Preliminary Analysis of Embodied Interactions between Science Communicators and Visitors Based on a Multimodal Corpus of Japanese Conversations in a Science Museum

Rui Sakaida[1], Ryosaku Makino[2], Mayumi Bono[1,3]

[1] National Institute of Informatics, [2] Faculty of Human Sciences, Waseda University,
[3] Graduate University for Advanced Studies (SOKENDAI)
[1] 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo, 101-8430, Japan
[2] 2-579-15 Mikajima, Tokorozawa city, Saitama, 359-1192, Japan
lui@nii.ac.jp, rmakino@aoni.waseda.jp, bono@nii.ac.jp

Abstract

This paper introduces preliminary analyses of embodied interactions, drawing on a multimodal corpus of Japanese conversations, which we video-recorded during scientific communications at a museum in Tokyo, the Miraikan. A comparison of similar cases extracted from our multimodal corpus shows both similarities and differences, not only in language use but also in bodily conduct in certain interactional sequences. We focus on a number of sequences, such as those where science communicators invite visitors to walk to the next exhibit, and our detailed analyses show that the practices of science communicators are context-free and context-sensitive interactional procedures, adapted and adjusted to the different situations communicators may encounter. After presenting our analyses, based on a corpus from a naturally occurring but partly controlled setting, we suggest that we can investigate both the generality and the situatedness of interactional practices. In the future, using such multimodal corpora, we will be able to both qualitatively and quantitatively analyze language use and non-verbal behaviors in situated activities.

Keywords: Multimodal corpus, Embodied interaction, Scientific communication

1. Introduction

This paper introduces a preliminary analysis of embodied interactions based on a multimodal corpus of Japanese conversations that we video-recorded in a specific social setting, i.e., scientific communication in a museum. We filmed face-to-face conversations between science communicators (SCs) and visitors at the Miraikan science museum in Tokyo (Figure 1). The data of this corpus, which we call the Miraikan Science Communication Corpus, were recorded in a very specific situation, unlike many classical corpora, which contain strictly controlled data. In this paper, we demonstrate that we cannot analyze both the generality and situatedness of our language use and non-verbal behaviors in conversations without drawing on such a unique corpus.

In the following sections, we first discuss why we should construct multimodal corpora recorded in complicated situations. Then, considering the need for such multimodal corpora, we explain the characteristics of our corpus. Next, we describe how we recorded the data and made annotations. Finally, we present our preliminary analysis of the corpus—the method was qualitative and comparative—suggesting the usefulness of the corpus.

Figure 1: Scientific communication in the Miraikan.

2. Background

As language resources, many corpora of Japanese conversations have been constructed and published in the field of corpus linguistics. For example, Den and Enomoto (2007) constructed a corpus of video-recorded three-party conversations. In each session, three participants were asked to sit around a table and talk about a topic selected at random, and started/ended by talking with the researcher. In this way, we can design a highly controlled dataset of both spoken language and non-verbal behaviors, such as eye-gaze and gestures.

Social scientists have long enthusiastically recorded and analyzed naturally occurring conversations, which would have occurred with or without being recorded by researchers (e.g., Schegloff, 1968; Sacks et al., 1974). It is not until we observe naturally occurring conversations that we can investigate the procedures by which conversations properly start or end (Schegloff and Sacks, 1973). Often, during conversations, the participants stand facing each other (Kendon, 1990) or even walk to another place (Mondada, 2012).

There is a methodological division between corpus linguistics and the social sciences. For researchers in corpus linguistics, highly controlled language resources are useful for analyzing the general characteristics of language use or non-verbal communication quantitatively or statistically. For social scientists, experimentally recorded corpora are not satisfactory for analysis of the variety of situational organizations of our language use or non-verbal behaviors. Therefore, many social scientists have recorded data suitable for their own research questions, instead of drawing on highly controlled corpora. However, conversations recorded by social scientists are often too complicated and improvisational to analyze the language use or non-verbal behaviors statistically or comparatively. In general, social scientists prefer to depend on and stick to the data that they have recorded themselves, rather than share their data with other researchers or research communities. Thus, each social scientist must record,

annotate or transcribe, and analyze his or her data individually, which takes a great deal of time and effort. Therefore, to connect the motivation of social scientists to analyze naturally occurring conversation with corpus linguistics, corpora of naturally occurring, but somewhat controlled, conversations must be constructed and published.

3. Characteristics of the Corpus

Considering the need for corpora of natural, but controlled, conversations, we made a corpus with the following characteristics.

First, the corpus was recorded in a semi-institutional setting, i.e., scientific communication in a museum. The social sciences have focused on and analyzed social interactions in highly institutional settings for many years (Drew and Heritage, 1992); for example, in medical settings (Heath, 1986), classrooms (Mehan, 1979), or courtrooms (Atkinson and Drew, 1979). Turn taking, sequence organization, and bodily orientation occur in different orders that are typical of each setting. By contrast, scientific communication in the Miraikan is not as institutionalized as the communication in other social settings, especially for the visitors (Bono et al., 2014). In general, people visiting museums or galleries not only interact with their companions but also react to strangers, negotiating which exhibits to look at or when to move to the next exhibit (vom Lehn et al., 2001; vom Lehn, 2013). In contrast to the usual modes of visiting museums, the characteristics of the interactions found in our corpus are unique. The visitors, not knowing what kind of communication they are supposed to conduct in the Miraikan, try to understand the sort of activity they are involved in, the membership category (Sacks, 1972) that is relevant for themselves, and how they should respond to the SCs.

In addition, conversations in the corpus are embedded in a complicated environment, i.e., an exhibition room in the Miraikan. Therefore, they involve a complicated embodied activity, i.e., appreciation and explanation of the exhibits. Our embodied actions in conversations are usually coupled with environmental elements (Streek et al., 2011). In particular, many of our gestures are coupled environmentally (Goodwin, 2007). For example, the meaning of a pointing gesture is made clear because it is tied to the objects being pointed out (Goodwin, 2003). Furthermore, conversations are not simply conducted while sitting around a table, which is the case in most conversation corpora. Rather, we frequently talk while standing and facing each other, operating or manipulating objects, conducting joint activities, or even walking together. Recently, spatial configuration (Kendon, 1990), objects (Nevile et al., 2014), multiactivity (Haddington et al., 2014), and mobility (Haddington et al., 2013) in conversations have been hot topics in interaction studies. In our corpus, by virtue of the very complicated environment of the exhibition room, the participants have to pay attention to various exhibits or move between exhibits. The SCs often point to the exhibit relevant to the current topic of conversation and encourage the visitors to pay attention to it. Or, after explaining an exhibit, the SCs may invite the visitors to walk to the next exhibit together.

4. Recording and Annotation

4.1 Recording

To construct the corpus, we asked the National Museum of Emerging Science and Innovation (the Miraikan) to allow us to video-record routine conversations between the science communicators (SCs) and visitors. However, to control the setting to a certain extent, the space used for the recordings was separated from the rest of the exhibition room. The themes of the exhibition area where we made the recordings were the "Spread of Space" and "Challenge the Universe with a Giant Telescope".

The staff of the Miraikan selected 15 expert SCs to participate in our recordings. The aim of our project and how we would manage the data was explained to all of the SCs and visitors who participated in the recordings, and they granted us permission not only to use the data for our own work but also to publish the videos, transcripts, and annotation data as a multimodal corpus.

The video-recordings were made over 10 days in February and March 2013, for about 1 hour per day. On each day, we asked two SCs to talk separately with three groups of visitors as naturally as possible. Although we did not tell the SCs which exhibit to explain or the route to the next exhibit, they usually followed the same route and explained the same series of exhibits. Most of the SCs started the talk with an explanation of the model of the solar system, and invited the visitors to the next exhibit, a model of the Subaru Telescope. Each group consisted of from 1 to 5 people. The average group had 2.26 visitors.
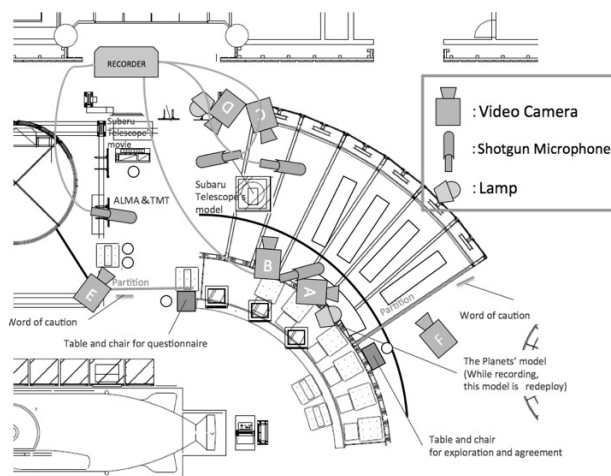

Figure 2: Layout of the cameras and microphones.

Figure 3: Screenshot of the merged data.

	18.000	00:05:18.200	00:05:18.400	00:05:18.600	00:05:18.800	00:05:19.000	00:05:19.200	00:05:19.400	00:05:19.600	00:05:19.800	00:05:20.000	00:05:20.2
scA [391]				mo ikko			tsukatteru no ga:, aaiu::					
v01 [102]												
v02 [105]												
scA.annotation face 1	turns right											
scA.annotation face 2	makes the visitors pay attention to the right											
scA.annotation hand 1							puts her right hand toward Subaru		holds her right hand toward Subaru			
scA.annotation hand 2							makes the visitors pay attention to Subaru		continues to make the visitors pay attention to Subaru			
scA.annotation body 1	turns right											
scA.annotation body 2	makes the visitors pay attention to the right											
scA.annotation feet 1	steps backward						steps backward					
scA.annotation feet 2	moves to avoid blocking the view of the visitors						moves to avoid blocking the view of the visitors					
scA.annotation meta	prepares to start explanation						starts to explain the next exhibit to the visitors					

Figure 4: Example of relevant annotation (Bono and Sunaga, 2016)

After the recording, participants were asked to give basic personal information, such as their age, writing it on their face sheets. The participants were from 5 to 66 years old, and the average age was 31.26 ($SD = 15.78$)[1].

The recording instruments included six video cameras[2] and seven microphones (Figure 2). Five video cameras were fixed around the separated area, while a professional camera operator recorded the front view of the participants with a mobile video camera. Similarly, four shotgun microphones[3] were fixed in position and several pin microphones[4] were attached to the participants' chests. After the recording, all the recorded sound was mixed, with the noise removed.

As a result, 15 expert SCs and 79 visitors participated in the recordings, and 35 sessions were recorded in total. The average length of the sessions was about 14 minutes. All of the video and audio data were merged into one file for each session (Figure 3).

4.2 Annotation

To facilitate use of the corpus, we annotated the data in several ways, using ELAN[5]. First, all of the utterances and vocal behaviors such as laughing or coughing were segmented into inter-pausal units (Koiso et al., 1998) and transcribed in Japanese[6]. In addition, the body movements of the SCs that were relevant to the current interactional practice were annotated. Inspired by Ethnomethodological Conversation Analysis, we described the action that each physical movement achieved. We call the annotation method "Relevant Annotation" (Bono and Sunaga, 2016). In this way, the face, hand, body, and feet movements of each SC were annotated and the action that all of the movements achieved was described in meta-level tiers (Figure 4). To date, Relevant Annotation has been used to target selected segments from 32 data files, and the total length of the annotated data is about 36 minutes.

5. Analysis: Synchronous Walking during Conversation

Thanks to the characteristics of the corpus, we can compare similar sequences of naturally occurring, but somewhat controlled, embodied interactions between SCs and visitors. At present, our analyses are not statistical or quantitative, but qualitative. Nevertheless, our analyses suggest that a
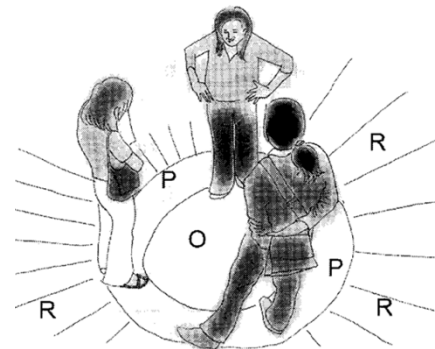
Figure 5: F-formation[7]

multimodal corpus of a natural, but semi-controlled, situation can be useful for comparative studies of both verbal and nonverbal behaviors. For example, it is common to focus on turn design (Drew, 2012), and by analyzing it, we can demonstrate the way in which interlocutors use a specific format of utterance. Joh et al. (2015) demonstrated that science communicators draw on a certain format of utterance, such as, "Do you know X?" to introduce new topics of conversation, simultaneously drawing the visitors' attention to the next exhibit to be shown. Also, we focus on participants' body movements and illustrate the way in which a particular body movement functions in conversations. Makino et al. (2015) found that a unique standing position, which we call the "H-formation", is employed to display readiness to begin explaining an exhibit. In this paper, we analyze how SCs and visitors together move from exhibit to exhibit.

In many studies of multimodal interactions, walking has been investigated as a typical means of establishing an encounter before beginning a conversation (Kendon, 1990; Mondada, 2009) and then breaking it up again after the conversation is completed (Heath, 1986; Broth and Mondada, 2013). Of course, we may walk *during* a conversation or interaction. As a matter of fact, in interaction studies, mobility has been the subject of the most pioneering research (McIlvenny et al., 2009, 2014; Haddington et al., 2013). Walking together is such an ordinary activity in our social life that we rarely regard it as complex, or as a particular accomplishment. However, many empirical studies have suggested that the organization of walking together is a complexly situated

---

issue for participants in naturally occurring social settings, such as in guided tours (De Stefani and Mondada, 2014), a supermarket (De Stefani, 2013), dance classes (Broth and Keevallik, 2014), or situations mediated by specific objects, such as revolving doors (Weilenmann et al., 2014).

Following Kendon (1990), people establish an F-formation, facing one another and engaging in overlapping transactional segments, when they stand in public places and interact (Figure 5). By contrast, when people move to a new place while maintaining the interaction, the F-formation dissolves, because it is difficult to walk forward while facing another person (De Stefani and Mondada, 2014). The dissolution of the original F-formation can lead to an absence of conversation or end the current conversation. For instance, it can be difficult for a hearer to display his/her availability or recepiency (Heath, 1986) to the speaker while walking; this can appear as a weak orientation toward the conversation. To elucidate how people can continue to talk even when walking, we observed the SCs and visitors twice as they moved between the two exhibits in the exhibition room, while maintaining their communication.

In the following analysis, we used the annotations explained above and presented additional annotations where necessary. For instance, excerpt 1 (Figure 6), which is analyzed below, was extracted from the example of the relevant annotation shown above (Figure 4). The annotations were edited and revised by the authors. In the transcripts, using the transcription system introduced by Mondada (2009, 2012), vocal utterances are given in black letters and body movements are in gray letters.

5.1 Excerpt 1

In excerpt 1 (Figure 6), an SC successfully invites the two visitors to the next exhibit, which is a model of the Subaru Telescope, in a multimodally organized and sequentially relevant way. In the time immediately preceding the excerpt, the SC had begun to explain the three ways in which researchers investigate the universe. In line 01, the SC says *konpyuutaa wo tsukatte shumireeshon wo suru no to:* (doing the simulation using a computer, and), which refers to the second way of studying the universe (the first way had already been referred to before the start of the excerpt.) After noting the second way, the SC utters *to:* (and), (line 01) and suspends her utterance. At the same time, the SC turns right and steps backward. Immediately thereafter, she looks at the Subaru, saying *mo ikko* (one more thing) (line 02). As the SC is saying this, the visitors respond to the multimodal invitation by the SC, turning to the right. At the end of her utterance, the SC points to the Subaru with her right hand, saying *tsukatteru no ga:, aaiu::* (they are using, that sort of) (line 02). *aaiu* is a kind of distal demonstrative adjective in Japanese, which can spatial-deictically refer to something at a distance from both the speaker and the hearer (cf. Hayashi, 2004) and project the referent itself will be uttered immediately.

```
01  SC   konpyuutaa wo tsukatte shumireeshon wo suru no to+:, (0.2)+(0.15)
          computer    P   use     simulation   P   do  thing and
          "Doing the simulation using a computer, and,"
     sc                                              +turns right and steps backward
     sc                                              +looks at Subaru
02  SC   mo† i*kko +(0.15) tsukatteru no ga:, aa*i†u:: (0.05)+(0.45) booenkyo+o.
          more one          using    thing P   that                  telescope
          "one more thing,"        "they are using, that sort of,"   "telescope."
     sc           +points to Subaru              +stops pointing and looks at Vs
     sc                                                       +starts to walk->
     v1        *turns right and steps backward  *steps forward a little
     v2     †turns right                         †starts to walk->
03       (0.25)
04  V1   (u*:+::n).
              huh
          "Huh."
     sc       +looks forward while nodding
     v1     *starts to walk->
05  SC   booenkyo+o tte mita koto ari masu ka:?
          telescope    P   saw  have_ever JD_PL P
          "Have you ever seen a telescope?"
     sc           +turns to Vs while walking->
06  V1   (0.35) <a[ri masu:>
                   have JD_PL
                "Yes, I have."
07  V2           [booenkyoo (0.3) [un, a, un un (0.7) un.
                  telescope        yeah oh yeah yeah     yeah
                 "A telescope...yeah, oh, yeah…yeah."
08  SC                            [oho
                                   "Oh."
09  V1   ufufu[fufufu.]
10  SC        [sa + su]ga: (0.05)+(0.3)+(1.2)†(0.25) mi*raikan ni irassyaru dake ari masu ne:.
              great                                  Miraikan to   come_PL  may_well JD_PL P
              "Great!"                               "You may well come to Miraikan."
     sc        +turns to Subaru briefly
     sc                  +turns to Vs
     sc                       +stops walking
     v1                               *stops walking
     v2                        †stops walking
```
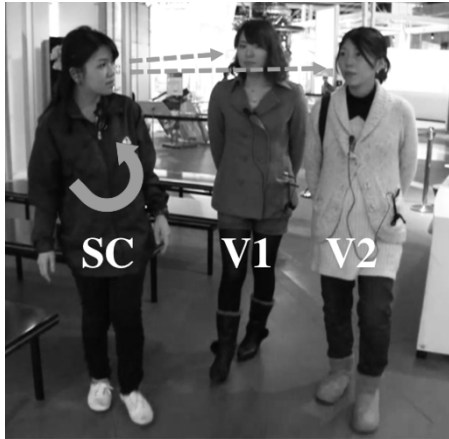
Figure 6: Transcript of excerpt 1.

Figure 7: Spatial formation while walking in excerpt 1.

When the SC is saying *aaiu::*, both of the visitors react to the SC in different ways. While V1 steps forward a little, V2 begins to walk. After saying *aaiu::*, during a pause, the SC stops pointing at the Subaru and looked at the visitors. Orienting herself to the visitors, the SC says *booenkyoo* (telescope) (line 02) and, as soon as she completes her turn, she begins to walk. After a silence in line 03, noticing that both of the interactants had begun to walk, V1 also begins to walk. In this way, the SC succeeds in inviting the visitors to pay attention and start walking to the next exhibit, with the topic of their conversation naturally transitioning from the general explanation of astronomy to a specific exhibit. Even after beginning to move to the next exhibit, the SC continus to orient herself to maintaining conversation with the visitors. Just after starting to walk, V1 responds to the SC, saying *u:::n* (huh) (line 04), to which the SC also responds with a nod, while looking forward. According to Goodwin (1981), if they look away from the speaker, hearers use vocal tokens or nods to avoid the loss of their display of recepiency. Here the SC orients to maintaining the conversation, even without gazing at her interactants. In line 05, the SC asks a question of the visitors, *booenkyoo tte mita koto ari masu ka:?* (Have you ever seen a telescope?). During this question, the SC turns to the visitors while continuing to walk forward, adopting a particular posture called body torque (Schegloff, 1998) (Figure 7). This posture embodies the SC's orientation to two actions: walking and talking. Maintaining the specific posture, the SC, gazing at the visitors, continues to talk with them until they reach the next exhibit (lines 06–10). Distributing her body orientation to several actions (Nishizaka, 2013), the SC skillfully ensures the continuance of their interactional space (Mondada, 2009) even while moving to the next exhibit.

5.2 Excerpt 2

In excerpt 2 (Figure 8), the SC's practice in the first half is, to some extent, similar to that of the SC in excerpt 1, although the SCs in the two excerpts are different. As in excerpt 1, the SC in this excerpt first points to the next exhibit, Subaru (line 04) and then starts to walk, inviting the visitors to move to it (line 05). However, we can observe a rather different practice in the second half.

```
01  SC  gunma tenmondai ga tabun okkii desu ne.
        Gunma observatory P probably big  JD_PL P
        "Gunma observatory is probably big."
02      (0.3)
03  V2  a, (0.3) a, sooiu [koto,
        oh      oh  so    thing
        "Oh, oh, it means that."
04  SC                   [itten  go meetaa (.) no: (0.45) kagami, +chou(do) dakara: (0.2)
                          one point five meter       P              mirror  exactly    so
                          "One point five meters',"                "mirror, so exactly,"
    sc                                                                           +looks at and points to Subaru
05  SC  are desu yo. (0.1)+(0.9)       +(0.15) kono†(0.25)*(0.15)
        that JD_PL P                            this
        "that is it."                          "This,"
    sc                   +walks aside +looks at the Vs
    v1                                         *starts to walk->
    v2                                        †starts to walk->
06  V1  (unn?)
        "Huh?"
07  SC  subaru to on+naji youna: katachi wo (0.75) shite masu.
        Subaru P similar  like   shape    P         be JD_PL
        "Subaru's shape is similar ((to that of Gunma observatory's telescope))."
    sc             +turns to Subaru and walks toward it->
08      (1.15)
09  SC  tsutsu jana+ku te,  moo kanzenni, (.) oopun desu kedo. (0.15)+(0.15)†(0.05)
        cylinder   not  P   exactly perfectly   open JD_PL though
        "It is no longer a kind of cylinders, and perfectly open though."
    sc             +turns to Vs and points to Subaru             +stops walking
    v2                                                                  †stops walking
10  SC  kore, subaru no (0.15)*(0.6) mokei nan'desu kedo.
        this  Subaru P               model  JD_PL   though
        "This is Subaru's model."
    v1                       *stops walking
```

Figure 8: Transcript of excerpt 2.

Figure 9: Spatial formation before walking in excerpt 2.

Figure 10: Spatial formation while walking in excerpt 2.

In the first part of the excerpt, the SC and the visitors are speaking of the way in which they can watch stars in the sky. In line 01, the SC, in saying *Gunma tenmondai ga tabun okkii desu ne* (Gunma observatory is probably big), answers the question, which had been asked by one of the visitors just before this excerpt, of where to go to watch the stars. V2, who had asked the question, responds to her response, saying *a, a, sooiu koto* (Oh, oh, it means that).

In line 04, the SC continues to explain in detail about the Gunma observatory. The SC refers to *itten go meetaa no kagami,* (a 1.5-meter mirror,) (line 04), with which the telescope of Gunma observatory is equipped. Immediately after saying this, the SC begins to look at and point to the model of the Subaru telescope, stating *chou(do) dakara are desu yo* (so exactly, that is it) (lines 04–05). *are* (that) is a Japanese distal demonstrative pronoun, and by using it speaker can implicitly request the interlocutor(s) to specify what it refers to, typically by connecting the speaker's body movement with the surrounding environment or any object around them. In excerpt 1, the SC used a similar distal demonstrative *aaiu* (that sort of), which is generally used as an adjective to strongly project the immediate occurrence of the noun it modifies, as mentioned in 5.1. On the other hand, *are* is a demonstrative pronoun, so it can be produced without mentioning the object it refers to when spatial-deictically used. Here, the SC points to the Subaru model with his right index finger, inviting the recipients to attend to the direction of the Subaru and recognize what the SC is pointing to (Goodwin, 2003). Just after speaking, the SC steps aside, toward the direction of Subaru, looking at the visitors, and then says *kono* (this) (line 05). Responding to these multimodal behaviors of the SC, the visitors begin to walk (line 05). *kono* (this) is a proximal demonstrative adjective, which implies that they should look at the

referent, i.e., the model of Subaru, more closely than from the current position. The SC's word selection and modification, that is, at first saying *are* (that) and then *kono* (this), successfully invite the visitors to walk toward the Subaru to look more closely.

In the latter half of this excerpt, the SC takes the visitors to the next exhibit in a rather different way to that in excerpt 2. Unlike in excerpt 1, the SC in excerpt 2 walks forward, without gazing at the visitors. The F-formation, which had been in effect before walking (Figure 9), is completely dissolved during the walk (Figure 10). Instead, the SC employs another way to continue their conversation. In line 07, the SC turns to the Subaru and starts to walk to it, producing a continuation of the utterance he had already begun before walking, *subaru to onnaji youna: katachi wo shite masu* (the Subaru's shape is similar). In line 05, the utterance of the SC, *kono*, projected that his turn-constructional unit (Sacks et al., 1974) was not yet completed and would lead to an explanation of "this". In fact, after a pause (line 05), the SC begins to continue the explanation of "this", with the relationship between the current topic of conversation and the next exhibit, i.e., Gunma observatory and Subaru observatory, becoming clear. In this sequence, the SC succeeds in taking the visitors to the next exhibit, in accordance with the visitors' previously expressed interest, that is, in watching stars in the sky. This practice appears to be an alternative to the practice in excerpt 1, i.e., looking at the visitors while continuing to talk.

6. Discussion

Our analysis in section 5 identified the following interactional practices. Specifically, the SCs and visitors were able to proceed in concert to the next exhibit while maintaining their conversation by engaging in the following behaviors in a step-wise fashion. First, the SC referred and pointed to the next exhibit, a model of the Subaru Telescope, thereby inviting the visitors to look at it. Then, the visitors looked at the exhibit in question[8]. Third, without saying anything about walking, the SC started to walk toward the next exhibit, inviting the visitors to follow. Fourth, the visitors followed the SC, demonstrating their implicit agreement with this plan of action. Fifth, the SC continued to talk to the visitors until they all reached the next exhibit.

Two aspects of these behavioral patterns seem to enable the SCs to maintain their conversation. First, before starting to walk, the SCs pointed to the next exhibit; this established the template for the new F-formation to be created at the next exhibit after the current F-formation had dissolved. Second, by not explicitly referencing the move to the next exhibit and by continuing to talk to the visitors while moving, the SCs implied that walking *per se* was not central to their activity; rather, walking was an interactional forum in which their main activity occurred. Consequently, even if the F-formation was dissolved, the SCs could maintain their scientific communication.

Nevertheless, there were several observable differences in the interactional practices between the two excerpts. First, the ways each of the SCs invited the visitors to attend to Subaru were different. In excerpt 1, the SC stepped

[8] In excerpt 2, from the beginning, the visitors were facing the Subaru, so it was not clear exactly when they were closely

attending to it, but they certainly did so, responding to the SC's invitation.

backward before pointing to Subaru so that the visitors could clearly look at it. On the other, in excerpt 2, the SC pointed to Subaru without stepping toward it, as he was not obstructing the visitors' view in that position. As a result, the SC's multimodal conduct in excerpt 1 more strongly projected the following sequence, i.e., starting to walk and moving to the next exhibit. One of the visitors (V2) in excerpt 1, actually began walking earlier than the others in both excerpts, even before the SC herself had begun walking. Secondly, the ways the SCs drew the visitors' attention to the conversation while walking were unique to each SC. In excerpt 1, the SC gazed at the visitors in a body-torqued position (Schegloff, 1998) while walking, maintaining an orientation toward their conversation. In excerpt 2, through the use of two different Japanese demonstratives, *are* and *kono*, and suspending his turn, the SC projected the continuation of his utterance concerning the exhibit they were walking to, even before walking. In such similar but different ways, the two SCs successfully invited their visitors to walk together and move to the next exhibit. In short, the practices of science communicators consist of context-free and context-sensitive interactional procedures (Sacks et al., 1974), which are adapted and adjusted to various situations they come across. In this way, comparing similar cases extracted from the multimodal corpus, we can see both similarities and differences not only in language use but also in bodily conduct in a certain interactional sequence.

7. Concluding Remarks

By demonstrating our analyses based on the multimodal corpus of a naturally occurring, but partly controlled setting, we investigated both the generality and the situatedness of interactional practices, exhibiting the possible uses of such corpora. In our daily lives, continuing a conversation may be difficult because of various situational factors. Nevertheless, we often manage to continue conversations in some way, in many instances, when extraordinary but quite reasonable multimodal practices are observed. Such behaviors are embedded in complicated environmental and/or social relationships, as in scientific communication in the Miraikan. In the future, by using similar multimodal corpora, we will be able to both qualitatively and quantitatively analyze language use and non-verbal behaviors in situated activities.

8. References

Atkinson, J. M. and Drew, P. (1979). Order in Court: The Organisation of Verbal Interaction in Judicial Settings. NJ: Humanities Press.

Bono, M. (2008). Units of analysis for understanding conversational structure: F-formation. *Journal of Japanese Society for Artificial Intelligence*, 23 (4): 545–551.

Bono, M., Ogata, H., Takanashi, K. and Joh, A. (2014). The practice of showing 'who I am': A multimodal analysis of encounters between science communicator and visitors at science museum. In Stephanidis, C. & Antona, M. (Eds.), *Universal Access in Human-Computer Interaction: Universal Access to Information and Knowledge, Lecture Notes in Computer Science*, 8514, 650–661, Berlin and Heidelberg: Springer.

Bono, M. and Sunaga, M. (2016). A proposal of annotation scheme for body movement based on participants' understandings. *JSAI SIGs Conference Papers*, SIG-SLUD-B503: 7–12.

Broth, M. and Mondada, L. (2013). Walking away: The embodied achievement of activity closings in mobile interaction. *Journal of Pragmatics*, 47: 41-58

Den, Y. and Enomoto, M. (2007). A scientific approach to conversational informatics: Description, analysis, and modeling of human conversation. In Nishida, T. (Ed.), *Conversational Informatics: An Engineering Approach*, 307–330. Hoboken, NJ: John Wiley & Sons.

De Stefani, E. (2013). The collaborative organisation of next actions in a semiotically rich environment: Shopping as a couple. In Haddington, P., Mondada, L. and Nevile, M. (Eds.), *Interaction and Mobility: Language and the Body in Motion*, 123–151, Berlin: De Gruyter.

De Stefani, E. and Mondada, L. (2014). Reorganizing mobile formations: When 'guided' participants initiate reorientations in guided tours. *Space and Culture*, 17 (2): 157–175.

Drew, P. and Heritage, J. (Eds.) (1992). Talk at Work: Interaction in Institutional Settings. Cambridge: Cambridge University Press.

Drew, P. (2012). Turn design. In Sidnell, J. & Stivers, T. (Eds.), *The Handbook of Conversation Analysis*, 131–149. NJ: Wiley-Blackwell.

Goodwin, C. (1981). Conversational Organization: Interaction between Speakers and Hearers. New York: Academic Press.

Goodwin, C. (2003). Pointing as situated practice. In Kita, S. (Ed.), *Pointing: Where Language, Culture, and Cognition Meet*, 217–242. NJ: Lawrence Erlbaum.

Goodwin, C. (2007). Environmentally coupled gestures. S. D. Duncan, J. Cassell, & E. T. Levy (Eds.), *Gesture and the Dynamic Dimension of Language: Essays in honor of David McNeill*, 195–212. Amsterdam: John Benjamins.

Haddington, P., Mondada, L. and Nevile, M. (Eds.) (2013). Interaction and Mobility: Language and the Body in Motion. Berlin: De Gruyter.

Haddington, P., Keisanen, T., Mondada, L. and Nevile, M. (Eds.) (2014). Multiactivity in Social Interaction: Beyond Multitasking. Amsterdam: John Benjamins.

Hayashi, M. (2004). Projection and grammar: Notes on the 'action-projecting' use of the distal demonstrative are in Japanese. *Journal of Pragmatics*, 36: 1337–1374.

Heath, C. (1986). Body Movement and Speech in Medical Interaction. Cambridge: Cambridge University Press.

Joh, A., Bono, M. and Takanashi, K. (2015). Constructing a dialogue at science museum: A practice of 'Do you know X?' from the perspective of interaction analysis. *Cognitive Studies*, 22 (1): 69–83.

Kendon, A. (1990). Conducting Interaction: Patterns of Behavior in Focused Encounters. Cambridge: Cambridge University Press.

Koiso, H., Horiuchi, Y., Tutiya, S., Ichikawa, A. and Den, Y. (1998). An analysis of turn-taking and backchannels based on prosodic and syntactic features in Japanese Map Task dialogues. *Language and Speech*, 41: 295–321.

Makino, R., Furuyama, N. and Bono, M. (2015). Spatial-orientational behavior for the narrative in a field: A case study on science communicators' standing position employed to display the readiness to start giving an explanation of the exhibit. *Cognitive Studies*, 22 (1): 53–68.

McIlvenny, P., Broth, M. and Haddington, P. (2009). Communicating place, space and mobility. *Journal of Pragmatics*, 41: 1879–1886.

McIlvenny, P., Broth, M. and Haddington, P. (2014). Moving together: Mobile formations in interaction. *Space and Culture*, 17 (2): 104–106.

Mehan, H. (1979). Learning Lessons: Social Organization in the Classroom. Cambridge: Harvard University Press.

Mondada, L. (2009). Emergent focused interactions in public places: A systematic analysis of the multimodal achievement of a common interactional space. *Journal of Pragmatics*, 41: 1977–1997.

Mondada, L. (2012). Garden lessons: Embodied action and joint attention in extended sequences. In Nasu, H. & Waksler, F. C. (Eds.), *Interaction and Everyday Life: Phenomenological and Ethnomethodological Essays in Honor of George Psathas*, 293–311. MA: Lexington Books.

Nevile, M., Haddington, P., Heinemann, T. and Rauniomaa, M. (Eds.) (2014). Interacting with Objects: Language, Materiality, and Social Activity. Amsterdam: John Benjamins.

Nishizaka, A. (2013). Distribution of visual orientations in prenatal ultrasound examinations: When the healthcare provider looks at the pregnant woman's face. *Journal of Pragmatics*, 51: 68–86.

Sacks, H. (1972). An initial investigation of the usability of conversational data for doing sociology. In Sudnow, D. (Ed.), *Studies in Social Interaction*, 31–73, New York: Free Press.

Sacks, H., Schegloff, E. A. and Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50 (4): 696–735.

Schegloff, E. A. (1968). Sequencing in conversational openings. *American Anthropologist*, 70 (6): 1075–1095.

Schegloff, E. A. and Sacks, H. (1973). Opening up closings. *Semiotica*, 8: 289–327.

Schegloff, E. A. (1998). Body torque. *Social Research*. 65 (5): 536–596.

Streeck, J., Goodwin, C. and LeBaron, C. (Eds.) (2011). Embodied Interaction: Language and Body in the Material World. Cambridge: Cambridge University Press.

vom Lehn, D., Heath, C. and Hindmarsh, J. (2001). Exhibiting interaction: Conduct and collaboration in museums and galleries. *Symbolic Interaction*, 24 (2): 189–216.

vom Lehn, D. (2013). Withdrawing from exhibits: The interactional organisation of museum visits. In Haddington, P., Mondada, L. and Nevile, M. (Eds.), *Interaction and Mobility: Language and the Body in Motion*, 65–90, Berlin: De Gruyter.

Weilenmann, A., Normark, D. and Laurier, E. (2014). Managing walking together: The challenge of revolving doors. *Space and Culture*, 17(2): 122–136.