

Preprint of the book chapter:

Jensenius, A. R., & Erdem, C. (2022). Gestures in ensemble performance. In R. Timmers, F. Bailes, & H. Daffern (Eds.), *Together in Music: Coordination, Expression, Participation*. Oxford University Press.

## Gestures in ensemble performance

Alexander Refsum Jensenius and Çağrı Erdem

### Abstract

Gestures, defined as meaning-bearing bodily actions, play important and varied roles in ensemble performance. This chapter discusses how the term “gesture” differs from physical “motion” and perceived “action.” The functional differences between sound-producing, sound-facilitating, sound-accompanying, and communicative actions are presented, alongside how these can be performed and/or perceived as meaning-bearing gestures. The role of gestures in ensemble performance is examined from four perspectives: (1) ensemble size and setup; (2) the musical degrees of freedom of the ensemble; (3) the musical leadership; and (4) the role of machines in the musicianship. It is argued that the use of gestures varies between different types of ensembles and musical genres. The common denominator is the need for meaning-bearing bodily communication between performers, with such gestures also playing an important part in the musical communication with the audience.

Keywords: gesture; ensemble; meaning; sound-producing action; communication

### Introduction

The topic of *gesture* has received growing attention among music researchers over recent decades. Some of this research has been summarized in anthologies on “musical gestures,” such as those by Gritten and King (2006), Godøy and Leman (2010), and Gritten and King (2011). There have also been a couple of articles reviewing how the term *gesture* has been used in various music-related disciplines (and beyond), including those by Cadoz and Wanderley (2000) and Jensenius et al. (2010). Much empirical work has been performed since these reviews were written, aided by better motion capture technologies, new machine learning techniques, and a heightened awareness of the topic. Still there

are a number of open questions as to the role of gestures in music performance in general, and in ensemble performance in particular. This chapter aims to clarify some of the basic terminology of music-related body motion, and draw up some perspectives of how one can think about gestures in ensemble performance. This is, obviously, only one way of looking at the very multifaceted concept of gesture, but it may lead to further interest in this exciting and complex research domain.

## From motion to gesture

Within this chapter, gesture is referred to as the *meaning* related to a perceived and/or performed action. In this context, “meaning” should not be understood as “meaningful,” but rather in the sense of communicating “something.” In a musical context, this could be information about the tempo, such as seen in the beat patterns of conductors. It could also be the expressive swaying of the upper body of a violinist in a string quartet, indicating the phrasing to her fellow musicians. How others experience such gestures are, of course, highly dependent on prior experience and cultural background. Thus, gesture is by definition a subjective term. This differentiates it from the related—but different—terms *motion*, *force*, and *action*. Motion refers to the physical displacement of an object in time and space, while force refers to the push or pull experienced in interaction with other objects. Both motion and force refer to physical phenomena, and can be studied objectively with various types of sensing devices (see, for example, Jensenius, 2018) for an overview of different methods for sensing music-related body motion). Motion and force are also interrelated: applying force to an object can set it in motion, and the motion of an object can lead to the experience of force.

While motion and force are terms that are drawn from the world of physics, and more precisely (bio)mechanics, we reserve action to describe the psychological experience of motion and force. An action can be understood as the *chunking* of continuous motion or force into what Godøy and Leman (2010) refer to as “cognitive units.” Such a unit is a piece of information that is held consciously in our focus of attention. Actions are often goal-oriented, that is, we think about them with respect to a particular goal: lifting a glass, opening a door, playing a key on a piano. It is usually easy to identify the goal, but it is more difficult to describe precisely when an action begins or ends. This has some implications for how we conceptualize and analyze actions. For example, think about an individual drum stroke, in which a drumstick is lifted and then dropped to hit the membrane of the drum. It is straightforward to identify when the excitation happens, but it can be very hard to say when the action began. This becomes more complex when multiple actions are

combined into action series, which leads to *coarticulation*, the merging of individual actions into larger shapes of actions (Godøy et al., 2010). A challenge from an empirical research perspective is that a motion capture recording will only inform about the continuous displacement of markers attached to a musician's body. It is non-trivial to segment such a continuous motion stream into actions, since this will rely on the perspective from which one is looking, the temporal and spatial resolution, and so on. Defining the meaning-bearing components of such actions—the gestures—relies on yet another level of abstraction. The reason we emphasize these differences between motion/force, action, and gesture, is that they are sometimes used synonymously in the literature. This causes confusion and reduces the power of the term gesture.

## *Gesture*

The power of the term gesture is that it goes beyond motion. In a linguistic context, it usually denotes bodily actions associated with speech, or what Kendon (2004, p. 7) referred to as “visible action as utterance.” McNeill (1992) showed how hand motion and facial expressions do not just randomly accompany speech, but are an integral part of the communication itself. He classified gesture into five functional categories:

- *Iconics* represent a particular feature of an object and can be described in terms of the shape and spatial extent of the gesture. Iconic gestures are often used to illustrate an action, for example imitating a knocking movement with a hand while saying “knocking on the door;”
- *Metaphorics* are similar to iconics but represent an abstract feature of an object. An example of a metaphoric gesture could be to say “something happened” while holding up the hands to refer to “something;”
- *Beats* occur together with spoken words to highlight discontinuities and stress specific words. Beats are typically carried out as in/out or up/down movements, such as a nod, and may be seen as emphasizing the most important words in a narrative;
- *Deictics* indicate a point in space, for example pointing in a specific direction while saying “over there;”
- *Emblems* are stereotypical patterns with agreed meaning, such as the goodbye or OK sign.

To explain the relationships between gesture and speech, McNeill (1992, p. 37) outlined what he calls the *Kendon continuum*. This continuum goes from *gesticulation* on one end, in which gestures always co-occur with speech, to *sign language* on the other end, in which the gestures are linguistically self-contained. In between are the two cases of what he calls *emblems* and *pantomime*. Similar relationships between action and sound can be found in the case of *musical* gestures. Gestures linked to the sound-producing actions of musicians, for example, are strictly related to musical sound. A conductor's gestures, on the other hand, can take on many different functions along such an imagined continuum.

### *Musical gesture*

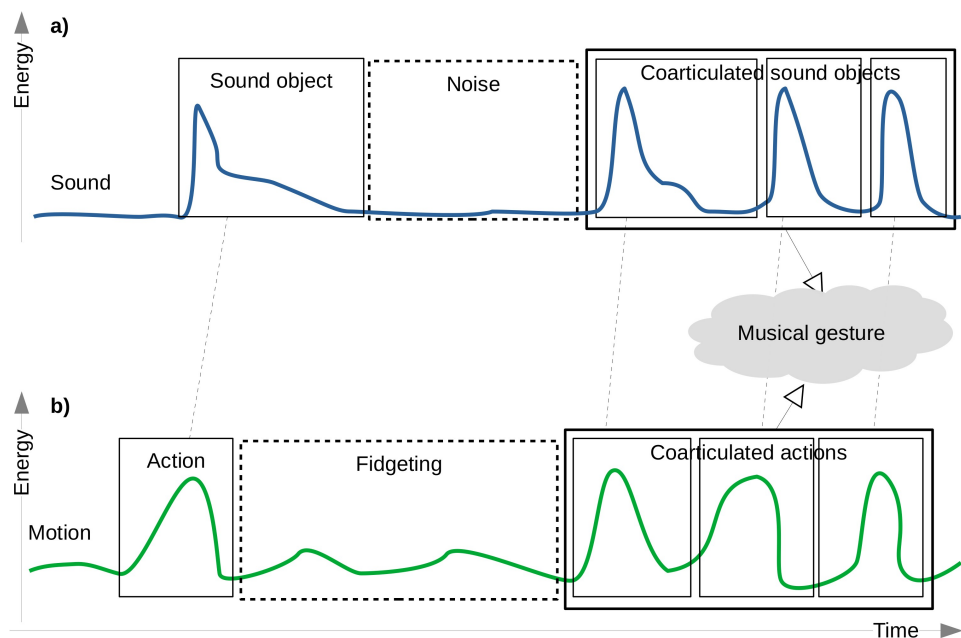
How can we think about gestures from a musical perspective? The term musical gesture has, over the years, been used in quite different ways. One approach is that of Hatten (2004, p. 95) who argues that a musical gesture is “significant energetic shaping through time.” He uses gesture in a metaphorical sense to describe motion-like qualities in the sound of music. This is quite different from the way the term is used to describe music-related body motion in some empirical music research. One definition that manages to combine these perspectives well—that is, between thinking about musical gesture as primarily related to sonic properties *or* to body motion—is the one presented by Gritten and King (2006, p. xx):

[A] gesture is a movement or change in state that becomes marked as significant by an agent. This is to say that for movement or sound to be(come) gesture, it must be taken intentionally by an interpreter, who may or may not be involved in the actual sound production of a performance, in such a manner as to donate it with the trappings of human significance.

This definition implies that there is a flow of communication between the performer and the perceiver, and that the performer's motion “becomes” a gesture only if it is understood as such by the perceiver.

An interesting question then arises concerning consciousness: does an action have to be carried out consciously to be experienced as a gesture? Following the argument of Gritten and King (2006), gestures could be performed unconsciously but still be valid if they are observed as significant by the perceiver. From these definitions and theories, it is clear that gesture is a highly subjective phenomenon.

In summary, the term musical gesture is related to both motion and sound (the physical), as well as actions and sound objects (the perceptual), as sketched in [Figure 14.1](#). In some cases, the experience of a musical gesture may be driven primarily by sound, other times primarily by motion. In many cases, however, the combination of motion and sound leads to the experienced gesture.



*Figure 14.1 A visual summary of how a musical gesture can be thought of as the combination of experienced sound objects (a) and actions (b). These actions and sound objects are perceived from the continuous stream of sound (a) and motion (b).*

### *Functional categories*

When it comes to understanding more about motion/force, action, and gesture in ensemble performance, it is helpful to examine different types. Jensenius et al. (2010) suggested dividing music-related motion into four categories: sound-producing, sound-facilitating, sound-accompanying, and communicative. It is only the fourth of these—the communicative—that could be categorized as gestural by definition. The three others can also be considered as gestures, but only if there is a meaning-bearing component expressed by the performer and/or experienced by the perceiver. For example, when a pianist hits a key with the finger, it involves motion and force and it can be experienced as a goal-directed action. However, it is not necessarily a gesture. It can be a gesture if the

action is performed with a particular type of expressivity, or there is some other communicative element to the action. One such example could be a pianist playing the final chord of a piece with a dramatic action, hence signaling that this is the end of the piece. Such a musical gesture has a clear communicative element and is based on both sound-producing and sound-facilitating actions.

## Gestures in ensembles

Ensemble performance is an excellent case for exploring musical gestures. After all, it is necessary to communicate to play together, and this often involves a combination of audition and vision. The type of communication, however, varies considerably, depending on several factors. In the following, we will investigate this from four perspectives: the *size and setup* of the ensemble, the *degrees of freedom* the musicians have over the music being played, how the *musical leadership* is distributed among musicians, and whether or not *machines* are involved in the musicianship.

### *Ensemble size and setup*

Keller (2014) argues that interpersonal synchrony and leader–follower relations in ensembles are established using primarily head nodding, body swaying, and gaze patterns. As such, they provide “visual cues” that support the coordination within the ensemble. He further argues that if co-performers are denied visual contact, the performers increase their level of body sway to regulate the performance timing. Such gesturing, however, depends to a large extent on the size and setup of the ensemble in question, as this directly impacts the level of attention and details in the gestural communication.

The setup of an ensemble is related to what Jensenius et al. (2010) refer to as the performance *scene* (the volume over which the musicians are spread), their performance *position* within that scene, and their gesture *space* related to that position. The performance scene may be visually defined in the form of a concert hall stage, but it could also be a socially constructed area in the middle of a busy city street. The musicians’ positions are defined within the boundaries of that performance scene. It may be tight and fixed, such as in the case of a string quartet sitting close together, even when they perform on a large stage. The musicians’ positions may also be large and flexible, as seen in the way members of a rock band move around on a large stage. Thus, the gesture space of the musicians, that is, the physical volume they have at their disposal from their performance position, varies a lot.

The gesture space for members of a string quartet is constrained to a small “box” around their chair position, while it could cover the entire stage (and beyond) for rock musicians. These two extremes influence the different types of gestures one would typically see during such performances. A string quartet sitting relatively close to each other in a semi-circle makes it possible to pick up subtleties of each other’s performance actions. This allows the musicians to use small gestures in their within-ensemble communication, in the form of a raised eyebrow, or a minute twist of the bowing arm.

### *Musical degrees of freedom*

Another parameter that regulates the type and level of gestures found in ensemble performance is what may be called the “musical degrees of freedom” of the performance, that is, how much of the performance each individual musician is controlling. In this context, freedom does not mean “freedom in speech,” but rather the number of independently variable factors affecting the performance. In a free improvisation group, for example, each musician has much freedom when it comes to controlling any aspect of the performance. This is quite different from a jazz trio playing a tightly synchronized Bebop tune or an orchestra playing a nineteenth-century concert. These differences in the musician’s degrees of freedom necessarily influence the way they gesture.

The degrees of freedom of a musical performance may correlate with whether a score is used in the performance, and the level of detail in the notation of the score. Musical scores come in many different flavours, and some are more open for flexibility on the performer’s side than others. However, even in performances based on very detailed score instructions, the musicians still have to create a meaningful performance as a group through the shaping of timbre/texture, the dynamics, and the timing of notes. Sawyer and DeZutter (2009) describe how the performances of a symphony orchestra can vary “from night to night,” despite the presence of explicit leadership and a musical score. This depends to a large degree on the continuous communication between performers, in which gestures may also play an important part. These gestures can be based on expressive elements of sound-producing actions, such as to indicate the beginning of a passage, or upper body swaying to shape a phrase. Other times they may be purely communicative such as head nodding or facial expressions.

Freely improvising ensembles relies even more on gestural communication during a performance. If there are no/few predefined musical elements, this requires a high level of awareness within the ensemble to follow each other (Becker, 2000). Then all sorts of cues—

including different types of gestures—are used to convey the intentions of the performers. Sometimes the performers may also use gestures to intentionally “sabotage” each other during the performance. One example of this can be found in “call and response” improvisation, during which the musicians may surprise each other, using both auditory and visual cues, to create more interesting and unexpected musical results (Schuiling, 2018, p. 46).

### *Musical leadership*

While most musicians’ primary focus is that of producing (musical) sound, the *musical leader* is also/instead focused on gestural communication with the rest of the ensemble and the audience. Such gesturing may be in the form of “indirect” sound production, through controlling individual onsets. It could also be more abstract, with the aim of creating a particular emotional feeling in the ensemble and audience.

Ensembles rely on different types of musical leadership such as the first violinist of a string quartet, the founder of a jazz band, the lead singer in a pop/rock band, or the conductor of a full-size orchestra. Davidson and Correia (2002, p. 243) describe how Annie Lennox takes on the role of leading the other members of her band in addition to her singing: “She is a narrator-interpreter in her use of illustrative and emblematic gestures with the co-performers and audience. She is a co-worker in her use of regulatory movements to coordinate musical entrances and exits.” Leante (2014) argues that *khyal* singers in North Indian classical music communicate the lyrics through iconics or metaphors, and perform abstract gestures along with the flow of the improvised sections. However, musical leadership is not always connected to a particular person. In a discussion of “shared intentionality,” McCaleb (2014, p. 91) describes how the leadership varies within a string quartet, what he refers to as the “fluidity of ensemble roles.” For this to happen, it is necessary that the musicians have knowledge of the musical material beyond their own part, and they also need to use non-verbal communication with their co-performers.

While such musical leadership is possible in small ensembles, larger ensembles often rely on a conductor as the *de facto* driving force in what Volpe et al. (2016) called “sensorimotor conversations” with the musicians. From a gestural perspective, the conductor is unique, being the only individual in the ensemble who does not have any direct influence on the sound production. They can instead be seen as an important communicator of emotional content, but they are also expected to provide temporal and structural information to the musicians. Using motion capture, Luck and Nte (2008) showed how conductors’ motion trajectories



induce the perception of temporal events. The role of the conductor is much more than only being a timekeeper, however, as orchestra musicians interact with each other in complex ways during performance (Ponchione-Bailey & Clarke, 2020). Wöllner and Auhagen (2008) have also shown that the perceived qualities of a conductor's gestures vary, depending on where the musicians sit in an orchestra. Their findings indicate that the frontal (woodwinds) and left-hand (first violinists) perspectives are better when it comes to receiving the "level of arousal" and "rhythmical clarity" of the conductor's gestures. This also shows the intrinsic relationship between the musical leadership and the size and setup of an ensemble.

The role of the conductor is institutionalized in Western (art) music traditions, and is largely focused around note-based scores. There are also cases in which the musical structure itself is conducted. One such example is John Zorn's *Cobra*, which features the composer as "prompting" the ensemble with a system of symbols printed on cue cards and lead sheets. This piece also relies on non-verbal interaction between the ensemble members: "*Cobra* exaggerates the degree to which all music performance depends on the communication of musical and other interpersonal signs—which may involve physical movement, including bodily and facial gestures, as much as sound" (van der Schyff, 2013, p. 6). An example of improvised leadership without the presence of a conductor can be found in traditional Turkish art. *Meşk* is a face-to-face teaching method in which the students emulate their master's way of performing. This leads to a participatory performance style, in which the musicians improvise according to the actions of the co-performers, but also with respect to the tradition. The performance can be seen as a "faithful repetition as a teaching method" (Özdemir, 2019, p. 148), which emerges as a performance concept for innovative improvisation, yet directed by the master and the tradition.

### *Machine musicianship*

Until now, we have focused on gesturing found in ensembles with human performers. However, what happens when *machines*—from mechanical devices to computers—enter the stage? What Robert Rowe (2001) coined as *machine musicianship* is today prevalent in many musical settings. This can be in the form of machine-based instruments played by humans, robots playing acoustic instruments, or machines conducting human musicians. In the most extreme cases, one may even find completely autonomous machine-based musicianship.

A common complaint among audiences that experience machine musicianship has been the lack of causal relations between what is seen and heard: "As far as I could tell, they were all just checking their e-mail"

(Trueman, 2007, p. 176). When performing with acoustic instruments, there is a direct energy transfer from musician to instrument, which can be seen/heard by other musicians and the audience. Machine-based instruments, and in particular digital musical instruments (DMIs), are based on the creation of *mappings* from action to sound. The creation of various types of “gestural controllers”—and meaningful mappings from these controllers to sound engines—have been a much-researched topic in the field of new interfaces for musical expression over recent decades (Jenseni & Lyons, 2017). This has been driven by the need for creating relationships between actions and sound that make sense, as well as an understanding of the power of gestures in music performance.

One form of machine musicianship that we find particularly interesting when it comes to the topic of ensemble gesture is that of *live coding*. This is a performance style in which the musician writes code on the computer, generating musical sound in real time. Since the musician typically sits in front of a computer during the entire performance, it has become common to project the written code on a screen behind the performer. Sometimes also a live video is shown of the musician’s finger typing. This helps in “humanizing” an otherwise quite disembodied performance style. With better Internet technologies available, live coding is now also spreading to collaborative performances, in which many people type code in shared “virtual rooms.” Here the performers’ telematic typing actions can be seen as having a meaning-bearing component, hence becoming gestures in their own right (see Chapter 20 for more on live coding music ensembles).

Over the last decades, there has been a growing interest in forming laptop orchestras that explore the potentials of machine-based performance in a larger ensemble setting. Interestingly, this has also led to the need for more well-defined musical leadership within the ensemble. Some laptop orchestras have solved this by having a human conductor, while in other cases, a computer is assigned a conductor role. Then the messaging can come in the form of text messages or symbols on the screen. There are also ensembles exploring combinations of human and machine conductors. For example, in Dan Trueman’s *PLahara*—inspired by the traditional North Indian *lahara* and composed for the 15-piece Princeton Laptop Orchestra (PLOrk)—the role of conductor is distributed among three conductors. The first uses a local network to add new layers and direct the tonal information of the piece. The second uses sign language to call particular sections, while the third uses a “knob box” to control various parameters of different instruments of PLOrk, over the network (Smallwood et al., 2008).

It is interesting to see that several laptop orchestras, and also other groups involving machine musicianship, incorporate various types of “gestural” controllers or actuators. In our machine-based performance practice, we have found it necessary to use gestural communication to play together. This includes using wearable sensors that capture motion or physiological data, live visuals, or physical objects that can “embody” the digitally produced sound. Without such visual references, it is often impossible to understand what is going on during a performance.

## Conclusion

As the above discussion has shown, gestures are an essential (and integral) part of ensemble performance. The types of gestures, and who performs them, differ widely depending on musical genre and culture. Factors such as the size and setup of the ensemble and their musical degrees of freedom influence the way musicians gesture. How the musical leadership is organized is another important factor. If there is no conductor, the musicians need to communicate more directly with each other. This leads to different types of performance gestures than when a conductor is taking on the main communication role. Finally, machine musicianship challenges traditional ideas of what an ensemble is and paves the way for many new performance styles. Despite these differences, however, one thing that all types of music performance share is the production of musical sound as well as the need for meaningful communication between musicians and audiences.

## References

- Becker, H. S. (2000). The etiquette of improvisation. *Mind, Culture, and Activity*, 7(3), 171–6.
- Cadoz, C., & Wanderley, M. M. (2000). Gesture-music. In M. M. Wanderley, & M. Battier (Eds.), *Trends in gestural control of music* (pp. 71–94). IRCAM.
- Davidson, J. W., & Correia, J. S. (2002). Body movement. In R. Parncutt, & G. E. McPherson (Eds.), *The science and psychology of music performance. Creative strategies for teaching and learning* (pp. 237–50). Oxford University Press.
- Godøy, R. I., Jensenius, A. R., & Nymoen, K. (2010). Chunking in music by coarticulation. *Acta Acoustica United with Acoustica*, 96(4), 690–700.
- Godøy, R. I., & Leman, M. (Eds.) (2010). *Musical gestures: Sound, movement, and meaning*. Routledge.
- Gritten, A., & King, E. (Eds.) (2006). *Music and gesture*. Ashgate.

- Gritten, A., & King, E. (Eds.) (2011). *New perspectives on music and gesture*. Ashgate.
- Hatten, R. S. (2004). *Interpreting musical gestures, topics, and tropes: Mozart, Beethoven, Schubert*. Indiana University Press.
- Jensenius, A. R. (2018). Methods for studying music-related body motion. In R. Bader (Ed.), *Handbook of systematic musicology* (pp. 567–80). Springer-Verlag.
- Jensenius, A. R., & Lyons, M. J. (Eds.) (2017). *A NIME reader: Fifteen years of new interfaces*. Springer.
- Jensenius, A. R., Wanderley, M. M., Godøy, R. I., & Leman, M. (2010). Musical gestures: Concepts and methods in research. In R. I. Godøy, & M. Leman (Eds.), *Musical gestures: Sound, movement, and meaning* (pp. 12–35). Routledge.
- Keller, P. (2014). Ensemble performance: Interpersonal alignment of musical expression. In D. Fabian, R. Timmers, & E. Schubert (Eds.), *Expressiveness in music performance: Empirical approaches across styles and cultures* (pp. 260–82). Oxford University Press.
- Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge University Press.
- Leante, L. (2014). Gesture and imagery in music performance: Perspectives from North Indian classical music. In T. Shephard, & A. Leonard (Eds.), *The Routledge companion to music and visual culture* (pp. 145–52). Routledge.
- Luck, G., & Nte, S. (2008). An investigation of conductors' temporal gestures and conductor—musician synchronization, and a first experiment. *Psychology of Music*, 36(1), 81–99.
- McCaleb, J. M. (2014). *Embodied knowledge in ensemble performance*. Routledge.
- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. University of Chicago Press.
- Özdemir, Ş. S. (2019). Discovering one's self through embodiment of tradition in meşk: An analysis of the mode of transmission in Turkish performative traditions. *Musicologist*, 3(2), 151–70.
- Ponchione-Bailey, C., & Clarke, E. F. (2020). Digital methods for the study of the nineteenth-century orchestra. *Nineteenth-Century Music Review*, 1–32.
- Rowe, R. (2001). *Machine musicianship*. MIT Press.
- Sawyer, R. K., & DeZutter, S. (2009). Distributed creativity: How collective creations emerge from collaboration. *Psychology of Aesthetics, Creativity, and the Arts*, 3(2), 81–92.

- Schuling, F. (2018). *The instant composers pool and improvisation beyond jazz*. Routledge.
- Smallwood, S., Trueman, D., Cook, P., & Wang, G. (2008). Composing for laptop orchestra. *Computer Music Journal*, 32(1), 9–25.
- Trueman, D. (2007). Why a laptop orchestra? *Organised Sound*, 12(2), 171–9.
- van der Schyff, D. (2013). The free improvisation game: Performing John Zorn's Cobra. *Journal of Research in Music Performance*, (Spring issue), 1–11.
- Volpe, G., D'Ausilio, A., Badino, L., Camurri, A., & Fadiga, L. (2016). Measuring social interaction in music ensembles. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1693), 20150377. <http://doi.org/10.1098/rstb.2015.0377>
- Wöllner, C., & Auhagen, W. (2008). Perceiving conductors' expressive gestures from different visual perspectives. An exploratory continuous response study. *Music Perception*, 26(2), 129–43.