Multimodality and Music Performance. The Lexicons of Gesture and Gaze in Orchestra and Choir Conductors

Isabella Poggi[1], Alessandro Ansani[1-2]

[1]Department of Philosophy, Communication and Performing Arts, Roma Tre University, Via Ostiense, 234, 00146 Rome, Italy, isabella.poggi@uniroma3.it

[2]Department of Psychology, Sapienza University of Rome, Via dei Marsi, 78, 00185 Rome, Italy, alessandro.ansani@uniroma1.it

Abstract: This paper examines recent studies within a general project, aimed at writing down the lexicons of a Conductor's body signals. Three observational studies focusing, respectively, on the semantic area of intensity and on the modality of gaze, through annotation of several fragments of orchestra and choir conduction in concert and rehearsal, discovered a total of 100 signals of intensity in various modalities, out of which, 21 intensity gestures and 20 conducting items of gaze. Subsequent perception studies on a subset of these signals reveal slight differences in interpretation by music Experts, Non-Experts and Amateurs, but confirm their comprehensibility, probably due to their underlying semiotic devices, the same holding also in everyday life gestures and gaze.

Keywords: Multimodality; lexicons; gesture; gaze; music; Conductor

1 Introduction

A Conductor's behaviour during music performance is a fascinating set of body movements that provides musically relevant information to performers but also makes part of the magic of that special kind of interaction and flow of emotions that holds in a choir or an orchestra. Research on the bodily communicative behaviour during music performance has investigated the musical and communicative functions of performers' movements, their influence on the listeners' subjective experience of music [1] and in co-performers' synchronisation [2] [3], gestures in singing and in musical teaching [4] [5], and the conductors' movements [6, 7, 8, 9]. Such body of research, beside affording application in musical teaching and performers' awareness, may be exploited in the construction of artificial systems that perform, understand, interpret or judge music, like systems for home conducting [10] or conductor virtual agents and

robots [11] [12]. But investigating the Conductor's movements by the same methods applied to human languages or other systematic communication systems could also bear on our knowledge of the underlying devices of human communication.

The goal of this paper is to demonstrate that the signals used in conducting make up communicative systems, i.e., systematic systems of signals produced by conductors and understood by musicians and singers, that include shared lists of signal–meaning pairs and rules for the construction of signals. To this aim, we overview six recent observational and empirical studies, on the systems of gestures and gaze signals in human conductors.

2 Communication Systems and their Rules

According to the model that drives our research [13], people communicate with each other by means of signals that convey meanings, and through various bodily and technological modalities they can produce a wide variety of signals: spoken and written words, gestures, facial expressions, head movements, glances, postures, but also pictures, movies, strikes and so forth. These signals, with their corresponding meanings, are organized into communication systems, only a subset of which can be called languages. Communication systems are sets of rules for setting a correspondence between signals and meanings, but we can distinguish two types of them. In "codified" communication systems, or "lexicons", a signal is linked to its meaning by a correspondence rule, i.e., a signal-meaning pair stored in the long-term memory of all interactants in a shared way, so that the Sender of a message after conceiving the meaning to be communicated (e.g. "animal flying by wings") searches one's lexical memory and finds out the signal devoted to convey it (*bird*), and the Addressee once received the signal searches one's lexical memory to retrieve the corresponding meaning. The words of a language, but also the symbolic gestures used in a specific culture, and even many gaze signals we produce in everyday life, make up are lexicons, that is, codified communication systems, whose lexical items can be found out and listed in a "gestionary" or a "gazeionary" [13]. "Creative" communication systems, instead, include only a small set of inference rules, generally rules of resemblance following which one can create new signals by imitating aspects of the meanings to convey: e.g., to communicate "animal flying by wings" to someone not sharing your language you might move hands as if they were wings. Pantomime and the iconic gestures invented on the spot in everyday interaction are creative communication systems [13]. The comprehensibility of signals between Sender and Addressee is made possible, for codified systems, by the signal-meaning pairs being shared in their memory, while for creative systems, by Sender and Addressee sharing the same rules of resemblance and imitation. Still, neither creative nor codified systems are

a language yet: we may call a language a communication system that includes both a lexicon (a codified set of words – or signs, for Sign Languages of the Deaf) plus a syntax (a set of rules to combine words into sentences) [13].

When it comes to the gestures, facial expressions, and gaze signals used by a conductor in music performance, our hypothesis is that they constitute communication systems, whether codified or creative, although not strictu sensu languages. In particular, we aim to demonstrate that their signals are shared between musicians and conductors, and that within them it is possible to find out lexicons and rules for the construction of iconic signals. To detect them, our goal was first to single out the conductors' signals making hypotheses on their meanings, and second to verify if other people, musicians and non-musicians, actually attribute them the meanings hypothesized.

3 Signals of Intensity

A strategy in the endeavor of analyzing the Conductor's communication systems is to focus on a single parameter of music performance – e.g., tone, rhythm, expressivity – and search, in whatever modality, signals conveying meanings in that area. Adopting this approach, a qualitative observational study [14] focused on the parameter of music intensity and singled out the gestures, postures and other signals that convey the meanings "forte", "piano", "crescendo", "diminuendo".

Ten fragments of orchestra and choir conducting (totally 122' 35": 27' 55" from 4 concerts and 94' 40" from 6 rehearsals) by an amateur and two professional conductors (Alessandro Anniballi, Leonard Bernstein, and Riccardo Muti) were analyzed, finding out 100 occurrences of intensity signals. Each signal was annotated in terms of the scheme in Table 1. Here, col. 1 contains the time in the video, col. 2 the intensity indication written on the musical score, 3 (for choir performance) the words sung when the Conductor's signal is produced; 4, the analyzed modality; 5, a description of the signal in terms of its parameters: e.g., the Conductor's posture is *"Bust forward, shoulders closed, head forward downward"*; col. 6 writes the "originary meaning" – the body movement from which the communicative meaning stems [13]: e.g., the described posture is that of someone making oneself smaller. Col. 7 contains the meaning conveyed: this posture means "softer", i.e., "make a 'smaller' sound". Col. 8 classifies the signal and col. 9 its underlying semiotic device: making oneself smaller is a "creative" iconic gesture (col. 8) that exploits an iconicity device, a transmodal shift from space to sound (col. 9), since a smaller body evokes a smaller (i.e., softer) music; a body taking less room recalls a sound taking less energy (Table 1).

The analysis resulted in a lexicon of signals of intensification and attenuation performed by gestures, head movements, facial expression, gaze, body movements, posture, and voice, with ones for intensification more frequent than those for attenuation: respectively, 50 signals for "forte" and 8 for "crescendo", 35 for "piano" and 7 for "diminuendo" (Table 2).

Table 1

An annotation scheme for signals of intensity

1 Time	2 score	3 words	4 Modality	5 Signal description	6 Originary meaning	7 meaning	8 Signal type	9 Semiotic device
3.18	*Più piano*	Je-e-su-u Chri-i-**i-ste**	Post-ure	*Bending trunk, head forward downward*	I make myself smaller →	Softer	Iconic	iconicity: space→sound Take less room=make a softer sound
			Gest-ure	*Bent elbows* *Lh. palm forward, move backward* *Lh. open makes precision grip*	refrain from I pick up something carefully	attenuate do some-thing subtle	Codified gesture Iconic gesture	Generic Codified Iconicity movement→ sound Smaller grip = softer sound
			gaze	*Eyebrows raised* *Eyeilds tense, squeezed half-closed*	I try not to be heavy	Be light,	Indirect iconic	Movement during action

Table 2

Signals of intensity in 7 modalities

	Forte	Piano	Cresc.	Dimin.	TOT
Gesture	33	23	5	2	63
Head Movements	3			1	4
Facial expression	5	1			6
Gaze	7	7		1	15
Posture		3		1	4
Body	2		3	1	6
Voice		1		1	2
TOT.	50	35	8	7	100

The 100 signals include 63 gesture tokens, within which 21 gestures types – recurrent gestures in which the same description of the signal corresponds to the same meaning – were singled out and classified into five categories, according to the semiotic device of their construction:

1) **generic symbolic gesture**: a symbolic (hence, codified) gesture also used by laypeople, exploited by conductors with the very same meaning as in everyday life: e.g., *Index finger over lips*, generally meaning "be silent", is used to ask for "piano" (Fig. 1);

2) **specific symbolic gesture**: a codified gesture used in conducting with a slightly different meaning than in everyday life: *hands, palm up, oscillating on wrist up-down* in its everyday reading [15, 13] means "come here", since it encourages the Addressee to come closer to the Sender, while in conduction it is used to mean "come on, play / sing louder";

3) **direct iconic gesture**: a creative gesture imitating movements or other kinds of change by a transmodal shift from another modality: e.g., *arms curve enlarging*, imitating a swelling body, ask for a "crescendo", a swelling sound; in this "transmodal iconicity", an analogy is set between the enlarging or widening of a physical shape and the progressive amplification of a sound: the idea of enlarging is shifted from a tactile and visual domain to an auditory one (Fig. 2);

4) **indirect iconic gesture**: an indirectly iconic gesture imitating a movement that is usually performed when producing an action or the resulting sound. In these cases, the Conductor performs a gesture that does not directly imitate the movement it refers to, or its transmodal analogue, but some movement that by inference may recall the desired intensity. Such indirect iconicity may pass through two different kinds of movement:

    a) **motoric attitude**: the gesture imitates a movement that is usually performed when producing another movement or the resulting sound. To mean "sforzato" (with effort) the Conductor may suddenly *clunch his fist*, thus imitating the movement people do when striving to perform some physical action;

    b) **emotion expression**: the gesture imitates the movements typically performed during the expression of an emotion that, when felt, induces the wanted type of motoric attitude or movement. For example, *hands in the shape of claws vibrating with high muscular tension* work as an indication for "forte", because tension is typical of an activating emotion like anger, and anger mobilises the energy required for playing or singing "forte" (Fig. 3).

In these gestures, sometimes the dynamic indication is conveyed by the gesture globally with all its parameters of handshape, location, orientation and movement, also including expressive sub-parameters of movement like amplitude, tension, fluidity [13]. However, sometimes it is not the whole gesture that bears the meaning of intensity, but one single aspect of it: for instance, "forte" is often conveyed by the gesture *both fists moving forward with high muscular tension* but, sometimes, only by the value "*high*" in the parameter of muscular tension, whatever the hand configuration and movement; or else only by the handshape "fist", which by itself bears a visual metaphor of strength.

Figure 1[1]
Generic symbolic gesture
Index finger before lips = "Be silent"

Figure 2
Direct iconic gesture
Arms curve enlarging to imitate a swelling body = "crescendo", a swelling sound [20]

Figure 3
Indir.iconicity through emotion expression
Claw vibrating with high muscular tension ="forte"

[1] Figures 1 – 5 are printed with permission of Alessandro Anniballi

As shown by Tables 3 and 4, the correspondence of intensity indications with the 21 gestures found is not random. Those with the *fist* handshape only convey "forte", while those exploiting the *claw* (*hand open, curve fingers, high muscular tension*) convey only "forte" and "crescendo". Conversely, the *precision grip* **(***thumb touches index*) conveys "piano" (and it would seem strange for it to mean "forte" or "crescendo"): the handshape is not indifferent to meaning, each one is the visual and motoric embodiment of a dynamic indication. Thus sometimes it is not necessary a gesture as a whole to convey a meaning of intensity, but also a gesture devoted to convey another type of musical indication – for example, a signal of "start" – may also contain, perhaps in a single value of one of its parameters, a nuance of intensity.

Table 3
Handshapes for "forte" and "crescendo"

Whole gesture *	Forte	Crescendo
Open hand	*right hand, palm up, oscillates on wrist from Musicians to Conductor, as if meaning "come on, come here, come forward"*	
Open hand curve loose		*Arms with open hands curve open outward while shoulders raise upward (Fig. 1)*
		Left hand open curve loose, palm to Conductor, rotates forward repeatedly in wider and wider rounds
		Both arms open curve loose, palm up, alternatively move fingers up and down gently towards Conductor
Extended index finger	*both index fingers pushed towards musicians*	
Fist	*right hand, palm to left, pushed forward towards musicians*	
	both hands, palms to each other, pushed forward towards musicians	
	right hand, palm up, moved forward with fluid movement towards musicians	
Claw (hand open, curve	*right hand, palm up, moves towards Musicians*	*right hand in claw shape, palm down,*

fingers, high muscular tension)		*raises upward*
	right hand, palm to Conductor, vibrating	
From fist to claw	*right hand in fist shape, palm up, pushed towards Musicians opens up in a claw*	

*Gestures ordered by handshape

Table 4
Handshapes for "piano" and "diminuendo"

whole gesture *	Piano	Diminuendo
Open hand	*Both hands open with close fingers, palms down, move inward–outward as if smoothing a surface*	
	Both hands open with close fingers, palms down, slightly move downward, as if keeping a surface down	
	Both hands open with close fingers, palms forward, move forward	
	Both hands open with slightly open fingers, palms forward, speedily oscillate on wrist left-right as if saying "no, no"	
	One or both hands open, palms down, alternatively move fingers up and down gently	
Open hand curve loose		*left arm with open hand palm down retracts backward*
Extended index finger	*Right hand, palm to left, extended index finger up, moved towards conductors mouth, to touch lips* (Fig. 1)	
Precision grip (thumb touches index)	*right hand or both hands with thumb and index finger touching, palms forward move forward in a fluid manner*	
V shape (index and middle fingers extended open)		*right hand in V shape, palm to Conductor, moves rightward progressively closing index and middle finger*

*Gestures ordered by handshape

That sometimes a single value is responsible for a specific meaning of intensity is even clearer from Table 5. Besides the handshapes of *fist* and *claw*, which bear a meaning of strength, and that of the *precision grip*, which evokes something to

pick with delicacy and accuracy, several values in the sub-parameter of hand movement are the carrier of systematically contrasting intensity meanings: gestures directed *towards musicians*, with *jerky movements*, *high muscular tension*, *wide amplitude*, *high quantity of movement*, and *vibrating* indicate "forte", while gestures directed *downward*, *fluid*, and of *low tension*, *amplitude*, and *quantity of movement* convey "piano".

Table 5
Single values in gesture parameters conveying "forte" and "piano"

Single value	Forte	Piano
handshape	*fist*	
	claw	
		precision grip
movement direction	*towards musicians*	*downward*
fluidity	*jerky movements*	*fluid movements*
tension	*high*	*low*
amplitude	*wide*	*low*
quantity of movement	*high*	*low*
manner of movement	*vibrating*	

3 The Lexicon of the Conductor's Gaze

A different strategy to study the Conductor's lexicons and the rules for the construction of communicative items is to collect all the signals of whatever semantic area in a single modality. Thus, after finding the Conductor's signals of various modalities in the semantic area of intensity, a qualitative observational study [16] investigated the Conductor's signals of all semantic areas in the modality of gaze. To test the hypothesis that gaze constitutes a lexicon, that is, a systematic and shared list of signal-meaning pairs conveying information and requests relevant for music making, a corpus was collected (99' 43") of video fragments of orchestra and choir conduction by Alessandro Anniballi: 5 from Gabriel Fauré's *Requiem* (2 in concert and 3 in rehearsal), 3 from Antonio Vivaldi's *Magnificat* (1 concert and 2 rehearsals), and one from Gioachino Rossini's *Petite Messe Solennelle* (1 concert fragment).

All fragments were annotated as in Table 6, where col. 1 identifies time and musical excerpt; col. 2, the words sung or keys played simultaneously to the analyzed gaze; col. 3 describes communication in other modalities (words, gestures, head movements); 4, the Conductor's gaze in terms of its parameters (e.g. eyes direction, eyebrows position…); col. 5 contains a verbal paraphrase of its literal meaning, and 6 one of the inferable contextual meaning. Col. 7 classifies

the gaze item in terms of a semantic typology (performative, emotion, cognitive state conveyed), and 8 its function of musical indication (intensity, start…). Col. 9 specifies the semiotic device exploited, among those found above for gestures.

At line 1, Col. 4, the Conductor *looks at tenors*: a bare request for attention in everyday communication (col. 5), here a request to prepare to start (6); then a gaze with a performative of request (7) and a function of technical indication, "start" (col. 8). This gaze is "specific codified" (col. 9), bearing a more specific meaning than in everyday communication.

At line 2, the Conductor *raises his eyebrows very high* (col. 4), imitating an upward movement (5) aimed at requesting (7) a higher note (6): an indication of musical pitch (8) conveyed through a semiotic device of "direct iconicity" (9), imitation of an audible "raising" by a visible "raising". "Indirect iconicity" is exploited instead on line 3: while asking tenors to sing an E by *extended index finger* (col. 2), the Conductor *sings* it himself (3) and *looks at them* with *frowning eyebrows* (4) to mimic an expression of effort (5), thus asking to strive maintaining that difficult note (6): here pitch indication (8) does not exploit the similarity between two movements (raising voice and raising eyebrows), but an expression of effort, that requests an effortful movement. Exploiting "motor induction indirect iconicity", this gaze requests the right pitch by evoking the technical movement (effort) necessary to produce it.

Table 6

Annotation scheme of the Conductor's gaze

1. Time	2. Voice /music	3. Other-modalities	4. Gaze	5. Literal meaning	6. Contextual meaning	7. Gaze Type	8. Mus. Func.	9. Semiotic device
1 R 0.02			*Looks rightward*	I request your attention	Be ready to start	Request	Start	Spec. Cod.
2 R 1.48	Ré-qui-em		*Raises eyebrows very high*	I imitate a raising	Start with the high note	Request	Pitch	Direct Icon.
3 R 2.44	*Piano*: E C	Sings "*mii*" (E); *upright index finger points to tenors*	*Frowns, looks at tenors*	I strive	Strive: keep the high note	Request	Pitch	Indir. Icon.
4 M 8.32	Maa-[gni-ficat]	*Shakes head + clutches and drops fists down*	*frowns + squints eyes*	I am angry	Play very loud	Request	Intensity	Indir. Icon.

The indirectly iconic gaze at line 4, where the Conductor *frowns* and *squints eyes* (col. 4) expressing anger (5), is a request (6) to sing aloud, an indication of intensity (8); the expression calls for the emotion, which in turn calls for energetic movement, hence loud sound: indirect iconicity (9).

The analysis revealed, for gaze signals, the same semiotic devices found for gestures: generic and specific codified, directly and indirectly iconic (see examples in Table 6). Further, when taking the actual musical effect into account, the analysis showed that, even across performances, the same meanings are recurrently attributed to the same gaze items.

In total, 17 gaze types were found in these fragments [16] (Table 7), later identified also in performances by other conductors, like Herbert von Karajan, Leonard Bernstein, Daniel Barenboim, and Riccardo Muti. In the scheme, col. 1 describes the considered gaze item, 2 its literal meaning, 3 a possible indirect meaning: e.g., at line 1, with the item "*gazes at X*" (col. 1) the Conductor may simply call for attention (2), but this attention request is more typically codified with the indirect meaning "prepare to start" (3). Then the item is classified (in col. 4) as to its function in conducting: except for the checking gaze (line 3) that is non-communicative, those of order, emphasis to ask for attention, praise, reproach have an **interactional** function (7 in total); gazes of start, intensity (play/sing soft, loud or *sforzato*) and higher note convey **technical** indications (6), while an attitude/emotional category includes both gazes asking for accuracy, concentration, and those displaying emotions (3): some (Outcome emotions [17]) are felt about the ongoing performance (see enjoy, n.16, that fulfils a motivational function), others (Meaning Oriented emotions) are those to be stamped into the music sung or played, e.g., feeling sad to make sad music, n. 13).

A third observational study [18] focused on the use of gaze by Leonard Bernstein, who credited such a relevance to gaze communication in conducting that, during the encore of a performance of Haydn's Symphony in 1989, he conducted the ensemble with his hands behind his back, using his face and gaze communication only. In this fragment, Bernstein's conducting gaze is the same as that of other conductors: 56 gaze items were found, conveying 74 between direct and indirect meanings: 22% indicating start, 16% expression, 15% intensity, 13% feedbacks, 6% other technical indications, and 28% emotion expression.

Figure 4
Warning gaze
Raised eyebrows with oblique gaze

Figure 5
Angry gaze = "forte"
Strong frown

Table 7
The lexicon of the Conductor's gaze

	1 Gaze type	2 Literal Meaning	3 Indirect meaning	4 Function
1	*Gazes at X*	Request for attention	Prepare to start	Technical (start)
2	*Gazes around at all musicians*	Broadcast request for attention		Interactional
3	*Looks at all musicians*	Checking gaze. (Non-communicative)		Self-information
4	*Raised eyebrows with oblique gaze*	Warning gaze (Fig.4)	I warn you about a difficult passage	Interactional
5	*Raised eyebrows with eyes open wide*	Emphasis	I ask for higher attention	Interactional
6	*Eyebrow frown with wide open eyes (+ extended index finger)*	Peremptory order		Interactional
7	*Wide open eyes fixing X*	Threating gaze (to prevent similar behavior)	I reproach you for your mistake	Interactional
8	*Raised eyebrows (+nodding)*	Appreciation + approval	I praise you	Interactional
9	*Continuous eyebrow frown (+ rocking head)*	Request to continue		Technical (tempo)
10	*Short single eyebrow raising*	Higher note		Technical (pitch)
11	*Raises eyebrows all along the musical fragment*	Imitation of light movement	Play/sing soft	Technical (intensity)
12	*Raises eyebrows (+ head in the shoulders)*	Caution gaze	Be accurate and precise	Attitude
13	*internal parts of eyebrows raised*	Sad gaze	Play/sing a sad way	Emotional (Meaning oriented emotion)
14	*Frown*	Angry gaze (Fig.5)	Feel/express anger → play loud	Technical (intensity)
15	*Squints eyes*	Imitation of effortful movement	Play/sing "*sforzato*"	Technical (intensity)

16	*Closed eyes*	Concentration	I do enjoy and want you to enjoy the pleasure of music	Emotional (Outcome emotion→ Motivational strategy)
17	*Squeezed eyes (+trunk retracting backward)*	Disgusted gaze	Outcome emotion → Neg. feedback	Interactional

4 Do People Understand the Conductor's Gesture and Gaze?

A set of signals, to constitute a lexicon, should be mutually understood by both Senders and Addressees. The above observational studies allowed us to find a set of conductors' gestures and gaze signals, and to make a hypothesis about the meaning of each of them. To test if these meanings are really shared, so as to make lexicons proper, three perception studies were carried out [18, 19, 20].

As the current paper relies on those perception studies in order to propose a theory in support of the lexical nature of the Conductor's gesture and gaze, for ease of readability, the statistical aspects of such studies will not be discussed. For all the analyses, tables, and graphs, the reader can refer to the single studies [18, 19, 20].

4.1 A Perception Study on the Conductor's Gestures

The first study, focusing on the Conductor's gestures, wondered 1) if people attribute them the same meanings as hypothesized by the above analysis; 2) whether to interpret them they look more at the hands or other parts of the Conductor's body; 3) if experts and non-experts in music differ in their interpretation.

In a between-subjects design, 77 participants (45 females; mean age: 39, SD = 19), among which 44 non-experts (28 females; mean age: 28) and 33 experts (17 females; mean age: 51) were submitted 8 video clips of 1" to 4" (medium length, 2") where two conductors, Herbert von Karajan and Alessandro Anniballi, both performed 4 of the gestures found in the observational study, each with one of the 4 meanings of intensity, "piano", "forte", "crescendo", or "diminuendo".

Each participant watched two gestures, one by von Karajan and one by Anniballi, without audio, in random order. The first gesture shown was followed by an open question asking to phrase its meaning verbally, the second by a multiple-choice question asking to rate the gesture, on a 5-point Likert scale, with respect to 8 possible randomized meanings: 4 of intensity ("play soft", "play loud", "play

progressively louder", "play progressively softer", corresponding to *piano*, *forte*, *crescendo*, *diminuendo*), and 4 distractors ("play more in tune", "play passionately", "scan notes well", "keep the time"). Then the participant rated, on a 5-point Likert scale, what part of the Conductor's body they had mainly looked at to understand the meaning: hand movements, head movements, facial expressions, gaze, mouth movements, trunk position, body movements. Finally, some questions assessed the participant's level of musical expertise[2].

To analyze the results of the first question, participants' free interpretation of the gesture stimuli, five types of open answers were distinguished, assessing their relative quantitative occurrence:

- MIOk (Musical Intensity Ok), gesture correctly interpreted: e.g., *"forte"* interpreted as "intensify the sound", "play louder": 55% of experts' answers, 43% of non-experts'
- MI (Musical Intensity), meaning of intensity, but not the right one: "vocal diminuendo" (instead of simply "piano"): 15% by experts, 14% by non-experts
- ME (Music–Else), meaning linked to musical features other than intensity (pitch, rhythm, tempo…): "play faster", "play using a higher pitch" (instead of "play louder"): 9% of experts', 25% of non-experts' answers
- AE (Attitude–Expressivity), meaning linked to feelings and their relative expressions: "majesty", "supplication", "tenderness"; 18% experts', 9% non-experts'
- O (Other), irrelevant or hard to interpret: e.g., "look at me": 3% experts, 9% non-experts'

The relatively higher frequency of MIOk answers (where the participants guessed not merely a musical and an intensity indication, but the right one) suggests that the stimuli are quite easily recognized, and not only by music experts. Conversely, less accurate meaning attributions (ME and O, i.e., meaning concerning music but not intensity, and Other meanings) are somewhat more frequent in non-experts. Moreover, answers MIOk and MI – the meanings provided by participants closest to the one hypothesized – sum up to 70% and 57%, respectively, for experts and non-experts: this shows that, even without the suggestion of a multiple-choice question, people generally understand the meaning of these gestures.

The results of the multiple-choice question about the second stimulus show that the gesture was attributed the hypothesized intensity in three cases out of four: *piano* mainly elicited the answer 'play soft' (3.40), *forte* mainly 'play loud' (3.86), *crescendo* 'play progressively louder' (3.74), *diminuendo* 'play soft' (3.45).

[2] The questionnaires used in all studies were anonymous and complied with Helsinki declaration for the ethics of research.

Moreover, answers polarize on the two ends of the intensity continuum: gestures for *crescendo* are more frequently confused with *forte*, those for *diminuendo* with *piano*, and this is more so with non-expert participants.

Concerning the body parts mainly looked at to interpret the stimuli, whatever the Conductor in the clip, Karajan or Anniballi, both experts and non-experts, but experts in a more clear-cut way, mainly reported looking at hands more than other parts of the body.

In general, the difference between experts and non-experts in the interpretation of the conductors' gestures is not significant. This result may be probably accounted for by their high expressivity, given that part of them, as found by the qualitative study, are generic or specific codified gestures, that is, ones with a meaning identical or similar to that of everyday life, and the others are directly or indirectly iconic.

4.2 A Perception Study on the Conductor's Gaze

A perception study similar to that on gestures was run on the Conductor's gaze, to test the interpretation of 8 gaze items and its possible differences in music experts and non-experts. In a between-subject study, 10 video clips of gaze during a concert were submitted to 177 participants (113 females, mean age: 22). Two of them were "neutral" gaze items, used for control goals, while the remaining 8 corresponded to the following items of Table 7: 1, start; 7, reproach; 10, high note; 11, play soft; 13, play sad/poignant; 14, play loud; 15, play sforzato; 16, concentration. Each gaze was produced by one of 4 different conductors: Alessandro Anniballi, Riccardo Muti, Herbert von Karajan, and Daniel Barenboim.

Ten different questionnaires were created, each containing two video clips without audio, the first with an open question asking to verbally phrase what the Conductor was communicating by his gaze, the second with a question asking to rate on a 5-point Likert scale how plausible was each of 13 possible meanings.

For the analysis of open and multiple-choice questions, both the meanings hypothesized for each stimulus and the participants' answers were classified, in terms of their functions, as a. general technical indications (e.g., "high note", "start"); b. intensity ("play soft", "play loud)"; c. attitude/emotion indications ("play poignant", "concentration"); then the class of each answer was compared to that of the stimulus meaning, resulting in a score from 0 to 3.

Concerning the open questions, the two control items obtained shallow scores, eliciting a wide range of meanings, and also the correspondence of some experimental stimuli to the hypotheses of Table 7 was not high. Yet, the relationships between the target meanings and ones proposed by participants provide insightful hints about the nature of gaze items. Some answers confirm the

relationship between the categories “intensity” and “attitude/emotion”: for example, the stimulus “forte” is interpreted only by one participant as “He communicates a lot of intensity”, but others phrase it as “he reproaches someone”, “anger”, “violence”, “strictness”, “grit”, “heat”; all phrasings alluding to high arousal movements, like those performed to play or sing *forte*. This finding confirms the hypothesis on the semiotic device of indirect iconicity: emotion expression generally evokes an intensity indication. Nevertheless, also for “high note”, while only 4 participants out of 23 interpret the stimulus as “tone raising”, many others mention emotions like “passion”, “involvement”, “enthusiasm”, “light-heartedness”. Even in this case, there seems to be an intriguing relationship between some activating or joyful emotions and an upward direction.

As to the multiple choice questions, the stimuli classified as attitude/emotion were not significantly recognized, but others resulted highly comprehensible: within “technical” indications the most comprehensible was “start”, while in the class of “intensity” the opposite meanings “soft” and “loud” were very well distinguished from each other. Furthermore, in these answers too, like for the open ones, confusion and mismatches can be accounted for by the relationship between the functions of “intensity” and “attitude-emotion”: “play loud” was often interpreted as “play with anger”, and “play soft” with “play in an accurate way”.

Finally, no significant difference was found between music experts’ and non-experts’ interpretations.

4.3 A Perception Study on Leonard Bernstein’s Gaze

Starting from the analysis of Bernstein’s conducting gaze, a third perception study was carried out [18] to assess the comprehensibility of his gaze items, with two variants with respect to the previous study. Here the objective was to test 1) how comprehensible are Bernstein’s gaze items; 2) if comprehensibility differs in three levels of music expertise: experts, non-experts, and amateurs; 3) if experts take advantage of audio-video as opposed to only-video stimuli presentation more than do non-experts and amateurs.

A questionnaire was submitted to 186 subjects, balanced for gender and age, 37% "experts" (strictly music professionals), 29% "amateurs" (playing or singing, but not on a professional basis), and 34% "non-experts", asking them to rate three clips of Bernstein’s gaze items: "start", "pay attention", "crescendo and accelerando". Each clip was repeated twice, first in video-only mode, then in audio-visual mode, and both times the participant had to answer a multiple-choice question that mixed up the expected meaning with four distractors; but in the first and the second question, the same alternative answers were in a different order.

As to the first research issue, the three items overall obtained 48% correct interpretations, which, considering that each answer was chosen among five

possible ones, is more than twice the level of chance. As to the second, the results show a strict relationship between participants' musical expertise and gaze items correct interpretation. Whether viewing stimuli in only video or audio-visual condition, experts systematically perform better in interpreting the gaze items, followed by amateurs (who, yet, show a remarkable difference between interpretation from video-only and audio-visual presentation) and finally non-experts. Third, amateurs and experts take advantage of the audio for a correct interpretation significantly more than do non-experts. Generally, in the audio-video condition, the three groups perform coherently with their expertise (i.e., the more expert, the better); but without audio, the amateurs fairly recognize only the signal for "start" and not the other two, while the scores of non-experts are low in both conditions.

4.4 Discussion

The works presented investigated two portions of the Conductor's signals: intensity gestures and gaze items. Two observational studies on corpora of conduction in rehearsal and concert singled out 21 intensity gesture types, 8 requesting "forte", 7 "piano", 4 "crescendo", 2 "diminuendo", and 17 types of gaze with meanings grouped around three different functions: interactional (praise, warning, reproach); technical (intensity, melody, start); and attitude/emotion (concentration, passion, sadness). More gaze items were also found in Bernstein's performance. Then for 8 gestures and 11 gaze items (8 by Anniballi and 3 by Bernstein), their comprehension was tested in perception studies, to assess whether the signal-meaning correspondences hypothesized were shared by participants and if there were differences in comprehension between experts, amateurs, and non-experts in music, and between video-only and audio-visual presentation.

The comprehension of intensity gestures is relatively high for the two ends of musical intensity, "piano" and "forte": participants mostly confuse only with their respective shifts of intensity, "diminuendo" and "crescendo", and no significant difference results between experts and non-experts in their interpretation. Such an easy recognition can be accounted for by an important aspect highlighted by the observational study: part of the Conductor's gestures is the same for everyday communication – generic codified, i.e., shared symbolic gestures, and specific codified, i.e., more specific uses of them – while others are directly or indirectly iconic; their direction, speed, tension, and amplitude are highly parallel and isomorphic to parameters of sound or of the motor actions required to produce it.

As to gaze signals, a first perception study reveals a fair level of comprehension of technical items like "start", "high note", "piano", "forte", while those of attitude and emotion tend to be more often interpreted as indirect technical indications. This study too found no significant difference in gaze comprehensibility between

experts and non-experts, showing somewhat continuity between everyday gaze signals and those for musical indication, which exploit the same mechanisms for signal creation as plain language, such as metaphor and metonymy, conveying the desired meanings through generic codified, emotional, and iconic expressions.

However, in the study on the perception of Bernstein's gaze, thanks to the more rigorous recruitment of participants that distinguished professional experts from amateurs, some significant differences emerge. While the "technical" meaning "start" is easily understood by all participants, the "interactional" meaning "pay attention" is more frequently misunderstood by non-experts, but correctly interpreted by both experts and amateurs. A possible account for this result might be that non-experts do not expect to find such "everyday life" meanings like the interactional ones in a musical performance, and rather try to interpret them as technical signals, while experts and amateurs, having some experience of the relationship between Conductor and performers, recognize these meanings both with and without audio.

As to stimuli presentation, experts interpret video-only clips better, but with audio correctness increases in both experts and amateurs: this might implicitly confirm the importance of expertise to grasp the right meaning of gaze signals.

Conclusions

This work reviewed several studies within the general area of outlining the multi-modal communicative repertoire of orchestra and choir conductors. The results warrant the hypothesis that gesture and gaze do produce specific lexicons, covering all the meanings a conductor may need to convey during a performance. In addition, these findings evidence the similarity of communicative devices underlying human communicative systems, in all modalities, from verbal language to everyday gestures and gazes.

Acknowledgement

I am indebted to the choir “Orazio Vecchi” in Rome for bearing my repeated grabbings of concerts and rehearsals, and profoundly grateful to Alessandro Anniballi for his insightful way of conducting.

References

[1] A. B. Kumar, S. J. Morrison: The Conductor As Visual Guide: Gesture and Perception of Musical Content, *Frontiers in Psychology* 7, 08 July 2016, https://www.frontiersin.org/article/10.3389/fpsyg.2016.01049, DOI=10.3389/fpsyg.2016.01049

[2] A. Williamon, J. W. Davidson: Exploring co-performer communication, *Musicae Scientiae*, 6(1), 53-72, 2002

[3] E. King, J. Ginsborg: Gestures and glances: Interactions in ensemble rehearsal, In *New Perspectives on Music and Gesture*, 177-201, 2011

[4] J. W. Davidson: Visual perception of performance manner in the movements of solo musicians, *Psychology of Music*, 21(2): 103-113, 1993

[5] L. Simones, F. Schroeder, M. Rodger: Communicating musical knowledge through gesture: Piano teachers' gestural behaviors across different levels of student proficiency, *Psychology of Music*, 43(5): 723-735, 2015

[6] G. Luck: Synchronising a motor response with a visual event: The perception of temporal information in a conductor's gestures, In *Proceedings of the 6th International Conference on Music Perception and Cognition (ICMPC6)* (CDROM), 2000

[7] K. Parton, G. Edwards: Features of conductor gesture: Towards a framework for analysis in interaction, In Stevens C., Schubert E., Kruithof B., Buckley K., Fazio S. (eds.) *Proceedings of The 2nd international conference on music communication science*, pp. 3-4, 2009

[8] P. Boyes Bräm, T. Bräm: Expressive gestures used by classical orchestra conductors, In *Proceedings of the Symposium on The Semantics and Pragmatics of everyday Gestures*, Weidler, Berlin, pp. 127-143, 2004

[9] D. Veronesi: 'The Guy Directing Traffic': Gestures in Conducted Improvised Music between Metaphor and Metonymy. Poster presented at *RaAM Workshop: Metaphor, Metonymy and Multimodality*, Amsterdam, June 4-5, 2009

[10] A. Friberg, R. Bresin, J. Sundberg: Overview of the KTH rule system for musical performance. *Advances in Cognitive Psychology*, Special Issue on *Music Performance*, 2(2-3), 145-161, Software tools for automatic music performance, 2006

[11] P. Bos, D. Reidsma, Z. Ruttkay, A. Nijholt: Interacting with a virtual conductor. In Harper R., Rauterberg, M., and Combetto, M. (eds), *Proceedings of the 5th International Conference on Entertainment Computing*, Cambridge, UK (September 2006), Springer Verlag, Berlin, pp. 25-30, 2006

[12] S. Cosentino, K. Petersen, Z. Lin, L. Bartolomeo, S. Sessa, M. Zecca, A. Takanishi: Natural human–robot musical interaction: understanding the music conductor gestures by using the WB-4 inertial measurement system, *Advanced Robotics*, 28(11), 781-792, 2014

[13] I. Poggi: *Mind, hands, face and body. A goal and belief view of multimodal communication*, Weidler, Berlin, 2007

[14] I. Poggi: Signals of intensification and attenuation in orchestra and choir conduction, *Normas*, 7(1), 33-47, 2017

[15] D. Morris: *Manwatching*, Jonathan Cape, London, 1977

[16] I. Poggi: Lo sguardo del maestro (The Maestro's gaze), In Corradi C. (ed.), *Cultura popolare, religione diffusa, analisi qualitativa: un sociologo*

italiano a cavallo tra due secoli. Studi in onore di Roberto Cipriani, Morlacchi, Perugia, pp. 233-252, 2018

[17] I. Poggi: Body and mind in the Pianist's performance, In Baroni M., Addessi M. R., Caterina R. Costa M. (eds.), *Proceedings of the 9th International Conference on Music Perception and Cognition, ICMPC*, Bologna, 22-26 August, 2006, pp. 1044-1051

[18] I. Poggi, L. Ranieri, Y. Leone, A. Ansani: The Power of Gaze in Music. Leonard Bernstein's Conducting Eyes. *Multimodal Technologies Interaction* 4(2), 20, 2020. Special issue *"Musical Interactions"* (ISSN 2414-4088) doi: 10.3390/mti4020020

[19] I. Poggi, F. D'Errico, A. Ansani: The Conductor's intensity gestures, *Psychology of music*, October 21, 2020, 1-20, https://doi.org/10.1177/0305735620963179

[20] I. Poggi, A. Ansani: The lexicon of the Conductor's gaze, In *Proceedings of MoCo (Movement Computing) 2018*, Genoa, June 28-30, 2018