

This is a "Post-Print" accepted manuscript, which has been published in

*Journal of Contemporary Ethnography.*

Please cite this publication as follows:

Pehkonen, Samu. Choreographing the Performer–Audience Interaction. *Journal of Contemporary Ethnography*. 2017, 46(6), 699-722. Article first published online: March 17, 2016. DOI: <https://doi.org/10.1177/0891241616636663>

You can download the published version at:

<https://doi.org/10.1177/0891241616636663>

### **Choreographing the Performer–Audience Interaction**

#### **Abstract**

The article introduces the concept of choreography, defined as situationally enacted participation and action framework that provides sequential structure for social interaction, for studying performer–audience interaction during musical performances. Performers develop a preferred type of interaction during a repeated series of concerts. Audiences become absorbed in the choreography through participation in the concerts and the circulation of the Internet videos from earlier concerts. As the audience learns to expect certain actions from the side of the performers, improvisation is required from the performers in order for the choreography to be successful. Attention is paid to the methods the performers use to produce “watchables” and to manage the audience responses. The spatial, temporal, and gestural elements of this enacted choreography are

analyzed sequentially using conversation analysis. The longitudinal data is composed of YouTube concert videos of Kings of Convenience performing a song, “I’d Rather Dance with You.”

### **Keywords**

interaction, choreography, joint action, audience response, pop music concert

### **Introduction**

Seoul, South Korea, April 2008. Erlend Øye and Eirik Glambek Bøe, two Norwegian singer-songwriters in their thirties from the folk-pop duo Kings of Convenience (KoC) count off their most famous song, “I’d Rather Dance with You,” a catchy up-tempo tune with seductive easy-going lyrics. The song is immediately orchestrated by the crowd of screaming girls in the audience. On an amateur concert video, screaming, cheering, and clapping systematically accompany the song until it ends. Having witnessed several videos of the band performing the same song in various countries over almost ten years, this event does not strike me as particularly noteworthy. After all, that is what Western popular culture has been about ever since Elvis and Beatlemania. The aim to manage emotional and bodily experience in live music concerts (Wicke 1990) and to establish a homogenized global culture (Frith 1996) has influenced further studies in the sociology of music, popular culture, and performativity (Bennett 2008; Stebbins 2007).

For sure, KoC, often referred to as Norway's answer to Simon and Garfunkel, operates firmly in the tradition of Western pop music. Erlend and Eirik play acoustic guitars and sing in harmony, occasionally accompanied by other musicians. The band, however, has the reputation of being a surprising live act when it comes to interacting with the audience. Sometimes they suddenly come up with absurd rules of conduct for audience behavior (for example, clapping during but also after a song may be prohibited but humming is allowed). They may engage in a lengthy conversation with their audience or pretend performing without any audience. This type of choreographed performer–audience interaction takes place also within a single song: The audience responses during “I’d Rather Dance with You” range from silence to loud cheering.

There is a distinctive choreography at work, defined as *situationally enacted participation and action framework that provides sequential structure for performer–audience interaction*.

How is the choreography of a particular performance enacted and played out successfully time after time in various parts of the world? While I assume that most of the South Korean audience members have never been to a KoC concert before, they are still capable of interacting according to the finely-tuned script. The screams and shouts in Seoul are not random but take place in the same routine order as a couple of months later in Helsinki where I witnessed the band playing live. Listening to my audio recording from the Helsinki concert, I observed how the audience members—including

myself—unintentionally played our part in the type of interactional game that seemed as important for the participants as the music that orchestrated the audience behavior.

Using the YouTube concert videos as a member-produced ethnographic online resource (Hallett and Barber 2014; Wittel 2000) and ethnomethodologically observable data (ten Have 2002; Laurier 2015), I describe how participants in a live concert coordinate their actions and interactions moment to moment and movement by movement and how they constitute the choreography as a joint practical achievement (Clark 1996). By enacting the choreography, individual participants become a collective audience responsive to the particular performer and the performance witnessed. The routinized interaction between the audience and the performer becomes predictable and the performer can use a knowledgeable audience to produce—and if needed, to modify—the expected choreography. The article suggests that learning “doing being fans” takes place not only through participation in the unfolding event of the performance (Sanders 1974) but also through member-produced artifacts circulating in the spaces of the Internet (Nieckarz 2005).

The concept of choreography will be developed as an alternative framework to study performer–audience interaction. In an ethnomethodological fashion (Garfinkel 2002), choreography highlights the orderliness of membership-bound competences that often escape from the generic spectacle-oriented analyses of live music concerts. The longitudinal empirical data set and multimodal interaction analysis (e.g., Broth 2011;

Broth and Keevallik 2014; Haviland 2011; Weeks 1990) is then introduced and discussed from the point of view of online ethnography. The results from the empirical analysis are discussed in two sections. The first summarizes the data indicating the existence of an established choreography. The second investigates the main methods used by the participants to contribute to the enactment of this choreography. The article concludes with a discussion on the ways in which the YouTube concert videos per se make certain aspects of interaction available for ethnographic inspection.

### **Musical interaction and Choreography**

The pop concert is a social gathering organized around the musical performance prepared and delivered by the musicians to the audience. These performances have been studied from various viewpoints ranging from the cultural ramifications of the global music industry (Auslander 2006; Frith 1996) to the individual emotions and perceptions of the audience and musicians (Doğantan-Dack 2012; Thompson, Graham, and Russo 2005). Yet, live popular music concerts also offer a versatile social environment to study how less ostentatious, mundane interactional achievements such as reward applauses are successfully put together in situ (Barkhuus and Jørgensen 2008; Kurosawa and Davidson 2005, 113).

I treat a live pop music concert as a social situation in which, according to Goffman (1964, 135), the concert acts as an environment of mutual monitoring possibilities and

the participants find themselves accessible to others present and vice versa. This mutual accessibility is created and sustained by performers and audience members who use their bodies to produce audible and visible cues (cheering, whistling, clapping, and gesturing) to maintain the flow in the situation. Previous research on musical interaction has contributed—from various methodological perspectives including mainly surveys, interviews and experiments—in important ways to our understanding of how performers accomplish playing music together (Weeks 1990; Gratier 2008; Veronesi 2014), how they communicate with the audience (Broughton and Stevens 2009; Kurosawa and Davidson 2005; Camurri et al. 2004), and how audiences evaluate musical performances (Platz and Kopiez 2013). There are, however, certain biases: Classical music and jazz performances are studied rather than pop music (Weeks 1990; Gratier 2008). There is also an emphasis on bodily techniques for playing musical instruments (Clayton 2005) or conducting (Parton 2014; Veronesi 2014), and when broader gestural analysis is conveyed, it is often concerned with how the musician extends his or her feelings toward the audience or represents the inner emotions written in the piece of music (Davidson 2006; Kurosawa and Davidson 2005; Moran 2013). My study, rooted in ethnomethodological conversation analysis, contributes to the existing research by offering a sequential investigation into the ways in which the performer– audience interaction and “doing being fans” are enacted in situ.

To address these interactional elements, the article applies an understanding of choreography that deviates from its common association with professional composing and arranging dances in advance. The word choreography occasionally appears in social scientific literature, usually pointing toward something pre-scripted and acted out more or less unconsciously in interactional settings (Aronsson 1998; Crossley 1996; Tulbert and Goodwin 2011) or used as an explicit performative political strategy (Foster 1998). Choreography can indeed be useful in bringing together diverse methodological discussions and making sense of various phenomena in the social sciences, but as a concept, I suggest, choreography must be organized around three criteria—situational enactment, multimodality, and membership-bound knowledge—to become an analytical tool that helps understand and describe the physical form and movement of bodies.

First, we need to acknowledge that choreographies consist of both the expected set of mutual actions and the improvised use of interactive methods in enacting those choreographies. Working from the ethnomethodological tradition, Whalen, Whalen, and Henderson (2002) study the organization of the work practices of call center employees as “improvisational choreography,” a notion that resolves some of the problems in the ways choreography is used as an analytical tool. “While choreographed action is commonly understood to mean a carefully arranged or directed sequence of steps and movements,” that is, following a pre-determined structure, and improvisation is seen about composing extemporaneously or fabricating “out of what is conveniently on

hand,” choreography and improvisation are not separate realms of action (Whalen, Whalen, and Henderson 2002, 241). For example, a singer who forgets the lyrics can improvise new ones in order to save the situation. However, these improvised lyrics are not any random words hanging in the air, but words picked up and made to be understood as improvised lyrics that still carry on the ongoing action of singing (Friedwald 2002, 98). Understanding “how [spaces], technologies and other resources can be carefully arranged to afford what must necessarily be a somewhat extemporaneous composition” (Whalen, Whalen, and Henderson 2002, 241) is a task that can be grasped by studying social events from the point of view of choreography.

Second, choreographies are constitutive of not just language but also of bodily, spatial, and material arrangements, an array of situatedly mobilized resources as pointed out in the conversation analytical literature on multimodality (Streeck, Goodwin, and LeBaron 2011). Foster (1998, 5–6) states that choreography “serves as a useful intervention into discussions of materiality and body by focusing on the unspoken, on bodily gestures and movements that, along with speech, construct [hybrid identities].” The spatial component is evident in the etymological origins of the Greek word *choreo*, meaning the being in, passing, entering to, or holding space. Choreography refers to “the writing of practices of being (corporeally) in space and inhabiting space” (Puumala and Pehkonen 2010, 54). Particular social situations (such as live pop concerts) stand out from other situations (such as karaoke or a drama festival) as identifiable because of



their distinct social, historical, and spatial arrangement. The same holds true with musical subgenres in which different behavior is expected from the audience regarding, for example, the physical distance between the stage and the floor (Simon 1997). The material basis of interaction (Goodwin 2000) allows performers to use their instruments not only to produce sounds but also to make visible their bodies toward the audience.

Third, choreography is about the epistemic and normative (Parton 2014) practices people deploy to form and maintain the identifiability and flow of actions. These practices rest partly upon individual skills but also upon interactional techniques, mundane members' methods (Garfinkel 1967) or grounding (Gratier 2008) to convey those skills as well as upon cultural conventions that govern the roles members find available to themselves in a given situation. While choreography builds on expected roles to be followed, in situ, reflexive actions and face-work (Goffman 1967) are required to maintain those roles. The smooth workings of choreography are based on knowledgeable participants who are able to explicate their knowledgeability. Any breach in showing this (a wrong note, misplaced laughter, or a clap) is a potential normative violation which may result in corrective turns in interaction. Choreography is, ultimately, about communication based on the membership-bound knowledge between participants.

## **Data and Method**

The data were collected in 2012 from YouTube and consisted of a total of 60 live video recordings of the song “I’d Rather Dance with You.” Twelve concert videos from ten different countries were chosen to cover the time span from the promotion of the song (three videos from 2004–2005) to the time the song had earned its reputation as the climax of the show (nine videos from 2008–2012). The two first performances were filmed and edited for television production and include multiple camera angles that periodically also capture the audience’s visual reactions and actions. All the other videos were filmed by members of the audience, usually with a cell phone. As a result, the quality is low compared to the professional footage since there are few possibilities to adjust the lighting and sound conditions while filming. Further, the fixed camera position provides no accurate visual information about all audience reactions. Thus, the data do not comprise all communication channels used in the interaction between the audience members and the performers.

In light of cyber-ethnographic research (Hallett and Barber 2014; Hine 2000), these do-it-yourself videos provide a rich and easily accessible repository of research material (Laurier 2015) into the membership-bound practices of fans. First, the unedited videos shared on YouTube capture the temporal, spatial, and sequential continuities of the live concert in a manner edited or often even ethnographic footage fails to do (Mondada 2006). This is because unedited videos naturally capture the event, showing what the

particular audience member finds worth filming and sharing with others, not from an ethnographer's prescribed or problem-oriented point of view (Strangelove 2010).

Second, if the audience member chooses to use the zoom or follow certain actions on the stage, these adjustments indicate the relevance of those sequences for the community of fans. Zooming on gyrating hips, for example, makes those movements available for the viewers of the videos as something that should be viewed with special attention. Third, the use of mobile devices to capture the concert is a common practice that makes it possible to include in the data multiple video recordings of the same event, often filmed from multiple angles thus resulting in more information about the context. This also resolves some of the problems a sole researcher finds when having to make decisions in advance about what to focus on while filming. Finally, YouTube videos are part of a continuously changing circulation of videos, comments and discussions that offers ethnographers with a valuable resource for the membership-bound practices.

My analysis began with the process of transcribing the twelve selected videos for sequential organization of the performances and the witness-able interaction (Mondada 2006). For this purpose, ELAN (EUDICO Linguistic Annotator) software was used, making it possible not only to annotate the precise occurrence and duration of the actions of performers and audience responses but also to compare the course of events in different performances. First, the position of actions and gestures performed on stage in relation to the song structure was measured. Then, the occurrence and level of

audience responses were depicted and timed in order to determine recurrent patterns of action as well as changes in the interaction over a selected time span. This resulted in a general understanding of the choreography where performers deployed certain gestures to solicit audience response. Especially relevant for intensive and affiliating audience participation were the beginning and end of the song as well as transitions to the chorus and solo parts, since these are the natural occasions for recognition and reward applauses and cheering (Barkhuus and Jørgensen 2008, 2927). More interestingly, the interaction in the C-part and the following A-part called for detailed analytical attention since this choreography showed negotiation on the actions audience and performers were supposed to engage in. These two sequences will be further described here with the help of conversation and multimodal analysis of interaction that aims “to document the precise ways in which talk, gesture, gaze, and aspects of the material surround are brought together to form coherent courses of action” (Stivers and Sidnell 2005, 1).

I pay attention to the methods used by the performers to produce *watchables*, the main targets of audience attention within a more complicated social situation. Instead of “just” gyrating, the performers produce watchables by explicating certain actions, such as “look, I’m gyrating!” Watchables are an effective resource in managing the level of audience reaction. While audience reactions in pop concerts are less controlled and the etiquette is looser than in classical music or theatre performances (Broth 2011), the reactions are still far from chaotic and unordered; instead, they tend to follow the

general rules of turn-taking in conversation such as “one party talks at a time” and “speaker-change recurs” (Sacks, Schegloff, and Jefferson 1974, 706), albeit not as straightforwardly as in face-to-face conversations. In the context of live music, these rules are typically, but not always, acknowledged by the audience placing its loud collective affiliating response so that it does not overlap with singing and by the parties acknowledging each other when they organize their interactional turns. I will first represent a sequence (from the second A-part) that shows this general tendency to regulate audience reaction as well as evidence for rule-breaking among audience and performers. Second, I focus on the preceding sequence (C-part) in search of evidence for the locally acceptable grounds for breaking the rule. I then analyze the fine-tuning choreography sequentially in terms of its spatial, temporal, and gestural elements in the case of a single set of repetitive sample data. The process of analysis and the results are illustrated with the help of still images, summary tables, and a traditional conversation analytic transcript.

### **Choreographies of Audience Response: Sequence 1**

In the longitudinal data, the choreography for the song is kept constant, with only minor variations in each concert. Variations occur mainly because the song is sometimes performed by the duo and sometimes features a violin player or a complete band. The song structure (ABCAB typical of Western pop music) remains in all the performances

studied, but changes in how many bars are played within each part are made. The song is started by Eirik playing the first accompaniment beats with guitar. He stands behind the fixed microphone stand throughout the whole song. He looks down or at the audience and only occasionally gazes at Erlend, who operates mainly on Eirik's right side. Erlend enters the front stage to sing the first verse, after which he typically moves around the stage at a pace ranging from lazy trailing to energetic dancing. The two contrasting bodies, Eirik's guitar playing but otherwise passive presence versus Erlend's directive presence, clearly make Erlend the main point of focus for the audience. When the audience reacts to the performance, it does so primarily to Erlend's actions.

After the first chorus (B-part) and an instrumental bridge, there is a C-part, originally recorded as a four-bar verse sung by Eirik in a smooth, even whispering voice: "The music is too loud, and the noise from the crowd / increases the change of misinterpretation." In live performances, the C-part is repeated by Eirik and Erlend, who joins in to repeat the two first bars in harmony. As soon as Eirik has finished the verse, Erlend takes over (and if accompanied by other musicians, they join in) singing the second A-part, "So let your hips do the talking," and producing 'his' famous hip movements (owing, of course, these gyrating moves to Elvis). He continues, "I'll make you laugh by acting like a guy who sings" and illustrates these seductive lyrics with a grooming gesture (possibly referring to Garfunkel's hair) placing his hand on the back of his head. In the data, there is variation in the way the gyrating and grooming gesture

is performed, but these actions are clearly produced as watchables, making them distinct from rhythmic movements or any random stroke of hair, for example.

As illustrated in Table 1, in eight out of the twelve concert videos analyzed, the audience reacts vocally to Erlend's gyrating hips and his grooming gesture. In two concerts (videos 1 and 6), the grooming gesture and audience reaction are missing. Further, in two concerts either the grooming (video 3) or the gyrating hips (video 12) gesture is performed, but the audience does not produce an audible reaction quickly enough to be qualified as a reaction to those particular gestures. Whereas typically the audience reaction sequences do not overlap with singing (except when the audience sings along), these particular reactions are sequentially ordered in a special manner. Gyrating hips and the grooming gesture seem to call for immediate reaction from the audience, and therefore, cheering, whistling, or shouting overlaps the accompanying singing. However, the general rule of turn allocation still holds as the duration of the overlap is minimized and the reaction to the first action fades as soon as the next line in the lyrics is performed. There is an established choreography of the performer–audience interaction in which the relatively fixed occurrence of Erlend's gyrating moves and grooming gesture are followed by distinctive audience reactions placed as minimally overlapping with the singing.

**#Table 1 here**

The timing of reactions varies over the period analyzed. In the early set of concerts (three videos), the reactions are somewhat spontaneous rewards by the individual audience members. In later concerts, Erlend's gestures are anticipated by the audience members as a group; as a result, the audience response becomes a coordinated action (Atkinson 1984, 371). There is about a 5-second timespan, corresponding three bars in the musical notation (#note1), for the audience to react to Erlend's gyrating hips: the first 3 seconds while Erlend is singing the phrase "So let your hips do the talking" and the remaining 2 seconds before the next phrase starts and the new gesture is set in motion. Measuring from the outset of the gyrating and grooming moves, in the first set of videos, the first audible audience responses to the gyrating start on average 3.95 seconds and for grooming 2.38 seconds (video 2) after the gestures. For gyrating, the response takes the form of single audience members' shouting or laughing gradually growing into collective cheering as more and more audience members join in. Reaction times are considerably slower than those measured in Atkinson's (1984, 372)

pioneering study on applause in public speeches where there is typically no gap between the end of the talk and the onset of the audience response as isolated whistles, shouts of "yeah," and claps are quickly taken over by collective clapping, or in Mann et al.'s (2013, 2) data on applauding oral student presentations in which the mean duration for the first person to begin clapping was measured as 2.1 seconds. Thus, the long reaction time and the isolated responses (or no response at all), although correctly placed, before



the collective burst of cheering indicate ambivalence about to what and how the audience should respond in relation to the performance.

In contrast to the first three videos, in the second set of videos (2008–2012), the audience engages in collective cheering that is mostly spot on, intensive response starting rapidly (the first reaction on average 1.61 seconds for gyrating and 1.54 seconds for grooming). The most active and receptive audiences, in South Korea and Mexico, seem to be aware of the gestures since the loud cheering peaks within a second. Even when the collective response is lacking, as in videos 10 and 11 for the gyrating, individual audience members show their anticipation of the oncoming actions by placing their shouts right after Erlend has started to move his hips. Interestingly, collective cheering as a form of expressing audience anticipation replaces other displays of affiliation, most notably clapping in rhythm. This is in contrast to Atkinson's data on applauding during speeches. Unlike vocal responses, clapping is not substantially constrained by breathing and thus quickly reaches maximum intensity (Atkinson 1984, 372). In live concerts, however, the audience can already be clapping, which leaves shouts, whistles, and other forms of vocal cheering the only available cues to make distinctive rewards. The predominance of vocal responses also partly explains the short duration of the response: The response lasts on average less than 4 seconds versus 8 seconds measured in the data by Atkinson (1984, 374) and 6 seconds in Mann et al. (2013, 2). Although the length of audience response is partly limited by the musical

flow, another reason for the short duration of response has to do with audience anticipation: The audience cuts off its vocal response just in time to make space for the next action and to minimize overlap with singing. Erlend's gyrating and grooming gestures are treated by the audience as two distinctive acts requiring separate feedback.

This section has shown that repeating the same choreography from concert to concert produces expectations for future interactions. The reaction times of the audible audience feedback are generally slower in the early concerts compared with more recent ones although there are exceptions to this general trend as the 2012 concert in Norway includes both the weakest reaction (only individual) to gyrating and the slowest reaction time for grooming (#note3). For example, while in a 2004 concert Erlend's hip dance movements resulted in audible individual audience reaction only 4.15 seconds after the action had started, in the sequence from the 2011 concert in Mexico, the equivalent time measure for collective response is 0.37 seconds. In the latter case, it is evident that the audible cheers and whistles for Erlend's hip dance are produced before they can reasonably be anticipated: Producing a laugh or scream after something that one could not expect takes a considerably longer time than to produce the same reaction to something that was already known to be on its way (Harbidge 2011).

In addition, the audience members' recordings illustrate the learning process and point out the potential of these artifacts for producing and recreating a community of fans. Especially in the videos that systematically aim at producing "a semi-professional"

concert video, the use of zoom indicates the expectations the audience member holds about the next action in the choreography. For example, zooming in on Erlend's hips some seconds before the actual hip movements (Figure 2) locally accompanies and projects the next possible action (Mondada 2006). By making certain actions watchable to viewers of the video, the videos become a medium for learning the choreography *ex situ*.

**#Figure2 here**

### **Choreographing Audience Responses: Sequence 2**

Thus far, I have described how the audience responds to the performance; but why do audiences respond in the first place to those particular gestures since not all gestures elicit an audible collective response? I return to a sequence of about 12 seconds taken from the C-part. Here Erlend and Eirik utilize and remold the song structure and its composition to produce a delayed peak in the performance. By delayed peak, I mean the transition from the first bars sung by Eirik in a soft voice—requiring a certain level of quiet from the audience to be properly heard—to the moment when Erlend's exaggerated hip movements and grooming gestures release the built-up tension as described in the previous section. Erlend's dance and posture serve as an indirect invitation for audience feedback and have proved to be an effective method for teasing out the preferred (as judged from performers' rhythmic nodding as an indication of

agreement) method of interaction without giving the audience direct verbal instructions. The natural flow of the sequence and the seemingly tight performer–audience interaction is, however, choreographed in terms of the space where it happens, the bodily and gestural methods it relies on, and the timing when it happens. The analysis is illustrated with still images from two live performances (Mexico City on 1 November 2011; Seoul on 15 April 2008) and conversation analysis transcriptions (Mexico City on 1 November 2011) depicting the choreographic production at particular points in relation to the lyrics and the performers taking turns (#note2).

### *Spacing*

The physical space of a live concert provides specific contextual resources for the participants (Haviland 2011; Mondada 2007). For their performance, the stage is typically arranged in a manner that allows Erlend and Eirik to be standing or seated behind fixed microphone stands with their guitars, close to each other at the center-front of the stage: Eirik on the left-hand side of the stage and Erlend on his right side. The microphone stands play a key role as elements of the contextual configuration (Goodwin 2000), not only by making the performers' singing and playing audible to the audience via a sound reinforcement system but also by fixing their appearance on the stage. Further, the audience has a fixed point of focus to concentrate on: Erlend and Eirik are seen as an equal duo. During the song "I'd Rather Dance with You," this

configuration is deliberately breached. Whereas Eirik plays accompanying guitar standing in one fixed place, Erlend takes off his guitar and typically removes the microphone from the stand, in this way offering the audience information about the forthcoming change in musical style and content (see Kurosawa and Davidson 2005, 125). The guitar arrangement makes it possible for Eirik to continue playing the introduction until Erlend decides to start singing. This opens up the stage to a more tactful interaction between the artists, as the rearrangement showcases Erlend as the lead vocal. The fact that the C-part is sung by Eirik, however, requires a collaborative effort to make Eirik watchable instead of Erlend, and to do this at just the right moment. Two sets of spatial techniques are used to accomplish this turn allocation.

The first and most frequently used strategy includes Erlend walking toward Eirik during the instrumental bridge and then stopping near him and pointing at him with his index finger (Image 1 in Figure 1). This act is reinforced by Erlend minimizing his presence in the performative space of the stage by gradually descending to the floor level (Image 6 in Figure 2). This is a clear indication to the audience that their focus should be on Eirik singing the next verse.

The second strategy works in an opposite manner. Here, during the instrumental bridge, Erlend moves away from the center-front of the stage to some remote location on the stage, leaving Eirik alone to occupy the central position. From an interactional point of view, however, this second strategy has disadvantages, because simply moving

aside does not tell the audience who or what should be their new point of focus. In the data, there are instances where the camera “erroneously” follows Erlend while Eirik is about to become temporarily the main character on the stage. The moving-away strategy thus does not have the capacity to maintain choreographic flow and to build up expectations to the forthcoming peak point when Erlend takes over to produce his hip movements. In the first strategy, during the two bars sung in harmony, Erlend stands up (Image 7 in Figure 2) and walks away from Eirik in order to produce enough space for his moves (compare Erlend’s location in Images 4 and 5 in Figure 1) and to make himself the main watchable. Again, the first strategy more successfully prepares the audience for what is coming next than the second strategy does.

Erlend’s “hip dance” (Image 5 in Figure 1)—varying from delicate seductive hip moves from side to side at a fixed stage location to a more extensive use of the space available for dancing around the stage—and the hair grooming pose (Image 7 in Figure 2)—representing “a guy who sings”—constitute the interactional peak of the performance. At this point the audience produces the strongest vocal reactions. It also marks a turning point in the performance: After this point, the self-generated and spontaneous audience participation—be it rhythmic clapping, shouting, or whistling—is not regulated by the performers.

## *Gesturing*

The central idea behind multimodality in interaction is that interactional order is produced and maintained not only verbally, but other modes of meaning-making (gaze, gesture, bodily arrangement, etc.) are equally important, especially in music performances when “the vocal” is largely reserved for expressing “the lyrics.” As mentioned, the existing literature on music and gesture does not always acknowledge gestures for their practical purposes intimately bound up with managing a particular situation (Laurier and Philo 2006). Some of the gestures evident here fall between *expressive* gestures conveying affective or emotional domain and *natural* gestures supporting verbal communication (McNeill 1992). These gestures also maintain the flow of the choreography, “the coordinated embodied actions of people and their perspectives upon the material, real-world setting within which they interact” (Streeck 2009, 5). The interactive work through gesturing is presented in the transcription (below, reference is made to specific lines in the transcription in Figure 1) of a sequence from the 2011 Mexico City concert.

### **#Figure 1 here**

The transcript starts from the beginning of the C-part, the sequence in which the floor is given to Eirik to sing a solo. For the performers, this type of turn allocation is embedded in the way the song is performed, but indicating the next speaker by pointing at him with an index finger is an effective method to make also the audience aware of the coming turn-taking. In line 1, Erlend stands still and points at Eirik while looking

down as part of an act to make himself irrelevant to the ongoing action (Image 1 in Figure 1). Erlend does not retract his pointing hand right away but instead produces three strikes in the air in sync with the beat and tempo of the song. While this type of rhythmic motion (nodding, foot tapping) is typical in experiencing and presenting music in general and could thus be interpreted as a routine part of the stage performance, in this case the motion has an interaction organizing end: It allows Erlend to maintain a position that neither puts him in the central position (looking down) nor makes him completely irrelevant (maintaining the rhythm) to the forthcoming action. This position on stage is held until another set of gestures occurs.

Another frequently used set of gestures by Erlend and Eirik is their coordinated request for the audience to stop singing along during the C-part. This request has appeared only within the last several years as a reaction to their audiences increasingly knowing the lyrics by heart and singing along. Singing along, however, produces a possible problem, as can be seen in the sample sequence.

Eirik—and the audience—start to sing the first bars of the sequence (Line 1). Right after having played a chord on the first beat of the second bar, Eirik lifts his right hand just a little above his guitar and leaves it off there for a half a second (Image 2; Line 1). Then, at the point of possible chord production (Line 2), Eirik seems to retract his gesture but maintains the same gesture for another half a second. Since this set of gestures seems meaningless, from the point of view of playing guitar (technical gesture)



and expressing the lyrics (interpretive gesture), it becomes understood as an instruction to the audience to restrain its level of participation. Note, however, that at this point there is no significant change in the audience participation: Some are still singing along.

When the verse is about to come to an end in line 2, Erlend and Eirik lift their free hands up to produce gestures almost similar to Eirik's first gesture (Image 3). Now, Erlend holds his left hand up at shoulder level. His palm is tilted forward toward the audience as an indication of "not yet." Eirik holds his right index and middle finger up at shoulder level. Holding up two fingers like this (and when interpreted with the timing of this gesture), Eirik indicates to the audience that the verse will be repeated. Erlend and Eirik collaboratively instruct the audience: Not yet, the verse will be repeated. This action is needed because typically the verse is repeated in live performances but not in the recorded version. In fact, some of the audiences collectively produce the wrong lyrics at the end of line 2 ("SO-").

When the early concerts are compared with more recent ones, the instructional segment has also modified Eirik's guitar playing. Originally, Eirik played two down strokes on each beat of the bar throughout the C-part. Now, to release his hand for instructing the audience, he can play only a whole note or a chord at the beginning of the bar. Thus, by singing along (too loudly), the audience unwillingly and accidentally also affects the musical outcome.

## *Timing*

Instructions for proper audience participation are partly a result of the fact that the audience has learned the recorded version by heart but not the live performance.

Although the audience members might be familiar with previous concert videos and thus aware of the oncoming watchables, the precise timing of their response need still to be locally accomplished. Looking at the timing of the spatial and gestural methods Erlend and Eirik use to instruct the audience before the repeated verse, one notices that the performers in fact anticipate these practical problems.

First, we look at the timing of the first “instructive” gesture in line 1. It is produced by Eirik, who, when singing solo, can be appreciated as having his turn and being the object of the audience’s focus. Eirik’s first gesture appears right after he produces the second chord in the verse. Producing the gesture before that would be too patronizing a move, leaving no chance for the audience to control its level of participation autonomously. In addition, had a similar gesture been produced first by Erlend, it would have breached the image of a harmonious duo, as the leading role had just been given to Eirik at the beginning of the C-part.

Second, the simultaneous apexes of their gestures are produced just before the repeated verse, in fact while they utter the word “interpretation” as the last word of the first verse (end of line 2), giving the audience a chance to follow the instructions.

Again, the timing of instructions maintains the harmonious interaction between the performers and the audience.

Third, the fact that no further instructive gestures are produced by the performers is relevant. The audience has sufficiently stopped singing along by the time the repeated verse “and the noise from the crowd” is reached in line 3. Eirik still raises his right hand up, but as he holds his hand close to his chest with the fingers together, he does not “extend” this gesture to the audience. Thus, this gesture becomes an interpretative gesture that allows him to repeat the same rhythmic pattern in his guitar playing. In turn in line 4, Erlend uses this moment to create space for initiating his coming “hip dance” that ultimately releases the tension and the audience bursts out screaming.

## **Conclusions**

Following the lines of ethnomethodological and conversation analytic literature, the choreographic analysis builds on members’ knowledge acquired in this case through online ethnographic inquiry. It brackets any “easy” background information that might explain why certain audiences react the way they do while others do not react at all.

Although it might be legitimate to question, for example, whether the type of concert (is the Kings of Convenience the only act of the show or just another band in the list of a festival program), the geographical proximity (why one of the weakest audience response is from the concert in their home country Norway) or the level of subcultural

awareness of musical norms (does the audience consist of “outsiders” or passionate fans) affect the behavior of the audience, these questions cannot be answered based on the existing data and methods. What the ethnomethodological approach to ethnographic data gives is a rich repertoire of practices and seen but unnoticed features of performer–audience interaction. Treating a concert as interactional choreography shows the existence of an interactional pattern (the audience reacts to certain actions that performers produce as watchables), the members’ resources to produce and account for the interaction, as well as the picking up of the choreography by individual members who through their practices of filming reproduce the choreography for the online community of possible concert-goers. It shows that even seemingly disorderly conduct in live concerts is produced in an orderly fashion and is a result of mutual learning.

Choreography should not be taken as a final product since learning also affects the enacted choreography. As pointed out, the audience singing along knows the song by heart in the form of the recorded version, not the live version. As Auslander (2006, 160) states, “the primary experience of the music is as a recording; the function of the live performance is to authenticate the sound on the recording.” The major difference in the case analyzed here lies in the repeated verse in the C-part. Instead of acknowledging the performers’ preparation to repeat the verse, the audience moves on to sing the next line “So let your hips...” When this false start is noticed by the audience, they stop singing altogether. This feature, the withdrawal by silence, also occurs in mundane

conversations when a speaker notices he or she has produced talk at an inappropriate place (before a possible turn transition point at the end of a completed phrase: Sacks, Schegloff, and Jefferson 1974). As in ordinary conversation, these mistakes are harmless and usually go unnoticed by the participants themselves (Schegloff, Jefferson, and Sacks 1977). However, the mistakes affect the choreography as was the case here with the performers attempting to solve the expected problem in advance.

Being a competent audience member and showing it to others at a live concert is a skillful interactional achievement orchestrated by the performers. Participants use interactional methods to manage where the audience's attention should be (watchables) and the form their reactions should take. It was shown how space, time, and gestures are deployed to produce and maintain the choreography in the case of a single but repeated performance. The micro-level interactionist perspective followed here is useful in correcting some of the overtly spectacle-oriented accounts of musical performances by taking notice of the ordinary, mundane elements that build up interaction in social situations. Treating audience reactions and responses as interactively produced acts also redirects the analytical emphasis away from dwelling on individual musical experiences toward understanding the members' methods and resources of witnessing music concerts as collective, shared phenomena in the Durkheimian sense of collective effervescence (Tyler 2008).

Using the concept of situationally enacted choreography instead of performativity based on subcultural norms and dwelling on the collective action rather than individual capacities help us understand how competences and expectations are played out in social situations. Although the artists from the Kings of Convenience might just have remolded traditions of popular music performance to suit a style of their own, the means for maintaining this performer–audience interaction can be and are used by other performers, although detailed analyses are difficult to find for comparison (see Davidson 2005). In this respect, future research should concentrate on single repetitive performances as they would allow the analyst to see not just a random outcome but also the process and development of the planned but situationally followed choreography.

Using the musical archives on Internet sites such as YouTube offers ethnographers and social scientists time-saving access to data to collect observations that serve as a starting point for further analyses of musical interaction as well as community production.

### **Endnotes**

#note1 The use seconds for measuring the audience responses is for analytical and comparative purposes only, whereas the participants themselves orient primarily to the musical flow measured in bars. There is a crucial difference between responses within musical performance compared with political speeches: the speaker can hold back his or

her next utterance to avoid overlap while in the musical performance the performers typically cannot affect the length of overlapping audience response.

#note2 Data have been transcribed according to slightly modified versions of conversation (Sacks, Schegloff, and Jefferson 1974) and multimodal analysis (Mondada 2006) conventions:

Gestures and actions' descriptions are delimited between two identical symbols (++ for Eirik, \* \* for Erlend, §§ for audience) and are synchronized with the corresponding sound times.

>> gesture or action described begins before the excerpt's beginning

+---> gesture or action described continues across subsequent lines

---->+ until the same symbol is reached

.... gesture's preparation

---- gesture's apex

,,,,,, gesture's retraction

↗↑ arrows indicate the approximate direction of a hand motion or gesture (e.g., R h ↓ indicates right hand, downward stroke)

♪ a note indicates a gesture needed to produce music

erl participant doing gesture is identified when he is not singing

im the exact point at which the image was taken is indicated with a number sign # showing its position within a turn.

#note3 A larger data set would be needed to make this prediction statistically relevant. The exceptional cases would also serve as starting point for further inquiries: Are the delayed or missing responses due to insufficiently produced watchables or ineffective learning process? These questions are beyond the bounds of the current paper.

## References

- Aronsson, Karin. 1998. "Identity-in-Interaction and Social Choreography." *Research on Language and Social Interaction* 31 (1): 75–89.
- Atkinson, J. Maxwell. 1984. "Public Speaking and Audience Responses: Some Techniques for Inviting Applause." In *Structures of Social Action: Studies in Conversation Analysis*, edited by J. Maxwell Atkinson, and John Heritage, 370–409. Cambridge: Cambridge University Press.
- Auslander, Philip. 2006. "Liveness. Performance and the Anxiety of Simulation." In *The Popular Music Studies Reader*, edited by Andy Bennett, Barry Shank, and Jason Toynbee, 85–91. New York: Routledge.
- Barkhuus, Louise, and Tobias Jørgensen. 2008. Engaging the Crowd – Studies of Audience–Performer Interaction. In *CHI 2008 Proceedings*, 2925–30.
- Bennett, Andy. 2008. "Towards a Cultural Sociology of Popular Music." *Journal of Sociology* 44 (4): 419–32.



- Broth, Matthias. 2011. "The Theatre Performance as Interaction between Actors and Their Audience." *Nottingham French Studies* 50 (2): 113–33.
- Broth, Matthias, and Leelo Keevallik. 2014. "Getting Ready to Move as a Couple. Accomplishing Mobile Formations in a Dance Class." *Space and Culture* 17 (2): 107–21.
- Broughton, Mary, and Catharine Stevens. 2009. "Music, Movement and Marimba: An Investigation of the Role of Movement and Gesture in Communicating Musical Expression to an Audience." *Psychology of Music* 37 (2): 137–53.
- Camurri, Antonio, Barbara Mazzarino, Matteo Ricchetti, Renee Timmers, and Gualtiero Volpe. 2004. "Multimodal Analysis of Expressive Gesture in Music and Dance Performances." In *Gesture-based Communication in Human-Computer Interaction, GW 2003, LNAI 2915*, edited by Antonio Camurri, and Gualtiero Volpe, 20–39. Berlin/Heidelberg: Springer-Verlag.
- Clark, Herbert H. 1996. *Using Language*. New York: Cambridge University Press.
- Clayton, Martin. 2005. "Communication in Indian Raga Performance". In *Musical Communication*, edited by Dorothy Miell, David Hargreaves, and Raymond MacDonald, 361–81. Oxford: Oxford University Press.
- Crossley, Nick. 1996. *Intersubjectivity: The Fabric of Social Becoming*. Thousand Oaks: Sage.

- Davidson, Jane W. 2005. "Bodily Communication in Musical Performance." In *Musical Communication*, edited by Dorothy Miell, David MacDonald, and Raymond Hargreaves, 215–37. Oxford: Oxford University Press.
- Davidson, Jane W. 2006. "'She's the One': Multiple Functions of Body Movement in a Stage Performance by Robbie Williams." In *Music and Gesture*, edited by Anthony Gritten and Elaine King, 208–25. Aldershot: Ashgate.
- Doğantan-Dack, Mine. 2012. "The Art of Research in Live Music Performance." *Music Performance Research* 5:34–48.
- Foster, Susan Leigh. 1998. "Choreographies of Gender." *Signs: Journal of Women in Culture and Society* 24 (1): 1–33.
- Friedwald, Will. 2002. *Stardust Melodies: The Biography of Twelve of America's Most Popular Songs*. New York: Pantheon Books.
- Frith, Simon. 1996. *Performing Rites: On the Value of Popular Music*. Cambridge: Harvard University Press.
- Garfinkel, Harold. 1967. *Studies in Ethnomethodology*. Englewood Cliffs: Prentice-Hall.
- Garfinkel, Harold. 2002. *Ethnomethodology's Program. Working Out Durkheim's Aphorism*. Lanham: Rowman & Littlefield Publishers.
- Goffman, Erving. 1964. "The Neglected Situation." *American Anthropologist* 66 (6): 133–6.

- Goffman, Erving. 1967. *Interaction Ritual. Essays on Face-to-face Behavior*. New York: Doubleday.
- Goodwin, Charles. 2000. "Action and Embodiment within Situated Human Interaction." *Journal of Pragmatics* 32 (10): 1489–522.
- Gratier, Maya. 2008. "Grounding in Musical Interaction: Evidence from Jazz Performances." *Musicae Scientiae* (Special Issue 2008):71–110.
- Hallett, Ronald E., and Kristen Barber. 2014. "Ethnographic Research in a Cyber Era." *Journal of Contemporary Ethnography* 43 (3): 306–30.
- Harbidge, Lesley. 2011. "Audienceship and (Non)laughter in the Stand-up Comedy of Steve Martin." *Participations. Journal of Audience & Reception Studies* 8 (2): 128–43.
- ten Have, Paul. 2002. "The Notion of Member is the Heart of the Matter: On the Role of Membership Knowledge in Ethnomethodological Inquiry." *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* 3 (3). Available at: [www.qualitative-research.net/index.php/fqs/article/view/834/1812](http://www.qualitative-research.net/index.php/fqs/article/view/834/1812) (accessed 10 August 2015).
- Haviland, John B. 2011. "Musical Spaces." In *Embodied Interaction: Language and Body in the Material World*, edited by Jürgen Streeck, Charles Goodwin, and Curtis LeBaron, 289–304. New York: Cambridge University Press.
- Hine, Christine. 2000. *Virtual Ethnography*. London: Sage.

- Kurosawa, Kaori, and Jane W. Davidson. 2005. "Nonverbal Behaviours in Popular Music Performance: A Case Study of The Corrs." *Musicae Scientiae* 9 (1): 111–36.
- Laurier, Eric. 2015. "YouTube: Fragments of a Video-tropic Atlas." *Area*. Available at: <http://onlinelibrary.wiley.com/doi/10.1111/area.12157/pdf> (accessed 26 February 2015).
- Laurier, Eric, and Chris Philo. 2006. "Cold Shoulders and Napkins Handed: Gestures of Responsibility." *Transactions of the Institute of British Geographers* 31 (2): 193–207.
- Mann, Richard P., Jolyon Faria, David J. T. Sumpter, and Jens Krause. 2013. "The Dynamics of Audience Applause." *Interface* 10 (85). Available at: <http://dx.doi.org/10.1098/rsif.2013.0466> (accessed 20 March 2014).
- McNeill, David. 1992. *Hand and Mind: What Gestures Reveal about Thought*. Chicago: University of Chicago Press.
- Mondada, Lorenza. 2006. "Video Recording as the Reflexive Preservation and Configuration of Phenomenal Features for Analysis." In *Video Analysis. Methodology and Methods. Qualitative Audiovisual Data Analysis in Sociology*, edited by Hubert Knoblauch, Bernt Schnettler, Jürgen Raab, and Hans-Georg Soeffner, 51–68. Bern: Peter Lang.
- Mondada, Lorenza. 2007. "Multimodal Resources for Turn-Taking: Pointing and the Emergence of Possible Next Speakers." *Discourse Studies* 9 (2): 195–226.

- Moran, Nikki. 2013. "Music, Bodies and Relationships: An Ethnographic Contribution to Embodied Cognition Studies." *Psychology of Music* 41 (1): 5–17.
- Nieckarz, Peter P., Jr. 2005. "Community in Cyber Space? The Role of the Internet in Facilitating and Maintaining a Community of Live Music Collecting and Trading." *City & Community* 4 (4): 403–23.
- Parton, Katharine. 2014. "Epistemic Stance in Orchestral Interaction." *Social Semiotics* 24 (4): 402–19.
- Platz, Friedrich, and Reinhard Kopiez. 2013. "When the First Impression Counts: Music Performers, Audience and the Evaluation of Stage Entrance Behavior." *Musicae Scientiae* 17 (2): 167–97.
- Puumala, Eeva, and Samu Pehkonen. 2010. "Corporeal Choreographies Between Politics and the Political: Failed Asylum Seekers Moving from Body Politics to Bodyspace." *International Political Sociology* 4 (1): 50–65.
- Sacks, Harvey, Emanuel A. Schegloff, and Gail Jefferson. 1974. "A Simplest Systematics for the Organization of Turn-Taking in Conversation." *Language* 50 (4): 696–735.
- Sanders, Clinton R. 1974. "Psyching Out the Crowd: Folk Performers and Their Audiences." *Journal of Contemporary Ethnography* 3 (3): 264–82.

Schegloff, Emanuel A., Gail Jefferson, and Harvey Sacks. 1977. "The Preference for Self-Correction in the Organization of Repair in Conversation." *Language* 53 (2): 361–82.

Simon, Bradford Scott. 1997. "Entering the Pit: Slam-dancing and Modernity." *The Journal of Popular Culture* 31 (1): 149–76.

Stebbins, Robert A. 2007. "The Sociology of Entertainment." In *21st Century Sociology: A Reference Handbook*, edited by Clifton D. Bryant and Dennis L. Peck, 178–85. Thousand Oaks, CA: Sage.

Stivers, Tanya, and Jack Sidnell. 2005. "Introduction: Multimodal Interaction." *Semiotica* 156 (1): 1–20.

Strangelove, Michael. 2010. *Watching YouTube: Extraordinary Videos by Ordinary People*. Toronto: University of Toronto Press.

Streeck, Jürgen. 2009. *Gesturecraft. The Manu-facture of Meaning*. Amsterdam/Philadelphia: John Benjamins.

Streeck, Jürgen, Charles Goodwin, and Curtis LeBaron. 2011. "Embodied Interaction in the Material World: An Introduction." In *Embodied Interaction: Language and the Body in the Material World*, edited by Jürgen Streeck, Charles Goodwin, and Curtis LeBaron, 1–26. Cambridge: Cambridge University Press.

Thompson, William Forde, Phil Graham, and Frank A. Russo. 2005. "Seeing Musical Performance: Visual Influences on Perception and Experience." *Semiotica* 156 (1): 203–27.

Tulbert, Eve, and Marjorie H. Goodwin. 2011. "Choreographies of Attention.

Multimodality in a Routine Family Activity." In *Embodied Interaction: Language and Body in the Material World*, edited by Jürgen Streeck, Charles Goodwin, and Curtis LeBaron, 79–92. Cambridge: Cambridge University Press.

Tyler. 2008. "Feeding the Animals: Collective Effervescence and Girl Talk." Available at: <https://vimeo.com/2502303> (accessed 20 February 2015).

Veronesi, Daniela. 2014. "Correction Sequences and Semiotic Resources in Ensemble Music Workshops: The Case of Conduction®." *Social Semiotics* 24 (4): 468–94.

Weeks, Peter. 1990. "Musical Time as a Practical Accomplishment: A Change in Tempo." *Human Studies* 13 (4): 323–59.

Whalen, Jack, Marilyn Whalen, and Kathryn Henderson. 2002. "Improvisational Choreography in Teleservice Work." *British Journal of Sociology* 53 (2): 239–58.

Wicke, Peter. 1990. *Rock Music: Culture, Aesthetics and Sociology*. Cambridge: Cambridge University Press.

Wittel, Anreas. 2000. "Ethnography on the Move: From Field to Net to Internet." *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* 1 (1).

Available at: [www.qualitative-research.net/index.php/fqs/article/view/1131/2518](http://www.qualitative-research.net/index.php/fqs/article/view/1131/2518)

(accessed 1 June 2014).



Nr	Place and date	Gyrating hips			Grooming hair		
		Form of audible audience reaction	Response time in seconds	Length of reaction in seconds	Form of audible audience reaction	Response time in seconds	Length of reaction in seconds
1	Germany, Oct. 2004	individual laugh & whistles collective clapping	4.15 (ind.) 5.69 (col.)	1.34 (ind.) 8.50 (col.)	no gesture, no reaction	-	-
2	USA, Feb. 2005	individual shouts	4.81 (ind.)	1.38 (ind.)	collective cheering clapping and laughing	2.38 (col.)	2.78 (col.)
3	Brazil, Oct. 2005	individual shouts collective cheering	2.89 (col.)	3.76 (col.)	no reaction	-	-
	Average: early set of videos	individual collective first reaction	4.48 4.29 3.95	1.36 6.13	collective	2.38	2.78
4	S. Korea, Apr. 2008	collective cheering	1.22 (col.)	4.09 (col.)	collective cheering	0.67 (col.)	3.34 (col.)
5	S. Korea, Apr. 2008	collective cheering	0.22 (col.)	4.52 (col.)	collective cheering	1.10 (col.)	4.51 (col.)
6	Sweden, Oct. 2009	individual shouts collective cheering	3.81 (ind.) 4.57 (col.)	0.46 (ind.) 3.78 (col.)	no gesture, no reaction	-	-
7	Philippines, Mar. 2010	collective cheering	2.42 (col.)	3.76 (col.)	individual shouts collective shouting	1.84 (ind.) 2.54 (col.)	0.67 (ind.) 2.36 (col.)
8	USA, Jun. 2010	collective cheering	2.05 (col.)	1.80 (col.)	collective cheering	1.68 (col.)	3.31 (col.)
9	Mexico, Nov. 2011	collective cheering	0.37 (col.)	4.78 (col.)	collective cheering	1.02 (col.)	6.48 (col.)
10	Norway, Apr. 2012	individual shouts	0.43 (ind.)	0.51 (ind.)	individual laughs collective cheering	2.81 (ind.) 5.06 (col.)	1.73 (ind.) 4.64 (col.)

11	Spain, Jun. 2012	individual shouts	2.32 (ind.)	1.56 (ind.)	collective cheering	1.63 (col.)	3.23 (col.)
12	Portugal, Jun. 2012	no reaction	-	-	individual shouts collective cheering	2.32 (ind.) 3.31 (col.)	0.52 (ind.) 2.68 (col.)
	Average: later set of videos	individual collective first reaction	2.19 1.81 1.61	0.84 3.78	individual collective first reaction	2.33 1.95 1.54	1.20 3.98

**Table 1.** Audience reactions to Erlend's gyrating hips and grooming gesture.

Transcription 1 KOC9 2.24-2:36

1 EIR #the +mus+ic's too \*loud\* and the +noi+se +from the #crowd\*  
+JR h ↓+ +JR h ↓+ +R h ↑ →palm pointing down

erl >>looking down, L h pointing at EIR, \*strikes ↗ in the air\*  
\*↗\* \*↗\* \*↗\*

aud THE MUSIC'S TOO LOUD AND THE NOISE FROM THE CROWD  
im # still 1 # still 2

2 EIR \*in+creases the chance of +mis+inter+pretation  
+R h ↑+ +JR h ↓+ +R hand ↑ 2 fingers up  
+.....+ +.....+ +.....+ +.....+ +.....+ +.....+

erl \*turn L, 2 steps toward EIR\* \*head up\* \*L hand ↑ tilted forward >  
\*.....+ \*.....+ \*.....+ \*.....+ \*.....+ \*.....+

aud INCREASES THE CHANCE OF MISINTERPRETA- SO  
im # still 3

3 EIR The+ mus+ic's too loud and the +noi+se from the +crowd  
+JR h ↓+ +JR h ↓+ +R h ↑ fingers together  
+.....+ +.....+ +.....+ +.....+

ERL The music is too loud and the noi\*se from the crowd  
then index finger up pointing at AUD\*  
..... ↗..... ↗..... ↗..... both hands drop down\*

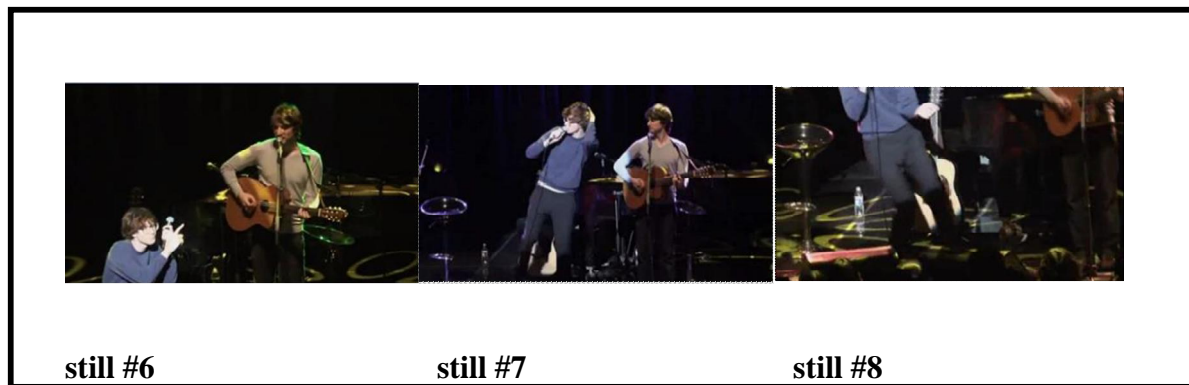
aud LET YOUR HIPPS (WOHOO) \*and the noise from the crowd\*  
im # still 4

4 EIR \*Increases the cha+nce of +mis+interpret+ation  
+JR h ↓+ +JR h ↓+ +.....+ +.....+

erl \*turns R, walks away, camera doesn't following, ERL's actions not available>  
aud \*increases the chance of misinterpretation\*

5 ERL So +let your §hips do the \*talking (2.4) §#  
until the camera refocus> \*legs wide apart, dancing moves  
eir +continues playing two R hand strokes per beat>  
aud so let your. § SCREAMING >>  
im # still 5

**Figure 1.** Transcription of “I’d Rather Dance with You” by Kings of Convenience, as performed in Mexico City on 1 November 2011. Still images taken from two videos (1–3 from [http://www.youtube.com/watch?v=-8ro-iiBZ\\_Q](http://www.youtube.com/watch?v=-8ro-iiBZ_Q) and #4-5 from <http://www.youtube.com/watch?v=DEh0lFK4I2E>).



**Figure 2.** Kings of Convenience performing “I’d Rather Dance with You” in Seoul, South Korea on 15 April 2008 (Still images 6–8 from <http://www.youtube.com/watch?v=xOtc3BqqoFg>).