, but the number of sequences are large. As for feedback, viewers' text-based comments/questions and gifts/hearts can be interpreted as feedback in broad sense. Broadcasters can only interpret all the information from these ways. Broadcasters use gestures or give vocal verbal replies to the comments and questions as feedback. Broadcasters sometimes may ask for feedback from the viewers. The viewers' feedback in live streaming can be ignored, interrupted, and delayed.

Broadcasters send information multimodally, but receive information with just the visual modality. However, viewers send information using two visual modalities (texts and images), while they perceive information with two sensory modalities (visual and auditory). The main reasons for these characteristics lie in live streaming technologies.

Compared with face-to-face communication, live streaming has lower flexibility when users choose modalities.

Multimodal communication is used in different types of live streaming. When live streaming is about chat, talk show, music or food, broadcasters always use auditory aspects (speech, song or non-linguistic voice), videos shot by front-facing camera to produce information during live streaming communication. While when broadcasting sports, travel, news activity or games, they usually combine auditory aspects (speech, song or non-linguistic voice) with videos shot by rear camera.

We pay special attention to body movements in live streaming when analyzing multimodal communication. There are 6 categories used frequently in live streaming: movements of arms and hands, facial gestures, head movements, distance, eye movements, and lip movements. Some body movements are used frequently in both countries, e.g. finger points to camera to emphasize; hand moves close to camera to show things in hand. According to our data, body movements in live streaming mainly occur in the upper body of broadcasters and are mainly used with speech in live streaming.

All the above are the patterns of communication existing in both countries, and therefore proposed as the general patterns of communication in live streaming.

7.2 Differences between Chinese and American Live Streaming

In this part, the differences between Chinese and American live streaming are summarized.

Due to the influence of commercial, technological, and psychological reasons, live streaming viewers in China are 432 times of viewers in the USA. Chinese live streaming has stronger interactivity than US live streaming. In China, there are much more viewers than in the USA. Some Chinese broadcasters broadcast for commercial benefits, so viewers are the center of broadcasting. Chinese broadcasters have stronger motivation to give feedback. American broadcasters broadcast for the satisfaction of

themselves.

According to the difference in degree of interactivity, we divide Chinese live streaming into 3 groups, talking live streaming, display-talking live streaming, and display live streaming. We use the same method to divide American live streaming, it clusters into two groups, obvious two-way interaction live streaming and obvious one-way interaction live streaming. By comparing the original categories existing on the different groups, we found that the biggest difference between China and USA in interactivity appears in the talk show. In China, the talk show is in the group with the highest interactivity. But in USA, the talk show is in the group of obvious one-way interaction live streaming, which has low interactivity. American talk show broadcasters share their own opinions to viewers more, while Chinese talk show broadcasters prefer to discuss the topics with viewers.

As for body movements, some of the body movements are used by only Chinese broadcasters or only American broadcasters. Heart-shaped gestures are only used by Chinese broadcasters, while stop gestures are only used by American broadcasters.

In addition, most Chinese broadcasters usually gave apologies or some explanations when they have delayed in answering questions, while most American broadcasters do not do this. American broadcasters mostly chose questions which were typed in the language they can understand. The differences of behaviors between China and America have their own cultural explanations in the results and analysis part.

To conclude, comparing Chinese and American live streaming, there are general patterns of communication specific in live streaming. Live streaming is a new medium with high interactivity, frequent turn changing, and short but large number of sequences. Because of the technological limitations, broadcasters can use more than one modality to broadcast, while the audience can only use text-based comments, sending gifts or tap hearts up to communicate with broadcasters. Live streaming has its own frequently used body movements and features. Although the common features of live streaming are same in both countries, unique culture based features can still be found in different areas. By studying patterns of communication in live streaming, this study contributes in: I) enrich the studies of live streaming (from communicative perspectives); II) study

communication patterns in a new medium; III) show how live streaming works; VI) demonstrate the differences and similarities between China and the United States. Furthermore, live streaming as a new medium, has a great potential of development with the help of new technologies. Live streaming and its real-time interactions can reduce the distance impact and improve the communicative efficiency. More and more news media are using live streaming to broadcast news, which makes the viewers close to the truth. The speed of information spreading are greatly accelerated, the information gap disappears, and the world will be more transparent. The study of communication patterns can help people develop its new functions and usages based on its features in future.

7.3 Limitations and Future Studies

Limitations of our study can be found in the observation work and analysis. The samples are enough for our study, but if they can be enlarged, the data can have a better performance of avoiding extreme values. Live streaming is very personal, and it is influenced a lot by personal traits, so more samples can help to reduce the personal traits, but also to increase the cultural characteristics. Observation can be more in detail, especially in the ICM part.

Communication of live streaming is still a new field of study, and communication is the main function of live streaming. With its technological limitations, live streaming has huge differences from face-to-face communication. Live streaming and other real-time communication available on the Internet have a great study value. Future studies can focus on Interactive Communication Management and Multimodal Communication.

References

- Ahlsén, E., Allwood, J., & Nivre, J., 2003. Feedback in different social activities. Nordic research on relations between utterances. *Copenhagen Working Papers in LSP*, Vol. 3, p 9-37.
- Allwood, J., 1985. Intercultural Communication, *Papers in Anthropological Linguistics 12*, Department of Linguistics, Göteborg University, Sweden, p 9–61.
- Allwood, J., 1999a. The structure of dialog (Gothenburg papers in theoretical linguistics, 82). *Göteborg: Dept. of Linguistics, Univ*, p1-24.
- Allwood, J., 1999b. Are there Swedish Patterns of Communication. In: *H. Tamuyra (Ed.) Cultural Acceptance Of Cscw In Japan & Nordic Countries*. Kyoto Institute Of Technology, p 90.
- Allwood, J., 2002. Body communication dimensions of expression and content. *Multimodality in language and speech systems*, p 7-26.
- Allwood, J. & Cerrato, L., 2003. A study of gestural feedback expressions. *First nordic symposium on multimodal communication*, p 7-22.
- Allwood, J., Cerrato, L., Jokinen, K., Navarretta, C., & Paggio, P., 2007. The MUMIN coding scheme for the annotation of feedback, turn management and sequencing phenomena. *Language Resources and Evaluation*, 41(3-4), 273-287.
- Allwood, J., Kopp, S., Grammer, K., Ahlsén, E., Oberzaucher, E., Koppensteiner, M., 2008a. The Analysis of Embodied Communicative Feedback in Multimodal Corpora - a prerequisite for behavior simulation, p255-272.
- Allwood, J., 2008b. Dimensions of embodied communication - towards a typology of embodied communication, p1-24.
- Allwood, J., 2008c. Multimodal Corpora. *An International Handbook*. Mouton de

Gruyter, Berlin, p 207-225.

Allwood, J. & Ahlsén, E., 2009. Gestures that precede and accompany speech--An analysis of functions and applicability for virtual agents in different activities. *NEALT PROCEEDINGS SERIES*, Vol. 6, p 3-9.

Allwood, J., & Lindb, M., 2010. On the Potential of some Pragmatic Concepts for the Web. *Systems, Signs & Actions*, Vol. 4(1), p 38-52.

Allwood J, Ahlsén E., 2012, 13. Multimodal Communication. *Handbook of technical communication*, p 435-460.

Allwood, J., 2013. A framework for studying human multimodal communication. *CRC Press, Taylor & Francis Group, Boca Raton, FL*, p 17-39.

Birkner, C., 2016. Scoping it Out. *Marketing News*, p 10-11.

Brouwer, B. 2015. What Live Streaming Means for Content Publishers. *EContent*, 2015, Vol.38(9), p.11(1).

Center, C. I. N. I., 2013. Statistical report on internet development in China. *China Internet Network Information Center*, 2017(1).

China in comparison with United States: <https://www.geert-hofstede.com/china.html>

United States in comparison with China: <https://www.geert-hofstede.com/united-states.html>

Cui, Q. X., 2016, The Mode Analysis and Future Development Trend of Network Live. *New Media Research*, Vol. 17, p 7-8.

David A. Shamma, D. A., 2009. Spinning online: a case study of internet broadcasting by DJs. In Proceedings of the fourth international conference on Communities and technologies, p 175-184.

Ding, Y. et al., 2011. Broadcast yourself: understanding YouTube uploaders.

In Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference, p 361-370.

Douyu: http://www.douyu.com/cms/about/about_us.html

Edelman, M., 2016. From Meerkat to Periscope: Does Intellectual Property Law

Hamilton, W. A. et al., 2014. Streaming on Twitch: Fostering Participatory Communities of Play within Live Mixed Media. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, p 1315-1324.

Hamilton, W. A., 2016. Rivulet: Exploring Participation in Live Events through Multi-Stream Experiences. In Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video, p 31-42.

Huya: <http://www.huya.com/zt/usc/#about>

Jensen, M., 2014. Smile as Feedback Expressions in Interpersonal Communication: A First Acquaintance Context.

Jia, Y., 2016. The prosperity and melancholy of Internet live show: interpersonal interaction, carnival, cultural trap. *Editorial Friend*, Vol. 11, p 42-48.

Justin D. Weisz, J. D., 2007. Watching together: integrating text chat with video. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, p 877-886.

Juhlin, O. et al., 2010. Mobile broadcasting: the whats and hows of live video as a social medium. In Proceedings of the 12th international conference on Human computer interaction with mobile devices and services, p 35-44.

Lim, S. et al., 2012, Getting closer and experiencing together: Antecedents and consequences of psychological distance in social media-enhanced real-time streaming video. *Computers in Human Behavior*, Vol.28, No. 4, p 1365-1378.

McDermott, I. E., 2015. See What I See Right Now. *Online Searcher*, Nov/Dec 2015, Vol.39(6), p.35-37.

Mishra, S., 2017. Communication Patterns.

Businesstopia: <https://www.businesstopia.net/communication/communication-patterns>

Panda Live: <http://www.panda.tv/news/1101414857/3700943925/3701130188.html>

Pires, K. & Simon, G., 2015. YouTube live and Twitch: a tour of user-generated live streaming systems. In Proceedings of the 6th ACM Multimedia Systems Conference, p 225-230.

Prohibit the Live Streaming of Commercial Sporting Events. *Columbia Journal of Law & the Arts*, Vol. 39, No. 4, p 469-495.

Powell, F., 2015. Broadcast yourself – live. *SOCIAL MEDIA*, Vol.29(7), p M27(1).

Tang, J. C. et al., 2016. Meerkat and Periscope: I Stream, You Stream, Apps Stream for Live Streams. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, p 4770-4780.

House, N.V., 2016, Periscope, Congress and the aliveness of video.

The Berkeley Blog: <http://blogs.berkeley.edu/2016/06/29/periscope-congress-and-the-aliveness-of-video/>

Siekkinen, M., et al., 2016, Anatomy of a mobile live streaming service: the case of periscope.

ResearchGate: <https://www.researchgate.net/publication/303221755>

Tang, J. C. et al., 2017. Perspectives on Live Streaming: Apps, Users, and Research. In Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, p 123-126.

Thorburn, E. D., 2014. Social media, subjectivity, and surveillance: Moving on from

occupy, the rise of live streaming video. *Communication and Critical/Cultural Studies*, Vol. 11, No. 1, p 52-63.

Youtube live <https://support.google.com/youtube/answer/2853700?hl=zh-Hans>

Yuan, A. Q. & Sun, Q., 2016. Regression and Transcendence: Network Directing under Visual Cultural Psychology. *New Media Research*, Vol. 16, p 54-58.

YY Live: <http://www.huanju.cn/s/introduction.html>

Zhang, C. & Liu, J. C., 2015. On crowdsourced interactive live streaming: a Twitch.tv-based measurement study. In *Proceedings of the 25th ACM Workshop on Network and Operating Systems Support for Digital Audio and Video*, p 55-60.

Appendices

Appendix 1 Observation Schema

Attention:

Number (A-China, B-The USA);

Gender (not sure=0, Female=1, Male=2, multi-gender=3);

Age (not sure=0, 0-20=1,20-30=2,30-40=3, >40=4);

Place (Indoor=1, Outdoor=2);

Place movement (change=1, not change=2);

Dress up (exposed=1, unexposed=2, not sure=3)

Number		Category		Content	
Time (0-24)		Duration (mins)		Platform	
Number of viewers		Boradcaster's ID		Gender	
Age		Place		Place (specific)	
Place movement		Dress up		Dressing and make-up	
Number of Feedback units					
Number of Questions					
Number of Answers					

Number of Sequences	
Body Movement	
Multimodal Communication	
Exceptions	

Appendix 2 Cluster Membership in USA

Cluster Membership			
Case	4 Clusters	3 Clusters	2 Clusters
1:Chat	1	1	1
2:Chat	1	1	1
3:Chat	1	1	1
4:Chat	2	2	1
5:Chat	1	1	1
6:Chat	1	1	1
7:Chat	1	1	1
8:Chat	1	1	1
9:Talk show	1	1	1
10:Talk show	3	3	2
11:Talk show	3	3	2
12:Talk show	2	2	1
13:Talk show	3	3	2
14:Talk show	3	3	2
15:Talk show	3	3	2
16:Talk show	2	2	1
17:Music	1	1	1
18:Music	4	3	2
19:Music	2	2	1
20:Music	4	3	2
21:Music	1	1	1
22:Music	1	1	1
23:Music	1	1	1
24:Music	4	3	2
25:Food	4	3	2
26:Food	3	3	2
27:Food	2	2	1
28:Food	4	3	2
29:Sports	4	3	2
30:Sports	4	3	2
31:Sports	4	3	2
32:Sports	3	3	2
33:Sports	1	1	1
34:Sports	3	3	2
35:Travel	4	3	2
36:Travel	2	2	1
37:Travel	1	1	1
38:Travel	3	3	2

39:Travel	3	3	2
40:Travel	2	2	1
41:Game	4	3	2
42:Game	4	3	2
43:Game	4	3	2
44:Game	4	3	2
45:Game	4	3	2
46:Game	4	3	2
47:News	4	3	2
48:News	4	3	2
49:News	4	3	2
50:News	4	3	2
51:Activity	4	3	2
52:Activity	4	3	2
53:Activity	4	3	2
54:Activity	4	3	2

Appendix 3 Cluster Membership in China

Cluster Membership			
Case	4 Clusters	3 Clusters	2 Clusters
1:Chat	1	1	1
2:Chat	1	1	1
3:Chat	2	2	1
4:Chat	2	2	1
5:Chat	1	1	1
6:Chat	1	1	1
7:Chat	1	1	1
8:Chat	3	3	2
9:Talk show	1	1	1
10:Talk show	1	1	1
11:Talk show	3	3	2
12:Talk show	4	3	2
13:Talk show	2	2	1
14:Talk show	1	1	1
15:Talk show	1	1	1
16:Music	2	2	1
17:Music	2	2	1
18:Music (chat)	1	1	1
19:Music	2	2	1
20:Music (chat)	1	1	1
21:Music	1	1	1
22:Music	1	1	1
23:Music	4	3	2
24:Music	2	2	1
25:Food	3	3	2
26:Food	2	2	1
27:Food	2	2	1
28:Food	3	3	2
29:Sports (chat)	2	2	1
30:Sports	2	2	1
31:Sports	3	3	2
32:Sports	3	3	2
33:Sports	3	3	2
34:Travel	4	3	2
35:Travel (chat)	1	1	1
36:Travel	3	3	2
37:Travel	3	3	2
38:Travel	4	3	2
39:Game	3	3	2

40:Game	4	3	2
41:Game	3	3	2
42:Game	2	2	1
43:Game	2	2	1
44:Game	3	3	2
45:News	3	3	2
46:News	3	3	2
47:News	3	3	2
48:News	3	3	2
49:Activity	3	3	2
50:Activity	3	3	2
51:Activity	3	3	2
52:Activity	3	3	2

Appendix 4 Division of Work

This is the joint thesis which is undertaken by Meimei Wang and Shuqian Zhou under the supervisor, Jens Allwood. We discussed and decided the topic of this thesis together. And then we collected the information of the development of live streaming and the relevant studies. After that, the theoretical framework and study method were decided by both of us. We also did the observation together. When writing the thesis, each author has the contribution to the project:

Meimei Wang is in charge of the abstract, introduction, and background;

Shuqian Zhou is responsible for the methodology, conclusion, and appendices;

The theoretical framework, results and analyses, and discussion sections are achieved by both of us.

After finishing writing, we revised the thesis according to the guidance of our supervisor. He gave us lots of constructive suggestions. We appreciate all of his efforts in our thesis.