THE SENSE OF SELF AND OTHER IN NONVERBAL COMMUNICATION: 
A STUDY USING VIDEO ANALYSIS OF DIALOG SCENES

Toshiki MATSUOKA

Graduate School of Education, Kyoto University, Japan

This research attempted to empirically investigate how individuals experience “self” and “others” according to the theory of sense of self from nonverbal aspects. In the survey, participants (N = 15) read a short story and discussed it with the researcher. All dialogs were video recorded and nonverbal behaviors of participants/the researcher in the obtained video data were coded based on analysis index. Quantified nonverbal behaviors were analyzed using moving correlation method, and time series changes of nonverbal behavioral synchronization between participants and the researcher were examined. Results revealed that in those with a high sense of self, active mutual exchange was maintained between self and others in the non-linguistic domain. Conversely, in those with a low sense of self, mutuality in the nonverbal area was weak. This result suggests the importance of nonverbal areas in considering the problems in sense of self of adults.

Key words: sense of self, experience of others, nonverbal interaction, synchronization, video-based analysis

INTRODUCTION

People subjectively organize and process their own psychological experiences and create a construct of “self” each time, and the subject of such organization is called “the sense of self.” Stern (1985) accumulated the results of infant research and was the first to try systematic theoreticalization of sense of self. Stern (1985) said, “While no one can agree on exactly what the self is, as adults we still have a very real sense of self that permeates daily social experience. It arises in many forms. There is the sense of a self that is a single, distinct, integrated body; there is the agent of actions, the experiencer of feelings, the maker of intentions, the architect of plans, the transporter of experience into language, the communicator and sharer of personal knowledge” (p. 5). Such sense of self is interwoven so that multiple regions (the sense of an emergent self, the sense of a core self, the sense of a subjective self and the sense of a verbal self) are accumulated since the earliest time, and each region continues to operate for a lifetime while interacting with others.

A noteworthy point of Stern’s theory is that the self is not established as a firm concept but develops gradually since early childhood. This made it possible to approach the problem of the self retroactively to the stage before language acquisition. These

Correspondence concerning this article should be addressed to Toshiki Matsuoka, Graduate School of Education, Kyoto University, Kyoto, Japan (e-mail: levaryssubed@gmail.com).
methodologies made it possible to consider psychotherapy, which had been often discussed as linguistic exchanges, in terms of nonverbal exchanges.

Examining the problem of self from the viewpoint of sense of self seems to be clinically meaningful. However, few studies have empirically investigated this mechanism. Matsuoka (2015) attempted to grasp aspects of sense of self in university students by creating the Sense of Self Scale (SoSS) and formally analyzing Thematic Apperception Test data. Results revealed that even in ordinary university students, there is a large individual difference in sense of self from the viewpoint of degree of integration. Some people have a highly conscious sense of self (high sense of self), whereas others showed weakness in cohesion of and ambiguous sense of self (low sense of self), and it was suggested that they formed one spectrum. In this spectrum, those with a high sense of self are considered to feel on a daily basis, their own persistence as a consistent body unit and live in a world of abundant subjectivity that is rich in emotional interaction. In addition, they can organize various emotional stimuli as their own experiences, and share narratives of experiences with others. On the other hand, those with a low sense of self have weak self-organization as a single body unit, and their persistence and their constructions of subjective worldviews are weak. It is also assumed that it is difficult for them to integrate into their sense of self complicated or ambiguous stimuli as their own experiences.

In Matsuoka (2015), the main focus was on how the individual experienced the “self” under the theory of sense of self. However, in order to explore the problem of developing a sense of self, it is not enough to consider one person externally, based on a self-reflection of verbal contents. It is said that creating a sense of self and sense of others is said to be a fundamentally parallel process (Stern, 1985), and it is speculated that one can only approach the essence of these concepts through interpersonal relationships. Furthermore, sense of self is originally created before language acquisition (Stern, 1985), and discussing it without nonverbal aspects will be problematic.

In recent years, clinical psychological practice is suggesting focus on nonverbal communication is an important therapeutic factor. The Boston Change Process Study Group (BCPSG) organized by Stern and others suggested that the change in psychotherapy is divided into two areas: “domain of conscious and explanatory language interpretation over explicit knowledge” and “domain of unconscious interactive and procedural knowledge over implicit knowledge,” and both are considered therapeutic. Of these, explicit knowledge refers to “interpretation” in so-called psychoanalytic psychotherapy and implicit knowledge is like “procedural knowledge related to relationships” including body sensations and nonverbal communication.

BCPSG (2010) argues about the significance of dealing with the treatment effect of “something implicit” and conducts considerable research by considering clinical cases. In case studies in Japan, incorporating the theory of Stern, Ono (1995) delineates the process of development of the sense of self through nonverbal emotional exchanges in psychotherapy, and Kaji (2013) uses physical nonverbal exchanges to approach the client. In addition, Nagaoka et al. (2011) demonstrated a correspondence between psychological change and nonverbal behavior in counseling using image analysis
method.

Based on the above, this research aimed to empirically investigate how individuals experience “self” and “others” according to the theory of sense of self from nonverbal aspects. The actual two-part scene including the change in the dialog structure was observed and the relationship between the two parties was examined in terms of the nonverbal communication. Synchronism of nonverbal behaviors is introduced as an indicator of the relationship between researcher and participant. Relationship in this study refers to those that are psychologically felt with an interactor and rather than formed since a long time, like a friendship; it is more likely to be influenced by a slight change in an interactor’s utterance, posture, or behavior and can be a factor that defines the next behavior of interactors (Duck & Sants, 1983). Komori and Nagaoka (2010) stated that synchronism of nonverbal behavior among interactors will be an effective indicator of acceptability and empathy in relationships. In infant observations in which Stern found the sense of self, micro changes (e.g., intensity, timing, or rhythm of behavior) that change every moment in the relationship was important. Therefore, focusing on the synchronicity of nonverbal behavior of researcher and participant was considered to be suitable for this research.

**Method**

**Participants**

A questionnaire survey was conducted using the SoSS (Matsuoka, 2015) with university students in the Kansai region, and 234 responses were received. The respondents with Top/Lower 25% SoSS scores were assigned to Group H/L, respectively, with 7 members in Group H (4 males, average age 22.86 years, SoSS average 90.33 points: numbering by H1–H7 for each participant) and 8 members in Group L (4 men, average age 20.75 years, SoSS average 64.13 points: numbering by L1–L8 for each participant). The research was approved by Kyoto University Clinical Psychology Research Ethics Review Board.

**Procedure**

The survey was conducted from August to October 2016. In the research room, the researcher and participants were seated, with a desk in between them, at about 1 m. The survey was made of two sections (section A and B), and lasted for about 45 minutes.

In section A (for about 30 minutes), the participants read the story aloud and discussed it with the researcher. As an experimental stimulus, the story “Young woman and sailor” was used. This story is an exercise that has been used for many years in the structured group encounter. In the story, five characters (a young lady, sailor, old man, the young lady’s fiancé, and fiancé’s best friend) are stuck on the uninhabited island and try to escape. After reading the story, participants of the exercise are asked to rank the characters in order of their favorability. The situation of section A was completely video recorded with the consent of the participants. Two video cameras were prepared for recording participants and researchers, which were arranged so that the movement of the body and gaze of the researcher and participant could be observed. According to Matsuoka (2015), the function of the sense of self is different depending on the degree to which the situation was structured. Based on this result, section A consisted of the following three phases with different levels of structuring of the dialog scene: structured, semi-structured, and unstructured. In the structured phase, participants provided definite answers. In the semi-structured phase, participants gave their opinions based on a theme and framework set by the researcher. In the unstructured phase, participants continued to talk freely.

In the structured phase (for about 5 minutes), the researcher handed a clipboard with the story to
Table 1. Nonverbal behavior analysis indices

<table>
<thead>
<tr>
<th>Behavior items</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaze</td>
<td>Regardless of the direction of the face, behavior that is judged to have seen the other’s face.</td>
</tr>
<tr>
<td>Gaze divergence</td>
<td>Something that diverts the line of sight. Regardless of the direction of the body, being in a posture that oriented the direction where the person cannot see the other’s face.</td>
</tr>
<tr>
<td>Laugh</td>
<td>It is judged to show the so-called smile face whether the person laughed.</td>
</tr>
<tr>
<td>Nod</td>
<td>Nodding when the other person is talking (It has no meaning of YES).</td>
</tr>
<tr>
<td>Utterance</td>
<td>Words such as “ah,” “um,” laughter, sighs and or so are excluded.</td>
</tr>
<tr>
<td>Vocalization</td>
<td>Speech other than “utterance,” such as “ah,” “um,” laughter, sighs, etc.</td>
</tr>
<tr>
<td>Silence</td>
<td>Not speaking for approximately 1 second or more.</td>
</tr>
<tr>
<td>Symbolic gesture</td>
<td>It is an action that acts as a direct substitute for the language, such as expressing victory with a V sign, yelling by swinging his head vertically, etc.</td>
</tr>
<tr>
<td>Illustrative gesture</td>
<td>Body movement directly linked to conversation that refines and clarifies the content of conversation. For example, an action to draw a falling curve with a fingertip to show how the object falls, and a gesture that shows the distance or degree by spreading hands.</td>
</tr>
<tr>
<td>Meaningless behavior</td>
<td>An action that is not directly related to conversation, meaninglessly moving a fingertip, a hand, or a body (The beginning and end of a series of actions that continue for a while).</td>
</tr>
</tbody>
</table>

participants and instructed, “Please read the story silently first.” When silent reading was over, participants were told to read it aloud for the researcher to understand. After reading aloud, participants were asked about the number of characters, their roles, and the grouping of the characters to see if they understood the content.

In the semi-structured phase (for about 15 minutes), the researcher instructed, “Please rank the five popular characters from the 1st to 5th place in terms of their favorability. Also, think about the reasons why you ranked them. Please complete the ranking and reasoning and let me know your ideas after 8 minutes,” and timed them with a stopwatch. During the eight minutes, the researcher remained silent while facing the participants. Eight minutes later, the participants indicated their ranking and the researcher inquired further about the reason for the ranking of all five characters.

In the unstructured phase (for about 10 minutes), at the end of the semi-structured phase, the researcher said, “This time I will not speak, so please tell me freely what you thought or felt from reading the story. You can stay silent, but please continue until I tell you to stop.” Then, participants were timed with a stopwatch. After that, the researcher heard the participants’ remarks giving responses except for “utterance” in the analysis index (described later). After 8 minutes, when participant finished talking, the investigator announced the end of section A.

In section B (for about 15 minutes), participants were interviewed about the experiences in section A. In order to control for the influence of the differences stemming from different interviewers, one researcher engaged with each participant throughout the interaction.
Data coding

We synthesized videos from two cameras for researcher and participants so that they were displayed in parallel on the same screen after time adjustment. Of these videos, 3 minutes from the structured phase (at the end), 4 minutes from the semi-structured phase (1 minute 30 seconds at the end of silence and 2 minutes 30 seconds after starting to answer), and 4 minutes from the unstructured phase (at the end of dialog) were analyzed.

Nonverbal behavior in the dialog video was measured with 1 s time sampling method based on analysis index. With reference to the nonverbal behavior analysis indices of Toda and Takamura (2003) and Isomura (2010), ten kinds of analysis indices such as “gaze,” “gaze divergence,” “laugh,” “nod,” “utterance,” “vocalization,” “silence,” “symbolic gesture,” “illustrative gesture,” and “meaningless behavior” were used (Table 1). Stern (1985) describes affect attunement, which is a cross-modal nonverbal communication, as intercommunication as a function of the sense of self. Therefore, it was thought that coding nonverbal behaviors belonging to multiple modalities at the same time is consistent with this research purpose.

RESULTS AND DISCUSSION

Analysis

In order to examine the temporal changes in the relationship between nonverbal behaviors occurrence of participants and researchers in each case, coding of the results of nonverbal behaviors at each stage were analyzed using the moving correlation method. The formula of Maeda, Nagaoka, and Komori (2007) was modified for analysis. In this study, the sum of nonverbal behaviors of researchers and participants was calculated, analysis was performed with a shift width of 1 second and a window length of 60 (1 minute). The moving correlation coefficient $R_o$ calculated under this condition was an index of synchronism of nonverbal behaviors of participants and researchers. In addition, in order to examine the influence of nonverbal behavior of researchers or participants on the subsequent nonverbal behaviors of participants or researchers, the relationship between the nonverbal behavior of the researcher and that of the participant after 1 second and the relationship between the nonverbal behavior of the participant and that of the researcher 1 second later were also examined with the moving correlation method. Correlation coefficients obtained by this analysis were taken as $R_i$ and $R_\Delta$, respectively. The formula for the three moving correlation coefficients is as follows, of which RES and PRT refer to researchers and participants, respectively.

$$R_o(t) = \frac{\sum_{i=0}^{n}(RES(t+i) - \mu_{RES})(PRT(t+\Delta t+i) - \mu_{PRT})}{\sqrt{\sum_{i=0}^{n}(RES(t+i) - \mu_{RES})^2} \sqrt{\sum_{i=0}^{n}(PRT(t+\Delta t+i) - \mu_{PRT})^2}}$$

$\Delta t = \{-1, 0, 1\}$

$R_o$: Degree of synchronization between the participant and the researcher at the same time

$R_i$: Degree of entrainment between the participant and the researcher after 1 second

$R_\Delta$: Degree of entrainment between the researcher and the participant after 1 second

In the following sections, the features obtained from the moving correlation analysis of each group will be examined for each phase.
Structured phase

Regarding $R_0$ in moving correlation analysis, no noticeable trend in each group was found for the value of the overall correlation coefficient. Looking at the time series change in the correlation, in Group H, the intersection of $R_0$ and $R_{-1}$ was seen in most cases (Fig. 1-a, b), whereas in Group L, such crossing was not seen much (Fig. 1-c, d). In Group H, the correlation coefficient $R_0$ increased at the intersection point of $R_1$ and $R_{-1}$.

As $R_{-1}$ is synchronization with the researcher following participant’s nonverbal behaviors and $R_1$ is synchronization with the participant following researcher’s nonverbal behavior, the intersection of these two indicators can be interpreted as an alternation in the relationship to be followed/follow-up between the two. In other words, this could be interpreted as an alternation in the initiatives among them. This intersecting phenomenon found in Group H shows that “delivery of nonverbal initiative” was a feature in this group and also suggests that certain mutual exchanges were conducted between the participant and the researcher in nonverbal areas. It is very interesting that this

![Graphs](Image)

**Fig. 1.** The result of moving correlation analysis of structured phase.
phenomenon was observed at the beginning of the survey and in the structured phase where the content of the dialog was almost fixed, suggesting that Group H might be open to nonverbal mutual exchange as a basic stance. The sense of self is thus formed in the domain of nonverbal mutual exchange at the earliest stage, and activities in this domain may be sufficiently performed in Group H with high sense of self even in adulthood. In addition, the intersection of $R_t$ and $R_{-1}$ can be interpreted as being the moment the entrained (leading) person is going to synchronize with the other (respond to the initiative), and patterns in which $R_0$ increased at this intersection point were observed in many cases in Group H. This phenomenon seems to be the result of one “approaching” the other at the nonverbal level.

On the other hand, the absence of such phenomena in Group L indicates that the relationship where one leads or follows the other consistently continues. This result suggests that in the structured phase, non-linguistic interactions were not yet performed in Group L, and the roles in interactions were still fixed.

In the interview, both groups described that this structured phase was the simplest task, such as “Because the answer was short, it was easy (H2)” and “It was easy to do because there is an answer (L3).” This phase was considered to be straightforward for both groups in that it could be managed without a subjective judgment process. Thus, it is interesting that even in this structured phase that was perceived as easy and clear by both groups, the experience at nonverbal level was very different. This may indicate that the differences in the function of sense of self in the dialog scenes of both groups were rooted in aspects that cannot be identified at conscious and linguistic levels.

**Semi-structured phase**

Regarding $R_o$, the rapid increase in the correlation coefficient (around 0.4 or more) from no correlation section, including the transition from silence to dialog scene was characteristic in Group L (Fig. 2-c, d). On the other hand, in Group H, the increase in $R_o$ at the transition from silence to conversation was relatively small (Fig. 2-b), or there were many cases where there was almost no change (Fig. 2-a).

The pattern of rapid increase in $R_o$ due to the transition shows that the researcher and participant pairs who were low in synchronism in the silence situation began to synchronize during transition to the dialog situation. The state of high entrainment can be interpreted as that where nonverbal behaviors are co-occurring among the researcher and participant, that is, the situation where the two parties are “together” at a nonverbal level. On the other hand, the state of low synchronization is a state where the nonverbal behaviors of both the parties do not co-occur (one does not respond to the other, or both act randomly), and it can be interpreted as “being alone” situation where one is separated from the opponent at a nonverbal level.

As the pattern of rapid increase in $R_o$ was observed remarkably in Group L, it seems that they had a tendency of being “alone” during the silent situation and the dialog situation is experienced as “to be together”, so the two situations were to be experienced as discontinuous for them. On the other hand, in Group H, it seems that the degree of discontinuation was milder than that in Group L, and silence and dialog were
continuously experienced at the nonverbal level. In other words, in Group H, even though there were silent scenes without verbal exchanges, it is thought that the participants experienced “being-with” the researcher.

Stern (1985) said that the ability “to be with others” is not an inherent one, but acquired in the process of developing a sense of self. Such “procedural knowledge on relationships” will work in parallel and interact with language knowledge in adults (Beebe et al., 2005). In Group H with a high sense of self, as this nonverbal procedural knowledge domain acted strongly in the understructure of interaction, regardless of the silence and dialog in the situation, participants would have been experiencing “being with others (the researcher)” continuously. On the other hand, it seems that in Group L, the experiences of “being with others” in the non-linguistic domain were sparse.

In the interview, the difficulty of the task of ranking was a feature reported only in Group L, and the commonly stated troubles were due to the lack of clear criteria for judgment, such as “I was wondering what my evaluation criterion was (L6).” Here, it

**Fig. 2.** The result of moving correlation analysis of semi-structured phase.
appears that the difficulty was because of having to make subjective decisions in the absence of objective criteria as seen in Matsuoka (2015). Regarding the situation, for example, a participant in Group L said that “The time limit was 8 minutes, but the answer was decided in a short time and I had too much time. I did not feel uncomfortable, but I already answered in my head, so I rested for a while (L2).” This remark also shows that the time spent in silence was experienced as a situation of “being alone,” separated from the researcher.

Unstructured phase

Focusing on the amplitude of $R_n$ in Group H, there were many cases where the amplitude was large (Fig. 3-a) or the relatively high correlation persisted (Fig. 3-b), whereas in Group L there were many cases where the amplitude was small and the wave was flat (Fig. 3-c) or the correlation increased or decreased monotonically (Fig. 3-d).

The larger the amplitude, the larger was the increase and decrease in the correlation coefficient. Here, the increase in the coefficient meant an increase in the synchronism and the decrease meant decline in synchronism or increase in asynchronism, so that larger amplitudes could be interpreted as repeat of more synchronizing and asynchronizing. In the results, it was confirmed that Group H had a pattern of larger amplitudes than Group L, that Group H repeated more dynamic synchronous and asynchronous patterns. On the other hand, Group L, seemed to be maintaining a relatively stable state compared to Group H.

As for $R_n$ in Group H, patterns of repeatedly increase and decrease in small increments (continuous micro peaks and valleys) or rapid increase after large decrease in value (macro mountains and valleys) were observed. The cause of such a change pattern was assumed to be a process of synchronism $\rightarrow$ asynchronism (divergence of two) $\rightarrow$ re-synchronism. This process could be understood as two people who were closely matched in a nonverbal communication to come to match again after having experienced a difference or gap.

When the synchronism between the two disappears and the self and the other diverge at nonverbal level, the person may be departing from where he being-with the other until then, and finally returning to one person separated from the other and experiencing being aware of self as others of others. In Utsumi (2012), the self emerges only at the break of consciousness, and he points out that the others make a break in continuity of consciousness. In the moment when the steady and sustained continuity of consciousness breaks down due to the experience of difference caused by “others,” the feeling of “self” will be vividly experienced in that person. The patterns seen in Group H suggest that such experiences have occurred in non-linguistic areas.

Furthermore, it is noteworthy that in the micro/macro pattern, not only had the synchronization decreased and the researcher and participant diverged, but there was also a movement to resynchronize. Here, it seems that the organization of sense of self is working, adjusting the experiences of differences in non-linguistic areas and restoring the relationship. Stern (1985) says that the experiences of differences in interaction are averaged between representations of interaction that have been generalized (RIGs) and
create new internal representations. Actively rewriting such representations related to the internal interactive experiences involves the process of reorganizing the experiences of differences from others in sense of self, and the pattern seen in Group H illustrates the state of its dynamic operation. It is also worth noting that there were micro and macro patterns in the mountain-valley pattern. The micro pattern is regarded as a movement to constantly adjust interactions by repeating fine tuning and asynchrony. It may be inferred that this is the basic organizational framework for that person’s interpersonal relationship. On the other hand, the macro pattern is considered to indicate organizing the experience of greater differences. There seems to have been more vivid experiences of self because of the large degree difference.

In contrast with the results of Group H as we have seen above, the pattern of Group L showed few repetitions of increase and decrease in synchronism, there were many monotonically decreasing or increasing sections or flat sections. From this result, it is presumed that the state of interaction in Group L is static compared to Group H, and that

![Fig. 3. The result of moving correlation analysis of unstructured phase.](image-url)
the organization did not work dynamically in experiences of differences. Under such circumstances, the role that one plays in interaction becomes fixed, and even if experiences of differences occur in the interactions, the experience never gets modified so that it is positioned newly in the relationship and will remain as a difference.

In the interview, it was said that this phase was the most difficult of the three in either group, and they commonly felt difficulties when they did not have any topics to talk about anymore. In Group H, it was often said that the situation where participants themselves talked unilaterally and the researcher did not reply was difficult, and the difficult situation was perceived in the context of the relationship. On the other hand, in Group L, it was reported that they had difficulty when there was no topic to talk about, such as “I didn’t know what to talk about from what perspective (L3)”, and the difficult situation was perceived in a context regardless of the relationship. The difference between these two experiences may be derived from the extent to which sense of self is open to interaction.

Conclusions

This study quantitatively showed the influence of nonverbal communication between two parties using video analysis and moving correlation analysis. Results revealed that in those with high sense of self, active mutual exchange was maintained between self and others in the non-linguistic domain. On the other hand, in those with low sense of self, mutuality in the nonverbal area was weak. In past case studies, it has been intuitively confirmed that psychotherapy involves sense of self in nonverbal areas as an important therapeutic factor. This study demonstrates that activities of sense of self are rooted in nonverbal areas even in adulthood, and suggests the importance of nonverbal areas in considering the clinical problems of sense of self of adults.

There were some limitations to the study. Although the timing at which nonverbal behavior occurred was measured, actual nonverbal behavior also included dimensions such as the form (the trajectory of motion and the waveform of sound) and the degree of the change, which are also considered to be responsible for nonverbal interaction. Additionally, there may be instability in terms of reliability, that there was only one coder. These should be examined by using video analysis by software together, or coding by plural coders and calculating the concordance rate of coding results. All of the findings obtained in this study were found through experimental situations. In order to apply these findings to practice, it is necessary to further refine the theory focusing on the operation of sense of self in the psychotherapeutic scene.

References


(Manuscript received 30 August, 2018; Revision accepted 14 November, 2018; Released online in J-STAGE as advance publication 5 September, 2019)