Aspect-Seeing in the Interactional Organization of Activities

Aug Nishizaka
Chiba University, Japan
augnish@chiba-u.jp

Vision in activity

Vision is one among the several important topics that Charles Goodwin has been studying. He radically de-psychologized vision by re-specifying it as what is publicly achieved in the unfolding distinct activity rather than what is privately lodged in the mind or the brain (Goodwin 1994, 1996; Goodwin, Goodwin 1996). Two general notes on vision are in order here.

First, as Ryle (1949) observed, the verb "see" is an achievement (rather than a process) verb; seeing is not something that occurs somewhere at or during a particular time. Therefore, the task of interaction studies of vision does not lie in locating or detecting any (revealed or concealed) process of "seeing" somewhere, whether in the interaction being examined or inside the participants in the interaction. What participants see can be a resource mutually and publicly available to them so that they can jointly accomplish the current activity that they engage in. That is, the analytic task is to elucidate how participants can ascribe seeing to each other in the joint accomplishment of the activity. Second, Wittgenstein (1953) distinguished between two types of seeing.

> The one: "What do you see there?" — "I see this" (and then a description, a drawing, a copy). The other: "I see a likeness between these two faces" — let the man I tell this to be seeing the faces as clearly as I do myself" (1953: 193).

One can show others what one sees by copying or pointing to it and describing its features. However, one cannot show the recipient the likeness between two (photographed) faces by copying or pointing to them and describing their features. In this manner, the recipient may still not able to see the likeliness, although he or she is able to see these two faces. The recipient may suddenly become able to see the likeliness, although there have not been any changes in the faces. Wittgenstein called this second type of seeing "aspect-seeing" (1953: 193). |

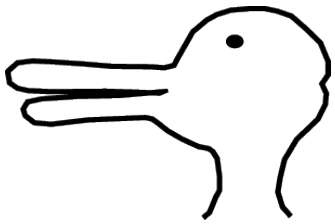


Figure 1 Duck-Rabbit

Wittgenstein (1953) characterized (a certain type of) aspect-seeing as "seeing-as"; aspect-seeing can be seeing something *as* something (such as seeing the two faces as alike). However, aspect-seeing is not the outcome of an additional operation (such as interpretation or judgment) on the first type of seeing (i.e., seeing *simpliciter*). In fact, when we see the well-known "duck-rabbit" picture (Figure 1), we do not first see the shape of the line and the position of the point and then interpret them *as* a duck or rabbit; rather, we cannot see the picture without seeing it under the aspect of a duck or rabbit. One sees an aspect rather directly. The analytic task, here once again, lies in elucidating the procedural ground for the ascribability of seeing, rather than inferring the process of the interpretation by which the participants achieve the aspect-seeing.

"Professional vision" (Goodwin 1994) is a kind of aspect-seeing. Students or laypersons become able to see the same archeological features or videotaped body movements under a different aspect; for example, they now see the same movements as aggressive, rather than victimized. To cite another example, one can show others an airplane with a specific number on it by pointing to it and describing its features, but one cannot show, by pointing to airplanes before one's eyes, the sameness and difference between the flights that different airplanes embody. Nevertheless, workers at an airport *see* the difference at an aircraft-parking apron (Goodwin, Goodwin 1996). They see the airplanes as flights with particular destinations.

Charles Goodwin clearly demonstrated that such aspect-seeing is related to the distinct activity in progress, which is historically constituted as well as locally organized. The sameness and difference between the flights is relevant to all who work at an airport; they could not work without seeing it. In this short essay, I provide some thoughts on the ascribability of seeing that is organized in manners relevant to the temporally unfolding activity.

Seeing-as

I analyze two interactional fragments in which at least retrospectively participants can be said to have *seen* something *as* something relevant to the activity in progress. The first example, videotaped early in 1990s, is the one that I analyzed | years ago

(Nishizaka 2000). An instructor (TEA) provides a lesson about how to use a Japanese word processor on a DOS/V (IBM compatible) machine. “Full-size” means the normal size for Japanese (and Chinese) characters and “half-size” for Roman letters. Using a typical Japanese word processor of that time, to obtain a half-sized Roman letters, one had to first convert full-sized Japanese letters into full-sized Roman letters (using function key 9) and then the latter into half-sized Roman letters (using function key 8).

(1) [Word Processor Lesson][1]

```
01 TEA:   jaa  ima oo- korette sa:: (.) <zenkaku no> ai bii emu
          then now      this   PART      full-sized  I   B   M
          And, now la(rge[?])- This is (.) a full-sized "IBM,"

02        ↓desu yo    ne::.
           is   PART  PART
          right?

03 STU:   hai
          Yes

04 TEA:   hankaku no ai bii e↓mu jaa  irete mite. (1.0) Ooyoo
          half-sized I   B   M   then input try      application
          Now, try to input a half-sized "IBM." (1.0)
          An application.

05        ( 7 . 2 )

06 TEA:   nn (.) °are°?
          Yeah    O?

07 STU:   a r e ?
          O?

08        ( 1 . 8 )

09 STU:   (°  ┌    ne°)
10 TEA:       └a aa aa::↓:. (0.2) soo des' ne
                Oh::              That's the way.

11 TEA:   >dakara< (0.6) ikinari (0.8) hachi ban    o
           so             first        eight number PART
12        oshi  tara ...
          press if
          So, if you press number 8 first, ...
```

| In lines 01-02, TEA points out what currently appears on the screen (i.e., a full-sized “IBM”), and in line 04, in contrast to this, provides a new task (i.e., inputting a half-sized “IBM”). Immediately after this task is given, the student (STU), who was looking at the screen, begins to look at the keyboard and attempt to complete the task (see Excerpt 1a below). After an unsuccessful trial is visible (lines 06-07), STU marks

the completion of the task (line 09), by possibly inviting confirmation (with the final particle *ne*), while raising her upper body (see Excerpt 1b below). Then, in line 10, TEA produces an emotional display of admiration almost simultaneously and proceeds to provide confirmation (i.e., a positive evaluation). In this small segment of the interaction, seeing the full-sized or half-sized "IBM" on the screen is ascribable in the manners relevant to the unfolding activity. However, the ascribability is accomplished in the very finely coordinated interactional detail. Excerpt 1a is a detailed transcription of lines 01-03 of Excerpt 1.

(1a) [Detail: 01-03]

```
                        Fig. 1-1   Fig. 1-2
                                ↓        ↓
01 TEA:    jaa  ima oo|- korette |sa:: (.) <zenkaku no> ai bii emu
           then now       this    PART      full-sized  I   B   M
           And, now la(rge[?])- This is (.) a full-sized "IBM,"
  tea.g:   MMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMM*
  stu.g:                       mmmmMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMM
   tea:                 |points at the      * "M" for "monitor screen"
                         screen -->|

02         ↓desu yo    ne::.
            is   PART  PART
           right?
  tea.g:   MMMMmxXXXXXXXX->>
  stu.g:   MMMMMMMMMMMMM->>

03 STU:    hai
           Yes
```

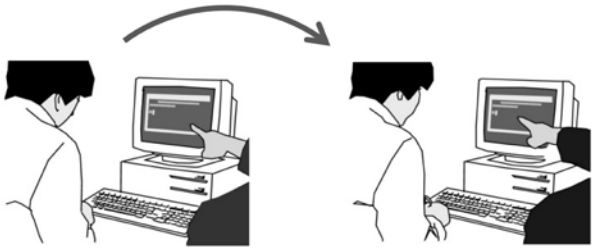


Fig. 1-1
STU looks downward when TEA begins to points to the monitor.

Fig. 1-2
STU looks up at the monitor.

TEA's request for confirmation in lines 01-02 and STU's production of a confirmation token in line 03 constitute a preliminary sequence for the sequence of task-giving and task-completion from line 04 onward. TEA's confirmation request in lines 01-02 is constructed such that STU is assured of the seeing of the full-sized "IBM" on the screen; TEA first cuts off the word *oo-* 'lar(ge)' and points | at the screen with a proximal deictic term (*kore* 'this'), thereby inviting STU to turn to the screen (Goodwin

1981). TEA then mentions the name of the object (full-sized “IBM”), which is now within STU’s visual field. In this context, STU giving confirmation in line 03, together with her gaze at the screen, provides TEA with a procedural ground for ascribing the seeing of the full-sized “IBM” on the screen to STU, and in fact, TEA proceeds to the base action of the task-giving in line 04, relying on what STU sees there.

Similarly, in line 09, STU’s verbal and nonverbal conduct claims the completion of the task and provides TEA with a procedural ground for ascribing the seeing of the half-sized “IBM,” and in fact, TEA proceeds to the positive evaluation of what appears on the screen. However, both participants see something more than the half-sized “IBM”; they see the “IBM” appearing on the screen *as* (or under the aspect of) STU’s successful completion of the task.

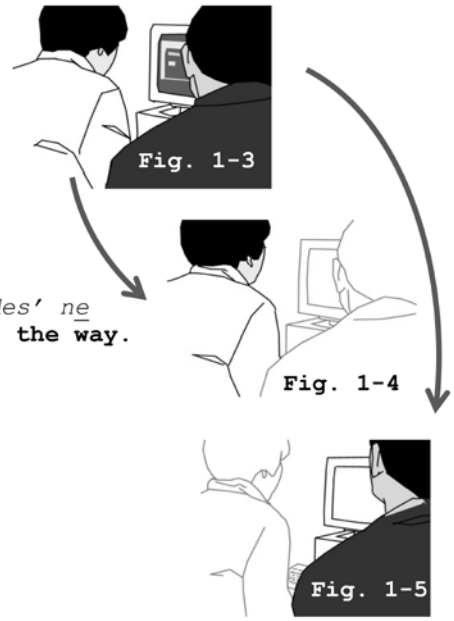
This aspect-seeing is also organized both in the order of action-sequencing and in the details of the local sequential order. First, TEA’s task giving in line 04 functions as an “instruction for seeing” (Goodwin 1996) as well. It specifies what is expected to appear on the screen ‘next’ and sets up a normative framework for how to see what appears there ‘next.’ The participants see what appears there ‘next’ within this framework. If something other than what is expected to appear does appear ‘next’, they see this something as the failure of the task completion, as they did in lines 06-07 (see Nishizaka 2000, for the detailed analysis). When after STU’s second try (line 08), what is expected appears ‘next’, they see this as success.

Second, TEA produces an emotional display of admiration only after STU marks the completion of the task by sharply raising her upper body while looking at the screen and nodding twice. Excerpt 1b is a detailed transcription of lines 08-10 of Excerpt 1.

(1b) [Detail: 08-10]

```
                                    Fig. 1-3
                                       ↓
08          (0 . 8) | (0 . 4) | (.)
  stu.g:        MMMMMMMMMMMMMMMMMMM->>
                    |presses  |presses
                     a key     a key

      Fig. 1-4   Fig. 1-5
          ↓          ↓
09 STU:   |(°  ┌  |  ne°)|
10 TEA:       └a |aa aa::↓:. (0.2) soo des' ne
                  Oh::              That's the way.
   stu:   |raises upper |
           body and     |
           nods twice ->|
   tea:          |raises upper body
```



| In response to STU's body movement, TEA also raises his upper body, produces the emotional display, and proceeds to the positive evaluation (line 10). Most likely, TEA has seen the half-sized "IBM" appear at the end of line 08, but he does not display admiration immediately after he sees it. What he expects, and is expected, to see is not merely the half-sized "IBM", but rather STU's successful completion of the task. Therefore, he only responds to what is visible on the screen after STU marks the completion of her task. Thus, the participants' (STU's and TEA's) seeing of the half-sized "IBM" as STU's successful task-completion is organized in the fine detail of the local order of interaction. 350

Integrated (multimodal) perception

Some perception is an integrated outcome of multiple sensory inputs (Nishizaka 2007). Husserl (1973 [1907]) described such perception in the following way.

> "[W]hen we see [a sheet of paper] and put hand [on it] at once, we have a mixed perception from two sides, where the seen part is not tactilely felt and the tactilely felt part is not seen. However, we have a mixed fullness such that only one kind of fullness belongs to those properly appearing segments" (1973 [1907]: 73).

The way in which segments of different modalities are integrated into a mixed perception is also organized in a way relevant to the unfolding activity. For example, what one perceives when pouring beer into a glass is an integrated perception of touch,

vision, and audition, also involving proprioception and kinesthesis. It results from holding the bottle, seeing the beer pour from the bottle's mouth into the glass, hearing the sound of the liquid pouring, feeling the weight of the bottle, and feeling the position and movement of one's hand and arm. Whatever the sensory sources of the perception, what is perceived, that is, the object of the perception, is a unity into which the sensory segments are integrated in a manner relevant to the ongoing activity, such as pouring beer.

As Merleau-Ponty (2012 [1945]: 238) observed, one can even "see" the rigidity of the glass. That is, one can visually perceive the tactile features of an object without touching it (i.e., without any tactile inputs). We can extend this observation and show that one can "see" proprioceptive features without moving or positioning one's body in *that* way. This section focuses on this special type of mixed perception, also a type of "aspect-seeing."

In instructional settings, an instructor often uses his or her body to demonstrate the correct manner of performing a body movement. Then, the student sees the instructor's movement under the aspect of how the student should move his or her own body, that is, under the aspect of kinesthesis or proprioception, although the student does not actually move or position his or her body. I cite one simple example that I analyzed elsewhere (Nishizaka 2017). In this example, | a professional violinist (TEA) 351
teaches a four-year-old child (CHI) how to use the bow to play quarter notes, represented by the sound of a dog barking *wan wan* 'bow wow'. TEA demonstrates how CHI should move her arm.

(2) [Violin Lesson]

```
01   TEA:     |koo<
              |like this.
     tea.g:   ((looking forward))
     chi.g:   XXXXX
     tea:     |touches the inside
               of her elbow

02   TEA:     (0 . 4)|(0 . 6)
     tea.g:   XXXXXXXXXXXXXXX ((her face is oriented forward))
     chi.g:   xxxtataTATATATA*    *"TA" for "TEA's right arm"
     tea:             |moves her
                       arm->>
```

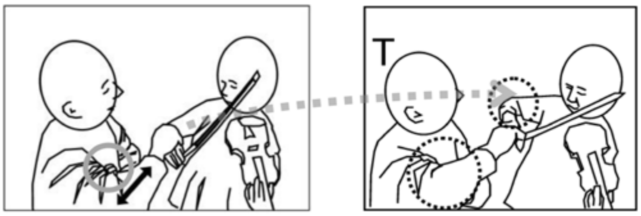


Fig. 2-1 Fig. 2-2

```
              ↓                ↓
03   TEA:     wan    wa|n    |wan
              Bow    wo|w    bow
     tea.g:   XXXXXXXXXXXXXXXxx
     chi.g:   TATATATATAtabbBB*   *"B" for "bow"
     tea:     ----------->|
     chi:             |raises her elbow
                       and places the bow
                       on the violin
```

In line 01, TEA touches the inside of her own right elbow with her left hand (see Figure 2-1, marked in a circle), while using a proximal demonstrative term (*koo* 'like this'), thereby inviting CHI's to look at the shape of the elbow. When TEA successfully directs CHI's gaze toward TEA's right elbow (indicated by "ta" and "TA" below line 02), TEA moves her own entire arm (lines 02-03) while producing sounds that represent quarter notes. CHI promptly raises her elbow and places the bow on the violin to play the notes, with her gaze focused on TEA's elbow. This conduct by CHI exhibits her following understanding: 1) the previous position of her elbow was incorrect, 2) the movement of TEA's arm is a demonstration of the correct movement of the entire arm (in contrast with the movement that CHI made actually with her elbow at a lower position and by using only her forearm [not her entire arm]), and 3) TEA's movement of her own arm is more than the movement of the arm; it demonstrates how to use the bow to play the notes. CHI sees TEA's arm movement under the aspect of how to move the arm correctly to play particular notes with a particular string, that is, under the aspect of the

kinesthesis of her arm movement. CHI does not move her arm when | she sees TEA moving her arm nor does she look at her own arm and compare it with the position of TEA's elbow when she raises her elbow. This aspect-seeing is embedded within the activity of rectifying the incorrect arm movement and demonstrating the correct movement.

The demonstration does not emerge in the interactional vacuum; it is performed in its juxtaposition with TEA's reproduction of CHI's incorrect performance (Excerpt 3). Immediately prior to Excerpt 2 (line 01 of Excerpt 2 is reproduced as line 03 of Excerpt 3), TEA moves CHI's arm in CHI's peripheral visual field with a proximal demonstrative term *kooyuu* 'like this' (see Nishizaka 2017, for a more detailed analysis).

(3) [Violin Lesson; immediately preceding (2)]

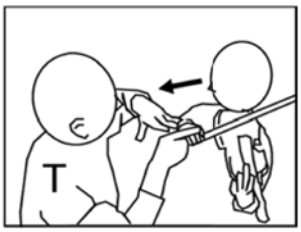


Fig. 3-1

```
                ↓
01  TEA:    |kooyuufuu ni  ugoite↓ta  °kara°|
            |this way PART moving_PAST because
            |Because ((it)) was moving like this.
    tea.g:  XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
    chi.g:  XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
    tea:    |pushes and pulls the bow
             with her right hand, while keeping
             her arm over CHI's arm ------->|

02          |( 0 . 8 )
    tea.g:  XXXxxx
    chi.g:  XXXXXXXXXX
    tea:    |slightly touches CHI's right elbow

03  TEA:    |koo<
            |like this.
```

It may be interesting what the demonstrative term (*kooyuu* 'this way') is understandable as referring to; is it the movement of TEA's hand, of the bow, or of CHI's arm? Observe that TEA raises her left arm over CHI's arm (Figure 3-1); the position of TEA's arm is perceivable as the normative position of the bow-arm, against which the actual position of CHI's arm should be assessed. In fact, TEA uses the past-tense construction to indicate that the movement that she is making is reproducing CHI's movement. Given

the bodily configuration in which TEA's utterance in line 01 (Excerpt 3) is embedded, together with the construction of TEA's entire utterance, the demonstrative term hearably refers to the movement of CHI's arm (rather than TEA's hand or the bow), and one sees TEA's movement | as a reproduction of what CHI was actually doing (this is another typical instance of "aspect-seeing"). 353

In the temporal and spatial juxtaposition with the recognizable reproduction of the incorrect movement, TEA's arm movement in lines 02-03 of Excerpt 2 is seeable as demonstrating the correct arm movement; it is temporally contrasted with TEA's immediately preceding reproduction of CHI's incorrect movement, and its spatial position is temporally connected to the position of TEA's left arm raised over CHI's right arm. In this bodily configuration, seeing how to move the right arm to play quarter notes is accomplished. Seeing how to move is an interesting phenomenon in that it is seeing under the aspect of kinesthesis without moving one's own body in the same way as the movement that one sees.

Concluding Remarks

Seeing under the kinesthesis is a phenomenon that has rarely been highlighted in interaction studies. However, it is important for organizing various activities in interaction. In this short essay, I have attempted to demonstrate two things. First, some aspects of Charles Goodwin's de-psychologizing of vision can be captured by reference to Wittgenstein's notion of "aspect-seeing." Second, Goodwin's focus on activity helps understand the organization of various kinds of vision that Wittgenstein thought about under the rubric of "aspect-seeing." Aspect-seeing is still a cover term for diverse kinds of perception that can change without changes to the sensory inputs. However, aspects are directly seen, and are not the result of interpretation, judgment, inference, or the like. Examining such aspects as relevant to the ongoing activity can lead to a systematic investigation of aspect-seeing.

The integrated "mixed" perceptions mentioned previously have not often been investigated in interaction studies, either. Their integrated nature can also be related to various organizations of activities, although I did not go into any detail of its interactional organization. Goodwin has opened a very rich area for interactional studies of perception, or for studies of interaction whose organization involves a variety of perception.

Note

[1] In the excerpts, each line is composed of two or three tiers. In the first, there is a [347]

Romanized version of the original Japanese. In the second tier, there are phrase-by-phrase glosses, where necessary. Finally, in the third tier, a rough English translation is provided. The first tier of the transcript utilizes Jefferson's (2004) transcription system. In the second tier, the following abbreviations are used: PART for "Particle," and PAST for "Past." Some excerpts include annotations of the embodied conduct of each participant below the English translation (i.e., in the extra tiers designated as "tea," "stu," and "chi"). The starting and ending points of the movements are indicated by the sign "|". Double arrows in these tiers indicate a continuation of the described conduct over the line. Some lines have annotations of the participants' gaze directions in the extra tiers designated as "tea.g:," "stu.g:," and "chi.g," which indicate the teacher's, the student's, and the child's gaze directions, respectively. "X" indicates that the gaze is directed toward the co-participant's face, and abbreviations used for where the gazes are directed are glossed in the margin of each excerpt. Lower-case letters in these tiers indicate the transition of gaze directions.

References

[353]

Goodwin, Charles 1981. *Conversational Organization: Interaction Between Speakers and Hearers*. New York: Academic Press.

Goodwin, Charles 1994. Professional vision. *American Anthropologist* 96(3): 606-33.

Goodwin, Charles 1996. Transparent vision. In: Ochs, Elinor; Schegloff, Emanuel A. Schegloff; Thompson, Sandra (eds.), *Interaction and Grammar.* Cambridge: Cambridge University Press, 370-404.

Goodwin, Charles; Goodwin, Marjorie-Harness, 1996. Formulating planes: Seeing as a situated activity. In: Middleton, David; Engeström, Yrjö (eds.), *Cognition and Communication at Work*: Cambridge University Press, 61-95. |

Husserl, Edmund 1973 [1907]. Ding und Raum. Husserliana Band XVI. The Hague: Nijhoff, 1-293. 354

Jefferson, Gail 2004. Glossary of transcript symbols with an introduction. In: Lerner, Gene H. (ed.), *Conversation analysis*. Amsterdam: John Benjamins, 13–23.

Merleau-Ponty, Maurice 2012 [1945]. *Phenomenology of Perception*. (Landes, Donald A., trans.) New York: Routledge.

Nishizaka, Aug 2000. Seeing what one sees. *Mind, Culture, and Activity 7*(1-2): 105-123.

Nishizaka, Aug 2017. The perceived body and embodied vision in interaction. *Mind, Culture, and Activity* 24(2): 110-128.

Ryle, Gilbert 1949. *The Concept of Mind*. London: Hutchinson.

Wittgenstein, Ludwig 1953. *Philosophical Investigations*. Oxford, UK: Blackwell.