



HAL
open science

Multimodal Interactional Practices in Live Streams on Twitter

Le Song

► **To cite this version:**

Le Song. Multimodal Interactional Practices in Live Streams on Twitter. Library and information sciences. Institut Polytechnique de Paris, 2024. English. NNT : 2024IPPAT019 . tel-04627171

HAL Id: tel-04627171

<https://theses.hal.science/tel-04627171v1>

Submitted on 27 Jun 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Multimodal Interactional Practices in Live Streams on Twitter

Thèse de doctorat de l'Institut Polytechnique de Paris
préparée à Télécom Paris

École doctorale n°626 Ecole doctorale Polytechnique de Paris (ED IP Paris)
Spécialité de doctorat: Sciences de l'information et de la communication

Thèse présentée et soutenue à Palaiseau, le 6 juin 2024, par

LE SONG

Composition du Jury :

Heike BALDAUF-QUILLIATRE Professeur, Université de Lyon, France (Département de langue et littérature allemandes)	Présidente/Examinatrice
Wyke STOMMEL Professeur, Radboud University (Département de Langue et Communication) Nimègue, Pays-Bas	Rapporteure
Anna SPAGNOLLI Professeur, University of Padua, Italie (Département de psychologie générale, Centre de recherche sur les technologies d'inspiration humaine)	Rapporteure
Enhua GUO Professeur associé, Ocean University of China (Faculté des Langues Étrangères) Qingdao, Chine	Examineur
Yumei GAN Professeur associé, Shanghai Jiao Tong University, Chine (École des Médias et de la Communication)	Examinatrice
Christian LICOPPE Professor, Institut Polytechnique de Paris (Télécom Paris), France	Directeur de thèse

TABLE OF CONTENTS

<i>Abstract</i>	2
<i>Résumé</i>	3
<i>Chapter 1</i>	1
<i>Introduction</i>	1
1.1 Live Streaming as an Emerging Mediated Interaction	3
1.1.1 The Evolution of Watching Video: Integrating Text Chat with Online Video.....	3
1.1.2 The era of mobile live streaming: A shift from desktop talking heads to mobile experience sharing....	4
1.1.3 The practices of using mobile live streaming in today’s society	7
1.2 Research in chat and video in EMCA	11
1.2.1 Computer-Mediated Chat	11
1.2.2 Video-Mediated Interaction	17
<i>Chapter 2</i>	1
<i>Multimodal Conversation Analysis as Method</i>	1
2.1 The Platform: Twitter Live	1
2.2 Data Collection and Data Corpus	7
2.3 Designing a multimodal transcription system for integrating video and text	9
<i>Chapter 3</i>	1
<i>The organization of live stream openings</i>	1
<i>Chapter 4</i>	1
<i>Noticing-based actions and the pragmatics of attention in expository live streams. Noticing ‘effervescence’ and noticing-based sequences</i>	1
<i>Chapter 5</i>	1
<i>“Oh my Dave, your face says it all!”: Exploring the Multimodal Practices of Tasting in Live Streaming</i>	1
<i>Chapter 6</i>	1
<i>Closing Live Video Streams: A Sequential Analysis</i>	1
<i>Chapter 7</i>	2
<i>Conclusion and future research</i>	2
7.1 Summary and synthesis of thesis	2
7.2 Theoretical and practical implications	5
7.3 Limitations and Future Expectations	6
<i>References</i>	1
<i>List of publications</i>	1

ABSTRACT

As an emerging form of mediated interaction, live streaming has become a rapidly growing practice that combines the technical and interactional features of video-mediated interaction and multi-party chat. Live streaming with mobile devices on multiple platforms has thus been a practice in which streamers and viewers interact in highly asymmetric forms—the streamer’s video display and the viewer’s written text. This doctoral dissertation focuses on live streams as interactional phenomena from a sequential perspective. Drawing on video-recorded data from ordinary users’ naturally unfolding activities in daily life-oriented live streams on Twitter (now ‘X’) and taking ethnomethodology and conversation analysis (EMCA) as its theoretical and methodological perspective, the thesis explores how the use of multiple (e.g., spoken, written and embodied) resources, as well as the manipulation of affordance of the devices in establishing the participation framework of live streaming interactions and achieving different joint actions stepwise. The dissertation consists of four main research articles, each focusing on a typical interactional phenomenon in live streaming. All of the articles have been published or are under review. Article I investigates the openings of live streaming. Unlike phone conversations with a canonical opening sequence, live stream openings appear more variable, with laminated participation frames. However, there is usually a recognizable "installation" phase where the stream activity begins. We also identified interactional concerns in the opening: the streamers’ wait for an adequate audience, their collective and individual management of viewers within a guest/host relationship, and the concern of participants regarding the immediate intelligibility of the stream. Article II discusses how streamers and viewers manage attention and engagement through noticing-based actions. It looks at how streamers and viewers produce noticing sequences and noticing-based sequences, and how the orientation towards noticing may lead to a distinctive form of ‘noticing effervescence.’ Article III inspects the activity of tasting in live streaming, re-examining tasting in this particular ecology as an interactive process that combines individual sensory experience with a public, witnessable, and intersubjective dimension. Article IV investigates the organization of closing sequences in live streaming. It shows that while participants can be seen to orient to the sequential organization of closings in ordinary conversation, they do so in a way that is particularly sensitive to the affordances of live video streams. The thesis thus provides a systematic analysis of the most characteristic interactional properties of live streaming.

Keywords: live streaming, multimodal conversation analysis, multimodality, computer-human interaction, mediated interaction, computer-mediated chat, video technology, Twitter

RÉSUMÉ

En tant que forme émergente d'interaction médiatisée, la diffusion en direct (live streaming) est devenue une pratique en pleine expansion qui combine les caractéristiques techniques et interactionnelles de l'interaction vidéo-médiatisée et du chat multi-participants. Le live streaming à l'aide d'appareils mobiles sur plusieurs plateformes est donc une pratique dans laquelle les diffuseurs (streamers) et les spectateurs interagissent sous des formes hautement asymétriques : l'affichage vidéo du diffuseur et le texte écrit du spectateur. Cette thèse de doctorat s'intéresse au live streaming en tant que phénomène interactionnel d'un point de vue séquentiel.

S'appuyant sur des données vidéo enregistrées d'activités advenant naturellement dans des live streams orientés vers la vie quotidienne sur Twitter (maintenant 'X') et sur l'ethnométhodologie et l'analyse conversationnelle (EMCA) comme perspective théorique et méthodologique, la thèse explore comment l'utilisation de multiples ressources (par exemple, parlées, écrites et corporelles), ainsi que la manipulation des 'affordances' des appareils permettent de produire les cadres de participation propres aux live streams et l'accomplissement de différentes actions conjointes de manière séquentielle.

La dissertation se compose de quatre articles de recherche principaux, traitant quatre phénomènes interactionnels caractéristiques des live streams. Le premier analyse les séquences d'ouverture des live streams. Contrairement aux conversations téléphoniques et leur séquence 'canonique' d'ouverture, les ouvertures de live streams apparaissent plus variables, avec de multiples cadres de participation stratifiés, bien qu'il y ait généralement une phase d'installation reconnaissable où l'activité de diffusion commence. La thèse identifie des préoccupations interactionnelles spécifiques aux ouvertures, à savoir l'attente d'un public adéquat par les streamers, la manière dont ceux-ci gèrent les interactions avec le public à la fois dans son ensemble, et de manière individuelle dans le cadre d'une relation invité/hôte, et les préoccupations affichées par les participants concernant l'intelligibilité immédiate du stream.

L'article II discute de la manière dont les diffuseurs et les spectateurs démontrent attention et engagement en formulant ce que les streams rendent remarquables. Il examine ainsi comment les streamers et les spectateurs produisent des séquences initiées par des remarques (noticings), et comment l'affinité particulière des live streams avec cette pratique à la fois attentionnelle et interactionnelle peut conduire à une 'effervescence attentionnelle' caractéristique.

L'article III inspecte l'activité de dégustation dans le live streaming, où la dégustation est accomplie comme un processus interactif et multimodal qui combine l'expérience sensorielle individuelle avec une dimension publique et intersubjective. Cette partie de la thèse met en lumière comment les streamers décrivent leurs sensations gustatives tout en interagissant avec leur audience, créant ainsi une expérience partagée qui transcende les limites de l'écran. Cette analyse révèle la complexité de l'activité de

dégustation en direct, qui nécessite une coordination fine entre la parole, les gestes et les expressions faciales pour transmettre l'expérience sensorielle de manière convaincante et engageante.

L'article IV enquête les séquences de clôture dans le live streaming. Nous montrons comment les participants s'y orientent vers l'organisation des clôtures caractéristique de la conversation ordinaire, mais d'une manière très sensible aux affordances des diffusions vidéo en direct. Les séquences de clôture sont marquées par des stratégies variées de signalement de la fin imminente de la diffusion, incluant des remerciements, des récapitulatifs des points saillants du stream, et des interactions finales avec les spectateurs. L'importance des adieux et des promesses de futurs streams est également examinée, soulignant comment les clôtures servent à maintenir l'engagement et la fidélité de l'audience.

La thèse fournit donc une analyse systématique des propriétés interactionnelles les plus caractéristiques du live streaming. En explorant les ouvertures, les remarques, les dégustations et les clôtures, elle met en lumière la manière dont ces interactions médiatisées créent de nouveaux types de relations sociales et de pratiques communicationnelles. Chaque article contribue à une compréhension plus approfondie de la dynamique interactionnelle dans le contexte unique des live streams, offrant des perspectives sur la manière dont les individus naviguent et négocient ces espaces numériques pour créer du sens et de la connexion sociale.

Les contributions de cette recherche sont multiples. Premièrement, elle enrichit le champ de l'analyse conversationnelle en appliquant ses principes à un nouveau domaine d'étude, celui du live streaming. Deuxièmement, elle propose des insights pratiques pour les streamers et les développeurs de plateformes, leur permettant d'améliorer l'interaction et l'engagement des utilisateurs. Enfin, elle ouvre des pistes pour des recherches futures sur d'autres formes d'interaction médiatisée, notamment en explorant comment les nouvelles technologies continueront de transformer les pratiques communicationnelles dans les années à venir.

En conclusion, cette thèse démontre que le live streaming n'est pas seulement une forme de divertissement, mais un phénomène social complexe qui mérite une attention académique approfondie. En combinant l'ethnométhodologie et l'analyse conversationnelle avec une observation minutieuse des pratiques réelles, elle offre une vision détaillée de la manière dont les individus co-crée des expériences interactives riches et dynamiques à travers les plateformes numériques.

Mots clés: diffusion en direct, analyse conversationnelle multimodale, multimodalité, interaction homme-ordinateur, interaction médiatisée, chat médiatisé par ordinateur, technologie vidéo

CHAPTER 1

INTRODUCTION

Central to my investigation is the concept of mediated interaction (Arminen, 2005; Arminen et al., 2016). Mediated interaction refers to engagements in which participants leverage the affordances¹ provided by mediating technologies to connect, communicate, and collaborate. In the forms of mediated interactions, live streaming has been an emerging phenomenon and a ubiquitous feature of everyday life. In the course of a day, we may interact with live streaming as we browse social media, buy stuff through e-commerce, have online courses, or play video games, to name just a few. Live streaming has garnered significant scholarly attention worldwide. In this dissertation, I embark on a journey into the world of live streaming, aiming to explore its interactional practices deeply embedded in the implications of social contexts.

This research defines live streaming as a highly asymmetric interaction where streamers make live appearances on video. At the same time, an active audience engages through publicly available text messages, with voice interactions occurring only sporadically. This digital theater of interaction unfolds in real time, with the act of producing messages restricted to the individual composing and sending them. However, this seemingly straightforward interaction has its challenges, as time lapses, known as lags, between messages sent and received introduce complexities in maintaining conversational continuity (Garcia & Jacobs, 1999). Live streaming defies simple categorization, manifesting in various forms across diverse platforms. Streamers, colloquially known as "content creators," transmit their lives, activities, or performances, bridging the gap between the virtual and the real. The level of interactivity afforded to viewers varies, allowing for written chat, voice, virtual gifts, and even financial contributions.

While not entirely novel, live streaming has undergone a remarkable resurgence, particularly with the advent of mobile live streaming applications (APPs) around 2015 (Garst, 2015). Today, live streaming occupies a pivotal role in the global social media landscape, straddling the realms

¹ Spagnoli et al. (2021) assert that affordances are the characteristics of a medium that suggest particular actions to the users. In the field of human-computer interaction, the emphasis has shifted away from the direct influence of a medium's technical traits on user behavior, towards an understanding of how these traits—referred to as affordances—take on meaning when they meet the users' needs. Affordances become meaningful when individuals discover ways to utilize them within certain activities (see Whittaker, 2003; Hutchby, 2001, etc. for a more detailed analysis).

of entertainment and e-commerce industries, most notably in some East Asian countries (Cunningham, Craig & Lv, 2019; Si, 2021). The statistics are staggering. Live streaming has become a competitive arena for internet giants seeking to attract subscribers. One example is Twitch, which was purchased by Amazon in 2014 and gained a significant number of young users. Twitch witnessed a sharp rise from 2.2 million active streamers in 2017 to 9.89 million in January 2021, a peak number until now (the number was 7.13 in September 2023) (Taylor, 2018; TwitchTracker, 2023²). Similarly, Periscope, a live streaming platform once owned by Twitter, was also striving to attract users in the social live streaming space by offering login and user recommendation features.

The competition in the live-streaming industry is not limited to Twitch and Periscope. Facebook also entered the race with its live streaming feature, Facebook Live. Furthermore, YouTube and TikTok have also enabled live streaming on its platform, allowing content creators to broadcast live to their audience. Globally, platforms like South Korea's AfreecaTV and various Chinese live streaming services, such as Douyin, Kwai, and Taobao, exert their influence. With so many options available, it is up to the users to decide which platform they prefer based on their specific needs and interests. Ultimately, live streaming is a reflection of the ever-evolving digital landscape, where innovation and convenience are key factors in attracting and retaining users.

It is essential to acknowledge the profound shift during the COVID-19 pandemic, where the rapid growth of live streaming has been catalyzed, prompting a shift towards online platforms for almost all areas, such as education and entertainment. During COVID-19, live streaming became a lifeline for education, enabling students and educators to engage remotely via video (Chen et al., 2021).

Today, live streaming is much more than just another way to watch videos online—it shapes how we experience and interact with everything happening around us. It is the tool we use to connect with others, share our stories, and be part of the important moments as they unfold. Beyond mere life live-streamed, it is life refracted through live streaming (Khobra & Gaur, 2020).

² See TwitchTracker: <https://twitchtracker.com/statistics/active-streamers>

In the following, I will explore the wide-reaching effects of live streaming, from how it changes the way we interact in real-time broadcasts to its significant influence on online shopping and its important role in political movements. I am bringing together different pieces of research, where each piece adds special meaning to our overall picture. Guided by many sources, both classic and new, these studies help us gain a much deeper understanding of the world of live streaming.

1.1 LIVE STREAMING AS AN EMERGING MEDIATED INTERACTION

1.1.1 THE EVOLUTION OF WATCHING VIDEO: INTEGRATING TEXT CHAT WITH ONLINE VIDEO

Live video streaming can be described as ‘the ability to broadcast video to a remote audience in the instant that it is captured’ (Juhlin et al., 2010; Rein & Venturini, 2018). After entering the new century (especially after the advent of YouTube in 2006), watching videos online was becoming increasingly popular. With the advent of video-on-demand services (Raman et al., 2018) and digital video recorders, how people consumed media was undergoing a fundamental change. People were less likely to watch shows at the same time, let alone in the same place. As a result, television viewing, which was once a social activity, has been reduced to a passive and isolated experience (Nathan et al., 2008).

Then, there came the transformative phase in digital media consumption, that is, the integration of real-time text chat into online video and TV programs. Online video platforms such as YouTube and Google Video and traditional media companies like ABC and CBS began trying to allow users to not only watch videos, movies, or television shows but also engage in discussions through text chat or even audio chat (Weisz & Kiesler, 2008) side-by-side with streaming video, making these video content more interactive (Weisz et al., 2007).

This shift, which began before the rise of mobile live streaming, fundamentally changed the way audiences interacted with video content. As Weisz et al. (2007) observed, the advent of mixing real-time chat and video brought forth the potential to turn passive video watching into an active and socially engaging experience. The authors empirically examined the influence of text chat and found that despite the potential distractions, text chat had a positive impact on social relationships. Weisz and Kiesler (2008) also compared viewers’ experience of text chat and audio text and found both text and audio chat enjoyable, with some preferring audio chat for socializing while watching videos with friends. Geerts et al. (2011) explored the importance of synchronization and the impact of “togetherness” when watching online videos together by

comparing the use of voice chat as opposed to text chat. Hamilton et al. (2014) explored the emergence of live video game streaming communities on platforms like Twitch and underscored the role of open Internet Relay Chat (IRC) live chat in fostering participatory online communities.

In conclusion, the integration of real-time chat, particularly text chat, into online video and TV programs marked a critical transition. It transformed solitary video consumption into a social experience, shaping the foundation for subsequent developments in mobile live streaming. This phase exemplified the adaptability of the digital entertainment industry in responding to evolving audience preferences for social interaction and engagement with video content.

1.1.2 THE ERA OF MOBILE LIVE STREAMING: A SHIFT FROM DESKTOP TALKING HEADS TO MOBILE EXPERIENCE SHARING

Live video streaming underwent a significant transformation toward being mobile, transitioning from traditional desktop "talking heads" to a mobile experience focused on sharing real-time experiences. In this part, I will explore the emerging and technological evolution of mobile live video streaming, which is the leading form and the focus of live streaming in this thesis.

The rise of mobile live video streaming can be attributed to the convergence of several technological advancements. Initially, applications for real-time video transmission over the Internet existed, with desktop video conferencing being a prime example (Rein & Venturini, 2018). However, the actual transformation occurred with the proliferation of modern mobile phones and the advent of 3G networks, enabling individuals to broadcast live video from their mobile devices (Rein & Venturini, 2018). Traditional cameras and video equipment are gradually losing the race with smartphones and small mobile devices that allow video, photo, and audio capturing on the go (de Sá et al., 2013) by manipulating affordances such as 'Pan-Tilt-Zoom' (Inkpen et al., 2013). The early generation of mobile streaming services includes Qik, Bambuser, Flixwagon, and kyte.com (Juhlin et al., 2010).

Before the popularity of mobile video communication, research primarily focused on "talking heads," which involved remote participants conferring or performing tasks together. However, Nardi et al. (1993) introduced the concept of "video-as-data," shifting the focus to video images of workspaces and work objects. This transition aimed to enhance task performance rather than provide telepresence, opening up new possibilities for video usage. Brubaker et al. (2012) emphasized the importance of shared experiences in video communication, challenging the

conventional "talking heads" approach. They discussed the need for multiple camera perspectives. They highlighted scenarios like sharing events in stadiums (Dezuli et al., 2013), where data is more critical than an individual's face, such as surgical operations or brainstorming sessions.

This shift towards mobile-based live streaming paved the way for collaborative live video production, where ordinary people, such as photography "amateurs," could work together to capture and share events in real time (Mughal & Juhlin, 2013). This made live streaming an ad hoc in-situ media (de Sá et al., 2013), even if they may struggle with camera handling techniques (Juhlin et al., 2010).

Although it had become popular, the quality of mobile streaming video then was limited by mobile technologies like bandwidth and devices. The introduction of Meerkat and Periscope in 2015 ignited a boom in the industry (Rein & Venturini, 2018). Since then, live mobile video streaming has begun to gain popularity worldwide, thanks in part to increasingly pervasive smartphone ownership, cellular broadband coverage, and the development of 4G (Singanamalla et al., 2018). Mobile live streaming and all sorts of mobile mediated interaction, especially mobile video-mediated interaction, have come into a boom in this period.

The appeal of mobile live streaming extends beyond its novelty. Users appreciate the authenticity, uncurated nature, and interactivity it offers (Tang et al., 2017). The fusion of mobile technology and live video streaming allowed for greater flexibility and accessibility in content creation (Tang et al., 2016). It becomes possible to stream and watch streams not only at home but also out in the real public world (Reeves et al., 2015). Unlike earlier mobile streaming services, Meerkat and Periscope leveraged more mature mobile platforms with higher-quality image and audio capture capabilities. This improvement allowed for more engaging and polished live broadcasts. Live streaming emerged as a versatile tool that accommodated various activities and settings, appearing to democratize the creation of real-time content (Tang et al., 2016).

Apart from the exponential growth of the mobile video market at the time (Morrison, 2015), the success of mobile live streaming services was also attributed to the entry of tech giants like Twitter, Google, Amazon, and Facebook into the social streaming market (Rein & Venturini, 2018). These platforms were pivotal in introducing the concept of mobile live streaming as a

growing form of participatory social media (Hamilton et al., 2016) and as a mainstream social activity (Tang et al., 2016).

Integration with internet giants such as Twitter further enhanced the audience reach of live streams. This integration was crucial in attracting an audience for impromptu live broadcasts (Tang et al., 2016). For example, Meerkat and Periscope capitalized on their integration with Twitter, allowing users to easily stream to their viewers and promote their streams via Twitter notifications (Morrison, 2015; Wagner, 2015; Rein & Venturini, 2018). Meerkat later suffered a blow when Twitter removed Meerkat's access to their social graph and bought its competitor Periscope (Bacheller, 2015). Periscope was able to dethrone Meerkat with Twitter's backing quickly and become the leading app that popularized mobile live streaming from phones (Kastrenakes, 2021).

Additionally, the rise of the "platform economy" or "platform capitalism" (Srnicsek, 2017) has further underscored the role of powerful platforms in reshaping the digital landscape (Rein & Venturini, 2018). Facebook's strategy in influencing the evolution of live video streaming is a testament to how influential players actively shape media landscapes through technological and financial initiatives (Rein & Venturini, 2018).

However, Periscope was not long-lasting. It was discontinued on 31st March 2021 due to declining usage, product realignment, and high maintenance costs (Periscope, 2019; Team Periscope, 2020), and all its data and functions remain on Twitter. Today, Twitter has the same page layout, functions, user accounts, and contents as Periscope but the Periscope app itself was shut down ("Periscope (service)", 2022).

This historical event seems to indicate the decline and unpopularity of live streaming. However, the fact is that the discontinuation of Periscope highlights how far video adoption can reach. As DeBois (2021) claims: "Twitter discontinues Periscope, but live streaming lives on." Social platforms such as Instagram and TikTok have begun to adopt streaming features, competing for user attention today by using live streaming. As more social media platforms introduced video capabilities, dedicated live streaming apps like Periscope became redundant, drawing fewer users and engagement over time.

As live streaming continues to evolve, the need for better support on mobile devices and the web to enhance the production quality of live streams, as well as the trend toward multi-stream

experiences, poses new challenges and opportunities, necessitating innovative solutions for future live streaming (Hamilton et al., 2016). The ongoing process of addressing technical challenges and harnessing its potential for social engagement is shaping its development and influence on society. This is also what we will investigate in the following section.

1.1.3 THE PRACTICES OF USING MOBILE LIVE STREAMING IN TODAY'S SOCIETY

Live streaming has experienced significant growth and diversification over the years. One of the notable impacts of mobile live streaming has been its role in sharing information. It has become a tool for disseminating real-time information and allowing viewers to engage with unfolding events, fostering interaction and mutual understanding among participants and distant viewers (Tang et al., 2017). This unique feature enables functions such as disseminating user-generated content, enriching self-presentation, fostering online communities, etc. (Wu et al., 2022), allowing for its utilization in various areas. However, originally a form of real-time communication, live streaming has evolved into a versatile platform for entertainment, education, commerce, politics, and so on. Streamers and viewers do all kinds of activities, including sharing thoughts and experiences, chatting, showcasing professional skills, broadcasting social events, playing video games, and eating in front of the camera. There have also been corresponding ethical and legal concerns. This sub-section will explore user practices in today's live streaming by examining key findings from various studies.

Live streaming has been a crucial form of real-time entertainment (Wang & Lu, 2023). Video game live streaming is a form of interactive video social media that combines traditional broadcasting and online gaming (Kim et al., 2019). Over the past decade, the popularity of video game live streaming has grown rapidly, prompting researchers to investigate its impact on various psychological variables. According to the Use and Satisfaction theory, Li et al. (2020) identified four factors that influence user participation (both streamers and audiences) in video game live streaming: streamer demand, audience demand, interaction behavior, and platform impact. This research provides insight into the mental health implications of Internet addiction. Live streaming platforms (e.g., Twitch) have attracted countless streamers, so anyone can find suitable channels and live content they want (Taylor, 2018). Recktenwald (2018) examined the dialogical and monological moves in game live streaming on Twitch, shedding light on the coordination of spoken-to-written communication. He identified different directions of communication. These directions include viewer-to-streamer, streamer-to-viewer, and viewer-

to-viewer. The study also revealed that each direction of communication serves a different social purpose.

As an important element for interacting, digital gifting has also been a crucial phenomenon in the interactions of live streaming. Lee et al. (2019) explored the motivates of viewers' gift-sending and found that viewers were motivated by the reciprocal acts of streamers, who would engage in various types of social interactions with gift senders during the live streams. Guan et al. (2021) explore what influences the purchase of virtual gifts in live streaming in China and reveal how viewers' social perceptions can be shaped by the IT-related factors of live streaming (e.g., responsiveness, two-way communication, and self-presentation, etc) and the cultural characteristics of China, and leads to flow experience and purchase intention.

The action of digital gifting is highly relevant to economic factors. The intersection of live streaming and economics (for example, e-commerce, digital labor, industry, and capitalism) emerges as a focal point in the research landscape. Research has also found that live streaming may lead to new employment opportunities, particularly for people in developing countries (Singanamalla et al., 2018). Dong et al. (2023) look into technological and social factors influencing consumers' purchase intentions. Hou et al. (2019) also investigate what factors can affect people's continuous watching and consumption intentions in live streaming in China and find that sex and humor appeals, social status display, and interactivity play considerable roles in the viewer's behavioral intentions. Fung et al. (2022) explore the platformization of people's everyday lives and platform labor on Douyin (TikTok's Chinese version). They argue that although it seems that live streaming viewers are closer than ever to the stream creators in terms of engagement, they are abstracted through platform algorithms, meaning that they are reduced to statistics provided to stream creators. Moreover, in light of its mobile and real-time broadcasting feature, live streaming is an ideal device for showing scenes of different places. Live streaming is continuously becoming an important tool for tourism destination marketing, and many destinations use it to attract tourists. The results indicate that the interactivity, vividness, authenticity, and immediacy features of destination live streaming positively influence consumers' sense of presence and trust, which in turn enhance their travel intentions (Zheng et al., 2022).

Besides its connections with economics, live streaming is also emerging as a critical role in education, particularly during the COVID-19 pandemic. The outbreak of COVID-19 has led to

a sharp transition from offline to online education in many countries and areas. Some offline on-site activities, such as folk clubs and fitness classes, have also been increasingly transformed into online via live streaming (Wang & Lu, 2023). Due to the pandemic, governments across the globe mandated school closures. This necessitated an abrupt shift to online/distance teaching (DeCoito & Estaiteyeh, 2022; Nikolopoulou & Kousloglou, 2022), and heightened the intensity of existing challenges of online education, such as knowledge sharing and learning (Lu et al., 2018; Fonseca & Barbosa, 2020), student attendance, and education equality. Research (Chen et al., 2021) showed that teachers utilized live video streaming on various social media platforms and repurposed several entertainment features to provide online teaching for improved student engagement. As a result, students even developed a stronger sense of connection with their teachers after experiencing certain online interactions.

Live streams have been gaining importance in Human-Computer Interaction (HCI) research and practice (Fonseca & Barbosa, 2021). Exploring the intertwining of gesture, voice, and AI Control in video streaming, Shamma et al. (2019) present a Wizard-of-Oz elicitation study to inform the design of interactions with smart 360° cameras or robotic mobile desk cameras for use in live streaming situations, moving forward the insights to the future personal video cameras for streaming and collaboration. Mobile live streaming has been used to enable immersive collaboration in task-oriented experiences such as remote shopping, virtual interactive tourism, auditing and verification, mobile crowdsourcing, and remote physical-world games, in which users appear as ‘avatars.’ For example, Singanamalla et al. (2018) explore the experiences in an ‘escape-the-room’ physical puzzle game where some team members remotely interacted with a video stream produced by other team members who were physically present in the escape room. Wu et al. (2023) conducted research to look into VR (virtual reality) streaming on Twitch and investigate the practices between streamers and viewers which combine interactions across displays and space. Their findings suggest that technical issues, a lack of smooth transitions between real and virtual environments, and unintentionally designed game scenes make it difficult for VR streamers to establish emotional connections (Guo & Fussell, 2020) with their audience and keep up a steady stream.

Protests and political demonstrations have been extensively covered through live streaming (Tang et al., 2017). Dougherty (2011) explored the emergence of the earlier generation of mobile live-streaming platforms like Qik and their impact on civic engagement. The study revealed that those who self-identified as activists, journalists, community leaders, or educators

were likelier to engage in live streaming, underscoring the medium's potential for civic participation.

It is worth noting that, despite gaining traction in North America, live streaming has not reached the popularity it has in some East Asia countries like China, where live streaming has a tremendous impact on society (Lu et al., 2018). There seems to be a mania for China's live streaming (Long & Tefertiller, 2020), and there are reasons behind this. One major difference is that compared to their Western counterparts, full-time streamers in China are more common (Lu et al., 2018), and Chinese live streaming apps allow people to broadcast a wider range of activities. Another significant difference is the virtual gift-giving feature (Xiang, 2016) as mentioned above. While American streamers earn money from ads, endorsements, direct donations, or subscription fees, Chinese streamers receive most of their income directly from fans in the form of virtual gifts. This means that the audience's level of engagement is crucial for success in the Chinese live streaming market (Lee et al., 2019). So live streaming in China highlights the lively social atmosphere and heated social interactions among streamers and viewers (Guan et al., 2021). Users also desire deeper functions for interaction in addition to the commenting, gifting, and fan groups that are already available (Lu et al., 2018). The success of TikTok and its short videos and live streaming functions might be a prime exemplar of how some of the Chinese platform's operation models could be applied and influence the world.

Ethical and moral concerns have also been a topic in live streaming research since live streaming is a synchronous platform that allows users to disclose information to a vast audience in real-time (Wu et al., 2022). It has not been clear how mobile technologies will affect social norms and policy debates regarding privacy, surveillance, and intellectual property (Faklaris et al., 2020). Extreme negative events such as live streaming suicide have occurred and questioned related platforms' ability to monitor shared content that can negatively affect the audience, although announcing suicides through live streaming may lead to the prediction of suicide (Shoib et al., 2022). Research has explored the legal and ethical implications of mobile live-streaming video apps through a review of public-policy considerations since the advent of mobile apps such as Meerkat and Periscope (Faklaris et al., 2016). Prototypes were designed to improve the privacy management on live streaming, increase users' awareness of privacy, and protect broadcasters' privacy (Alamiri & Blustein, 2016), improving jurisprudence, public policy, and applied ethics in related areas (Faklaris et al., 2016).

In conclusion, the literature on live streaming practices demonstrates its diverse applications and impacts across various domains, including education, commerce, entertainment, politics, and ethics. Live streaming continues to evolve, offering new opportunities and challenges and shaping the way we engage with content and each other in the digital age.

1.2 RESEARCH IN CHAT AND VIDEO IN EMCA

In this section, I will present a review of studies in the field of mediated interactions from the perspective of EMCA. I will unfold this section into two parts: computer-mediated chat and video-mediated interaction, which are highly relevant to the two main media forms that constitute live streaming, i.e., chat and video.

1.2.1 COMPUTER-MEDIATED CHAT

Not only have our smartphones—and the apps we use on them—opened up a world of exciting new communication channels, but they have also expanded the settings in which those channels can be used. Apps like WhatsApp, Line, WeChat, Facebook Messenger, X, Instagram, TikTok, Twitch, Blued, Tinder, and Grindr have given significance and functionality to old forms of communication, especially text messages. Text chat is now accessible in multiple situations, and the media that can be included in messages has also changed. In addition to text, messages can now include images, audio, stickers, and video clips, which can be shared amongst dispersed groups and displayed in threaded conversations (Brown et al., 2018).

Research on computer-mediated chat from the perspective of EMCA has gone through a path from comparing text-based typed chat with face-to-face speech conversations to studying typed chat in new media contexts. Reed and Ashmore (2000) claim that chat is defined by its 'informality', not its modality. They note that chat includes both informal textual interaction and informal voiced interaction: postcards, letters, informal emails, telephone, and dinner table conversation. Despite its 'informality' feature, however, chat, rather than being trivial, may be central to the organizational work of everyday serious talk (McHoul & Rapley, 2000). While computers can facilitate online chat, the reverse is also true: chat can make supposedly formal conversation, like psychological testing, possible.

Likening to 'strangers in a strange land' (Rintel & Pittam, 1997), with the advancement of communicative technology, interactions in Quasi-Synchronous Computer-Mediated Communication (QS-CMC) (Garcia & Baker Jacobs, 1999; Segerstad, 2002; Kelsey & Amant,

2008), such as email, real-time chat protocols, asynchronous discussion forums, web pages, and the like (Herring, 2010), began to attract attention by researchers from EMCA, particularly in the comparison with traditional face-to-face (FTF) interactions, speech and conversations. After the popularity of mobile phones but before the prevalence of 4G, typed exchanges such as text messaging on mobile phones have been, and continue to be, the most typical kind of computer-mediated ‘conversations’ (Herring, 2010). People even claim that 2010 is the year technology replaced talking (Jayson, 2010). After the prevalence of 4G and ubiquitous networks, typed exchanges are still popular. However, they began to integrate more and more into other forms of mediated interactions, such as problem presentations in e-mail (Stommel & Van der Houwen, 2016), WhatsApp (König, 2023; Mostovaia, 2021), WeChat (Pan and Gan, 2022), video-mediated class interaction (Gudmundsen, 2023), knowledge status in Skype videoconference (Dooly & Tudini, 2022), live streaming, “visual addressivity” in a multimodal 3D virtual environment (Naper, 2011), and AI chatbots (Eisenmann et al., 2023; Lee, 2021).

The research of text-based chat is based on the research of spoken conversations. Benwell and Stokoe (2006) believe that participants in virtual worlds do not create entirely new ways of communicating but rather recreate and adapt features of communication from the offline world. For example, by comparing Facebook chat with spoken interaction, Meredith and Stokoe (2013) propose that the study of online text interaction treats it as a speech-exchange system with its own normative rules, which are adapted from and show similar orientations to the speech exchange system of ordinary, mundane interaction. Also, regarding the research method on text-based chats (Neuage, 2004; Meredith, 2015; Sandel et al., 2019), Koivisto et al. (2023) note that to date, studies in digital CA usually build on observations obtained from the conversation analytic research of spoken interactions. According to Meredith (2019), CA can add to broader conversations about the role of technology in interaction by providing detailed evidence about the characteristics of online interaction. Garcia and Baker Jacobs (1999) note that we were able to see what information each participant had at the time they decided to write, post, edit, or erase because we used video recordings of participants’ computer screens to capture the interactional process of the conversation’s production. As Have (2000) points out, since more and more people are spending time ‘online’ chatting with friends or anyone available, it is appropriate to study computer-mediated conversation from the same procedural perspective that we used to study face-to-face conversation and telephone conversation.

Internet Relay Chat (IRC) in the mid-to-late 1990s is believed to be the Internet's first widely popular QS-CMC system (Rintel et al., 2006) and has received lots of research. IRC is a protocol for real-time text messaging between internet-connected computers that was created in 1988. As a 'naturally occurring chat machine' (Reed and Ashmore, 2000), IRC is distinct from other CMC forms such as e-mail and discussion lists. IRC users type messages that are relayed by a server, often within one second, to other users who are also present in the same channel. Participants treat the messages that accumulate on the channel screen as 'conversation' (Vallis, 2001). Studies (Rintel and Pittam, 1997; Rintel et al., 2006; Vallis, 1999) have compared the opening and closing phases on IRC with face-to-face casual group interactions and suggested that interaction management is similar, but that the content, structure, and ordering of the strategies are subject to adaptation. Rintel et al. (2003) also inspect the non-response situations on IRC and explore four intersections of participant-action and system-occasioned non-responses. They propose an extension to Pomerantz's (1985) 'pursuing a response' problems. Applying membership categorization analysis (MCA) that applies Sacks' concepts of category, Rhyll Vallis (2001) explicates the orderliness of how members accomplish moral identities in their talk despite the seeming 'anonymity' provided in chat-room talk. Schönfeldt and Golato (2003) look into repairs in German web chats and discuss when, why, and how participants adjust the repair practices from ordinary conversation to suit the particular conditions of the medium of web chats.

Turn-taking has been a striking issue among research topics in QS-CMC studies (Keng Wee Ong, 2011; Herring, 2011). Garcia and Jacobs (1999) compared the turn-taking systems in QS-CMC and oral conversation and discovered that the turn-taking system of QS-CMC is substantially different from the turn-taking system of oral conversation (Sacks, Schegloff, & Jefferson, 1974). Due to limitations on turn-taking, Herring (1999, 2011) notes that text-only CMC is interactionally incoherent. She reveals a high degree of disrupted adjacency, overlapping exchanges, and topic decay in text-only CMC. However, she also claims that the ability of users to adapt to the medium and the advantages of 'loosened relevance' (Herring, 2013) account for the popularity of CMC despite its relative incoherence.

Besides IRC and turn-taking, several key areas (Meredith, 2019), such as sequence organization (Nilsen & Mäkitalo, 2010), assessment, response design, repair, facework, knowledge status, openings, and closings, have also been examined in media contexts like social media, video communication, chatbots, etc. In the field of social media, Mostovaia (2021) looked into the

other-initiations of repair in German WhatsApp chats and discussed how different other-initiation repair forms shape the negotiation of responsibility for the trouble source, manage epistemic rights and Goffmanian face work (1967). The ‘face’ and ‘politeness’ issues are also investigated by Golato and Taleghani-Nikazm (2006) by analyzing requests and assessments in web chats. Spagnolli et al. (2021) look into the design of message extensions in the instant messaging app Skype and examine the practices parties extend prior turns and how they treat them. Pan and Gan (2022) have inspected the “continuing assessments” in online dating on WeChat. They showed how the continuing assessments are sequentially organized in a conversational context and how continuing assessments are applied to end the dialogue and to manage face work. König (2023) examined response designs in WhatsApp chats and showed how they contextualize different stances of agreement in text-based interaction. In research of chat in video, studies focus on how participants make use of the chat function in video-supported mutually coordinated interaction like Zoom, in which multiple synchronous communication modalities such as text chat, video, images, and audio are synchronously available to the participants (Thorne & Hellermann, 2022). By investigating interactions in Skype videoconference student meetings, Dooly and Tudini (2022) imply how platform features like parallel text chats are used by the students and teachers to sequentially organize their knowledge status and pedagogical intersubjectivity. Gudmundsen (2023) looks into text chat in a Zoom digital language café and shows how engaging the chat in a video-mediated setting emerges as a social practice and develops as part of the participants' context-specific interactional achievement. Oittinen (2024) examines chat as an organizational and participatory resource in video-mediated activities and investigates the use of the chat interface in Zoom online courses. The study shows how written turns posted in the publicly available channel become included in spoken interaction, highlighting the chat interface as an important resource for other-than-current speakers to adjust their participant status *in situ*. Licoppe and Morel (2018) examine the interactional practices in live streaming, in which viewers interact with streamers visible onscreen in real time by sending text comments. They find the streamer’s ‘read aloud and respond’ pattern in (selectively) responding to the viewers’ comments in chat and managing the issue of addressivity. A similar pattern is found also in messaging in co-present situations. Brown et al. (2018) explore the role messaging plays in interaction and talking with those who are co-present. They find that both sent and received messages share similarities with reported speech in conversation. They note that messages can become a topical resource for local conversation by verbatim reading aloud or adaptive summaries.

Apart from Vallis (1999), Rintel et al. (2006), and Rintel and Pittam's (1997) studies on the opening and closing phases of IRC mentioned above, scholars have also looked into the openings and closings of text-based interactions on other media platforms. Using the framework of conversation analysis, Markman (2009) examines the structure of interaction in chat-based team meetings, focusing on the openings and closings. The author describes how the computer-mediated chat disrupts the temporal flow of conversation and makes beginning and ending difficult. Stommel and te Molder (2015) and Stommel (2016) present a conversation analysis of opening sequences and initial phases of online chat counseling and discuss how affordances of digital communication media may work as constraints when the participants orient to face-to-face interactional norms. Christodoulidou (2018) employs conversation analysis to study e-counseling and to assess consultee satisfaction in e-counseling chat endings and implies that satisfied consultees end the chat with enthusiastic appreciation and dissatisfied consultees resist closing the chat. De Rijk and Stommel (2023) investigate initiating post-match chat interaction on the dating app Tinder. They found that initiating interaction is often done by launching a first topic immediately. The claim is critical because the Tinder topic initiations are overwhelmingly recipient-designed and often contain information-seeking first-pair parts, designed to elicit a response (Licoppe 2021). Lee (2021) examines patterns of users' first turns interaction with a service chatbot that was developed in South Korea, expanding CA as a methodology to investigate human-chatbot interaction.

Gender is a crucial topic in CMC chat research. Issues such as sexual identity authentication and embodiment in language use (Del-Teso-Craviotto, 2008), gender plays (Rellstab, 2007), and trading sex pics (Slater, 1998) have been widely examined. It is claimed that to engage in the complexities of online desire and eroticism, users must first authenticate themselves as members of a particular gender or sexual group. Because there are no visual or auditory cues, authentication is inevitably a discursive act in this context. It is accomplished through language techniques like the age/sex/location schema, self-descriptions, and screen names. Therefore, it is argued that social practices—rather than being an external process imposed upon people—are what lead to authentication. The argument is consistent with Roca-Cuberes et al. (2023) conclusion in their investigation of the relationship initiation and formation in Tinder chat conversations; that is, relationships are ongoing routine accomplishments arising in mundane sociorelational contexts. The analysis shows that the interactional device through which the revelation of personal and intimate information is prompted is the 'elicited self-disclosure sequence.'

One pervasive use of dating mobile applications such as Tinder and Grindr is the initiation and accomplishment of relationships between strangers based on location awareness (Licoppe et al., 2016). As a core feature of the platform, the chat function receives attention from EMCA researchers. Christian Licoppe does a series of chats on mobile dating app studies using a conversation analytic approach, focusing on a joint understanding of the orderliness of what social media users do online and of the language ideologies that may support such orderliness (Licoppe et al., 2016; Licoppe, 2019; Licoppe, 2021). Licoppe et al. (2016) examine Grindr casual hook-ups as interactional achievements, analyzing the unique profile-matching sequences-based interactional practices that Grindr users have created to get around the relational affordances of electronic conversation. The studies also reveal how proximity awareness is embedded in dating apps and location-aware mobile technology, and sexual impulse becomes fused into a hybrid actant (Licoppe et al., 2017). Comparing the interactional practices on Tinder, Licoppe (2019) finds that Grindr and Tinder users take almost opposite conversational stances regarding the organization of casual hookups. That is, gay Grindr users need to balance between using computer conversations to set up quick sexual encounters and avoiding leading to more intimate relationships. Heterosexual users on Tinder are looking to achieve topically rich chat conversations with strangers. All these are reflected through the unique linguist ideologies that both users have created.

Online chat counseling, particularly for medical purposes, has also caught scholars' attention. Stommel and colleagues have done a series of studies on online chat counseling sessions. They examine the organizational and interactional tasks of trouble management and formulations in institutional chat counseling (Stommel & Van der Houwen, 2013; Jager & Stommel, 2017), as these have been found to play an important interactional role in face-to-face counseling. Stommel and te Molder (2015) and Stommel (2016) present a conversation analysis of opening sequences and initial phases of online chat counseling and discuss how affordances of digital communication media may work as constraints when the participants orient to face-to-face interactional norms. Stommel et al. (2017) explored how hyperlinks are used in chat counseling and how empathy is displayed in these interactions (Stommel & Lamerichs, 2016). Spagnolli et al. (2023) look at the conversation with Digital Health Assistants (DHAs) in Apps and find that the DHA-patient encounter only includes three phases compared to the human physician-patient encounter: the opening, the history-taking, and the closing. They propose a design

checklist for designing DHA and provide suggestions, such as the task of the DHA being a data collecting assistant and avoiding self-attributing a clinical role.

To sum up, some key features can be found in computer-mediated chats. For example, the ‘disrupted turn adjacency,’ ‘loosened relevance’ of messages, the abridgment and extension of turns and sequences, etc. Also, some critical research topics can be identified from these studies, for example, the opening and closing phases of the chat, assessments, repair, recipient design, and the actions that users achieve through instant messaging (medical consultation, dating, etc.). These features and research topics will facilitate and inspire our research in live streaming research. I will now turn to another constitutive modality of live streaming, video-mediated interaction.

1.2.2 VIDEO-MEDIATED INTERACTION

In comparison to co-present interaction, video-mediated technologies allow social interaction in a ‘fractured ecology’ (Luff et al., 2003) and constrain the ways through which the co-participants can achieve and maintain intersubjectivity and reciprocity of perspectives (Koivisto et al., 2023). In video-mediated interaction (VMI) (Mlynář, González-Martínez, & Lalanne, 2018; Due & Licoppe, 2021; Haddington & Oittinen, 2022), distant co-participants have restricted access to each other’s visual environment, as they can only monitor the part of the environment that is currently framed on the screen (Arminen et al., 2016). This can limit the possibilities of co-orientation and collaboration and create a need for additional workspaces.

Also, the current video technology makes it impossible to use specific embodied resources such as eye contact to allocate turns in multiparty interaction. Asymmetrical visual access also restricts the use of pointing and other gestures, as well as posture and body movement, as interactional resources. Health and Luff (1992, 1993) note that video-mediated presence reveals asymmetries in interpersonal relations which, as far as we are aware, are neither found within face-to-face interaction nor in other technologically mediated forms of communication such as telephone calls. Dealing with visual boundaries (Koivisto et al., 2023), therefore, becomes a crucial issue in video-mediated interaction. As an overarching feature of video-mediated interaction, participants have been observed to display orientation to the visual talking heads arrangement (such as a headshot) as the default frame for video interaction. That is, departures from it (such as shots of one’s environment) are treated as accountable actions and inspected by others for relevance (Licoppe & Morel, 2012).

CA examines technology-mediated interaction, such as VMI, from an emic perspective. CA is not concerned with the features or shortcomings of the technology itself, but rather with how the technology and mediation impact and shape the organization of talk and action within the context of communication. This approach emphasizes the relevance and significance of VMI in communication, as demonstrated by research conducted by Arminen et al. in 2016. As personal, professional, and recreational video communication applications have proliferated and diversified during the past 10 or 15 years (see, e.g., Harper et al., 2022; Due & Licoppe, 2021), the multimodal organization of interaction in video-mediated environments has received increasing scholarly attention.

The literature review below will unfold surrounding several main research domains of EMCA video-mediated studies; that is, videoconference, video call, video game, online consultation, online education, video tutorials, etc. The studies demonstrate how interactional and relational work is accomplished within the affordances and constraints of video technology, offering a rich understanding of the implications of this technology for human sociality.

First, the literature on video conferencing reveals a deepening understanding of how video technologies intersect with human communication, particularly in professional settings. Previous studies highlight the practices in which individuals and professionals navigate the affordance and constraints presented by video conferencing environments.

Mondada's work (2007) extends the understanding of video conferencing into the domain of telesurgery, demonstrating how video technology expands traditional teamwork by including remote experts in real-time operations. Seuren et al. (2024) have analyzed how physiotherapists and patients communicate during video consultations. The study highlights the importance of tailored communication strategies to overcome the inherent challenges of remote assessments, such as managing safety and clinical risk and ensuring visibility for exercises and movements. Licoppe and colleagues' research from 2013 to 2017 explores the use of video conferencing in French courtrooms, focusing on the multimodal and multilingual complexities of such interactions. Their studies look into how camera movements and video frames contribute to the creation of a shared space of action and the representation of participants, such as interpreters, within legal proceedings. Cappellini and Azaoui (2017) focus on telecollaboration and the co-construction of normative evaluation sequences within educational settings. This comparative study provides insights into the interactional features facilitated by desktop video conferencing

systems and how they affect learning outcomes. Cipolletta et al. (2017) investigate the formation of therapeutic relationships online, emphasizing the elements that characterize effective video-mediated therapy sessions. Their findings contribute to a better understanding of how therapeutic rapport and client engagement can be established and maintained in a remote setting. The works of Ekberg et al. (2019, 2020) examine the use of physical objects in speech and language therapy for young children via video conferencing. Their research offers practical strategies for engaging children in therapeutic interventions and elucidates the differences between telehealth and in-person sessions.

Second, video calling is also an important research object in EMCA video-mediated interaction studies. The literature examines how video calls have evolved from a fixed and static form of communication to one characterized by mobility and multimodality, referred to as the “mobility turn.”

Starting with de Fornel (1996), the investigation into the interactional frameworks of videophone interactions sets the stage for understanding how users adapt to technological interfaces to create meaningful social exchanges. This foundational work shows the locally produced organization of videophone communication, where users must adapt their conversational strategies to the limitations and possibilities of the medium. Licoppe and Morel's series of studies from 2009 to 2012 look into the collaborative work that goes into producing meaningful shots during mobile video calls and Skype interactions. They discuss the characteristic features of “video-in-interaction,” such as the orientation towards the “talking heads” format, the scrutiny of video images for their interactional relevance, and the obligations of the video handler in multiparty interactions. Their work underscores the mobility turn, where the use of lightweight, portable devices has allowed for an expansion of the communicative space beyond fixed locations, allowing participants to share their environments.

Harper, Rintel, Watson, and O’Hara’s studies from 2013 to 2017 explore various facets of video calling, from storytelling and category formation in couple’s calls to the concept of the “interrogative gaze,” which discusses how video communication invokes an order of behavior that participants manage collaboratively. Their work illustrates the coordination of appearances and greetings in video call openings, highlighting the routine nature of video calling in contemporary social life. Rintel’s (2013a, b) papers discuss how couples in long-distance relationships navigate the challenges of technological mediation, using audio/video distortions as resources within their relational communication. This points to a nuanced understanding of

technology as an active participant in shaping relations, rather than merely a conduit for communication.

Licoppe's further studies (2014, 2017) on showing objects in video calls detail the collaborative achievement involved in making objects visible and relevant within interactions, thereby sustaining "intimacy-at-a-distance." The work by Licoppe et al. (2017) on the handling and manipulating of artifacts in video-mediated settings elaborates on the multimodal nature of these interactions, where body positioning and gaze play critical roles. The contributions of Rosenbaun and Licoppe (2017, 2019), Zouinar and Velkovska (2017, 2019), and Busch (2018) expand on the theme of maintaining personal relationships and intimacy through video-mediated communication by showcasing objects and engaging in shared participation frames. Their analyses reveal how video calls are used by family members as a technology-mobilized resource for maintaining intimacy despite geographical distances.

Gan and deSouza (2022) further this understanding by exploring how family members manage participation and visibility during video calls. Their work contributes significantly to reconceptualizing the relationship between visibility, technology, and mobility in video-mediated communication, especially noting the nature of participation, where individuals may move in and out of the camera's view as an element of their interactional involvement. Gan's (2021) further exploration into multisensoriality in video calls emphasizes the embodiment and sensorial interactions in the constitution of intimacy across distances.

Third, the investigation of video games in EMCA has yielded substantial insights into the intersection of play, interaction, and social organization in the last decade. Video gaming has served as a medium for video-mediated interaction in multifaceted ways, such as entertainment and social learning.

Lorenza Mondada's series of studies serves as a cornerstone in this field. In her 2011 work, Mondada reevaluates the use of imperative verbs in French within the context of video games, arguing for a more nuanced understanding of directives. Rather than seeing imperatives as inherently rude, Mondada positions them as responsive and context-sensitive instructions in the gaming interaction. Her subsequent papers in 2012 and 2013 explore the temporal and embodied aspects of gaming interaction. Mondada's 2012 study on virtual football games offers a granular look at how players navigate the fluid boundaries between in-game action and the social space of the co-present players. The 2013 paper extends this by examining how real-time

coordination and mobile actions within video games necessitate and generate timely communication.

Continuing this exploration, Laurier and Reeves (2014) offer an examination of the role of cameras in video games, relating them to narrative, architecture, and human-computer interaction. Their analysis shows how game cameras influence the visibility arrangements and the spatio-temporal experience of gameplay, contributing to a deeper understanding of the practical and perceptual elements of play. The collaborative study by Reeves, Greiffenhagen, and Laurier (2017) draws together disparate research threads to present an EMCA perspective on video gaming as a practical accomplishment. The authors employ a "tactical zoom" approach, progressively narrowing the focus from the broad context of multiplayer online gaming spaces to the interactions between players and spectators at the screen and finally to avatar conduct within the game itself. This comprehensive study foregrounds the material "work" of gameplay, highlighting the complexity and situated nature of video gaming interactions.

Baldauf-Quilliatre and colleagues have also done a series of research on video games, focusing on issues such as embodied movement, the meaning of the screen, turn reciprocity and encouragements, etc. Lydia Heiden, Heike Baldauf-Quilliatre, and Matthieu Quignard (2023) examine the interactional functions of cursor movements in a screen-based video game, where pointing and showing by hand is not available. They claim that cursor movements (straight pointing movement, swaying/circling movement, and "moving away") can contribute to the interaction as a display of attention, as a device for turn-holding/turn-claiming/leaving the floor, or as a particular resource in argumentation. Baldauf-Quilliatre and Colón de Carvajal (2020) explore how participants encourage each other in videogame interactions through a comparison of encouragement with instructions. They argue that encouragements show another sequence structure than instructions, occur in particular contexts and are characterized by a specific turn design. In the virtual world of the videogame, the switch of turns of speech is highly dependent on the actions. By examining the organization of turn-taking, Baldauf-Quilliatre and de Carvajal (2015) investigate the issue of whether the avatar is considered as a participant by the players and the place the players of video games give to the avatars. They identified four forms of exchange based on whether a turn or action is addressed to a player, an avatar, a player and his avatar together, or an unclear referent. They have shown that the avatar can be considered as a real participant by the players in video games. As the important embodiment of video, the screen itself has also been studied from an EMCA perspective. Baldauf-Quilliatre and Ursi

(2023) look into the screen-centered family game interaction and highlight three constitutive dimensions of the screen: spatial, material, and temporal. They note that spatially, participants can navigate between different positions in front of the screen; materially, the screen can be mobilized either as a technological device or as a fragile object; temporally, local accomplishments on the screen can be prospectively or retrospectively assessed or embedded within the general flow of the session.

Research has also paid attention to adolescents and children who play video games. Arja Piirainen-Marsh (2012) zooms in on the home environment to explore how adolescents engage with video games. Through a multimodal perspective, this chapter focuses on social play and how the discourse of the game becomes a tool for structuring interaction and participation among players. The work of Sabria Salama Jawhar et al. (2024) provides an insightful case study of children's online pre-gaming interaction, with a particular focus on Roblox. Employing Multimodal Conversation Analysis, the authors dissect the cooperative learning strategies exhibited by preteen players. They underscore the educational benefits of gaming, noting how players engage in mutual scaffolding and share knowledge to solve problems collaboratively.

Fourth, the incorporation of VMC in healthcare consultations has led to the popularity of online video consultations. Research has focused on understanding how these technologies reshape the interactional fabric of medical encounters, shedding light on the ways in which participants collaborate to create a shared interactional space and accomplish clinical tasks.

Wyke Stommel and colleagues (2020, 2021) provide a microanalysis of companions' participation in video consultations, demonstrating how their presence and involvement are contingent upon their visibility within the video frame. These studies reveal the distinctions between video-mediated and face-to-face consultations, especially regarding the participation framework and the establishment of shared focus during examinations and assessments.

Hansen's work (2022) looks into interpreter-mediated hospital encounters in Norway. Through multimodal conversation analysis, Hansen examines how interpreters use embodied actions to signal and manage communication troubles within video-mediated environments. These embodied actions are akin to recruitments and are pivotal in initiating repair mechanisms in interactions that are both linguistically and technologically complex. In her 2020 study, Hansen identifies the challenges in organizing the visual ecology within these video-mediated interpreting sessions. The study highlights how participants rarely negotiate the visual setting,

despite the incongruence of views potentially causing interactional troubles. The research elucidates how participants navigate this visual materiality to achieve the goal of multilingual interpreting within hospital settings.

Brian Due and Simon Bierring Lange (2020) address the adaptation of physical examinations to the constraints and affordances of video-mediated consultations. Their research reveals two distinct gestural highlighting practices within video-mediated physiotherapy consultations: ‘mimicable body part highlighting’ and ‘direct body part highlighting.’ These practices illustrate the creative ways in which patients and healthcare professionals navigate the limitations of sensory access to conduct successful history-taking and symptom assessments.

Lucas Martinus Seuren and his colleagues (2021) focus on the impact of latency on the organization of turn-taking in video-mediated interactions. Their analysis provides insight into how participants negotiate turn-taking in the face of delayed auditory and visual cues. Despite the challenges, these encounters offer a unique setting to study the mechanisms of social interaction and the participants' adaptability. The same research team, in their 2020 study, expands on the topic by examining the linguistic and ethnographic dimensions of video consultations in various medical services. They identify key interactional challenges such as opening the video consultation, managing conversational flow disruptions, and conducting remote examinations. Interestingly, despite the technological and operational issues, the study finds that clinicians and patients effectively employ established communication strategies to navigate these challenges.

Fifth, all forms of online education within video-mediated interactions are increasing especially due to Covid-19 in the past years. Studies using EMCA have provided significant insights into the communicative and interactive features that define this educational format.

Some research has focused on online education in schools and families. Davidson, Danby, and Thorpe (2017) look into the multimodal meaning-making processes of young children as they engage with YouTube videos in a preschool setting. Through conversation analysis, they reveal the collaborative nature of viewing, where the children and their teacher use a variety of resources to construct shared understandings of video content. Helen Melander Bowden and Johanna Svahn (2020) explore the interactional work involved in video-mediated homework support. They demonstrate how tutors and students collaboratively navigate the shared digital workspace of an online platform despite geographical separation and unequal access to physical

artifacts. The study emphasizes the need for participants to establish shared points of reference and negotiate problem-solving trajectories, suggesting that the design of online educational platforms should accommodate such collaborative attributes.

Some research explores the second-language (L2) education via video. Jakonen, Dooly, and Balaman (2022) contextualize technology's role in second-language education and explore how classrooms can be reconceived as material-technological ecologies that extend beyond their physical confines. This broader conceptualization calls for an integration of multimodal and ethnomethodological perspectives in educational research and practice. Sert, Jenks, and Çimenli (2022) contribute to this body of work by investigating topic maintenance in virtual exchanges. They introduce the concept of “rolling the ball back” as a discursive practice that facilitates the extension of topics and engagement in L2 interactions, providing a potential tool for enhancing online language education. Uskokovic and Talehgani-Nikazm (2022) examine the embodied practices of L2 speakers during word searches in video-mediated interactions. Their work reveals how gestures and verbal cues can signal the need for a pause, allowing learners to manage conversational turns and navigate the challenges of online language learning. Guo and Zhang (2021) look into the organization of teacher-initiated turn-taking in online German L2 sessions during COVID-19 and propose that video-conferencing undermines the availability of embodied resources in establishing co-present participation frameworks, compared to face-to-face classroom interaction.

Other research has focused on the multimodal practices of sustaining the mutual interaction space. Rusk and Pörn (2019) address the issue of latency in video-mediated language learning environments. They examine how learners use various interactional resources to maintain a shared understanding despite technological delays. Their work shows that clearly defined conversational roles facilitate interaction in such settings, with implications for the design and facilitation of online language learning. Tuire Oittinen (2023) focuses on the practice of highlighting within video-mediated learning activities. By analyzing instances of screen sharing, Oittinen highlights the coordinated use of cursor movements and on-screen annotations in tandem with verbal and visual cues to maintain attention and negotiate content. This research contributes to our understanding of how digital interfaces serve as collaborative tools in online education.

EMCA researchers have also been interested in other video-mediated forms. For example, we found literature on video tutorials and live streaming.

Studies concerning instructional videos for practical tasks offer insights into how users engage with this medium for learning and task execution. Heinemann and Möller (2015) contribute to this field by examining the virtual accomplishment of knitting via video tutorials. Their focus is on the challenges that novice knitters face when they are devoid of in-person expert guidance. The research reveals that learners often fail to recognize errors when they first occur, typically noticing them only after subsequent problems arise. Tuncer, Brown, and Lindwall's studies (2020, 2021) look into how users interact with online instructional videos to accomplish practical tasks. Their research highlights the dance between user control, as evidenced by the strategic pausing of videos, and the temporality of tasks at hand. Pauses are utilized for various purposes: locating objects necessary for the task, preparing to enact a step, maintaining pace with the video, and addressing unexpected issues.

Depending on video interactions in real time, the issue of dealing with time emerges. Perry, Juhlin, and Engström (2014) investigate the complex coordination challenges faced by multi-camera live television production teams, particularly in sports TV. The study examines how production teams manage the integration of historical and real-time visual content during live broadcasts. One of the central tasks is the sense-making involved in selecting from multiple visual feeds, including the rapid incorporation of instant replay footage recorded just moments before the broadcast. Licoppe and Morel (2018) explore the structured nature of interactions in Periscope live video streams using conversation analysis. They identify key elements of these interactions: the talking heads orientation common to video communication, the expectation for streamers to attend to messages from viewers as much as possible, and a 'loose' organization of viewer responses resembling multi-party chats. A particular interaction pattern is noted where streamers 'read aloud and respond' to messages, which is aligned with the asymmetrical communication feature of Periscope.

In the following, I will review mainly video-mediated interaction EMCA works relevant to this thesis's topics. That is, video-mediated interactions that investigate the opening, closing, noticing, and food evaluation issues.

First, research on the opening sequences of video-mediated interactions offers critical insights into how social and institutional activities transition from non-interaction to interaction in a

digital space. These studies reveal the varied practices individuals use to establish co-presence, coordinate actions, and engage in activities across different contexts.

With the advent of video-mediated communication, openings in contexts such as video court hearings (Licoppe and Dumoulin, 2010), video consultations (Pappas and Seale, 2011; Stommel et al., 2019), video tele-homecare (Ilomäki and Ruusuvuori, 2020), video conferences (Oittinen and Piirainen-Marsh, 2015; Mondada, 2015; Hoffmann and Fele, 2023), video blogs (Frobenius, 2011), video calls between friends and family members via Skype (Licoppe, 2017) or WeChat (Gan et al., 2020) have been investigated. Licoppe and Dumoulin (2010) examine the particularities of opening court hearings via videoconference, exploring how traditional institutional speech acts are adapted or become irrelevant in a distributed setting. Their ethnographic and sequential analysis highlights the impact of technology on how openings are produced in legal contexts. Frobenius (2011) studies the opening sequences of video blogs, finding that speakers adapt and borrow strategies to compensate for the absence of a live interlocutor. These openings serve not just to initiate interaction but to encourage audience engagement and contribute to identity construction within the computer-mediated communication (CMC) context. Research on video consultations (Pappas and Seale, 2011) discusses the complex negotiations of skills and roles required in telemedicine, where the novelty of the medium adds layers of complexity to the communication between healthcare professionals and patients. The study underlines the need for specific training to navigate these complexities effectively. Lorenza Mondada (2015) looks into the openings of videophone meetings in a medical setting, identifying distinct phases such as 'pre-openings' and 'beginnings' that are separate from the openings proper, which are influenced by the connection and co-presence of participants. The work of Oittinen and Piirainen-Marsh (2015) on technology-mediated business meetings outlines the multimodal practices necessary for establishing co-orientation and entry into the meeting space, emphasizing the coordination of verbal, bodily, and technological acts. Santos Arantxa Muñoz (2016) examines the initial moments of multi-party videoconference meetings, highlighting how participants signal their availability and manage co-presence online. Licoppe (2017) analyzes Skype conversation openings, noting the importance of sequential organization, such as appearing/noticing sequences and the greeting practices that establish a joint interactional frame. Research into video consultations (Stommel et al., 2019) explores the extra interactional work required at the start of postoperative consultations, focusing on the delicate balance between other-attentiveness and the transition to medical agendas. Ilomäki and Ruusuvuori (2020) examine tele-homecare openings for older

adults, observing how participants navigate the boundaries of co-presence through practices familiar with phone conversations and other Skype interactions. Gan et al. (2020) explore the orchestrated openings in video calls between Chinese migrant parents and their left-behind children, demonstrating the emotional significance of these openings and how they contribute to family bonding across distances. Siitonen et al. (2022) focus on waving as an interactional resource in video-mediated breaks, showing how this gesture manages co-presence and supports relationship maintenance work during virtual encounters. Sabine Hoffmann and Giolo Fele (2023) tackle the issue of managing missing participants in the opening phases of a videoconference, revealing the practices that accommodate and adjust to such absences.

Collectively, these studies shed light on the complexity of interactional work in the digital realm, offering perspectives on how individuals manage openings in various forms of video-mediated communication, and showing the stepwise unfolding of the opening phase requires the orchestrated coordination of verbal and bodily resources like “talking heads” (Licoppe and Morel, 2012), waving gestures (Siitonen et al., 2022), as well as the affordances of the technological artifacts utilized, and revising the notion of “opening” of interaction by examining “pre-openings” as distinct from “openings” per se, highlighting the role of video technology in shaping openings.

Second, in recent years, research on noticings has extended to video-mediated interactions. Zouinar and Velkovska (2019) analyze how showings and noticings are used for introducing, changing topics and sustaining talk in mediated conversation in the image-based family and friends video calling. Oittinen (2020) focuses on moments in which either an auditory or a visual problem emerges in multiparty video-mediated meetings and how the participants orient to these troubles through embodied noticings. The research emphasizes the features of embodied noticings in mediated settings, and they note that embodied noticings involve subtle and visual bodily displays, e.g., gaze aversion, gestures, facial expressions, and so forth, projecting the orientation to obtain recipients’ attention to an other-than-current activity (Kaukomaa et al., 2014; Kääntä, 2014; Keisanen, 2012).

Third, to date, very few studies on eating and video-mediated interaction have used a conversation analytic approach. However, some studies have already examined eating activities via VMI from the perspective of examining spoken and embodied practices. Filming people’s eating-related activities and making them into pre-recorded edited videos, or so-called eating

shows (Rüdiger, 2020), occupies a unique intersection between studies on eating activities and video-mediated interaction, capturing the interest of researchers from diverse backgrounds. Eating shows can take different media forms, such as YouTube vlogs and TikTok short videos. Rüdiger (2020) examines the use of language and digital food discourse in North American YouTube eating shows, investigating how the structure and discursive performance contribute to the creation of intimacy and informality (Rüdiger, 2022), as well as how they blur the sense of immediacy and distance (Rüdiger, 2021). Different from eating shows as pre-recorded edited videos, eating broadcasts (or “mukbang”) normally refer to live streams with synchronous interactions (text messages) from viewers. The popularity of eating broadcasts on multiple live-streaming platforms has obtained considerable attention from researchers in recent years. Several studies in conversation analysis have examined the multimodal structures and functions of eating in video-mediated contexts like live streaming (Choe, 2019; Choe, 2020). Furthermore, empirical methods have been employed to explore viewers’ motivations and practices, aiming to understand why people watch others eat food in live streaming. Results suggest that viewers mainly treat eating streams as a primarily private activity, as a mealtime companion, and as a leisure activity (Anjani et al., 2020). Viewers’ motivations include the streamer’s attractiveness (Pereira et al., 2019), multisensory experience, virtual satiation, physiological experiences, spectacle of the streamer’s performative eating (Anjani et al., 2020; Kim, 2020), the desire for remote digital commensality (Spence et al., 2019), an escapist compensatory strategy (Kircaburun et al., 2020), and even sexual reasons. Some conversation analysis studies have focused explicitly on the multimodal aspects of sensory experiences and socialization related to eating (although not specifically in video-based contexts) (Ochs et al., 1996; Wiggins, 2002; Sneijder and te Molder, 2006; Wiggins & Keevallik, 2020). These studies aim to demonstrate that eating behaviors and enjoyment are not solely physiological sensations but also interactive achievements.

Fourth, studies have also investigated closings in video-mediated interactions. They aim to show why closing matters (Raymon & Zimmerman, 2016) and how closing is achieved step-by-step. Ilomäki and Ruusuvoori (2020) not only examine openings in video-mediated tele-homecare as mentioned above, but they also look into closings in it. They note that managing visuality and time-oriented talk is important in both openings and closings and that openings and closings mirror each other in this institutional setting. They demonstrate that closings result in a terminal exchange, and both participants disengage from the encounter: the clients by walking away and the nurses by closing down the connection. They also show that tele-

homecare is an interesting hybrid of institutionality and informality. In his study, Oittinen (2021) uses conversation analysis to examine how technology-mediated meetings are brought to a close. He focuses on the various resources that the chair and participants use to transition from meeting talk to a coordinated exit. Since bodily-visual leave-taking behaviors are not easily accessible, other mutually recognized practices must be implemented to initiate and advance closings. This requires the close coordination of participants' vocal and embodied conduct in the physical setting, which must be rendered publicly intelligible via the screen at specific moments. Oittinen's analysis portrays closings as emergent, collaborative accomplishments that require negotiation and alignment of multimodal turn constructions and behaviors in situ.

To sum up the literature review of both computer-mediated chat and video-mediated interaction, prior studies have demonstrated novel interactional practices have emerged in these environments to utilize and adjust to the parallel existence of two highly different modes of participation. These distinctive features are oriented to the highly asymmetrical communicative affordances of live streams. So far, there are only a few live-streaming studies using the method and theory of conversation analysis, apart from the works by Recktenwald, Licoppe, and Morel as aforementioned.

This dissertation explores the interactional features of live streaming. It is not a monograph but a collection of four distinct articles, each examining a specific aspect of this digital phenomenon. All of the articles have been published or are under review. Article I looks into the initiation phase of live streams, revealing a flexible "installation" phase, unlike the fixed openings in phone conversations. It identifies key interactional challenges such as streamers' anticipation of an adequate audience, collective management of viewer interactions, and the formation of a guest/host relationship, all while maintaining the intelligibility of the stream content. Article II is accepted by a journal. Article II examines the mechanisms of attention and engagement, highlighting the role of noticing-based actions in fostering 'noticing effervescence' within the asymmetric engagement of expository streams. Article III is under review. Article III reinterprets tasting as an interactive spectacle in live streaming, outlining the configuration of "tasting heads," gaze patterns, facial reactions, and the responses of both streamers and viewers. Article IV has been published. Article IV investigates the closing live streams, emphasizing the adaptation of traditional conversational closure techniques to the digital streaming context.

These works contribute to our understanding of how technology redefines interactional practices and sequential organization in live-streaming environments. By examining how participants employ a variety of resources to navigate actions within this digital realm, the dissertation aims to enhance our comprehension of mediated interactions. It paves the way for future research into the interaction practices of live streaming and other communicative technologies, offering insights into the evolving landscape of digital communication.

CHAPTER 2

MULTIMODAL CONVERSATION ANALYSIS AS METHOD

I use multimodal conversation analysis (Hazel et al., 2014; see also Schegloff, 2007) as my research method. My analysis aims to identify, describe, and explain the orderly and recurrent practices adopted by participants during the tasting activities. I closely examine the multimodal resources that emerge within the technologized contextual configurations, including gestures, gaze, body postures, and the embodied manipulation of the phone and the App.

For the conversation analytic approach, the next-turn proof procedure by Sacks, Schegloff, and Jefferson (1974) is an essential guide. To gain insights into the question "Why that now" (Schegloff & Sacks, 1973: 299), analysts can examine how a turn is related to the previous turn and how it is responded to. In other words, people display their understanding of the preceding turn in their next turn, which can then be used for analysis.

Working with live streaming, the application of the next-turn proof procedure can be challenging, because live streaming interactions are intertwined between the streamer's talk and the viewer's text messages, and are highly influenced by the 'disrupted turn adjacency' featured by multi-party chat (Herring, 1999). It has two outcomes for live streaming interactions. First, overlapping, gaps, and silence in live streaming are more complex since they occur between talk and text and even action and text. Second, it becomes harder to identify what is the previous actions as the previous action may not be in the just-prior place. In the following, I will unfold how I implement this approach in detail through my platform selection, data recording, and transcription design.

2.1 THE PLATFORM: TWITTER LIVE

The streaming platform I selected is Twitter (now "X") Live. On Twitter Live, broadcasters can stream live videos while engaging with an active audience that can send messages. To access video streaming, Twitter users can tap the camera icon in the tweet composer on their smartphones and then select "go live." They can also add a description that will appear as a tweet describing the stream before going live. Livestream users can invite guests to join the

broadcast, which can help establish an ongoing live streaming schedule and build a small audience (DeBois, 2021).

Twitter Live bought Periscope in 2015 and was connected to Periscope before 31st March 2021. Since they were connected, they had the same page layout, functions, user accounts, and contents (“Periscope (service),” 2022). During that time, while Periscope remained a standalone application, its live content was fully incorporated into the main Twitter app. Whenever Periscope videos were shared on Twitter, they appeared on users' feeds just like Twitter's native videos (Rein and Venturini, 2018). On 31st March 2021, Twitter disconnected Periscope due to declining usage, product realignment, and high maintenance costs (Periscope, 2019; Team Periscope, 2020), and users can still stream on Twitter.

Twitter Live streams are broadly similar to the video streaming platforms that have been rapidly expanding in the last few years, such as YouTube Live, Facebook Live, and TikTok Live, though each platform may provide streamers and their audiences with distinctive affordances. We have focused on streams documenting everyday life and experiences, though other uses of live streaming have developed, for instance, for telemarketing or for documenting cultural and political events. We also collected a sample of similar streams on other platforms to confirm that the sequential phenomena we describe here are more generally relevant to live streaming.

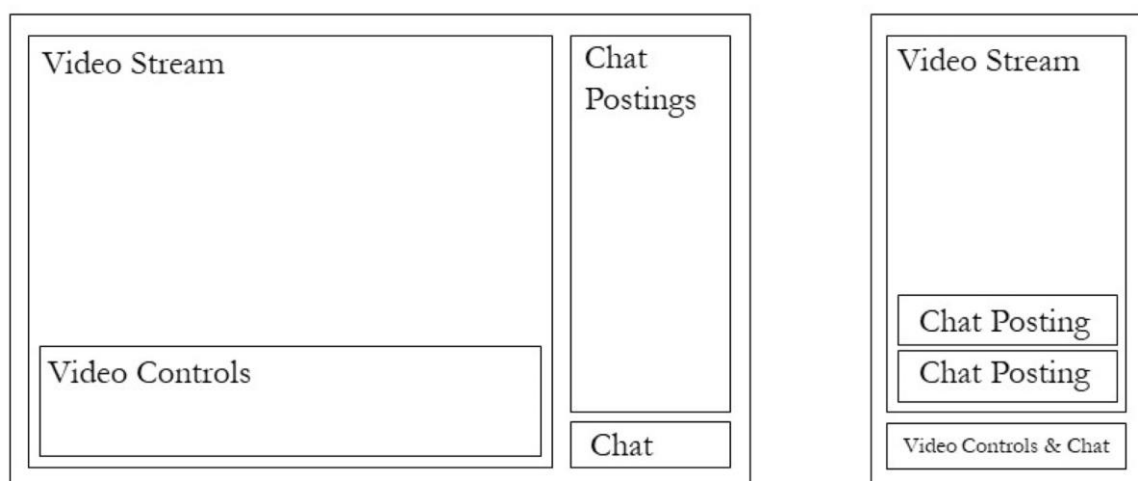


Fig 1 : Schematic illustration of the Twitter viewing area on the stationary device (left) and mobile phone (right) (from Dang-Anh, 2021)

Anyone can watch streams or their replays on Twitter by just opening the site, and you can search for a streamer that you like or browse randomly to get a stream and replay (Fig 2). You

do not have to subscribe to a streamer’s account to watch their streams, although by subscribing you can find their streams more easily next time. Then, choose one stream you like and click in, and you will get an interface with live video and viewers’ live comments superimposed on it. You can type and send your own comments in the bottom entry bar and send out virtual hearts by clicking the heart shape symbol at the bottom right (Fig 3). At the very bottom right corner, you can also see the number of current viewers. To launch your own stream, just put in the title you want for your stream on the top input field “What’s happening?” and then click the red “Go LIVE” button at the bottom (Fig 4).



Fig 2

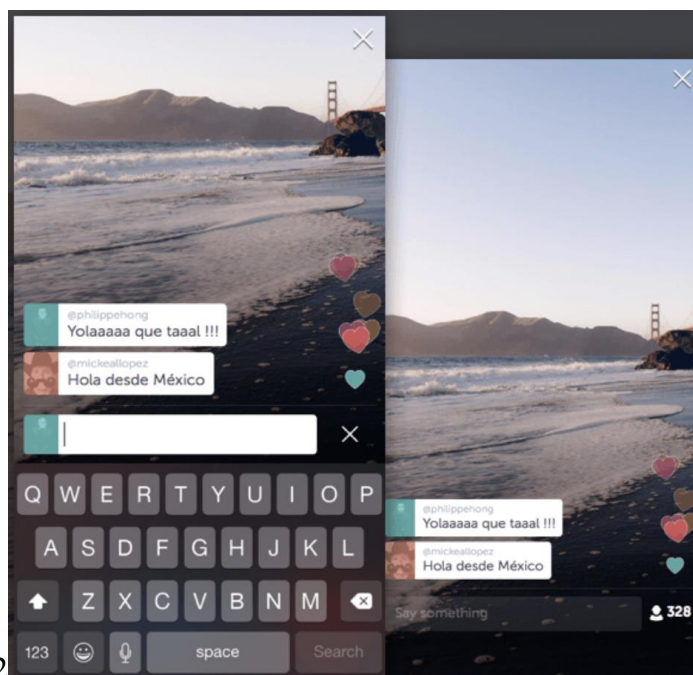


Fig 3

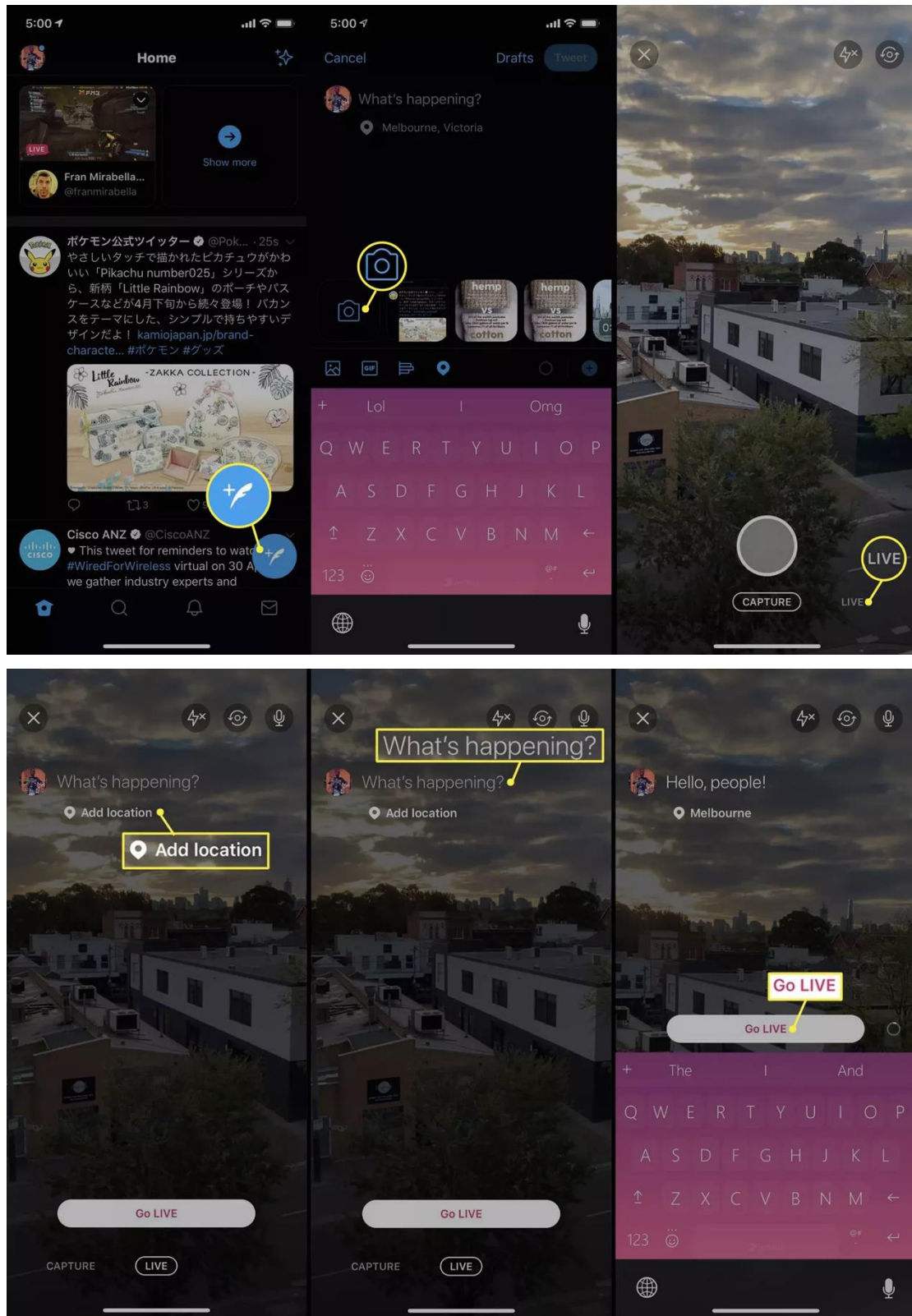


Fig 4

Regarding messages from the audience, they appear on the bottom of the screen (both to streamers and all viewers), where they stay for a few seconds before disappearing. The ability

of the streamer to treat them in this lapse of time may be constrained by their involvement in the setting.

There are multiple contexts and devices (smartphone, drone, GoPro, etc.) that can be used to live stream. While it is often used in a stationary state and a fixed location (like home), it is also used extensively outdoors in mobile settings (e.g., driving, walking). Smartphones are the most used type of device in our data corpus.

Smartphones are the most used type of device in our data corpus. Features and functions vary by device. For mobile phone live streaming, almost all the functions of the phone's camera could be used in the broadcasting, for example, zooming in, zooming out, and flipping the camera. Flipping the camera is a widespread action in streamer's live broadcasting, and it means flipping the phone's rear and front camera by pressing the flipping button on the screen (or by just double-clicking the screen or by using gestures for some devices and platforms) (Fig 5), to achieve the image switches between the streamer themselves and the environment. We will mention it again in the analysis part of the paper.



Figure 5. The flipping camera function in a typical mobile phone

Getting the audience to be aware of live streams is an issue for live streaming platforms, which they address differently (Fig 6). Twitter allowed one to find streams by researching titles, or more originally, through the use of localization. Streams are geotagged, and potential viewers could navigate a world map, and active streams (as well as recorded streams of the day) would become visible. For other platforms, the way to address these problems can be very different. Take TikTok live streams, for example; it is either through subscription to the chain of a regular

viewer or through the mediation of an algorithm using the viewer's previous history of the platform that viewers are alerted to or proposed live streams (Fig 7).

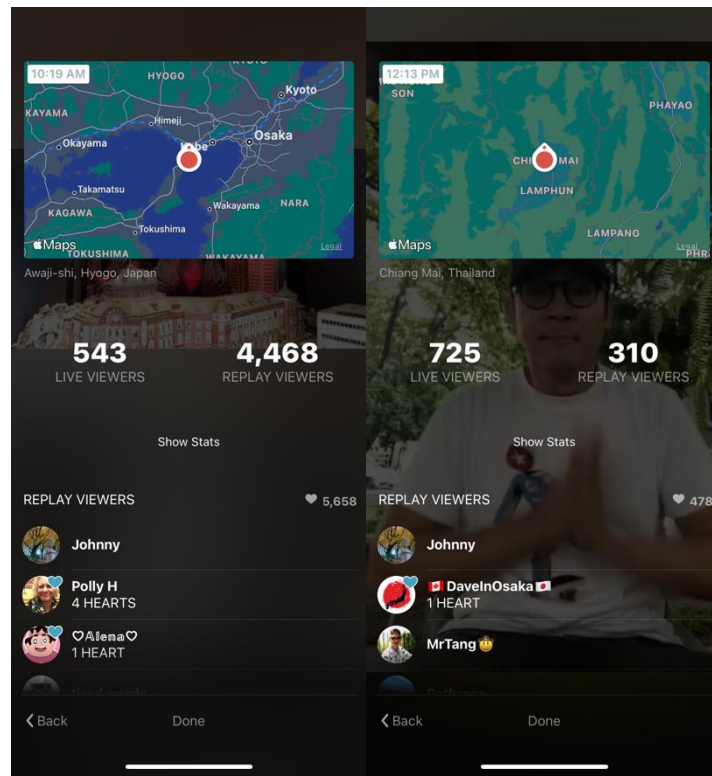


Figure 6. Engaging viewers is a task for every streamer (the backstage stats of viewers for live streams on Twitter. Data retrieved on 14 May 2023)

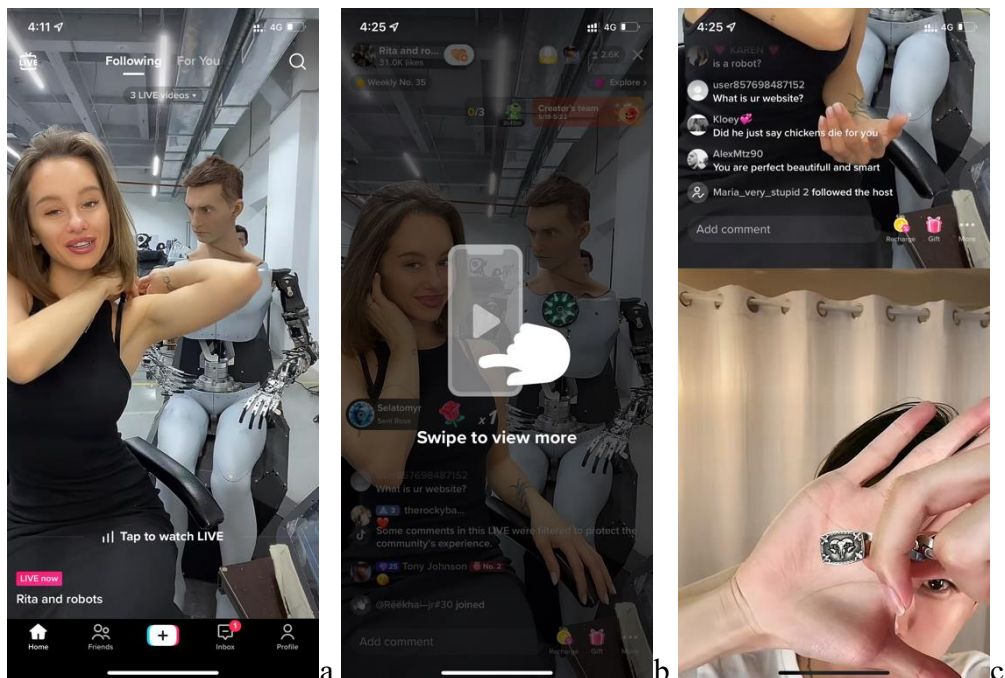


Fig 7: a) Preview mode of a TikTok stream before you click in and formally enter the stream. On the screen you can see “Tap to watch LIVE”; b&c) the “Swipe” mode of watching live streams on TikTok

My research finds the following Twitter Live characters to be of great significance:

- Archived past streams with live comments provide researchers with a convenient way to access relevant video material for analysis.
- The majority of streams have a reasonable length, usually no more than an hour.
- Streamers come from all over the world and provide a diverse range of content.
- The interface layout is easy to understand, concise, and intuitive.

2.2 DATA COLLECTION AND DATA CORPUS

I need to record the streams to establish my data corpus for multimodal conversation analysis.

The recorded streams provided me with a) a sense of the settings in which they used live streaming applications; b) a sense of how interactions are going between streamers and viewers, or between the use of the live apps and other activities (which would be directly observable on the video) or outside the screen (which would be indirectly observable through the absence of the streamers on the recordings); c) rich data on the production of message-mediated interactions, recorded as they actually unfolded in time. The recordings provide a kind of access that is close to that of the streamer or their audience (though it may happen that some messages do not appear, or that the lag between video and messages are different).

The streams themselves are public and available online without a subscription. In line with French regulations, recording and using data on public platforms is allowed if the privacy of everybody recorded is preserved. I have done systematic anonymization and pseudonymization work on the data presented in the paper. I have changed every account name in the data transcriptions and blurred bystanders, car plates, etc., for each screenshot we use, so nobody should be recognizable in our exploitation of the data.

There are two main criteria for my selection of streamers and streams: a) The main language used by the streamers and viewers is English; b) I decided not to allow for streams shorter than 2 minutes, and the majority of our streams are in the 10'-30' range. The first criterion is to ensure the representativeness of our data since the most used language on Twitter is English; the second is to ensure that there are enough viewers to elicit interesting episodes of interactivity

(more than five active users is usually enough) for every single video. Apart from these criteria, streamers and live broadcasts are selected at random.

I have recorded 92 broadcasts from 34 streamers in our corpus (Fig 8). They were recorded from 2020 to 2021, with a total duration of over 20 hours. Given that we needed to study the opening and closing sequences of the streams, I recorded most of our data from the archived past streams, which were already in their complete length. I started ethnographic participation (Hsiao, 2022) with a list of random streamers on Twitter Live for almost a year before conducting this study. I initially recorded more than 92 streams and 34 streamers but ended up with these numbers when I believed they contained all types of streams and had reached the expected total length under our selection criteria. Most of the streams are daily life-featured. I watched their streams, read their posts, and followed their activities and streams not only on Twitter but also on platforms such as YouTube and Happs, if they also stream there.

To prevent loss, I store the recorded videos on three different external hard drives and number them. I name each streamer's videos in a folder with their account name (e.g., ChantalTV) or an easily recognizable alias (e.g., Boston Girl, Swiss) and prefix them with numbers, such as 01. Using the account name makes it easier to search, while adding numbers helps with organization. Subsequently, any newly recorded videos of that streamer are placed in their respective folder and renumbered (Fig 9). The benefit of this approach is that if a streamer's recorded videos change, you can simply modify their respective folder without affecting the overall numbering. Finally, I include a screenshot of the streamer's account homepage in each folder, displaying their account name, self-description, number of followers, number of following, and the number of live streams as important reference information for future use.

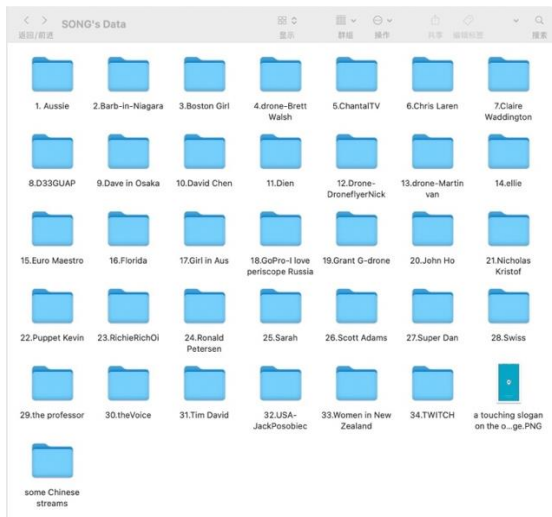


Figure 8. Screenshot of the data corpus

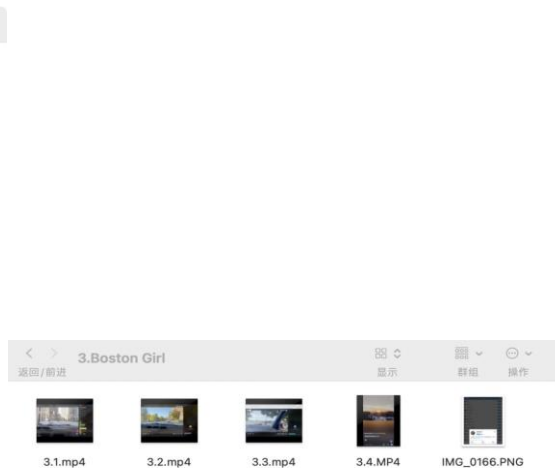


Figure 9. Content in each file

2.3 DESIGNING A MULTIMODAL TRANSCRIPTION SYSTEM FOR INTEGRATING VIDEO AND TEXT

Transcription of data is necessary and a widely acknowledged challenge for online multimodal conversation analysis (Dutt & Graham, 2023). For live streaming, the transcription system goes beyond purely ordinary talk, video call (Licoppe & Morel, 2012; Gan et al., 2020), or written chat (Meredith, 2015). Interactional practices in live streaming have been less studied from a conversation-analytic perspective (but see Licoppe & Morel, 2018; Recktenwald, 2018; Choe, 2019, 2020; Hsiao, 2022; Dutt & Graham, 2023; Song & Licoppe, 2023; Song & Licoppe, 2024). Two asymmetric modalities characterize the interactional features of live streaming. The first is the audio-visual presence of a streamer. Since the streamer appears live in video, participants in live streams orient to the normativity of the “talking heads” format, which was observed to characterize video-mediated interaction (Licoppe & Morel, 2012). The second is computer-mediated chat (see Paulus et al., 2016), which is the prescribed modality for viewers’ contributions. In live streaming, messages that are responsive to some prior action usually appear with a delay. This is the live streaming analog to the phenomenon of “disrupted turn adjacency” observed in multi-party chat interactions (Herring, 1999). Streamers can be treated as accountable for non-responding (Licoppe & Morel, 2018), and the audience may feel it legitimate to ‘pursue a response’ (Pomerantz, 1984).

To analyze this complex interplay, this thesis presents a multimodal transcription system that coordinates the visual, verbal, and textual elements of live streaming. The system primarily

focuses on synchronizing the video image with the streamer’s spoken and bodily actions, alongside the viewers' chat or typed actions, offering a holistic view of the communicative event.

Current methodologies in transcribing live streaming content can be broadly categorized into two styles. The first style separates the transcriptions of the streamer and the viewers, often presented in a table-like, multi-column layout. Scholars such as Recktenwald (2017), Choe (2019), and Hsiao (2022) have utilized this approach, making it easy to distinguish between modalities with fewer symbols embedded (see figs. 10-12). However, this style has limitations in demonstrating the ‘overlapping’ of the modalities more precisely.

Timestamp	Game Events	Streamer	Chat
218 :[00:01:00]			<the_tacos_doctor> @tsm_bjergsen theoddone he sucks so bad why do u play with him?
:			
239 :[00:01:06-6]		like I was on some of that good stuff.	
:			
252 :[00:01:10-1]		holy shit that charm was awful.	
253 :[00:01:11]			MOD: obskunia
254 :[00:01:11]			<tylaaa69> say ok if youre a rammus
255 :[00:01:12]			<maroter> Bjerg been spending to much time with goodguygary and now he all about that 420
256 :[00:01:12]			<bustedproject> @Stryker135 I dont understaned
257 :[00:01:12-2]		oddone is fucking dead though.	
258 :[00:01:12-6]	Riven /Megazero\ kills Olaf /Odone\		
259 :[00:01:13]			<alairaine> sounds like it was a fun experience
260 :[00:01:14-4]		{gaze shift to secondary monitor}	
261 :[00:01:15]			<rexdanknightftw> @Killer_ranger_2217 🤔
262 :[00:01:16]			<frozenliquidz> LOL that good good

Figure 10: Recktenwald’s (2017) transcription for Twitch game streaming


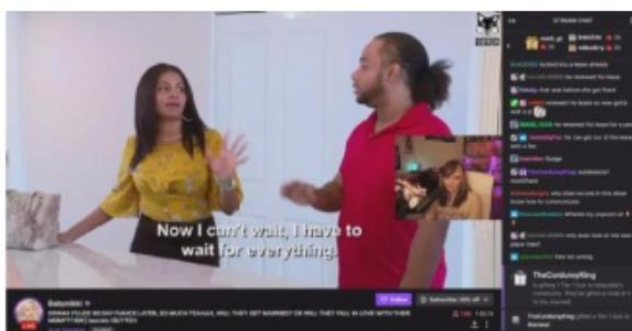
Image	BJ ChangHyun		Viewers			
	Line	Spoken action	Bodily action	Line	Typed action	ID
	5		Eyes looking at the chat screen	C	/굿/ ‘/Good/’	ltk145
				D	재밌었는데 ‘It’s fun though’	bibi333
	6	음 떡볶이도? ‘Umm teokbokki?’				

Figure 11: Choe’s (2019) transcription for Mukbang on AfreecaTV

Timestamp	((Paralinguistic features or body language))	Excerpt 2-4:
28:48	1. 他說	E. 擦擦擦 : 🙄🙄
Nida:	He [Somer] said,	F. 45 - 乾坤編 : 剛進來的時候真的覺得你很奇怪跟其他我看的不一樣
28:48	2. 我覺得差別-是在-	I truly thought you were weird when I started to watch your channel. You were different from other streamers.
	"I think the difference [between my feelings then and now]-is-"	G. 94 007 : 自然梗就很好笑
28:53	((pausing for approximately five seconds before uttering))	A punchline of a joke that comes naturally is great.
	3. 因為-大部份-	H. 45 - 乾坤編 : 結果你留下我了
	"because-most-"	It turned out you made me stay (on your channel).
28:53	4. 他說(.)大部份-He said(.)	I. 龜頭 : 我也是~了解你之後覺得你很幽默
	"most-"	Same here. I think you are humorous after getting to know you.

Figure 12: Hsiao's (2022) transcription for streams on Lang Live, in which she separates the streamer's actions (left) from the viewers' messages (right).

In contrast, the second style, which integrates the transcriptions of both streamers and viewers into a single narrative (see Licoppe & Morel, 2018; Dutt & Graham, 2023, fig. 13), better addresses issues such as 'overlapping' and sequential flow. The disadvantage is that it could look complex to understand when video, text, talk, and symbols are embedded together. Special symbols need to be created to show the connectedness of the modalities (see fig. 14).



12	40:29	NIKKI:	He's gas lighting HE:::::::::R! (1.4)
56	40:58		↳ halloneon: what does gaslighting mean? @Babynikki
64	41:05	NIKKI:	→ (1.3) (Vid.Pause) ((reading chat)) "Wha'does gaslighting mean,"=
65a			→ Gaslighting basically just mea- just uh how- (0.3) how?- (0.6)
65b			hold on. (0.4) ((grabs phone))

Figure 13: Dutt and Graham's (2023) transcripts of the TV show re-boasting on Twitch. We can see that the streamer's talk (lines 12, 64, 65a, b), embodied actions (vid. Pause, reading chat in line 64, grabs phone in 65b), as well as the viewer's text message (line 56), are integrated together. The authors use the symbol "↳" to mean textual messages are nested below the spoken utterance ongoing at the moment of their arrival.


1. WAR: combine ça coûte ton télescope
how much does it cost your telescope
2. (3.0)
3. S: combine coûte mon telescope euh:*:: hhh hhh
how much costs my telescope e:::r
4. HAM: * **tu peux pas zoomer plus**
you cannot zoom more
5. S:*alors il y en a pour plus de mille cinq cent euros
so there is more than one thousand *five hundred euros
6. GHA:***ton telescope a beaucoup servi 'image est pas très net**
your telescope has been used a lot the image is not very clear
7. EMR: ***Oww shit**
8. S: pis après tu comptes *le materiel (.) les lentilles et
and then you have the whole set the lenses and
9. EMR: * **My dream**
10. S: tout le tralala qui vont avec
everything which goes with it
11. (2.0)

Figure 14. Licoppe and Morel's (2018) transcription of streams on Periscope. Bold texts are used to make streamer and viewer's actions distinct. The symbol '' is used to indicate the 'overlapping' of text and talk.*


We favor and adopt the second approach because it suits mobile platforms. Most of our streams are played on mobile phones, where text messages are superimposed over video images. The transcription thus mirrors the actual interactional framework viewers encounter.

Screenshots are usually incorporated into the transcription to capture critical moments more effectively. They can be placed outside the text (see Fig. 13 and Fig. 16 below) or embedded within the transcript to highlight significant instances of the streamer's embodied actions (see Fig. 15).

92 → dread_dormammu: it's about exploiting someone's
 93 trust/intimacy as well



94 41:33 NIKKI:
 → ((gazing down at phone)) "Gaslighting, to manipulate someone by
 → psychological means



95 → into questioning their own sanity."

Figure 15: screenshots are inserted within the flow of the transcript.

Timestamps are another crucial element used in live streaming transcripts to show sequentiality and temporality. By integrating timestamps, the system allows for granular analysis of the temporal dynamics within the live stream, such as response latencies or the pacing of interactions.

We believe that a transcript should adopt established conventions to describe common features where possible (Recktenwald, 2017). No transcription scheme is ever final or complete. It should be possible to amend a transcription to adapt it to different purposes. We adopt Gail Jefferson's (2004) transcription conventions for talk and Lorenza Mondada's (2018) conventions for embodied actions. Courier New font in nine-point is deployed for the transcripts. Figure 16 below serves as an example of my transcription in my thesis, and Table 1 provides detailed information regarding my transcription conventions. In all my transcripts, 'S' represents the streamer. The position of a message in a transcript indicates the moment it first appears on the screen. We added the exact timestamps for some extracts to give the transcripts a better demonstration of sequentiality, which is also facilitative for our analysis. Besides, some particular conventions for live streaming have referred to Licoppe and Morel's work in 2018.

For example, we use the symbol ‘§’ to indicate the ‘overlapping’ of the message’s appearance with the streamer’s talk. We use normal font words to show the streamer’s utterances and conducts and **bold font** to show the audience’s text comments with their pseudonymized account names. Finally, it remains a question of how to present the time that messages stay on the screen and if we should transcribe it or not. In many situations, we see that it also impacts interaction, but a fine way should be proposed to avoid the transcription becoming too ‘narrow’ (Recktenwald, 2017; Dressler, 2000:27) and unnecessary.

```

1      S: We'll say #6.1, good night or have a great day,
2          and go$odbye for Grisons.
3  BOX:      §thanks for the View 🙌
4          (3.0)
5          Goodbye for Switzerland.
6          Talk to you soon.
7          Bye everyone.
8          Bye #6.2
9  VIC: I'll be looking forward to it. bye
10         (28.0) ((moves camera slightly to the right #6.3))
11     S: ((disconnects))

```

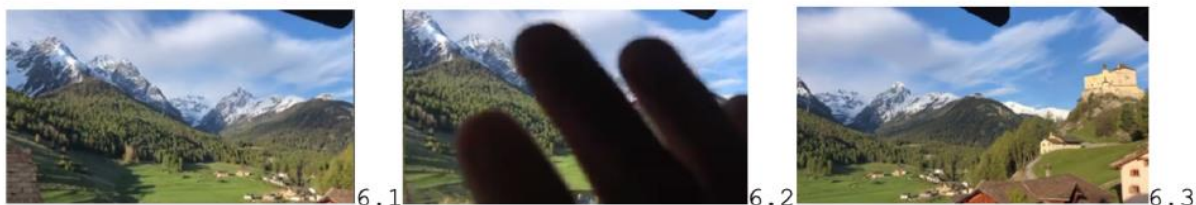


Figure 16: a typical transcript in my thesis

S	the streamer
normal font	the streamer’s utterances and actions
bold font	the audience’s text comments
,	intonation continuation
.	downward intonation
\ and /	the falling and rising intonation, respectively
(.)	inconspicuous short pauses of the streamer’s talk (usually less than 0.2 seconds)
(2.0)	the silence time of the streamer’s talk when more than 0.2 seconds

.hh	to inhale or laugh
§	indicates the moment when the audience's messages come up
---> ---->	asterisk marks the moment and the phase (when the action continues longer than just a moment) of the action; dotted line indicates the continuation of the action when the action continues more than one line; The description of the action is normally presented in the subsequent line in gray in double brackets, and is also marked by an asterisk
#	indicates where the screenshot appears
£ £	the content within the symbol is said when laughing
=	indicates the connection of two turns
(())	indicates non-verbal behavior

Table 1. Some of the main transcription conventions for live streaming in this thesis

CHAPTER 3

THE ORGANIZATION OF LIVE STREAM OPENINGS

Abstract

We present here the first empirical study of the organization of live stream openings. We demonstrate that beyond their apparent complexity and variability, they feature recurrent interactional episodes and sequences, with significant variations between the 'expository' and 'experiential' genres of live streaming. Our analysis links these sequences with particular participation frames and membership categorization relevancies: streamer-initiated collective greetings and announcement sequences with the streamer/audience categorical framing, interpersonal greeting sequences that we show are relevant not only during the openings but throughout the live stream with the guest/host framing, and initially-positioned noticing by streamers as enacting a collective viewing framework. Furthermore, we illustrate how, throughout all kinds of live streams, participants display a strong concern for the immediate intelligibility of live streams during openings. This provides a better understanding of the interactional concerns and expectations of participants in different genres of live streaming.

Keywords: live streaming; Twitter; opening; sequence organization; multimodal conversation analysis; mediated interaction; video

Introduction

Live streaming is a relatively new practice wherein streamers record and broadcast themselves live on a growing number of dedicated platforms for an audience of online viewers who may log in and log out relatively freely. Understanding live streaming as an interactional phenomenon presents three main difficulties.

First, the uses of live streaming are remarkably varied and still diversifying. Teenagers use live streaming to display themselves and chat with their audiences (Licoppe & Morel, 2018). Live streaming is also utilized to share activities or performances, such as online video gaming, which played a major role in the initial development of live streaming, with platforms such as Twitch (see, for instance, Recktenwald, 2017; Jodén & Strandell, 2021), food tasting (for example, Choe, 2019; Anjani et al., 2020; Rüdiger, 2022), or teaching during COVID (Chen et

al., 2021; Marulanda et al., 2022). In all cases, some form of human activity is transformed into a live and public experience by being live-streamed. Additionally, live streaming is utilized to promote commercial goods, such as a form of telemarketing that draws large audiences in China (Si, 2021), or as a resource for political activists to document and broadcast live their participation in the unfolding of significant political events to engage online audiences (see, for instance, Kavada & Trere, 2020; Fang, 2023).

Second, partly because of the sheer variety of settings and activities being live-streamed, it has not proven easy to achieve a typology of live streams. However, an important distinction is between “expository” streams and “experiential streams” (Tang et al., 2016):

- In expository streams, the streamers “are directly explaining their opinions or showing their skills through chatting, demonstrating, or hosting “how-to,” ask me anything, or talk show sessions. These activities give the streamer a chance to expound upon their personal perspectives that constitute their brand directly”.
- In “experiential streams,” streamers want to share something they think is interesting, like an event, scenery, social gathering, or something funny.

One reason why this broad distinction is helpful is that we can find interactional markers tying each genre to specific interactional practices. For example, in expository streams (or during expository moments in a given stream), the streamer strives to engage conversationally with the audience. It is expected to appear in a ‘talking head’ configuration (Licoppe & Morel, 2018). Conversely, experiential streams are usually done (and expected to be done) so that the streamer visually shows the perceptible features of the relevant experience and removes herself visually from the screen. As we will try to show later, the distinction is also relevant with respect to the organization of openings.

Third, live streaming is a highly asymmetric type of encounter, for while viewers can watch and hear the streamers live, the streamer can only see written notifications (for instance, when a viewer logs in to the stream) and written chat messages sent by the viewers. So, live streaming will reshape and articulate expectations and practices relevant to both video-mediated interaction (Due & Licoppe, 2021) and multi-party online discourse (for a review, see Paulus et al., 2016). This articulation of different modalities and temporalities of interaction makes live streaming an exciting phenomenon from an ethnomethodology and conversation analysis (EM/CA) perspective. This approach has been used to tackle issues such as transcription and the organization of engagement between the streamer and her online audience (Recktenwald,

2017), the lamination of participation frames (Choe, 2019; 2020), the multimodal character of live streaming interaction, and especially the orientation towards the ‘talking heads’ format and the use of the ‘read-and-respond’ practice (Licoppe & Morel, 2018), temporal coherence and the ‘pivotal’ or ‘choral’ role of some sequences (Recktenwald, 2017; Dutt & Graham, 2023), the importance of ‘noticing-based actions’ in live streams (Song & Licoppe, 2024) and the organization of closing sequences (Song & Licoppe, 2023).

EM/CA has brought into focus the sequential organization of openings and their importance for understanding various encounters as interactional phenomena, and we utilize that approach here for the first time to examine the openings of live streams. Openings serve as one of the places where participants' concerns regarding the type of live-streaming encounters they wish to achieve become salient and observable as concrete practices and sequences. However, as we will demonstrate, the openings of live streams exhibit a high degree of variability and complexity. There is no ‘canonical’ opening sequence in live streams, as there is for phone conversations (Schegloff, 1986). Similarly, although the openings of video-mediated encounters (see, for instance, for institutional settings Licoppe & Dumoulin, 2010; Mondada, 2015; Oittinen & Piirainen-Marsh, 2015; Ilomäki & Ruusuvuori, 2020; and regarding video-mediated communication between family and friends, Licoppe, 2017; Gan et al., 2020) and chat-based encounters (see, for instance, Rintel et al., 2006; Stommel, 2016; Marmorstein, 2023) have been studied in some detail, many observations made in such settings are not applicable to live streams due to the hybrid nature of the latter.

To bring some order to the complexity of live stream openings, we have analyzed them in relation to the concerns expressed by participants, the participation frames (Goffman, 1974), and the membership categorization (for membership categorization analysis, see, for instance, Hester & Eglin, 1997; Fitzgerald & Housley, 2015) that they enacted and deemed relevant. This approach to the openings of live streams required us once again to differentiate between expository (or conversational) streams and experiential (or display) streams, but also enabled us to refine this distinction further.

2. Data and Methodology

2.1 Twitter Live

We chose Twitter Live as our platform of focus. Past streams on Twitter Live are conveniently archived along with their comments, facilitating researchers to access pertinent video material

for analysis. Streamers on this platform are dispersed globally, contributing to a diverse array of content. The Twitter Live interface empowers viewers to engage actively by typing and sending comments through the bottom entry bar. Additionally, viewers can express appreciation by sending virtual hearts, accessible via the heart-shaped symbol in the bottom right-hand corner. Adjacent to the heart symbol, the interface displays the current number of viewers (Fig 1). When launching a stream, users have the option to provide a title for their broadcast.

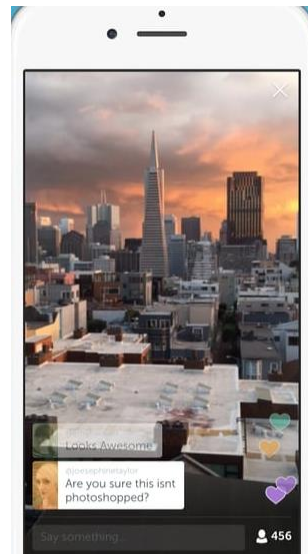


Fig 1. A Typical Stream.

Smartphones emerged as the predominant device type in our data corpus. The versatile functionalities of smartphone cameras, such as zooming in, zooming out, and flipping the camera, are frequently utilized during broadcasts. In terms of audience messages, they typically appear at the bottom of the screen for both streamers and viewers, remaining visible for a brief duration before disappearing. The streamer's ability to address these messages within this timeframe may be constrained by their involvement in the setting and the number of engaged viewers. A sizable audience of active viewers, typically exceeding one hundred, coupled with a rapid influx of messages, can pose challenges for the streamer in responding promptly. Conversely, when the number of viewers is fewer and messages linger longer on screen, systematic responses from the streamer become feasible, fostering the development of interactional sequences (Licoppe & Morel, 2018).

1.2 Data Collection and Analysis Method

Our corpus comprises 92 broadcasts from 34 different streamers, providing a rich diversity of streams that proved sufficient for our analytical purposes. These streams were recorded between 2020 and 2021, accumulating a total duration of over 20 hours. To facilitate the study of closing sequences, we primarily sourced our data from archived past streams. Our focus was primarily on English-speaking streams, which are predominant on Twitter. We specifically targeted streams with typical durations ranging from 10 to 30 minutes, ensuring that a substantial number of viewers were present to foster engaging episodes of interactivity. Typically, streams with more than five active users meet this criterion.

The streams were publicly accessible online without requiring a subscription. In accordance with French regulations, recording and utilizing data from public platforms is permissible as long as the privacy of all recorded individuals is safeguarded. Consequently, we meticulously anonymized and pseudonymized the data featured in our paper. This involved altering all account names in the data transcriptions and obscuring bystanders, car plates, and any other identifiable information in the screenshots we utilized. While the recordings offer a level of access akin to that of the streamer or their audience, it is worth noting that occasional discrepancies may occur, such as certain messages not appearing or variations in the lag between the video and the messages.

We employ multimodal conversation analysis (Hazel et al., 2014; Schegloff, 2007) as our research methodology, with our transcription conventions detailed in the Appendix. Within our transcripts, "S" represents the streamer, while bold text denotes messages from the audience, pseudonymized with their account names. The message's position in the transcript indicates when it first appeared on screen, with timestamps included to illustrate temporal dynamics. Following the next-turn-proof procedure (Sacks et al., 1974), our study identifies, describes, and interprets the systematic and recurring multimodal practices utilized by both streamers and audiences in doing and treating noticings.

3. A continuous flux of incoming viewers: the pervasive relevance of greetings

While greetings and greeting sequences typically occur primarily at the onset of live streams, it is crucial to note that viewers can join the stream at any moment. Therefore, from their perspective, their engagement in the stream commences upon connecting. Consequently, it

remains pertinent for some viewers to offer greetings upon joining, resulting in greeting messages consistently appearing throughout the stream, even well after its beginning. Let us consider Extract 1.

In this live stream, the streamer captures and shares his experience of Chinese New Year celebrations in New York as he walks through the parade.

Extract 1. Title: Happy Chinese New Year

1 S: ((silently filming))(10.0)



image 1.1

2 53:56 NEY: **Hi Tim!**

3 (4.0)

4 S: Hi NEY (1.0) welcome (0.5) to Chinatown.

5 Happy Chinese New Year.

6 (20.0)

7 S: ((scroll for old comments))

8 (10.0)

9 54:40 NEY: **Hi hiiii. Happy year of the pig**

The greeting message in line 2 appears an hour after the beginning of the stream, following a period of silence from the streamer. In line 4, the streamer acknowledges the greeting and proceeds to welcome the viewer to Chinatown, extending a New Year's wish in lines 4-5. Subsequently, the viewer responds with an enhanced salutation and reciprocates the New Year greeting in line 9. Notably, the streamer's response, "Welcome to Chinatown," evokes the mannerisms of a host welcoming a guest to a party. This action frames the interaction as the inception of an encounter, enacting the streamer as the host and the viewer as a guest. This example illustrates how greetings can manifest at any juncture within the stream, maintaining relevance, and potentially initiating a more extended interpersonal exchange.

In the following example, the streamer is sharing his tasting experience at a food festival in Japan, and approximately 5 minutes have elapsed since the beginning of the stream.

Extract 2 (Title: "Osaka food festival part 2: The eating continues, #Japan travel")

1 S: It's important it's not too thick#

2 4:48 CYN: **@KAN hello**



image 2.1

3 S: and the quality of the chicken \$is excellent.
4 **4:52 CYN: Hey Dave**
5 S: *it's not kind of (.) doesn't have (.) the sinewy kind
*((moves fingers as if assessing the feel of something))



image 2.2

6 S: of thing that's hard to swallow
7 **4:58 KAN: @CYN Hy CYN**
8 S: Hi (.) Hi CYN
9 **5:00 CYN: lovely day there**
10 S: the screen is super bright now so I can't read
11 S: people's comments
12 (2.0) ((masticates))
13 S: mm there's a sweetness in there

The situation here is somewhat more intricate. Initially, the viewer greets another viewer in Line 1, then greets the streamer in Line 4, addressing them by their first name, indicating a degree of familiarity. The streamer reciprocates the personal greeting in Line 7. Following this, the streamer issues both a generic greeting, "Hi," and a personalized one, "Hi Cyn," in quick succession in Line 22. This action appears to be a form of correction, possibly stemming from difficulty in grasping the actual organization of this multiple sequence of greetings as he looks down at the screen, with the streamer producing such an account immediately after (Lines 10-11).

This example suggests that familiarity with someone may further emphasize the relevance of greeting sequences when viewers join the stream. Goffman noted that while one might require a reason to interact and greet a stranger in public, there is an obligation to engage and greet acquaintances (Goffman, 1966). Live streams also resemble public spaces, and a similar perspective seems pertinent, explaining why an incoming viewer might greet both the host and another viewer they know. Therefore, acquaintanceship strengthens the likelihood of initiating greeting sequences when a new or familiar viewer joins the stream.

Another phenomenon that enhances the relevance of such delayed greetings for viewers who are joining or have recently joined is the "talking head" format (Licoppe and Morel, 2018). It often occurs that a streamer, who has been sharing their surroundings, switches the camera to

talk, suddenly becoming visible as a talking head. While this signals and establishes a different and more conversational form of interaction with the viewers, it also makes delayed greetings more likely and relevant. This bears a resemblance to video-mediated communication, where the initial appearance of a participant renders a greeting sequentially relevant, forming a distinctive adjacent pair known as the "appearance-greeting" sequence (Licoppe, 2017).

Consider the following case: The streamer was documenting his hiking journey when, without prior notice, he switched the camera and appeared on the screen.

Extract 3 (Title: #Lavaux #wineterrasses #Lemanlake #Switzerland #travel)

1 11:05 DIS: looks gorgeous
2 S: ((flips camera))

Image 3.1

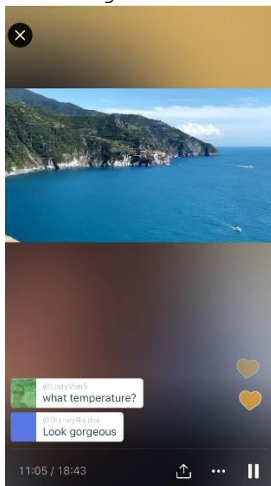


Image 3.2



3 S: ((waves)) hello:
4 (8.0)
5 11:17 ABE: hellooi
6 S: so: (.) ya I it's warm but it's not hot
7 11:20 ABE: wear mask
8 S: \$it's pretty (.) at the moment
9 11:22 THE: Similar to Greece?
10 11:23 ABE: ((smiley))
11 11:24 AWS: Hey!! There he is::
12 (3.0)
13 S: no you don't have to
14 11:29 RUR: □□
15 (2.0)
16 11:34 DAN: Maybe he's converted to be a Roman Catholic
17 11:34 AWS: Good to see ur face!
18 S: you can if you're- if you're (near) to people
19 and cannot keep so (full) distance

After appearing, the streamer greets and waves (Line 3), displaying his own orientation towards the fact that his appearance makes a greeting relevant next. The return greeting in line 11 is especially noteworthy because it is followed by a remark that emphasizes the streamer's

appearance as a noteworthy event ("there he is," line 11), followed by a second message that further enhances the acknowledgment with a positive assessment ("Good to see ur face," line 17). Consequently, such visual appearances by streamers are perceived as noticeable and making relevant greetings, even if the stream has been ongoing for some time.

In summary, greetings maintain relevance throughout the livestream as new viewers participate. While streamers might opt to overlook such greetings without explicit repercussions, responding to them serves as a means for streamers to demonstrate their engagement with viewers and continually present themselves as hosts. The relevance of initiating and responding to these "late" greetings is heightened by acquaintanceship. Their occurrence isn't purely arbitrary; certain interactional events, such as the streamer's sudden visual appearance, position them as appropriate next actions for viewers, particularly those who have recently joined. With this in mind, we can now examine how the beginnings of live streams are structured, as these initial moments are a context where opening-related actions hold relevance and where new viewers continuously join the stream.

4. Live stream genres and openings: Two contrasted cases

When exploring live stream platforms, one encounters a wide range of configurations. On one end of the spectrum, we observe live streamers visually showcasing and promoting themselves to an audience of unseen viewers, actively engaging in conversation with them. On the opposite end, there are streams seemingly documenting external events for an online audience, where the streamer minimizes their presence: in extreme cases, they are not visible and silent, and their existence is only deducible indirectly from their camera work and the stream's title. Some researchers have attempted to delineate these differences by introducing distinctions such as "expository streams" and "experiential streams" (Tang et al., 2017) or "conversational streams" and "display streams" (Song & Licoppe, 2023), though these classifications remain somewhat broad. Instead of adhering strictly to these genre differentiations, we will follow Erving Goffman's approach. Goffman argued that the efforts made by participants at the onset of activity episodes to establish them as such provide valuable insights into how participants perceive and engage with these forthcoming activities (Goffman, 1974).

4.1 The opening of “expository” or “conversational streams”

In the following example, the streamer arranges both himself and his device in such a way that he becomes visible upon connecting (Image 4.1) and adopts a "talking heads" format when interacting with the audience (Images 4.6 and 4.7 below).

Extract 4 (Title "What's the future like?")

1 0:00 S: ((opens the stream in the image of looking somewhere else))#
Image 4.1



2 0:04 **ORA** joined

3 0:05 **LEX** joined

4 0:07 **CRL** joined

5 S: ((sits up straight))
Image 4.2



6 S: ((strikes hands together# and rubs hands together for some seconds#))



7 0:10 **PAU** joined

8 S: ((puts hands down#, leans forward towards the screen and smile#))

Image 4.5

Image 4.6



9 0:13 S: Good afternoon people
 10 0:14 **LUI joined**
 11 S: Hey PAU (smiles) (2.0)
 12 0:16 **LRW joined**
 13 S: hey hey hey, it's Monday again, waiting for \$my coffee.
 14 0:22 **PAU:** \$Hello □
 15 (1.0)
 16 0:23 **YAN joined**
 17 S: How is everyone doing,
 18 (1.0)
 19 S: hey, LRW ((reaches out for screen))
 20 0:28 **LIZ joined**
 21 0:31 **YAN: Hi JOH**
 22 0:34 **DOU: Hey**
 23 S: Hey HEN (laughter) \$YAN (laughter)
 24 0:37 \$TJO joined
 25 0:39 **RIC joined**
 26 S: hey DOU, how is it going,
 27 0:42 **SOF joined**
 28 S: err, mm so I had a bar/becue dinner \$(.) in my home last night,=

Immediately upon connecting, the streamer exhibits a demeanor of detachment. He averts his gaze (as depicted in Image 4.1) and remains somewhat distanced from the screen for approximately 13 seconds, though he remains primarily focused on his mobile phone (as shown in Images 4.2 to 4.5). During this interval, a few individuals join the stream, prompting notifications to appear on the screen. Consequently, when the streamer subsequently leans forward (Line 8, Image 4.6) and offers a collective greeting (Line 9), his embodied conduct conveys an increased engagement with the live stream. Thus, his previous actions can be interpreted retrospectively as a form of loose involvement, indicating that he was waiting for an audience to assemble, demonstrating his concern not to begin until an 'adequate' enough audience had gathered.

His collective greeting (“Good afternoon people”) enacts an orientation toward the audience as a whole while also establishing a categorical distinction between himself as the streamer and the audience as a collective of equivalent viewers. Immediately following this, he proceeds to address an individual viewer who has joined, using the first name displayed on her nametag

(Line 11). Subsequently, he shifts focus once more, addressing the audience collectively to make a casual remark about his current situation, emphasizing informality and everydayness with references to Monday morning and coffee (Line 13). The viewer reciprocates his greeting shortly thereafter (Line 14). The fact that actions are performed during the response and the initial message and that the response does not directly address them is a recurring characteristic of interactions between streamers and viewers in live streams (Dutt & Graham, 2023). This phenomenon mirrors the concept of "disrupted turn adjacency" observed in multi-party chat interactions (Herring, 1999).

The streamer, notably, does not acknowledge the return greeting, indicating an orientation towards minimal greeting sequences with viewers during this initial phase. However, he proceeds with a general inquiry directed at the entire audience ("How is everyone doing," Line 17). Subsequently, he leans towards the screen and alters the focus by individually greeting other new viewers, utilizing parts of their nametags for identification (Lines 19-26). This pattern shifts when he changes his 'footing' once more, engaging in an activity relevant to the stream's theme and the entire audience by sharing a personal story (Line 28). This utterance marks the end of the opening and the story in progress as the current 'business' of the stream.

A notable aspect of this opening sequence is the intricacy of two distinct interactional frames. Goffman's concept of laminated participation frames and 'footings' has been demonstrated to be applicable to live streaming contexts (Choe, 2020), and we aim to extend this insight by establishing connections between footings, participation frames, and categorical relevancies. In the preceding excerpt, we thus observe the interweaving of a 'public speaker'/'audience' framework, where the streamer addresses the viewers collectively as a single entity, the 'audience,' and a 'host'/'guest' participation framework, in which the streamer orchestrates individual greeting exchanges with incoming viewers, effectively welcoming them on a personal level within the public sphere. This approach adeptly addresses another common concern of streamers in this type of live stream: balancing the stream as a public performance while still demonstrating an inclination towards engaging individually with incoming viewers (those who send messages). This situation bears resemblance to the concerns of a party host at the beginning of an event, who must greet incoming guests while simultaneously kicking off the evening's activities.

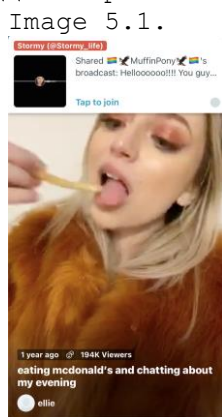
In this example, the streamer does much interactional work to enact the live stream as a 'conversational' one: the streamer talks, initiating utterances in the first position; he tells a

personal story in a sequential position, which is recognizable as the start of the proper business for this stream (telling personal stories is the kind of thing you do in a conversation); he displays an orientation towards treating the viewers' messages rather systematically. Adding the fact that he waits at the start for a significant 'audience' to gather, and appears continuously as a talking head, the opening appears to be managed so as to frame the stream to come as a 'conversational' performance.

In the following case, the title explicitly frames the upcoming live stream as a conversational experience focused on eating a burger and engaging in chat. This title allows viewers to anticipate a video format where the streamer will be prominently featured as a talking head, conversing while eating, and engaging in dialogue to the extent that her eating allows. It is worth noting that the streamer in question wields significant influence on social media platforms and consistently draws a large audience during her live streams. As a result, there is a substantial influx of messages, making it impractical to provide a full transcription of the interactions (we have endeavored to include as much content as possible without occupying excessive space). The extract specifically covers the initial minute of this 16-minute-long stream.

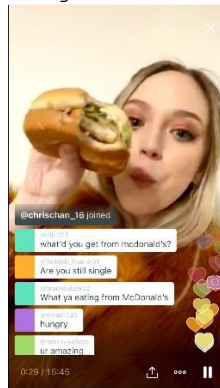
Extract 5 (Title: "eating McDonald's and chatting about my evening")

- 1. 0:00 S: ((opens the stream while eating French fries))#
- 2 ((multiple notifications of incoming viewers))



- 3 0:08 S: hmm cool what's up guys?
- 4 (notifications)
- 5 0:14 S: like (.) I should maybe go in the room
- 6 (notifications)
- 7 0:20 S: alright(.) let's pop f- my god there's a lot of people
- 8 watching this already
- 9 (notifications)
- 10 0:26 S: ((looking right)) are you okay?
- 11 0:27 **JOB: what did you get from mcdonald's?**
- 12 S. ((looks to screen))
- 13 0:27 **TED: Are you sill single**
- 14 0:28 **BRA: what ya eating from mcdonald's**

13 0:29 S: ((shows her burger))
Image 5.2



14 0:29 KYR: **hungry**
15 0:30 MIS: **ur amazing**
16 0:31 S: ((moves camera towards roommate))
17 0:31 PSY: **Heyyyy! I want McDonalds too!**
18 0:32 JRR: **Hiiiiiii**
19 0:32 S: ((moves camera back to talking heads))
20 0:32 S: I'm eating chicken
21 0:33 ALW: **ok go girl**
22 0:35 COL: **hi ellie**
23 0:35 S: and French fries
24 0:36 MIC: **Hi**

.....((lots of incoming messages which she appears to read without responding, while providing some audible receipt tokens -"mm"))

25 0:53 SAM: **Erica?**
26 S: Cu\$z??really
27 0:54 MAR: **\$Ugh I want some**
28 0:55 THE: **Say hi ellie**
29 S: which is wh\$y I am behaving this wa\$y
30 0:56 BNE: **\$In love with you**
31 0:57 MAR: **\$I'm so happy**
32 0:58 CCH: **Mcchicken classic!!!**
33 (1.0)
34 S: but what's up yo\$u guys, (.)
35 0:59 JAC: **\$Howdy!**
36 S: I feel li\$ke I haven't done this in a rea\$lly really long
37 time
38 1:00 SAN: **\$Your roommate is fire**

The streamer's initial interaction lacks a greeting or any indication of the relevance of a greeting sequence. Her first utterance ("hm cool, what's up guys," Line 3) is crafted to invite viewers (addressed collectively as "guys") to introduce topics, suggesting that she will monitor incoming messages and engage with some of them. Approximately 20 seconds into the stream, she observes the influx of viewers (Line 6-7) with an utterance designed to convey sudden realization (it interrupts an emerging turn and is preceded by the change-of-state token "my god"). This type of noticing and assessment presents another way for streamers to demonstrate their concern with the size and engagement of their audience. It is a common feature of the

openings of conversational streams, especially for streamers who do not have a specific topic prepared and wait for viewers to join in.

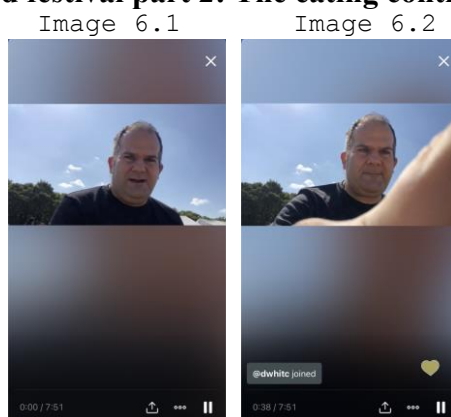
Although she acknowledges the growing audience size, she initially refrains from direct engagement. Instead, she turns her head and directs a question to her roommate beside her (Lines 5-6). Only afterward does she glance at the screen and respond to a message for the first time, showcasing her burger (Line 13, Image 5.2) in response to a query about her meal choice. Subsequently, she describes her eating activity (Lines 20, 23, 29) without responding to incoming messages. Following a brief, untranscribed segment where she seems to concentrate on messages (leaning in, eyes moving, vocalizing a few token acknowledgments, "mm"), a viewer requests her to greet them ("say hi Ellie," line 28). While what the request precisely addresses in the previous stream is uncertain, it is noteworthy that a) it constitutes the viewer's first message (indicating its perceived urgency and relevance); b) it follows several greeting messages she has not responded to (such as lines 20, 24); c) it follows a period where she appears to acknowledge messages (through vocalizations) without responding. Nonetheless, the request reflects the viewer's expectation of engagement with messages and underscores the perceived relevance of a greeting from the streamer during openings. She does not respond to this message either, opting instead to reiterate her invitation for viewers to speak up and initiate topics (Line 34). Her lack of involvement in greeting sequences is treated as troublesome, and she is made accountable for it.

The 'conversational' dimension of the two last streams is reinforced by the fact that the initial topic is disconnected from the immediate, 'live' situation. In the first extract, the first topic is a personal story that happened to the streamer earlier. In the second extract, the streamer displays herself as having nothing to tell while wishing to 'talk' and offering viewers an opportunity to initiate a first topic. In both cases, the opening enacts an interactional framework in which the 'business' of the stream is to talk about 'talkable' stuff. Another type of stream is made of those streams in which the streamer is still expected to talk and engage with an audience (in that sense, such streams retain a 'conversational' feel). However, the streamer is observably engaged in some activity or performance for which she is responsible and which she 'owns.' The purpose or topic of such live streams is to make the performance publicly accessible and provide the viewers with opportunities to discuss it.

In our next example, the stream's title denotes it as part of a series ("part 2") and frames it as documenting the eating experience at the Osaka food festival. It is worth noting that sharing,

eating, and tasting are significant genres of live streaming, which have been studied to some extent (Choe, 2019; Anjani et al., 2019; Rudiger, 2022). Let's delve into how this particular stream unfolds at the beginning.

Extract 6. Title: “Osaka food festival part 2: The eating continues, #Japan travel”



01 0:00 S: #6.1We're back at it again
 02 (2.0)
 03 in (.) the Osaka (.) food stall festival.
 04 Let me see ·if I can see
 05 (unintelligible words)
 06 Hi guys
 07 (1.5)
 08 **0:18 ARI: hello Dave**
 09 S: So (.) I'm here, hi, how are you?
 10 Hope you can see me clearly. I think I've set
 11 this a little bit clear#6.2, actually should be
 12 in very high resolution.
 13 (1.0)
 14 S: An:d I'm going to sample (.) some
 15 more (.) foods
 16 (8.0)
 17 0:42 S: .hh so I'm here and I go:t I got three dishes
 18 here today I got three dishes (1.0)
 19 **0:48 MIM: Hello, Dave**
 20 S: Hi/. Hi\$, everybody (0.8)
 21 **0:50 DWH: \$Clear**
 22 **0:51 REN: Watching from Brazil**
 23 S: Hello: Clear. OK. Hello, Brazil, how's it going?
 24 >Thanks for coming in<. I can't see bubbles.
 25 I can only see people who are chatting. So (.)
 26 if you're not chatting, I can't \$see you.
 27 **1:02 RJU: \$Hi Dave**
 28 \$Hey, what's up,
 29 **1:03 MSP: \$Hi Dave !!**
 30 S: Hi MSP how are you?
 31 So another day:=
 32 **1:08 HID: Hello David**□□□□□□□□
 33 S: =I got-
 34 **1:09 ARI: \$very clear**
 35 S: \$er I got all my food ready,
 36 S: I've got er three dishes here, the ones we didn't
 37 eat yesterday I think I did show them to you.
 38 Hi HID.

39 Ah: and these are good are unique ones. These
40 ones I saw yesterday that I thought were quite
41 good, so I thought I'd try them. So I came here
42 a little bit early to get the dishes first (.)

Just after the video connection, the streamer's first utterance is an announcement ("We are at it again the Osaka food festival", Lines 1-2). It is anaphorically tied to a previous stream ("again"), thus strengthening the sense of it as one in a series. It also enacts a collective "we," that is, the streamer and the viewers sharing some experience. It frames the intelligibility of the scene by stating and formulating the location in a way that also has some topical relevance. Then he does visual repair work (Lines 4-5) and greets ("Hi guys," line 6).

This greeting enacts a new streamer-audience participation framework (as opposed to the previous "we"), contrasting the streamer with a collective audience ("guys"). Then, as he keeps on talking about what he will be doing, several streamers greet him personally, using his first name in a way that shows some degree of familiarity (Lines 8, 19, 27, 29, 32). When the greeters join in, they can understand the stream is beginning from the streamer's preliminaries, which makes it particularly relevant for them to greet: there is a relative clustering of viewers' greetings at the beginnings of live streams, even if, as we saw above, viewers' greetings may be done throughout the stream. The streamer displays his engagement with individual viewers by responding to all these greetings, often adding how-do-you-dos to his return greeting (Lines 8-9, 19-20, 27-28, 29-30), and sometimes breaking the progressivity of a turn in progress to return a greeting (Lines 32,38). These greeting sequences create interpersonal moments within the framework of streamer-audience participation.

The interplay between both is made salient in line 20, where a single "hi" is followed by a collective one (Hi. Hi everybody) before the streamer goes on with his prefacing work. The first greeting appears as a return interpersonal greeting, while the second re-enacts the streamer-performer/audience participation framework, which is relevant to his ongoing prefacing work. Some viewers' initial turns directly address this participation framework, such as "watching from Brazil" (Line 22). While this message is treated as a greeting (see the streamer's response, "Hello Brazil how is it going?", line 23), it is initially designed so as to state a mode of participation ("watching") and a location, thus enacting the newcomer as one more spectator in an audience. There is a tension between collective welcoming and public performance management on the one hand and interpersonal greeting sequences with newcomers on the other. In a face-to-face setting, let us say a party, both would lead to temporally well-distinct

participation episodes, as when the host disengages from the party and comes to the door to welcome new guests before getting back. Here, since messages arrive instantaneously and in a way that cannot be controlled entirely by the participants, the temporal interplay between the two types of participation frameworks is very fine-grained, as evidenced by a turn such as “Hi. Hi, everybody.”

In summary, the beginnings of these 'conversational' or 'expository' streams, where the streamer presents as a 'talking head' and engages in 'just' talking with or performing for the audience, constitute complex collaborative events. There's a layering of participation frameworks that render various types of activities, sequences, and concerns relevant during the opening phase. Thus far, we have identified:

- a) The waiting for an adequate audience. This is marked by a lapse in the streaming activity. Usually, the streamer does not speak about the stream yet. He can either display signs of disinvolvement or attend the screen while only greeting newcomers as they join in. This usually comes first, right after connection. Such an ostensive waiting episode may also be absent without apparently being missed, though when present, it clearly displays a concern with the gathering of a significant audience.
- b) The streamer's collective greeting of the audience. It strongly enacts the categorical contrast between the streamer/performer and all viewers (treated as equivalent). Such sequences are characterized by collective terms of address and display a concern with the constitution of the viewers as an 'audience.'
- c) Individual greeting sequences between the streamer and (some or most) incoming viewers. They display a concern with engaging individually with viewers and enacting an interpersonal and individual host/guest relationship.
- d) A topical framing of the stream, addressed to the viewers as an audience. It displays the streamer's potential concern with the stream's immediate intelligibility and relevance.

One analytical challenge arises from the considerable variability observed regarding the presence or absence of these activities and sequences, as well as their timing and order. Unlike phone conversation openings, which can often be described as orienting to a "canonical" opening sequence (Schegloff, 1986), live stream openings exhibit complexity, characterized by laminated participation frames.

4.2 Mediating access to an outside event

In the streams we are presently examining, the streamers are not personally responsible for the events featured in the stream, and they often are not held accountable for the events unfolding, unlike the previous section. Instead, they operate as a mediator, granting access to the events they attend and possibly participate in. However, the stream is not primarily focused on the streamer’s personal experiences. In this regard, these streams are less 'conversational' or 'expository' compared to the previous section.

Consider the following example—the title of the stream orients towards the observation of birds in Australia.

Extract 8. Title “Wild cockatoo visits #Australia”

01 0:00 S: ((opens the stream))
 02 # (5.0) ((zooms in on the bird))#



Image 8.1



Image 8.2

03 0:05 **POL** joined
 04 0:05 **DET** joined
 05 0:05 S: I think (.)
 06 0:06 **LAD** joined
 07 0:06 **KEL** joined
 08 0:07 S: it could be Snow Flake.
 09 (2.0) ((pans slowly left))
 10 0:09 S: ((whistles)) Hello/
 11 (2.5) ((pans left and other bird appears))#



Image 8.3

12 0:14 **KOA** joined
 13 0:14 S: I wonder. Oh/ (.) there's another bird too

14 (1.5)
 15 0:17 DET: **Snow flake**
 16 S: two birds. (10.0)
 17 0:23 TAN: **Hey Peter and Bert**
 18 0:25 LAD: **SNOW PUFF**
 19 0:27 DET: **Incredible**
 20 S: Hello, good morning, how are you? Ah: it's not- it's Shy-
 21 it's Shy Apple. Envy you, (8.5)
 22 0:37 TAN: **Bert**
 23 S: Shy Apple is very uh, (.) very tentative. I have to lay
 24 \$down the seeds for her to come. (2.0)

In this stream, the video connection initially displays a view of a window (Image 8.1), and within the first few seconds, the streamer zooms in on the window, visually highlighting a bird on the balcony (Image 8.2). His first verbal statement is a claim of recognition, where he identifies the bird as "Snow Flake," although this claim is highly qualified (prefaced by "I think" in line 5 and "it must be" in line 8). This claim is contextually bound to the environment (Goodwin, 2018), as it can only be understood in relation to his prior camera work and the resulting image. The title of the stream, along with this particular multimodal action as the initial action in the stream, frames the focus of the stream as being about the observable birds. Furthermore, as he pans the camera to the left and another bird comes into view, he notices it ("Oh, there is another bird too," line 13), with the initial-positioned "change-of-state" marker "oh" indicating both the visually prompted nature of the subsequent statement and its informative content (Heritage, 1984). Viewers orient to the framing of the stream as a showing of the birds the streamer discovers and makes available by producing first messages that are responsive to the streamer's action and relevant to such a framing: repeating the name of the bird (Line 15), providing another identification (Line, 18), and an assessment of the event (Line, 19).

The streamer keeps engaged with the audience by greeting back viewers individually (Lines 17 and 20), though there is no collective greeting or welcoming sequence. He still acts as a host greeting individual guests, but not anymore as a performer working up her audience. He rather acts a) a "mundane video director" (Licoppe & Morel, 2014), which gives visual access to features of her environment, thus made noticeable to the audience, and b) as an onlooker himself, watching, noticing, and commenting on what he makes available to online viewers. He, therefore, behaves as another viewer, albeit a privileged one with more direct access to the bird show.

More generally, as live streams focus on documenting events that appear independent of the streamer's agency, streamers tend to fade into the background, sometimes even from the beginning of the stream. In the following extract, similar to the previous one, the live stream opens with a shot of the environment (Image 9.1). However, the streamer remains unseen, and she speaks very little. This occasions the following opening sequence:

Extract 9 (Title: “Dolphins maybe”)

Image 9.1



1 0:00 S: ((Opens without verbal utterances on a continuous
shot of a dolphin jumping and following the boat#9.1))

2 1:20 **EJW: Hi BAR³**

3 1:30 **HOC: Hey BAR, are you swimming with the dolphins?**

4 1:37 S: **Hi👋👋👋👋⁴**

5 2:00 **EJW: @BAR what part of Florida**

6 2:01 **MAR: He's loving it**

7 2:07 S: Gulf Coast

8 2:32 **LOU: Hiya BAR**

9 2:41 S: Hi LOU

10 2:46 **EJW: My mom lived on the gulf coast near Tampa**

11 3:00 S: where did he go?

12 3:04 **CAM: They can swim very fast indeed!**

13 3:05 **LOU: @EJW Hello EJW**

14 3:18 **EJW: Have you seen manatees**

15 3:30 **EJW: @LOU hi LOU**

16 3:36 **CAM: They are quick!**

17 4:03 **NIC: Wow**

18 4:07 S: he is swimming as fast as the boat

19 4:09 **NIC: Definitely Dolphins**

20 4:16 **HAL: I love dolphins**

21 4:19 **HOC: They want tips**

22 4:27 **LIS: this is awesome!!**

³ In the extract, 'BAR' is the streamer's pseudonymized name.

⁴ The streamer herself sends a message with emojis at Line 3, and that's why we also make the texts bold for this line.

At the outset, the streamer begins the stream without any verbal utterances, maintaining a continuous shot of a dolphin jumping and following the boat. This video footage, coupled with the stream's title, clearly indicates the stream's purpose: to showcase the appearance of dolphins to potential viewers. Many viewers seem to align with this objective, as evidenced by their initial messages, which are often prompted by the event captured in the video. These messages include appreciative assessments (Lines 17, 22, 23) or observations related to the observable behavior of the dolphins (Lines 6, 16, 21). The streamer herself contributes only one utterance, a question based on observation: "Where did he go?" This question is directed either to herself or to other co-present participants and is triggered by a potential concern (if the dolphins were to disappear, it would defeat the purpose of the live stream). She responds to only one observation ("They are quick!", Line 6) with a second observation-based assessment of her own: "He is swimming as fast as the boat" (Line 18).

Regarding greetings, the streamer does not offer a collective greeting or welcome to her entire audience. Instead, a few participants initiate greetings in the first position in a manner that suggests familiarity (Lines 2, 3, 8), which is further confirmed by two participants exchanging greetings with each other in a separate interaction (Lines 13 and 15). The streamer responds to two of these greetings, once verbally (Line 9) and once, notably, by sending a message of her own (Line 4).

This case bears similarities to the previous one, with the streamer assuming dual roles: granting viewers access to the live event she is attending and functioning as a viewer herself. However, in her role as a viewer, she remains largely in the background: she refrains from commenting on what she sees by initiating observations in the first position, and only does so when she notices something potentially troublesome. She appears to assert minimal privilege in her position as a viewer, behaving as though she were little more than another member of the audience.

The disappearance of the streamer from the interactional scene at the beginning of the stream reaches its pinnacle in the following example, where the streamer, aside from his video work, appears strangely absent from the scene as the live stream begins. In this instance, the streamer, who is a Trump supporter, records a Donald Trump rally (the arrival of Air Force One) just before the 2020 election. Such streams belong to a growing genre where streamers broadcast political events or demonstrations as they unfold (Kavada & Trere, 2020; Fang, 2023). As a well-known political critic with a significant following on Twitter, the streamer's broadcasts

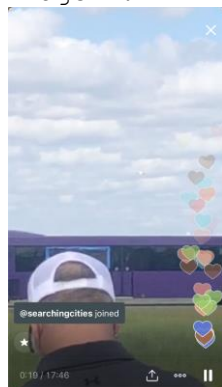
attract many potential viewers. Despite not engaging in many conversations or responding to messages, the streamer still manages to garner substantial viewer interactions. Within this genre, many messages, particularly the first ones, are designed as observations, treating the stream as a spectacle (Song & Licoppe, 2024).

Extract 10 (title: INCOMING)

01 0:00 S: ((opens the stream))#
Image 2.1



02 ((zooms in, pans left and right, while many notification of incoming viewers appear))#
Image 2.2



03 0:25 JEF: Good morning Truth detective.
04 0:26 MVM: Hey Will!!!
05 0:28 KEE: □□□□□□
06 0:29 MDM: □□□□□□
07 0:30 CKR: Greenville?
08 0:30 LIA: Wow
09 0:31 ELO: nervous energy
10 0:33 YOF: What is it, JAC?
11 0:34 BOB: □
12 0:35 DKV: □
13 0:36 BAN: be careful
14 0:37 LOE: Where are we?
15 0:38 SCU: this is the most entertaining 24 hours of all time lol
16 0:39 STE: The Deep state doing its thing in GA and VA
17 0:40 LIA: Thanks JAC!
18 0:41 MOD: POSO MOD SQUAD! MODDING WHILE WORKING OUT! □□□□□□□□□□
19 0:42 ROS: What are we looking at?

The opening sequence, if we may still call it that, is reduced to its bare minimum. The stream starts with two viewers greeting the streamer, probably orienting to the fact that it is just starting. After that, incoming viewers do not greet. Their first message is usually some comment on the stream, showing that they orient almost immediately towards what happens on the screen rather than to their joining in as marking the beginning of an encounter (which would warrant a greeting). An incoming viewer who would connect 30 seconds into the stream might not even get the sense that it has just started if the two initial greetings had already disappeared. Therefore, rather than having an opening, we could describe this live stream as simply beginning immediately after the video starts.

One reason for that is the relative absence of the streamer, which has receded far away from the perceptible scene. Invisible and silent, his presence can only be inferred, first because of the fact of the stream itself, and second, because of the camera work (such as the zooming he does in line 2). Viewers seem to orient to that self-imposed backgrounding of the streamer. The message in line 13, "*be careful,*" enacts the streamer as someone who can hear but is interactionally engaged elsewhere in this other interactional scene, which he makes publicly available. Almost all other first messages from viewers orient in some way to what is being shown either by providing some explicit form of appreciation (Lines 8, 11, 12) or by noticing something discernible in the audio-visual display and using it as a resource to do various actions, what we could call 'noticing-based' actions (Song and Licoppe, 2024): "*nervous energy*" (Line 9), "*the deep state doing its thing in GA and VA*" (Line 16).

The immediate orientation of viewers towards the audio-visual display is highlighted by messages that indicate sense-making difficulties: "Greenville?" (Line 7), "What is it JAC?" (Line 10), "Where are we?" (Line 14), "What are we looking at?" (Line 19). By expressing their inability to comprehend what the stream is about upon joining, viewers demonstrate a significant concern with the immediate "scenic intelligibility" (Jayyusi, 1988) of the stream and the expectation that they should be able to make sense of what is happening immediately, and particularly at the beginning. While this concern exists in all streams, it is particularly pronounced in this type of documentary stream because there are few clues provided by the streamer: the title is ambiguous, and the streamer does not speak.

This provides for a particular type of interactional framework and encounters, very different from the conversational stream in extract 1, where viewers could feel they were 'meeting' with the streamer. Here, the quasi-absent streamer makes publicly available a supposedly interesting

scene he participates in and is thus able to document, for an audience of viewers to comment on publicly. So, except for the few that directly address the streamer, most messages from viewers seem to be produced for other viewers to read, pick up on, and possibly respond to. Within such an interactional configuration, an opening sequence with greeting exchanges, a position for the first topic, etc., appears irrelevant.

Conclusion

Having compiled a corpus of live streams on Twitter, we approached their openings from an EM/CA (Ethnomethodology and Conversation Analysis) perspective. One immediate empirical observation and challenge we encountered was the variability of these openings. Unlike phone conversations, for instance, there is no canonical opening sequence. However, a significant portion of the streams seemed to require a recognizable juncture to start the proper business of the stream, thereby rendering the notion of an opening relevant. Furthermore, participants exhibited a set of recurring concerns in these live streams, manifested through distinctive sequences or interactional practices that could be observed and analyzed. The common characteristic of these live streams was their conversational or expository nature, with the streamer prominently featured, whether they intended to engage in dialogue or showcase a performance for which they were accountable and responsible—essentially, content that they 'owned.'

We were able to identify four primary concerns enacted through distinct interactional practices and categorical relevancies in these live stream openings. First, streamers exhibited an ostensible concern for gathering an adequate audience, often manifested in the form of a perceptible 'waiting' episode immediately after connecting to the stream. Second, streamers demonstrated a concern with managing the audience as a collective entity, evidenced by the use of collective terms of address (such as an "I-you" participative framework and collective greetings). Third, both streamers and viewers engaged in a guest/host relationship dynamic, where there was an expectation for both parties to personally greet each other when new viewers joined. This framing positioned the live stream as a kind of hosted encounter and gave rise to interpersonal greeting sequences, the occurrence of which was contingent upon the timing of new arrivals. Fourth and finally, both streamers and viewers showed a concern for the intelligibility of the live stream, often demonstrated through an utterance in which the streamer contextualized the content of the stream within the "I-you" participative framework. While the

production of such an utterance typically signaled the beginning of the stream, interpersonal greeting sequences could continue long after, as new viewers joined continuously.

The variability of such livestream openings can be attributed to three key factors. Firstly, some of these concerns and the interactional episodes that enact them may be absent without giving the impression of being missing, as is the case with the waiting sequence. Secondly, those elements that are normatively expected, such as greeting sequences and the immediate intelligibility of the live stream (along with the framing utterances that establish it), may occur in different orders. Lastly, the variability of the openings is further compounded by the randomness in the way viewers join the stream. This makes the production of interpersonal greeting sequences relevant, potentially anywhere during the opening and even beyond the opening, as demonstrated in section 3.

Furthermore, we observed a distinct category of live streams that exhibited markedly different openings or even lacked a recognizable opening episode beyond the video connection. What distinguished these live streams from the previous ones is that the streamers did not seem to promote themselves or their personal accomplishments but rather aimed to document events in which they may have been involved without emphasizing their personal participation. In these streams, interpersonal greeting sequences and announcements framing the content of the stream may be absent. In these types of openings, the streamer may adopt different participation frames and categorical relevancies. For example, they might act as another viewer (albeit with privileged access), contributing observations about what they see unfolding (an action typically associated with the category of 'viewer'). In extreme instances of these 'experiential' or documentary streams, the streamer effectively removes themselves from the interactional scene; their presence is only felt as a mediating camera operator, providing audiovisual access to events for viewers to discuss among themselves. In such cases, there may not be a distinct opening episode, but viewers show strong normative expectations regarding the immediate intelligibility of the stream's content.

These findings may require further nuance through additional studies, particularly across different platforms, as platform affordances could influence the form, presence, and distribution of the interactional episodes observed here. However, the approach and observations made in this study may be sufficiently generalizable. Linking participation frames and categorical relevancies with interactional sequences and participants' concerns is a powerful enough

approach to be applicable across platforms. It has enabled us to comprehend the complexity of live stream openings and to establish genre distinctions in a manner that is grounded in empirical data rather than third-party interpretations. Consequently, the systematic analysis of live stream openings adds depth to earlier, more informal distinctions, such as between 'expository' or 'conversational' streams and 'experiential' or 'display' streams.

APPENDICES. TRANSCRIPT CONVENTIONS

This paper's transcription conventions mainly adopt Gail Jefferson's (2004) conventions for talk and Lorenza Mondada's (2018) conventions for streamers' talk and embodied actions. The main convention symbols involved in this article are as follows:

\ and / represent the falling and rising intonation, respectively;

↑ indicates a sudden rise in intonation;

(.) means inconspicuous short pauses (usually less than 0.2 seconds);

.hh means inhale or laughter;

§ indicates the moment or the period when the audience's messages come up, or when an action happens;

*--->

---->* indicates the continuation of the action;

indicates where the screenshot appears;

£ £ means that the content within the symbol is said when laughing;

() words in parentheses are unintelligible;

[indicates the overlapping of speech;

= indicates the connection of two turns;

(()) indicates non-verbal conduct.

Reference

- Anjani, L., Mok, T., Tang, A., Oehlberg, L., & Goh, W. B. (2020). Why do people watch others eat food? An empirical study on the motivations and practices of mukbang viewers. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/3313831.3376567>
- Chen, Z., Cao, H., Deng, Y., & Gao, X. (2021). Learning from home. A mixed-methods analysis of live streaming-base remote education experience in Chinese colleges during the

- COVID-19 pandemic. In Proceedings of CHI '21: CHI Conference on Human Factors in Computing Systems. DOI:[10.1145/3411764.3445428](https://doi.org/10.1145/3411764.3445428)
- Choe, H. (2019). Eating together multimodally: Collaborative eating in mukbang, a Korean livestream of eating. *Language in Society*, 48(2), 171–208. <https://doi.org/10.1017/s0047404518001355>
- Choe, K. (2020). Talking the cat. Footing lamination in a Korean livestream of cats mukbang. *Journal of Pragmatics* 160, 60-79.
- Due, B. L., & Licoppe, C. (2021). Video-Mediated Interaction (VMI). *Social Interaction. Video-Based Studies of Human Sociality*, 3(3). <https://doi.org/10.7146/si.v3i3.123836>
- Dutt, S. and S. Graham (2023). "Video, talk and text: How do parties communicate coherently across modalities in live videostreams?" *Discourse, Context & Media* 55, . <https://doi.org/10.1016/j.dcm.2023.100726>.
- Dutt, S. and S. Graham (2023). "Video, talk and text: How do parties communicate coherently across modalities in live videostreams?" *Discourse, Context & Media* 55, . <https://doi.org/10.1016/j.dcm.2023.100726>.
- Fang, K. (2023) The social movement was live streamed: a relational analysis of mobile live streaming during the 2019 Hong Kong protests, *Journal of Computer-Mediated Communication*, Volume 28, Issue 1, January 2023, zmac033, <https://doi.org/10.1093/jcmc/zmac033>
- Gan, Y., Greiffenhagen, C., & Licoppe, C. (2020). Orchestrated openings in video calls: Getting young left-behind children to greet their migrant parents. *Journal of Pragmatics*, 170, 364–380. <https://doi.org/10.1016/j.pragma.2020.09.022>
- Goffman, E. (1966). *Behavior in public places: Notes on the Social Organization of gatherings*. Free Press.
- Goffman, E. (1974). *Frame Analysis. An Essay on the Organization of Experience*. New York, Harper and Row.
- Goodwin, C. (2018). *Co-operative action*. Cambridge University Press.
- Hazel, S., Mortensen, K., & Rasmussen, G. (2014). Introduction: A body of resources – CA studies of Social Conduct. *Journal of Pragmatics*, 65, 1–9. <https://doi.org/10.1016/j.pragma.2013.10.007>
- Herring, S. (1999). Interactional coherence in CMC. *Journal of Computer-Mediated Communication*, 4 (4).

- Ilomäki, S., & Ruusuvoori, J. (2020). From appearances to disengagements: Openings and closings in video-mediated tele-homecare encounters. *Social Interaction. Video-Based Studies of Human Sociality*, 3(3). <https://doi.org/10.7146/si.v3i3.122711>
- Jayyusi, L. (1988). Toward a socio-logic of the film text. *Semiotica*, 68(3-4), 271-296. <https://doi.org/10.1515/semi.1988.68.3-4.271>
- Jodén, H., & Strandell, J. (2021). Building viewer engagement through interaction rituals on Twitch.tv. *Information, Communication & Society*, 25(13), 1969–1986. <https://doi.org/10.1080/1369118x.2021.1913211>
- John Heritage, (1984), "A change-of-state token and aspects of its sequential placement", In *Structures of Social Action: Studies in Conversation Analysis* (J. Maxwell Atkinson, John Heritage, eds.), Cambridge, U.K., Cambridge University Press, pp. 299–345.
- Kavada A., Treré E. (2020). Live democracy and its tensions: Making sense of livestreaming in the 15M and occupy. *Information, Communication & Society*, 23(12), 1787–1804. <https://doi.org/10.1080/1369118X.2019.1637448>
- Licoppe, C. (2017). Skype appearances, multiple greetings and ‘coucou’ The sequential organization of video-mediated conversation openings. *Pragmatics*, 27(3), 351-386.
- Licoppe, C., & Dumoulin, L. (2010). The “Curious Case” of an Unspoken Opening Speech Act: A Video-Ethnography of the Use of Video Communication in Courtroom Activities. *Research on Language & Social Interaction*, 43(3), 211–231. <https://doi.org/10.1080/08351811003741319>
- Licoppe, C., & Morel, J. (2012). Video-in-Interaction: “Talking Heads” and the Multimodal Organization of Mobile and Skype Video Calls. *Research on Language & Social Interaction*, 45(4), 399–429. <https://doi.org/10.1080/08351813.2012.724996>
- Licoppe, C., & Morel, J. (2014). Mundane video directors in interaction: Showing one's environment in Skype and mobile video calls. In *Studies of video practices* (pp. 135-160). Routledge.
- Licoppe, C., & Morel, J. (2018). Visuality, text and talk, and the systematic organization of interaction in Periscope live video streams. *Discourse Studies*, 20(5), 637–665. <https://doi.org/10.1177/1461445618760606>
- Marulanda, J.F.F.; Collazos, C.A.; Hurtado, J.A. Evaluating an Immersive Virtual Classroom as an Augmented Reality Platform in Synchronous Remote Learning. *Information* 2023, **14**, 543.
- Michal Marmorstein, (2023), "Why say ‘hi’?: Framed openings in Hebrew WhatsApp messaging", In *Conversation Analytic Perspectives to Digital Interaction: Practices*,

- Resources, and Affordances (Aino Koivisto, Heidi Vepsäläinen, Mikko T. Virtanen, eds.), Helsinki, Finnish Literature Society, pp. 43–69.
- Mondada, L. (2015). Ouverture et préouverture des réunions visiophoniques. *Réseaux*, (6), 39-84.
- Oittinen, T., & Piirainen-Marsh, A. (2015). Openings in technology-mediated business meetings. *Journal of Pragmatics*, 85, 47–66.
<https://doi.org/10.1016/j.pragma.2015.06.001>
- Paulus, T., Warren, A., & Lester, J. N. (2016). Applying conversation analysis methods to online talk: A literature review. *Discourse, Context and Media* 12, 1-10.
<http://dx.doi.org/10.1016/j.dcm.2016.04.001>
- Recktenwald, D. (2017). Toward a transcription and analysis of live streaming on Twitch. *Journal of Pragmatics*, 115, 68–81. <https://doi.org/10.1016/j.pragma.2017.01.013>
- Richard Fitzgerald, William Housley, eds., (2015), "Advances in Membership Categorisation Analysis", London, Sage Publications, pp. 216.
- Rintel, E. S., Mulholland, J., & Pittam, J. (2006). First Things First: Internet Relay Chat Openings. *Journal of Computer-Mediated Communication*, 6(3), 0–0.
<https://doi.org/10.1111/j.1083-6101.2001.tb00125.x>
- Rüdiger, S. (2022). Intimate consumptions. Eating shows and the performance of informality. *Internet Pragmatics*, 5(1), 111-142.
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language*, 50(4), 696.
<https://doi.org/10.2307/412243>
- Schegloff, E. A. (1986). The routine as achievement. *Human Studies*, 9(2–3), 111–151.
<https://doi.org/10.1007/bf00148124>
- Schegloff, E. A. (2007). *Sequence organization in interaction: A primer in conversation analysis I* (Vol. 1). Cambridge University Press.
- Si, R. (2021). *China livestreaming e-commerce industry insights*. Springer Singapore.
- Song, L., & Licoppe, C. (2023). Closing live video streams: A sequential analysis. *Discourse, Context & Media*, 53, 100698. <https://doi.org/10.1016/j.dcm.2023.100698>
- Stephen Hester, Peter Eglin, (1997), "Membership categorization analysis: an introduction", In *Culture in Action: Studies in Membership Categorization Analysis* (Stephen Hester, Peter Eglin, eds.), Washington, D.C., University Press of America, pp. 1–24.

- Stommel, W. (2016). Information giving or problem discussion? formulations in the initial phase of web-based chat counseling sessions. *Journal of Pragmatics*, 105, 87–100. <https://doi.org/10.1016/j.pragma.2016.09.001>
- Tang, J. C., G. Venolia, and Quinn, K. (2016). "Meerkat and Periscope: I Stream, You Stream, Apps Stream for Live Streams." Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, 4770-4780.

CHAPTER 4

NOTICING-BASED ACTIONS AND THE PRAGMATICS OF ATTENTION IN EXPOSITORY LIVE STREAMS. NOTICING ‘EFFERVESCENCE’ AND NOTICING-BASED SEQUENCES

Abstract

This paper focuses on live streams as interactional phenomena from an ethnomethodology and conversation analysis (EMCA) perspective. It discusses how streamers and viewers manage attention and engagement and how noticing-based actions constitute a powerful and crucial resource to produce and display these, finely tuned to the asymmetric communicative affordances of live streams, which, in the case of expository live streams, enact a streamer on display for the perceptual consumption of remote viewers. It discusses how streamers and viewers may produce noticing sequences and noticing-based sequences. It also discusses how the orientation towards noticing in expository live streams may become contagious and lead to a ‘noticing effervescence.’

Keywords

Live Streaming; Noticing; Attention; Multimodal Conversation Analysis; Contagion; Effervescence

1. Introduction

Live streaming is a rapidly growing practice in which streamers may record and broadcast themselves live on dedicated platforms, for an audience of online viewers who may log in and log out relatively freely. This is an extremely asymmetric type of encounter, for while viewers can watch and hear the streamers live, the streamer can only see written notifications (for instance, when a viewer logs in to the stream) and written chat messages sent by the viewers. A precursor to live streaming was instantiated by “webcam girls” in the early 2000s (White, 2003). But live streaming per se took off in the 2010s, with the now-defunct platforms Meerkat and Periscope (Tang et al., 2016) and the initially specialized and influential Twitch platform, which allowed video gamers/viewers to watch video games live and interact with players/streamers (Johnson & Woodcock, 2018). Today, live streaming has spread, and streams are broadcast on numerous platforms (YouTube, Instagram, Facebook, Twitter, TikTok, etc.).

It has been especially successful in China, where commercially oriented streams may enjoy hundreds of thousands of viewers (Li et al., 2018). The uses of live streaming have also considerably diversified, and the engagement in live streaming, i.e., the articulation of streamers' conduct and viewers' written chat messages, have begun to receive attention in very different live streaming settings, such as live video game playing (see, for instance, Recktenwald, 2018; Jodén & Strandell, 2021; Johnson, 2022), having online classes during the COVID-19 pandemic (Chen et al., 2021; Marulanda et al., 2022), promoting goods (Si, 2021), introducing local unique cultural activities (Lu et al., 2019), and watching others eat (for instance, Choe, 2019; Anjani et al., 2020; Rüdiger, 2022).

Live streaming encounters are embedded in the kind of “economy of attention” that characterizes social media, in which success is assessed in terms of audience numbers, and such numbers are “monetized” (Citton, 2017). It is thus common in our data to see streamers complain that nobody is there. However, unlike many other social media, live streams involve an interactive ‘live’ performance, and success is not only assessed as a number of ‘views’ or viewers but also in terms of viewers’ *engagement*, that is, how viewers respond continuously to what is going on. Such audience response may take the form of ‘gifts’ (such as hearts, stars, flowers, etc., varying according to the platform) and, more importantly for our research, of written messages. From the viewer’s perspective, such a positive form of engagement has been shown to be stimulated by the immersive and immediate character of live streams qua live video broadcast (in part), and by the interactional features that can emerge from its communicative affordances (Wang & Li, 2020), and particularly the relevance, tone, and topicality of the messages, and the fact that viewers’ messages are attended by streamers (Haimson & Tang, 2017).

From the streamer’s perspective, a significant motivation appears to be “personal branding”: Streamers use live streaming “as a way of building an authentic personal brand that gave viewers a comprehensive view of their personality” (Tang et al., 2016). This makes relevant a particular “moral economy” for live streaming, which becomes visible in our data when some viewers complain about a stream having no or too few *active* viewers or, conversely, about a stream having many viewers and messages and not deserving it because the streamer does not appear to say or show anything or to behave in a way that might be worthy of such attention. Streams and streamers are, therefore, oriented toward eliciting various responses from viewers that are somehow relevant to what happens in the stream.

Regarding interactional practices in live streaming, which have been less studied from an interactional or conversation-analytic perspective (but see Recktenwald, 2018; Choe, 2019; 2020), we would like to highlight two constitutive features of live streaming related to the two asymmetric modalities of interaction that characterize it.

The first is video-mediated interaction, or VMI (see, for instance, Due & Licoppe, 2021), on the streamer's side. Since the streamer appears live in video, participants in live streams orient to the normativity of the "talking heads" format, which was observed to characterize VMI (Licoppe & Morel, 2012). The normative expectation there is that the video participant should face the screen with her head fully visible, as close as the setting allows, with her head fully visible, and her gaze appearing to attend the screen. When transposed to live streaming, this means that when the streamers appear to 'talk' or engage mainly conversationally, they should be visible on screen. If what goes on relates to some perceptually available activity or environment feature, the streamer would be expected to show it. The relevance of such a "talking heads" orientation brings some light to another important distinction regarding the genres of live streams. Tang et al. (2016) thus distinguish "expository" streams and "experiential streams":

- In expository streams, the streamers "are directly explaining their opinions or showing their skills through chatting, demonstrating, or hosting "how-to", ask me anything, or talk show sessions. These activities give the streamer a chance to directly expound upon their personal perspectives that constitute their brand".
- In "experiential streams", streamers want to share something they think is interesting, like an event, scenery, social gathering, or something funny.

From an interactional perspective, it is relevant that in expository streams (or during expository moments in a given stream), the streamer appears to engage conversationally, and therefore, such streams are usually done (and expected to be done) in a talking head configuration. Conversely, experiential streams are usually done (and expected to be done) so that the streamer visually shows the perceptible features of the relevant experience.

The second modality is chat-like text messaging (for a review of computer-mediated communication within a CA perspective, see Paulus et al., 2016), which is the prescribed modality for viewers' contributions. As is the case in chat interaction, the production of written responses by the viewers is not visible to the streamer, and both the viewer and the streamer

have little control over the timing of their message's appearance on the streamer's screen. In live streaming, the moment at which messages become visible and legible cannot be adjusted to the sequential organization of the streamer's ongoing, 'live' talk and embodied conduct. Messages that are responsive to some prior action usually appear with a delay, while other things have been done in between, and without any sensitivity to the activities in progress. This is the live streaming analog to the phenomenon of "disrupted turn adjacency" (Herring, 1999) observed in multi-party chat interactions (Herring, 1999). However, because streamers are permanently on display, the reasons for their lack of response are less ambiguous than those of the audience members, or with respect to what can be observed in quasi-synchronous multi-party chats (Rintel et al., 2003). As their behavior is visible, technical problems or various forms of distraction can readily be inferred. Streamers are treated as accountable for non-responding (Licoppe & Morel, 2018), and audience members may feel it legitimate to 'pursue a response' (Pomerantz, 1984) in various ways.

Within the assemblage of communicative affordances that characterize live streaming, the achievement of viewers' engagement appears as a crucial concern. We want to focus on the main interactional practice through which such an engagement is achieved and displayed, i.e., the production by viewers of messages that are designed so that they appear responsive to some prior perceptible feature in the stream, which they either formulate explicitly or on which they build implicitly to do other things. While the production of such 'noticing-based actions' is pervasive in all kinds of streams, the actual details of these actions are sensitive to the type of stream, in particular, whether it is an expository or experiential stream. Lacking the space for an exhaustive treatment, we have chosen to focus here on expository streams, where noticing-based actions often have a ritual dimension, for they affect the streamer's face. We will show how the expository live stream setting makes relevant noticing-based actions and, in turn, how the production of such noticing-based actions shapes the live streams as attentional displays where streamers and their environments appear to be fair game for noticing all around.

2. Noticings in Conversation Analysis: A Quick Overview

According to the conventional practices of producing a noticing (Schegloff, 2007; 2019; but see also Goodwin & Goodwin, 2012; Pillet-Shore, 2020), noticing can be taken to involve three participant orientations (Kesselheim et al., 2021).

(1) Registering and formulating some perceivable feature of the setting (this is the second

part of the "retro-sequence," which ties back to some more or less specified and accessible earlier perceptual realization)

(2) Making it relevant, here and now.

(3) Making relevant some response, and particularly a (re)orientation of attention of the recipient towards the noticeable.

Regarding (1), Schegloff has introduced an essential distinction between the two senses of 'noticing'. First, 'noticing' as a perceptual and cognitive event, which he proposes to call "cognitive noticing", and which may be private or not (when one may be seen 'doing noticing' in that sense). Second, "noticing" as an interactional move, the vehicle of which may be an utterance (as in line 2 here), or some embodied conduct, which he proposes to call "interactional noticing" (Schegloff, 2007: 87). As with the organization of repair, with noticings, "it is only by their subsequent component that some prior are rendered a first" (Schegloff, 2007, *ibid.*). Noticings as interactional moves are designed to be intelligible as occasioned by, and responsive to, some prior perceptual realization. As such, they can be understood as second pair parts in "retro-sequences" (Schegloff, 2007: 219).

Regarding (2) and (3), we may use 'registering' as a placeholder to account for the retrospective work that an interactional noticing does with respect to some prior perceptual event (Pillet-Shore). However, as initially noted by Sacks, interactional noticings are geared to the production of further talk, for which they often figure as the first pair part of adjacent pair sequences. In that respect, they can be understood to do a variety of things, "just-noticing", but also announcing, assessing, complimenting, complaining, calling to order, initiating a showing sequence, etc.

While an explicit feature of the CA literature, this tension between treating noticings as such or subsuming them in the study of other actions has been left mostly as an open question. As we will see, when studying live streaming as a phenomenon, viewer engagement is achieved through contributions that are designed as attentional displays, and therefore as triggered by some earlier perceptual event in the stream. But simultaneously, such viewer contributions seem to explore perceptually the stream to notice different things, in a way that achieves a variety of different actions, complimenting and criticizing in particular. Because of that, we found it useful to think of such audience messages as "noticing-based actions," which makes conditionally relevant, varied responses. It is interesting to note that while "noticing-based actions" to be

understood as such must index some prior event that provides for their occasioning, such a backward-looking tie may be more or less salient, or more or less embedded. We think that, beyond the particulars of our empirical study of live streaming, such a notion may be interesting in CA to study perceptually-occasioned actions and social feedback as a whole, without having to either highlight it as noticing, or reducing it to the study of, let's say, compliments or complaints.

The conversation-analytic literature on noticings, which deals mostly with co-present interaction, has noted that in such settings, there was a relationship between the production of noticings and the overall structure of the encounter. Noticings tend to occur more frequently in openings (Pillet-Shore, 2018), though not exclusively. While such an orientation also seems relevant in live streaming, in this latter setting, members are joining the stream all the time. From the perspective of these incoming viewers, the live streaming encounter begins when they join, so it may be more relevant to notice things then. So even if there is an orientation to do noticings in beginnings in live streams as well, since live streaming continuously provides for new encounters within the stream, noticings are more likely to be spread throughout it.

Managing (joint) attention by noticing has been an essential topic in EMCA studies. According to Schegloff (2007: 219), as an action-in-conversation, noticing works by mobilizing attention on the features in the immediate, physical, and sequential environment, such as prior talk or anything interesting or unusual, and projects some further action. Subsequent research on noticings from an EMCA perspective was mainly conducted in face-to-face settings like homes (Pillet-Shore, 2020), classrooms (Kidwell & Zimmerman, 2007; Böckler & Sebanz, 2013; Mondada, 2014; Kunitz, 2018), cars (Goodwin & Goodwin, 2012; Rauniomaa et al., 2018), surgery rooms (Campion & Langdon, 2004), etc. They generally discussed how, through resources such as talk and subtle bodily displays, e.g., gaze aversion, gestures, facial expressions, and so forth, noticers and recipients collaborate to achieve a joint orientation towards some “noticeable,” that is, something “observable-and-reportable, i.e., available to members as situated practices of looking-and-telling” (Garfinkel, 1967: 1).

In recent years, research on noticings has extended to mediated interactions. Zouinar and Velkovska (2019) analyzed how showings and noticings are used for introducing, changing topics, and sustaining talk in mediated conversation in image-based family and friends video calling. Oittinen (2020) focused on moments in which either an auditory or a visual problem

emerges in multiparty video-mediated meetings and how the participants orient to these troubles through embodied noticings. This research thread emphasizes the specific features of embodied noticings in mediated settings. Our paper aims to build on and extend the emerging EMCA literature on noticings in mediated settings and, more specifically, to show how ‘noticing-based actions’ may be a constitutive feature of live streams.

3. Data and Methodology

The platform we selected was Twitter (now “X”). Past streams are archived with their comments, making it easy for researchers to find relevant video material for analysis. Streamers are distributed worldwide, and contents are diversified. The Twitter Live interface allows viewers to type and send comments in the bottom entry bar and send virtual hearts by clicking on the heart-shaped symbol in the bottom right-hand corner. Just underneath the heart symbol, one can also see the number of current viewers (Fig. 1). To launch one’s stream, one can choose to enter a title.



Figure 1. A Typical Stream

Many functions of the smartphone’s camera may be used during broadcasting, for example, zooming in, zooming out, and flipping the camera. Regarding messages from the audience, they appear on the bottom of the screen (for both streamers and viewers), where they remain for a few seconds before disappearing. Messages are, therefore, chased by other incoming messages (like in multi-party chat configurations), so there is a limited time to read them for all parties.

Reading and responding can require alertness and skillful behavior on the part of streamers, particularly in mobile settings. Conversely, streamers are less accountable for not responding when there are many messages and viewers.

We included 92 broadcasts from 34 streamers in our corpus, which allowed us a significant variety of streams and proved adequate for the phenomena we will describe. The streams were recorded from 2020 to 2021, with a total duration of over 20 hours. Streams were either recorded live, or from archived past streams. There were two main criteria for our selection of streamers and streams: a) the primary language used by the streamers and viewers had to be English; b) we decided not to choose streams shorter than 2 minutes, and the majority of our streams were within the 10'-30' range. The first criterion was to ensure the representativeness of our data, as the most commonly used language on Twitter is English; the second was to ensure that there were enough viewers to elicit interesting episodes of interactivity (more than five active users were usually enough) for every single video (though there could also be “quiet” episodes with no or limited interaction during the stream).

The streams themselves were public and available online without a subscription. According to French regulations, recording and using data on public platforms is allowed as long as the privacy of everyone recorded is preserved. We systematically anonymized and pseudonymized the data that is presented in the paper. We changed every account name in the data transcriptions and blurred bystanders, car plates, etc., for each screenshot we used so that nobody could be recognized in our data exploitation. The recordings provide access similar enough to that of the streamer or their audience (though it may happen that some messages do not appear or that the lag between the video and the messages is different).

We use multimodal conversation analysis (Hazel et al., 2014; Schegloff, 2007) as our research method (see the Appendix for a detailed introduction to our transcription conventions). Moreover, in all our transcripts, “S” stands for the streamer; bold text represents the viewers’ messages with their account names prefaced, which have all been pseudonymized. The position of a message in a transcript denotes the moment when the message first appears on the screen. We added the exact timestamps for some of the extracts that show the beginning of the streams, i.e., Extract 2, and Extract 3, to give the transcripts a better demonstration of sequentiality, which is also facilitative for our analysis.

4. Noticing-based 'effervescence' in an expository live stream

Noticing and noticing-based actions may occur anywhere in social encounters. This is one of the main ways in which human sociality displays "local sensitivity" to the lived environment (Bergmann, 1990). However, live streams may provide a particularly congenial environment for noticing-based actions because a) participants are concerned with audience engagement, and noticing-based actions constitute a powerful resource for viewers to initiate contributions and display engagement; b) the video field, for the selection of which the streamer is accountable, provides a kind of focused field of visual relevance, so that what can be seen there or what happens there is especially mentionable; c) especially with expository streams, the specific asymmetry of communicative affordance highlights the spectacular dimension of the stream: the streamer and her environment can be seen as 'on display' for the visual consumption of an invisible audience. I want to discuss in this section a consequence of the interplay of that particular affinity for noticing-based actions in live streams and the collective interactional features that characterize this social media by showing an instance of a stream (see Fig 2) that becomes almost 'saturated' with viewers' noticing-based actions.

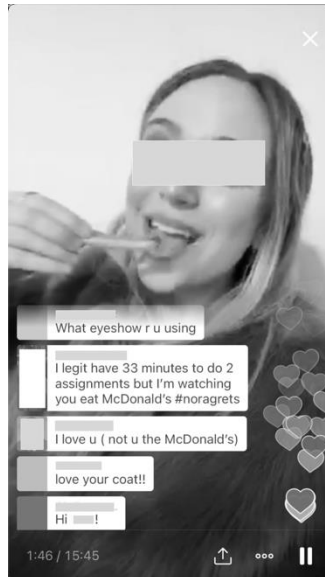


Figure 2. Screenshot of the stream with viewers' noticing-based actions

In this stream, the streamer keeps eating a McDonald's meal (hamburger and fries). She says very little and does not respond to most messages, which is understandable due to the high volume of messages and her engagement in an activity that prevents her from talking. Although this may seem unexciting, her stream attracts a rather large audience of more than three

thousand viewers, leading to a constant flow of responsive messages rapidly appearing on the screen. Since she mostly remains silent, the initiation of topics is primarily left to the audience. Interestingly, among the numerous messages she receives, a significant fraction (roughly two-thirds of the messages) consists of written messages designed as 'noticing-based actions' occasioned by and focusing on perceptible features made available in the live stream. A sequential analysis is not fully relevant here (as the streamer typically does not respond to viewers' noticing-based, sequence-initiating actions). This is a more general issue. While the development of interactional sequences appears to be a potential concern of live stream participants, and noticing sequences may be achieved, as we will see in the next sections, various contingencies may prevent the development of such sequences as is the case here with the sheer number of messages.

However, this particular stream remains particularly interesting for two reasons. From a CA perspective, by providing us with a wealth of perceptually-occasioned messages, it allows us to study their design from an action formation viewpoint, and to give some flesh to our argument for a) the relevance of the concept of 'noticing-based' action and b) that the noticing part may be more or less salient and embedded in the design of such 'noticing-based' actions. From an empirical perspective more focused on an understanding of live streams as interactional events, it will give us insight into a) how deeply streamers may be vulnerable to compliments and criticisms in expository streams and b) how a peculiar kind of collective dynamics—an 'effervescence' of noticing—may develop in live streams, that illustrates the particular affinity between live streams and noticing-based actions.

Considering such messages as "possible" actions (Robinson, 2016: 9-10), we will attempt to provide some evidence for the variety of messages and how they target multiple aspects of the streamer's appearance, conduct, and environment. So, what does the audience notice in our example? What is noticeable there? As we will see, almost everything. First, viewers' attention targets her eating:

- The burger: *"this Mc chicken looks gigantic in your hands lmao"*; *"No Mc nuggets ???? Bold move"*;
- The crispiness of the fries: *"That fry sounded crunch tho"*; *"Why do those fries look so crispy ???"*; *"Alright seriously why are those fries so crunchy"*; *"When the fries have a crunch you know it's real"*.

We can see that the grammatical design of the messages may vary a lot, with complete and incomplete clauses, questions, directives, etc. This is in line with CA research on assessments and compliments, which noted how such actions could be designed in very different ways in talk (Golato, 2005; Shaw & Kitzinger, 2012). We may also consider that these messages index some prior perceptual realization (that is why we may treat them as noticing-based); they vary considerably in terms of the embeddedness of the noticing. We can thus contrast “That fry sounded crunch tho,” which is designed as an explicit, salient noticing, with “When the fries have a crunch you know it’s real,” where the noticing of the sound made when eating the fried is implicit and embedded, and the message is designed to do something else.

Considering her eating, a few messages are designed as corrective instructions regarding her conduct and the way she eats. Such a design embeds a noticing of the conduct in question and enacts it as the kind of conduct which violates some rules of good conduct. Such a design leans towards making complaints and criticisms to the noticed conduct, the noticing of which appears embedded in the instructions: “*eat with your mouth closed*” and “*breathe through your nose.*”

A good number of messages deal with the streamer’s appearance, which is as visually salient as her eating in this talking head configuration. These are usually designed as assessments of features belonging clearly to her personal territory, and, therefore, potentially working as compliments (when displaying a positive stance). They also display some variety in their grammatical design:

- As a whole: “*U cute*”; “*You’re cute actually*”; “*Ur amazing*”; “*Your amazing Ellie*”; “*You are the cuuutest. Love Mc Donalds*”; “*Look cute tonight*”; “*hot*”; “*you’re the woman of my dream; An absolute queen*”; “*You look identical to Elizabeth Olson*”
- Her clothing: “*Cute jacket*”; “*I like your jacket*”; “*Unreal jacket*”; “*Biz jacket is better.*”; “*Dig the fur, solid Saturday look*”; “*alice orange fur coat*”; “*Is it real fur?*”; “*love your coat !!*”; “*Alice is dressed like a mcchicken and I dig it*”.
- Her make-up: “*Love that eye shadow*”; “*I loveeeee your make up*”; “*I like the eye shadow*”; “*YOUR EYE SHADOW (heart emojis)*”; “*How do you do your brows lol*”; “*What eyeshow r u using*”; “*What eyeshadow pallet is that*”; “*Where is ur makeup from*”; “*Eyeshadow brand!*”.

Almost everything one can perceive about how she looks and she behaves seems noticeable in this way, leading to some rather unusual potential compliments: “*You’ve a very unique jaw/mouth structure, it’s quite attractive/appealing*”.

Another series of noticing-based messages deals with her visible environment and the events that may happen there. It is a smaller set, for, in a talking head configuration, the local ecology of the streamer is less salient than the streamer by video shot design :

- Her flat: *“Ya need some artwork on the wall”*; *“It’s lit as hell”*.
- Local events, such as, following the brief appearance of a young woman on the screen beside her, *“your roommate is fire”*.

Various other aspects of the context may provide noticing-based messages, such as the time and the date of the live stream: *“Alice why you home so early”*.

Finally, a significant group of noticing-based actions targets the live streaming event itself, and the experience it provides:

- The size of the audience, which appears worthy of noticing: *“These comments are going really fast lmao”*; *“Damn you got 3k people just watching you eat a cheeseburger that’s clout”*.
- The live stream experience: *“hottest Mac Donalds eating ever”*; *“U eat McDonalds so attractive This drunken eating video is waaaayyyy better than Hasselhof’s”*; *“this is sickly entertaining lol”*.
- Moral criticisms of the live stream: *“3.5 k gross”*; *“Simp nation”*; *“Do you enjoy all this attention?”*.

This last group of messages noticing the streaming event itself nicely displays the moral economy of live streaming we referred to earlier. A high level of audience and viewer engagement, as is the case here, provides a positive contribution and value to the streamer as a person (“that’s clout,” as one message has it). Indeed, one of the very few things the streamer says is a thanking and appreciative turn-at-talk, which acknowledges such a positive contribution to her face: *“so many of you I really appreciate you it’s really nice (.) all your comments”*. However, it has to be deserved by the quality of the viewing and listening experience that is provided in the stream and the messages above, and some viewers clearly doubt this is the case, making her vulnerable to a particular form of criticism, being just a kind of empty attention-seeker.

Three notable aspects stand out here. Firstly, there’s a striking diversity in the noticeables targeted by the messages. These encompass anything perceptible in the stream, ranging from the streamer’s appearance, clothing, and conduct, to persons, things and events in her

environment, and more. Predictably, the volume of messages appears to be roughly correlated with the visual salience of potential noticeables. Consequently, there is a concentration of messages dealing with the streamer's physical appearance, clothing, and conduct (if she is engaged in some visible activity), all sorts of things that are particularly salient visually in a talking head configuration.

Secondly, nearly all noticing-based messages seem to point to aspects for which the streamer holds personal accountability. Consequently, most of these actions can potentially be interpreted ritually, either as face-enhancing or face-threatening, similar to assessments of such noticeables, which are often perceived as compliments or criticisms. Notably, a significant number of these noticing-based messages are from strangers, suggesting a claim to entitlement in noticing-based actions that would be considered intrusive or even offensive on the part of strangers in face-to-face situations. This is evident in messages expressing a negative stance on behavior, such as "*eat with your mouth closed*," or even apparent compliments like "*u cute*."

Thirdly, the abundance of noticings in the form of noticing-based messages, accumulating and resonating with prior messages, is notable. There likely exists a gendered dimension to this intensity, especially given that it occurs with a female streamer, reflecting the gendered performances and performativity observed on streaming platforms (Zhang & Hjorth, 2017). However, this phenomenon appears to be more general, characteristic of streams with a high number of engaged viewers. Such streams are considered potential sites for "choral response" (Dutt & Graham, 2023) and collective "emotional amplification," reminiscent of Durkheim's concept of collective effervescence (Luo et al., 2020; Guo & Fussell, 2020), as well as collective imitation concerning the nature and stance of messages (Seering et al., 2017). Aligned with this perspective on live streaming, we can describe the example here as illustrating a form of "noticing effervescence." In this scenario, the viewer's orientation towards noticing intensifies collectively, leading to contagion and amplification of potentially face-enhancing and face-threatening actions concerning the streamer.

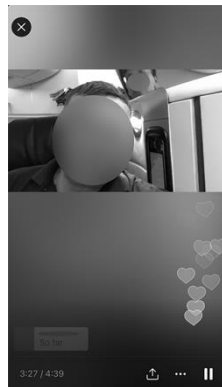
5. Noticing sequences in live streams

Live streams are characterized by a concern from streamers to elicit viewers' responses that are relevant to what is going on. One powerful way for viewers to engage, and to make that engagement visible is for them to formulate and make relevant some perceptible feature/event

of the stream, that is, to ‘notice’ some feature of the stream (when the gist of the message is the formulation of such an occasioned perception) or to embed it more or less deeply in what therefore appears as a ‘noticing-based’ action. Though the sequential organization of live streams is looser than in face-to-face interaction, with the relaxation of the requirement of joint attention and a lessened pressure to respond, streamers and viewers nevertheless achieve noticing sequences. Let us consider, for instance, the following example. The streamer is on a plane about to leave, which can be inferred both from the title of the stream and the fact he is still able to use his phone.

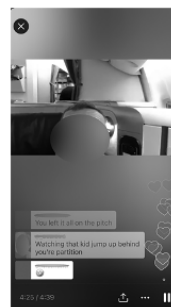
Extract 1 (Title: “Time to leave for #Paris”)

Image 1.1



1 **Nor:** No Hawai'i layovers 😊 next trip..
 2 S: ((brings the camera close to his face))
 3 **The:** Bon Voyage n see u soon from Milann
 4 S: Hawai'i lay (.) \$overs OK next time why not
 5 **Ron:** \$Taking off?
 6 (2.5)
 7 S: yeah now well (.) Paris first
 8 S: yeah\$ (.) er lea\$ving:: Lima Peru right now to go
 9 **Nor:** \$You left it all on the pitch
 10 France to Paris then \$home to erm-
 11 **Dea:** \$Watching that kid jump up behind you're
 partition
 12 **Nor:** 😊
 13 (4.0)
 14 S: ((slightly lifts the camera to check the space behind him#))

#Image 1.2



15 S: er that's no kid that's a man (1.0) \$ (hh) (hh) (hh) (hh) (hh)
 16 **Chi:** \$have a good flight
 17 **Nor:** 😊
 18 S: anyway take care (0.5) (hh) (hh) (hh) (hh)
 19 S: see you around (.) OK (.) ciao for now(.) bye.
 20 **Ron:** @Ann Hi Ann
 21 S: ((disconnects))

The viewer provides various written responses that are fitted to the situation (see line 5 for instance), but we would like to focus here on Dea's message "Watching that kid jump up behind you're⁵ partition" (Line 11). The message formulates a perceptible event (a kid jumping behind the streamer), thus turning it into a noticeable. Part of the noticeability rests on the incongruity of the noticed conduct, which makes it accountable. It appears as a 'retro sequence', though, in terms of relationship to some prior perceptual realization, the design of the message frames the viewer as 'watching the event', that is as if he were still in the act of perceiving, somehow abolishing the temporal distance between the prior perception and the verbal formulation. The streamer initially orients to this written message as a noticing. First, he moves his phone to make visible the partition behind him (which is an equivalence to turning his head to visually check the noticeable; image 1.2). However, the event is not observable anymore. By correcting one aspect of the formulation of the noticeable, and stating it is not a kid (Line 15), the streamer

⁵ "You're" should probably read "your" with a misspelling due to automatic text suggestion.

shows he now orients to the initial message as an informing: he treats the noticeable event as established, and the initial message as a telling opening a sequential slot to agree or disagree with the proposed formulation.

This sequence shows how noticings (in the CA sense) may still be a relevant category to understand interactions in live streamings. Viewers' written messages may be designed and recognized as doing noticing. However, there are significant differences, especially with respect to the temporal organization of actions in live streaming. First, while they may still be designed and treated as a response to earlier perceptual realizations, that is, as second pair parts of 'retro sequences', the temporal relationship between the prior and the noticing is very loose in live streamings. Considering the time necessary to write the message and the lack of control regarding its appearance, there might be a long time between the prior and the noticing. In the example above, the delay is long enough for the noticeable not to be available anymore, and to be consequential (it opens up the possibility of eventually treating the noticing as an informing). Second, CA research has shown that noticings tended to cluster at the beginning of encounters, which makes discoverable features of the setting (Schegloff, 2007; Pillet-shore, 2020). Here, viewers may come in all the time, and notice as they get into the stream, which marks for them the beginning of the encounter. The streamer may, therefore, be the recipient of noticings throughout the stream, though viewers still orient towards the relevance of noticings at the beginning of encounters.

6. Information requests done as a 'noticing-based action'

Clearcut cases of noticings, such as extract 1 are rare in our corpus. CA research has shown that noticings could do all sorts of actions besides noticing (Schegloff, 2007), and this seems also true in live streams. In expository live streams in particular, unlike the previous example in which the noticing targeted a noticeable event for which the streamer was not accountable, many noticings from the audience target noticeables that are "owned" by the streamer (Sacks, 1992; Pillet-Shore, 2020), and are often therefore potentially intelligible as face-enhancing (e.g. complimenting, praising) or face-threatening (e.g. complaining, criticizing) actions. In the following example, the streamer, framed as a 'talking head', has just started his stream, and begun to tell a story (Lines 3-5), in a configuration that is characteristic of expository streams.

Extract 2 (Title: “What’s the future like ?”)

Image 2.1



- 1 **0:34 REA: Hey**
2 S: hey DOU, how is it going,
3 Err, so I had a bar/becue dinner (.) in my home last night,
4 cuz I got (.) one of those American barbecue grills
5 It's like (.) =
6 **0:55 PAU: What's the symbol on your shirt**
7 S: =a fi:ve (.) burner (.) thing (0.5) it's a huge one (.) you know,
8 (hh)(hh)it felt like an American doing that burger-flipping (.)
9 thing
10 S: ((leans forward))
11 S: e::r *it's just* a skateboard T-shirt it's a skateboard brand
 *hand goes to T-shirt design# and down again

#Image 2.2



- 12 S: (0.5) ((humidifies lips)) (.) e:r (.) yeah and then it rained in
13 the middle of the night

Then, one viewer, Pau, provides a written contribution that appears irrelevant to the story being told. It is designed as a recognitional question: “What’s the symbol on your shirt” (Line 6). The question presupposes a noticing sequence: it indexes some prior perceptual realization (paying attention to the logo, which was visually salient in this kind of ‘talking head’ video framing). However, the temporal relationship with this prior noticing remains vague and implicit from

the streamer's point of view: since the logo was always visible, such a perception could have occurred at any point before. In the response line 11, it is oriented to as an information request. Interestingly the response has two parts, and involves an "action repair" (Drew et al., 2013). In the first part, "it-s just a skateboard T-shirt", through the modalization "just", the streamer seems to minimize the noticeability of his T-shirt. It shows how 'noticing' something verbally allows the recipient a slot to align (or disalign) with the formulation of the noticeable and its noticeability. In the second part of the answer, the utterance is self-repaired so that the modalization 'just' disappears, thus eventually aligning with the noticeability of the T-shirt logo, and producing a straightforward answer to an information request. The self-correction in line 11 displays the streamer's orientation toward providing an answer that fits the question. Such an identification-oriented question, while keeping the noticeable and its formulation salient can and is understood as an information request, which presupposes and builds upon a noticing. It is already an instance of a 'noticing-based action.'

It is also observable that the noticing-based information request is done so that it could appear interruptive. It is topically irrelevant with respect to the storytelling. Also, the message appears instantaneously in the middle of a turn construction unit, which is part of the streamer's story-in-progress. This is another instance of how the phenomenon of "disrupted turn adjacency" observed in multi-party, asynchronous computer-mediated chats (Herring, 1999), may be relevant to live streaming. Moreover, unlike with talk, where talking and hearing are usually simultaneous, the streamer can be seen to read it later when he leans forward (Line 10). While the leaning forward is done at a transition-relevant point, it is also done at a moment when the story is recognizably incomplete. After reading the message, the streamer immediately responds to it (Line 21). Such a routine treatment of a turn- and topic-interrupting action to complete a noticing-based sequence shows the live streaming situation's moment-by-moment openness to all sorts of viewer responses, particularly noticing-based responses. Even if there are normative constraints regarding what might be a legitimate comment, this openness is a consequence of the concern for eliciting viewer engagement and displays of attention in live streaming situations.

Extract 3 provides a very different example of how what is written to look as a prototypical noticing in a message can constitute a resource for doing other things. Here, the streamer is driving and just opening her stream with several viewers almost immediately connecting to it (Lines 1-5).

Extract 3 (Title: “Tuesday”)⁶

Image 3.1



1 0:04 **SNO joined**
2 0:11 **PIC joined**
3 0:12 **JOA joined**
4 0:13 **RAN joined**
5 0:14 **MUK joined**
6 S: He:y\$ Mr. PIC, how are you \$going?
7 0:15 \$**RAN joined**
8 0:16 \$**SOM joined**
9 (2.0)
10 S: how is \$your arv going
11 0:19 **PIC:** \$**Oops your driving. G'day and bye**
12 S: Ah hahaha bye hehehe
13 Yes, don't encourage bad behavior
14 (10.0)

One of them, PIC, then messages her with a first component of the message designed to look like a canonical noticing turn in talk (Line 11). The initially positioned “oops” can be read to work both as a change-of-state token (Heritage, 1984), indexing some realization and as an expressive response cry displaying that something is amiss with the speaker/writer. As the fact that the streamer is driving is obvious, the noticing is intelligible, as occasioned by her perceptual ‘discovery’ of the situation when entering the stream. It is followed by a formulation of the noticeable, here her perceived current activity, “your driving.” However, this ‘noticing’ is followed in the message by another component, which does greeting and exiting in the same move. The succession of the noticing, a greeting, and closing in the same message frame the latter actions as made relevant by the earlier, explicit noticing, i.e., as ‘noticing-based actions’. The noticing appears retrospectively as a resource to account for the action the message eventually accomplishes, i.e., disengaging from the stream. All this implicitly relies for its

⁶ “Arv” in line 10 means “afternoon” in oral Australian English.

intelligibility on a social norm, that one should not be streaming (or more generally engaging in any attention-requiring activity) while driving, which gives a reproachful undertone to the noticing itself. The streamer displays such a reading of the message by addressing its multiple sequential implications. First, she laughs and greets, thus aligning with the viewer's exit (Line 12). Then, she provides approval for the way the exit has been made accountable and provides her own account for the exit (the obligation not to support potentially safety-endangering behavior), an account that also involves an admission of her own guilt.

While the message eventually does 'leave-taking', it is designed and understood as a noticeable 'leave-taking' occasioned by the noticing of a conduct that would be incompatible with further streaming engagement. One could say the exit is a noticing-based action, which requires some noticing to be done and understood. This is very explicit here, but in other cases, the noticing may be more implicit and embedded in the action a given message might be doing 'on the record.' Another important case of noticing-based actions in expository live streams is that of compliments and criticisms, which are massive phenomena in live streams.

7. Noticing-based compliments and criticisms in live streams

In most cases and all types of live streams, viewers' noticings target perceptible features of the live streaming situation, which they make relevant. The corresponding messages may often be understood as both a noticing and/or an assessment. It is only the streamer's response that may distinguish between the two kinds of actions. In our corpus, written noticing/assessments from viewers are quite common in 'experiential streams' where the streamer shows something happening, and the audience usually engages by commenting on it. With expository streams (or expository episodes within a stream), the situation is a little different for the streamer, and often her personal surroundings are 'on show' so that a lot of what is made potentially noticeable pertains to what Goffman has called "territories of the self", that is ecological features "that are ordinarily patrolled and defended" by the person, over which she may have "claims", and therefore assumes accountability and responsibility (Goffman, 1971, 29). When bearing on such noticeables, assessments will usually be understood as expressing some positive (resp. negative) stance with respect to the streamer's face. This is exactly what complimenting (resp. criticizing) is about (Pillet-Shore, 2015). Let us consider a first example (Extract 4). The transcript starts 1'30'' into the stream.

Extract 4 (Title: “Come Smoke”)

- 1 MEL: What's u PPP I like ur necklace
- 2 MEL: The jade
- 3 YOW: ⚡⚡
- 4 S: Right on #(.) appreciate it.

#Image 4.1



In this instance, MEL connected relatively soon after the beginning of the stream and relatively soon after her entrance. She sends a ‘how-do-you-do’ type question, followed by a unit about the streamer’s necklace in the same message (Line 1). The message is noticing-based, indexes some prior perception of the necklace, which is visually salient in the video frame, and makes the necklace into a stream-relevant ‘mentionable’. It is designed as a positive assessment of what the streamer wears (which is one of the typical “personal territories” of an individual). It is, therefore, a noticing-based action, potentially intelligible as a compliment. The viewer then self-corrects her message, respecifying the noticeable, i.e., the jade it visually appears made of (Line 3). The streamer responds to both components of the message, and orients to the mention of the necklace as a compliment, through his thankful response, expressing how he is positively affected by it (Line 4). In CA studies of compliments, it has been shown that responding to them involved managing cross-cutting preferences, the preference for responding to positive assessments with more positive ones (Pomerantz, 1984), and the dispreferred character of self-praising (Pillet-Shore, 2015). The streamer’s response here manages nicely the reception of the compliment while avoiding self-praise or smugness.

Up to the moment when the noticing message appears, the streamer has said almost nothing. In the preceding minute and a half, he has just shown the joint that he is smoking, a gun, and then a bottle of alcohol, apparently going for a ‘gangster rap’ atmosphere. Therefore, there is no apparent and salient activity or ongoing topic to recognize. By showing a succession of objects, he seems to guide the audience into noticing his behavior, though MEL targets a different

noticeable. In conversation analysis, noticings are a resource to keep the conversation going, and it has been observed that in video-mediated communication, they constitute a powerful resource for topic initiation (Zouinar & Velkovska, 2019). In streams where streamers appear as talking heads and do little, noticing-based messages from viewers seem to play a similar role.

While all this suggests some aspects of CA research on assessment and compliment sequences may be relevant to live streams, the setting is very different from a face-to-face conversation. Viewers are invisible, often anonymous, and often unacquainted with the viewers. They cannot be held accountable for the kind of facework they are doing in the same way as a co-present participant. Let us consider the following example.

Extract 5 (Title: “Tuesday”)

Image 5.1



01 S: but yes, \$overall=
02 **ALO:** \$you're sexy as hell
03 S: =good working day (.) Ah:: Tuesday's (1.0)
04 not *as bad
05 *((looks down at screen))
06 S: OH: thank you.
07 (7.0)
08 S: ((looks at screen))
09 S: what is everyone up to/ It is currently four
10 eleven in the afternoon in Sydney Australia
11 (1.0)
12 S: it just stopped raining

In this stream, the streamer (the same one as in extract 3) has been complaining about how few people are connected to her stream. To get things going, she starts to describe her day, wrapping up her telling with an overall assessment (Lines 1 and 3-4). In the middle of this, the viewer ALO messages her, “you’re sexy as hell” (Line 2). The design of the message combines features of a noticing (it topicalizes the way she appears visually at present, which indexes some prior

realization) and a highly positive assessment (“sexy as hell”) so that this noticing-based action is potentially intelligible as a compliment. Indeed, once she appears to have noticed the message herself (Line 5), she treats it as such, first uttering a change-of-state token and then thanking the viewer.

The compliment itself is potentially a delicate matter as it targets something personal (her appearance) and sexualizes it so that it can potentially be treated as flirting, which relies a lot on such sexualized compliments (Speer, 2017). At the same time, it is apparently coming from a stranger (she usually marks the recognition of her acquaintances when responding but does not do so with this viewer). She does not appear to treat the message as offensive (her thanking response is affiliative, at least on the surface). However, some aspects of the design of that response suggest that she may be sensitive to the delicacy of the matter. First, the initial change-of-state token and the prosody mark a surprise. Even though several messages and topics “come out of the blue” during her stream, she usually responds directly without marking surprise in this way. Second, her response is minimal and does not project further talk. She does not contribute any more talk for several seconds. She follows this silence with a complaint about the lack of audience response (which is why she started talking about her day in the first place). However, she just had that one from ALO, which makes it seem like she is not even considering it before eventually resuming her story (Lines 9-12). This marks the previous exchange as one not to be pursued, worthy only of minimal treatment on her part. This lack of further engagement suggests that the streamer treats the message as somewhat inappropriate.

This ambivalence (acknowledging the message while minimizing further talk) may point towards a more general phenomenon in streams. First, streamers make themselves perceptually available (and intentionally so) on streams, but their appearance is made especially salient by the asymmetry of perceptual access between streamers and viewers. Second, because viewers remain perceptually invisible and largely anonymous (they mostly use pseudonyms), as is the case with many interactive online media, they may feel emboldened to notice the streamers’ appearance, conduct, and environment since they will be less accountable for doing so than would otherwise be the case in other types of encounters, such as face-to-face ones.

In another stream in our data, with so much audience that it would take up too much space to analyze all the messages it contained, a “troll” started to send abusive messages, some construed as noticings of their appearance and treated as derogatory. In this case, the two young female

streamers, who were trying to respond to all of their messages, showed in various ways that they were taking these abusive messages as improper and even offensive. They ignored the messages (while responding to the preceding and following ones), displayed their indignation, and evoked the possibility of blocking the viewer, which they eventually did. This shows how the sanction system regarding improper noticing in live streaming combines responses relevant to face-to-face encounters (ignoring, complaining, and confronting the offender) with responses that reflect the general organization of sanctions in social media, such as threatening to block and actually ‘canceling’ the offender. However, it took a lot of abusive messages for the two streamers to block such a ‘troll’ which suggests that streamers who display themselves and strive to engage online with unknown and invisible others may be ready to give their audience much leeway regarding the entitlement to notice and the propriety of noticing, the kind of license they probably would not be ready to give in face-to-face encounters or symmetric video-mediated communication. On the viewers’ side, audiences seem to actively exploit that license when noticing streamers’ personal territory.

8. Conclusion

Streams and streamers are, therefore, oriented toward eliciting various responses from viewers, that appear responsive to what the stream makes perceptually available. Our study shows that a powerful resource for viewers to achieve and display engagement in live streams is to send messages designed as noticing-based actions, which index some prior perceptual realization regarding features or events in the stream, to achieve prospectively various social actions, and particularly assessing, complimenting, or criticizing. We have introduced the notion of ‘noticing-based action’ rather than using the term ‘noticing’ to account for the fact that, in many cases, what is noticed and treated as noticeable is not formulated explicitly and may be deeply embedded in the design of the messages. This notion might prove useful for CA, and we plan to explore it in other settings in future work. Such an action format casts viewers as onlookers, framing the streaming activity as a spectacle—a particular framing favored by live stream affordances that make live video content available to a remote and mostly anonymous audience. While this observation is relevant to live streaming in general, our focus here is on expository live streams, where the streamer is expected to be mostly presented as a ‘talking head.’ Utilizing EMCA, we have provided instances of noticing-based actions and sequences, specifically examining noticing sequences in section 5, a noticing-based information request sequence in section 6, and discussed noticing-based compliment and criticism sequences in section 7.

In expository streams, 'noticeables' often encroach upon the streamer's 'personal territories.' This was evident in the sequences we analyzed in section 7, and notably pronounced in the 'effervescent' stream of section 4. In this case, noticing-based messages look into various personal dimensions of the streamer on display—her physical appearance, conduct, clothing, environment, etc. This dynamic introduces a strong ritual component in the production of noticing-based messages. When these messages target 'personal' noticeables, they can be interpreted as either face-enhancing or face-threatening actions. Consequently, as we saw in section 7, noticing-based messages designed as assessments are often understood as compliments or criticisms, rendering the streamer vulnerable in distinctive ways. The combination of the noticing orientation of viewers in live streams and the framing of viewers as an anonymous and remote audience creates an environment where viewers can notice things in ways that might be considered face-threatening and even abusive in other settings. Additionally, this framing implies that holding viewers accountable would involve considerable effort on the part of streamers. The recurrence of noticing-based messages in live streams offers evidence of streamers adopting an onlooker role and highlights the potency of such messages in achieving/displaying engagement. The specific case examined in section 4 illustrates how this orientation can intensify, creating a 'noticing effervescence' that becomes contagious. Local factors may contribute to this phenomenon, such as the streamer's minimal activity in the stream, making noticing-based actions the primary resource to initiate new topical lines. This aligns with findings in video-mediated communication (Zouinar & Velkovska, 2019), emphasizing the significance of noticing-based actions. The collective 'noticing effervescence' observed suggests that the noticing orientation tracked through the analysis of noticing-based messages by viewers could be a constitutive feature of live streaming as a communicative event. Social media platforms have been described as embedded in an "economy of attention" that monetizes quantitative audiences, as is particularly the case with vloggers and influencers (Citton, 2017). Some of the messages in section 4 highlighted a moral side to this economy of attention in the case of live streaming. Some viewers remarked that this 'effervescent' stream had many viewers and messages without deserving it because the streamer did not say or show anything, nor did he or she behave in a way that might be worthy of such attention. Such a concern goes beyond a preoccupation with the number of viewers, which streamers may display when they wait for enough viewers to connect (as in extract 2) or when they complain more explicitly there are few viewers. It suggests a moral valuation underpinning in which large audiences should be 'deserved' on the basis of what the streamer says, shows, or does, and which therefore links the moral worthiness of the streamer as such with the success of her

stream in terms of viewer numbers and engagement, particularly in the form of noticing-based responses. Such a moral economy of live streaming and the way it is articulated in actual live streaming interactions deserve further research.

Data availability

Anonymized transcripts of the data underlying this article may be shared on reasonable request to the corresponding author.

Appendix. Transcript Conventions

This paper's transcription conventions mainly adopt Gail Jefferson's (2004) conventions for talk and Lorenza Mondada's (2018) conventions for embodied actions. Courier New font in nine-point is used for the transcripts. Besides, some particular conventions for live streaming have referred to Licoppe and colleagues' works (Licoppe & Morel, 2018; Song & Licoppe, 2023). The main convention symbols involved in this article are as follows:

normal font: the streamer's utterances and actions;

bold font: the audience's text comments;

, means intonation continuation;

. means downward intonation;

\ and / represent the falling and rising intonation, respectively;

(.) means inconspicuous short pauses of the streamer's talk (usually less than 0.2 seconds);

(2.0) means the silence time of the streamer's talk when more than 0.2 seconds;

§ indicates the moment when the audience's messages come up;

*--->

---->* asterisk marks the moment and the phase (when the action continues longer than just a moment) of the action;

indicates where the screenshot appears;

= indicates the connection of two turns;

(()) indicates non-verbal behavior.

References

Anderson, S. L., & Johnson, M. R. (2022). Gamer identities of video game live streamers with disabilities. *Information, Communication & Society*, 25(13), 1900-1915.

- Anjani, L., Mok, T., Tang, A., Oehlberg, L., & Goh, W. B. (2020). Why do people watch others eat food? An empirical study on the motivations and practices of mukbang viewers. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/3313831.3376567>
- Bergmann, J. (1990). On the local sensitivity of conversation. Ivana Markova and Klaus Foppa (Eds.). *The Dynamics of Dialogue*. Hemel Hempstead, Harvester Wheatsleaf: pp. 201-226.
- Böckler, A., & Sebanz, N. (2013). Linking joint attention and Joint Action. *Agency and Joint Attention*, 206–215. <https://doi.org/10.1093/acprof:oso/9780199988341.003.0013>
- Campion, P., & Langdon, M. (2004). Achieving multiple topic shifts in primary care medical consultations: A conversation analysis study in UK general practice. *Sociology of Health and Illness*, 26(1), 81–101. <https://doi.org/10.1111/j.1467-9566.2004.00379.x>
- Chen, Z., Cao, H., Deng, Y., & Gao, X. (2021). Learning from home. A mixed-methods analysis of live streaming-base remote education experience in Chinese colleges during the COVID-19 pandemic. In *proceedings of CHI '21: CHI Conference on Human Factors in Computing Systems*. DOI:[10.1145/3411764.3445428](https://doi.org/10.1145/3411764.3445428)
- Choe, H. (2019). Eating together multimodally: Collaborative eating in mukbang, a Korean livestream of eating. *Language in Society*, 48(2), 171–208. <https://doi.org/10.1017/s0047404518001355>
- Choe, K. (2020). Talking the cat. Footing lamination in a Korean livestream of cats mukbang. *Journal of Pragmatics* 160, 60-79.
- Citton, Y. (2017). *The ecology of attention*. John Wiley & Sons.
- Drew, P., Walker, T., & Ogden, R. (2013). Self-repair and action construction. *Conversational repair and human understanding*, 30, 71-94.
- Due, B. L., & Licoppe, C. (2021). Video-mediated interaction (VMI). *Social Interaction. Video-Based Studies of Human Sociality*, 3(3). <https://doi.org/10.7146/si.v3i3.123836>
- Dutt, S. and S. Graham (2023). "Video, talk and text: How do parties communicate coherently across modalities in live videostreams?" *Discourse, Context & Media* 55, . <https://doi.org/10.1016/j.dcm.2023.100726>.
- Fang, K. (2023) The social movement was live streamed: a relational analysis of mobile live streaming during the 2019 Hong Kong protests, *Journal of Computer-Mediated Communication*, Volume 28, Issue 1, January 2023, zmac033, <https://doi.org/10.1093/jcmc/zmac033>
- Garfinkel, H. (1967). *Studies in Ethnomethodology*. Englewood Cliffs, NJ: Prentice-Hall.

- Garfinkel, H. (2002). *Ethnomethodology's Program. Working Out Durkheim's Aphorism.* Lanham, Rowman & Littlefield.
- Goffman, E. (1971). *Relations in public: Microstudies in public.* Penguin Books.
- Goffman, E. (1974). *Frame Analysis. An Essay on the Organization of Experience.* New York, Harper and Row.
- Goffman, E. (1983). "The Interaction Order." *American Sociological Review* 48(1): pp. 1-17.
- Golato, A. (2005). *Compliments and Compliments Responses: Grammatical Structure and Sequential Organization.* Amsterdam/Philadelphia, John Benjamins.
- Goodwin, M. H., & Goodwin, C. (2012). Car talk: Integrating texts, bodies, and changing landscapes. *Semiotica*, 2012(191). <https://doi.org/10.1515/sem-2012-0063>
- Guo, J., & Fussell, S. R. (2020). A Preliminary Study of Emotional Contagion in Live Streaming. Conference Companion Publication of the 2020 on Computer Supported Cooperative Work and Social Computing. doi:10.1145/3406865.3418309
- Haimson, O., & Tang, J. (2017). What Makes Live Events Engaging on Facebook Live, Periscope, and Snapchat. In CHI '17: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, May 2017, 48–60. <https://doi.org/10.1145/3025453.3025642>.
- Hazel, S., Mortensen, K., & Rasmussen, G. (2014). Introduction: A body of resources – CA studies of Social Conduct. *Journal of Pragmatics*, 65, 1–9. <https://doi.org/10.1016/j.pragma.2013.10.007>
- Herring, S. (1999). Interactional coherence in CMC. *Journal of Computer-Mediated Communication*, 4 (4).
- Jefferson, G. (2004). Glossary of transcript symbols with an introduction. *Conversation Analysis*, 13–31. <https://doi.org/10.1075/pbns.125.02jef>
- Jodén, H., & Strandell, J. (2022). Building viewer engagement through interaction rituals on Twitch. *Information, Communication & Society*, 25(13), 1969-1986.
- Johnson, M. (2022). Humor and comedy in digital game live streaming. *New Media & Society*, <https://doi.org/10.1177/14614448221095>
- Johnson, M. R. & Woodcock, J. (2018). The Impacts of Live Streaming and Twitch.tv on the Video Game Industry. *Media, Culture and Society*, 41(5), 670-688. <https://doi.org/10.1177/0163443718818363>
- Joseph Seering, Robert Kraut, and Laura Dabbish. 2017. Shaping Pro and Anti-Social Behavior on Twitch Through Moderation and Example-Setting. In Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (Portland,

- Oregon, USA) (CSCW '17). ACM, New York, NY, USA, 111--125.
<https://doi.org/10.1145/2998181.2998277>
- Kaisanen, T. (2012). "Uh Oh we're going there". Environmentally-occasioned noticings of trouble in in-car interaction". *Semiotica* ; 191 (1/4), pp. 197-222.
- Kavada A., Treré E. (2020). Live democracy and its tensions: Making sense of livestreaming in the 15M and occupy. *Information, Communication & Society*, 23(12), 1787–1804.
<https://doi.org/10.1080/1369118X.2019.1637448>
- Kesselheim, W., Brandenberger, C., & Hottiger, C. (2021). How to notice a tsunami in a water tank: Joint discoveries in a science center. *Gesprächsforschung - Online-Zeitschrift zur verbalen Interaktion* 22, 87-113
- Kidwell, M., & Zimmerman, D. H. (2007). Joint attention as action. *Journal of Pragmatics*, 39(3), 592–611. <https://doi.org/10.1016/j.pragma.2006.07.012>
- Kunitz, S. (2018). Collaborative attention work on gender agreement in Italian as a foreign language. *The Modern Language Journal*, 102, 64-81.
- Laanesoo, K., & Keevallik, L. (2017). Noticing breaches with nonpolar interrogatives: Estoniankes(“who”) ascribing responsibility for problematic conduct. *Research on Language and Social Interaction*, 50(3), 286–306.
<https://doi.org/10.1080/08351813.2017.1340721>
- Li et al. (2018). You Watch, You Give, and You Engage: A Study of Live Streaming Practices in China. *In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems 2018*
- Licoppe, C., & Morel, J. (2012). Video-in-interaction: “Talking heads” and the Multimodal Organization of Mobile and Skype Video calls. *Research on Language & Social Interaction*, 45(4), 399–429. <https://doi.org/10.1080/08351813.2012.724996>
- Licoppe, C., & Morel, J. (2018). Visuality, text and talk, and the systematic organization of interaction in periscope live video streams. *Discourse Studies*, 20(5), 637–665.
<https://doi.org/10.1177/1461445618760606>
- Luff, P., Heath, C., Kuzuoka, H., Hindmarsh, J., Yamazaki, K., & Oyama, S. (2003). Fractured ecologies: creating environments for collaboration. *Human-Computer Interaction*, 18(1-2), 51-84.
- Luo, M., Hsu, T. W., Park, J. S., & Hancock, J. T. (2020). Emotional Amplification During Live-Streaming: Evidence from Comments During and After News Events. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW1), 1-19. doi:10.1145/3392853

- Marulanda, J.F.F.; Collazos, C.A.; Hurtado, J.A. Evaluating an Immersive Virtual Classroom as an Augmented Reality Platform in Synchronous Remote Learning. *Information* **2023**, *14*, 543.
- Maynard, D. W., & Zimmerman, D. H. (1984). Topical talk, ritual and the social organization of relationships. *Social Psychology Quarterly*, 47(4), 301. <https://doi.org/10.2307/3033633>
- Mondada, L. (2014). Pointing, talk, and the Bodies. From Gesture in Conversation to Visible Action as Utterance, 95–124. <https://doi.org/10.1075/z.188.06mon>
- Mondada, L. (2018). Multiple temporalities of language and body in interaction: Challenges for transcribing multimodality. *Research on Language and Social Interaction*, 51(1), 85–106. <https://doi.org/10.1080/08351813.2018.1413878>
- Oittinen, T. (2020). Noticing-prefaced recoveries of the interactional space in a video-mediated business meeting. *Social Interaction. Video-Based Studies of Human Sociality*, 3(3). <https://doi.org/10.7146/si.v3i3.122781>
- Paulus, T., Warren, A., & Lester, J. N. (2016). Applying conversation analysis methods to online talk: A literature review. *Discourse, Context and Media* 12, 1-10. <http://dx.doi.org/10.1016/j.dcm.2016.04.001>
- Pillet-Shore, D. (2015). “Compliments”. In K. Tracy, C. Ilie, & T. Sandel (Eds.), *The International Encyclopedia of Language and Social Interaction*. Wiley. DOI:[10.1002/9781118611463.wbielsi127](https://doi.org/10.1002/9781118611463.wbielsi127)
- Pillet - Shore, D. (2020). When to make the sensory social: Registering in face-to-face openings. *Symbolic Interaction*, 44(1), 10–39. <https://doi.org/10.1002/symb.481>
- Pomerantz, A. (1984). Agreeing and disagreeing with assessments : some features of preferred/dispreferred turn shapes. *Structures of Social Action*. J. M. Atkinson, and Heritage, J. Cambridge, Cambridge University Press: pp. 57-101.
- Rauniomaa, M., Haddington, P., Melander, H., Gazin, A.-D., Broth, M., Cromdal, J., Levin, L., & McIlvenny, P. (2018). Parsing tasks for the mobile novice in real time: Orientation to the learner's actions and to spatial and temporal constraints in instructing-on-the-move. *Journal of Pragmatics*, 128, 30–52. <https://doi.org/10.1016/j.pragma.2018.01.005>
- Recktenwald, D. (2018). The discourse of online live streaming on Twitch: communication between conversation and commentary. Ph.D dissertation. Available at <https://theses.lib.polyu.edu.hk/handle/200/9795>.
- Robinson J. D. (2016). *Accountability in Social Interaction*. Oxford: Oxford University Press.

- Rüdiger, Sofia. 2022. "Intimate Consumptions: YouTube Eating Shows and the Performance of Informality." *Internet Pragmatics* 5(1): 115-142. <https://doi.org/10.1075/ip.00070.rud>
- Schegloff, E. A. (2007). Sequence organization in interaction: A primer in conversation analysis I (Vol. 1). Cambridge university press.
- Shaw, R. and C. Kitzinger (2012). "Compliments on a Home Birth Helpline." *Research on Language and Social Interaction* 45: 213 - 244.
- Si, R. (2021). China livestreaming e-commerce industry insights. Springer Singapore.
- Song, L., & Licoppe, C. (2023). Closing live video streams: A sequential analysis. *Discourse, Context & Media*, 53, 100698. <https://doi.org/10.1016/j.dcm.2023.100698>
- Speer, S. (2017). Flirting. A designedly ambiguous action? *Research on Language and Social Interaction*, 50:2, 128-150.
- Tang, J. C., G. Venolia, and Quinn, K. (2016). "Meerkat and Periscope: I Stream, You Stream, Apps Stream for Live Streams." Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, 4770-4780.
- Tannen, D. (2006). Intertextuality in action: Re-keying family arguments in public and in private. *Text Talk* 26 (4/5), 597-617.
- Vázquez Carranza, A. (2016). Remembering and noticing. *Spanish in Context*, 13(2), 212–236. <https://doi.org/10.1075/sic.13.2.03vaz>
- Wang M, & Li D (2020) What motivates audience comments on live streaming platforms? *PLoS ONE* 15(4): e0231255. <https://doi.org/10.1371/journal.pone.0231255>
- White, M. (2003). Too close to see. Men, women, and webcams. *New Media & Society*, 5(1), 7-28.
- Zhang, G., & Hjorth, L. (2017). Live-streaming, games and politics of gender performance: The case of Nüzhubo in China. *Convergence: The International Journal of Research into New Media Technologies*, 25(3) 135485651773816. doi:10.1177/1354856517738160
10.1177/1354856517738160
- Zouinar, M., & Velkovska, J. (2019). Talking about things. *Benjamins Current Topics*, 87–118. <https://doi.org/10.1075/bct.103.04zou>

CHAPTER 5

“OH MY DAVE, YOUR FACE SAYS IT ALL!”: EXPLORING THE MULTIMODAL PRACTICES OF TASTING IN LIVE STREAMING

ABSTRACT

Research on the activity of tasting, examined through the lens of conversation analysis, reveals that in face-to-face contexts, tasting is an interactive and sequential process that combines individual sensory experience with a public, witnessable, accountable, and intersubjective dimension. This perspective can be extended to the realm of live-streamed tasting, where streamers demonstrate and communicate the taste of food products to online audiences in real time. By adopting multimodal conversation analysis to scrutinize the unfolding sequence moment by moment, my study aims to demonstrate 1) Three practices to achieve the configuration of "tasting heads," wherein the current taster's face is displayed on-screen. 2) Patterns of gaze withdrawal from the screen and subsequent gaze back to the screen during the tasting process. 3) The performative and animated facial expressions. 4) The structure of responses from both streamers and viewers after the tasting.

Keywords: tasting, multimodal conversation analysis, live streaming, human-computer interaction, video-mediated interaction

1 INTRODUCTION

As a form of video-mediated interaction (Due and Licoppe, 2021), live broadcasts⁷ have attracted significant attention from researchers worldwide. In recent years, interaction on live broadcasts has become the focus of numerous studies. Various topics related to live streaming have been explored, particularly emphasizing the interactive activities facilitated through this medium. Examples include playing video games (Recktenwald, 2018; Johnson, 2022; Jodén & Strandell, 2021), conducting online classes during the Covid-19 pandemic (Florez Marulanda, 2022), promoting and selling products or services (Si, 2021), showcasing local cultural activities (Lu et al., 2019), and the practice of watching others eat.

⁷ In this paper, I do not distinguish between “live streaming” and “live broadcasting”, “streamer” and “broadcaster”, “audience” and “viewer”. Therefore, they have equal meaning when appearing. Also, I use “message” or “comment” to refer to the live texts sent by the viewers, which may be preferably called “danmaku” or other names in different societies and communities.

Among these, live broadcasting of people eating has gained tremendous popularity. According to recent research (Choe, 2019; Kang et al., 2020), eating broadcast has a Korean background and was initially known and popularly named as “mukbang” (a combination of the Korean words for 'eating' and 'broadcast'). It has quickly become a global phenomenon, gaining popularity worldwide but also raising concerns about negative social impacts such as food misuse and wastage⁸ (Tidy, 2020; Xu et al., 2022), and the gender politics of female ‘big stomach kings’ with starking slim bodies in some countries (Qu, 2021).

Rather than examining general eating activities or food-related topics in live streams, this paper focuses explicitly on the tasting practices in live streaming, adopting a perspective rooted in conversation analysis. From the perspective of conversation analysis, tasting and eating are distinct in several ways. For example, tasting involves a silent exclusive focus on the multisensorial experience, whereas eating is characterized by multiactivity and embedded in talk (Mondada, 2018a). In this research, I aim to show how the activity of doing tasting in live streaming is a collaborative, interactional achievement that differs in organizational terms not only from doing eating, but also from other settings such as face-to-face. As an accountable action visibly performed with the body, I also show how the tasting activity is displayed for, and seen by the co-participants (Mondada, 2020b) in the ecology of live streaming.

Despite their distinction, the inherent connections between tasting and eating are hard to ignore. Therefore, to study tasting in the context of live streaming, it is necessary to start a review from previous eating and live streaming studies.

1.1 Eating, media and live streaming

Eating has always been a significant topic within interaction studies. Goffman (1966: 61) considers eating a form of auto-involvement that requires minimal attention in itself. The existing literature on food-eating research appears closely intertwined with the concept of "commensality." Commensality refers to eating together in a social context, such as with family, friends, and colleagues (Fischler, 2011). It highlights the positive social interactions associated

⁸ In many cases, mukbang has the exaggerated act of devouring large quantities of food (Anjani et al., 2020).

with shared meals, influencing factors like food intake, food choice, and food satisfaction (Sobal, 2000; Sobal & Nelson, 2003; Højlund, 2015; Spence et al., 2019).

As communication devices become integrated into individuals' solitary or group eating activities (Panicker et al., 2020; Lemke & Schifferstein, 2021; Avgustis & Oloff, 2023), the role of technology is increasingly prominent in shaping the eating experience. Mobile devices (Latzer et al., 2015), particularly, have exerted far more influence on people's eating activities than traditional media such as television (Morris & Katzman, 2003) due to their pervasive, portable, and movable attributions.

For example, smartphones and tablets can significantly impact eating, both positive and negative. They may enhance the enjoyable social aspects of dining (Moser et al., 2016; Ferdous et al., 2016a, b), producing "foodporn" (Ibrahim, 2015; McDonnell, 2016; Mejova et al., 2021) across all social media platforms by "camera eats first" people (Yong et al., 2020; Gambetti & Han, 2022; Ahmed et al., 2023). They may contribute to healthy weight management (Andersen et al., 2021) since smartphone applications become novel methods for digitally capturing and tracking information-dense dietary intake (Caspersen et al., 2015), monitoring and intervening in eating patterns (Manoogian et al., 2022). However, they may also significantly increase caloric intake due to using a smartphone or iPad to play video games, chat, or watch Netflix during mealtimes (da Gonçalves et al., 2019). These phenomena of eating with the communication devices involved have been termed "digital remote commensality" (Spence et al., 2019) or "computational commensality" (Niewiadomski et al., 2019).

Besides, filming people's food and eating-related activities and making them into pre-recorded edited videos, or so-called eating shows⁹ (Rüdiger, 2020), occupies a unique intersection between studies on eating activities and video-mediated interaction, capturing the interest of researchers from diverse backgrounds. Eating shows have different forms, themes, and situations (Leer & Krogager, 2021), including video recipes (Melville et al., 2023), traveling & eating shows, food-sampling reaction shows (e.g., McDonald's new product sampling, or Chinese stinky tofu or chicken feet trying), etc. Eating shows can take different media forms, such as YouTube vlogs and TikTok short videos. Rüdiger (2020) examines the use of language and digital food discourse in North American YouTube eating shows, investigating how the

⁹ In this paper, I draw a distinction between 'eating shows' as pre-recorded edited asynchronous videos, and 'eating broadcasts' (mukbang) as synchronous livestreams.

structure and discursive performance contribute to the creation of intimacy and informality (Rüdiger, 2022), as well as how they blur the sense of immediacy and distance (Rüdiger, 2021).

Different from eating shows as pre-recorded edited videos, eating broadcasts (or “mukbang”) normally refer to live streams with synchronous interactions (text messages) from viewers. The popularity of eating broadcasts on multiple live-streaming platforms has obtained considerable attention from researchers in recent years. Empirical methods have been employed to explore viewers’ motivations and practices, aiming to understand why people watch others eat food in live streaming and results suggest that viewers mainly treat eating streams as a primarily private activity, as a mealtime companion, and as a leisure activity (Anjani et al., 2020), and that viewers’ motivations include the streamer’s attractiveness (Pereira et al., 2019), multisensory experience, virtual satiation, physiological experiences, spectacle of the streamer’s performative eating (Anjani et al., 2020; Kim, 2020), the desire for remote digital commensality (Spence et al., 2019), and even sexual reasons, or as an escapist compensatory strategy (Kircaburun et al., 2020),

Commercial and economic aspects play a crucial role in eating broadcasts. Pay items like virtual gifts are utilized between viewers and broadcasters, creating the production and consumption relationship in a digital space and unique cultural meanings beyond capitalism (Bruno and Chung, 2017). Eating broadcasts have reshaped online commerce, with food live-streaming shopping becoming a new form of marketing that may lead to overconsumption (Lee & Wan, 2023). In fact, the promotion effect of streamers is so strong that even the regular onscreen food consumption of game broadcasters can influence the food choices, whether healthy or unhealthy, of emerging adult viewers (Micallef et al., 2022).

The health issue is another research focus for eating broadcast studies. Research has been conducted to show how watching eating broadcasts related to eating disorders and internet addiction, causing problems such as overlong watching (Kircaburun et al., 2020) and overeating (binge drinking and obesity-related eating) (Kim et al., 2021). However, for some people, watching mukbang could be a destructive force that motivates restrictive eating or loss-of-control eating, reduces the sense of loneliness, and improves well-being and mental health (Strand & Gustafsson, 2020).

1.2 Tasting Research from the Perspective of Conversation Analysis

Topics surrounding food, eating, drinking, and tasting have captured the attention of researchers exploring conversation analysis. Several studies in conversation analysis have examined the multimodal structures and functions of eating and drinking in face-to-face conversations (Laurier, 2008; Pomerantz and Mandelbaum, 2016; Hoey, 2018; Ogden, 2021; Majlesi et al., 2022), as well as in video-mediated contexts like live streaming (Choe, 2019; Choe, 2020). Some conversation analysis studies have focused explicitly on the multimodal aspects of sensory experiences and socialization related to eating (Ochs et al., 1996; Wiggins, 2002; Sneijder and te Molder, 2006; Wiggins & Keevallik, 2020). These studies aim to demonstrate that eating behaviors and enjoyment are not solely physiological sensations but also interactive achievements.

There are significant differences between eating and tasting regarding their sequential structures from a conversation analytic perspective (Mondada, 2023). Mondada has made notable contributions to the study of multimodal unfolding practices of tasting in various face-to-face situations, such as cheese gourmet shops (Mondada, 2018a; Mondada, 2019a), professional cheese tasting training sessions (Mondada, 2020b, 2021), and beer tasting experiments (Mondada, 2019b). She also explored the audible smelling action in tasting activities (Mondada, 2020a). Her research investigates how the body engages in sensory explorations involving systematic embodied conducts, sequentially organized and publicly visible, before eventually producing verbalized assessments and descriptions, shared with the co-participants, constructing the intersubjectivity of tasting as a collectivized activity. Temer (2017) has examined the multimodal sequence organization of food assessments in Chilean Spanish interactions, exploring topics such as interaction unit boundaries (Temer and Ogden, 2021), eye gaze, gestures, and prosodic features in food tasting. Goodwins and Pomerantz have also conducted seminal work on assessments, which provides valuable insights relevant to the study of tasting activities, although their focus extends beyond tasting alone. Goodwin and Goodwin (1987) argue that assessments are structured interactive activities that make shared understanding visible and serve to secure the recipient's participation. Pomerantz (1985) investigates the agreement and disagreement in assessments and reveals the types of structure of assessment pairs. She found that when the speaker assumes the recipient has access to the referent, they share the assessment and may seek agreement; when the speaker assumes the recipient lacks access to the referent, the speaker reports assessments and the recipient assesses reporting, displaying understanding. The intertwining of facial expressions and verbal assessments has been examined. Research shows how interactants arrange faces and talk to

display stances and that facial expressions can stretch the temporal boundaries of an action (Ruusuvuori and Peräkylä, 2009).

Previous research on tasting has primarily focused on face-to-face situations and remains limited in scope. To contribute to the existing research on tasting, this paper concentrates on the tasting phenomenon in live streaming and explores the interaction patterns within this context. One notable difference in tasting during live streams is that viewers cannot have specific or complete sensory access (for example, touch or smell) to the food being tasted. This raises practical, interactional, and relational concerns that require examination. The paper addresses two main research questions: a) What is the multimodal sequential organization of tasting in live streaming? b) What are the distinguishing characteristics of tasting practices in live streaming compared to face-to-face situations?

The study utilizes the theory and method of multimodal conversation analysis, analyzing data derived from a corpus of over 90 Twitter live video streams. The paper will show the configuration of "tasting heads" in Section 3.1; the unique gaze patterns and performative facial expressions in Section 3.2; and the structure of responses after the tasting in Section 3.3.

2 MATERIALS AND METHOD

2.1 Twitter Live

For my research, I chose Twitter Live as the platform of focus. This platform offers the advantage of archiving past streams and their associated comments, providing researchers with readily accessible video material for analysis. Additionally, streamers on Twitter Live are geographically dispersed, resulting in diverse content available for study.

The Twitter Live interface features an intuitive design that actively engages viewers with the stream. Users can initiate their own streams. They can also participate by typing and sending in-time-appearing text messages through the bottom entry bar. Furthermore, they can express their appreciation for the stream by sending virtual hearts, which can be achieved by clicking on the heart-shaped symbol in the interface's bottom right-hand corner. Adjacent to the heart symbol, viewers can also observe the real-time count of current viewers, providing a visual indicator of the stream's popularity. Replays of past streams are available and easily accessible with both the video and the messages, except for sending new messages.

Viewers' messages appear at the bottom of the screen for streamers and viewers. Messages typically appear on screen for a few seconds before being replaced by other incoming messages, similar to multi-party chats. This time frame and the streamer's actual streaming environment may constrain their ability to address the messages.

Moreover, as the predominant device used for Twitter live streaming, the smartphone's camera functionality offers various features streamers can utilize during broadcasting, such as zooming in, zooming out, and flipping the camera.

2.2 Data Collection and Analysis Method

The findings are based on the analysis of video recordings of live-streaming interactions. The study relied on conversation analysis to examine the streamer's verbal talk, nonverbal actions, and the viewers' text messages, as they unfold step by step in real-time. The study is based on the premise that social actions are multimodal accomplishments (e.g. Goodwin C., 1981, 2000).

2.2.1 Multimodal Conversation Analysis as Research Method

I employ multimodal conversation analysis (Hazel et al., 2014; Schegloff, 2007) as the research method. My analysis aims to identify, describe, and explain the orderly and recurrent practices adopted by participants during the tasting activities. I closely examine the multimodal resources that emerge within the technologized contextual configurations, including gestures, gaze, body postures, and the embodied manipulation of the phone and the App.

For the conversation analytic approach, the next-turn proof procedure by Sacks, Schegloff, and Jefferson (1974) is an essential guide. To gain insights into the question "why that now" (Schegloff & Sacks, 1973: 299), analysts can examine how a turn is related to the previous turn and how it is responded to. In other words, people display their understanding of the preceding turn in their next turn, which can then be used for analysis.

2.2.2 the Establish of a Data Corpus of Recorded Streams

For conversation analysis on videos, the creation of a data corpus of recorded videos is necessary. For my corpus, I recorded a total of 92 broadcasts from 34 different streamers, resulting in a total duration of over 20 hours (Figure 1). The recorded streams provided me with a) a sense of the settings in which they use the mobile live streaming applications, b) a display

of how these video and message-mediated interactions are going between streamers and viewers, and other simultaneously ongoing activities that are observable.

I employed two main considerations to establish criteria for streamer and stream selection. Firstly, streamers and viewers had to primarily use the English language to ensure the representativeness of the data, as English is the most commonly used language on Twitter. Secondly, I opted not to include streams shorter than 2 minutes, with the majority of selected streams falling within the 10 to 30-minute range. This criterion aimed to ensure that enough observable interactions were generated for a stream.

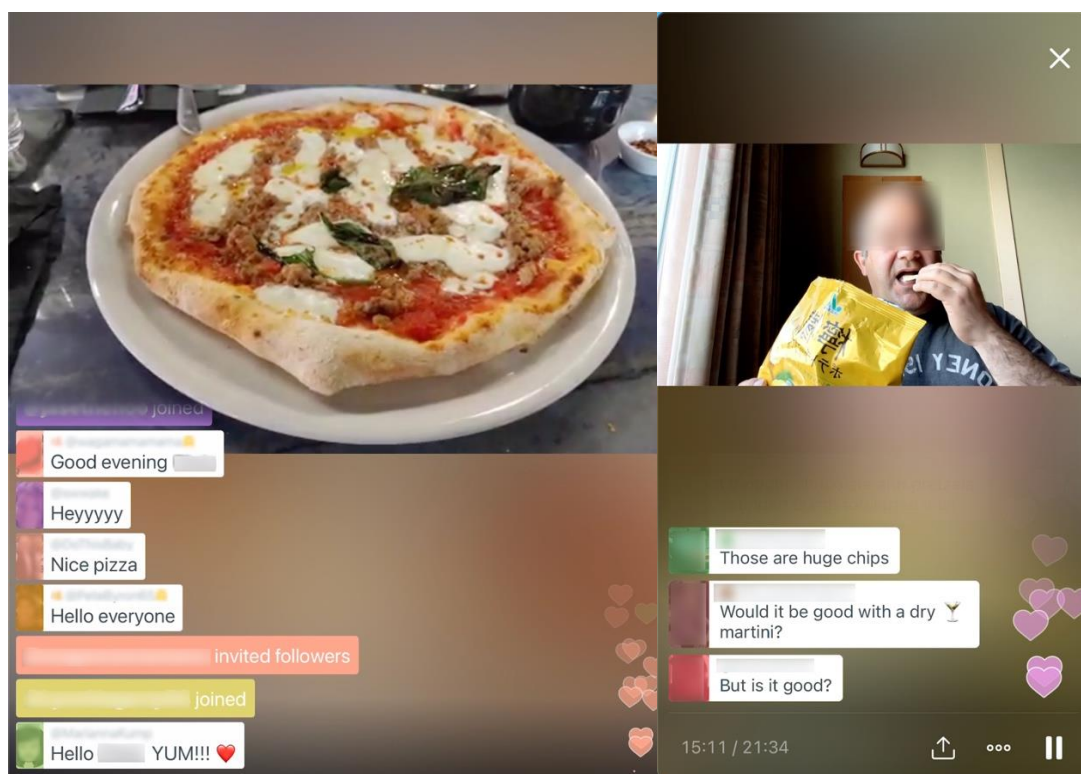


Figure 1. Food streams in my data corpus

2.2.3 Transcription and Data Ethics

Transcription of data is also a necessary element for conversation analysis. I did multimodal transcription to my data. Detailed information regarding my transcription conventions can be found in the Appendix. In all my transcripts, "S" represents the streamer, and bold texts denote viewers' messages with their pseudonymized account names. The position of a message in a transcript indicates the moment it first appears on the screen.

The streams analyzed in this study were publicly available online without any subscription requirements. Following regulations, recording and using data from public platforms is permitted as long as the privacy of individuals recorded is preserved. To maintain anonymity, I systematically anonymized and pseudonymized the data presented in the paper. Account names in the data transcriptions were changed; bystanders, car plates, and other potentially identifiable information were blurred in the screenshots used for analysis.

3 RESULTS

3.1 Putting the Current Tasting People's Face Onscreen: "Tasting Heads" Configuration as Participants' Achievement

The term "tasting heads" is derived from the concept of "talking heads." In video-mediated interaction, "talking heads" refers to the practice of displaying the face of the current speaker on-screen (Licoppe and Morel, 2012), particularly in mobile video calls. It is important to note that this does not imply that participants will always adhere to a talking heads arrangement in video communication. However, when they deviate from or inadequately display the face of the current speaker, they may be held accountable for their actions.

The concept of the talking heads maxim in a live-streamed tasting interaction operates similarly but with distinct characteristics. In my corpus, when engaging in a tasting activity during live streaming, participants often utilize phone cameras to display the faces of those currently tasting the food. This visual projection allows for the showcasing of facial expressions, which are considered accountable for both producing their own responses to the food and eliciting the audience's assessments. As a result, I refer to this configuration as the "tasting heads" configuration. According to the observations of my corpus, I will explore the "tasting heads" configuration with a focus on three main practices to achieve it in my corpus: putting the camera in a fixed position, flipping the camera, and handing the camera from one to another¹⁰.

3.1.1 Putting the camera in a fixed position

The practice of putting the camera in a fixed position is commonly observed in a prepared tasting, which often refers to food tasting shows (Choe, 2019). In such a stream, a streamer sets the stream's pre-figured activity as a food sampling event, intending to showcase their responses

¹⁰ Extracts in this section usually represent the tasting activities before the proper tasting. Their continuations or full versions will appear in the following sections for analysis with different focus if it is the case.

after tasting the food. They aim to provide assessments of the food they have sampled and elicit comments from the audience.

Extract 1 exemplifies a "tasting heads" configuration in a prepared single-streamer tasting activity, where the streamer temporally deviates from his face to respond to audience requests to show the food and positions the camera back to his face for the upcoming food-tasting experience.

Extract 1

1 **BAR:** You need the camera pointing down#

2 **LIG:** @SUM omg ewwy 🤢

3 S: That is, it is apple sauce, hehe, It's totally apple sauce

4 I need the camera po-

5 Ok, well, you want to see the consistency of it

6 There it is, right there#1.2



1.1



1.2

7 **KAN:** @MAR Hi Mark

8 S: Do you want me to point it \$down? I'll point it down.#1.3#1.4

9 **HID:** \$Lol 🤔



1.3



1.4

10 **BRU:** Dave please think about creating the series "The Wild & Weird. in Japan" with Dave!

11 **SUM:** @LIG 🤔🤔🤔

12 S: there it is, if you wanna see it#1.5



1.5

13 The Wild and Weird in Japan, OK, I could do that. Yeah
14 I could make it like \$(.) a a series in my channel

15 MIL: \$@DRO good thanks

16 S: I could do a series in my channel

17 So there we got these pieces of apple

18 An::d let me mix this in a little bit more

19 BRU: **Yes a series!**

20 S: Yeah, that that might be, that might work

21 Hey Mark how is it going,

22 We are trying,

23 DRO: **How much was this noodles in a box?**

24 S: Just-just to explain

25 Ah- it was like nothing, it was like (.) super cheap\$, yeah

26 BRU: \$Needs some
apple sauce?

27 S: Mark we \$are trying#1.6, (1.0) apple sauce,=



28 KAN: \$No source?

29 S: =apple \$pie noodles

30 MAR: \$@BAR hiya thanks for the broadcast 🙌🙌sorry 😊
I couldn't. join live

31 S: um, (2.5) no, yeah the sauce I put in their \$case,

32 LIG: \$I thought
trying weird stuff was already a series 😊?

33 S: So it was the apple sauce I put the sauce in

34 So why you can make series on YouTube.

35 So alright, let me try it

36 if you don't mind, I'm going to put it back on my face#1.7



37 **SUM: I dunno about apples & noodles...**

38 S: Yeah I might do that


39 **BAR: @MAR sure 🍷**

40 S: Apple noodles, okay, let's try, hehehe let's try

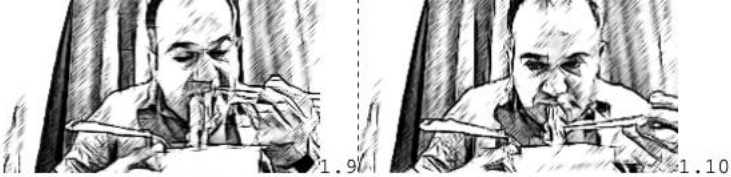
41 (4.0)

42 Okay

43 ((picks up and blows on the noodles))#1.8



44 ((withdraws gaze from screen and bites))#1.9#1.10



45 **BRU: Mmmmmmm 🍷🍷🍷🍷🍷🍷🍷🍷🍷🍷!**

In this extract, the streamer, who is from Canada but living in Japan, regularly produces a food sampling series featuring interesting Japanese food on his live streaming channel. In this particular instance, he is about to sample a type of Japanese apple pie flavor instant noodles. He strategically positions the camera before him, leaving him with a talking head onscreen (Image 1.1) and two free hands to introduce and taste the noodles.

Right before the extract, the streamer introduces what he will sample today. At Line 1, a viewer named BAR initiates a recruitment, which is a term used in conversation analysis to describe the project of getting others to do things (Kendrick and Drew, 2016; Kendrick, 2021; Choe, 2019). The viewer requests the streamer to point the camera down, as the audience cannot see the images of the noodles being introduced from the screen. The streamer notices the audience's message (Line 4) and treats it as a request. He responds with an embodied alignment, adjusting the camera to make the noodles visible on the screen (Image 1.2), and says "there is right there" (Line 6). The streamer further reinforces this action by repeating the message in an interrogative manner and self-answering it (Line 8). He continues to adjust the camera (Images 1.3, 1.4) until the noodles are clearly presented on the screen (Image 1.5) while stating "there it is, if you wanna see it" (Line 12). Notably, the streamer's camera motion aligns with the audience's request, indicating that the action is in response to the audience's demand, rather than something he decided to do on his own.

After resolving the camera issue, the streamer proceeds to respond to other messages and introduce the product. The term "alright" at Line 35 marks a temporary completion of the prior topic, and he follows it with "let me try it" in the next turn-constructive unit (TCU), signaling the beginning of his tasting sequence. Just before the tasting, he places his face back on the screen (Line 36, Image 1.7).

The viewer's request at Line 1 demonstrates the expectation that streamers should present the tastables that will be tasted. While the display of the tastables is also essential in the sequence, the core of the tasting activity lies in the demonstration of the "tasting head," signifying the streamers' facial expressions and responses during the tasting experience.

The next extract is also an example of "tasting heads" configuration achieved by the camera in a fixed position but in a multiparty setting.

Extract 2

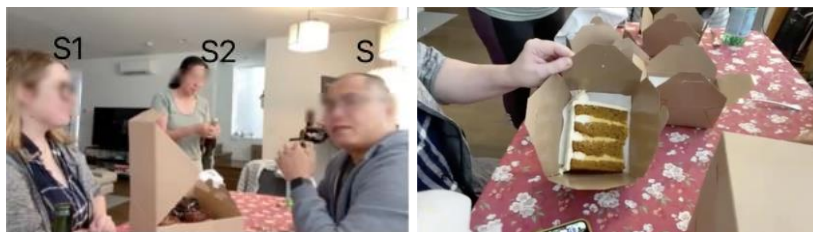


Figure 5 (left). S fixes the phone in a position before they begin the tasting
Figure 6 (right). They are showing the cake before the tasting

In this extract, "S" (the streamer) fixes the phone in a particular position before commencing the tasting activity. This ensures that the camera is set up optimally to capture the tasting process, as well as the responses of the streamer during the tasting. By positioning the phone in advance, the streamer can engage in the tasting process smoothly and efficiently while maintaining the "tasting heads" configuration for the audience to observe.

01 **Not: Who could have foreseen such a vibe on Tuesday. We did it!**

02 S: Alright, so: we are starting with the pumpkin,


03 (1.0)

04 S1: I'm most [excited for this one,

05 S: [Alright

06 And there's \$a slice of pumpkin, so you can see it?

*#2.1turns to screen and holds a bit of cake to screen----->



2.1

07 **Cur: \$As a longtime Cowboys fan, I think I have to swallow my pride and become an Eagles fan**

08 S1: ((raises a bit of cake to mouth and is ready to eat))


09 S2: pump[kin and nine pound butter

10 S: [Alright

11 ((looks at S1))

12 >[dididi<

*#2.2puts hand to stop S1 from putting the cake into mouth-->



2.2

13 S2: [holy crap, there's some kinda pumpkin cream in it.

14 S1: ((stops putting the cake into mouth and holds the spoon))

15 >yeah yeah yeah<, the cream in the middle is different

16 from (unintelligible word).

The main distinction between Extract 1 and Extract 2 is that the latter represents a multiparty stream and involves multiple streamers. As shown in Figure 5, the man on the right is the main streamer, referred to as "S" (the stream account owner). The individuals labeled "S1" and "S2" are respectively his friend and family. In this stream, they gather to try some cakes to celebrate Joe Biden's election.

All three participants remain seated in a relatively still position throughout the stream. The main streamer, "S," positions the camera at an angle that allows all three of them to be visible on-screen (Figure 5). This setup forms a circular arrangement known as an F-formation with the phone (Kendon, 1990: 209; Tong et al., 2016; Pathi et al., 2019; Barua et al., 2020). This F-formation ensures that all participants are included in the frame and contributes to a coherent and inclusive view of the tasting activity for the audience.

3.1.2 Flipping the Camera in an ongoing course of activity

Besides well-prepared tasting shows as in the two cases above, some tasting activities in live streams have been identified as occasioned tasting, embedded in a more global course of daily activities, such as cooking, traveling, selling, or buying food (Mondada, 2019a; Mondada, 2020b; Streeck, 2011). These activities normally happen in a mobile setting.

In these cases, the streamer may not have initially set up the camera in a fixed position of prime view. Instead, the streamer adjusts the camera to showcase their experience of tasting the food, creating the tasting head configuration. Flipping the camera is a practice observed to achieve the tasting heads configuration.

The following case is an example of this. In this extract, viewer ‘RUS’ is interacting with the streamer with a voice connection¹¹, and no text messages come up during the extract.

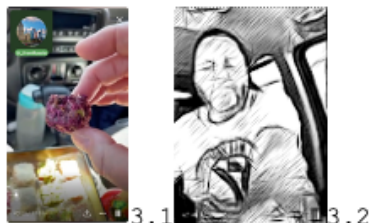
Extract 3

¹¹ Extract 3 and Extract 7 (extracted from the same stream) are exceptional cases in which the viewer ‘RUS’ interacts with the streamer by voice via a voice connection. So, in the transcripts of these two cases, this viewer’s talks are also in bold. This is rare in my corpus since, in live streams, most viewers interact with streamers by text messages, even though voice talk and even video-to-video talk appear to be a trend in some live streaming platforms due to the affordances of their functions.

1 S: *#3.1Wow, it's (.) fully, fully covered in this stuff
*pinching the delight in front of the cam----->

2 **RUS: Looks like (.) hibi-hibiscus* petals, ye[ah,**
----->*

3 S: ((flips the camera to talking head))#3.2



4 S: [Right

5 **RUS: probably**

6 S: Yeah

7 S: ((Showing the food))#3.3 but I think it's ro::se >petal<



8 **RUS: Rose?**

9 (2.0)

10 \$(S puts the snack into mouth#3.4)

11 \$(S bites part of it and withdraws gaze from scn #3.5)



12 **RUS: Rose should smell like rose. Does it smell like rose?**

In this stream, the streamer has just purchased a box of Turkish delights from the supermarket. Right before the extract, he opens the box and takes one of the delights with his fingers, displaying it to the camera (Image 3.1). The streamer then flips the screen to show his own face (Line 3, Image 3.2) before proceeding with the tasting, which starts from Line 10.

In some cases, the flipping action may occur after the actual tasting has taken place. Extract 4 serves as an example of this case. The streamer is a tour guide in New Zealand. She also broadcasts on Twitter from time to time to show the landscape of New Zealand to viewers around the world. In this stream, she is guiding viewers to see a small town and is ready to have

some lunch in a café just beside the road. Prior to this extract, the streamer initiated a sequence of showing and provided a detailed description of her lunch, which consists of a mussel fritter and a potato fritter (Figure 7). After this showing sequence, the streamer flips the screen to display her face and then proceeds with the tasting. This sequence illustrates how the streamer can switch the camera orientation between showing the food and showing their facial expressions during the live stream.

Extract 4



Figure 7. The mussel fritter (left) and potato fritter (right).

1 S: the bread, I don't need that, but this is mussel fritter inside, (.)
2 can see all the yummy bits of mussel in there?
3 So I'm having a little bit of onion and a few herbs, (.) and lots of
4 egg and a little bit of flour just to bind it all together #4.1



4.1

5 Oh, yum
6 ((laughs))
7 Yum yum haha
8 S: ((flips the camera to talking head))#4.2



4.2

9 S: And it tastes of mussels that's what you want when you get a mussel
10 fritter, you want it to taste of mussels. And then the potato fritter
11 is, hmm,#4.3
12 ((puts into mouth with laughter))#4.4



4.3



4.4

13 H/mn, look
*#4.5puts the bitten potato close to camera----->
14 Hm, really good, crunchy crunchy, (.) and soft inside
----->



4.5

15 Okay (.) that's awesome. ((laughs))

In this stream, the streamer starts by showing and introducing her lunch (Lines 1-2, Image 4.1), which consists of a mussel fritter and a potato fritter. As she continues to eat, her responses and assessments are audible through utterances like 'oh', 'yum', and laughter (Lines 5-7), all produced while the screen image still displays the food.

At Line 8, the streamer flips the screen to show her face while she continues to provide her assessments (Lines 9-10) with a delightful expression (Image 4.2). The streamer then picks up the potato fritter at Line 12, displaying it in front of the screen (Image 4.3), and takes a bite (Image 4.4). After the bite, she says 'hmm, look' at Line 13, and puts the bitten potato close to the camera (Image 4.5) while making additional assessments (Line 14). She appears to be

recruiting her audience to join in the tasting experience (Choe, 2019) by sharing the food visually and through her enthusiastic responses.

The practice of flipping the camera to achieve tasting head configuration reveals how an activity like tasting can be embedded in live streaming and is relevant and consequential for the organization and progressivity (Mondada, 2018a) of the streaming activities. It also displays how streamers manipulate the affordances of the digital device to achieve the action of showing their tasting experience.

3.1.3 Handing the camera from one to another in a multiple streamer setting

In some multiple-streamer tasting streams, the camera's position is not fixed but instead allows for mobility. Extract 5 provides an example of such a case. In this extract, three streamers are friends gathering in a teahouse in Hong Kong. S1, a local Hong Kong resident, invites the other two to savor Oolong tea (Figure 10). The main streamer (S) and S2 sit on one side of the table (Figure 8), while S1 sits on the opposite side, facing them (Figure 9). This arrangement forms a circular F-formation (Kendon, 1990: 209), where the three individuals are positioned in a way that fosters face-to-face interactions.

Extract 5



Figure 8 (left). S and S2 sit on one side of the table.

Figure 9 (right). S1 sits on the other side of the table.



Figure 10. They are served with three cups of oolong tea.

1 SOL: Quien me contesto lavan el te?
2 S2: ((takes a sip))#5.1
3 ((slightly nods to screen direction))#5.2
4 ((takes another sip))#5.3



5.1



5.2



5.3

5 S: How was it
6 HEA: @JOH hi John! 🍷
7 S2: ((slightly nods))
8 °uh-huh° (1.0) good
9 KAR: @SOL yo
10 S: °Okay°
11 S1: ((fetches the phone)) #5.4



5.4

12 S1: ((makes S present in the screen in talking head mode)) #5.5



5.5

13 S: let's see let's see let's see
14 S: ((lifts the cup with eyes on the cup))#5.6

15 ((takes a sip of the tea with eyes rolling up))#5.7
16 ((tastes with tea in mouth))#5.8#5.9



5.6



5.7



5.8



5.9

17 S1: Hi Heather, I'm with JL right now, =

18 S: =Umm=

19 S1: =Jon

*moves camera to S2#5.10-->

20 S1: ((moves screen back to S's face))#5.11



5.10



5.11

21 S: It's nice

*#5.12turns back to tea---->

22 S: ((takes another sip)) #5.13

23 ((tastes with eyes gazing nowhere))#5.14#5.15



5.12



5.13



5.14



5.15

24 S: I like it#5.16



5.16

25 S1: ((laughs))

26 You like it

27 **EMI:** @SOL kar. ina
 28 S: I like it
 29 I \$like it la/
 *turns to S2-→
 30 **HEA:** \$This is so cool!
 31 **ALL:** ((laugh))
 32 S1: So, um, (.) it's darker, define[tely
 33 S: [yeah, \$that's definitely darker
 34 **KAR:** \$Emily .!! Que bueno que
 lo lavaron asi ya sabre tomarlo

In this extract, S2 is sampling the tea with his cup while S, sitting beside S2, holds the phone and films S2 in a tasting head screen frame (Lines 2-4). S then asks for S2's assessments (Line 5), and S2 responds with an embodied action of nodding (Line 7) and an utterance "uh-huh good" (Line 8). S acknowledges the information with an "okay" (Line 10), indicating his readiness to taste in his turn on this occasion.

The ensuing camera motion and the talking heads intertwine and unfold as follows: S1 fetches the phone from S's hand and positions S on the screen in a talking head configuration. After handing over the phone, S begins his tasting (Lines 13-16, Images 5.6-5.9) and provides his verbal assessments (Lines 18, 21). He turns his head toward S2 (Line 18) while producing the assessments, seemingly seeking agreement. As S1 holds the phone, he can see the comments on the screen. He notices a greeting message (Line 6) and responds with a reciprocated 'hi,' mentioning that he is with JL (S) and Jon (S2) (Lines 17, 19). He moves the camera to S2 (Image 5.10) when mentioning S2's name (Line 19). However, the focus on S2's face is short-lived. S1 quickly moves the camera back to S's face (Line 20, Image 5.11) just before S says 'it's nice' (Line 21). S takes another sip (Line 22), and S1 records his tasting face and responses (Lines 22-24, Images 5.13-5.15).

To sum up, in section 3.1, the analysis reveals the strong orientation towards featuring the current tasting face on-screen before or after the taste to capture the tasting activities, particularly, the facial expressions and assessments that are going to be analyzed in the next sections.

3.2 The gaze and facial expressions

3.2.1 Withdrawing gaze from the screen when tasting

In the example of tasting in cheese shop encounters, Mondada (2009a; 2018a) describes how both seller and customer orient to an ‘individual interactional space’ in which the customer disengages the gaze (but does not disengage from the interaction through keeping a bodily orientation to the seller), which allows them to taste in silence to fully focus on the tasting experience.

This individual interactional space, characterized by gaze patterns, still exists and is found in my live streaming data, although the contexts and orientations have changed. We may call it a unique ‘digital individual space’ within the streamer and their viewers/screen wherein streamers withdraw their gaze from the screen when they begin to taste and gaze back at the screen after the tasting, particularly before assessments.

This individual space does not mean that the streamer goes into and immerses in a private space by themselves. It is an interactional achievement. To get to and establish this individual space, the streamer has to claim a silent tasting moment for the senses among the two seemingly conflicted and intertwined ongoing courses of actions: 1) monitoring, reading and responding to messages and 2) the activity of tasting the food by themselves. This orientation is more obvious for single-streamer, prepared tasting streams, and streams with more messages coming up. By doing so, the streamers orient to tasting as a situated activity that requires silence and concentration, and it is achieved by both streamers and viewers as a joint accomplishment. The next extract demonstrates how this is achieved by an ordered distribution of multimodal resources.

Extract 6



Figure 11. The streamer shows the fried chicken before this extract

- 1 **Tra:** Hey Dave□□□□□
beautiful day
- 2 S: It's very dark the screen#6.1 now suddenly
*#S's hand covers the screen-->

3 S: Hey: howzi going?
----->*



6.1

4 S: It is#6.2 a beautiful day >yeah#6.3 OK< let's#6.4 try
the
5 first one(.)#6.5 if th/is(.)broadcast \$fails for some
reason

6 **Kan:**

\$@ L Hi Laura



6.2



6.3



6.4



6.5

7 **DoT: Does Pablo's cheesecake have a good stall?**

8 S: =I will↑come back

9 **MsP: @nai hi!!!!**

10 S: .h so just be patient

11 **DoT: Food**

12 **DDe: Dave, Dr dee, Keizo**

13 S: Does Pablo's Cheesecake ha:ve(1.0) a gu:d↑

14 **Kan: @hat おはよう (Morning!)**

15 S: I can't read that. Good s::-(.)

16 **Oh** they don't hava-\$e heh they don't have a stall here=

17 **Bez: \$How are the infection rates there?
Is**

it going down or spiking up again?

18 S: =But I do have a Pablo's Cheese\$cake product, I am
19 gonna show you (0.5)

20 **DrD: \$@kan hi Keizo san**

21 S: Ah they are going down. Infection rates are going down
22 here, so they are getting better.h which is why this is
23 going on, if it wasn't ah-if it wasn't getting better I
24 can't guarantee \$you what will go on, so

25 **Kan: \$@MsP Hi□**

26 S: OK, let's try this#6.6

27 Ah, where is my screen

28 ah there we go#6.7 ((laugh)) the fried chicken,



6.6



6.7

29 I don't know what sauce is on here

30 So I am gonna (1.0) I am gonna sample with\$out really

31 knowing what sauce it is

32 **DrD:**

§@DEA hi dean

33 S: ((withdraw gaze from screen))#6.8



6.8

34 §(10.5)

35 S: §((puts into mouth))#6.9

36 §((bites and eyes up gazing nowhere))#6.10

37 §((bites a small part of the chicken))#6.11

38 §((chews)) #6.12



6.9



6.10



6.11



6.12

39 **Kan:** §@Hid こんにちは (how are you?)

40 S: §((gazes back at screen))#6.13

41 §((slightly nods))#6.14



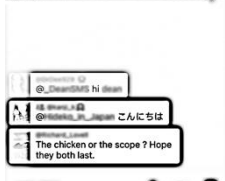
6.13



6.14

42 **Ric:** §The chicken or the scope? Hope they both last.

43 S: Oh yeah that's good#6.15



6.15

44 I don't know what sauces that is

45 ((looks up gazing nowhere))#6.16



6.16

46 (2.0)

47 S: It might be like just like- (.) a meat sauce of chicken
48 a meat sauce.

49 S: ((puts into mouth the whole remaining part))#6.17

50 ((chews while staring at the screen)) #6.18



51 S: The chi#6.16cken or the scope, ho#6.17pe they both
52 last((laugh)), yeah



The streamer is visiting a food festival in Japan and he is sampling a type of Hakata fried chicken to his audience. Before the sampling, the streamer introduces the site and the food (fig) with camera images. At Lines 4-5, he produces the tasting announcement ‘let’s try the first one’ the first time, prefaced by a change-of-state token ‘OK’. Right before this announcement, he is responding to a message at Line 1. He greets back (Line 3) and shows alignment to the viewer’s assessment of the weather (Line 4). But he appears to try to minimize the response by almost repeating ‘beautiful day’ with a simple agreement token ‘yeah’. He speeds up his talk when he utters ‘yeah’ and ‘OK’, going from one turn to the next with almost no silence in this turn transitional point. Besides, he has already diverted his gaze from the screen to the food when he has not finished responding to the message (Image 6.2). These verbal and embodied actions display his work in setting a boundary and finding timing for initiating the tasting.

However, he fails in going into the tasting and moves into the turns about the concern of the streaming condition (Lines 5, 8, 10). After these, he continues to respond to two messages that appear at Line 7 and Line 17 until he proposes the tasting announcement once again prefaced by the change-of-state token ‘OK’ at Line 26.

The streamer bears the pressure of responding to messages as much as possible, the normative orientation in live streaming interactions (Licoppe & Morel, 2018). He demonstrates strong orientations of reading and responding to messages in distinct and multimodal ways. For example, he expresses concerns about the streaming condition itself (Lines 5, 8, 10). He tries to read and monitor the screen (Image 6.1; Lines 27-28) and produces accounts for not seeing clearly (Lines 2 & 15). In case of missing messages, he keeps a pattern of intermittent gazes between the screen and things in hand (e.g. Images 6.3 & 6.4 at Line 4).

The announcement at Line 26 marks the boundary of the projected tasting with his previous interactions, such as responding to messages. After the announcement, he adjusts his camera to make the chicken present well on the screen (Lines 27-28, Images 6.6-6.7). The streamer then withdraws his gaze from the screen when he is finally going to sample the chicken, adopting a diverted posture from the screen (Line 33, Image 6.8), displaying a ‘not attending to the viewers’ configuration for the time being, and going into the ‘individual space.’ He puts the chicken into his mouth (Line 35; Image 6.9) and bites while lifting his eyes and gazing into the distance before him (Line 36; Image 6.10). He bites a small part of the chicken (Line 37, Image 6.11) and begins chewing (Line 38, Image 6.12); while chewing, his eye keeps gazing off the screen. When he gazes back at the screen (Line 40; Image 6.13), he first slightly nods (Line 41; Image 6.14) and then produces the first verbal assessment, ‘Oh yeah that’s good’ (Line 43).

The streamer’s gazing back at the screen marks the end of the tasting and the disengagement of the individual interactional space for the moment. However, the streamer displays his not attending to the messages when tasting by suspending responding to the message at Line 42. When he says ‘Oh yeah that’s good’ at Line 43 he gazes at the screen (Image 6.15). The viewers could treat the gazing as he is supposed to see the screen and the messages. But the streamer shifts the gaze away quickly and once again engages in the individual space at Line 45 (Image 6.16) when he tries to sample what the sauce is. He demonstrates the gaze withdrawal posture again, looking nowhere, focusing on the taste before he produces his descriptions of the sauce (Lines 47-48). He then gazes back to the screen and puts the remaining part of the chicken into his mouth (Line 49; Image 6.17). The activity here has transferred from tasting to eating. He chews overtly while staring closely at the screen (Line 50; Image 6.18). Only then does he begin reading (Images 6.16 & 6.17) and responding (Lines 51-52) to the already appeared message at Line 42. Here, it forms a contrasting comparison regarding the multimodal organization between tasting and eating—the streamer does not need to withdraw his gaze to an individual space to focus on the senses when eating.

Actually, it is normal that not all messages are treated, as is the case especially with messages less relevant to the streamer, such as in Lines 20, 25, 32, and 39, in which the viewers are interacting with each other by ‘@’ their account names. But the streamer’s suspension of responding to messages shows his coordination of multiactivities (Raymond & Lerner, 2014; Kamunen, 2019) and preservation of an embodied silent tasting moment and enhances the claim

that the individual interactional space is a normative practice in the tasting activity (Mondada, 2018a) even in the ecology of live streaming, although shaped by different orientations.

The same gazing away and gazing back patterns can also be seen in other extracts, for example, Line 44 and Images 1.9 & 1.10 in Extract 1, Line 11 and Image 3.5 in Extract 3, Line 47 and Image 9.1 in Extract 9.

From the analysis, I aim to show how this digital interactional space with an individual character is achieved and shaped by gaze patterns stepwise. It is also important to note that the streamer and viewers have different rights and entitlements in this unique space created by live streaming. The streamer has the right to engage in a sensory experience and taste food while the viewers have the right to monitor the streamer. Essentially, the streamer is being watched by default in the context of live streaming. This makes it a space of intersubjectivity where private personal sensing practices are provided to watch and be shared through the configuration of a tasting head.

3.2.2 Animated Facial Expression prior to verbal responses

The streamer's performative eating, such as the loud gustatory sounds and the exaggerated facial and verbal expressions, provides a spectacle that draws in viewers. (Anjani et al., 2020; Kim, 2020). My data reveals the presence of animated and exaggerated facial expressions during the tasting experience in live streaming. These facial expressions serve as embodied assessments and noticings of the taste experience. When streamers taste food during the live stream, their facial expressions become an embodied means of conveying their immediate sensory perceptions. The facial expression makes the tasting experience witnessable, observable, and accountable, enforcing its role as a subtle device for securing shared understanding and affiliation (Ruusuvaori and Peräkylä, 2009) and achieving intersubjectivity.

The achievement of the tasting head configuration through camera motion is crucial in demonstrating these expressions effectively. By flipping the camera to the face, the streamer allows the audience to witness their responses in real-time. Extract 7 is an example of this, and it is a continuation of Extract 3 (starting from Line 10).

Extract 7

10 S: \$((puts the snack into mouth#7.1))
11 \$((bites part of it and withdraws gaze from scn #7.2))



12 RUS: Rose#7.3 should smell#7.4 like ro#7.5#7.6se.#7.7 D[oes it smell
13 like rose?

14 S: [Oh my gosh,
15 hehe ((slightly taps the food several times))#7.8



16 S: ((sniffs the food#7.9))



17 S: Yeah, and it tastes like rose. It has-it has real rose
18 petals on it. It's like jelly, kind of.

After taking a bite of the snack (Images 7.1-7.3), the streamer demonstrates a surprising look with his eyes wide open to a climax and eyebrows abruptly raised and curved (Images 7.4-7.6). These exaggerated and performative expressions precede his verbal response cry "oh my gosh" at Line 14, constituting his embodied response to the tasting. In Lines 17-18, the streamer then proceeds to respond to the viewer's question 'Does it smell like rose' at Lines 12-13, after he sniffs the snack first (Line 16; Image 7.9).

In the next case, which is the continuation of Extract 5 ("W" in the transcript means the waitress who is just beside and serving them). In this multi-streamer setting, a similar performative expression appears on the screen after the tasting. The difference is that one onsite participant's question about the taste appears in the midst of the streamer's animated expressions after the tasting. We will see how the streamer coordinates his embodied responses and verbal responses.

Extract 8

1 S: ((takes a sip))#8.1



8.1

2 S: ((keeps tea in mouth and taste)) #8.2#8.3#8.4



8.2



8.3



8.4

3 S: ((gradually lifts eyes, gazing back to the screen))#8.5



8.5

4 S1: What [does it taste like

5 S: [Oh:

6 ((gazing at screen and still holds the cup)) #8.6



8.6

7 S: I'm spreading #8.7#8.8#8.9



8.7



8.8



8.9

8 ALL: ((laugh))#8.10=



8.10

9 W: =really?=
10 ALL: =((laughs))

In Extract 8, occurring in the same stream as Extract 5 but featuring a different type of tea, the streamer S tastes a flower tea instead of oolong tea. After the tasting, S gradually lifts his eyes with an odd and performative expression (Line 3; Image 8.5). At this time, S1 produces a question about the tasting of the tea (Line 4). S1's question projects a verbal answer about the taste. Instead of providing a direct description or assessment, S just produces a response cry 'oh' (Line 5). This responding with a minimal 'oh' responds to the question yet minimizes the verbal response. By doing so, S appears to give way to his embodied actions as he continues his 'performance'. Keeping that odd expression, S looks back at the screen (Line 6, Image 8.6),

demonstrating a lifting motion with his arms while holding the cup (Images 8.7, 8.8), and utters an even more exaggerated verbal response ‘I’m spreading’¹² (Line 7). His face also becomes animated with laughs (Image 8.9). These verbal and embodied responses evoke laughter (Line 8), displaying people’s understanding of them and that they treat the responses as funny conduct.

The finely tuned organization of the facial expressions and verbal responses reveals not only the importance of facial expressions in streamers’ tasting responses but also the order of the responses; that is, the facial expressions (and other embodied actions) normally occur prior to the verbal response.

Mondada (2018a) has claimed that the occurrence of animated facial expressions that project an assessment is one of the practices in the normative sequential organization of tasting widely observed in different settings. The finding also appears to apply to the ecology of live streaming through the analysis above. However, not all the facial expressions in my data are as performative and exaggerated as in the two cases above. Some of them are just animated embodied actions such as nods (see other extracts in this paper, e.g., Extract 6).

The animated facial expression adds a playful and engaging element to the watching of a tasting experience. It is not an exclusive phenomenon for tasting in live streaming, but the pattern of first gazing back to the screen and then displaying facial expressions (e.g., Lines 16 & 17 in Extract 6; Images 7.3 & 7.4 in Extract 7; Lines 3-7 in Extract 8) exhibits a unique practice in the setting of live streaming.

In summary, for section 3.2, we can see that the streamer’s gaze becomes a crucial device to frame to beginning and ending of the tasting itself. Right after the tasting, the streamer re-engages with the viewers in various ways even before speaking. These include shifting their gaze back to the screen, becoming more expressive with a more animated face and head, incorporating nods and facial expressions, and using these embodied movements to signal the resumption of the conversation and, particularly, the responses of the tasting, which is what I will turn to next.

3.3 Streamer’s Verbal Assessments and Viewers’s Text Responses

¹² With his action of lifting his arms as wings (Images 8.7, 8.8), S means that he is about to fly.

In food-trying videos, the taster's reactions are one of the most expected parts by viewers. Assessments, descriptions, noticings, questions, response cries, and embodied responses such as facial expressions analyzed above etc. are important outcomes of a tasting experience. Verbal and embodied resources from the streamer, and text messages from the audience are mobilized for the action formation of the responses after the tasting.

Assessments are essential productions of a tasting activity. In live-streamed tasting, streamers introduce the food, taste it, and assess it all by themselves. The streamer assumes the audience lacks access to the food. When making an assessment, a streamer claims knowledge of the subject being assessed (Goodwin and Goodwin, 1987). Their assessments are, at the same time, broadcasted and treated as news reporting by the audience. While the audience, who normally can't taste the food themselves, responds to (assess, notice, inquiry, describe, etc.) the streamer's assessments, or what they see on screen, a certain topic, the watching experience, through text messages. These practices result in a response structure of assessment reporting from the streamer and understanding displaying from the audience.

The audience's messages thus become a device where congruent understanding is made visible (Goodwin and Goodwin, 1987).

Based on observations from my data, I will illustrate this response structure in two distinct participation frameworks (Goodwin and Goodwin, 2004)—the single-streamer tasting and the multi-streamer tasting.

3.3.1 The response structure in a single-streamer tasting

Extract 9 is the continuation of Extract 1, in which the streamer, a Canadian living in Japan, is sampling an apple pie flavor Japanese instant noodles.

Extract 9

46 **MEK:** Worst thing I ever ate in my travels: bondegi. Boiled
silkworm cocoons in Seoul, Korea 🤢

47 S: ((gazes back to screen)) #9.1



9.1

48 S: Oh,

49 (7.0)

50 \$((keeps chewing))

51 ((puts a piece of apple into mouth)) #9.2



9.2

52 \$(5.0)

53 \$((keeps chewing)) #9.3



9.3

54 S: it really does \$taste like apple pie

55 **DRO:** \$Lmao

56 S: Because the noo\$dles are hot,

57 **MAR:** \$You're always having adventures with food



58 (2.0)

59 S: ah, you know the heat of the noodles

60 **PAU:** @MEK ewww

61 S: And I guess you know the \$noodles are also \$carbohydrate

62 **KAN:** \$@SUM Even in Japan, I've never
seen

63 **BRU:** \$Oh my Dave
your face says it all!

64 **HID:** Hahaha 😂

65 S: so you got the=

66 **CNM:** Love it

67 S: =you got the flavor of the apple \$sauce, =

68 **LIG:** \$oh on YouTube got ya
there's a category for that it falls in with muck bang

69 S: =which is, you know, heated by the noodles,

70 and then the carbon of the, (.) the carbon, you know the noodles are

71 carbon, (.) but they have a saltiness to them

Notably, DRO's question is the first message to receive a response from the streamer after the tasting, highlighting the relevance and importance of interrogative turns in driving subsequent actions. Additionally, MEK and PAU (Lines 46, 60, 81), KAN and SUM (Lines 62, 83) engage in topic discussions with each other using "@+name" cues, while LIG's comment is in response to a prior topic (Line 68). These messages demonstrate the characteristic "loosened relevance" and "disrupted turn adjacency" commonly seen in live streaming and other text-mediated communication (Herring, 1999, 2013).

3.3.2 The response structure in a multi-streamer tasting

In multi-streamer tasting streams, the response structure takes on a slightly different contextual configuration than in streams with only one streamer. In multi-streamer tasting, streamers are in a mediated setting (Meyrowitz, 2014), where streamers are physically present while the audience participates virtually. Streamers have access to the food. Different streamers demonstrate and compare their affiliating or disaffiliating assessments of the tasting experience. This assessment sharing is treated as news reporting by the audience, of whom the access to the food is assumed impossible. The audience then displays their understanding by responding to the stream through text messages. Extract 10 is such an example illustrating this, and it is the full version of Extract 2, in which the streamers are sampling newly bought cakes to celebrate the election of Joe Biden.

Extract 10

01 **Not: Who could have foreseen such a vibe on Tuesday. We did it!**

02 S: Alright, so: we are starting with the pumpkin,
03 (1.0)

04 S1: I'm most [excited for this one,

05 S: [Alright

06 *And there's \$a slice of pumpkin, so you can see it?

*#10.1turns to screen and holds a bit of cake to screen----->



10.1

07 **Cur: \$As a longtime Cowboys fan, I think I have to swallow my
pride and become an Eagles fan**

08 S1: ((raises a bit of cake to mouth and is ready to eat))

09 S2: pump[kin and nine pound butter

10 S: [Alright

11 ((looks at S1))

12 >*[dididi<

*#10.2puts hand to stop S1 putting the cake into mouth-->



10.2

13 S2: [holy crap, there's some kinda pumpkin cream in it.

14 S1: ((stops putting the cake into mouth and holds the spoon))

15 >yeah yeah yeah<, the cream in the middle is different
16 from (unintelligible word).

17 S: ((turns left and looks at screen)) #10.3

18 S2: I'm sensitive to that, (.) OK

19 S: Alright

20 S2: Ch[↑eers,

21 S1: [Ch↑eers,

22 S: [Ch↑eers,#10.4



10.3



10.4

23 S2: Yeah↑[:::

24 S1: [What a great day

25 S: What a great day

26 S2: ((puts cake into mth with eyes open))
27 S1: ((closes eyes and puts cake into mth))
28 S: ((closes eyes and puts cake into mth))#10.5
29 S1: ((chews with eyes closed))
30 S: ((chews with eyes closed))
31 S2: ((looks at S1 and chews)) #10.6



32 S: ((opens eyes and mutually gazes with S2))
33 S1: ((opens eyes and looks at S3))#10.7
34 S: ((mutually gazes with S1))#10.8
35 S: I mean it's very good
36 **Jor: Atlanta here! I love pretty close to Devindra, actually**
37 S1: I like this very much
38 S: ((turns to and looks at screen))#10.9



10.7



10.8



10.9

39 S2: You don't love pumpkin in general=

40 S: =Yeah=

41 S2: =Pumpkin \$spicies

42 Jor: \$*live

43 S: This is a good, very good [traditional

44 S1: [It tastes a lot like 'cherry cake'.

45 S: [It tastes a lot-it tastes aa lot like cherry cake.

46 S2: [It does.

47 ALL: ((looks at screen))#10.10



10.10

48 S: >But it's very good<=

49 S1: =Yeah=

50 S2: =[yeah

51 S: [>It's very good<.=

52 S2: =>It's [very good.<

53 S: [Okay

54 S1: Okay.

55 S: Alright, move on.

56 ((drags another box))

57 S2: *It's extremely [good.

*looks to screen---->

In this extract, the three streamers engage in a simultaneous cake tasting, deviating from the turn-taking approach seen in the tea-tasting stream above. At Line 12, when S1 is ready to taste the cake, S intervenes by raising his left hand to suspend her from proceeding, indicating to wait for S2 to join in the tasting. Three reasons underlie this phenomenon. Firstly, the fixed camera position captures all their faces simultaneously, creating a cohesive visual presentation. Secondly, allowing anyone to taste the cake first could disrupt the structure of the subsequent assessment production. Since the production of assessments to the recipient relies on the assumption that access to a referent is the precondition for assessing it, tasting together would shape the ensuing sequence into knowledge-sharing assessments of the food with agreement or disagreement comparing, rather than the one who tastes first offering a news reporting type of assessments of the food. Another significant aspect of tasting together is the production of

responses as a collective display. The ceremonial cheer-up right before the taste at Lines 20-22 (Image 10.4) accentuates the sense of togetherness in the cake-tasting activity.

After the joint tasting (Lines 26-31), a moment of prompting for a response emerges as the streamers open their eyes and exchange shifting mutual gazes (Lines 32-34; Images 10.7-10.8). These practices go beyond the projection of intuitive assessment and guide the tasters to search for the proper assessment and description of it. S utters the first assessment designed with an 'I mean' prefaced positive assessment 'it's very good' (Line 35). The use of 'I mean' seems to mitigate the vulnerability of a first assessment, which runs of the risk of being downgraded by others (Mondada, 2009). The first assessment is quickly followed by a second assessment from S1 that exhibits alignment (Line 37). Interestingly, S2's response takes the form of a question (Line 39), indicating a degree of question and inquiry into S's assessment. The subsequent exchange of comments (Lines 43-46) leads to a consensus, and they quickly reach a shared assessment of the cake as 'very good' (Lines 48-53), a practice effectively concluding the episode of cake tasting.

Despite the streamers' limited direct interaction with the audience, they demonstrate an orientation toward attending to the audience. Before the tasting, S holds up a spoonful of cake to the screen (Line 06; Image 10.1), accomplishing two actions: showcasing the cake to the audience and inviting them to virtually partake in the experience (Choe, 2020), fostering a sense of shared engagement. Throughout the unfolding sequence, the main streamer S turns his head to the screen at different points (Line 17, Image 10.3; Line 39, Image 10.9; Line 47, Image 10.10), a body torque that keeps two orientations, highlighting the contrast between live-streamed and face-to-face tasting activities. The configuration of attending to the audience underscores the interactive nature of the live-streamed tasting event, despite the physical distance between the streamers and their virtual viewers.

In sum, this section focuses on the outcomes of the tasting experience from the two participating parties, the streamers and the viewers. A preliminary response structure is proposed based on the participation frameworks.

4 Discussion and Conclusion

As video-mediated interaction has become a new normal (Due and Licoppe, 2021) in the current society, screens have become deeply intertwined with people's routine life, such as eating. As an individual but also publicly witnessable activity (Mondada, 2019a), eating has been exposed to various screens. It is, therefore, crucial to unravel the eating practices in the digital landscape.

This paper tries to combine the study of tasting and the study of video-mediated interaction from the perspective of multimodal conversation analysis. The work addresses the role of technology in promoting remote dining experiences through the consumption of interactive and multisensory video streams. It explores the intersubjective sharing of sensorial experience in digital contexts. It aims to demonstrate how sensory practices are interactively and intersubjectively organized in live streaming. I have shown practices and patterns in the live streaming tasting activities in three sub-sections.

First, the "tasting heads" configuration emerges as a significant phenomenon, where streamers strategically position their faces on-screen to show their immediate responses. This configuration enhances the live-streaming experience by allowing viewers to witness and participate in the tasting process, creating a sense of togetherness even in remote settings. Second, the observable gaze withdrawal before tasting and subsequent gazing back at the screen for assessments indicate the streamers' conscious efforts to engage with the virtual audience. Notably, facial expressions during the tasting are shown to be performative and even exaggerated, serving as embodied responses to the tasting. These expressions add a visual and emotive dimension to the live-streamed tasting, intensifying the interactive experience for both streamers and viewers. Third, the study unveils the unique response structure in two participation frameworks. Streamers provide initial assessments as reporting, and the audience responds with messages, displaying their understanding of the tasting experience. In contrast, in multiparty broadcasts, streamers' assessments become a means of knowledge-sharing about the food between themselves and, at the same time, news reporting for the audience.

This study contributes to the broader fields of conversation analysis, multimodal interaction, and digital communication. The exploration enriches our understanding of the ever-changing ways in which humans connect and interact in the digital age. While this study offers valuable insights into the emerging practices of live-streamed tasting, further exploration and refinement are warranted. For example, the description and distinction of 'single-streamer stream' vs. 'multiple-streamer stream,' and 'prepared tasting' vs. "occasioned tasting" in my analysis could

be refined according to the participation frameworks of the streams, as well as the filmed nature of the settings (Broth et al., 2014) in future studies. Future research could also explore deeper into the variations and nuances of tasting configurations and their design implications for audience engagement.

5 APPENDICES. TRANSCRIPT CONVENTIONS

This paper's transcription conventions mainly adopt Gail Jefferson's (2004) conventions for talk and Lorenza Mondada's (2018) conventions for embodied actions. Besides, some particular conventions for live streaming have referred to Christian Licoppe and Julien Morel's work (2018) and Le Song and Christian Licoppe's work (2023). The main convention symbols involved in this article are as follows:

\ and / represent the falling and rising intonation, respectively;

↑ indicates a sudden rise in intonation;

(.) means inconspicuous short pauses (usually less than 0.2 seconds);

.hh means inhale or laughter;

§ indicates the moment or the period when the audience's messages come up, or when an action happens;

*--->

---->* indicates the continuation of the action;

indicates where the screenshot appears;

£ £ means that the content within the symbol is said when laughing;

() words in parentheses are unintelligible;

[indicates the overlapping of speech;

= indicates the connection of two turns;

(()) indicates non-verbal conduct.

REFERENCES

- Adjacency pairs, preference organization, and assessments. (2013). An Introduction to Interaction. <https://doi.org/10.5040/9781350284821.ch-006>
- Ahmed, S., Ting, D. H., Sharif, T., & Abedin, M. Z. (2023). Implication of 'camera eats first' construct: Unraveling the potentials of digital images in social media on information sharing. *Electronic Commerce Research*. <https://doi.org/10.1007/s10660-023-09737-x>
- Andersen, T., Byrne, D. V., & Wang, Q. J. (2021). How digital food affects our analog lives: The impact of food photography on Healthy Eating Behavior. *Frontiers in Psychology*, 12. <https://doi.org/10.3389/fpsyg.2021.634261>

- Anjani, L., Mok, T., Tang, A., Oehlberg, L., & Goh, W. B. (2020). Why do people watch others eat food? An empirical study on the motivations and practices of mukbang viewers. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/3313831.3376567>
- Avgustis, I., & Oloff, F. (2023). Mundane Food Photography. *Social Interaction. Video-Based Studies of Human Sociality*, 6(2). <https://doi.org/10.7146/si.v6i2.125292>
- Broth, M., Laurier, E., & Mondada, L. (Eds.). (2014). *Studies of video practices: Video at work*. Routledge.
- Barua, H. B., Pramanick, P., Sarkar, C., & Mg, T. H. (2020). Let me join you! real-time F-formation recognition by a socially aware robot. 2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN). <https://doi.org/10.1109/ro-man47096.2020.9223469>
- Bruno, A. L., & Chung, S. (2017). mökpang: Pay me and I'll show you how much I can eat for your pleasure. *Journal of Japanese and Korean Cinema*, 9(2), 155–171. <https://doi.org/10.1080/17564905.2017.1368150>
- Casperson, S. L., Sieling, J., Moon, J., Johnson, L., Roemmich, J. N., & Whigham, L. (2015). A mobile phone food record app to digitally capture dietary intake for adolescents in a free-living environment: Usability study. *JMIR mHealth and uHealth*, 3(1). <https://doi.org/10.2196/mhealth.3324>
- Choe, H. (2019). Eating together multimodally: Collaborative eating in mukbang, a Korean livestream of eating. *Language in Society*, 48(2), 171–208. <https://doi.org/10.1017/s0047404518001355>
- Choe, H. (2020). Talking the cat: Footing lamination in a Korean livestream of cats mukbang. *Journal of Pragmatics*, 160, 60–79. <https://doi.org/10.1016/j.pragma.2020.02.009>
- Due, B. L., & Licoppe, C. (2021). Video-mediated interaction (VMI). *Social Interaction. Video-Based Studies of Human Sociality*, 3(3). <https://doi.org/10.7146/si.v3i3.123836>
- Fischler, C. (2011). Commensality, society and culture. *Social Science Information*, 50(3-4), 528–548. <https://doi.org/10.1177/0539018411413963>
- Florez Marulanda, J. F. (2022). Student satisfaction pilot experience with synchronous classroom live streaming styles during the covid-19 pandemic. *IEEE Revista Iberoamericana De Tecnologias Del Aprendizaje*, 17(3), 301–306. <https://doi.org/10.1109/rita.2022.3191285>
- Goffman, E. (1966). *Behavior in public places: Notes on the Social Organization of gatherings*. Free Press.

- Goodwin, C. (2019). *Co-operative action*. Cambridge University Press.
- Goodwin, C., & Goodwin, M. H. (1987). Concurrent operations on talk. *IPrA Papers in Pragmatics*, 1–54. <https://doi.org/10.1075/iprapip.1.1.01goo>
- González Temer, Verónica. (2017). A Multimodal Analysis Of Assessment Sequences In Chilean Spanish Interaction. 10.13140/RG.2.2.15563.03364.
- Goodwin, C., and Goodwin, M.H. (2004) ‘Participation’, in A. Duranti (ed.) *A Companion to Linguistic Anthropology*, pp. 222–243. Oxford: Basil Blackwell.
- Gambetti, A., & Han, Q. (2022). Camera eats first: Exploring food aesthetics portrayed on social media using Deep Learning. *International Journal of Contemporary Hospitality Management*, 34(9), 3300–3331. <https://doi.org/10.1108/ijchm-09-2021-1206>
- Herring, S. C. (2013). Relevance in computer-mediated conversation. In S. C. Herring, D. Stein, & T. Virtanen (Eds.), *Handbook of pragmatics of computer-mediated communication* (pp. 245-268). Berlin: Mouton. Prepublication version: <http://ella.slis.indiana.edu/~herring/relevance.pdf>
- Herring, S. C. (1999). Interactional coherence in CMC. *Journal of Computer-Mediated Communication*, 4 (4). <http://jcmc.indiana.edu/vol4/issue4/herring.html>
- Højlund, S. (2015). Taste as a social sense: Rethinking taste as a cultural activity. *Flavour*, 4(1). <https://doi.org/10.1186/2044-7248-4-6>
- Hazel, S., Mortensen, K., & Rasmussen, G. (2014). Introduction: A body of resources – CA studies of Social Conduct. *Journal of Pragmatics*, 65, 1–9. <https://doi.org/10.1016/j.pragma.2013.10.007>
- Ibrahim, Y. (2015). Food Porn and the invitation to gaze. *International Journal of E-Politics*, 6(3), 1–12. <https://doi.org/10.4018/ijep.2015070101>
- Jefferson, G. (2004). Glossary of transcript symbols with an introduction. *Conversation Analysis*, 13–31. <https://doi.org/10.1075/pbns.125.02jef>
- Jodén, H., & Strandell, J. (2021). Building viewer engagement through interaction rituals on Twitch.tv. *Information, Communication & Society*, 25(13), 1969–1986. <https://doi.org/10.1080/1369118x.2021.1913211>
- Kamunen, A. (2019). How to disengage: Suspension, body torque, and Repair. *Research on Language and Social Interaction*, 52(4), 406–426. <https://doi.org/10.1080/08351813.2019.1657287>
- Kendon, A. (1990). *Conducting interaction: Patterns of behavior in focused encounters*. Cambridge University Press.

- Kendrick, K. H., & Drew, P. (2016). Recruitment: Offers, requests, and the organization of assistance in interaction. *Research on Language and Social Interaction*, 49(1), 1–19. <https://doi.org/10.1080/08351813.2016.1126436>
- Kendrick, K. H. (2021). The ‘other’ side of recruitment: Methods of assistance in social interaction. *Journal of Pragmatics*, 178, 68–82. <https://doi.org/10.1016/j.pragma.2021.02.015>
- Kim, Y. (2020). Eating as a transgression: Multisensorial performativity in the carnal videos of mukbang (eating shows). *International Journal of Cultural Studies*, 24(1), 107–122. <https://doi.org/10.1177/1367877920903435>
- Kim, J., Choi, S., Kim, H., & An, S. (2021). Binge drinking and obesity-related eating: The moderating roles of the eating broadcast viewing experience among Korean adults. *International Journal of Environmental Research and Public Health*, 18(15), 8066. <https://doi.org/10.3390/ijerph18158066>
- Kircaburun, K., Yurdagül, C., Kuss, D., Emirtekin, E., & Griffiths, M. D. (2020). Problematic mukbang watching and its relationship to disordered eating and internet addiction: A pilot study among emerging adult mukbang watchers. *International Journal of Mental Health and Addiction*, 19(6), 2160–2169. <https://doi.org/10.1007/s11469-020-00309-w>
- Kircaburun, K., Harris, A., Calado, F., & Griffiths, M. D. (2020). The psychology of Mukbang watching: A scoping review of the academic and non-academic literature. *International Journal of Mental Health and Addiction*, 19(4), 1190–1213. <https://doi.org/10.1007/s11469-019-00211-0>
- Kang, E., Lee, J., Kim, K. H., & Yun, Y. H. (2020). The popularity of eating broadcast: Content analysis of “Mukbang” YouTube videos, media coverage, and the health impact of “Mukbang” on public. *Health Informatics Journal*, 26(3), 2237–2248. <https://doi.org/10.1177/1460458220901360>
- Laurier, E. (2008). Drinking up endings: Conversational Resources of the café. *Language & Communication*, 28(2), 165–181. <https://doi.org/10.1016/j.langcom.2008.01.011>
- Lemke, M., & Schifferstein, H. N. J. (2021). The use of ICT devices as part of the solo eating experience. *Appetite*, 165, 105297. <https://doi.org/10.1016/j.appet.2021.105297>
- Licoppe, C., & Morel, J. (2012). Video-in-interaction: “Talking heads” and the Multimodal Organization of Mobile and Skype Video calls. *Research on Language & Social Interaction*, 45(4), 399–429. <https://doi.org/10.1080/08351813.2012.724996>

- Licoppe, C., & Morel, J. (2018). Visuality, text and talk, and the systematic organization of interaction in Periscope live video streams. *Discourse Studies*, 20(5), 637–665. <https://doi.org/10.1177/1461445618760606>
- Lu, Z., Annett, M., Fan, M., & Wigdor, D. (2019). "I feel it is my responsibility to stream". *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/3290605.3300459>
- Latzer, Y., Spivak-Lavi, Z., & Katz, R. (2015). Disordered eating and media exposure among adolescent girls: The role of parental involvement and sense of empowerment. *International Journal of Adolescence and Youth*, 20(3), 375–391. <https://doi.org/10.1080/02673843.2015.1014925>
- Lee, D., & Wan, C. (2023). The impact of mukbang live streaming commerce on consumers' overconsumption behavior. *Journal of Interactive Marketing*, 58(2–3), 198–221. <https://doi.org/10.1177/10949968231156104>
- Leer, J., & Krogager, S. G. S. (2021). *Research methods in Digital Food Studies*. Routledge.
- Micallef, D., Schivinski, B., Brennan, L., Parker, L., & Jackson, M. (2022). 'what are you eating?' can online game microcelebrities influence what emerging adults eat? *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4233222>
- Morris, A. M., & Katzman, D. K. (2003). The impact of the media on eating disorders in children and adolescents. *Paediatrics & Child Health*, 8(5), 287–289. <https://doi.org/10.1093/pch/8.5.287>
- Melville, N., Fairchild, R., & Evans, E. W. (2023). Food Safety Communication in YouTube video recipes. *British Food Journal*, 125(7), 2628–2643. <https://doi.org/10.1108/bfj-02-2022-0143>
- Manoogian, E. N., Wei-Shatzel, J., & Panda, S. (2022). Assessing temporal eating pattern in free living humans through the mycircadianclock app. *International Journal of Obesity*, 46(4), 696–706. <https://doi.org/10.1038/s41366-021-01038-3>
- Mcdonnell, E. M. (2016). Food porn: The conspicuous consumption of food in the age of digital reproduction. *Food, Media and Contemporary Culture*, 239–265. https://doi.org/10.1057/9781137463234_14
- Mejova, Y., Abbar, S., & Haddadi, H. (2021). Fetishizing food in Digital age: #foodporn around the world. *Proceedings of the International AAAI Conference on Web and Social Media*, 10(1), 250–258. <https://doi.org/10.1609/icwsm.v10i1.14710>

- M. Hoey, E. (2018). Drinking for speaking: The Multimodal Organization of drinking in conversation. *Social Interaction. Video-Based Studies of Human Sociality*, 1(1). <https://doi.org/10.7146/si.v1i1.105498>
- Meyrowitz, J. (2014). *No sense of place: The impact of electronic media on social behavior*. Oxford University Press.
- Majlesi, A. R., Ekström, A., & Hydén, L.-C. (2022). Assessments in assisted eating activities. *Communication and Medicine*, 17(2), 134–149. <https://doi.org/10.1558/cam.18648>
- Margutti, P., Tainio, L., Drew, P., & Traverso, V. (2018). Invitations and responses across different languages: Observations on the feasibility and relevance of a cross-linguistic comparative perspective on the study of actions. *Journal of Pragmatics*, 125, 52–61. <https://doi.org/10.1016/j.pragma.2017.12.010>
- Mondada, L. (2009). The embodied and negotiated production of assessments in instructed actions. *Research on Language and Social Interaction*, 42(4), 329–361. <https://doi.org/10.1080/08351810903296473>
- Mondada, L. (2018a). The multimodal Interactional Organization of tasting: Practices of tasting cheese in Gourmet Shops. *Discourse Studies*, 20(6), 743–769. <https://doi.org/10.1177/1461445618793439>
- Mondada, L. (2018b). Multiple temporalities of language and body in interaction: Challenges for transcribing multimodality. *Research on Language and Social Interaction*, 51(1), 85–106. <https://doi.org/10.1080/08351813.2018.1413878>
- Mondada, L. (2019a). Contemporary issues in conversation analysis: Embodiment and materiality, multimodality and multisensoriality in social interaction. *Journal of Pragmatics*, 145, 47–62. <https://doi.org/10.1016/j.pragma.2019.01.016>
- Mondada, L. (2019b). Rethinking bodies and objects in social interaction: A multimodal and multisensorial approach to tasting. *Discussing New Materialism*, 109–134. https://doi.org/10.1007/978-3-658-22300-7_6
- Mondada, L. (2020a). Audible sniffs: Smelling-in-interaction. *Research on Language and Social Interaction*, 53(1), 140–163. <https://doi.org/10.1080/08351813.2020.1716592>
- Mondada, L. (2020b). Orchestrating multi-sensoriality in tasting sessions: Sensing Bodies, normativity, and language. *Symbolic Interaction*, 44(1), 63–86. <https://doi.org/10.1002/symb.472>
- Mondada, L. (2021). Language and the sensing body: How sensoriality permeates syntax in interaction. *Frontiers in Communication*, 6. <https://doi.org/10.3389/fcomm.2021.664430>

- Mondada, L. (2023). Tasting vs. eating: The Methodic and situated differentiation of embodied multisensorial activities in social interaction. *Complexity of Interaction*, 29–64. https://doi.org/10.1007/978-3-031-30727-0_2
- Nielsen, F. (2019, May 14). How to live stream an event on Twitter Live: 14 steps to follow. *Billetter Blog*. <https://billetter.co.uk/blog/how-to-live-stream-event-twitter/>
- Niewiadomski, R., Ceccaldi, E., Huisman, G., Volpe, G., & Mancini, M. (2019). Computational commensality: From theories to computational models for social food preparation and consumption in HCI. *Frontiers in Robotics and AI*, 6. <https://doi.org/10.3389/frobt.2019.00119>
- Ochs, E., Pontecorvo, C., & Fasulo, A. (1996). Socializing taste. *Ethnos*, 61(1-2), 7–46. <https://doi.org/10.1080/00141844.1996.9981526>
- Ogden, R. (2021). Swallowing in conversation. *Frontiers in Communication*, 6. <https://doi.org/10.3389/fcomm.2021.657190>
- Pathi, S. K., Kristoffersson, A., Kiselev, A., & Loutfi, A. (2019). F-formations for social interaction in simulation using virtual agents and Mobile Robotic Telepresence Systems. *Multimodal Technologies and Interaction*, 3(4), 69. <https://doi.org/10.3390/mti3040069>
- Pomerantz, A. (1985). Agreeing and disagreeing with assessments: Some features of preferred/dispreferred turn shapes. *Structures of Social Action*, 57–101. <https://doi.org/10.1017/cbo9780511665868.008>
- Pomerantz, A.M., & Mandelbaum, J. (2016). Advocating healthy eating through interpersonal interaction. Talk presented at the 102nd convention of the National Communications Association. 10-13 November 2016. Philadelphia, USA.
- Panicker, A., Basu, K., & Chung, C.-F. (2020). Changing roles and contexts: Symbolic interactionism in the sharing of food and eating practices between remote, intergenerational family members. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW1), 1–19. <https://doi.org/10.1145/3392848>
- Pereira, B., Sung, B., & Lee, S. (2019). I like watching other people eat: A cross-cultural analysis of the antecedents of attitudes towards Mukbang. *Australasian Marketing Journal*, 27(2), 78–90. <https://doi.org/10.1016/j.ausmj.2019.03.001>
- Qu, L. (2021). Waste on the tip of the tongue: Social Eating Livestreams (chibo) in the age of Chinese affluence. *Asiascape: Digital Asia*, 8(1–2), 43–69. <https://doi.org/10.1163/22142312-12340131>
- Raymond, G., & Lerner, G. H. (2014). A body and its involvements. *Multiactivity in Social Interaction*, 227–246. <https://doi.org/10.1075/z.187.08ray>

- Ruusuvuori, J., & Peräkylä, A. (2009). Facial and verbal expressions in assessing stories and topics. *Research on Language & Social Interaction*, 42(4), 377–394. <https://doi.org/10.1080/08351810903296499>
- Rüdiger, Sofia. 2022. "Intimate Consumptions: YouTube Eating Shows and the Performance of Informality." *Internet Pragmatics* 5(1): 115-142. <https://doi.org/10.1075/ip.00070.rud>
- Rüdiger, Sofia. 2021. "Digital Food Talk: Blurring Immediacy and Distance in YouTube Eating Shows." *Anglistik* 32(2): 111-130. <https://doi.org/10.33675/ANGL/2021/2/9> (Open Access)
- Rüdiger, Sofia. 2020. "Dinner for One: The Use of Language in Eating Shows on YouTube." In: Rüdiger, Sofia and Susanne Mühleisen, eds. *Talking about Food - The Social and the Global in Eating Communities*. Amsterdam/Philadelphia: John Benjamins. 145-165. <https://doi.org/10.1075/impact.47.08rud>
- Stephenson, B. (2021, December 12). Your beginner’s guide to live streaming on Twitter. *Lifewire*. <https://www.lifewire.com/twitter-live-stream-4685835>
- Strand, M., & Gustafsson, S. A. (2020). Mukbang and disordered eating: A netnographic analysis of online eating broadcasts. *Culture, Medicine, and Psychiatry*, 44(4), 586–609. <https://doi.org/10.1007/s11013-020-09674-6>
- Song, L., & Licoppe, C. (2023). Closing live video streams: A sequential analysis. *Discourse, Context & Media*, 53, 100698. <https://doi.org/10.1016/j.dcm.2023.100698>
- Schegloff, E. A. and Sacks, H. 1973. Opening Up Closings. *Semiotica* 8: 289–327.
- Schegloff, E. A. (2007). *Sequence Organization in interaction: A Primer in conversation analysis*. Cambridge University Press.
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50(4), 696. <https://doi.org/10.2307/412243>
- Sneijder, P., & te Molder, H. F. M. (2006). Disputing taste: Food pleasure as an achievement in interaction. *Appetite*, 46(1), 107–116. <https://doi.org/10.1016/j.appet.2005.10.002>
- Sobal, J., & Nelson, M. K. (2003). Commensal eating patterns: A community study. *Appetite*, 41(2), 181–190. [https://doi.org/10.1016/s0195-6663\(03\)00078-3](https://doi.org/10.1016/s0195-6663(03)00078-3)
- Sobal, J. (2000). “Sociability and the meal: facilitation, commensality, and interaction” in *Dimensions of the meal: The science, culture, business, and art of eating*. ed. H. Meiselman (Gaithersburg, MD: Aspen), 119–133.

- Spence, C., Mancini, M., & Huisman, G. (2019). Digital commensality: Eating and drinking in the company of Technology. *Frontiers in Psychology*, 10. <https://doi.org/10.3389/fpsyg.2019.02252>
- Streeck, J. (2011). The changing meanings of things: Found objects and inscriptions in social interaction. *Embodied interaction. language and body in the material world*, 67-78.
- Temer, V. G., & Ogden, R. (2021). Non-convergent boundaries and action ascription in multimodal interaction. *Open Linguistics*, 7(1), 685–706. <https://doi.org/10.1515/opli-2020-0170>
- Tidy, J. (2020, August 20). Mukbang: Why is China clamping down on eating influencers? BBC News. Retrieved November 22, 2022, from <https://www.bbc.co.uk/news/technology-53840167>
- Tong, L., Serna, A., Pageaud, S., George, S., & Tabard, A. (2016). It's not how you stand, it's how you move. *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services*. <https://doi.org/10.1145/2935334.2935343>
- Wiggins, S. (2002). Talking with your mouth full: Gustatory mmms and the embodiment of pleasure. *Research on Language & Social Interaction*, 35(3), 311–336. https://doi.org/10.1207/s15327973rlsi3503_3
- Wiggins, S., & Keevallik, L. (2020). Enacting gustatory pleasure on behalf of another: The multimodal coordination of infant tasting practices. *Symbolic Interaction*, 44(1), 87–111. <https://doi.org/10.1002/symb.527>
- Xu, J., Qu, L., & Zhang, G. (2022). Governing social eating (chibo) influencers: Policies, approach and politics of influencer governance in China. *Policy & Internet*, 14(3), 525–540. <https://doi.org/10.1002/poi3.318>
- Yong, J. Y. Y., Tong, E. M. W., & Liu, J. C. J. (2020). When the camera eats first: Exploring how meal-time cell phone photography affects eating behaviours. *Appetite*, 154, 104787. <https://doi.org/10.1016/j.appet.2020.104787>

CHAPTER 6

CLOSING LIVE VIDEO STREAMS: A SEQUENTIAL ANALYSIS

Abstract

Research has argued that “ordinary conversation” and its organization are foundational to “institutional” talk (Drew & Heritage, 1992), and that institutional forms can be characterized as constraints on such a sequential organization. Such an argument can be extended to technology-mediated interaction, in which participants may orient jointly to “standard” conversational sequences and the technology’s interactional “affordances” to achieve interactional ends. We discuss here how such an interactional goal (closing a technology-mediated encounter) may be achieved in live video streaming. We argue that while participants can be seen to do that in a way that orients to the organization of closings in ordinary conversation, they do so in a way that is sensitive to the affordances of live video streams. We find that: a) Most of the streams involve closing sequences, thus framing the live stream as a social encounter for which closing is relevant; b) A partial and relaxed orientation to talking heads configuration in live stream closings; c) A four-part closing sequence; d) Two different topic development in closing the live streams—a topic message comes from the audience and some “mentionable noticeables” initiated by the streamer.

Key words

Live streaming; Closing sequence; Multimodal conversation analysis; Twitter Live; Mediated interaction

1. Introduction: Closing a mediated interaction

By “mediated interaction” we mean the interactions in which the participants involved orient to the affordances that the mediating technologies offer (Arminen, 2005: 215; Arminen et al., 2016). Functionalities of the mediating technologies produce alterations to the sequential organization of communication. Even minor alterations like short delays of messages may impose systematic variations on the closing sections, and restrain the achievement of

intersubjectivity in action formation and action ascription (Arminen, 2005: 215; Deppermann, 2022).

As a form of video-mediated interaction (Due & Licoppe, 2021), live streaming has been investigated by researchers significantly. Live streaming in this research refers to the highly asymmetric form of interaction, in which a streamer appears live on video, with an active audience being able to respond to them through publicly available text messages (mostly) or voice (only a few in our data corpus) (Licoppe & Morel, 2018)¹⁴.

Live streaming itself is not new, but the use of mobile live streaming applications (APPS) is not long in human media history. Since around 2015, the burgeoning of live streaming APPs has brought live streaming into the limelight (Garst, 2015). Today, live streaming has taken a crucial role in the world's social media landscape, and it has almost become both an entertainment industry and an e-commerce industry, particularly in some East Asian countries (Cunningham, Craig & Lv, 2019; Si, 2021). Existing research on live streaming has focused on different streaming platforms and miscellaneous topics (Licoppe & Morel, 2018; Rein & Venturini, 2018; Lu et al., 2018; Recktenwald, 2018; Choe, 2019, 2020; Anjani et al., 2020; Johnson, 2022; Marulanda, 2022; Zelenkauskaitė & Loring-Albright, 2022). Among them, some researchers have focused on how participation is multimodally managed by viewers and streamers in live streams, by using interactional linguistics along with insights from Ethnomethodology and Conversation Analysis (EMCA) (see Maynard & Clayman, 1995 and Llewellyn & Hindmarsh, 2010) studies. For example, Licoppe and Morel (2018) propose the “read aloud and respond” pattern in Periscope's live streaming interactions, and Recktenwald (2018) contributes to a transcription and analysis of game live streaming on Twitch and shows how dialogical and monological moves coordinate and relate to the unfolding gameplay in spoken-to-written communication. Choe (2019) examines how sociable eating is jointly and multimodally achieved in the Korean live eating shows “mukbang” and finds that users coordinate their talk, written comments, and embodied actions, thus establishing engagement and social agency. In another mukbang research, by inspecting eating shows of cats, Choe (2020)

¹⁴ In this article, we do not distinguish between “live streaming” and “live broadcasting,” “streamer” and “broadcaster,” “audience” and “viewer.” Therefore, they have equal meaning when appearing. Also, we use “message” or “comment” to refer to the live texts sent by the viewers, which may be preferably called “danmaku” or other names in different societies and communities.

explores the notion of “footing lamination”, and claims that there is a complexity of footing dynamics where different footings interact with each other in cats mukbang interactions.

Previous research from the perspective of EMCA has involved studies in manifold mediated interactions and corresponding media platforms. Such studies have focused on the practices of closing the interaction, and these include the traditional study of closing sequences using telephone conversations as data, for example the very initial and seminal papers by Schegloff and Sacks (1973), Jefferson (1973) and Button (1987, 1990), in which they proposed a canonical and archetype closing sequence in telephone conversations which consists of four turns and is organized in two adjacency pairs, and topics such as sequence-closing device (Park, 2010), turn-initial high-grade assessments (Antaki, 2002), politeness strategies (Coppock, 2005) in telephone closings. Besides the telephone, various practices of closing a mediated interaction have emerged with the development of mediating technologies. Studies have also covered closings in talkback radio conversations (Döpke et al., 1994), the activity of an airline cockpit (Nevile, 2005), instant messages (Raclaw, 2008), video-mediated tele-homecare for older adults (Ilomäki & Ruusuvuori, 2020), online meetings (Oittinen, 2021) and disengagements between human and humanoid robots (Licoppe & Rollet, 2020), etc. All the topics have been investigated to show why closing matters (Raymon & Zimmerman, 2016), and how closing is achieved step-by-step.

It can be seen that not only are there few EMCA studies on live broadcasting, but even fewer studies focus on how a live broadcast is closed moment-by-moment, that is, its closing sequence. Our paper thus studies the closing sequences of this unique and widely known mediated interaction in the method of multimodal conversation analysis and aims to fill a gap in the research of closing a mediated interaction. It also contributes to the research in live streaming in an EMCA vision.

We aim to demonstrate that: a) Most of the streams involve closing sequences, thus framing streams as social encounters for which closing is relevant (Section 3.1 and Section 3.2); b) A partial and relaxed orientation to talking heads configuration (Licoppe & Morel, 2012)¹⁵ exists

¹⁵ In video-mediated communication (VMC) (in which both parties interact ‘live’ through audio and video), the talking heads configuration involves the user's face is fully visible and oriented towards the camera. It is treated as a default mode: when there is nothing relevant to look at, it is expected that people should appear in a talking

in streams' closings sequences (Section 3.3); c) There appears a unique four-part closing organization of live streaming that relates to the “disrupted turn adjacency” feature (Herring, 1999) (Section 3.4); d) Closing sequences are sensitive to topic talks either triggered by a topic message comes from the audience or elicited by the streamer due to her topicalizing some ‘noticeable’ (Section 3.5).

2. Data and methodology

2.1 Twitter Live

The streaming platform we selected is Twitter Live (connected to Periscope before 31st March 2021; see Team Periscope, 2020). On Twitter Live, the broadcasters stream live videos while hosting a conversation with an active audience that can send messages.

Twitter Live streams are broadly similar to the video streaming platforms that have been rapidly expanding in the last few years, such as YouTube Live, Facebook Live, and TikTok Live, though each platform may provide streamers and their audiences with distinctive affordances. We have focused on streams documenting everyday life and experiences, though other uses of live streaming have developed, for instance, for telemarketing or for documenting cultural and political events. We also collected a sample of similar streams on other platforms to confirm that the sequential phenomena we describe here are more generally relevant to live streaming.

The following characters of Twitter Live are important for our research:

- Past streams are archived with live comments, which makes it easy for researchers to find relevant video material for analysis.
- Most streams had a reasonable length (rarely more than an hour).
- Streamers are distributed worldwide, and contents are quite diversified.
- The layout of the interface is clear, concise, and intuitive.

Anyone can watch streams or their replays on Twitter by just opening the site, and you can browse randomly to get a stream or search for a streamer that you like (Fig 1). You do not have to subscribe to a streamer’s account to watch their streams, although you can find their streams more easily by subscribing next time. Then, choose one stream you like and click in, and you will get an interface with live video and viewers’ live comments superimposed on it (Fig 2).

head format (Licoppe & Morel, 2012).

You can type and send your own comments in the bottom entry bar and send out virtual hearts by clicking the heart shape symbol at the bottom right. At the very bottom right corner, you can also see the number of current viewers. To launch your own stream, just put in the title you want for your stream on the top input field “What’s happening?” and then click the red “Go LIVE” button at the bottom (Fig 3).

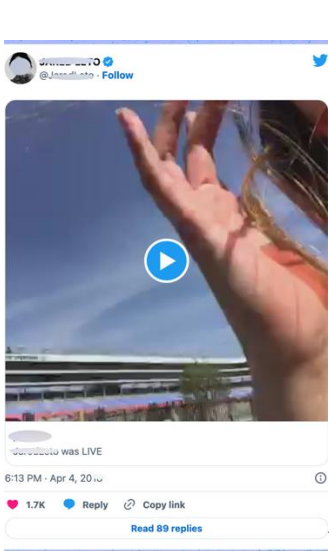


Fig 1



Fig 2



Fig 3

There are multiple contexts and devices (smartphone, drone, GoPro, etc.) that can be used to live stream. While it is often used in a stationary state and a fixed location (like home), one also finds much outdoor usage in mobile settings (e.g., driving¹⁶, walking, see Fig 4, Fig 5, Fig 6). Smartphones are the most used type of device in our data corpus.



Fig 4



Fig 5



Fig 6

Features and functions vary by device. For mobile phone live streaming, almost all the functions of the phone’s camera could be used in the broadcasting, for example, zooming in, zooming out, and flipping the camera. Flipping the camera is a widespread action in streamer’s live broadcasting, and it means flipping the phone’s rear and front camera by pressing the flipping button on the screen (or by just double-click the screen or by using gestures for some devices

¹⁶ In this case many drivers as we have observed have affixed the mobile phones in front of them (Fig 4), which makes them able to read and respond, at least to some extent. Depending on countries such a practice may be illegal or not (in France for instance, the law focuses on drivers with mobile phones IN HAND), but it may be deemed to be rather unsafe in all cases.

and platforms) (Fig 7), to achieve the image switches between the streamer themselves and the environment. We will mention it again in the analysis part of the paper.



Fig 7

Regarding messages from the audience, they appear on the bottom of the screen (both to streamers and all viewers), where they stay for a few seconds before disappearing. The ability of the streamer to treat them in this lapse of time may be constrained by their involvement in the setting. Moreover, one must distinguish between ‘hot streams’ with a large audience and flux of messages and ‘cold streams’, where both proceed at a leisurely pace (Licoppe & Morel, 2018). In the former case, the on-screen messages are quickly chased by new incoming ones, so they stay visible for a very short time, making it difficult for the streamer to treat most of them.

2.2 Data Corpus and Method

We have included 92 broadcasts from 34 streamers in our corpus. They were recorded during the year 2020 to 2021, with a total duration of over 20 hours. Given that we needed to study the closing sequences of the streams, we recorded most of our data from the archived past streams, which were already in their complete length. We recorded more than 92 streams and 34 streamers initially but ended up with these numbers when we believed they contained all types of streams and had reached the expected total length under our selection criteria.

There are two main criteria for our selection of streamers and streams: a) The main language used by the streamers and viewers be English; b) We decided not to allow for streams shorter than 2 minutes, and the majority of our streams are in the 10’-30’ range. The first criterion is to ensure the representativeness of our data since the most used language on Twitter is English; the second is to ensure that there are enough viewers to elicit interesting episodes of interactivity

(more than five active users is usually enough) for every single video (though there may be “quiet” episodes with no or limited interaction along the stream). Apart from these criteria, streamers and live broadcasts are selected at random.

The streams themselves are public and available online without a subscription. In line with French regulations, recording and using data on public platforms is allowed if everybody's privacy is preserved. We have done systematic anonymization and pseudonymization work on the data presented in the paper. We have changed every account name in the data transcriptions and blurred bystanders, car plates, etc., for each screenshot we use, so nobody should be recognizable in our exploitation of the data. The recordings provide a kind of access that is close to that of the streamer or their audience (though it may happen that some messages do not appear or that the lag between video and messages is different). Messages appear and stay on the screen before being chased by other incoming messages (a bit like what occurs in multi-party chat configurations). There is a limited time to read them for all parties, so reading and responding may require alertness and skillful behavior, particularly in mobile settings.

While closings are almost always recorded, it is not so with beginnings that may have yet to be recorded or may also have been cut out of the archive for technical reasons. Moreover, when recorded, openings often occur where there is no audience yet (the audience only catches up gradually as the stream goes on). For these reasons, we have chosen to focus on closing sequences as a phenomenon of interest. With streams involving a significant audience, closings involve collaboration between it and the streamer. Moreover, as a “leave-taking” practice, closing live streams should somehow reflect the kind of encounter that has just occurred (Goffman, 1966).

Regarding closings, our dataset yields the following distribution (Table 1):

Type of closings	Number
With closing sequences	61
Without closing sequences	4
Closing sections not recorded	27

Table 1: Number of closing sequences in our corpus

We use multimodal conversation analysis (Hazel et al., 2014; see also Schegloff, 2007) as our research method (see the Appendix for a detailed introduction of our transcription conventions.

In all our transcripts, “S” stands for the streamer; texts in bold represent the audience’s live messages with their account names prefaced, and we have pseudonymized all account names. The position of a message in a transcript denotes the moment when the message appears on the screen.). Based on the next-turn-proof procedure (Schegloff & Sacks, 1973), our analysis identifies, describes, and accounts for the orderly and recurrent practices that the streamers and the audience use in the closing sequences of the live-streamed interactions. We will closely investigate the multimodal resources emerging in the specific technologized contextual configurations, such as gestures, gaze, body postures, and embodied manipulations of media technologies (Mondada, 2019).

3. Analysis

3.1 Live video streams without a recognizable closing sequence

We will first consider the few streams without recognizable closings, that is, without recognizable closing-relevant verbal utterances or recognizable closing-relevant embodied actions (e.g., hand waving) from the streamers. We have listed and described them in Table 2.

Streams without Closing Sequences	Brief Introduction	Talking Head or Not	
CASE 1	From the same streamer	The US Capitol broadcasting (Fig 8)	“Non-talking head” mode all through the stream
CASE 2		Donald Trump rally broadcasting	“Non-talking head” mode all through the stream
CASE 3	The Chinese New Year broadcasting in New York’s Chinatown	“Non-talking head” mode when disconnecting	
CASE 4	Breakfast broadcasting	“Non-talking head” mode all through the stream	

Table 2: Basic information of streams without closing sequences

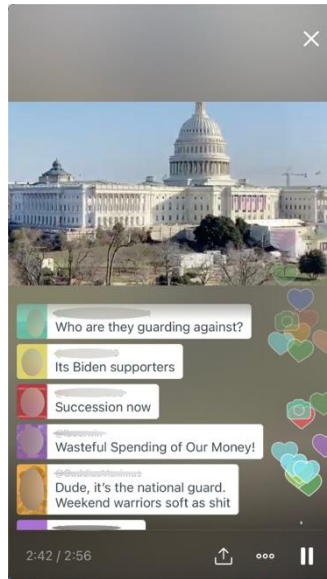


Fig 8

In the first two cases, the streamer is silent and invisible from beginning to end, even though he¹⁷ receives a constant flux of messages until the final disconnection. The streamer either moves his phone around to give viewers a panoramic view of the Capitol events (Extract 1), or he walks around a Trump rally to show the event (Extract 2). While the majority of incoming messages are political opinions, passionate slogan-like statements, and responses to other messages (which can be ignored by the streamer), there are still a few messages directly addressed to the streamer, which he disattends, thus showing that he is not oriented towards his stream being an interactive encounter between him and his audience. Both streams thus appear to belong to a particular and quickly developing sub-genre for live streams, the proximal broadcasting of political and often conflictual collective events as they unfold on the ground (Gerbaudo, 2017; Martini, 2018; Kavada & Trere, 2019; Yang & Kang, 2021). Because of their orientation towards live broadcasting, they seem to make closing irrelevant, first because the broadcasted event has no clear ending, and second because the streamer seems to erase himself visually and interactionally (at least in part) from the stream by not interacting directly with the active audience. For such broadcast-oriented political streams, mere disconnection appears as a relevant option.

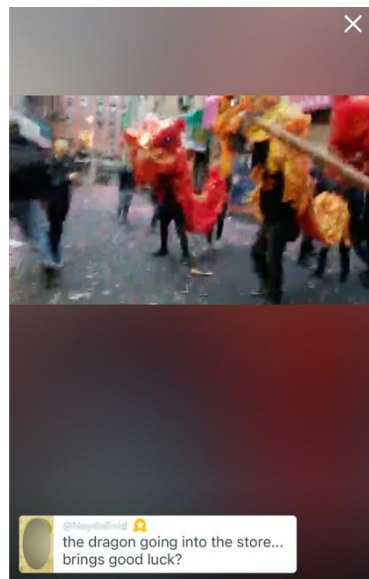
Also, both of these streams are “hot streams”, with a lot of incoming messages, which makes it very salient and striking the way streamers ignore them. Our next two cases are “cold streams”, with very few incoming messages. Though one might argue that the lack of audience might account for a lack of closing sequence, this does not seem to be the case for we have observed

¹⁷ From this streamer’s other streams, we know it is a male streamer.

a fair amount of cold streams with closings. The broadcasting genre seems to be more relevant and both types share a lot of common features regarding the genre.

Case 3 (Extract 1) documents a “Dragon and Lion Dance” performance in New York’s Chinatown during the lunar Chinese New Year. The streamer rarely shows himself and almost always holds the phone to show the performance. He talks mostly to describe and comment on the performance itself.

Extract 1



1 S: #1.1To get a little street alleyway here in Chinatown, extremely
2 historic.
3 (16.0)
4 **NEY: the dragon going into the store..brings good luck?**
5 \$ (25.0)
6 S: \$ ((stops walking to film)) #1.2
7 S: ((disconnects))



1.1

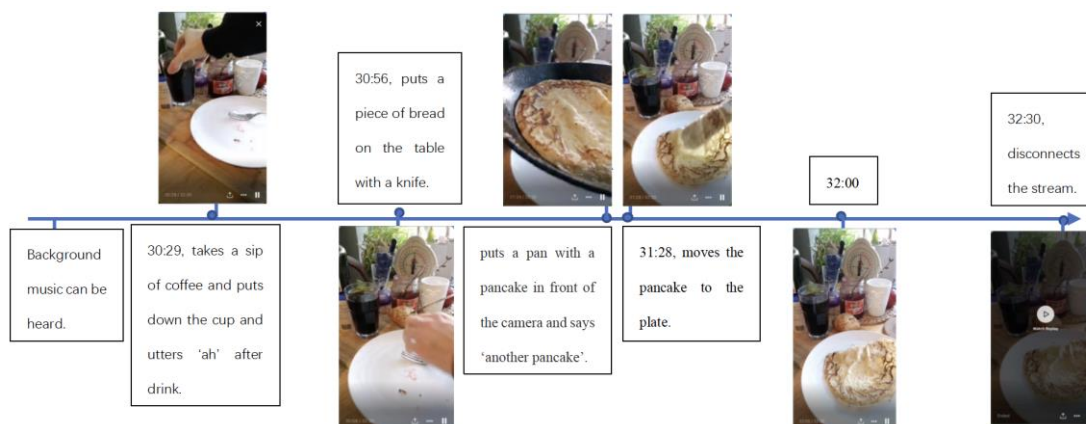


1.2

Case 4 (Extract 2) is somewhat different beyond superficial similarities. It involves the streamer documenting a mundane activity (the preparation of breakfast). He is not visible, speaks very

little, and has no incoming messages from a potential audience. The lack of incoming messages makes it appear that no encounter with a relevant audience has occurred. In turn, this lack of engagement might account for the lack of closings, in line with Goffman’s point that the way leave-taking is done can be seen as an insider’s comment on the type of encounter that has just occurred (Goffman, 1966).

Extract 2



In brief, the few streams without closing sequences belong to broadcast-oriented “TV news-like” streams or streams with little audience engagement. Both cases suggest that, conversely, closings are expected when engagement with the audience occurs (as in most cases in our corpus). Is this the case? How are closings actually done? These are the two questions we will turn to now.

3.2 Closing live streams. A typical example.

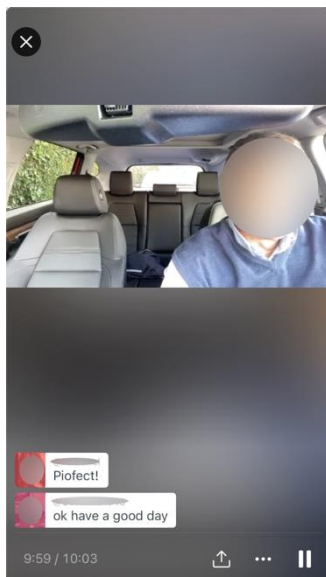
The seminal analysis of closings in phone conversations highlights three constitutive features of closing sequences (Schegloff & Sacks, 1973), that is, they are expected to be initiated by the caller, they are done as adjacent pairs in pre-closings and closings, etc., and that they are a collaborative achievement (with the called party being able to align or disalign with projected closings). We will keep these in mind as a guideline to flesh out the particulars of closing sequences in streams.

Interestingly, vlogs, which may be considered as asynchronous precursors to live streams, display an orientation to closings through ritualized closing actions (Frobenius, 2014). We observe here that most streams involve a recognizable closing sequence. In all cases, it is

initiated by the streamer. This has to do with the asymmetries in the forms of engagement and the affordances of the live stream platform. First, it is the streamers' choice to initiate and broadcast a stream, so in a sense, they own the stream, and, therefore the rights to end it. Second, the platform's functionalities make it so that only the streamers have the technical resources to initiate and end the streaming in practice.

In this section, we will analyze a typical example of closing a stream to highlight the specific concerns streamers and their audiences orient to at the end of streams. In this example, the streamer lives in Switzerland and regularly streams his various trips to show and document his experience for a significant audience (there is a steady stream of incoming messages).

Extract 3



- 1 S: ((pulls over #3.1))
- 2 ((flips the camera))
- 3 ((waves hand #3.2))
- 4 Hello.
- 5 So guys I'll see you later (.) on my way back,
- 6 have a good morning or have a good day.
- 7 Take care.
- 8 ((waves hand #3.3))
- 9 Bye.
- 10 Rur: Piofect!
- 11 Coc: ok have a good day
- 12 S: Ciao.
- 13 ((Disconnects))



We want to highlight three particular features here. First, the sequence occurs after the streamer shows and comments on his environment while driving, with the mobile device affixed to the top of the windscreen on his right. He pulls over as he arrives at his destination (which projects some future involvements) and flips back the camera towards himself to produce a “talking head” configuration. This normative orientation implies that users expect openings and closings to be done in a talking head, for these sequences is usually to open or close, not to show. Live streaming is a different proposition from traditional video-mediated communication, for only one party (the streamer) is interacting live through video. However, something akin to the talking heads orientation in video-mediated communication may still be relevant, with the camera flip as an index. The camera flip at Line 2 may be understood retrospectively as signaling a change in participation status, making it relevant to show himself again as a talking head and as a preliminary move to his leave-taking. This is reinforced by the greeting and waving (Lines 3-4), which indexes his visual re-appearance¹⁸ as if a new encounter within the encounter is beginning. This suggests the talking head format for closing might be relevant in live streaming interaction as a kind of change in visual “footing” (Goffman, 1981) preliminary to the closing of the encounter, but to what extent?

Second, there are closing turns, initially done unilaterally. The streamer provides pre-closing and closing turns and gestures in succession (Lines 5-9). Then two messages come up quickly. The first one, “Piofect!”¹⁹ (Line 10), is potentially responding to the streamer’s future arrangement talk (Line 5), and the second one, later, “ok have a good day” is aligning with his closing project and producing a return closing. Instead of disconnecting at this point, the streamer displays a sensitivity to the messages, looking at the screen and responding to them with a new closing (“Ciao,” Line 12), differently designed from his previous ones, before eventually disconnecting. This example shows that the way the closings are done may be sensitive to the likelihood, actual occurrence, content, and timing of incoming messages, which cannot be anticipated. When the messages are oriented to the closings, this raises the issues of ignoring them (and producing them unilaterally) or treating them (and producing the closing sequence collaboratively). When they deal with different topics (either responding to earlier ones or introducing new ones), with a timing often “off” concerning the streamers’ turns, the

¹⁸ It is worth noticing that the organization of ‘symmetric’ video-mediated is highly sensitive to such visual (re-)appearances, with the “appearance-greeting” sequence (Licoppe, 2017), and that, as pointed by one peer reviewer, greetings following the visual reappearance of the vlogger may also feature in vlogs.

¹⁹ This word can be understood as a variant form for “perfect”, whether the misspelling is intentional or not.

same concerns occur regarding responses. If treated, this might launch a new topic during the closing. We will now examine these different aspects in more detail, starting with camera flips, and the talking heads orientation in streams.

3.3 A partial orientation to talking heads in live stream closings

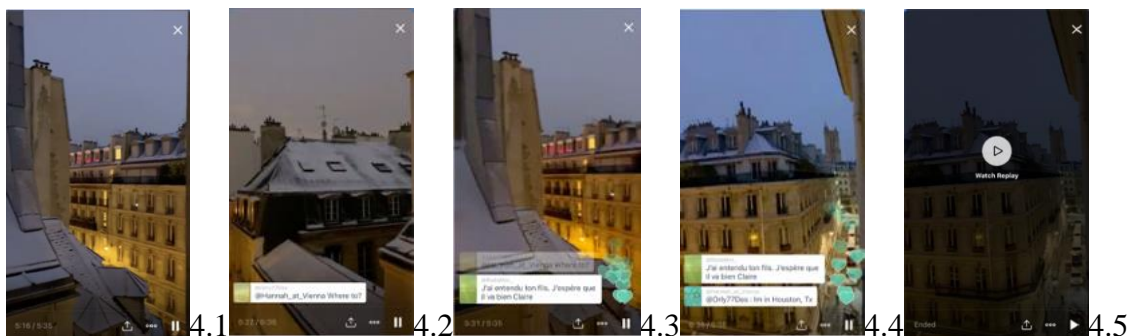
The camera flip can be seen as a preliminary and closing-implicative move in the closings, with the streamers appearing as talking heads as they engage in the closing sequence. The situation would be reminiscent of the “talking heads” normative within that interpretation. Moreover, in an earlier corpus of our data from Periscope, a similar orientation had been observed for live streaming (Licoppe & Morel, 2018). However, streams in that corpus mainly were about people using live streaming to engage and talk with their audience. For this corpus here, the situation seems somewhat different. A more critical fraction of the streams are oriented towards documenting activities unfolding in time around the streamers.

Suppose we consider streams with closing sequences in our corpus. Roughly half of them (30 out of 61) end with the streamer being visible (mostly talking to the camera somehow). Extract 3 above is a typical example of that. Alternately, half of those end with the streamer keeping invisible, usually going on showing what they have been showing, as in Extract 4 below.

Extract 4²⁰

²⁰ For clarification, as evidence by the “@ + tagname” in the message which makes a message intelligible as sent to another viewer, ORL and HAN are interacting with one another, ‘sideways’ (Line 4 and 10), but in line with a topic elicited by the streamer earlier. The commenters can interact with each other by “@” their account names, as is shown in Lines 4 and 10 between ORL and HAN. By this way they give their comments a specific recipient. A more common case is that of RED interacting with the streamer (Line 7) and the streamer responding to her (Line 8).

1 S: #4.1 I remember when I was working out in the middle of nowhere they
 2 would stop.
 3 Uhm, § we don't have to go to work, because buses couldn't run.
 4 ORL: § @HAN Where to? #4.2
 5 S: Cuz they didn't have, um,
 6 they couldn't § drive on road.
 7 RED: § J'ai entendu ton fils. J'espere que il va bien
 Claire
 8 S: Right. #4.3 Merci.
 9 Alright, I gotta go, bye everyone.
 10 HAN: @ORL Im in Houston, Tx #4.4
 11 S: ((Disconnects)) #4.5



In the first type of cases, where streamers make themselves visible during closings, streamers have usually made themselves visible to engage with the audience several times before. In the second type, where they close without becoming visible, they have never or very rarely appeared. Therefore, there is a correlation between how closings are visually done and how the whole stream is designed regarding participation status.

Though for practical reasons, we could not record all openings, the data also suggests a correlation with how openings are done: closings tend to be done similarly to how openings are done in terms of the visibility of the streamer. Finally, we do not get messages in our corpus asking the streamer to make themselves visible, as often happens in 'symmetrical' video-mediated communication, with both parties appearing in the video (see Licoppe & Morel, 2012).

The latter point suggests there is no normative orientation towards the visibility of streamers' *qua* talking heads in live streaming threads, which would constitute a constitutive feature of the live stream interaction order. It is not expected that the streamer, as a speaker, should be visible when there is nothing else relevant to show, especially during openings and closings. However, this is not to say talking heads are irrelevant in live streaming since we still see streamers closing

the streams with faces in about half of the streams with closings (as mentioned above). So we could speak of a relaxed relevance of the talking heads configuration in live streaming.

It might also be significant in that respect to distinguish between two ideal-typical genres of streams:

- The “interpersonal live stream” is where a most visible and often localized streamer talks about engaging with their online audience, with significant concern for finding mentionables.
- The “documentary live stream,” in which a usually mobile or active streamer film their experience while providing a kind of running commentary: mobility, activity, political engagement, etc.

The former makes the talking heads format very relevant in closings (and elsewhere) (we could speak of a “strong” orientation to the talking heads format), while the latter makes its use unnecessary or occasional in closings and elsewhere. We could speak of a “relaxed” orientation to the talking heads format, for, even in these streams, the talking heads format may not be completely absent in closings, as the following extract shows.

Extract 5

1 S: Just in case the taste is not so good, so,
2 uh hahaha.

3 **MER: enjoy it!**

4 S: >Lucky I had coffee before I left town<. #5.1
5 So:
6 ((flips the camera to talking head))
7 Hi and bye/ ((puts up hand))#5.2
8 (2.0) ((waves)) #5.3

9 S: Awesome.
10 A:nd,=
11 S: ((flips the camera back)) #5.4
12 =I'm gonna let you go for now.
13 And look forward to seeing you all along the journey,
14 (1.0)
15 That's beautiful, look, (.) yeah.
16 We're in the Bombays now. Salima!
17 Awesome, we're in the Bombays
18 Pretty cool countryside.
19 >\Okay<, catch you all later, thank you, bye for now.
20 B:ye, and thanks for the coffee mall,
21 Thank you \$. Okay, bye

22 **BAR:** \$You look gorgeous as usual 😊

23 S: Ah, hahahaha, bye\$, hahaha.h hahaha

24 **SAL:** \$ wow gorgeous
25 Really, yeah
26 Here we go we're in the Bombays
27 the Bombays hills, sort of scooting around them

28 **SAL: @BAR I agree**
29 Okay, see you all later
30 Thank you.
31 Joe:: and Jadean, sorry just leaving,
32 But let me stand up, >that's very nice of you<.
33 >Here you go<,
34 Ca\$tch you all later

35 **SAL:** \$👋

36 **OOM: Bye**
37 Oh, this side, they've got horses ((turns camera to the other side

of the train)) #5.5
 38 Sorry ah hahahaha
 39 They've got some nice horses,
 40 Awww, and a big tree
 41 Okay, this one ((turns back)) #5.6
 42 Okay, I'm off now. Okay, see you later,
 43 Bye, thank you
 44 ((disconnects))



5.1



5.2



5.3



5.4



5.5



5.6

In this extract, the streamer documents a train trip in New Zealand, and she has yet to show herself (Image 5.1). As she initiates the closing sequence, she briefly flips the camera and waves (Images 5.2 and 5.3), flipping back to her environment in her next closing turn and keeping it that way. She even shows the view again in Images 5.5 and 5.6. Throughout all this, she manages closings, apparently treating a few messages²¹, and eventually ends the call in this mode (with her face off the screen). So, even in documentary mode, flipping the camera to display oneself as a talking head, albeit momentarily, is still an option for the streamer, though not necessarily expected.

These two ideal types, the “interpersonal live stream” and the “documentary live stream,” lie at the two extremes of a continuous spectrum, with many streams combining features of both to some extent. Considering our early work on Periscope streams as mentioned above (Licoppe & Morel, 2018), where a majority of the streams had an “interpersonal character,” and this corpus, where “documentary streams” seem to dominate, it is an open question for further research whether this is an artifact of data collection, or whether there is some temporal trend operating, in which live streams are currently becoming more documentary²².

²¹ The laughter in line 23 is a response to the message in line 22.

²² The only differences between the data collections of these two corpora are that the 2018 one was collected around three years earlier than this one, and the 2018 one consists of streams mainly in French while this one is mainly in English.

3.4 The four-part organization in live streams' closing sequences

As noted earlier, the initial study of closings in phone conversations has shown that closings are typically initiated by the party calling the other, involve the collaborative production of relevant adjacent pairs (pre-closings, closings), and are provided an opportunity for the formulation of “unmentioned mentionables,” so that closings may become quite extended (Schegloff & Sacks, 1973). While the closings of streams usually involve recognizable pre-closings and closing turns, they are shaped by different interactional concerns that we want to unpack here.

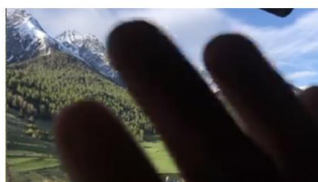
First, streams are somehow “owned” by the streamers, who are not just talking, but broadcasting. Moreover, streamers have many more interactional resources than viewers (who can only contribute discrete written messages without control over their precise timing). It is, therefore, almost always the case that streamers initiate closings. However, in “cold streams”, the relative rarity of the audience’s messages makes it possible that though made of recognizable pre-closing or closing turns, closings seem to be carried chiefly by the streamers. Extract 6 provides an example.

Extract 6

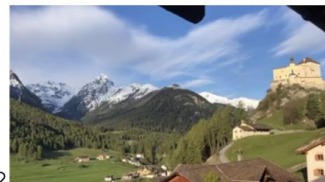
- 1 S: We'll say #6.1, good night or have a great day,
2 and goSodbye for Grisons.
3 **BOX:** §thanks for the View 👍
4 (3.0)
5 Goodbye for Switzerland.
6 Talk to you soon.
7 Bye everyone.
8 Bye #6.2
9 **VIC: I'll be looking forward to it. bye**
10 (28.0) ((moves camera slightly to the right #6.3))
11 S: ((disconnects))



6.1



6.2



6.3

In this extract, the streamer initiates closings (Lines 1-2). The message in Line 3 could be understood as a response. Then the streamer initiates a succession of closings with different designs, a “goodbye for Grisons” (Line 2), a “goodbye for Switzerland” (Line 5), and a “talk to you soon” (Line 6), which heralds future encounters, a “bye” to all (Line 7), which retrospectively frames the previous interaction as an encounter with a collective of undifferentiated viewers, and a “bye” which may be understood as a final closing (Line 8). Indeed, the streamer seems to move to other activities before eventually disconnecting, though a return closing message arrives (Line 9), which is left untreated. The possibility of producing a volley of first-pair part closings such as Lines 5-8, illustrates a potential concern for streamers in cold streams. They may have to handle the closings with no response or minimal and out-of-sync response from the audience, unlike closing sequences in conversational settings, where participants collaborate moment by moment to the production of closings.

Even in “hot streams”, with many incoming messages, it may happen that the streamer appears to carry the burden of doing closing by himself, by ignoring incoming messages, as in Extract 7.

Extract 7

1 S: Alright\$, thanks\$ for watching,=
2 **KEL:** \$in my fridge? Dead bodies
3 **BIG:** \$Your hair looks better now
4 S: =helping me to make for choices.
5 And I appreciate that\$.
6 **3AC:** \$embrace the drunkenness
7 S: I'm going to \$bed.
8 **MAR:** \$Deep cleanse?
9 \$(1.5)
10 **LIN:** \$\$Such a chore
11 S: I'm going to bed\$.
12 **CLE:** \$sing some tswift
13 S: An:d I appreciate you guys, and I love you\$.
14 **JAK:** \$Slide in my DMs
15 S: And you g\$uys are the best.
16 **DYL:** \$Cool
17 S: And\$ I hope you have a wonder\$ful even\$ing.
18 **JOB:** \$Good night ❤️❤️❤️
19 **AND:** \$Put your fox jacket in hall of fame
20 **HAN:** \$Stream yummy
21 S: An:: \$d I'll catch yo::u\$ on\$ the next one\$. hh
22 **MAR:** \$Noo
23 **RAC:** \$Queen of Chicago
24 **MAY:** \$What's in ur fridge? Wrong answer only
25 **MSH:** \$Night ily
26 S: ((reaches out and clicks closing buttons)) #7.1



27 **MAR:** Love you too
28 S: Ehm:::, got\$ it.
29 **ADA:** \$Cool
30 S: ((disconnects))

In this extract, the streamer produces a series of pre-closing and closing turns (Lines 7, 11, 13, 15, 17, 21), in the midst of incoming messages which she ignores. While some seem irrelevant to previous topics and to closings, others appear to be closing-relevant (Lines 18, 25, 27). This,

however, is an unusual situation²³. When messages relevant to closings are incoming, there seems to be an orientation towards treating them, as shown in Extract 8 below.

Extract 8



²³ For responding to closing messages, among the 61 streams with closing sequences in our corpus, there are only four streams with no definite responses (all hot streams, from one streamer), and five streams getting no messages from viewers when closing (so it is hard for streamers to attend to any message). Apart from these, the rest of the streams have all got some way of attending to closing messages.

1 S: Okay guys, I'm- (1.0) I'm going to ah- (2.0) get my
2 day started. Thanks for joining me. (.) Hope you
3 enjoy the birds.
4 \$ (30.0)

5 **det: \$have a great day**

6 **TPA: \$thanks Pete**

7 **Lad: \$@Pas my favorite is Shark Bite**

8 **Pas: \$Thanks and have a good day!**

9 **Koa: \$Enjoy your day. Bye.**

10 S: take care guys

11 **det: ciao**
12 (2.0)

13 **Lis: you too**

14 S: \$I-

15 **Lad: \$Thank you Peter! Have a very productive day!**

16 S: I hope everyone has a great morning, (.) \$evening,
17 **Lad: \$Bye**
18 (2.0)

19 S: \$lunch, (.) or a great night's sleep whatever the
20 case may-

21 **Pol: \$Bye**

22 S: might be (0.5) \$e::hm

23 **Lad: \$Elbow bumps everyone**

24 **Pol: Be well**

25 S: Yeah. \$Take care guys. Elbow bumps everybody.
26 Bye bye.

27 **Pas: \$@Lad seems interesting**

28 S: Ciao ciao Fiona. Sorry for the missing comments.
29 (0.5)
30 I apologize.
31 (1.0)
32 Take care guys. Bye bye.
33 (0.5)
34 S: ((disconnects))

Before closing and disconnecting (Lines 32-34), the streamer apologizes for not replying to all messages (“sorry for the missing comments”, Line 28). This shows that the normative orientation towards responding to relevant messages, which has been previously observed in streams (Licoppe & Morel, 2018), also operates during closing sequences. In the spirit of this

apology, several of the streamer's contributions may potentially be understood as responses to closing messages, such as Line 10, where the "take care guys" seems to respond to the volley of closing messages which precede (Lines 5, 6, 8, 9); Line 25, in which the initially positioned "yeah" makes it appear responsive to what precedes in general, while the design of the end-positioned expression "elbow bumps everybody" mirrors a previous message at Line 23, to which it appears responsive; and Line 28, the "ciao ciao Fiona" both highlights some degree of acquaintanceship²⁴ and frames the encounter as involving a few acquaintances, and a collection of unknown viewers, which possibly echoes the "ciao" at Line 11.

This orientation of the streamer to respond to closing messages must be nuanced, however. The hotter the stream, the more messages will be coming, so it becomes apparent the streamer will not be able to respond to all the messages in practice, giving a streamer more leeway to ignore closing messages or selectively respond to them. In hot streams, normativity and practicality may operate at cross-purposes regarding the treatment of incoming closing messages from the audience.

This normative orientation also gives a particular sequential aspect to closing sequences in streams. Instead of a succession of adjacent pairs, as in ordinary conversation, we often observe a four-part organization. The streamer initiates closings in the form of one or several first-pair part closing utterances, and some (usually delayed) responsive closing message (or messages) comes up. Because the streamer is expected to respond to messages, the streamer responds to such messages with closings of his or her own, which therefore appear to be third-positioned.

Such four-part closing sequences may need to be made clear by time delays. A general feature of streams is a particular form of the phenomenon of "disrupted turn adjacency" (Herring, 1999), which has initially been described in multi-party, asynchronous chat interaction. Since messages do not appear on screen until their authors hit the sending key (Berglund, 2009), and messages are posted in the order received by the system, without regard for what they are literally responding to (Herring, 1999), it is often the case that other messages will be sent in the meantime, disrupting and separating logically related turns from each other (Herring, 2013; Berglund, 2009). This results in "disrupted turn adjacency", which is described as one of the

²⁴ He uses her given name, which is not inferable from her messaging identifier.

obstacles to sequential coherence in CMC and a new norm of “loosened relevance” in computer-mediated conversations (Herring, 1999, 2013).

A similar phenomenon occurs in live streams. For the same reason, the audience's messages do not always appear on screen instantly after the talk or the embodied conduct in the video to which they are responsive, leading to a live stream-specific form of “disrupted turn adjacency”. This phenomenon is of course more pronounced in “hot streams”. In the case of closing messages, the “loosened relevance” manifests itself in making it equivocal to determine which initial closing from the streamer a delayed closing message from the audience is responding to, and it might also be unclear to which closing message from the audience a third-positioned closing from the streamer is responding to. The problem is amplified by the formulaic character of many closings. However, recipient design, as with “elbow bumps everybody” in the previous extract, is a powerful resource to manage such potential equivocality.

Such a sequential organization displaying four-part closing sequences (general closing, closing responsive message, recipient designed streamer’s third-position closing, repetition of generic closing by streamer) appears as a powerful resource to manage acquaintanceship (as with the third-positioned “ciao Fiona” in the extract above (Line 28). It is impossible for us to determine whether this is an instance of prior acquaintances following streamers, or of regular and previously unknown followers becoming acquaintances. However, the latter phenomenon may be reinforced by the development of “live streaming channels”, where the streamer is expected to stream regularly. Many closing sequences display at least an utterance foreshadowing a future stream: “And look forward to seeing you all along the journey” (Extract 5, Line 13); “talk to you soon” (Extract 6, Line 6); “I’ll catch you on the next one”(Extract 7, Line 21); and, below, “take care and hope we’ll see in our next scope” (Extract 9, Line 48); “see you around from Montreux (Extract 11, Lines 33 and 38”); “I’ll probably be back in like, half an hour” (Extract 12, Line 18). Such utterances evoke those turns making future arrangements observed before final closings in ordinary conversations (Schegloff & Sacks, 1973).

Regular streams facilitate the development of acquaintanceship between streamers and members of the audience, and streamers orient towards such acquaintanceship as a relevant feature of closing sequences, as in the extract below.

Extract 9



1 S: #9.1 I'll actually shut down here (.) so hhh



9.1

2 (4.0)

3 S: WHoo::hh

4 (1.0)

5 S: Okay h

6 (1.0)

7 S: alright guys hh

8 (1.9)

9 S: take care hh

10 (1.0)

11 Pat: Deep breathe

12 S: thanks for coming in and having a chill

13 S: San:d I'm gonna have a little bit=

14 QLa: \$Thanks Dave

15 dev@Kat: omg, living in California we know what that's like

16 **Kat@_Na:** -it was crazy
 17 S: =\$rest and I'm gonna head back to Osaka so hh
 18 **Pat:** \$gracias
 19 **Kat:** Bye Dave
 20 (3.0)
 21 S: Do we still have anybody coming in? h
 22 **yan:** BE SAFE
 23 S: uh:
 24 **cir:** Bye
 25 **Pat:** ❤️👍
 26 S: Denis (.) Laura h (.) Ciro (0.5) \$Molly
 27 **Lou:** \$Bye for now
 28 (1.5)
 29 **Kat:** GREAT to see you 🥳👏
 30 S: who else is in here?
 31 **dev:** Thanks Dave
 32 S: Dina h
 33 **CGi:** 🙌❤️
 34 **Qla:** Bye bye 🙌🙌
 35 S: Louis h
 36 S: uh::who else is in here?
 37 **lon:** 🤔
 38 S: Mocy h
 39 **rik:** 👍
 40 S: \$an:d hh
 41 **Pat:** \$paty
 42 (2.0)
 43 **NNa:** bye
 44 **Mss:** 🙌❤️🙌
 45 S: Sherly
 46 S: thanks for coming in guys
 47 **moa:** 🙌
 48 S: take care and hope we'll see in our \$next scope
 49 **Kat@Lou:** \$- bye!
 50 S: bye for now
 51 S: ((disconnects))

The streamer first initiates generic pre-closings and closings in which he retrospectively frames the audience as an undifferentiated collective by using the address term “guys” (Line 7), and more specifically as a “visiting” viewer “thank you for coming in” (Line 12). Then, several audience members respond with closing messages (Lines, 14, 18, 19). The streamer then

displays his paying attention to these messages. It highlights their relevance with “do we still have anybody coming in?” (Line 21), thus appearing to invite more of these messages. More messages come in (Lines 22, 24, 25), which, because of their timing, may now look as a response to the streamer’s request. Then the streamer acknowledges some of these closing messages with third-positioned receipts, designed as address terms, and more specifically, first names: “Dennis, Ciro, Laura, Molly”. Interestingly, some of these names cannot be inferred directly from messages, so he displays a high degree of acquaintanceship with the audience here. Then the streamer repeats the invitation for viewers to manifest themselves so that he may greet them personally (Lines 30, 36). He goes on to greet more viewers personally (Lines 32, 35, 38, 45). Eventually, he breaks with this pattern, thus marking the individualized acknowledgment sequence as over, and produces a new generic closing sequence (Lines 46-51), reframing the audience as an undifferentiated whole by re-using the address term “guy”, before eventually disconnecting. So what we observe here is the orientation of the streamer towards greeting individually known members of the audience in the closing sequence (as is also the case, though less systematically, in Extract 8). It is not possible from the data itself to say whether this is a normative expectation by the audience themselves. In any case, we see that within the four-part closing sequence:

1. Streamer’s generic closing
2. Audience members’ closing messages
3. Recipient-designed acknowledgment and/or closing by streamer
4. Repetition of generic closing by streamer

Step 2 and step 3 provide for a kind of inserted sequence that constitutes the main resource for managing acquaintanceship (and also enacting it) in live stream closing sequences. Such sequences display the normative expectation that streamers should treat relevant messages from the audience as much as possible during streams and in managing their closings.

3.5 The sensitivity of live stream closing sequences to topical talk

In ordinary conversation, closing sequences allow introducing yet unmentioned topics (Schegloff & Sacks, 1973), often giving rise to extended closing sequences. Though superficially similar, the situation is a bit different in live streams. Live streams are not conversations between acquaintances to update one another. They are directed by a streamer so that when the latter initiate closings, there is much less of a sense that this is the last chance for

discussing some unmentioned topic than in ordinary conversations. However, two sorts of things may occur recurrently. First, and particularly in hot streams, topical messages from the audience relating to an earlier topic or raising another point related to the stream may happen during the closing sequence. Because of the orientation of the streamers to treat messages, they may select one or several of these messages to respond to, so this occasions a sequential development. Second, since the streamers are documenting their experience, some noteworthy event may happen to them which they topicalize, initiating a new sequential development themselves during the closing sequence. We will discuss these two cases in the next sections.

3.5.1 The sensitivity of the closing sequence to incoming messages from the audience

The situation here involves streams with incoming messages and the streamer responding to them or selecting some of them to respond to. Since it takes time for the audience to write and send messages, this mostly takes the appearance of a potentially disrupted adjacent pair, i.e., the message and its response, which appears embedded in the closing sequence, as in Extract 10 below.

Extract 10

1 S: Alright guys, I'm gonna shut down the scope.
 2 (2.0)
 3 S: OK bye bye. Mike, gong bye bye #10.1



10.1

4 Son: \$Bye bye.
 5 DAN: \$Is testing up to standards though?
 6 S: See ya in a bit. Is testing\$ up to standards though?
 7 Mic: \$Very good..US is high!!! 😊
 8 S: /Yeah, medical systems in Thailand are actually super good so,
 9 yeah, we are not worried\$ about it.
 10 Isa: \$Well behaved
 11 S: Alright, bye guys.
 12 See ya another time.
 13 ((disconnects))

The streamer first initiates a pre-closing at Line 1, and then a generic closing (Line 3). He also invites his son to the closing (Line 3) (“gong bye bye” means “say bye bye” in Cantonese) and

his son joins with a closing (Line 4). He then produces another generic closing “see ya in a bit” (Line 6). At this moment, a new message comes up (Line 5), The streamer reads it out (Line 6) and then responds to it²⁵ (Lines 8-9). He ignores the other messages and produces the final closing (Lines 11-12) before disconnecting. So the message at Line 5, and the “reading + response” pattern (Lines 6, 8, 9) provide an instance of a topical adjacent pair embedded in the closing sequence.

However, it may also happen in that situation that topical developments in the closing sequence get more elaborate, especially if the incoming messages relate to a previous discussion, as in Extract 11 below. The streamer (in Switzerland) and some members of the audience have been discussing the Covid situation in different countries (Lines 1-6), then the streamer provides a “so” which signals the potential end of the topic (Line 8), then a brief topical elaboration (Line 11), and finally another topic-closing utterance (Line 12). He then flips the camera towards himself (Line 14), waves (Line 16), and initiates a pre-closing utterance (Line 17).

Extract 11

²⁵ Reading-responding is a standard way to treat messages in live streams (Licoppe & Morel, 2018).

1 **WAG: "SAFE" for the moment**
2 **mie: Very nice**
3 S: £S:afe for the moment£, hahaha. Somehow, yeah, (.) sure.
4 (2.0)
5 Well, they used to be allowed. You're in Japan. Japan is on the
6 safe list.
7 (1.0)
8 So,
9 (3.0)
10 S: ((pans slowly to the left #11.1))
11 S: you could actually travel:h.
12 Alright(.), so::
13 S: ((pans fast to the right #11.2))
14 S: ((flips the camera #11.3))
15 **WAG: Really??**
16 S: ((waves #11.4))
17 S: I'll see y-
18 yeah, Japan is on the safe list as far as I know. (.)=
19 **x_a: It's a very serene place**
20 S: =Yeah you'll have to check that one more but the EU
21 safe list e:r (0.5) we have Japan on it. (.) I think so.
22 (1.0)
23 **WAG: The situation is not good here**
24 S: yeah I know, but it is not as bad as it er (0.5) \$as
25 it makes it close.
26 **mar:** **\$grazie**
27 S: well whatever I'll #11.5 have to look maybe they
28 changed it already again. They've been changing so
29 often right now, so: hh #11.6
30 Okay, see you around
31 S: ((waves #11.7))
32 **WAG: Wow ok**
33 S: see you around from er
34 (1.0)
35 S: \$from e:::r #11.8
36 **Ton: \$See u**
37 **mie: Bye see you later I hope**

38 S: Montreux (.) ya (.) ciao
 39 (0.5)
 40 S: bye bye
 41 **x_a: See you**
 42 S: ((disconnects))



As he makes these pre-closing moves, a message from one member of the audience in Japan (WAG) with whom he was discussing Covid rules comes up (“Really??”, Line 15), projecting further topical elaboration. The streamer cuts his pre-closing turn during its production and self-repairs it into a topical response (Lines 18, 20, 21). A new message from WAG arrives (with a message in between from somebody else, which the streamer ignores), which initiates new topical developments in response from the streamer (Lines 24-25, 27-29), before relaunching the closing sequence. So, in this case, we do not have just a topical adjacent pair embedded into the closing sequence, but a whole topical development starts before the first orientation to closing and overflows into the closing sequence.

3.5.2 The sensitivity of the closing sequence to mentionable events on the streamer’s side

Since the streamer primarily documents or shares their experience, their streaming activity is sensitive to events happening to, or around them. This provides a particular form of sensitivity of the closing sequence in live streams to what we might call “mentionable noticeables”. In this configuration, the streamer initiates a topical development after potentially orienting toward closings. Extract 11 is a typical example. The streamer is driving and responding to messages (Lines 1-4), then she orients to closings through her utterance “Alright I’m home” (Line 5). The initially positioned “alright” signals the end of some prior episode of activity, while “I’m home” accounts for the potential end of the stream (which has been focused on her car trip). We cannot tell whether her audience understands it this way for there is no closing message from her yet, and she immediately pursues by noticing some remarkable event happening in front of her—a guy wearing a hat with tail-like stuff (Lines 5-6).

Extract 12

1 **GAd:** and people with pre existing conditions
2 S: And people pre, \$yeah.
3 **GAd:** \$I could technically get it if I wanted to
4 S: Oh, shit, you should get it.
5 Alright, I'm home. Does this dude wear like a
6 fucking \$like Russian-looking \$hat? =
7 **fka:** \$What kind of car is that
8 S: ((clicks the screen to flip the cam))#12.1



9 **Cmk:** \$Anyone 16 n up here
10 S: =Can you guys#12.2 check this dude out, please, like,



11 S: ((zooming in))#12.3



12 S: Uh £it's got a fucking tail on it, yeah, it has a tail on
13 it£, (1.0)
14 S: oh my god, he is fucking mad, jeez cunt,
15 oh, that killed me.
16 Okay\$, I'll see you guys later, ummmmm
17 **GAd:** \$too much effort right now having to prove it 😊
18 S: I'll probably be back in like, half an hour,
19 **Cmk:** **Awesome** 😊
20 S: hehehe, bye.
21 S: ((Disconnects))

As she utters her noticing, she flips the camera away from her and zooms in on the noticeable thing to make it visually available to her audience (Lines 8, 10, 11) while commenting on the event (Lines 12-15). At least part of her audience has oriented to what she is doing, for the “awesome” message (Line 19) can be understood as an assessment of what she is noticing and

making visually available. Eventually, she resumes the closing sequence by making arrangements (indicating when she expects to go on streaming, Lines 16 and 18) and a final closing (Line 20). So, within that orientation of live streaming to document her experience, closing sequences in live streams still allow for mentioning some noticeable event, if it is occasioned there and then.

4 Discussion and Conclusion

This paper describes how live streams ended in a corpus of 92 streams from Twitter Live. We have shown that except for a few instances of live streams more oriented towards broadcasting some public event, most live streams involve closing sequences, engaging the streamer and their audience, and thus retrospectively framing the live stream as a whole social encounter for which leave-taking is relevant. Roughly half of these closing sequences are done in a talking heads orientation (with camera flipping). Thus, the talking heads orientation appears loosely relevant and less a constitutive feature of live-streamed interaction than in two-sided video-mediated communication.

We have observed that streamers typically initiate closings, displaying that live stream encounters are “owned” by streamers and that their organization is sensitive to the flux of messages that may come from the audiences. There is a normative expectation that streamers should treat relevant messages as much as possible, which still seems relevant to some extent in the closing sequences throughout the streams we have shown. This provides for a characteristic four-part closing sequence in streams, where a streamer initiates a generic closing sequence, and closing responses from the audience come up as messages. Eventually, streamers reply to these closing responses through recipient-designed closings. The recognizability of these sequences may be blurred because messages come out of sync with timing over which participants do not have control, leading to the phenomenon of “disrupted turn adjacency” initially observed in multi-party chat interaction. We have also shown how this four-part closing sequence constitutes a powerful resource for streamers to acknowledge and manage acquaintanceship when leave-taking (instead of treating the audience as an undifferentiated group).

Finally, we have shown how the closing sequences of live streams provide an opportunity for further topical development in two different ways. First, when a “delayed” topic (concerning sequential relevance) comes from the audience during the closing sequence and is treated by the streamer (who is expected to do so when relevant), in its most basic form, this kind of topical

development provides for an adjacent pair (topical message, response by streamer) embedded in the closing sequence, but it can also lead occasionally to more elaborate topical development. Second, since streams are often oriented toward making the streamer's experience recognizable and shareable, streamers are sensitive when some "mentionable noticeables" happens to them or in their environment, even after potentially orienting towards closings. Different from the first one, this topical development often takes the form of a first-positioned noticing by the streamer, as the streamer is the one who initiates it.

As Goffman (1966) has argued, the way we do leave-taking actually provides a comment from the participants themselves on the kind of encounter that has just occurred. Our analysis of closing sequences in streams shows how they are thus retrospectively framed as a specific kind of video-mediated, multimodal, social encounter where participants have extremely asymmetric resources for communication and which are oriented towards making the streamers' experiences recognizable and shareable (and also, therefore, participating to a more significant trend of online behavior allowing for self-disclosure), repeatedly, to an online audience therefore increasingly made of regulars (and potentially constituting acquaintances). Such an encounter appears organized around an expectation of mutual engagement, that is, that the audience would respond to what is shown by the streamer with messages, and that the streamer would respond as much as possible to relevant messages.

A potential limitation of this study is related to the proliferation of live-streaming platforms with various and distinctive affordances for stream presentation and selection, stream access and audience response. However, the organization of closings we have identified here depends mainly on the articulation of live video and chat, which is common to all platforms. We have collected a random sample of similar streams on other platforms to confirm similar closing sequences, which is the case. Though a more systematic study is needed, there are also some interesting differences. Some affordances of these platforms, such as the 'channel' logic (which on Tiktok involves the need for viewers to subscribe to get access and relatively frequent broadcasts) seem to promote a higher degree of acquaintanceship.

Moreover, some aspects of our discussion, especially our suggestion of ideal-typical genres, must be reconsidered in light of these developments. Also, live streams increasingly use the possibility that several participants may be live on video simultaneously (in a sort of zoom-like videoconference situation) in front of a remote online audience. The development of multi-

streamer threads may affect the organization of closing sequences, which calls for further empirical research of the kind we have done here.

References

- Anjani, L., Mok, T., Tang, A., Oehlberg, L., & Goh, W. B. (2020). Why do people watch others eat food? an empirical study on the motivations and practices of mukbang viewers. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/3313831.3376567>
- Antaki, C. (2002). “lovely”: Turn-initial high-grade assessments in telephone closings. *Discourse Studies*, 4(1), 5–23. <https://doi.org/10.1177/14614456020040010101>
- Arminen, I. (2005). *Institutional interaction: Studies of talk at work*. Ashgate.
- Button, G. (1987). Moving out of closings. In: Button G and Lee JRE (eds) *Talk and Social Organization*. Philadelphia, PA: Multilingual Matters, pp. 101–151.
- Button, G. (1990). On varieties of closings. In: Psathas G (ed.) *Interaction Competence*. Lanham, MD: University Press of America, pp. 93–148.
- Berglund, T. Ö. (2009). Disrupted Turn Adjacency and Coherence Maintenance in. *Instant Messaging Conversations*. *Language@Internet*, 6, article 2.
- Clayman, S. E., & Maynard, D. W. (1995). *Ethnomethodology and conversation analysis. Situated order: Studies in the social organisation of talk and embodied activities*, 199(5), 1-30.
- Coppock, L. (2005). *Politeness strategies in conversation closing*. Unpublished manuscript: Stanford University.
- Choe, H. (2019). Eating together multimodally: Collaborative eating in mukbang, a Korean livestream of eating. *Language in Society*, 48(2), 171–208. <https://doi.org/10.1017/s0047404518001355>
- Choe, K. (2020). Talking the cat. *Footing lamination in a Korean livestream of cats mukbang*. *Journal of Pragmatics* 160, 60-79.
- Cunningham, S., Craig, D., & Lv, J. (2019). China’s livestreaming industry: Platforms, politics, and precarity. *International Journal of Cultural Studies*, 22(6), 719–736. <https://doi.org/10.1177/1367877919834942>
- Deppermann, A. (2022). *Action ascription in interaction*. Cambridge University Press.
- Drew, P., & Heritage, J. (1992). *Analyzing Talk at Work: An Introduction*. In P. Drew, & J. Heritage (Eds.), *Talk at Work. Interaction in Institutional Settings* (pp. 3-65). Cambridge: Cambridge University Press.

- Döpke, S., Brown, A., Liddicoat, A., & Love, K. (1994). Closings in Talkback Radio. *Spoken Interaction Studies in Australia*, 11, 21–46. <https://doi.org/10.1075/aralss.11.02dop>
- Frobenius, M. (2014). *The pragmatics of monologue: interaction in video blogs* (dissertation). Saarbrücken.
- Garst, K. (2015, October 14). How live-streaming is changing the face of Social Media. *Salesforce Canada Blog*. Retrieved June 15, 2022, from <https://www.salesforce.com/ca/blog/2015/10/live-streaming-and-social-media.html>
- Goffman, E. (1966). *Behavior in public places: Notes on the Social Organization of gatherings* (by) Erving Goffman. Free Press.
- Goffman, E. (1981). *Forms of talk*. University of Pennsylvania Press.
- Hazel, S., Mortensen, K., & Rasmussen, G. (2014). Introduction: A body of resources – CA studies of Social Conduct. *Journal of Pragmatics*, 65, 1–9. <https://doi.org/10.1016/j.pragma.2013.10.007>
- Herring, S. C. (2013). Relevance in computer-mediated conversation. In S. C. Herring, D. Stein, & T. Virtanen (Eds.), *Handbook of pragmatics of computer-mediated communication* (pp. 245-268). Berlin: Mouton. Prepublication version: <http://ella.slis.indiana.edu/~herring/relevance.pdf>
- Herring, S. C. (1999). Interactional coherence in CMC. In *Proceedings of the 32nd Annual Hawaii International Conference on Systems Sciences*. 1999. HICSS-32. Abstracts and CD-ROM of Full Papers (pp. 13-pp). IEEE.
- Arminen, I., Licoppe, C., & Spagnolli, A. (2016). Respecifying mediated interaction. *Research on Language and Social Interaction*, 49(4), 290–309. <https://doi.org/10.1080/08351813.2016.1234614>
- Ilomäki, S., & Ruusuvuori, J. (2020). From appearings to disengagements: Openings and closings in video-mediated tele-homecare encounters. *Social Interaction: Video-Based Studies of Human Sociality*, 3(3). <https://doi.org/10.7146/si.v3i3.122711>
- Johnson, M. R. (2022). Humour and comedy in Digital Game Live Streaming. *New Media & Society*, 146144482210951. <https://doi.org/10.1177/14614448221095160>
- Jefferson, G. (2004). Glossary of transcript symbols with an introduction. *Conversation Analysis*, 13–31. <https://doi.org/10.1075/pbns.125.02jef>
- Jefferson, G. (1973). A case of precision timing in ordinary conversation: Overlapped tag-positioned address terms in closing sequences. *Semiotica*, 9(1). <https://doi.org/10.1515/semi.1973.9.1.47>

- Licoppe, C., & Morel, J. (2012). Video-in-interaction: “Talking heads” and the Multimodal Organization of Mobile and Skype Video calls. *Research on Language & Social Interaction*, 45(4), 399–429. <https://doi.org/10.1080/08351813.2012.724996>
- Licoppe, C., & Morel, J. (2018). Visuality, text and talk, and the systematic organization of interaction in periscope live video streams. *Discourse Studies*, 20(5), 637–665. <https://doi.org/10.1177/1461445618760606>
- Due, B. L., & Licoppe, C. (2021). Video-mediated interaction (VMI). *Social Interaction. Video-Based Studies of Human Sociality*, 3(3). <https://doi.org/10.7146/si.v3i3.123836>
- Licoppe, C., & Rollet, N. (2020). « je dois y aller ». analyses de séquences de clôtures entre humains et robot. *Réseaux*, N° 220-221(2), 151–193. <https://doi.org/10.3917/res.220.0151>
- Licoppe, C., 2017. Skype appearances, multiple greetings and ‘coucou’ The sequential organization of video-mediated conversation openings. *Pragmatics* 27 (3), 351–386. <https://doi.org/10.1075/prag.27.3.03lic>.
- Lu, Z., Xia, H., Heo, S., & Wigdor, D. (2018). You watch, you give, and you engage. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/3173574.3174040>
- Llewellyn, N., & Hindmarsh, J. (Eds.). (2010). *Organisation, interaction and practice: Studies of ethnomethodology and conversation analysis*. Cambridge University Press.
- Mondada, L. (2018). Multiple temporalities of language and body in interaction: Challenges for transcribing multimodality. *Research on Language and Social Interaction*, 51(1), 85–106. <https://doi.org/10.1080/08351813.2018.1413878>
- Mondada, L. (2019). Contemporary issues in conversation analysis: Embodiment and materiality, multimodality and multisensoriality in social interaction. *Journal of Pragmatics*, 145, 47–62. <https://doi.org/10.1016/j.pragma.2019.01.016>
- Marulanda, J. F. F. (2022). Student satisfaction pilot experience with synchronous classroom live streaming styles during the COVID-19 pandemic. *IEEE Revista Iberoamericana de Tecnologías del Aprendizaje*, 17(3), 301-306.
- Nevile, M. (2005). ‘checklist complete’. or is it? closing a task in the airline cockpit. *Australian Review of Applied Linguistics*, 28(2), 60–76. <https://doi.org/10.1075/aryl.28.2.05nev>
- Oittinen, T. (2021). Material and embodied resources in the accomplishment of closings in technology-mediated business meetings. *Pragmatics. Quarterly Publication of the International Pragmatics Association (IPrA)*, 32(2), 299–327. <https://doi.org/10.1075/prag.19045.oit>

- Park, I. (2010). Marking an impasse: The use of anyway as a sequence-closing device. *Journal of Pragmatics*, 42(12), 3283–3299. <https://doi.org/10.1016/j.pragma.2010.06.002>
- Raymond, G., & Zimmerman, D. H. (2016). Closing matters: Alignment and misalignment in sequence and call closings in institutional interaction. *Discourse Studies*, 18(6), 716–736. <https://doi.org/10.1177/1461445616667141>
- Rein, K., & Venturini, T. (2018). Ploughing digital landscapes: How facebook influences the evolution of live video streaming. *New Media & Society*, 20(9), 3359–3380. <https://doi.org/10.1177/1461444817748954>
- Raclaw, J. (2008). Two Patterns for Conversational Closings in Instant Message Discourse. *Colorado Research in Linguistics*, 21. <https://doi.org/10.25810/5e6r-fk53>
- Recktenwald, D. (2018). *The Discourse of Online Live Streaming on Twitch: Communication between Conversation and Commentary (Dissertation)*. The Hong Kong Polytechnic University, Hong Kong.
- Schegloff, E. A. (2007). *Sequence organization in interaction: A primer in conversation analysis*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511791208>
- Schegloff, E. A., & Sacks, H. (1973). Opening Up Closings. *Semiotica*, 8, 289-327. <http://dx.doi.org/10.1515/semi.1973.8.4.289>
- Si, R. (2021). *China livestreaming e-commerce industry insights*. Springer Singapore.
- Team Periscope. (2020, December 15). Farewell, periscope. Medium. Retrieved August 17, 2022, from <https://periscope.medium.com/farewell-periscope-164db2742b7c>
- Zelenkauskaitė, A., & Loring-Albright, G. (2022). Facebook Live is not “liked”: Construction of liveness and the reception of video livestreaming. *New Media & Society*, 146144482210781. <https://doi.org/10.1177/1461444822>

CHAPTER 7

CONCLUSION AND FUTURE RESEARCH

7.1 SUMMARY AND SYNTHESIS OF THESIS

In an era characterized by a proliferation of communication technologies, the way individuals initiate interactions, maintain interactions and close interactions across various platforms has become a topic of growing scholarly interest. Mediated interactions can be distinct from face-to-face ones in many aspects. Goffman (1966) emphasizes the importance of face and eye contact for face-to-face encounters. Once a set of participants have avowedly opened themselves up to one another for an engagement, an eye-to-eye ecological huddle tends to be carefully maintained, maximizing the opportunity for participants to monitor one another's mutual perceivings for eye contact opens one up for face engagement (Goffman, 1966: 95). He also notes that encounters also demand participants' mutual openness. To initiate an encounter, there has to be a setting of mutual openness, in which the participants act as open persons (Goffman, 1966: 129) and put themselves in open regions (Goffman, 1966: 132) or positions for unacquainted to greet each other, open themselves up for encounters. Moreover, encounters are considered to be the moment when co-participants appear on the social scene with mutual appearances in symmetrical resources, which means either face-to-face, voice-to-voice, text-to-text, or video-to-video, like encounters in face-to-face, and technology-mediated engagement like telephone conversations, WhatsApp type chatting, skype video chatting WeChat video chatting, online business meeting (although online conference platforms like Zoom also afford asymmetric interaction when participants choose their voices muted and faces not showing.). However, none of the features above are totally established in the live streaming setting. Live streaming is the asymmetric interaction form that combines the video interaction of the streamer and the internet relay chat style interaction of the viewers. By the affordance of live streaming, the streamers can have their face, voice, and environment shown via video, but can only interact with mostly acquainted viewers who never show their faces or voices. In live streaming, most viewers only send text messages with an account avatar, an account name, and a joining notification. If the viewers do not even send a message, the streamer has to talk or show things before a group of people who almost hide themselves. It is like one is acting under cover while the other is in the open.

Live streaming is, therefore, a highly asymmetrical setting of participants in non-mutual openness. The core idea of this thesis is to examine the foundational multimodal practices and

sequential organization that constitute live streaming interactions. This study uses ethnomethodology and conversation analysis to analyze how live streams function as a unique form of communication. The research focuses on four key areas: the beginning, noticing, tasting, and closing sequences of live streams. By synthesizing empirical findings from these areas, the study aims to provide an understanding of the communicative features of live streaming.

The first article looks into how live streams on Twitter begin and highlights the unique nature of these starts compared to more predictable openings like in phone conversations. It turns out there's quite a variety in how streamers kick things off, yet many do follow a pattern where the stream seems to officially start at a certain point that everyone recognizes. Streamers often wait for more viewers to show up right after they go live, showing they care about having an audience. They also try to create a group atmosphere by addressing everyone together and then shifting to a more personal tone, greeting new viewers as they come in, which gives the feeling of hosting a friendly gathering. The streamer usually makes it clear what the stream will be about, setting the stage for the audience. Interestingly, while some of these practices are commonly expected, their absence is not always felt as something is missing—like the waiting part. And because viewers keep joining in randomly, the greetings can happen throughout the beginning of the stream and not just at the start.

In the experiential stream, the streamer takes on a more passive role, simply showing events without making it about their own presence or contributions. Here, the usual greetings and explanations might be skipped, and the streamer might comment as if they were just another viewer with a better view.

While these findings from Twitter streams give good insight, we suggest that the variety of features on different platforms could influence how these starts are shaped. Still, the method of tying together the streamers' and viewers' actions and concerns should be widely applicable across platforms. This approach has allowed for a deeper understanding of live stream beginnings, moving beyond simple categorizations to a more evidence-based view of live streaming as a social interaction.

The second article discusses how noticing works on live streaming platforms. We talk about the concept of "noticing-based action," where viewers send messages to the streamer about things they notice, which could be anything from a game they are playing to something in the

background. This helps viewers engage with the stream and shows they are paying attention. The study found that these messages often talk about personal things like how the streamer looks or acts. This can make the streamer feel good if the messages are positive, but it can also be uncomfortable or even seem rude, depending on what's said.

The article also mentions how some streams become very active with these noticing messages, which we call "noticing effervescence," where a lot of viewers start commenting on things they notice all at once. It seems that this happens, especially when the streamer is not doing much, so the viewers focus more on small details and the messages themselves. Finally, the article touches on the idea that some viewers think a streamer should "deserve" the attention they get. If a stream gets lots of viewers and messages, but the streamer is not doing much, some people might question why they are getting so much attention. We think this idea of "deserving" attention is interesting and should be looked into more.

The third article looks at how tasting, a personal yet publicly observable activity, is shown and shared through screens in today's world, where interacting via video is quite normal. The study used detailed analysis to explore how tasting food is shared and understood during live streams. It found that streamers often set up their cameras to capture their facial reactions to tasting, which helps create a sense of community with their viewers. It also noticed that streamers will look away after they taste something and then look back at the camera when they are ready to share their reaction, which is a way to involve their virtual audience. Sometimes, these reactions are extra expressive, making the shared experience more intense and lively.

The research also looked at how streamers and viewers communicate during these streams. Streamers will often give their first thoughts on what they are tasting as if they are reporting news, and viewers will reply with messages showing they get what the streamer is experiencing. In streams with more than one person, these initial thoughts become a way for the streamers to share knowledge about the food with each other and report to the audience at the same time.

This study helps us understand how people connect and communicate about shared experiences in new ways, thanks to digital technology. It is a contribution to fields like conversation analysis and digital communication. However, the paper suggests that more research is needed to get into the finer details of how these tasting sessions are presented and how they might be made

even more engaging for the audience, looking into different types of streams and how the setting is filmed.

The last study explored how live streams on Twitter Live come to a close. It found that, except for some streams that just show public events, most have a way of ending that makes the whole thing feel like a social event with a proper goodbye. Streamers are the ones who typically start the process of ending the stream, which shows they are in charge. They do this by responding to viewers' messages, which is expected as a way to interact personally with the audience. This interaction can sometimes be messy because messages do not always come in on time.

During the end of a stream, there can be some last-minute discussions. Either the audience brings up a new topic, or the streamer notices something worth mentioning. This adds a bit more interaction even as things are winding down. The way a stream ends can give us a clue about the nature of the interaction, which often involves the streamer sharing personal experiences with an audience that gets to know them over time.

The study notes that different live streaming platforms have unique features that can affect how streams end. For example, with more streams having multiple streamers on at once, like a video conference, this might change how streams come to a close. So, more research is needed to understand all these differences.

7.2 THEORETICAL AND PRACTICAL IMPLICATIONS

We have both theoretical and practical implications. This study adds a lot to our understanding of how people communicate online, especially when it comes to live streaming. It shows that the methods and ideas we use to study face-to-face and telephone interactions still apply when we are looking at the digital world (Have, 2000). These ideas help us see that the way people interact online is not just random; it follows patterns and uses different kinds of communication, like visual cues and how the audience gets involved. It also talks about getting the audience involved. This study believes that involvement is not just something that happens because the content is good; it is also about how the streamer and audience interact using different modes like sound, sight, and sometimes even taste. When it comes to 'sensory sociality,' which is about how we experience things together using all our senses. The study looked at live streams where people taste food and found that these digital spaces can actually create shared sensory experiences, which is something new compared to older forms of communication. Lastly, the

study helps us to understand 'affordances'—the idea of what you can do with a particular technology. The way people use features of live streaming platforms can change how we think about these technologies and what they are for. People might use features in unexpected ways or come up with new ways to connect and communicate.

For practical implications regarding the domain of human-computer interaction (HCI), this study provides helpful tips for people who make content for live streams and those who design the platforms they use. For creators, it is about making viewers more involved and building active communities. For designers, it is about making tools that fit the way people talk and share in live streams. Creators could: 1) Start their streams in a way that makes both the group and individuals feel noticed. This builds a sense of community from the get-go; 2) Pay attention to what viewers point out and use these moments to keep the conversation going. This back-and-forth keeps viewers interested; 3) Really get into the experience of sharing sensory stuff on camera. Emotional reactions may make viewers feel more connected; 4) End their streams in a way that gives viewers a last chance to interact, leaving them feeling satisfied and looking forward to the next stream. Designers could: 1) Make tools that let streamers start and end their streams in various ways, tailored to their audience's needs. 2) Give streamers instant data about how engaged their viewers are so that they can tweak their approach in real-time. 3) Create features that let viewers share how they are feeling more vividly, like better quality video for streams that focus on tastes or ways to react that go beyond just text. 4) Make it easier for streamers to spot and respond to important viewer messages, maybe with AI that highlights the comments that really matter.

7.3 LIMITATIONS AND FUTURE EXPECTATIONS

This collection of articles does have limitations. The study is limited by its focus on a specific set of live streaming platforms, which have their unique sets of features and user demographics. Specifically, for example, we only focus on live streams on Twitter, we only look at a limited number of the essential practices in live streaming, and we have so far an immature and developing way of categorizing the genres of live streams, etc. The limitations pave the way for further scholarly exploration.

1. More multimodal practices deserve exploration, far more than what has been presented in this thesis. Neville (2015) describes the interest in multimodal communicative resources as the “embodied turn,” and there is a wide range of embodied features that

are discussed in contemporary research (Mondada, 2014; Sindoni, 2014; Norris, 2014), including gaze, eye movement, posture, facial expressions, gestures, pointing, body torque, proxemics (i.e., the distance to the camera or between participants), kinesics (i.e., other types of body movement), etc. All these multimodal resources can be found relevant in the activities of live streaming. Within the scope of our data corpus and research plan, the practices of showing actions/activities, and flipping the camera, noticing in experiential streams are critical phenomena that deserve further research, although we have mentioned them more or less in the articles. In live streaming, both streamers and viewers manipulate multiple resources to show things in a certain framework. And flipping the camera is an extra common function for today's mobile phones, and users adopt it in different scenarios. Besides, some key concepts regarding live streaming interactions could be inspected, and live streaming is a perspicuous context for exploring these concepts. For example, multiactivity, affordances, joint attention, etc. Moreover, the observation of how live streaming is produced is also important. Video recording of the production of streams, like research done on video calling (Gan, 2020), can provide a distinct perspective on examining live streaming. We have some video recordings of how a streamer broadcasts on public streets in Paris, with all the devices *in vivo* (Brown et al., 2013); further data analysis could be applied to the recordings.

2. Comparisons across platforms and different stream types. It would be good to look at different streaming services to see how their specific tools and the types of communities they host impact the way people interact during live streams. Also, apart from streams in which viewers can have direct interaction with people in the streaming, there is an important type of stream in which viewers can not interact directly with people in it but can interact among themselves via texts. For example, cat eating streaming (Choe, 2020), nature/city view 7/24 streaming, and grand events/live news streaming (football match, presidential election, Olympic opening, etc.). Interactions in such types of streams are different. Frameworks and footings of viewers' messages vary according to the changing images of the streaming.
3. New features of live streaming platforms. Live streaming is always changing, especially as new features pop up. How a platform is set up can really shape the kind of communities that form and how streamers and viewers interact with each other. For example, the user 'subscribe' function and the embeddedness of the algorithm in TikTok Live and platforms alike can seriously affect what streams viewers will watch. Also, it

appears that streaming platforms are trending toward enabling viewers to interact with streamers via voice and even video rather than just traditional text messages. Besides, more and more live streaming platforms are enabling multi-streamer mode. When more than one person is streaming at the same time, things get complex. Traditional ideas about who talks, when and how the audience gets involved might be re-examined. This deserves a closer look. Long-term studies would be great for keeping an eye on how these changes affect the way people do live streaming.

4. New tech like VR and AR in live streaming. As virtual and augmented reality become more common, it is worth studying how these might change the way we experience and interact with live streaming, particularly on a social and sensory level.
5. Other video-mediated interactions. With the research done in this thesis as a foundation, we could extend our studies beyond just live streaming. Other video-mediated technologies may have similar interactions and the comparison work is worthy.

In conclusion, this thesis has explored the territory of live streaming interactions, unraveling the multimodal and sequentially organized practices of this technology-mediated phenomenon. By applying the perspective of EMCA, the study aims to demonstrate the stepwise communicative practices that constitute live streaming. The findings contribute to the understanding of the ways in which people engage, communicate, and connect in a digitalized world. It shows and, to some degree, extends the possibilities of the EMCA approach, as noted by Paul ten Have:

“Now that an increasing number of people spend various amounts of their time 'online', chatting with friends or whoever is available, it is time to study Computer-Mediated Conversation (CMC), as we previously studied face-to-face conversation and Telephone (Mediated) Conversation, using the same procedural perspective.” (Have, 2000)

REFERENCES

- Arminen, I., Licoppe, C., & Spagnolli, A. (2016). Respecifying mediated interaction. *Research on Language and Social Interaction*, 49(4), 290–309.
<https://doi.org/10.1080/08351813.2016.1234614>
- Alamiri, D., & Blustein, J. (2016). Privacy awareness and design for Live Video Broadcasting Apps. *HCI International 2016 – Posters' Extended Abstracts*, 459–464.
https://doi.org/10.1007/978-3-319-40548-3_76
- Antaki, C. (2000). Two rhetorical uses of the description “chat.” *M/C Journal*, 3(4).
<https://doi.org/10.5204/mcj.1856>
- Arminen, I., Licoppe, C., & Spagnolli, A. (2016). Respecifying mediated interaction. *Research on Language and Social Interaction*, 49(4), 290–309.
<https://doi.org/10.1080/08351813.2016.1234614>
- Baldauf-Quilliatre, H., & de Carvajal, I. C. (2015). Is the avatar considered as a participant by the players? A conversational analysis of multi-player videogames interactions. *PsychNology Journal*, 13(2).
- Baldauf-Quilliatre, H., & Ursi, B. (2023). Playing together on a large screen: Spatiality, materiality, temporality and the complexity of interaction. *Complexity of Interaction*, 255–295. https://doi.org/10.1007/978-3-031-30727-0_8
- Baldauf-Quilliatre, H., & Colón de Carvajal, I. (2020). Encouragement in videogame interactions. *Social Interaction. Video-Based Studies of Human Sociality*, 2(2).
<https://doi.org/10.7146/si.v2i2.118041>
- Benwell. (2006). *Discourse and identity*. Edinburgh University Press.
- Brown, B., McGregor, M., & Laurier, E. (2013). iPhone in vivo. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*.
<https://doi.org/10.1145/2470654.2466132>

- Brown, B., O'hara, K., Mcgregor, M., & Mcmillan, D. (2017). Text in talk. *ACM Transactions on Computer-Human Interaction*, 24(6), 1–25. <https://doi.org/10.1145/3152419>
- Brubaker, J. R., Venolia, G., & Tang, J. C. (2012). Focusing on shared experiences. *Proceedings of the Designing Interactive Systems Conference*. <https://doi.org/10.1145/2317956.2317973>
- Busch, G. (2018). How families use video communication technologies during intergenerational Skype Sessions. *International Perspectives on Early Childhood Education and Development*, 17–32. https://doi.org/10.1007/978-981-10-6484-5_2
- Busch, G. R. (2018). 'happy birthday grandpa.' *Research on Children and Social Interaction*, 2(1), 74–97. <https://doi.org/10.1558/rcsi.34849>
- Cappellini, M., & Azaoui, B. (2017). Sequences of normative evaluation in two telecollaboration projects: A comparative study of multimodal feedback through desktop videoconference. *Language Learning in Higher Education*, 7(1).
- Cech, C. G., & Condon, S. L. (2004). Temporal properties of turn-taking and turn-packaging in synchronous computer-mediated communication. *37th Annual Hawaii International Conference on System Sciences, 2004. Proceedings of The*. <https://doi.org/10.1109/hicss.2004.1265282>
- Chen, X., Chen, S., Wang, X., & Huang, Y. (2021). "I was afraid, but now I enjoy being a streamer!" *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW3), 1–32. <https://doi.org/10.1145/3432936>
- Christodoulidou, M. (2018). Consultee satisfaction in ending chats of an e-counseling service. *Discourse Studies*, 20(4), 461–487. <https://doi.org/10.1177/1461445618754431>
- Çimenli, B., Sert, O., & Jenks, C. (2022). Topic maintenance in video-mediated virtual exchanges: Rolling the ball back in L2 interactions. *System*, 108, 102834. <https://doi.org/10.1016/j.system.2022.102834>

- Cipolletta, S., Frassoni, E., & Faccio, E. (2017). Construing a therapeutic relationship online: An analysis of videoconference sessions. *Clinical Psychologist*, 22(2), 220–229.
- Dang-Anh, M. Situational involvement and footing shifts in mobile live video streams. *Gesprächsforschung*, (22), 304-336.
- Davidson, C., Danby, S. J., & Thorpe, K. (2017). “uh oh”: Multimodal meaning making during viewing of YouTube videos in preschool. *Multimodal Perspectives of Language, Literacy, and Learning in Early Childhood*, 233–255. https://doi.org/10.1007/978-3-319-44297-6_12
- de Fornel, M. (1996). The interactional framework of videophone interactions. *Réseaux*, 4(1), 68–99.
- de Rijk, L., & Stommel, W. (2023). Where to start? Initiating post-match chat interaction on Tinder. In A. Koivisto, H. Vepsäläinen, & M. T. Virtanen (Eds.), *Conversation Analytic Perspectives to Digital Interaction: Practices, Resources, and Affordances* (pp. 127–147). Helsinki: Finnish Literature Society.
- de Sá, M., Shamma, D. A., & Churchill, E. F. (2013). Live Mobile Collaboration for VIDEO PRODUCTION: Design, guidelines, and requirements. *Personal and Ubiquitous Computing*, 18(3), 693–707. <https://doi.org/10.1007/s00779-013-0700-0>
- DeBois, P. (2021). *Twitter discontinues periscope, but live streaming lives on*. CMSWire.com. <https://www.cmswire.com/digital-marketing/down-periscope-twitter-discontinues-the-livestreaming-app/>
- DeCoito, I., & Estaiteyeh, M. (2022). Transitioning to online teaching during the COVID-19 pandemic: An exploration of stem teachers’ views, successes, and challenges. *Journal of Science Education and Technology*, 31(3), 340–356. <https://doi.org/10.1007/s10956-022-09958-z>

- Del-Teso-Craviotto, M. (2008). Gender and sexual identity authentication in language use: The case of chat rooms. *Discourse Studies*, 10(2), 251–270.
<https://doi.org/10.1177/1461445607087011>
- Dezfuli, N., Günther, S., Khalilbeigi, M., Mühlhäuser, M., & Huber, J. (2013). Costream@home. *Proceedings of the 2nd International Workshop on Socially-Aware Multimedia*. <https://doi.org/10.1145/2509916.2509927>
- Dezuli, N., Huber, J., Churchill, E. F., & Mühlhäuser, M. (2013). CoStream: Co-construction of shared experiences through mobile live video sharing. *Electronic Workshops in Computing*. <https://doi.org/10.14236/ewic/hci2013.7>
- Dressler, R. A., & Kreuz, R. J. (2000). Transcribing oral discourse: A survey and a model system. *Discourse Processes*, 29(1), 25–36.
https://doi.org/10.1207/s15326950dp2901_2
- Dooly, M., & Tudini, V. (2022). ‘we should google that’: The dynamics of knowledge-in-interaction in an online student meeting. *Classroom Discourse*, 13(2), 188–211.
<https://doi.org/10.1080/19463014.2021.2023596>
- Dougherty, A. (2011). Live-Streaming Mobile Video. *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*.
<https://doi.org/10.1145/2037373.2037437>
- Due, B. L., & Licoppe, C. (2021). Video-mediated interaction (VMI). *Social Interaction. Video-Based Studies of Human Sociality*, 3(3). <https://doi.org/10.7146/si.v3i3.123836>
- Due, B. L., Lange, S. B., Nielsen, M. F., & Jarlskov, C. (2019). Mimicable embodied demonstration in a decomposed sequence: Two aspects of recipient design in professionals' video-mediated encounters. *Journal of Pragmatics*, 152, 13-27.
<https://doi.org/10.1016/j.pragma.2019.07.015>

- Due, B. L., & Lange, S. B. (2020). Body Part Highlighting: Exploring two types of embodied practices in two sub-types of showing sequences in video-mediated consultations. *Social Interaction. Video-Based Studies of Human Sociality*, 3(3). <https://doi.org/10.7146/si.v3i3.122250>
- Eisenmann, C., Mlynář, J., Turowetz, J., & Rawls, A. W. (2023). “machine down”: Making sense of human–computer interaction—Garfinkel’s research on Eliza and lyric from 1967 to 1969 and its contemporary relevance. *AI & SOCIETY*. <https://doi.org/10.1007/s00146-023-01793-z>
- Ekberg, S., Danby, S., Theobald, M., Fisher, B., & Wyeth, P. (2019). Using physical objects with young children in ‘face-to-face’ and telehealth speech and language therapy. *Disability and Rehabilitation*, 41(14), 1664–1675.
- Ekberg, S., Houen, S., Fisher, B., Theobald, M., & Danby, S. (2020). Engaging young children in speech and language therapy via videoconferencing. In C. M. Hayre, D. Muller, & M. Scherer (Eds.), *Everyday Technologies in Healthcare* (pp. 175-192). CRC Press.
- Faklaris, C., Cafaro, F., Blevins, A., O’Haver, M. A., & Singhal, N. (2020). A snapshot of bystander attitudes about mobile live-streaming video in public settings. *Informatics*, 7(2), 10. <https://doi.org/10.3390/informatics7020010>
- Faklaris, C., Cafaro, F., Hook, S. A., Blevins, A., O’Haver, M., & Singhal, N. (2016). Legal and ethical implications of mobile live-streaming video apps. *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*. <https://doi.org/10.1145/2957265.2961845>
- Favario, L., Siekkinen, M., & Masala, E. (2016). Mobile live streaming: Insights from the Periscope Service. *2016 IEEE 18th International Workshop on Multimedia Signal Processing (MMSP)*. <https://doi.org/10.1109/mmisp.2016.7813395>

- Fitzgerald, R., & Rintel, S. (2013). From lifeguard to bitch: How a story character becomes a promiscuous category in a couple's video call. *Australian Journal of Communication*, 40(2), 101–118.
- Fonseca, L. M., & Barbosa, S. D. (2020). Using live streams to share knowledge. *Proceedings of the 19th Brazilian Symposium on Human Factors in Computing Systems*. <https://doi.org/10.1145/3424953.3426637>
- Fonseca, L., & Barbosa, S. (2021). Knowledge sharing live streams: Real-time and on-demand engagement. *Proceedings of the 23rd International Conference on Enterprise Information Systems*. <https://doi.org/10.5220/0010401104410450>
- Francis, D., & Hester, S. (2007). *An invitation to ethnomethodology: Language, Society and Social Interaction*. Sage.
- Frobenius, M. (2011). Beginning a monologue: The opening sequence of video blogs. *Journal of Pragmatics*, 43(3), 814–827.
- Fung, A., Ismangil, M., He, W., & Cao, S. (2022). If I'm not streaming, I'm not earning: Audience relations and platform time on Douyin. *Online Media and Global Communication*, 1(2), 369–386. <https://doi.org/10.1515/omgc-2022-0001>
- Goffman, E. (1966). *Behavior in public places: Notes on the Social Organization of gatherings*. Free Press.
- Gan, Y. (2021). Capturing love at a distance: Multisensoriality in intimate video calls between migrant parents and their left-behind children. *Social Interaction: Video-Based Studies of Human Sociality*, 4(3).
- Gan, Y., & deSouza, D. K. (2022). Moving Out of View: Managing Participation and Visibility in Family Video-Mediated Communication. *Social Interaction: Video-Based Studies of Human Sociality*, 5(2).

- Gan, Y., Greiffenhagen, C., & Licoppe, C. (2020). Orchestrated openings in video calls: Getting young left-behind children to greet their migrant parents. *Journal of Pragmatics*, 170, 364–380. <https://doi.org/10.1016/j.pragma.2020.09.022>
- Garcia, A. C., & Baker Jacobs, J. (1999). The eyes of the beholder: Understanding the turn-taking system in quasi-synchronous computer-mediated communication. *Research on Language and Social Interaction*, 32(4), 337–367. https://doi.org/10.1207/s15327973rls3204_2
- Geerts, D., Vaishnavi, I., Mekuria, R., van Deventer, O., & Cesar, P. (2011). Are we in sync? *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/1978942.1978986>
- Golato, A., & Taleghani-Nikazm, C. (2006). Negotiation of face in web chats. *Mult*, 25(3), 293–321. <https://doi.org/10.1515/multi.2006.017>
- Guan, Z., Hou, F., Li, B., Phang, C. W., & Chong, A. Y. (2021). What influences the purchase of virtual gifts in live streaming in China? A cultural context-sensitive model. *Information Systems Journal*, 32(3), 653–689. <https://doi.org/10.1111/isj.12367>
- Gudmundsen, J. (2023). “Can you write it.” *Social Interaction. Video-Based Studies of Human Sociality*, 6(2). <https://doi.org/10.7146/si.v6i2.132542>
- Guo, E., & Zhang, F. (2021). From offline to online: The reduced embodiment in teacher-initiated turn-taking in GSL instructions. *Learning Technologies and Systems*, 357–369. https://doi.org/10.1007/978-3-030-66906-5_34
- Guo, J., & Fussell, S. R. (2020). A preliminary study of emotional contagion in live streaming. *Companion Publication of the 2020 Conference on Computer Supported Cooperative Work and Social Computing*. <https://doi.org/10.1145/3406865.3418309>
- Hutchby, I. (2001). Technologies, texts and affordances. *Sociology*, 35(2), 441–456.

- Haddington, P., & Oittinen, T. (2022). Interactional spaces in stationary, mobile, video-mediated and virtual encounters. In A. H. Jucker & H. Hausendorf (Eds.), *Pragmatics of Space* (pp. 317-362). De Gruyter Mouton.
- Haimson, O. L., & Tang, J. C. (2017). What makes live events engaging on Facebook live, periscope, and Snapchat. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/3025453.3025642>
- Hamilton, W. A., Garretson, O., & Kerne, A. (2014). Streaming on twitch. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/2556288.2557048>
- Hamilton, W. A., Tang, J., Venolia, G., Inkpen, K., Zillner, J., & Huang, D. (2016). Rivulet. *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video*. <https://doi.org/10.1145/2932206.2932211>
- Hansen, J. P. (2020). Invisible participants in a visual ecology: Visual space as a resource for organising video-mediated interpreting in hospital encounters. *Social Interaction. Video-Based Studies of Human Sociality*, 3(3). <https://doi.org/10.7146/si.v3i3.122609>
- Hansen, J. P. (2022). Recruiting repair: Making sense of interpreters' embodied actions in a video-mediated environment. *Discourse Studies*, 24(6), 719–740. <https://doi.org/10.1177/14614456221112261>
- Harper, R., Rintel, S., Watson, R., & O'Hara, K. (2017). The 'interrogative gaze': Making video calling and messaging 'accountable'. *Pragmatics*, 27(3), 319–350.
- Harper, R., Watson, R., & Licoppe, C. (2017). Interpersonal video communication as a site of human sociality. *Pragmatics*, 27(3), 301–318.
- Harper, R., Watson, R., & Woelfer, J. P. (2019). The Skype Paradox: Homelessness and Selective Intimacy in the Use of Communications Technology. In R. Harper, R. Watson,

- & C. Licoppe (Eds.), *Skyping the Family: Interpersonal Video Communication and Domestic Life* (pp. 147–174). John Benjamins.
- Have, P. ten. (2000). Computer-mediated chat. *M/C Journal*, 3(4).
<https://doi.org/10.5204/mcj.1861>
- Heath, C. C., & Luff, P. (1993). Disembodied conduct: Interactional asymmetries in video-mediated communication. In G. Button (Ed.), *Technology in Working Order: Studies of Work, Interaction, and Technology* (pp. 35–54). Routledge.
- Heath, C., & Luff, P. (1992). Media space and communicative asymmetries: Preliminary observations of video mediated interaction. *Human-Computer Interaction*, 7(3), 315–346.
- Heinemann, T., & Möller, R. L. (2016). The virtual accomplishment of knitting: How novice knitters follow instructions when using a video tutorial. *Learning, Culture and Social Interaction*, 8, 25–47. <https://doi.org/10.1016/j.lcsi.2015.11.001>
- Hellermann, J., & Thorne, S. L. (2022). Collaborative mobilizations of interbodied communication for cooperative action. *The Modern Language Journal*, 106(S1), 89–112. <https://doi.org/10.1111/modl.12754>
- Herring, S. C. (1999). Interactional coherence in CMC. *Journal of Computer-Mediated Communication*, 4 (4). <http://jcmc.indiana.edu/vol4/issue4/herring.html>
- Herring, S. C. (2010). Computer-mediated conversation: Introduction and overview. *Language@Internet*, 7, article 2. <http://www.languageatinternet.org/articles/2010/2801>
- Herring, S. C. (2013). 10. relevance in computer-mediated conversation. *Pragmatics of Computer-Mediated Communication*, 245–268.
<https://doi.org/10.1515/9783110214468.245>
- Hoffmann, S., & Fele, G. (2023). Dealing with missing participants in the opening phases of a videoconference. *Pragmatics*.

- Hou, F., Guan, Z., Li, B., & Chong, A. Y. (2019). Factors influencing people's continuous watching intention and consumption intention in live streaming. *Internet Research*, 30(1), 141–163. <https://doi.org/10.1108/intr-04-2018-0177>
- Hsiao, C. (2022). Authenticating discourses of “being oneself” on monetary-motivated livestreams. *Discourse, Context & Media*, 47, 100592. <https://doi.org/10.1016/j.dcm.2022.100592>
- Ilomäki, S., & Ruusuvuori, J. (2020). From appearings to disengagements: Openings and closings in video-mediated tele-homecare encounters. *Social Interaction: Video-Based Studies of Human Sociality*, 3(3). <https://doi.org/10.7146/si.v3i3.122711>
- Inkpen, K., Taylor, B., Junuzovic, S., Tang, J., & Venolia, G. (2013). Experiences2Go. *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*. <https://doi.org/10.1145/2441776.2441926>
- Jager, M., & Stommel, W. (2017). The risk of metacommunication to manage interactional trouble in online chat counseling. *Linguistik Online*, 87(8). <https://doi.org/10.13092/lo.87.4179>
- Jakonen, T., Dooly, M., & Balaman, U. (2022). Interactional practices in technology-rich L2 environments in and beyond the physical borders of the classroom. *Classroom Discourse*, 13(2), 111–118. <https://doi.org/10.1080/19463014.2022.2063547>
- Jawhar, S. S., Amir, A., Huq, R.-u., & Stewart, S. (2024). Collaborative problem solving and literacy practices: A conversation-analytic case study of children's online pre-gaming interaction. *Social Interaction: Video-Based Studies of Human Sociality*, 7(1).
- Jayson, S. (2010, December 30). 2010: The year technology replaced talking. *USA Today*. Retrieved February 10, 2024 from http://www.usatoday.com/yourlife/parenting-family/2010-12-30-1AYEAR30_CV_N.htm

- Jenks, C. J. (2009a). Getting acquainted in skypecasts: Aspects of social organization in online chat rooms. *International Journal of Applied Linguistics*, 19(1), 26–46. <https://doi.org/10.1111/j.1473-4192.2009.00211.x>
- Jenks, C. J. (2009b). When is it appropriate to talk? managing overlapping talk in multi-participant voice-based chat rooms. *Computer Assisted Language Learning*, 22(1), 19–30. <https://doi.org/10.1080/09588220802613781>
- Jenks, C. J., & Brandt, A. (2013). Managing mutual orientation in the absence of physical Copresence: Multiparty voice-based chat room interaction. *Discourse Processes*, 50(4), 227–248. <https://doi.org/10.1080/0163853x.2013.777561>
- Juhlin, O., Engström, A., & Reponen, E. (2010). Mobile Broadcasting. *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services*. <https://doi.org/10.1145/1851600.1851610>
- Heiden, L., Baldauf-Quilliatre, H., & Quignard, M. (2023). The role of cursor movements in a screen-based video game Interaction1. *People, Technology, and Social Organization*, 207–229. <https://doi.org/10.4324/9781003277750-14>
- Kastrenakes, J. (2021, March 31). *Periscope shuts down today*. The Verge. <https://www.theverge.com/2021/3/31/22360717/periscope-shut-down-twitter-live-streaming>
- Kelsey, S., & Amant, S. K. (2008). *Handbook of Research on Computer Mediated Communication*. Information Science Reference.
- Keng Wee Ong, K. (2011). Disagreement, confusion, disapproval, turn elicitation and floor holding: Actions as accomplished by ellipsis marks-only turns and blank turns in quasisynchronous chats. *Discourse Studies*, 13(2), 211–234. <https://doi.org/10.1177/1461445610392138>

- Khobra, U., & Gaur, R. (2020). Live (LIFE) streaming: Virtual Interaction, virtual proximity, and streaming everyday life during the covid- 19 pandemic. *Rupkatha Journal on Interdisciplinary Studies in Humanities*, 12(5).
<https://doi.org/10.21659/rupkatha.v12n5.rioc1s26n2>
- Kim, S., Lee, G., Lee, S., Lee, S., & Lee, J. (2019). Game or live streaming?: Motivation and social experience in live mobile quiz shows. *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*. <https://doi.org/10.1145/3311350.3347187>
- Kleinberg, S. (2015, April 1). *Live streaming: The next big thing in social media*. Chicago Tribune. <https://www.chicagotribune.com/lifestyles/ct-so-social-livestreaming-meerkat-periscope-20150401-column.html>
- Koivisto, A., Vepsäläinen, H., & Virtanen, M. T. (2023). *Conversation Analytic Perspectives to Digital Interaction. Practices, Resources, and Affordances*.
<https://doi.org/10.21435/sflin.22>
- König, K. (2023). Response design in WhatsApp chats: contextualising different stances of confirmation and agreement in text-based interaction. In A. Koivisto, H. Vepsäläinen, & M. T. Virtanen (Eds.), *Conversation Analytic Perspectives to Digital Interaction: Practices, Resources, and Affordances* (pp. 95–124). Helsinki: Finnish Literature Society.
- Laurier, E., & Reeves, S. (2014). Cameras in video games: Comparing play in Counter-Strike and the Doctor Who adventures. In M. Broth, E. Laurier, & L. Mondada (Eds.), *Studies of Video Practices: Video at Work* (pp. 181–207). Abingdon, Oxon UK: Routledge.
- Lee, Y. (2021). Patterns of users' first turns with a service chatbot: A conversation analytic perspective. *Korean Journal of Applied Linguistics*, 37(Special), 117–160.
<https://doi.org/10.17154/kjal.2021.7.37.special.117>

- Lee, Y.-C., Yen, C.-H., Wang, D., & Fu, W.-T. (2019). Understanding how digital gifting influences social interaction on live streams. *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services*. <https://doi.org/10.1145/3338286.3340144>
- Li, Y., Wang, C., & Liu, J. (2020). A systematic review of literature on User Behavior in video game live streaming. *International Journal of Environmental Research and Public Health*, 17(9), 3328. <https://doi.org/10.3390/ijerph17093328>
- Licoppe, C. (2015). Video communication and ‘camera actions’: The production of wide video shots in courtrooms with remote defendants. *Journal of Pragmatics*, 76, 117–134.
- Licoppe, C. (2017). Showing objects in Skype video-mediated conversations: From showing gestures to showing sequences. *Journal of Pragmatics*, 110, 63–82.
- Licoppe, C., 2017. Skype appearances, multiple greetings and ‘coucou’ The sequential organization of video-mediated conversation openings. *Pragmatics* 27 (3), 351–386. <https://doi.org/10.1075/prag.27.3.03lic>.
- Licoppe, C. (2019). Liquidity and attachment in the mobile hookup culture. A comparative study of contrasted interactional patterns in the main uses of Grindr and tinder. *Journal of Cultural Economy*, 13(1), 73–90. <https://doi.org/10.1080/17530350.2019.1607530>
- Licoppe, C. (2021). The spectre of ‘ghosting’ and the sequential organization of post-match tinder chat conversations. *Analysing Digital Interaction*, 155–176. https://doi.org/10.1007/978-3-030-64922-7_8
- Licoppe, C., & Dumoulin, L. (2015). “Is there someone in my videoconference room?”: Managing remote witnesses in distributed courtrooms. In B. Dupret, M. Lynch, & T. Berard (Eds.), *Law at Work: Studies in Legal Ethnomethods* (pp. 191–219). Oxford University Press.

- Licoppe, C., & Morel, J. (2009). The Collaborative Work of Producing Meaningful Shots in Mobile Video Telephony. In Proc. MobileHCI 2009 (pp. 254-263).
- Licoppe, C., & Morel, J. (2012). Video-in-interaction: “talking heads” and the multimodal organization of mobile and Skype video calls. *Research on Language and Social Interaction*, 45(4), 399–429.
- Licoppe, C., & Morel, J. (2014). Mundane video directors in interaction. Showing one's environment in Skype and mobile video calls. In M. Broth, E. Laurier, & L. Mondada (Eds.), *Studies of Video Practices: Video at Work* (pp. 135-160). Routledge.
- Licoppe, C., & Morel, J. (2018). Visuality, text and talk, and the systematic organization of interaction in periscope live video streams. *Discourse Studies*, 20(5), 637–665. <https://doi.org/10.1177/1461445618760606>
- Licoppe, C., & Tuncer, S. (2019). The initiation of showing sequences in video-mediated communication. *Gesprächsforschung: Online-Zeitschrift zur verbalen Interaktion*, 20, 545-571.
- Licoppe, C., & Veyrier, C.-A. (2017). How to show the interpreter on screen? The normative organization of visual ecologies in multilingual courtrooms with video links. *Journal of Pragmatics*, 107, 147–164.
- Licoppe, C., Luff, P. K., Heath, C., Kuzuoka, H., Yamashita, N., & Tuncer, S. (2017). Showing Objects: Holding and Manipulating Artefacts in Video-mediated Collaborative Settings. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17* (pp. 5295–5306). ACM Press.
- Licoppe, C., Rivière, C. A., & Morel, J. (2016). Grindr casual hook-ups as interactional achievements. *New Media & Society*, 18(11), 2540–2558. <https://doi.org/10.1177/1461444815589702>

- Licoppe, C., Rivière, C. A., & Morel, J. (2017). Proximity Awareness and the privatization of sexual encounters with strangers. *Place, Space, and Mediated Communication*, 45–65. <https://doi.org/10.4324/9781315394183-4>
- Licoppe, C., Verdier, M., & Dumoulin, L. (2013). Courtroom interaction as a multimedia event: The work of producing relevant videoconference frames in French pre-trial hearings. *The Electronic Journal of Communication / La Revue Electronique de Communication*, 23(1-2), 1–28.
- Long, Q., & Tefertiller, A. C. (2020). China’s new mania for live streaming: Gender differences in motives and uses of social live streaming services. *International Journal of Human–Computer Interaction*, 36(14), 1314–1324. <https://doi.org/10.1080/10447318.2020.1746060>
- Lu, Z., Heo, S., & Wigdor, D. J. (2018). StreamWiki. *Proceedings of the ACM on Human–Computer Interaction*, 2(CSCW), 1–26. <https://doi.org/10.1145/3274381>
- Lu, Z., Xia, H., Heo, S., & Wigdor, D. (2018). You watch, you give, and you engage. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/3173574.3174040>
- Luff, P., & Heath, C. (1998). Mobility in collaboration. *Proceedings of the 1998 ACM Conference on Computer Supported Cooperative Work*. <https://doi.org/10.1145/289444.289505>
- Markman, K. M. (2009). “so what shall we talk about”: Openings and closings in chat-based virtual meetings. *Journal of Business Communication*, 46(1), 150–170. <https://doi.org/10.1177/0021943608325751>
- McHoul, A., & Rapley, M. (2000). Still on holidays Hank? --. *M/C Journal*, 3(4). <https://doi.org/10.5204/mcj.1857>

- Melander Bowden, H., & Svahn, J. (2020). Collaborative work on an online platform in video-mediated homework support. *Social Interaction. Video-Based Studies of Human Sociality*, 3(3). <https://doi.org/10.7146/si.v3i3.122600>
- Meredith, J. (2015). Transcribing screen-capture data: The process of developing a transcription system for multi-modal text-based data. *International Journal of Social Research Methodology*, 19(6), 663–676. <https://doi.org/10.1080/13645579.2015.1082291>
- Meredith, J. (2019). Conversation analysis and online interaction. *Research on Language and Social Interaction*, 52(3), 241–256. <https://doi.org/10.1080/08351813.2019.1631040>
- Meredith, J., & Stokoe, E. (2013). Repair: Comparing Facebook ‘chat’ with spoken interaction. *Discourse & Communication*, 8(2), 181–207. <https://doi.org/10.1177/1750481313510815>
- Mlynář, J., González-Martínez, E., & Lalanne, D. (2018). Situated organization of video-mediated interaction: A review of ethnomethodological and conversation analytic studies. *Interacting with Computers*, 30(2), 73–84.
- Mondada, L. (2007). Operating together through videoconference: Members’ procedures for accomplishing a common space of action. In S. Hester & D. Francis (Eds.), *Orders of Ordinary Action: Respecifying Sociological Knowledge* (pp. 51–67). Ashgate.
- Mondada, L. (2011). The situated organisation of directives in French: Imperatives and action coordination in video games. *Nottingham French Studies*, 50(2), 19–50.
- Mondada, L. (2012). Coordinating action and talk-in-interaction in and out of video games. In R. Ayaß & C. Gerhardt (Eds.), *The Appropriation of Media in Everyday Life* (pp. 231–270). Amsterdam/Philadelphia: John Benjamins.
- Mondada, L. (2013). Coordinating mobile action in real time: The timely organisation of directives in video games. In P. Haddington, L. Mondada, & M. Nevile (Eds.),

- Interaction and Mobility: Language and the Body in Motion (pp. 300–342). Berlin: De Gruyter.
- Mondada, L. (2018). Practices for showing, looking, and videorecording: The interactional establishment of a common focus of attention. *Embodied Activities in Face-to-Face and Mediated Settings*, 63–104. https://doi.org/10.1007/978-3-319-97325-8_3
- Morrison, K. (2015, May 13). *Beyond Periscope and Meerkat: The State of live-streaming video*. Adweek. <https://www.adweek.com/performance-marketing/beyond-periscope-and-meerkat-the-state-of-live-streaming-video/>
- Mostovaia, I. (2021). Other-initiations of repair in German WhatsApp chats. *Discourse, Context & Media*, 40, 100470. <https://doi.org/10.1016/j.dcm.2021.100470>
- Mughal, M. A., & Juhlin, O. (2013). Context-dependent software solutions to handle video synchronization and delay in collaborative live mobile video production. *Personal and Ubiquitous Computing*, 18(3), 709–721. <https://doi.org/10.1007/s00779-013-0701-z>
- Muñoz, S. A. (2016). Attending multi-party videoconference meetings: The initial problem. *Language@Internet*, 13(3).
- Naper, I. (2011). Conversation in a multimodal 3D virtual environment. *Language@Internet*, 8, Article 7.
- Nardi, B. A., Schwarz, H., Kuchinsky, A., Leichner, R., Whittaker, S., & Scلابassi, R. (1993). Turning away from talking heads. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '93*. <https://doi.org/10.1145/169059.169261>
- Nathan, M., Harrison, C., Yarosh, S., Terveen, L., Stead, L., & Amento, B. (2008). CollaboraTV. *Proceedings of the 1st International Conference on Designing Interactive User Experiences for TV and Video*. <https://doi.org/10.1145/1453805.1453824>
- Neuage, T. (2004). *Conversational Analysis of Chatroom Talk* (thesis). *Conversational analysis of chatroom talk*.

- Nikolopoulou, K., & Kousloglou, M. (2022). Online teaching in covid-19 pandemic: Secondary School Teachers' beliefs on teaching presence and school support. *Education Sciences*, 12(3), 216. <https://doi.org/10.3390/educsci12030216>
- Nilsen, M., & Mäkitalo, Å. (2010). Towards a conversational culture? how participants establish strategies for co-ordinating chat postings in the context of in-service training. *Discourse Studies*, 12(1), 90–105. <https://doi.org/10.1177/1461445609346774>
- Oittinen, T. (2023). Highlighting as a referential and collaborative practice in multiparty video-mediated learning activities. *Classroom Discourse*, 1–23. <https://doi.org/10.1080/19463014.2023.2259020>
- Oittinen, T. (2023). Including written turns in spoken interaction: Chat as an organizational and participatory resource in video-mediated activities. *Research on Language and Social Interaction*, 56(4), 269–290. <https://doi.org/10.1080/08351813.2023.2272524>
- Oittinen, T., & Piirainen-Marsh, A. (2015). Openings in technology-mediated business meetings. *Journal of Pragmatics*, 85, 47–66. <https://doi.org/10.1016/j.pragma.2015.06.001>
- Paulus, T., Warren, A., & Lester, J. N. (2016). Applying conversation analysis methods to online talk: A literature review. *Discourse, Context and Media* 12, 1-10. <http://dx.doi.org/10.1016/j.dcm.2016.04.001>
- Pan, S., & Gan, Y. (2022). Continuing assessments in online dating: Enabling relational development between potential romantic partners in WeChat Conversations. *Discourse Studies*, 24(5), 545–565. <https://doi.org/10.1177/14614456221108597>
- Pappas Y, Seale C. The opening phase of telemedicine consultations: an analysis of interaction. *Soc Sci Med*. 2009 Apr;68(7):1229-37. doi: 10.1016/j.socscimed.2009.01.011. Epub 2009 Feb 7. PMID: 19201514.

- Pekarek Doehler, S., & Balaman, U. (2021). The routinization of grammar as a social action format: A longitudinal study of video-mediated interactions. *Research on Language and Social Interaction*, 54(2), 183–202. <https://doi.org/10.1080/08351813.2021.1899710>
- Perry, M., Juhlin, O., & Engström, A. (2014). Dealing with time, just in time: Sense-making and clip allocation in multiperson, multistream, live replay TV production. In *Studies of Video Practices* (pp. 262-285). Routledge.
- Piirainen-Marsh, A. (2012). Organising participation in video gaming activities. In R. Ayaß & C. Gerhardt (Eds.), *The Appropriation of Media in Everyday Life* (pp. 195–230). Amsterdam/Philadelphia: John Benjamins.
- Pomerantz, A. (1985). Pursuing a response. *Structures of Social Action*, 152–164. <https://doi.org/10.1017/cbo9780511665868.011>
- Raman, A., Tyson, G., & Sastry, N. (2018). Facebook (a)live? *Proceedings of the 2018 World Wide Web Conference on World Wide Web - WWW '18*. <https://doi.org/10.1145/3178876.3186061>
- Reed, D., & Ashmore, M. (2000). The naturally-occurring chat machine. *M/C Journal*, 3(4). <https://doi.org/10.5204/mcj.1860>
- Reeves, S., Greiffenhagen, C., & Laurier, E. (2017). Video gaming as practical accomplishment: Ethnomethodology, conversation analysis, and play. *Topics in Cognitive Science*, 9(2), 308–342.
- Reeves, S., Greiffenhagen, C., Flintham, M., Benford, S., Adams, M., Row Farr, J., & Tandavantij, N. (2015). I'd hide you. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/2702123.2702257>
- Rein, K., & Venturini, T. (2018). Ploughing digital landscapes: How facebook influences the evolution of live video streaming. *New Media & Society*, 20(9), 3359–3380. <https://doi.org/10.1177/1461444817748954>

- Rellstab, D. H. (2007). Staging gender online: Gender plays in swiss internet relay chats. *Discourse & Society*, 18(6), 765–787. <https://doi.org/10.1177/0957926507082195>
- Rhyll Vallis. (2001). Applying membership categorization analysis to chat-room talk. *How to Analyse Talk in Institutional Settings*. <https://doi.org/10.5040/9781350933927.ch-007>
- Rintel, E. S., & Pittam, J. (1997). Strangers in a strange land interaction management on internet relay chat. *Human Communication Research*, 23(4), 507–534. <https://doi.org/10.1111/j.1468-2958.1997.tb00408.x>
- Rintel, E. S., Mulholland, J., & Pittam, J. (2006). First things first: Internet relay chat openings. *Journal of Computer-Mediated Communication*, 6(3), 0–0. <https://doi.org/10.1111/j.1083-6101.2001.tb00125.x>
- Rintel, S. (2013a). Tech-tied or tongue-tied? Technological versus social trouble in relational video calling. In Proceedings of the Forty-Sixth Hawaii International Conference on System Sciences (pp. 3433–3352). IEEE.
- Rintel, S. (2013b). Video calling in long-distance relationships: The opportunistic use of audio/video distortions as a relational resource. *The Electronic Journal of Communication / La Revue Electronique de Communication*, 23(1-2).
- Roca-Cuberes, C., Gibson, W., & Mora-Rodriguez, M. (2023). Relationship initiation and formation in post-match tinder chat conversations. *Discourse & Communication*, 17(4), 462–493. <https://doi.org/10.1177/17504813231156118>
- Rong, E. Z., Zhou, M. M., Lu, Z., & Fan, M. (2022). “it feels like being locked in a cage”: Understanding blind or low vision streamers’ perceptions of content curation algorithms. *Designing Interactive Systems Conference*. <https://doi.org/10.1145/3532106.3533514>
- Rosenbaun, L., & Licoppe, C. (2017). Showing ‘digital’ objects in web-based video chats as a collaborative achievement. *Pragmatics*, 27(3), 419–446.

- Rosenbaun, L., & Licoppe, C. (2019). Showing ‘Digital’ Objects in Web-Based Video Chats as a Collaborative Achievement. In R. Harper, R. Watson, & C. Licoppe (Eds.), *Skyping the Family: Interpersonal Video Communication and Domestic Life* (pp. 119–146). John Benjamins.
- Rusk, F., & Pörn, M. (2019). Delay in L2 interaction in video-mediated environments in the context of Virtual Tandem Language Learning. *Linguistics and Education*, 50, 56–70. <https://doi.org/10.1016/j.linged.2019.02.003>
- Spagnolli, A., Genovese, S., & Mori, M. (2021). Similarities and differences across settings: The case of turn continuations in instant messaging. *Analysing Digital Interaction*, 133–154. https://doi.org/10.1007/978-3-030-64922-7_7
- Spagnolli, A., Cenzato, G., & Gamberini, L. (2023). Modeling the conversation with Digital Health assistants in adherence apps: Some considerations on the similarities and differences with familiar medical encounters. *International Journal of Environmental Research and Public Health*, 20(12), 6182. <https://doi.org/10.3390/ijerph20126182>
- Sandel, T. L., Ou, C., Wangchuk, D., Ju, B., & Duque, M. (2019). Unpacking and describing interaction on Chinese WeChat: A methodological approach. *Journal of Pragmatics*, 143, 228–241. <https://doi.org/10.1016/j.pragma.2018.08.011>
- Schönfeldt, J., & Golato, A. (2003). Repair in chats: A conversation analytic approach. *Research on Language & Social Interaction*, 36(3), 241–284. https://doi.org/10.1207/s15327973rlsi3603_02
- Searles, D. K. (2018). ‘look it daddy.’ *Research on Children and Social Interaction*, 2(1), 98–119. <https://doi.org/10.1558/rcsi.32576>
- Segerstad, Y. H. af. (2002). (dissertation). *Use and adaptation of written language to the conditions of computer-mediated communication*. Department of Linguistics, Göteborg University, Göteborg.

- Seuren, L. M., Gilbert, A., Ramdharry, G., Walumbe, J., & Shaw, S. E. (2024). Video analysis of communication by physiotherapists and patients in video consultations: A qualitative study using conversation analysis. *Physiotherapy*, 123, 30–37.
- Seuren, Lucas M., Wherton, J., Greenhalgh, T., & Shaw, S. E. (2021). Whose turn is it anyway? latency and the organization of turn-taking in video-mediated interaction. *Journal of Pragmatics*, 172, 63–78. <https://doi.org/10.1016/j.pragma.2020.11.005>
- Seuren, Lucas Martinus, Wherton, J., Greenhalgh, T., Cameron, D., A’Court, C., & Shaw, S. E. (2020). Physical examinations via video for patients with heart failure: Qualitative study using conversation analysis. *Journal of Medical Internet Research*, 22(2). <https://doi.org/10.2196/16694>
- Shamma, D. A., Marlow, J., & Denoue, L. (2019). Interacting with smart consumer cameras: Exploring gesture, voice, and AI control in video streaming. *Proceedings of the 2019 ACM International Conference on Interactive Experiences for TV and Online Video*. <https://doi.org/10.1145/3317697.3323359>
- Shaw, S. E., Seuren, L. M., Wherton, J., Cameron, D., A’Court, C., Vijayaraghavan, S., Morris, J., Bhattacharya, S., & Greenhalgh, T. (2020). Video consultations between patients and clinicians in diabetes, cancer, and heart failure services: Linguistic Ethnographic Study of video-mediated interaction. *Journal of Medical Internet Research*, 22(5). <https://doi.org/10.2196/18378>
- Shoib, S., Chandradasa, M., Nahidi, M., Amanda, T. W., Khan, S., Saeed, F., Swed, S., Mazza, M., Di Nicola, M., Martinotti, G., Di Giannantonio, M., Armiya’u, A. Y., & De Berardis, D. (2022). Facebook and suicidal behaviour: User experiences of suicide notes, live-streaming, grieving and preventive strategies—a scoping review. *International Journal of Environmental Research and Public Health*, 19(20), 13001. <https://doi.org/10.3390/ijerph192013001>

- Siitonen, P., Helisten, M., Siromaa, M., Rauniomaa, M., & Holmström, M. (2022). Managing co-presence with a wave of the hand: Waving as an interactional resource in openings and closings of video-mediated breaks from work. *Gesture*, 21(1), 82-114. <https://doi.org/10.1075/gest.00000.000>
- Singanamalla, S., Thies, W., & Scott, C. (2018). Avatar. *Proceedings of the 3rd International Workshop on Multimedia Alternate Realities*. <https://doi.org/10.1145/3268998.3269003>
- SitePoint. (n.d.). Radware Page. <https://www.radware.com/security/ddos-knowledge-center/ddospedia/irc-internet-relay-chat/>
- Slater, D. (1998). Trading sexpics on IRC: Embodiment and authenticity on the internet. *Body & Society*, 4(4), 91–117. <https://doi.org/10.1177/1357034x98004004005>
- Srnicek, N. (2020). *Platform capitalism*. Polity.
- Stommel, W. (2016). Information giving or problem discussion? formulations in the initial phase of web-based chat counseling sessions. *Journal of Pragmatics*, 105, 87–100. <https://doi.org/10.1016/j.pragma.2016.09.001>
- Stommel, W. J., & Stommel, M. W. (2021). Participation of companions in video-mediated medical consultations: A microanalysis. *Analysing Digital Interaction*, 177–203. https://doi.org/10.1007/978-3-030-64922-7_9
- Stommel, W. J., van Goor, H., & Stommel, M. W. (2020). The impact of video-mediated communication on closed wound assessments in postoperative consultations: Conversation analytical study. *Journal of Medical Internet Research*, 22(5). <https://doi.org/10.2196/17791>
- Stommel, W., & Lamerichs, J. (2016). Empathy displays in Dutch chat counselling: Showcasing a microanalysis of online data. *Applying Linguistics in Illness and Healthcare Contexts*. <https://doi.org/10.5040/9781350057685.0014>

- Stommel, W., & te Molder, H. (2015). When technological affordances meet interactional norms: The value of pre-screening in online chat counseling. *PsychNology Journal*, 13(2-3), 235-258.
- Stommel, W., & van der Houwen, F. (2013). Formulations in “trouble” chat sessions. *Language@Internet*, 10(3).
- Stommel, W., & Van der Houwen, F. (2016). Counseling and new media technologies: A comparison of problem presentations in e-mail and in Chat. *Communication and Medicine*, 12(2–3). <https://doi.org/10.1558/cam.18298>
- Stommel, W., Licoppe, C., & Stommel, M. (2020). “difficult to assess in this manner”: An “ineffective” showing sequence in post-surgery video consultation. *Social Interaction. Video-Based Studies of Human Sociality*, 3(3). <https://doi.org/10.7146/si.v3i3.122581>
- Stommel, W., Paulus, T. M., & Atkins, D. P. (2017). “Here’s the link”: Hyperlinking in service-focused chat interaction. *Journal of Pragmatics*, 115, 56–67. <https://doi.org/10.1016/j.pragma.2017.02.009>
- Stommel, W., van Goor, H., & Stommel, M. (2019). Other-Attentiveness in Video Consultation Openings: A Conversation Analysis of Video-Mediated Versus Face-to-Face Consultations. *Journal of Computer-Mediated Communication*, 24(6), 275–292. <https://doi.org/10.1093/jcmc/zmz015>
- Tang, J. C., Kivran-Swaine, F., Inkpen, K., & Van House, N. (2017). Perspectives on live streaming. *Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. <https://doi.org/10.1145/3022198.3022200>
- Tang, J. C., Venolia, G., & Inkpen, K. M. (2016). Meerkat and Periscope. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/2858036.2858374>

- Taylor, T. L. (2018). *Watch me play: Twitch and the rise of Game Live Streaming*. Princeton University Press.
- Thorne, S. L., & Hellermann, J. (2022). Coda: The interactional affordances and constraints of technology-rich teaching and Learning Environments. *Classroom Discourse, 13*(2), 231–239. <https://doi.org/10.1080/19463014.2022.2071959>
- Total Twitch streamers*. TwitchTracker. (2023). <https://twitchtracker.com/statistics/active-streamers>
- Tuncer, S., Brown, B., & Lindwall, O. (2020). On pause: How online instructional videos are used to achieve practical tasks. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. <https://doi.org/10.1145/3313831.3376759>
- Tuncer, S., Lindwall, O., & Brown, B. (2020). Making time: Pausing to coordinate video instructions and practical tasks. *Symbolic Interaction, 44*(3), 603–631. <https://doi.org/10.1002/symb.516>
- Uskokovic, B., & Talehgani-Nikazm, C. (2022). Talk and embodied conduct in word searches in video-mediated interactions. *Social Interaction. Video-Based Studies of Human Sociality, 5*(1). <https://doi.org/10.7146/si.v5i2.130876>
- Vallis, R. (1999). Members' methods for entering and leaving IRCbar: a conversation analytic study of internet relay chat. In K. Chalmers, S. Bogitini, & P. Renshaw (Eds.), *Educational Research in New Times*. Flaxton, Australia: PostPressed.
- Virtanen, M. T., & Niemi, J. (2023). Thanking and positive assessments in video-mediated workshops: Managing creativity exercises remotely. In *Conversation Analytic perspectives to digital interaction: Practices, resources, and affordances* (pp. 208-237). Suomalaisen Kirjallisuuden Seura.

- Whittaker, S. (2003). Theories and methods in mediated communication. In: A. Graesser, M. Gernsbacher & S. Goldman (Eds.), *The handbook of discourse processes* (pp. 253–293). Erlbaum.
- Wagner, K. (2015, March 26). *Twitter launches Livestreaming app Periscope -- its own version of meerkat*. Vox. <https://www.vox.com/2015/3/26/11560692/twitter-launches-livestreaming-app-periscope-its-own-version-of>
- Weisz, J. D., & Kiesler, S. (2008). How text and audio chat change the online video experience. *Proceedings of the 1st International Conference on Designing Interactive User Experiences for TV and Video*. <https://doi.org/10.1145/1453805.1453809>
- Weisz, J. D., Kiesler, S., Zhang, H., Ren, Y., Kraut, R. E., & Konstan, J. A. (2007). Watching together. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/1240624.1240756>
- Wu, L., Liu, Q., Zhao, J., & Lank, E. (2023). Interactions across displays and space: A study of virtual reality streaming practices on Twitch. *Proceedings of the ACM on Human-Computer Interaction*, 7(ISS), 242–265. <https://doi.org/10.1145/3626473>
- Wu, Y., Li, Y., & Gui, X. (2022). “I am concerned, but...”: Streamers’ privacy concerns and strategies in live streaming information disclosure. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW2), 1–31. <https://doi.org/10.1145/3555104>
- Zheng, S., Wu, M., & Liao, J. (2022). The impact of destination live streaming on viewers’ travel intention. *Current Issues in Tourism*, 26(2), 184–198. <https://doi.org/10.1080/13683500.2022.2117594>
- Zouinar, M., & Velkovska, J. (2017). Talking about things: Image-based topical talk and intimacy in video-mediated family communication. *Pragmatics*, 27(3), 387-418.

LIST OF PUBLICATIONS

- **JOURNAL**

Song, L., & Licoppe, C. (2024). Noticing-based actions and the pragmatics of attention in expository live streams. noticing ‘effervescence’ and noticing-based sequences. *Journal of Pragmatics*, 226, 1–16. <https://doi.org/10.1016/j.pragma.2024.04.001>

Song, L., & Licoppe, C. (2023). Closing live video streams: A sequential analysis. *Discourse, Context & Media*, 53, 100698. <https://doi.org/10.1016/j.dcm.2023.100698>

Song, L., & Licoppe, C. (2024). The organization of live streaming openings (under review)

- **CONFERENCES IN COMPUTER SCIENCE AND ROBOTICS**

Le Song (2024): “Oh my Dave, your face says it all!”: Exploring the Multimodal Practices of Tasting in Live Streaming (major revision, the ACM International Conference on Mobile Human-Computer Interaction 2024)

Le Song and Zhegong Shangguan (2024): The Moment That The Driver Takes Over: Examining trust in full self-driving in a naturalistic and sequential approach. In: Proceedings of the 22nd European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centered Computing on the Design of Cooperation Technologies - Exploratory papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.48340/ecscw2024_ep04

Le Song and Zhegong Shangguan (2024): Multimodal Practices to Sustain Multiactivity When Live Streaming. The ACM International Conference on Interactive Media Experiences (IMX), Stockholm, Sweden, 12-14 June, 2024, <https://doi.org/10.1145/3639701.3663639>

Le Song and Christian Licoppe (2024): The “Silent Tasting Moment”: Withdrawing Gaze from the Screen as an Interactional Achievement When Tasting Food in Live Streaming. The Seventeenth International Conference on Advances in Computer-Human Interactions (ACHI), 26-30 May, 2024. Barcelona, Spain

Le Song and Mei Li (2024): Trainspotting and Sharing: Talking and Waiting To See Passing Trains in A 24-Hour Railway Station Live Stream, 26th ACM International Conference on Multimodal Interaction, (4-8 Nov 2024) (under review)

Zhegong, Shangguan, Yang Liu, **Le Song**, Tingcheng Li, and Adriana Tapus, 2024. Using a Pneumatic Tactile Steering Wheel to Enhance the Multi-Modal Takeover Request In Smart Vehicle. In Proceedings of 16th International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications (AUTOMOTIVEUI '24). ACM, California, CA, USA, 16 pages (under review)

IEEE RO-MAN 2024 Workshop (member of organizers)

Ethics Challenges in Socially Assistive Robots and Agents: Legality, Value Orientation, and Future Design for Human-Robot Interaction (HRI), California, USA, August 2024

(Organizers: Adriana TAPUS, Jim Tørresen, Roger Andre Søråa, Yu Cheng, **Le Song**, Zhegong Shangguan)

- **CONFERENCES IN MEDIA, COMMUNICATION, AND LINGUISTICS**

Digital Meeting for Conversation Analysis 2022, October 31-November 4, 2022 Online

Conference paper: Chinese Directives and Action Coordination in Mobile Interaction
Mei Li, **Le SONG**, Citing Li

The 72nd International Communication Congress (ICA) Paris, France, May 26-30, 2022
Extended abstract: Exploring the Multimodal Organization of Tasting Sequences in Live Video Streaming
Le SONG & Enhua GUO

The 17th International Pragmatics Conference (2021) June 27-July 2, 2021 Winterthur, Switzerland (online)
Abstract: The Organization of Closing Sequences in Periscope-like Live Video Streams
Le SONG & Christian Licoppe

- **PUBLICATIONS IN CHINESE**

汉语移动交际中的指令性行为研究，语言学研究，2024，修改录用
李梅，宋乐，李茨婷
(Chinese Directives in Mobile Interaction, Linguistic Research, 2024
Mei Li, **Le SONG**, Citing Li)

《翻转摄像头：手机直播中的“感官传递”与“空间建构”》
获2021年上海交大-国际传播学会（ICA）新媒体论坛“最佳学生论文”并入选《2021年新媒体国际论坛》正式出版文集
宋乐，甘雨梅
(2021 SJTU-ICA New Media Forum October 2021 Shanghai, China, online
Best Student Paper Award, Flip Camera: Sensory Transmission and Space Construction in Mobile Live Broadcasting
Le Song & Yumei Gan)

Titre: Pratiques interactionnelles multimodales dans les diffusions en direct sur Twitter

Mots clés: diffusion en direct, analyse conversationnelle multimodale, multimodalité, interaction homme-ordinateur, interaction médiatisée, chat médiatisé par ordinateur, technologie vidéo

Résumé: En tant que forme émergente d'interaction médiatisée, le streaming en direct est devenu une pratique en pleine expansion. Cette thèse de doctorat se concentre sur les flux en direct en tant que phénomènes interactionnels d'un point de vue séquentiel. En adoptant l'ethnométhodologie et l'analyse de la conversation (EMCA) comme perspective théorique et méthodologique, la thèse explore comment l'utilisation de multiples ressources et la manipulation des affordances des dispositifs permettent d'atteindre différentes actions conjointes. La dissertation se compose de quatre principaux articles de recherche. Tous les articles ont été publiés ou sont en cours d'examen. L'article I examine les ouvertures du streaming en direct. Contrairement aux conversations téléphoniques avec une séquence d'ouverture canonique, les ouvertures de flux en direct apparaissent plus variées, avec des cadres de participation laminés. L'article II traite de la manière dont les streamers et les spectateurs gèrent l'attention et l'engagement à travers des actions basées sur la notation. Il examine comment les streamers et les spectateurs produisent des séquences de notation et des séquences basées sur la notation, et comment l'orientation vers la notation peut conduire à une forme distinctive d'« effervescence de notation ». L'article III inspecte l'activité de dégustation dans le streaming en direct, réexaminant la dégustation dans cette écologie particulière comme un processus interactif qui combine une expérience sensorielle individuelle avec une dimension publique, observable et intersubjective. L'article IV examine l'organisation des séquences de clôture dans le streaming en direct. Il montre que bien que les participants s'orientent vers l'organisation séquentielle des clôtures dans les conversations ordinaires, ils le font d'une manière qui est sensible aux affordances des flux en direct. La thèse fournit ainsi une analyse systématique des propriétés interactionnelles les plus caractéristiques du streaming en direct.

Title : Multimodal Interactional Practices in Live Streams on Twitter

Keywords : live streaming, multimodal conversation analysis, multimodality, computer-human interaction, mediated interaction, computer-mediated chat, video technology

Abstract: As an emerging form of mediated interaction, live streaming has become a rapidly growing practice. This doctoral dissertation focuses on live streams as interactional phenomena from a sequential perspective. Taking ethnomethodology and conversation analysis (EMCA) as its theoretical and methodological perspective, the thesis explores how the use of multiple resources and the manipulation of affordance of the devices in achieving different joint actions. The dissertation consists of four main research articles. All of the articles have been published or are under review. Article I investigates the openings of live streaming. Unlike phone conversations with a canonical opening sequence, live stream openings appear more variable, with laminated participation frames. Article II discusses how streamers and viewers manage attention and engagement through noticing-based actions. It looks at how streamers and viewers produce noticing sequences and noticing-based sequences, and how the orientation towards noticing may lead to a distinctive form of 'noticing effervescence.' Article III inspects the activity of tasting in live streaming, re-examining tasting in this particular ecology as an interactive process that combines individual sensory experience with a public, witnessable, and intersubjective dimension. Article IV investigates the organization of closing sequences in live streaming. It shows that while participants orient to the sequential organization of closings in ordinary conversation, they do so in a way that is sensitive to the affordances of live streams. The thesis thus provides a systematic analysis of the most characteristic interactional properties of live streaming.