

## **What to Learn: The Embodied Structure of the Environment**

*Aug Nishizaka*  
Department of Sociology  
Meiji Gakuin University

The purpose of this article is to demonstrate, through the detailed analysis of videotaped violin lessons, that when one learns something, one learns a restructuring of the world. The issues I address in this article include how the restructuring of the environment is accomplished in the actual course of instruction giving, how this restructuring of the environment is intersected with the organization of participation in interaction, and how performance after the instruction is organized to display sensitivity to the achieved structure of the environment such that the teacher can refer to this sensitivity in her evaluation of the performance. In conclusion, I explicate how this study can contribute to further research on learning.

### **SEEING STRUCTURES OF THE WORLD**

In this article, I focus on several very small fragments of Japanese interaction in violin lessons that a 4-year-old child takes from a professional violinist. In the course of the detailed analysis of these fragments, I show how the participants perceive and make use of structures of the environment to restructure that environment. The achieved structure is thereby involved as an integral part of what the child learns. This is not a feature

---

Many thanks to Dom Berducci, Chuck Goodwin, Makoto Hayashi, and Tomoyo Takagi for all their valuable comments at various stages of the development of the analysis and argument in the following pages. Particularly, without Chuck's encouragement, I would not have been able to finish this article. Dom's very detailed comments were extremely helpful for putting it into shape.

Correspondence should be sent to Aug Nishizaka, Department of Sociology, Meiji Gakuin University, Minato-Ku, Tokyo 108–8636, Japan. E-mail: [augnish@soc.meijigakuin.ac.jp](mailto:augnish@soc.meijigakuin.ac.jp)

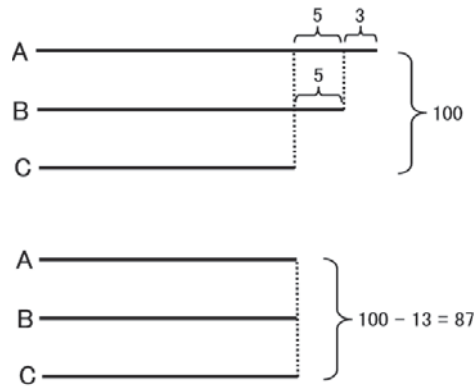


FIGURE 1 The perceptual structuring of three lines in a calculation.

unique to violin lessons. To make the point clear, a familiar example for how children learn to make a simple calculation may be helpful. The following is a question: If three children, *A*, *B*, and *C*, have 100 pencils in total, but *A* has three more than *B*, and *B* has five more than *C*, then how many does each of them have? Some elementary school children who are not familiar with equations at all are taught to solve this kind of question by using three parallel lines. Children are supposed to learn how to perceptually structure those lines, that is, to see a specific perceptual configuration (or *Gestalt*) of the lines, which is something *more* than the physical arrangement of three lines parallel to each other. This perceptual restructuring of lines helps children develop their understanding of the meaning of this particular calculation. Children around the age of 10 or 11 years learn to perceive the specific structure of the lines. If one works out the structure of the parallel lines, one will get 87 by cutting off the parts sticking out, and then one will get the number of pencils that *C* has by dividing the remainder, that is, 87, by 3 (see Figure 1).

Similarly, calculations with equations involve perceptual structuring of written signs. If one attempts to solve simultaneous equations, one will find immediately that one needs to perceive some specific structures of strings of signs beyond their physical arrangement. One may write down the following equations to solve the preceding question:

$$X + Y + Z = 100$$

$$X - Y = 3$$

$$Y - Z = 5$$

For these equations to make sense, one needs to perceive various symbolic structures such as the distinction of both sides of each equation,  $X$ ,  $Y$ , and  $Z$  being of one class and  $+$  and  $-$  of another and so on. Furthermore, one also needs to see some other structures of those equations to operate on them. If one perceives the following structure, marked with a box and a brace

$$\left. \begin{array}{l} \boxed{X} + Y + Z = 100 \\ \boxed{X} - Y = 3 \\ Y - Z = 5 \end{array} \right\}$$

then one goes on to delete  $X$ s. Then, one may see the following structure:

$$\left. \begin{array}{l} 2Y + \boxed{Z} = 97 \\ Y - \boxed{Z} = 5 \end{array} \right\}$$

and goes on to delete  $Z$ s, and so on. Because one has learned the inscription system of equations and is able to perceptually structure strings of signs by following the system, one can operate on equations. There is no substantial difference between drawing three lines and operating on equations in the sense that one needs to learn to properly perceptually structure these physical objects—that is, an inscription on a sheet of paper—to use these objects.

Children may be able to learn how to perceptually structure the objects properly by themselves, but it would be much easier to learn if taught by adults. Indeed, some children may study it by reading, but they ultimately learn it through mediated interaction with adults. In what follows, I attempt to show how new perceptual structures of the surroundings are achieved in and through the actual development of interaction between a teacher and a child and how these achieved structures are related to the achievement of learning in interaction.

## LEARNING IN INTERACTION

There have been studies of interaction that have dealt with learning. Some conversation analytically inspired studies have elucidated the social organization of interaction in the classroom through the detailed analysis of

naturally occurring interaction between teachers and students in terms of the organization of turn taking (McHoul, 1978), action sequencing at a variety of levels in the classroom (Mehan, 1979, 1985, among others), and the like.<sup>1</sup> Mehan (1979), for example, observed that classroom lessons are built from the basic sequence he designated as I–R–E, that is, an initiation (I) by a party (i.e., a teacher), a response (R) to it by another (i.e., a student or students), and an evaluation (E) of the response by the original party. These studies have focused on the specific organization of a process or activity in the classroom. In this article, although I also attempt the detailed analysis of interaction, I rather address the issue of *what* one should learn when one learns. This may sound strange because it is precisely how to calculate or how to play quarter notes that one learns in the previous example or in the analysis to follow, respectively. However, as I suggested previously, learning how to calculate or how to play quarter notes includes, as an integral part, learning to perceptually restructure the world properly. In what follows, I show how the restructuring of the world is interactively accomplished.

The issue I address is divided into two parts: how the teacher's instruction as to how to play quarter notes on the violin is interactively accomplished, and when the child (the learner) plays to the accompaniment of her mother's piano following the instruction 1 month after the instruction was given, how the fact of "the child's performing following the given instruction" is accomplished in the course of their joint performance in front of the teacher. It is, indeed, by reference to this accomplished fact that the teacher evaluates the child's performance positively. In what follows, my main focus is on the first issue. I examine how participants restructure their interactional environment in the course of their interaction while making use of the various pregiven structures of that very environment. Then, I attempt to show how participants are *visibly* jointly oriented to that achieved environmental structure in their joint performance so that the teacher can appreciate their joint orientation. The teacher's positive evaluation is based on what the child actually did in front of the teacher, not on any kind of inference from the child's behavior about what takes place inside the child. Indeed, whatever (experiential or neurological) processes or events take place inside the child is irrelevant to the evaluation of the child's completion of a task. Similarly, the ascription of learning to someone is only because he or she completes with some regularity what he or she was not able to complete before, and what takes place inside the individual is also irrelevant to the ascription of learning.

I also attempt to show in the course of the analysis that the restructuring of the environment and the ascription of learning belong to the local order of activity. This may remind some readers of Lave's (1988; Lave & Wenger, 1991) argument, which may be summarized in the following way: Something is wrong with school education to the extent that it cherishes the basic conception that there must be a "core" (e.g., in arithmetic) to be learned that should be the same across all variety of contexts of activity. This conception is mistaken. It supposes that if only one learns that "core," one can apply it in different situations almost automatically. However, Lave (1988) claimed, learning is always achieved in a particular way specifically appropriate to each particular context of activity. In addition, although school education is caught up in a mistaken conception, it also forms a particular context for learning calculations in the same way as supermarkets form a particular context for learning calculations. Lave's (1988) argument is very relevant to mine in this article. The notion of a decontextualized core to be learned may, indeed, slip easily into the conception that one carries the learned core somewhere *inside oneself* (in the nervous system or in the mind) over to various contexts of activity if one must keep it somewhere after one learns it. The purpose of this article is to elucidate a local order of activity in which the structuring of the world to be learned is achieved and visibly oriented to in the completion of the assigned task. Even though school education is practiced based on a mistaken conception, insofar as participants ascribe learning among themselves there *intelligibly*, they may also be expected to structure and restructure the environment of their activity appropriately to the local order of the activity. How is this (re-)structuring accomplished? This is the issue that I address in the following analysis.<sup>2</sup>

## LEARNING HOW TO PLAY THE VIOLIN

### Data

The data I analyze here are videotaped fragments from violin lessons in which a professional violinist teaches a 4-year-old child how to play quarter notes. Both participants' native tongue is Japanese. I videotaped their weekly lessons for 13 months, although I was not able to record all the lessons during the time period. In the exchange cited as Extract 1 following, the teacher, designated as T, instructs the child ("Hana-chan") on

how to play quarter notes for the first time. Note that *wanwan* in the extract represents a pair of quarter notes. The teacher uses animal sounds to represent notes: *wanwan*, the sound of dogs, represents quarter notes; *nyaago*, the sound of cats, half notes; and *gawoo*, the sound of lions or tigers, whole notes. The teacher mentioned *wanwan* en passant in the previous lesson, but the exchange reproduced as Extract 1 is designedly the first opportunity for her to introduce what exactly to do to play quarter notes. Indeed, regardless of whether or not this is actually the first opportunity, this opportunity is designed as *first* by the teacher in her talk, as one will see.

Immediately before the exchange in Extract 1, the child was playing whole notes, which she has been practicing for weeks. The teacher marks the end of the preceding performance by saying “thank you” to the child (at line 01) and taking the bow and then introduces a new task for the day, proposing that they do *wanwan*, that is, quarter notes (at line 02). The teacher gives instruction on how to play quarter notes (*wanwan*) at line 09. Note that the teacher highlights this instruction as what is to be specifically remembered at the moment by framing it with “*oboete ne*” (“remember this, OK?” at line 05) and “*oboera reru ka na?*” (“Can you remember?” at line 12) from the *front* and the *rear*, respectively. I also note that the child’s mother observes this lesson off camera from one corner of the room:

(1) [Violin Lesson 2 31]<sup>3</sup>

- 01 T: *hai arigato, hana chan, (.) jaa:: ne, kyoo, (.)*  
OK thankyou PN P then P today
- 02 *hana chan no sukina wanwan de ikoo ka.*  
PN P P like bowwow P let’s go IR  
“OK. Thank you, Hana-chan, then, today, let’s try *wanwan* [bowwow],  
(because you) like ((dogs)).”
- 03 (0.6)
- 04 H: *i::(h)hi(hhh [h])*
- 05 T: [wan, (.) wan, te iu no wa ja ne, oboete ne:::  
bowwow P called P then P remember P  
“The thing called *wanwan* is, then remember this, OK?”
- 06 (0.6)
- 07 T: *hai::, hana chan mite:::*  
OK? PN P look  
“Hey, Hana-chan, look ((at this)).”
- 08 (1.4)
- 09 T: <yumi no mannaka de::,> wanwan hi(ite) mi mashoo.  
bow of middle at bowwow play try let’s

- 10 “At the middle of the bow, let’s play *wanwan*.”  
(0.6)
- 11 ((sounds))
- 12 T: *oboera reru ka na::::?*  
memorize can IR P  
“Can you remember?”
- 13 (0.8)
- 14 T: *oboera eru. (.) yoshi.jaa, <chotto> hiite mi mashoo.*  
memorize can good then a bit play try let’s  
“You can remember. (.) Good. Then let’s try to play.”
- 15 (0.8)

The preceding description of the fragment glosses over the subtle interactional organization of the instruction sequence. The instruction is jointly occasioned by the participants in a very subtle and artful way.

### Transition

First, I focus on three identifiable interactional objects at line 01: “*arigato*” (“thank you”), “*jaa*” (“then”), and “*kyoo*” (“today”). Saying “thank you” exhibits the understanding that something beneficial to the speaker has been completed successfully. In this context, that is, when the teacher says “thank you” to the student, it reveals the teacher’s view that the child has completed an assignment in a way desirable to the teacher. The teacher’s “thank you” marks the completion of one task, the one that the child has been engaging in. Produced after “thank you,” “*jaa*” (“then”) can be heard to mark what follows it as *next*, highlighting a disjuncture between what follows it and that which precedes it. What job is done with “*kyoo*” (“today”) here? Note that the exchange starts about 30 min after the lesson started. Many things have been done in this lesson previous to its beginning. Indeed, by this time, the participants have completed more than half of the lesson. Introducing what they should do now, nevertheless, by explicitly marking it as what they are going to do “today” at this point in the lesson, distinguishes what follows it from what precedes it and reveals that what follows it is *the* task specifically for “today,” whereas what precedes it was a review of what they did in previous lessons. Saying “today” at this point strongly indicates that what follows it will be a *new* task that has not been brought up before “today.” Taken together, these three objects at line 01 project the introduction of a task that is designedly unfamiliar to the

child and project the launch of teaching what the child has never been taught.

### Proposal

The design of the proposal for doing quarter notes at line 02 also deserves close examination. Whether or not the child remembers that the teacher mentioned *wanwan* in the previous lesson, she may now understand that *wanwan* means some notes simply because *wanwan* is as typical an animal sound as *gawoo* (grrr) and *nyaago* (mewl), which mean a whole note and two half notes, respectively. The teacher uses another item of the same class as she has used to refer to previous tasks so that the child may be able to hear “*wanwan*” as a new task of the same kind as represented by *gawoo* and *nyaago*. Therefore, although the teacher uses the verb “*iku*” (“try”), which functions like a *pro-term*, one can easily understand that she is proposing that they practice playing some notes “today.” Note also that the typical expression of the sound of an animal is often used for representing the animal itself, particularly by and for young children. It is obvious that when the teacher says “*hana-chan no sukina wanwan*,” literally meaning “*wanwan*, which you [Hana-chan] like,” the “*wanwan*” also signifies dogs. The entire utterance at line 02 can be paraphrased as “Today let’s try *wanwan* [quarter notes] because you [Hana-chan] like *wanwan* [dogs].” Insofar as “*wanwan*” can be heard to come from the same class as the other sounds for notes from previous lessons, the double meaning of “*wanwan*” at line 02 should not cause the recipient to misunderstand.

The teacher’s proposal at line 02 makes conditionally relevant its acceptance or rejection in response to it. The silence at line 03 (the child is wiggling her body with a smiling face during the silence) and the laughter at 04 together delay a conditionally relevant response to the proposal, that is, break the contiguity of the *first* and a possible *second* pair part of a sequence type, and may project a possible *dispreferred* response, which is usually a disagreement or rejection (see Pomerantz, 1984; Sacks, 1987). However, the teacher goes on to the instruction of quarter notes as if her proposal has been accepted.

The last thing that one can presume here is that the teacher presupposes the child’s incompetence for interaction. Indeed, the child exhibits her interactional competence in making some response (if not a condition-



ally relevant one) after an action that makes relevant a response from the recipient was addressed to her. Although the child starts to smile immediately after the teacher says “*arigato*” (“thank you”), however, she only starts to wiggle her body in a marked way following the teacher’s proposal, and she never laughs aloud during that proposal.<sup>4</sup> Note also that the proposal at line 02 is constructed such that the reason for it is incorporated into its design. Regardless of whether or not this particular child actually likes dogs, it is generally expected that children of this age should like animals such as dogs and cats. Mentioning that the child likes dogs is hearable as the reason for the proposal, and therefore, the proposal is hearable to be designed as an action that is beneficial to the recipient with respect to the latter’s private, or individual, preferences. Then the child’s laughter, produced immediately after the proposal, can be heard to exhibit an affiliative attitude toward the design of the proposal. That is, it is hearable as getting tuned in, if not an agreement, to what the teacher proposed to do. Indeed, although it may be difficult to see only from the transcript reproduced as Extract 1, the laughter produced at line 04 sounds (and looks from the video) like an expression of delight rather than hesitation. This expression of delight may even display the child’s acknowledgment of the teacher’s thoughtfulness about the child’s preference exhibited in the teacher’s proposal. Although this laughter still delays a now conditionally relevant response, the teacher can take the laughter as a sign of a possible acceptance of her proposal. In this way, the child’s conduct at lines 03 to 04 appears to be very subtly organized to fit in the current status of the sequential organization in interaction.

### Focus of Instruction

The teacher then moves on to instructing how to play quarter notes. At line 05, the teacher marks *wanwan* as a *new* thing by introducing it with the expression “*te iu no*” (“a thing called”) whose use presupposes the possible recipient unfamiliarity of the marked thing. She then moves on to a request for memorization (“*oboete ne*” [“remember”]), which projects what is to be specifically remembered, that is, learned, about this new thing to come after it, although this expression does not fit the grammatical structure initiated by the preceding part of the utterance (i.e., the utterance at line 05 goes like this: “The thing called *wanwan* is, then remember”). As I mentioned previously, this request for memorization and the inquiry about the child’s memorization at line 12 highlight what comes between them as the

focus of the instruction that was projected at line 02 and started at line 05. The new thing that appears in the teacher's instruction at line 09 is "*yumi no mannaka de*" ("at the middle of the bow"), which is pronounced slowly and articulately, therefore emphatically.

I return to the utterance at line 07. I note two points here. First, the middle of the bow is marked as the focal point for what is to be specifically remembered or learned here. Insofar as the middle of the bow becomes the focal point for remembering and learning, it also becomes the focal point for the current activity, that is, the activity of the lesson in playing quarter notes. Although the child should not necessarily look at the middle of the bow as such, the child is now supposed to properly display her orientation to it and organize her participation in the activity through doing so. Second, the middle of the bow is also organized as the focal point for the child's perception of the bow's movements. The child should also learn to perceive, in a proper way, proper movements of the bow that she makes for herself when she plays the violin. Perceiving movements of the bow properly in relation to its middle is now constituted as an integral part of playing the instrument properly.<sup>5</sup> Indeed, when the teacher (positively) evaluates the child's performance 1 month after this lesson, it appears that the evaluation is organized such that it displays the sensitivity to the propriety of the child's perception of movements of the bow.

## STRUCTURING OF THE BOW

### An Environmentally Coupled Gesture

The next issue for instruction is how one plays notes at the middle of the bow. The bow has a pregiven structure. It is long and slender, and therefore, its middle is located midway between both ends (Figure 2a). To see the point being made, compare it with a circular thing whose middle is to be located at its center (Figure 2b). What can be referred to as the middle is constrained by the physical structure of the thing being mentioned. However, it may still be difficult to see exactly how one is supposed to use the bow only by hearing someone say "Play at the middle of the bow." How exactly should the student play at "the middle of the bow"? How should she actually move the bow?

When the teacher gives an instruction as to how to play quarter notes by saying at line 09, "At the middle of the bow, let's play *wanwan*," she

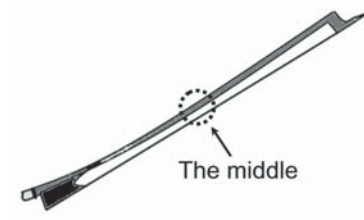


FIGURE 2a The middle of the bow.

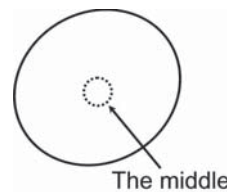


FIGURE 2b The middle of a circular shape.



FIGURE 3 The teacher's hand gesture with two extended fingers.

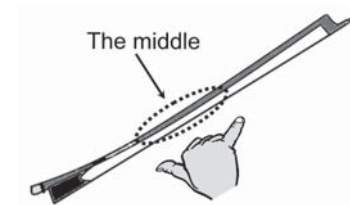


FIGURE 4 The middle of the bow, that is, the range of playing.

raises her right hand with both the thumb and the little finger extended (Figure 3). The teacher puts her thumb and little finger of her right hands over the bow exactly when she says “*mannaka*” (“middle”; Figure 4). As represented as Figure 3, the teacher’s hand gesture with two extended fingers may not have any meaning by itself. However, if this gesture is juxtaposed with the structure of the bow and the talk (“the middle of the bow”), the meaning of the gesture becomes clear. One can see through this juxtaposition that the gesture shows the range indicated by the extended thumb and finger as the range for playing *wanwan*, that is, quarter notes (Figure 4). That is to say, the unique structure of the bow (e.g., its long slender shape) provides a basis for the meaning of the hand gesture. Moreover, the bow, through that juxtaposition, obtains a new structure, that is, a three-part structure, one of whose parts constitutes the range for playing quarter

notes. The configuration of the hand provides a basis for the restructuring of the bow, dividing it into three parts, which are asymmetrically marked because the extended thumb and finger are connected to each other through the hand, that is, the middle part of the bow is directly indicated by the entire hand. Furthermore, the meaning of the expression *play at the middle of the bow* is also achieved through that juxtaposition. The expression now means that one should play using that range of the bow indicated by the teacher's thumb and little finger rather than meaning that one should play around the midpoint, for example. Thus, the structure of the bow, the talk, and the gesture elaborate each other through their juxtaposition such that new meanings are achieved.

Goodwin (2003a) called this practice of juxtaposing a gesture, talk, and the structure of the environment a "symbiotic gesture" (for this practice in various settings, see also Goodwin, 2000a; Hutchins & Palen, 1997; Nishizaka, 2003). Goodwin also occasionally calls the same practice an "environmentally coupled gesture." The basic idea is that the meaning of an environmentally coupled gesture is organized in conjunction with the structure of the environment as well as talk.

### **Embodied Perceptual Structures**

Goodwin (2003a) drew our attention to the case in which a student archaeologist on an excavation with her teacher, after moving a trowel just above the soil being investigated to highlight a feature on it, lowers the trowel into the soil as she moves her hand. The gesture of moving a trowel over the feature is another example for an environmentally coupled gesture; its meaning is only achieved in conjunction with the structure of the soil and the talk produced at the same time as the gesture ("There just sort of circling"). After producing this gesture and talk, in responding to the teacher's request for confirmation ("Do you see right here"), the student, while saying "Right there," moves the trowel into the soil and along the boundary of the feature. Now the student "transduces the shape that is the focus of her gesture from one medium (the moving hand) into another (the soil itself)" (Goodwin, 2003a, p. 34). This gesture leaves a permanent trace on the environment whose structure is being highlighted. Goodwin (2003a) noted that it is useful analytically to "treat these two events as points on a continuum" (p. 34) from iconic gestures that in no way invoke the immedi-

ate surroundings through “symbiotic gestures” (i.e., environmentally coupled gestures) to gestures that change the world being talked about. Certainly, moving a trowel that was lowered into the soil may not be called a gesture. Rather it should be called an operation on the soil. However, for the participants themselves, both moving a trowel over the soil and lowering it into the soil are two practices of highlighting the boundary of the feature before their eyes.

The student archeologist’s action of moving a trowel *over* the soil and the violinist’s hand gesture being examined here are similar in the sense that both crucially depend on the structure of the environment for the achievement of their meaning and are produced to make visible a structure of the environment (the boundary of a feature on the soil and the three-part structure of the bow, respectively). However, the three-part structure of the bow is also a newly created structure, created through the teacher’s gesture. (There is no difference in color or texture on the surface of the bow as opposed to the case with the feature on the soil.) The teacher’s gesture, although the achievement of its meaning depends on the environment, does transform perceptually, not physically, this very environment (i.e., the bow) and gives it a new embodied, perceptual structure, embodied through an actual gesture and talk. It may be said to be located somewhere on the *continuum* of gestures between the two gestures of the student archeologist. What is striking in this respect is that 1 month later, a small green mark can be seen (from the video) to have been attached to a place slightly above the middle point, that is, around the point that the teacher’s little finger extended to. That is to say, after the bow was perceptually restructured in the instruction, the physical operation on the bow (attaching a green mark) has *transduced* the new, achieved structure of the bow from one mode (perceptual) to another (physical). Both the teacher’s hand gesture in the instruction and the green mark attached to the bow are intended to make visible a structure of the bow, but now the hand gesture has been *inscribed* into the physical structure of the bow. Notice that the green mark on the bow as such does not make any more sense than the hand gesture with two fingers extended does by itself. The green mark makes sense only insofar as it is perceived as the inscription of the hand gesture. This small artifact is an incarnated history of instruction (see Hutchins, 1995, pp. 168–169; Pea, 1993, among others). Probably, the inscription stabilizes the new, achieved structure for various usages, although I do not pursue here what is done by inscription.

### Structures of the Bow in the History of Lessons

Then why did the teacher mention in her talk only the “middle” of the bow to show the *range* (rather than the *point*) of the bow that should be used for playing the notes? Why did the teacher not specify in her talk both ends of the range? If she was more explicit about the ends of the range in her talk, the range might be indicated more clearly. Indeed, the one (lower) end is actually referred to as “a little above the silver” (there happens to be a silver band on the lower part of the bow stick; see Figure 5) by the teacher later in the same lesson. The other (upper) end could also have been referred to as “a little above the midpoint” or designated by a small sticker or string, as they will actually do later. However, the restructuring of the bow with the expression “the middle” is very finely accommodated to this stage of the history of their lessons. The “middle” is very well marked out and contrasted with the *tip* and the bottom (the part close to the grip of the bow). Indeed, one has to use the entire range of the bow (i.e., from the top to the bottom of the bow) to play whole notes and half notes, which the child has been practicing for several weeks. Mentioning the “middle” brings into prominence the most crucial difference between *this* task and previous tasks. Now that the two ends of the range of playing are closer to the midpoint of the bow than previously, it can be more difficult to properly locate the departure and end points for moving the bow than when using the entire bow.<sup>6</sup> Both ends of the range at the “middle” do not have any physical marks on the bow, whereas when using the entire bow, both ends of the range for playing are simply both ends of the bow hair. That is to say, *this* task is proposed as one to be assigned only after previous tasks have been (more or less) completed. Therefore, the bow is uniquely structured within the history of lessons and not only within *this* lesson (Figure 5).

In this section, I have explored how the teacher’s utterance at line 09 of Extract 1 is organized such that it intelligibly constitutes an instruction for playing a type of note. I have shown how her talk, gesture, and the structure of the environment elaborate each other to achieve their meaning and thereby to engender the intelligibility of the instruction.<sup>7</sup> On the other hand, bear in mind that the intelligibility of the instruction is also achieved in and through the actual development of interaction. Indeed, in the previous section, I showed how the instruction sequence has been initiated in interaction. For example, the environment that an environmentally coupled gesture is to be juxtaposed with has to be *jointly* oriented to by the participants. Furthermore, to achieve the intelligibility of the instruction, the partici-

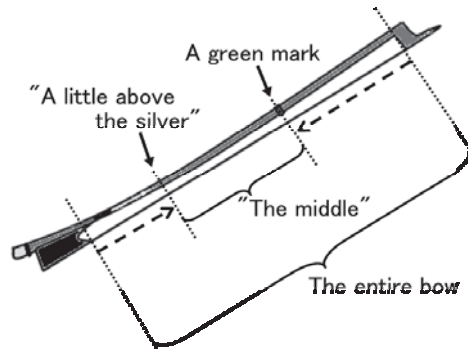


FIGURE 5 The structuring of the bow in the local history of lessons.

participants have to distribute their orientations to each other’s body and the environment in a way appropriate to each stage of the history of interaction. The participants have to display to each other their orientations and jointly manage them by designing their talk and gestures appropriately. The joint management of orientations, exhibited in their talk and gestures, can be called the *organization of participation in interaction*.<sup>8</sup>

### A FOCAL POINT IN PARTICIPATION

#### Focal Point as an Achievement

Earlier, I described a practice by which the middle of the bow was brought to the foreground and highlighted as the focal point for learning and an activity that the participants currently engage in. The practice was *framing* the focal point forward and backward:

- (1a)
- 05 T: [wan, (.) wan, te iu no wa ja ne, oboete ne:::.  
 bowwow P called P then P remember P  
 “The thing called *wanwan* is, then remember this, OK?”
- 06 (0.6)
- 07 T: hai::, hana chan mite::.  
 OK? PN P look  
 “Hey, Hana-chan, look ((at this)).”
- 08 (1.4)
- Forward Framing  
 ↓

09	T:	<yumi no mannaka de:;> wanwan hi(ite) mi mashoo. bow of middle at bowwow play try let's
		“At the middle of the bow, let’s play wanwan.”
10		(0.6)
11		((sounds))
12	T:	oboera reru ka na:::? memorize can IR P
		“Can you remember?”

Backward Framing

We saw earlier that highlighting through the practice of framing is an ongoing contingent accomplishment in the sense that the *forward* framing was only produced when instruction was properly started after the preceding activity was properly completed and the proposal for the instruction was adequately accepted. Furthermore, even if the forward framing is produced, what is designedly framed does not necessarily (or automatically) follow. In Extract 1, what immediately follows the forward framing (“*hai, hana chan mite*” [“Hey, Hana-chan, look”] at line 07) is what is hearably not the item that has been projected by it. *Being the item projected* is also an ongoing contingent accomplishment in interaction.

Now I turn to what happens from line 05 through line 09. Note, first of all, that the teacher starts her utterance at line 07 with “*hai*” plus a summons (“Hana-chan”). The “*hai*” at a turn-initial position, which does not hearably respond to the prior turn, can be a resource for requesting the redistribution of the recipient’s orientations (or, to use a more vernacular but content free formulation, a resource for attempting to *get someone’s attention*). The similar thing holds true for a summons. It may be said that a summons can be a resource for requesting the recipient to redistribute their orientations in a way that is incompatible with the current distribution of orientations (see Schegloff, 1968, 1995, 2002). The teacher initiates the current turn with a nonresponsive “*hai*” plus a summons to mark it as getting off the trajectory of the ongoing activity. The question to be asked is as follows: Why does the teacher need to get off the trajectory?

After the child responds to the teacher’s proposal with laughter at line 04, the teacher, while starting the utterance at line 05 (“*wanwan te iu no wa ...*” [“The thing called *wanwan* is ...”]), shifts her gaze to the bow, and then, synchronized with the expression “*te iu no*” (“called”), brings her hand to the bow with her thumb and little finger extended. At almost the same time as the teacher begins this gesture, however, the child starts to look away from the teacher toward the violin at her chest (see Figure 6). That is, while the teacher is oriented to the bow, the child starts to be ori-



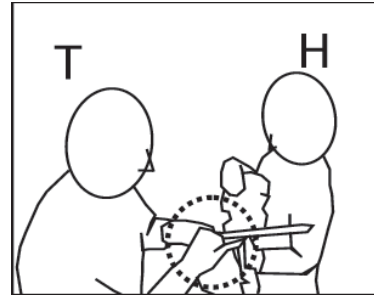
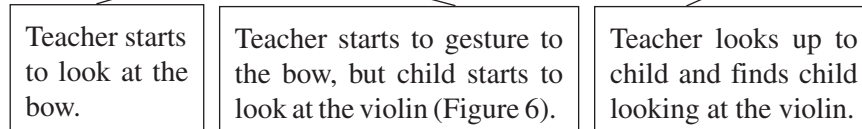


FIGURE 6 05: T: *te iu no wa*  
 Teacher starts the instruction, gesturing to the bow.  
 Child looks at the violin, not the bow. T = teacher; H =  
 Hana-chan.

ented to an object in a different direction. After saying “*oboete ne*” (“remember this, OK?”), the teacher looks up to the child and finds the child looking at the violin. At line 07, as one now sees, the teacher requests the child to look at the bow that the child is now supposed to be oriented to:

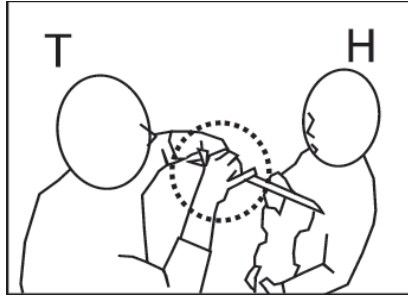
(1b) [05–06]

05 T: [wan, (.) wan, te iuno wa ja ne, oboete ne:::  
 bow-wow P means P then P remember P



06 “The thing called *wanwan* is, then remember this, OK?”  
 (0.6)

After the pause at line 06, that is, immediately before line 07, the teacher raises the bow in front of the child’s face to bring it into the child’s line of sight, encouraging the child to “look” at it (Figure 7). Not only is the bow made salient by it being positioned in the child’s line of sight, it is also raised to the level of their mutual orientation. In this arrangement, even if one is looking at the bow, one can also see where the other is looking. This arrangement makes the mutual coordination of the participants’ orientations to each other and the environment much easier because everything relevant to the development of interaction, including their body parts that display their current orientations most acutely—eyes—, is positioned inside each other’s visual field, that is, each other’s most acute orientational

FIGURE 7-09: *wanwan*

Teacher raises the bow up to the level of their mutual orientation and gestures to the bow in the same way as before. T = teacher; H = Hana-chan.

field. With the bow made salient in this way, the teacher moves on to requesting the child to look at the bow at line 07. During the pause at line 08, the child starts to look at the bow. At this time, not only does she look at the bow, but she also does *looking at the bow* by opening her mouth sharply and holding it open in a marked way:

(1c)

07 T: *hai::, hana chan mite:::*  
OK? PN P look

Teacher raises the bow to the level of their mutual orientation.

“Hey, Hana-chan, look ((at this)).”

08

(1.4)

Child does *looking at the bow* by opening her mouth sharply, and teacher sees child doing *looking at the bow*.

Holding one’s mouth open “idiomatically” indicates that one is absorbed in something and so absorbed that one has lost any concern about others who might be there. (One may be reminded of young children who watch television with their mouth open. They are so absorbed in the program that they cannot remain being attentive to their appearances.) The child opens her mouth *so* wide and *so* sharply that she rather appears to “quote” that behavior to show that she is so concentrated on seeing the bow. In doing so, the child is *doing seeing* the bow. I note that in this context it is insufficient to only factually see the bow. What is important is to see things in the way that other participants can see that one sees them, that is, to *do seeing* things. It may be surprising that even a child at the age of 4 years

knows this and uses available resources for practicing it. The teacher, after seeing the child looking at the bow, resumes her abandoned instruction as to how to play *wanwan* (i.e., quarter notes), keeping the bow at the level of mutual orientation, and puts her hand with two fingers extended over the bow (Figure 7).

After the teacher asks the child whether she can remember what the teacher just explained, the child nods intently during the silence at line 13. Why does she nod so intently? Probably the child is not only claiming strongly that she can remember. She rather appears to display her sensitivity to her (or *their* as I show following) previous failure that she now realizes, claiming that she understands an important point to have been brought out and that she has seized it. Clearly, the actual appearance of the item that has been projected by forward framing is a contingent joint accomplishment in interaction. The instruction giving that has been initiated cannot be completed until the child's orientations have been distributed properly, that is, until the child *does seeing* the bow before the teacher's eyes. The teacher had to establish that both she and the child were jointly oriented to the bow before giving the instruction. It is only after the joint orientation to the bow was secured that the teacher brought her right hand with two fingers extended toward the bow in a salient way. The production of the instruction, or the appearance of what has been marked as the focus of the instruction, is the participants' joint accomplishment (also) in this sense.<sup>9</sup>

The trouble that the teacher encountered was the child's failure to look at the bow, but is it appropriate to describe the child's failure to look at the bow as *failure*? If it is to be appropriate to say that she *failed* to look at the bow, there had to be established the expectation that she should look at it, an expectation not cherished privately but exhibited in public. Relatedly, what was *problematic* to the participants about bringing a particular object in the environment into the focus of interaction in interaction? This question is related to the general issue of the organization of participation in interaction, that is, the proper (re-)distribution of the participants' orientations in interaction. To address these questions, I now turn to lines 01 through 05 of Extract 1.

### **Organization of Participation: Interaction Versus the Bow**

The teacher, while saying "thank you" at line 01, takes the still moving bow with her right hand and pulls it toward herself as if to stop the child's further playing (see Figure 8). When the teacher calls the child's name, the

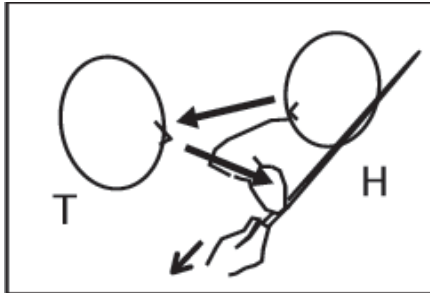
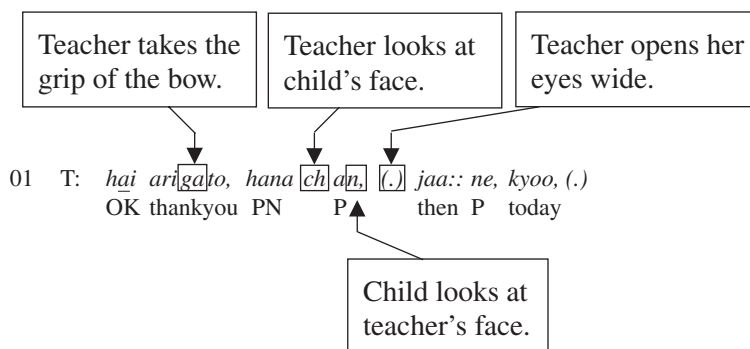


FIGURE 8 01: (.)  
Teacher and child are mutually oriented.  
Teacher is pulling the bow toward herself. T =  
teacher; H = Hana-chan.

child looks up at the teacher's face. The teacher's pulling the bow and calling the child's name indicates publicly that the teacher is now attempting to move out of the current activity. It appears that the child's looking up is responsive to this attempt. In the current activity, that is, playing whole notes, it has been the most appropriate for the child to distribute her orientations such that she should look at the contact point of the bow and a string: That is, the most acute orientation should be shown toward that point. This distribution of orientations is distinctive to that activity so that if the activity is being stopped, it is reasonable to return to the basic arrangement for a speaker and a hearer. When the teacher calls the child's name, the teacher also shifts her gaze from the bow to the child, and here a state of mutual orientation is established. The teacher marks this established mutual orientation by opening her eyes wide momentarily and then moves on to the proposal for "today's" task:

(1d) [01-02]



Generally, it may be improper to keep looking at the other's face. However, the basic arrangement for a speaker and a hearer is one in

which the basic orientation is allocated to each other, displayed on the upper body, or at least the lower body (see Kendon, 1990; Schefflen, 1973). Furthermore, particularly at the beginning of a joint activity, it may be specifically appropriate for participants to look at each other to show and find each other's availability for any activity that each may initiate, as when the summoned looks at the summoner to show their availability for the activity that the summoner intends to initiate. When the teacher sees the child looking at her, it is reasonable to *do seeing* the child looking at her, by opening her eyes wide, to mark their return to the basic arrangement for a speaker and a hearer before moving on to the proposal for "today's" task.

The teacher, after seeing that they are looking at each other, shifts her gaze back toward the bow precisely at the moment when she is starting to introduce "today's" task by saying "*jaa ne kyoo*" ("then today"). The teacher also pulls the bow further toward herself (taking the bow from the child's hand) and puts her (the teacher's) left hand to the bow. This makes the teacher slightly lean toward the bow, with her gaze directed to it. In doing so, the teacher makes prominent her orientation to the bow. The teacher's looking at and leaning toward the bow is designed to be seen by the child and to induce the child's gaze toward it. The teacher is here attempting to *do being oriented to the bow* together with the child. When she turns her gaze back to the child saying "*hana chan no*" at line 02 and finds that the child still looks at her, she erects the bow in front of herself. Then she moves on to mention the point of the task (*wanwan*). What is salient here is that the teacher waves the erected bow precisely when she mentions *wanwan*; this conduct is designed to contribute to the prominent connection of the bow to the task, at the periphery of their mutual orientation (see Figure 9), while, on the other hand, highlighting the task represented by "*wanwan*":

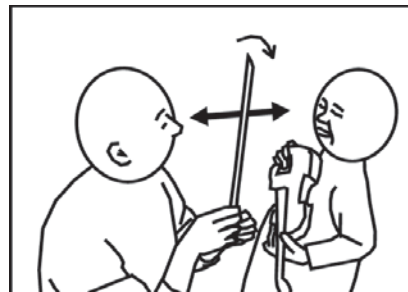
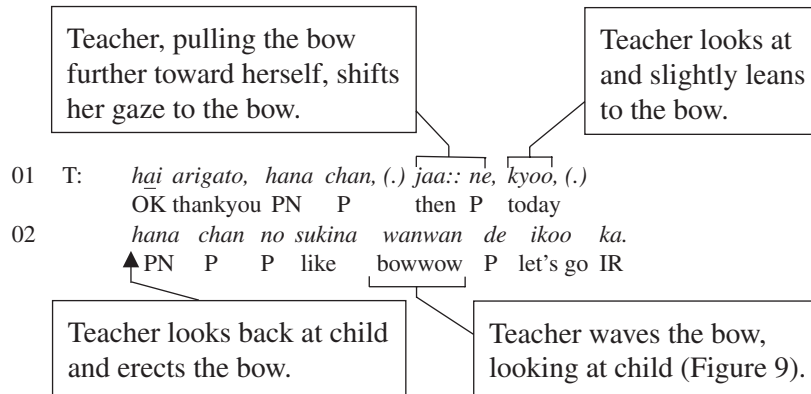


FIGURE 9 02: *wanwan*  
Teacher waves the bow at the periphery of their mutual orientation.

(1e)

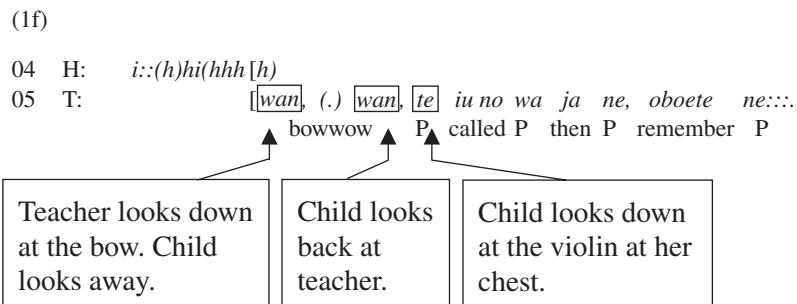


“OK. Thank you, Hana-chan, then, today, let’s do *wanwan*, ((because you)) like ((dogs)).”

I note that it is still natural to propose the task in the basic arrangement for a speaker and a hearer rather than in one in which the basic orientation is allocated to the bow. This is simply because in the proposal the task (*wanwan*), it is only named and is not performed. The participants should be mainly involved with the bow only when the instruction for playing *wanwan* is given after the proposal is accepted. On the other hand, insofar as the task named *wanwan* is teaching/learning how to use the bow to play *wanwan*, the task may also have to be related to the bow during the proposal in one way or another. The peripheral visual (or orientational) field, as I suggested provides the teacher with a resource for establishing the relation between the bow and the task (see Heath, 1986, for the peripheral visual field as a resource for the management of orientations).

After the proposal is accepted, during which the bow has been made adequately prominent and a relation should have been established between the bow and the task, the teacher shifts her gaze down to the bow and starts the instruction. This conduct is designed to induce the child’s gaze to the bow. Taken together, it appears that at this moment, there is an expectation that the child should look at the bow. If the child does not do so, it may be appropriate to describe that she *failed* to do so. However, the peripheral visual field is vulnerable as a resource for making a particular object adequately prominent. This may be a general problem of the organization of participation in interaction, namely, how to reorganize participation from

the current standing arrangement to another one. Indeed, precisely when the teacher starts to look down at the bow during the child’s laughter at line 04, the child happens to look away from the teacher momentarily, wiggling her body while laughing. The child finds, when she looks at the teacher again, that the teacher retreats from the basic arrangement for a speaker and a hearer, which is, in turn, the basic arrangement of interaction as such (as was suggested by Sacks, Schegloff, & Jefferson, 1974, who claimed that “conversation,” namely, interaction between “speakers” and “hearers” rather than interaction between a “doctor” and a “patient,” a “teacher” and “students,” and the like, is the basic speech exchange system). This may have occasioned the child to retreat from interaction as such. The child then starts to look down at the violin at her chest (see Figure 6):



“The thing called *wanwan* is, then remember this, OK?”

In view of what goes on from 01 through 05, it may be more appropriate to describe the child’s failure to look at the bow as *their* failure to coordinate their orientations toward a particular thing in the environment rather than *her* failure. If one of the possible sources of this failure lies in the general problem of the reorganization of participation, then one now sees another possible reason why the teacher raises the bow to the level of mutual orientation when she finds that the child does not look at the bow. She brings the bow not only into the child’s line of sight but into the less peripheral visual field in the basic arrangement of a speaker and a hearer.

In this section, I have examined the local order of interaction in which the new (perceptual) structure of the bow is contingently jointly achieved in interaction. Now I turn to the issue of how the child’s performance 1 month later displays the orientation to this achieved structure.

## MOVEMENT AND ORIENTATION OF THE BOW

During the lesson 1 month after the one in which the instruction was given, the child plays quarter notes to her mother's accompaniment on the piano. Immediately after they finish their performance, the teacher positively evaluates it by saying "joozu da 'tta:." ("It was good"), clapping her hands. I attempt to show that the teacher's positive evaluation may be not only responsive to the sounds that the child has made but to the child's and the mother's joint orientation to the structure achieved in the previous lesson. Indeed, although framed by the sound of the piano, all the stretches of sound that the child produces do not have the same duration. However, the child and the mother are *doing being jointly oriented to that achieved structure* before the teacher's eyes when they play quarter notes together. The teacher is expected to be sensitive to this visible joint orientation in her evaluation of the child's performance. In the following analysis of the child's performance with her mother, I focus on how they organize their performance.

Before going on to the analysis of the performance, I need to describe briefly two more embodied structures of the bow, which I have presupposed but have not yet described in my previous analyses: the perceptual structure of the bow's movement and the intrinsic orientational structure of the bow.

### Perceiving the Movement of the Bow

The following fragment is extracted from the first lesson. The child was practicing whole notes (*gawoo*) after receiving the instruction for playing quarter notes. The teacher was assisting the child in playing whole notes with her right hand supporting the grip of the bow. Now the teacher pulls the bow away from the violin and calls the child's name:

(2) [Violin Lesson: 2 42]

01 T:    hana chan sa .hh gawo:::~::~: 'tte,  
          PN  P  P           grrr    P  
02       koko made iku aida ni,  
          here to   go while  
          "Hana-chan, while ((the bow is)) going to this point  
          with 'Grrr'."



- 03 (0.6)  
 04 T: *miete'ru?*  
       see  
       “Do you see?”  
 05 (0.6)  
 06 T: *yumi ne::.*  
       bow P  
 07 (0.8)  
 08 T: *ko'k kara koko made iku aida ni::?*  
       here from here to go while  
 09 (0.6)  
 10 T: *sense:: ga yottsu te o tataki mas'.*  
       teacher P four hand P clap JD  
       “While you are going from this point to this, (0.6) I will  
       clap my hands four times.”

The teacher is now suggesting that the child play whole notes to the teacher's clapping her hands. At line 01, the teacher, looking at the bow, slides her left index finger upward along the bow while saying “*gawoo*” (Figure 10). She synchronizes the arrival of her index finger at the top of the bow with her saying “*koko*” (“here”) at line 02. The finger points to the top of the bow sharply when it arrives there. The gesture of sliding the finger is another example of an environmentally coupled gesture. The movement of the index finger has its meaning only accomplished through its juxtaposition with the unique structure of the bow and what the gesturer says throughout the entire gesture, that is, “*gawoo*.” One sees that the movement of the finger draws *the movement of the bow* to be made during one whole note (*gawoo*).

Does the movement of the finger in fact draw the movement of the bow? The issue to be addressed here is what the supposed subject of “*koko made iku*” (“is going to this point”) is. (I note that in Japanese, it is very common for subject terms not to be expressed.) It is obvious that “*koko*”

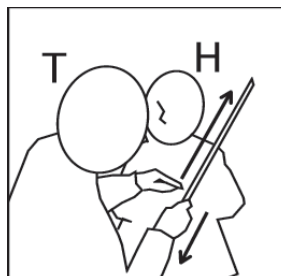


FIGURE 10 01: *gawoo*  
 Teacher slides her left index finger upward along the bow. At the same time, she pulls her right hand on the grip of the bow down.  
 T = teacher; H = Hana-chan.

(“here”) refers to the tip of the bow and that the something that “is going” is going to the tip of the bow. Indeed, the teacher pulls the bow with her right hand by its grip (Figure 10). It may be a string, or the contact point of the bow hair with a string, that the left finger represents. If one follows formal logic, moreover, one may infer that the subject is the *string* given that the thing, which “is going,” is going to a part of the bow and that it is impossible for a thing to go to its own part. However, it is still more natural for me to hear that what “is going” is the bow, even though this hearing contradicts laws of physics. The *string* is rather the last thing that I hear as the subject simply because it is the bow, not the strings, that moves.<sup>10</sup>

If one is instructed to draw the bow to one end, it will be very easy to understand what one should do. Following the instruction, one will draw the bow until the bow hair contacts the string at one end of the bow. Then one may perceive the movement of the bow as the movement *toward* this end regardless of its physical movement. One may perceive the movement of the bow as if one moves the bow from its bottom to the top and returns it to its bottom and the like. This way of perceiving the movement may still make sense in perceptual terms if not in physical terms. In real-life situations, human perception is not governed by formal logic or physical terms.

### **Intrinsic Orientation of the Bow**

The physical features of a bow, especially its long, slender shape, permit one to meaningfully talk of two ends. (A circular shape would not allow it.) A bow naturally has two ends based on its shape. Moreover, those two ends are perceptually structured such that one end, what is called the *tip* or *point* in some practice books, is the top or the upper end and the other end, the part close to the grip (or the frog), is the bottom or the lower end, probably based on the default position of a bow in use. (When one plays the violin, usually the tip is at a higher position than the part close to the grip.) Recall that the teacher refers to one end of the range for playing quarter notes as “a little *above* the silver.” This orientational structure, the distinction of the upper and lower, is intrinsic to bows, as is the distinction of the upper and lower on a human body. Even if a bow is located such that the tip is closer to (the center of) the earth than the grip, it is appropriate to say “the upper part is now down,” as it is appropriate to say of one who stands on one’s hands that the upper part is now down (I follow Levinson, 2003, in

naming this kind of orientational structure *intrinsic*). The distinction of the upper and lower is intrinsic to this particular thing, regardless of the relative position of each part from the earth.

The embodied perceptual structure of the bow that was achieved 1 month before has been inscribed into the bow with a green mark at a place a little *above* the midpoint of the bow (see Figure 5). Now the child is supposed to move the bow at the range between this green mark and that place “a little above the silver.” The bow is therefore supposed to be going *up* to the green mark and coming *down* to “a little above the silver.”<sup>11</sup> The child and the mother appear to be jointly oriented to this structure (of the movement) of the bow during their performance. Here I turn to how they are oriented to this structure.

## ORGANIZATION OF THE CHILD'S PERFORMANCE

### *Doing Seeing the Proper Movement of the Bow*

In their joint performance, the mother persistently looks at the child and beats time with her head while playing the piano. Because of the layout of the room, the child plays the violin standing behind the mother, who plays the piano, and the mother has to have her body twisted to look at the child while she, the mother, plays (Figure 11). This “body torque” (Schegloff, 1998) seems to exhibit the mother's strong, rather than tempo-

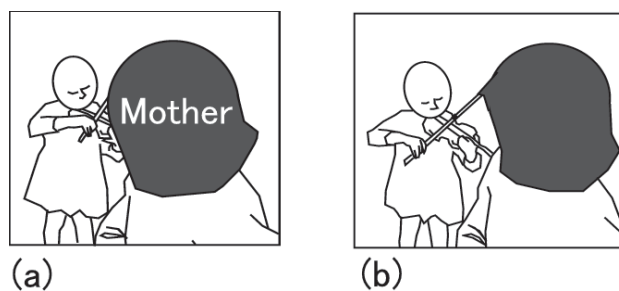


FIGURE 11 The mother looks at the child while accompanying the child on the piano, with her body “torqued.” Compare the position of the bow and the mother's head between Figure 11a and Figure 11b. When the bow is at the lower end of the range (i.e., the bow contacts a string at the lower end of the range), the mother's head is lowered, and when the bow is at the upper end, the mother's head is raised.

rary, commitment to watching the child's performance in the sense that she takes the trouble to look at the child. Of course, the mother's basic involvement is allocated to playing the piano, displayed by the orientation of the lower body, which is fixed on the piano chair (see also Kendon, 1990; Schefflen, 1973). However, it is practically impossible to orient the entire body to the child while playing the piano. To play the piano, it is even preferable to be entirely oriented to it. The mother, while playing the piano, assumes a twisting posture, nevertheless, to look at the child. In doing so, she *does seeing* the child intently. She does so in such an emphatic way that it appears her *doing seeing* the child is also designed to be seen on the peripheral visual field of the child who is continuously looking at the point where the bow contacts the string or the movement of the bow.

What part of the child or her actions does the mother see or *do seeing* then? Obviously not the movement of the child's eyebrows. The only intelligibly relevant thing that she *does seeing* so intently is the child's drawing the bow or simply the movement of the bow. The structure of the bow that has been achieved is not only used as a resource for the child's playing but also provides her mother, who accompanies the child on the piano, with *projectability* of the child's performance in progress. The achieved structure is now inscribed physically into the bow with a green mark attached on a place a little above the midpoint. Together with another available physical mark, the silver part, the green mark makes it possible for the mother to see *where the child is* during each note. The bow coming to the green mark or the silver part projects the coming completion of each current note. The mother can now *see* the coming end of each stretch of sound before it actually occurs, making it easier for her to time her piano to the child's violin and, at the same time, to induce the child to time her violin to the mother's beats (with the piano and head movements), especially when the violin is going off beat (indeed, the child was going off beat at the 13th note). With the green mark and the silver part as markers of both ends of the range for quarter notes, the child herself can see where she (the child) is during each note and accommodate her drawing the bow to the mother's beats. All these things are done by the participants transparently to each other. The fact that the child plays quarter notes following the instruction that was given 1 month before, in this sense, is a joint product of the child's and the mother's mutual timing.

The teacher, who has been observing the child's performance, remarks immediately after the performance is finished as follows:

(3) [Violin Lesson 3 20]

- 01 T: *hai. °waa::°joozu da 'tta::[:*  
 OK wow! good JD PAST  
 “OK. Wow! It was good!”
- 02 H: *[(yoka 'tta de: [:su)*  
 good PAST JD  
 “It was good.”
- 03 T: *[sugoi ne,*  
 excellent P
- 04 *yappa okaasan tsuku to chigau na:: .hhh yoshi,*  
 as assumed mother together when differ P OK
- 05 *jaa tsugi mo ikoo:*  
 then next too let's do  
 “Excellent! Mother's accompaniment makes difference, as expected. OK. Let's do the next one, too.”

The teacher evaluates the child's performance positively and mentions the mother's accompaniment as a possible reason for the achievement. (Note that the expression “*yappa*” [“as expected”] at line 04 reveals that she regards what follows it as expected. This is why the statement at lines 04 through 05 is hearable as a reason for the positive evaluation.) Given that the remark about the mother's accompaniment at lines 04 through 05 was produced in response to what the teacher saw the child and the mother do transparently to each other during the performance, it appears that the teacher's evaluation exhibits her sensitivity to the collaboratively produced character of the child's performance. Incidentally, I note en passant that the teacher's evaluation is very emphatic. The expression of surprise at line 01 (“*wa::*” [“wow”]), together with the emphatic assessment term at line 03 (“*sugoi*” [“excellent”]), indicates that the performance was the best one that had ever been provided by the child. In this evaluation, not only does the teacher evaluate the performance positively, but she also claims with the expression of surprise that this high level of performance has been accomplished *for the first time*.

### Following of the Structure of the Bow's Movement

In the performance, the child played four bars in four-four time, that is, 16 quarter notes in total. The mother strikes one beat with her head to one

note for each of the first 4 notes. However, from the 5th note (until the 14th), she strikes one beat for every 2 notes with her head; laying down and raising up her head on the 5th; and keeping it up at the 6th; and so on. This bodily movement of the mother corresponds to the perceptual structure of the bow's movement. There are two points to be made. First, the fifth stretch of sound begins with the lower end of the range for playing (the place "a little above the silver"), whereas the sixth begins with the upper end (the green mark), and therefore each 2 notes form one cycle in terms of the structure of the bow's movement—that is, one cycle of the bow starting from the lower end, going through the upper end, and returning to the lower end. The way of the mother's beating time with her head appears to be sensitive to this cycle.

Second, it may not be accidental that when the mother strikes one beat for one cycle, she tilts her head *down* on the fifth stretch and keeps it *up* on the sixth, and so on, while beating time with her shoulders and the piano. The up-and-down movements of the mother's head may correspond to the perceptual structure of the bow's movement in the sense that the relative positions of the mother's head are publicly and visibly connected to the intrinsically structured positions of both ends (the *lower* and *upper* ends) of the range for playing or the bow's intrinsic structure with a distinction between *lower* and *upper*. Note that actually both ends contacting a string are at the same level in physical terms insofar as the bow moves on the string that remains stationary (see Figure 11).

Certainly, because the child actually plays 16 notes in four-four time (i.e., four bars each having four quarter notes), and the stresses are on the 1st and 3rd notes in each bar, the mother's head movements may only mark those stresses. Even if this is so, however, the mother still monitors the bow's movement in a marked way and accommodates her head movements, the movement of her shoulders, and the sound of the piano to the bow's movement, thus doing *being sensitive* to the perceptual structure of the bow's movement.<sup>12</sup> Notably, the mother tilts her head down on the very final (16th) stretch of sound, although this stretch is not one with a stress. Note that the final stretch starts at the upper end of the range of the bow and ends at the lower end, that is, the bow "returns" to the departure point of the whole performance at the completion of the final stretch. If the mother kept her head up on the final stretch, her head would remain up when the bow returns to the lower end. By tilting her head down on the final stretch, she permits the relative position of her head to agree with the bow's structured position where the bow contacts the string when the final stretch, and there-

fore the entire performance, ends. She thus marks with that head movement the bow's return to the departure point at the completion of the final cycle.

To sum up, when the instruction for playing quarter notes is given, the bow's new structure is achieved. The child's performance 1 month later with her mother is organized in a way that exhibits its sensitivity to the structure of the bow that was achieved in previous lessons. The child's seeing the bow's achieved structure during her performance is available to the mother as a resource for organizing the joint production of the child's performance, and the entire organization of the joint production of the child's performance is available to the teacher as a resource for her evaluation. I have attempted to demonstrate that all this is visible and accessible, that is, transparent, to all the participants.

Indeed, to ascribe learning to others, one does not examine what goes on inside the skull. The ascription of learning to others can only be based on what lies before our eyes. What the ascription of learning is based on is not only physical structures found in the surroundings but also various embodied perceptual structures in and of the environment. These embodied structures are, insofar as they are embodied, that is, achieved through gestures and talk, accessible to all the participants in interaction.

### CONCLUDING REMARKS

Some may wonder what was going on during the 1 month between the lesson in which the instruction for playing quarter notes was given and the lesson in which the child performed with her mother. When exactly did the child learn to play quarter notes? This question is unanswerable, however. Certainly, many things happened during that time, but is it possible to specify the moment of learning or becoming able to properly play quarter notes? Is it the moment when the child finishes playing 16 notes without a break for the first time? Apart from the possibility that it may be only a fluke, it is hard to answer affirmatively simply because the success in playing must be evidence for her having already become able. Should one then look further for the moment of learning sometime before that moment when she played successfully for the first time? It is now evident that such inquiry would be fruitless, not because it is technically difficult to locate the moment of learning but because it does not make any sense to speak of *the moment of learning*. Learning (as an achievement as opposed to a process or activity) is simply not *something* whose location and time of occur-

rence one can intelligibly inquire about in the same way as the location and time of a process and an event.<sup>13</sup>

I have described what goes on in a very restricted context, that is, violin lessons given to a young child. What I described may seem only relevant to some special species of learning, in particular learning involving training in the use of some physical tools or instruments. However, as I demonstrated at the beginning of this article, even some algebraic calculations involve structuring of the environment. The utility of this article may be direction in how to see the situation of learning. Now we can search for the structuring and restructuring of the environment that is achieved and oriented to with learning. Even students who engage in a task involving no training in the use of an instrument, such as the memorization of European capitals, may be oriented to a perceptual structuring of the environment (the environment may include an atlas, a table with names of countries on one column and names of capitals on another, and the like). In this respect, perhaps, this article should be read itself as an instruction for perceptual (re)structuring of the world, rather than a *theory* of learning (see Garfinkel, 2002).

Clearly, mental concepts such as thinking, memory, vision and the like, are “polymorphous” (Coulter, 1979). The same must be true of the concept of learning. Then, it would not make any sense to speak of what learning in general looks like. However, I believe this article contributes to showing one direction that further research on learning can take. The description in this article has provided one “clear view” (Wittgenstein, 1953) of some aspects of learning.

## NOTES

- 1 In this light, Wootton (1997) is a prominent figure. Based on a long-term, video-assisted observation of a child and the detailed analysis of the video recordings, Wootton formulated some significant patterns of development of interaction competence.
- 2 Certainly, those anthropologists who have attempted to reconceptualize learning through the observation of actual settings (see also Hutchins, 1990, 1995; Jordan, 1993, chap. 7; Wenger 1998, among others), it seems, intend to locate the primordial site for learning in a variety of activities that people actually engage in rather than formal settings, and contextualize learning in those activities to obtain a completely different conception of learning. Through looking at violin lessons, a formal setting for learning, I examine, rather, the *local* order of interaction, that is, the order in particulars and details



of the interactional context of activity, in the same vein as the conversation analytically inspired students of interaction.

- 3 All the extracts cited in this article are composed of three parts: At each numbered line, there is a romanized original Japanese transcript, and below this are phrase-by-phrase glosses. Finally a rough English translation is added after each turn. In the original transcript a transcription system developed by Gail Jefferson is used (see Jefferson, 2004, for its most recent version). In phrase-by-phrase glosses the following abbreviations are used:
 

IR	Interrogative
JD	Judgmental
P	Particle
PN	Proper name
- 4 The child's starting to smile while turning toward the teacher right after the teacher says "thank you" appears to mark the transition from playing the violin, during which the child is supposed to be adequately concentrating on the current business, to what I call later the basic arrangement for a speaker and a hearer.
- 5 For the relation between visual perception and activity, see Goodwin (1994, 1996, 2000b, among others), Goodwin and Goodwin (1996), and Nishizaka (2000).
- 6 The expression *the departure and ending points for moving the bow* may sound queer in terms of formal logic. I return to this point.
- 7 In this context, Goode's (1994) description of the life of "children born deaf and blind" should be mentioned. Goode provided detailed observations on how those children understand and make themselves understood in the structured environment of their home. The observations are good examples of the accomplishment of the meaning of their own and their "interlocutors'" conduct in and through its juxtaposition with the structure of the environment. For this point, see also a series of Goodwin's (1995a, 2003b) work on the "social life of aphasia."
- 8 I owe my discussion of participation in this article to Goodwin's (1981, 1984, 1995b) work on participation frameworks, Erickson and Shultz's (1981, 1982) concept of participation structure, and Kendon's (1990) discussion of "F-formation." Incidentally, Goodwin's (1994, 2000b, 2003a, 2003c) analyses of data from an archeological excavation are most relevant to this article in that they have rich implications for the consideration of learning.
- 9 Insofar as when the teacher attempted to forward frame the focus of the instruction at line 05, not only did the child fail to display any orientation to what the instruction is to be about (i.e., the bow), but also it appears that she had retreated from interaction, so it may be said that the forward framing was ineffective. As I noted previously, however, the child's mother observes this lesson at one corner of the room off camera. The teacher's utterance is designed to be also listened to and heard by the mother, and the teacher's utterance "remember this, OK?" still may serve as forward framing to the mother.
 

Incidentally, if the teacher had not encountered the trouble (with distribution of orientations) that I described, the instruction might have gone without repeating "*wan-*

wan” at line 09 (“*yumi no mannaka de wanwan hiite mi mashoo*” [“At the middle of the bow, let’s play *wanwan*”]). Bringing in the topic, “*wanwan*” (quarter notes), here again may reveal the teacher’s sensitivity to the possible ineffectiveness of her attempt to mark the topic at line 05 (“*wanwan te iu no wa*” [“the thing called *wanwan* is”]).

- 10 I had an opportunity to ask a child at the age of 11 who also had some experience with the violin what she hears “is going” on the videotape. She answered without any hesitation that the bow “is going.” A possible alternative hearing may be that the subject is *you*, although it may still sound formally logically nonsensical to say while ... [you are] going to this point.”  
Incidentally, I noticed another similar case in the same lesson in which the teacher slides her left index finger downward along the bow and then says “have/has to return.” The very same question can be asked: What does the movement of the teacher’s finger draw or what is the subject of the expression “have/has to return”? Here again I would most naturally see that the movement of the finger draws the bow’s movement and hear that the bow is said to have to “return.”
- 11 Practice books for the violin differentiate drawing a bow “up” and “down” in terms of the physical movement of the hand with the bow. When one draws a bow “up,” the bow moves from the *upper* to the *lower* part.
- 12 One may even say that the fact that the practice is organized in four-four time exhibits its sensitivity to that perceptual structure of the bow’s movement.
- 13 Ryle (1949/1963) distinguished between “achievement verbs” and “process verbs.” The verb *learn* is ambiguous in this respect. When it is used as a process verb, it refers to an activity or a process that takes place for certain duration at a certain location such as taking a class or a lesson, trying to memorize Chinese characters by copying them from a textbook, and so on. On the other hand, when it is used as an achievement verb, it registers a success of becoming able to perform an arithmetical calculation, just as the verb *win* registers the success of coming first. Ryle argued about “mental conduct words” that if one ignores the fundamental difference between process verbs and achievement verbs in their use and treat the latter as a species of the former (i.e., as a process, state, or event), one is inclined to look for what happens when someone learns, understands, remembers, and so forth. One is then inclined to look for it in some hidden place because it is not visible. See also Wittgenstein’s (1953, §§ 143–155) discussion on *understanding* (another achievement word).

## REFERENCES

- Coulter, J. (1979). *The social construction of mind: Studies in ethnomethodology and linguistic philosophy*. London: Macmillan.
- Erickson, F., & Shultz, J. (1981). When is a context? Some issues and methods in the analysis of social competence. In J. L. Green & C. Walleet (Eds.), *Ethnography and language in educational settings* (pp. 147–160). Hillsdale, NJ: Ablex.

- Erickson, F., & Shultz, J. (1982). *The counselor as gatekeeper: Social interactions in interview*. New York: Academic.
- Garfinkel, H. (2002). *Ethnomethodology's program*. Oxford, England: Rowman & Littlefield.
- Goode, D. (1994). *A world without words: The social construction of children born deaf and blind*. Philadelphia: Temple University Press.
- Goodwin, C. (1981). *Conversational organization: Interaction between speakers and hearers*. New York: Academic.
- Goodwin, C. (1984). Notes on story structure and the organization of participation. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action* (pp. 225–246). Cambridge, England: Cambridge University Press.
- Goodwin, C. (1994). Professional vision. *American Anthropologist*, 96, 606–633.
- Goodwin, C. (1995a). Co-constructing meanings in conversations with an aphasic man. *Research on Language and Social Interaction*, 28, 233–260.
- Goodwin, C. (1995b). Seeing in depth: Space, technology and interaction on a scientific research vessel. *Social Studies of Science*, 25, 237–274.
- Goodwin, C. (1996). Transparent vision. In E. Ochs, E. A. Schegloff, & S. Thompson (Eds.), *Interaction and grammar* (pp. 370–404). Cambridge, England: Cambridge University Press.
- Goodwin, C. (2000a). Action and embodiment within human interaction. *Journal of Pragmatics*, 32, 1489–1522.
- Goodwin, C. (2000b). Practices of color classification. *Mind, Culture and Activity*, 7, 19–36.
- Goodwin, C. (2003a). The body in action. In J. Coupland & R. Gwyn (Eds.), *Discourse the body and identity* (pp. 19–42). Houndsmill Hampshire, NY: Palgrave/Macmillan.
- Goodwin, C. (2003b). Conversational frameworks for the accomplishment of meaning in aphasia. In C. Goodwin (Ed.), *Conversation and brain damage* (pp. 90–116). Oxford, England: Oxford University Press.
- Goodwin, C. (2003c). Pointing as situated practice. In S. Kita (Ed.), *Pointing: Where language, culture and cognition meet* (pp. 217–241). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Goodwin, C., & Goodwin, M. H. (1996). Formulating planes: Seeing as a situated activity. In D. Middleton & Y. Engestrom (Eds.), *Cognition and communication at work* (pp. 61–95). Cambridge, England: Cambridge University Press.
- Heath, C. (1986). *Body movement and speech in medical interaction*. Cambridge, England: Cambridge University Press.
- Hutchins, E. (1990). The technology of team navigation. In J. Galagher, R. Kraut, & C. Edigo (Eds.), *Intellectual teamwork* (pp. 191–220). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Hutchins, E. (1995). *Cognition in the wild*. Cambridge, MA: MIT Press.
- Hutchins, E., & Palen, L. (1997). Constructing meaning from space, gesture and speech. In L. B. Resnick, R. Saljo, C. Pontecorvo, & B. Burge (Eds.), *Discourse, tools and reasoning: Essays in situated cognition* (pp. 23–40). New York: Springer-Verlag.
- Jefferson, G. (2004). Glossary of transcript symbols with an introduction. In G. H. Lerner (Ed.), *Conversation analysis: Studies from the first generation* (pp. 13–23). Philadelphia: John Benjamins.

- Jordan, B. (1993). *Birth in four cultures: A crosscultural investigations of childbirth in Yucatan, Holland, Sweden, and the United States* (4th ed.). Prospect Heights, IL: Waveland.
- Kendon, A. (1990). *Conducting interaction*. Cambridge, England: Cambridge University Press.
- Lave, J. (1988). *Cognition in practice*. Cambridge, England: Cambridge University Press.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, England: Cambridge University Press.
- Levinson, S. C. (2003). *Space in language and cognition: Explorations in cognitive diversity*. Cambridge, England: Cambridge University Press.
- McHoul, A. (1978). The organization of turns at formal talk in the classroom. *Language in Society*, 7, 183–213.
- Mehan, H. (1979). *Learning lessons: Social organization in the classroom*. Cambridge, MA: Harvard University Press.
- Mehan, H. (1985). The structure of classroom discourse. In T. van Dijk (Ed.), *Handbook of discourse analysis 3: Discourse and dialogue* (pp. 119–131). New York: Academic.
- Nishizaka, A. (2000). Seeing what one sees: Perception, emotion and activity. *Mind, Culture, and Activity*, 7, 105–123.
- Nishizaka, A. (2003). Imagination in action. *Theory & Psychology*, 13, 177–208.
- Pea, R. D. (1993). Practices of distributed intelligence and designs for education. In G. Salomon (Ed.), *Distributed cognitions* (pp. 47–87). New York: Cambridge University Press.
- Pomerantz, A. (1984). Agreeing and disagreeing with assessments: Some features of preferred/dispreferred turn shapes. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action* (pp. 57–101). Cambridge, England: Cambridge University Press.
- Ryle, G. (1963). *The concept of mind*. Middlesex, England: Penguin Books. (Original work published 1949)
- Sacks, H. (1987). On the preferences for agreement and contiguity in sequences in conversation. In G. Button & J. Lee (Eds.), *Talk and social organisation* (pp. 54–69). Clevedon, England: Multilingual Matters.
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50, 696–735.
- Schefflen, A. E. (1973). *Communicational structure: Analysis of a psychotherapy transaction*. Bloomington: Indiana University Press.
- Schegloff, E. A. (1968). Sequencing in conversational openings. *American Anthropologist*, 70, 1075–1095.
- Schegloff, E. A. (1995). *Sequence organization*. Unpublished manuscript.
- Schegloff, E. A. (1998). Body torque. *Social Research*, 65, 335–596.
- Schegloff, E. A. (2002). Opening sequences. In J. E. Katz & M. A. Aakhus (Eds.), *Perpetual contact: Mobile communication, private talk, public performance* (pp. 326–385). Cambridge, England: Cambridge University Press.
- Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. Cambridge, England: Cambridge University Press.
- Wittgenstein, L. (1953). *Philosophical investigations*. Oxford, England: Basil Blackwell.
- Wootton, A. J. (1997). *Interaction and the development of mind*. Cambridge, England: Cambridge University Press.