Towards a Coherent Terminology and Model of Instrument Description and Design

Tellef Kvifte
Dept. of musicology
University of Oslo
P.O.Box 1017 Blindern
0315 Oslo, Norway

tellef.kvifte@imv.uio.no

Alexander R. Jensenius
Dept. of musicology
University of Oslo
P.O.Box 1017 Blindern
0315 Oslo, Norway

a.r.jensenius@imv.uio.no

ABSTRACT

In this paper, we discuss the need for a framework for the description of musical instruments and their design, and discuss some possible elements in such a framework. The framework is meant as an aid in the development of a coherent terminology that may be used to describe, compare and discuss different musical instruments and musical instrument designs. The different perspectives of the listener, the performer, and constructor are discussed, and various levels of description introduced.

Keywords

Musical instruments, design, mapping, gestures, organology.

1. INTRODUCTION

In the literature on musical instruments and musical instrument construction, one central theme is the relation or mapping between the gestures[1] used to control the instrument and the sound that comes out of it. In the later years, an increasing number of articles describe aspects of this relation, and many of them base the discussion on specific examples of newly constructed controllers and/or instruments. In many cases, however, the discussion does not address more general principles, and even if the instruments described are interesting, the discussions do not necessarily add to a broader understanding of musical instrument construction. Part of the problem is a lack of consensus on terminology, and this paper is an attempt to start a discussion of necessary, fruitful and convenient terminology in the study of musical instrument description and construction. Through this, we advocate a field of study that might be called theoretical organology.

---

[1] By 'gesture' we will refer to bodily gestures, and not use the more metaphorical sense of 'musical gesture' as a characterization of segments of the music itself.

The construction of new instruments and mappings raises a number of considerations. One set concerns the listener: what are reasonable and natural mappings? What kind(s) of sounds do we expect when we see a certain gesture? Another set concerns the performer: what are convenient and exciting mappings that stimulate creativity?

This is not to say that new mappings should follow such expectations, but it would be interesting to have relatively detailed data about this to be able both to make mappings that conform to *and* to make mappings that conflict with the mentioned expectations. In both cases a set of conceptual tools are needed; and if we are to be able to put together insights from the great number of studies already done in the field, we need to start discussing how we might be able to coordinate terminology used in our work.

2. THE THREE PERSPECTIVES

There is of course a close connection between overall perspective and terminology. We have already mentioned two different perspectives, that of the performer and that of the listener. A third is that of the instrument constructor. In the following, we will describe three versions of models of the overall performer-instrument system corresponding to these three perspectives; then we will discuss approaches to descriptions of various parts of these models.

2.1 The Listener

Seen from the perspective of a listener, it is relevant to be able to characterize the overall relation between a gesture and the related sound, as illustrated by the simple diagram in Figure 1.

Gesture ⟶ Musical sound

Figure 1: Gesture - Sound

The musical sound may in turn invoke some kind of activity in the listener or audience, like dance or, in a concert hall, just quiet sighs followed by applause at appropriate moments. In Figure 11 there is therefore a distinction between Gesture P(erformer) and L(istener).

The gestural activity of the audience may in turn influence the musicians and their gestures; and we will have a closed loop of information flow, like in Figure 3.

In this paper, where we the focus on the description and

Gesture P ⟶ Musical sound ⟶ Gesture L

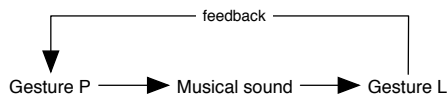**Figure 2: Musical sound also evoking new gestures**

Figure 3: Closed information loop

construction of musical instruments rather than the interplay between musicians and audience, this larger loop is not in the center of attention. Here, we will concentrate on how the listener perceives the interplay between the musician, the instrument, and the resulting sound, as seen from a smaller or greater distance. We want to characterize how the the actual connection between the gestures and the musical sound; the *mapping* from gestures to sound, is perceived.
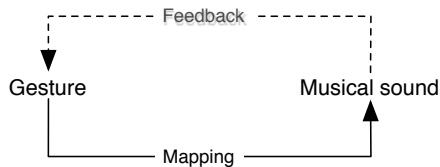

Figure 4: Model of the listeners perspective

The possible feedback part of the loop is, as indicated in Figure 4, not in the center of interest under this perspective.

Seen from a distance, smaller details of finger movements will probably not be as important as larger bodily movements, and one may not be able to distinguish easily between sound-producing gestures and other movements. Purely expressive and/or optional gestures may nevertheless be experienced as relevant to the sound by the spectator, and it is not obvious where to draw the analytical line. As Wanderley et. al observe in their paper on clarinetists' ancillary gestures: "These findings show that clarinetists' ancillary gestures are not randomly produced or just a visual effect, but rather they are an integral part of the performance process." [9, 98]

To account for this, descriptions of the mapping between gesture and sound at this level might therefore include several levels of resolution.

2.2 The performer

In the perspective of the performer, we add a *device* and include the feedback. The performing subject is only implicitly represented in the figure, as the agent performing the Gesture, and receiving the Feedback.

In the literature, we find several models that amplifies on different aspects in this basic scheme. One rather abstract model is found in [6], and includes just the performer's perceptions and intentions about the playing on the one hand, and what is called the *instrument control* on the other (see Figure 6). Implied in the model is that the feedback from the instrument control to the performer is in terms of sound, understood as music. This model is focused primarily on
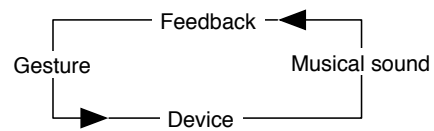

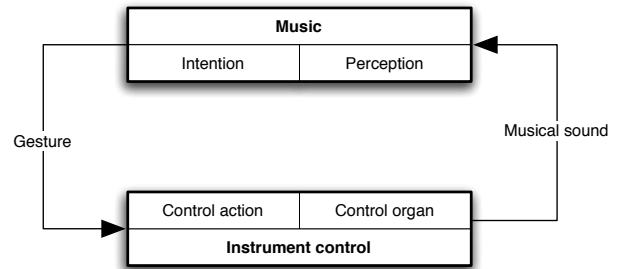**Figure 5: A first approximation of a model of the performers perspective**

Figure 6: The playing technique perspective after [6]

the performer, to the extent that the instrument as such is not present; only the parts of the instrument that are sensitive to control. These are called *control organs* in this model, and we will use this term to denote instrument parts like keyboards, buttons, finger-holes etc., as well as sensors of various kinds that are put to use for controlling sound-producing hard- and software.

Choi's model in Figure 7 is quite similar, but a bit more detailed, as the instrument here also includes the sound-producing parts, and the instrument control, here called *interface*, is exemplified somewhat. Both models lack feedback other than the sound; haptic and tactile feedback from the interface is not considered.
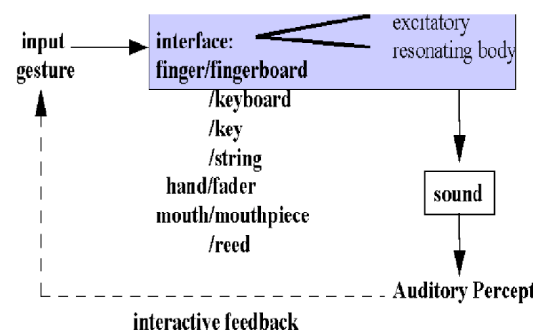

Figure 7: Interactivity of solo performer with musical instrument. After [2]

A still more abstract model is found in [4] as shown in Figure 8.

Here we also find the same view of the instrument control, where performer actions and input devices corresponds to the control actions and control organs in Figure 6. Here, however, the performer and feedback loop are not shown;
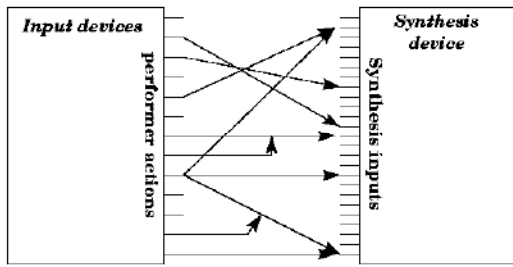
Figure 8: Mapping of performer actions to synthesis parameters. After [4]

instead the sound producing device is included. But this model highlights the mappings between performer actions and the sound producing device as a possibly rather complex system of connections.

A more comprehensive model should include the mappings gesture/control organ and control organ/sound, but in the musical information flow, the inner workings of the sound module is not of interest to the performer, so that part is reserved for the constructors perspective.
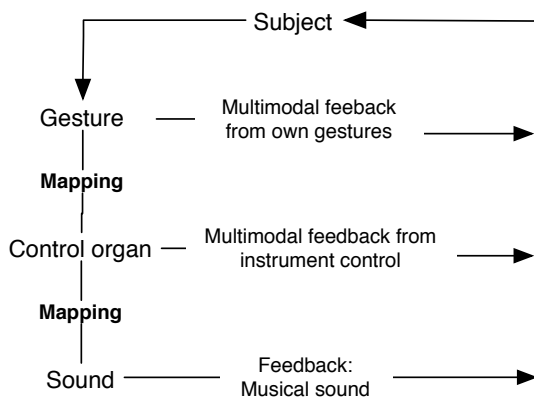

Figure 9: A model of the performers perspective

Also important, is to show explicitly the feedback from the different parts of the system. The performer may receive and use feeedback from his own gestures, from the contact with the control organs of the instrument, as well as from the actual sound produced. Sometimes the feedback is essential it is literally impossible to play a theremin without hearing the sound and pitches produced. In other cases feedback may be helpful, but not essential, like the visual feedback from a piano style keyboard, as is illustrated by e.g. professional, blind pianists.

2.3 The Constructor

The constructor needs a far more detailed view of the system. There are several models in the literature, each focusing on different aspects. A relatively comprehensive model might look like the one in Figure 10.

This figure highlights the fact that there are many different mappings, involving different kinds of parameters, and
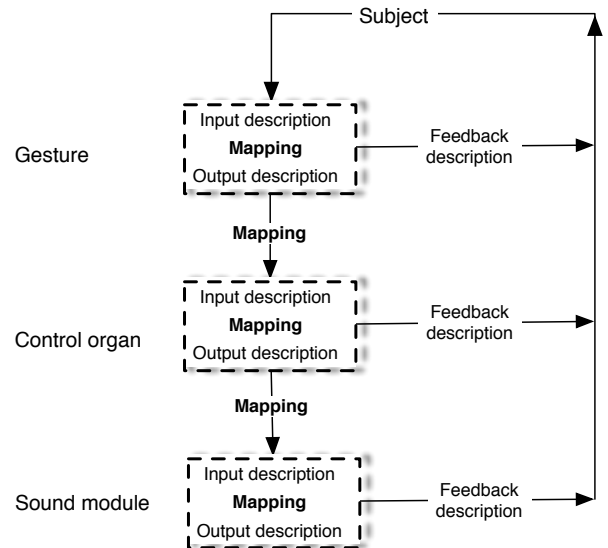

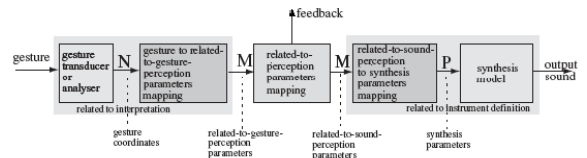Figure 10: Model of the constructors perspective

Figure 11: Mapping chain after [1]

also that the mappings may be viewed as chained: the output of one mapping is the input to the next. Also, the phenomena to be described are quite diverse; from space-time trajectories (gestures) to interface layout to sound synthesis descriptions to descriptions of sound as music. To a constructor all of these kinds of descriptions may be relevant at some point; at the same time precisely the connection between the diverse set of descriptions is at the center of interest. This model is quite similar to the one found in [1], where the interesting concept of related to-perception parameters is introduced, (see Figure 11), but slightly more explicit in the number of descriptions and mappings.

Unlike Arfib's, however, the model in Figure 10 in itself does not say anything about actual parameters to form the descriptions, or about the mappings between them.

2.4 Technical vs. musical construction, or the role of information and energy

In models of the kinds shown here, we will find accounts of both information flow and flow of physical energy, and it is not necessarily obvious whether Figure 10 refers to information or energy, or both. One reason for this, is that the concept of 'energy' is used in many different ways, both as a concept of physics, but also to describe perceived qualities of music, like when we talk about 'energetic playing', 'forceful sounds' etc.

On closer inspection, it is obvious that the upward flow

in the figure refers to information exclusively, and not to physical energy, as the only entities flowing are *feedback descriptions*. This is of course not to say that physical energy is not involved in the process, but only to point out that what we are interested in, also concerning energy, is the information conveyed.

When a key is struck on a piano, the performer will of course use energy to do this, and the energy is musically relevant in the sense that the energy used will determine the force with which the hammer strikes the strings, that in turn will determine the energy of the sound produced, that in turn will determine the energy reaching the ear of the performer and listener. This energy, however, is not the same as perceived loudness, when loudness is seen as a musical parameter. While the physical energy level decreases with the square of the distance from the sound source, the perceived musical loudness may be almost invariant over quite a distance from the sound source. Relevant to the performing subject is not the actual energy used to produce a certain energy level in sound output, but the *difference* in energy needed to produce a *difference* in a perceived musical parameter.

In general: as far as the figure is seen as a model of music-making, one should only be concerned with information. The *musical construction of an instrument* is a matter of information processing.

The energy as such, however, is clearly relevant in the context of *technical* construction. The piano key must be constructed so that differences in output levels, corresponding to musically meaningful differences, are physically feasible for the performer. And the chain of energy that carries the information around the circuit must be geared to the equipment used at the different stages.

Such considerations may be kept separate from the musical construction of the system, and we believe the musical construction will become clearer if this separation is made.

3. PARAMETERS

3.1 Parameter types

The parameters involved in the description may be organized in the following general types:

- Gestural parameters

- Technical parameters

- Musical parameters

Typically, the input parameters of the whole system will belong to the first type, while the output will, as argued above, best be described in terms of musical information.

The different types of parameters must be described in different ways, and the challenge is to find ways to connect the descriptions to find common or corresponding properties in the different kinds of description.

3.2 Parameter description

In this section, we will concentrate on properties that are valid across different kinds of parameters and their description in detail. This concerns level of specificity, the distinction discrete/continuous, and the concept of measurement levels. These are all considerations that may be helpful when looking for how properties of parameters at one point are reflected in properties of parameters at another point in the chain.

3.2.1 Levels of specificity

It is obvious that there is a need for descriptions on different levels, or with different degree of detail. The gestures used by the performers might be best described in rather broad terms in the perspective of the listener, while one will need a far greater amount of detail when describing them from the performers perspective and the constructor may need even greater amount of detail.

3.2.2 Musical parameters

One of the real challenges in this connection, is to define relevant description of the musical output. At a very general level, the parameters *pitch, loudness, timbre* and *duration* may be used. As soon as one wants more specific descriptions, however, there is a large number of possible descriptions, and meaningful descriptions are very much depending on musical style, as well as on experience with the instrument in question

The descriptions needed and felt relevant, develops with experience, and with development of new instruments. From the very start of synthesizer construction, a large number of new parameters, especially for timbre, became available for exploration and incorporation in a musical practice through the new user interfaces. More recently, a similar development concerning the musical manipulation of time in various software packages, widens our musically relevant parameters in the field of rhythm, tempo and time.

In this context, we will touch upon only the most general level.

3.2.3 Levels of measurement

One way to characterize parameters in a general way, is through the concept of *levels of measurement* (see [3]). One may distinguish four such levels:

- Nominal level: Values may only be distinguished from each other, and not ordered.

- Ordinal level: Values may also be ordered in a sequence.

- Interval level: Values may be ordered, and there is a way to measure distance between values.

- Ratio level: Values may be ordered; distances measured, and there exists an absolute zero value, so that division and multiplication of values are meaningful operations.

At a general level, these four levels may be associated with the four general musical parameters in the following way:

- Nominal level: Timbre. There is no generally accepted way to characterize timbre as a perceptual entity in an ordered sequence. Should an oboe sound come between a piano and a flexaton, or before the piano? On the other hand, it is possible to construct scales for *aspects* of timbre, like e.g. brightness, and thus treat part of the timbre variable at a higher level of measurement.

- Ordinal level: Loudness. It is obvious how to order loudness levels, but not necessarily how to describe precise *intervals* of loudness.

- Interval level: Pitch. Musical pitches are ordered into classes with perceptually meaningful comparable distances, like half tone, whole tone etc. But there is no meaningful zero point.

- Ratio level: Durations may be ordered and measured, and there is an obvious zero (no duration), and ratios are meaningful — a half note is half the duration of a whole note.

The implication for the control of musical instruments is that an input parameter has to be on at least as high a level of measurement as the output parameter it is meant to control. But there is not always a need for controlling an output variable at the maximum level, as will be explained in more detail below.

3.2.4 Discrete vs. continuous parameters

The perception of the four general musical parameters may be described as having two different dimensions, *discrete* and *continuous*. The discrete dimension is tied to categorization or sets of concepts; the continuous to gradual variations. Pitch may be perceived both as belonging to discrete pitch-classes (c, d etc), *and* as a continuous entity as in vibrato and glissandi. Similar considerations may be done for the three other general parameters.

There are several implications of this for the connection between input (gestural) parameters and output (musical) parameters. First of all, a continuous output parameter needs a continuous input control if it shall be possible to control it in detail. But it is also possible to use a discrete input to *trigger* a preprogrammed continuous variation of output like a vibrato on an synthesizer. On the other hand, a discrete output may be controlled by a discrete input, like pitch classes controlled by a keyboard, but there are also numerous examples of a continuous controller controlling a discrete output parameter, like a trombone slide controlling discrete pitch classes (as well as continuous pitch variations).

Output

	Discrete	Continuous
Discrete	Piano keyboard: discrete Pitch	LFO triggers in synths. Rare in acoustic instr.
Continuous	Trombone slide: discrete pitch	Trombone slide: continuous pitch Striking force: loudness

(Input — left vertical label)

Figure 12: Combinations of discrete and continuous input and output

The various combinations of discrete and continuous variation for input and output for different musical parameters give different demands and possibilities for the performer, as is generally acknowledged by teachers and students of performance. Nevertheless, we have not much general knowledge on the effects of the different combinations.

3.3 The description of mappings

3.3.1 The mapping chain

In the three perspectives described above, the mapping chain is descibed in increasing detail, with just one mapping from gesture to sound in the listeners perspective to five different mappings in a chain in the constructors perspective.

All the mappings in the last chain will at some point have to be described during the construction of an instrument. They will, however, in a sense be subordinate to the overall gesture-to-sound mapping, that is the defining mapping for the instrument — given this mapping, the technical solution of the intervening mappings is not important to the workings and identity of the instrument. In the following, only the overall gesture-music mapping is considered.

3.3.2 Basic mapping strategies

There are three basic approaches to the actual mapping. The traditional way is to describe and construct a fixed or static mapping, where the relations between input and output parameters stay the same. Traditional acoustic instruments are usually well suited to such descriptions.

Another possibility is a variable mapping, where mappings may be changed by the performer. Common commercial synthesizers are good examples.

Finally, mappings may be the outcome of a dynamic (learning) process where the performer chooses gestures to fit possible sound, or the mapping is modified by the actual behaviour of the performer (see e.g. [1] and [8]).

All three cases need a common way of describing mappings.

3.3.3 A general mapping description problem

Several authors mention that mappings may take different forms with regard to how many input parameters are controlling how many output parameters. In [6], the term coupling is used, differentiating between control couplings and sound variable couplings, corresponding to the many-to-one and one-to-many example in Figure 13 respectively. (See also [1] and [4]).
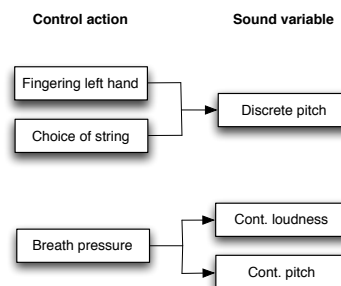
Control action	Sound variable
Fingering left hand	Discrete pitch
Choice of string	
Breath pressure	Cont. loudness
	Cont. pitch

Figure 13: Examples of many-to-one and one-to-many mappings of input to output.

While there are a number of high-level characterizations available, there is a need for a more detailed typology of

mappings. Traditional acoustical instruments may be a good starting point, because they represents a quite diverse set of possibilities, that are well known through long practice.

One way is to make general overviews of mappings for a number of instruments, where control actions are mapped to the generalized musical parameters as described above. In [6], this is done for 23 instruments, and a few general points are made.

Clarinet	Pitch		Loudness		Timbre	
	C	D	C	D	C	D
Fingering		☐				
Form of mouth cavity					☐	
Lip control	☐	☐	☐		☐	☐
Mouthpiece position			☐		☐	☐
Wind pressure	☐	☐	☐			☐
	☐	☐	☐		☐	

Figure 14: Mapping chart for clarinet. C and D means Continuous and Discrete (see text for explanation)

With such a basis, it might be possible to identfy gestural ensembles as an entity at the same time slightly more detailed and general than the gestures used to control a single instrument. Bowing, string-stopping, finger-hole-fingering, piano-type-keyboarding, button-accordion-buttoning etc. might be examples of such ensembles. Such entities will represent a specification of the control of single instruments, and at the same time be useful over a range of instruments; possibly useful also as a basis for a more general typology of gesture-to-sound mappings in musical instruments.

Also relevant to such a typology, is the more detailed model found in [7], where an analysis of a single musical tone is used to develop terminology for the description of what is called the musical control space (p171). An overview, showing the five components a musical tone described, is shown in Figure 15.

There is, however a problem of a general kind here. While the input parameters may be relatively well-defined gestures, the same is not the case with the output parameters. As touched upon earlier (3.2.2), meaningful descriptions of musical output in more detail, depends heavily on musical style and experience with the musical instrument one is discussing.

This means that the development of a more general mapping theory, also depends on a development in *music theory*, in the sense that we need a theory of relevant musical entities, what they are, and how they are related. Without some kind of consensus in this field, no coherent theory of mappings is possible.

4. CONCLUSIONS

As is obvious from the few articles cited so far, there is no general agreement on terms. This holds on almost every level of description, and almost every aspect of the models presented here. It is an open question on what level and to what extent the field would benefit from a better coor-
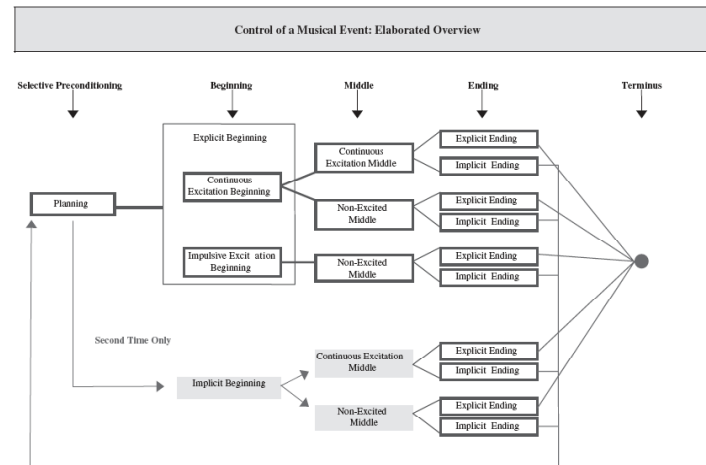

Figure 15: Overview of five components of a musical tone after [7]

dination of terms, but we believe that models of the kind presented here may help in locating a diverse terminology on a common map, to better see existing agreement blurred by diverse terminology, as well as highlighting real differences of opinion.

This paper has not adressed questions of more detailed descriptions of useful parameters and mappings. In [5], we discuss these questions further.

5. REFERENCES

[1] D. Arfib, J. Couturier, L. Kessous, and V. Verfaille. Strategies of mapping between gesture data and synthesis model parameters using perceptual spaces. 2002.

[2] I. Choi. Gestural primitives and the context for computational processing in an interactive performance system. In M. Wanderley and M. Battier, editors, *Trends in Gestural Control of Music*, pages 139–172. Ircam - Centre Pompidou, 2000. 352.

[3] O. Hellevik. *Forskningsmetode i sosiologi og statsvitenskap*. Universitetsforlaget, Oslo, 1971.

[4] A. Hunt, M. M. Wanderley, and R. Kirk. Towards a model for instrumental mapping in expert musical interaction. In *Proceedings of the 2000 International Computer Music Conference*, pages 209–212. International Computer Music Association,, San Fransisco, 2000.

[5] A. R. Jensenius and T. Kvifte. Towards a gesture description interchange format. *Submitted to NOME06, Paris.*

[6] T. Kvifte. *Instruments and the Electronic Age*. Solum Forlag, Oslo, 1989.

[7] D. J. Levitin, S. McAdams, and R. L. Adams. Control paramteres for musical instruments: a foundation for new mappings of gesture to sound. *Organised Sound*, 7(2):171–189, 2002. 344.

[8] D. J. Merrill. *FlexiGesture: A sensor-rich real-time adaptive gesture and affordance learning platform for electronic music control*. Master of science in media arts and sciences, MIT, 2004.

[9] M. Wanderley, B. W. Vines, N. Middleton, C. McKay, and W. Hatch. The musical significance of clarinetists' ancillary gestures: An exploration of the field. *Journal of New Music Research*, 34(1):97–113, 2005.