

Verbal and Nonverbal Skills in Open Communication: Comparing Experienced and Inexperienced Radio Duos

Noriko Suzuki¹(✉), Yu Oshima², Haruka Shoda^{2,3}, Mamiko Sakata²,
and Noriko Ito²

¹ Faculty of Business Administration, Tezukayama University,
7-7-1 Tezukayama, Nara 631-8501, Japan
nsuzuki@tezukayama-u.ac.jp

² Department of Culture and Information Science, Doshisha University,
1-3 Tatara Miyakodani, Kyotanabe, Kyoto 610-0394, Japan

³ Japan Society for the Promotion of Science, 5-3-1 Kojimatchi,
Chiyoda-ku, Tokyo 102-0083, Japan

Abstract. This paper examines how the difference in talk skill for open communication affects the orientation of the verbal and nonverbal behaviors of the talk partner or audience. An experiment was carried out using multiple radio duos having different levels of talk skill, i.e., experienced and inexperienced. The experiment's task was conducted in a pseudo-radio setting under three conditions: audience-present talk, audience-absent talk, and audience-absent/post-talk sessions. The speech and body gestures of all participants were video-recorded and analyzed. The results suggest that the different levels of experience in radio talk are expressed in different speech and gesture orientations. These findings seem applicable to the speech- and gesture-expression model for conversational robots, especially for nursing-care robots designed to talk with other robots or cohabitants.

Keywords: Open communication · Radio talk skill · Orientation of verbal and nonverbal behaviors · Experienced and inexperienced radio duos

1 Introduction

The development of domestic robots has progressed rapidly in recent years, targeting the needs of the coming aging society with fewer children. We will live at home with multiple robots for different purposes, i.e., nursing care, home care, and training, just like the movie droids “C-3PO and R2D2” in “Star wars” or the robotic boy “David” in “Artificial Intelligence: AI.” What would you think about such robots interacting silently with each other via a network right in front of you? The response of many people might be a feeling of alienation or anxiety. To solve this perception problem, domestic robots should have a model

of expressing explicit verbal and nonverbal behaviors to put users at ease [4]. The authors give attention to the mechanism of open communication as a key for expressing participant-oriented behaviors [2].

Open communication is a type of performer-to-audience communication in which the audience perceives indirect messages from direct conversation among performers [7]. Cooking shows and domestic comedies on TV program are typical examples of open communication. In particular, the authors focus on radio-duo shows as a distinct form of open communication. A radio-duo show is usually not watched directly by the audience. Nevertheless, the audience can perceive indirect messages from the conversation between the partners of a radio duo without any visual information. In other words, the radio duo can communicate to an invisible audience, although the duo partners seem to neglect the presence of the audience while talking to each other. The skill of radio talk in open communication might be a useful analog for increasing the ability of domestic robots to express participant-directed behaviors.

The authors conducted an experiment to explore radio talk skills by comparing experienced with inexperienced radio duos. A pseudo-radio session was selected as the task. Seven experienced and seven inexperienced radio duos took part in the experiment. In this paper, the partners in an experienced radio duo have at least one year of on-the-job experience on a university's radio show. The partners in an inexperienced radio duo have no experience in talking to each other on any radio show, although they may be acquaintances.

Conventional studies in multiparty interaction have shown that the presence or the absence of an audience affects the amount of speech or the speech orientation of the performer in a comic duo [9]. On the other hand, related works in gesture have shown that representational or beat gesture was produced more frequently in a face-to-face setting than in a separated setting [1,6]. However, few research works have taken into account the relationship between the difference in talk experience with an audience and the orientation of verbal and nonverbal behaviors.

A within-participants experimental design was used in three situations: audience-present talk, audience-absent talk, and audience-absent/post-talk sessions. Audience-absent/post-talk session was defined as closed communication, although both audience-present talk and audience-absent talk sessions are defined as open communication. The turn duration, speech intervals, frequency of back channels, duration of representational and beat gestures for each participant, and overlaps and gaps between the partners of the radio duo were annotated. The authors performed a two-way analysis of variance to examine the effects of skill (i.e., experienced versus inexperienced) and session (i.e., audience-present, audience-absent, post-talk) and to conduct a correlation analysis between a post-experiment questionnaire on attention given to the audience and the verbal/nonverbal behaviors of the performers.

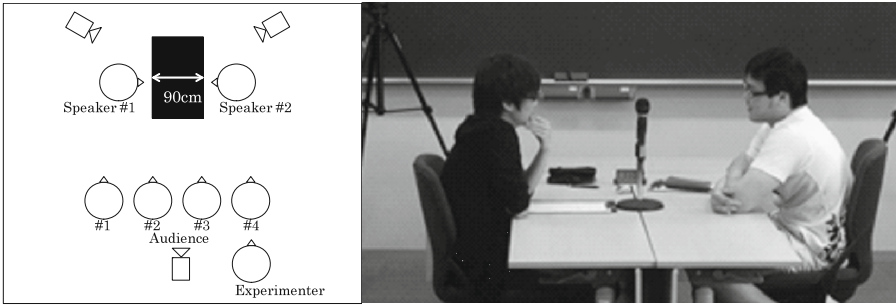


Fig. 1. Experimental setting: recording setup in audience-present talk session (left) and sample scene (right)

2 Method

2.1 Participants

A total of 28 graduate and undergraduate students (Mean age: 19.6 years, SD: 1.2) participated in the experiment, and they were assigned to either an “experienced radio duo” ($n = 14$, 7 pairs) or an “inexperienced radio duo” ($n = 14$, 7 pairs). In this paper, an experienced radio duo means that each partner has at least one year of on-the-job experience on a university radio show. The inexperienced radio duo means that the partners have no experience in talking to each other on any radio show, although they may be acquaintances.

2.2 Procedure

Each radio duo was instructed to sit down and face each other across the desk (Fig. 1). They participated in the pseudo-radio sessions: audience-present talk for 10 min, audience-absent talk for 10 min, and audience-absent/post-talk sessions for 10 min. In audience-present talk session, radio duo talked each other to four numbers of audience in front of the radio duo. It was a kind of open communication. In audience-absent talk session, radio duo talked each other to audience without physically being. They were instructed that audience listened their talk in a separate room. It was also a kind of open communication. Here, “post-talk” refers to a brief discussion after the simulated radio show in which the duo’s partners evaluated their performance. It was a kind of closed communication as against open communication. The first two talk sessions were counter-balanced. All radio duos in both talk sessions discussed the same topic, i.e., the item they would take to a deserted island. We video-taped the upper body of each participant with three video cameras (HDR-XR550V, Sony) and four wireless microphones (ECM-AW3T, Sony). After three sessions, participants answered a questionnaire on such items as attention to the audience.

2.3 Parameters

For 3 min within the 10-min recording of each radio duo, we extracted verbal and nonverbal behaviors by using the annotation software ELAN (EUDICO Linguistic Annotator [3]). We measured the duration of each of the conversations and the gestures made according to the following criteria.

Turn Duration Per Minute (in sec.): We measured the turn duration per minute through a talk session. The turn duration means the length of speech turn while a partner talks until the other partner begins to talk. Back channels are not included in a speech turn.

Speech Interval Per Minute (in sec.): We measured the speech interval per minute between two participants through a talk session. The speech interval includes both response latency and speech overlap.

Frequency of Back Channels Per Speech: We measured the number of back channels per speech. The back channels are utterances such as “Yeah” or “Umm”.

Frequency of Representational Gesture Per Speech: We extracted the number of each participant’s representational gestures that express semantic content related to the speech by virtue of the hands’ shape, placement, or motion (e.g., [8]).

Frequency of Beat Gesture Per Speech: We extracted the number of each participant’s beat gestures that express simple, rhythmic gestures that do not convey semantic content (e.g., [5]).

2.4 Predictions

Verbal Behaviors: We predict that the difference in radio talk experience will affect the verbal behaviors of the radio duos as well as the presence of audience does (e.g., [9]).

Nonverbal Behaviors: We predict that the difference in radio talk experience will also affect the nonverbal behaviors of the radio duos as well as the presence of audience does (e.g., [1,6]).

3 Results

3.1 Verbal Behaviors

The authors performed two-way analysis of variance to examine the effects of skill (i.e., experienced versus inexperienced) and session (i.e., audience-present, audience-absent, post-talk).

For the turn duration, the main effect of skill and that of session were not significant, but the two-way interaction was significant: $F(1,26) = 3.91, p = .04, \eta_p^2 = .13$ (skill \times session) (Fig. 2). A post-hoc *t*-test with Bonferroni’s correction showed that an inexperienced radio duo expressed longer turn duration per minute in the audience-present session than in the audience-absent session

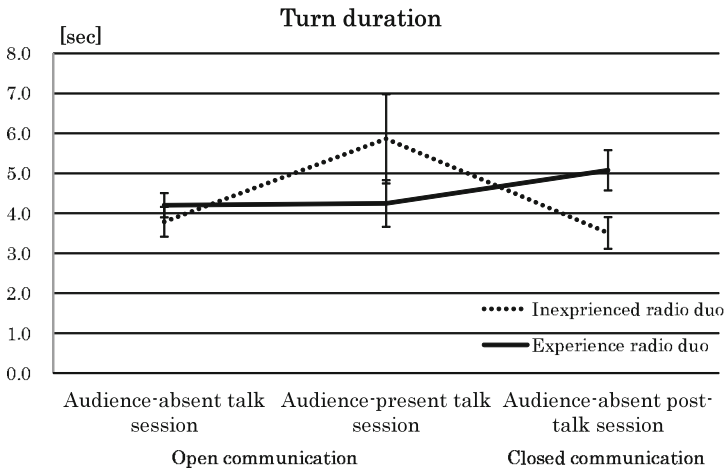


Fig. 2. Results of turn duration per minute

($p = 0.09, d = 0.83$), and in the audience-present session than in the audience-absent/post-talk session ($p = 0.07, d = 0.90$).

For the speech interval, the two-way interaction was not significant, but the main effect of skill and that of session were significant: $F(1, 26) = 6.74, p = .02, \eta_p^2 = .21$ (skill) and $F(1, 26) = 41.40, p < .001, \eta_p^2 = .61$ (session) (Fig. 3). A post-hoc t -test with Bonferroni's correction showed that both inexperienced and experienced radio duos expressed a longer speech interval per minute in the audience-present session than in the audience-absent session ($p = 0.09, d = 0.82$) and in the audience-absent/post-talk session than in the audience-present session ($p < 0.01, d = 2.51$).

For the frequency of back channel, the main effect of skill and the two-way interaction were not significant, but the main effect of session was significant: $F(1, 26) = 9.09, p < .001, \eta_p^2 = .26$ (Fig. 4). A post-hoc t -test with Bonferroni's correction showed that both inexperienced and experienced radio duos expressed a larger number of back channels per speech in the audience-absent session than in the audience-absent/post-session ($p = 0.02, d = 1.08$) and in the audience-present talk session than in the audience-absent/post-session ($p = 0.04, d = 1.24$).

3.2 Nonverbal Behaviors

The authors performed two-way analysis of variance to examine the effects of skill (i.e., experienced versus inexperienced) and session (i.e., audience-present, audience-absent, post-talk).

For the frequency of beat gesture, the main effects of session and two-way interaction were significant: $F(1, 26) = 6.22, p = .007, \eta_p^2 = .19$ (session), and $F(1, 26) = 6.13, p = .004, \eta_p^2 = .19$ (skill \times session), respectively (Fig. 5).

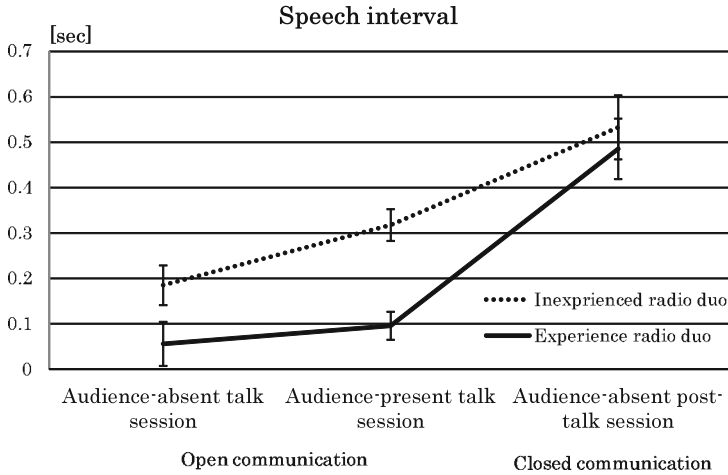


Fig. 3. Results of speech interval between radio duo partners per minute

In the results of the simple main effect, there was a significant difference in experienced radio duos through three sessions ($p = 0.03, d = 0.38$). A post-hoc t -test with Bonferroni's correction showed that experienced radio duos expressed longer frequency of beat gesture per speech in the audience-absent session than in the audience-absent/post-talk session ($p = 0.05, d = 1.16$) and in the audience-present talk session than in the audience-absent/post-talk session ($p = 0.02, d = 1.37$).

For the frequency of representational gesture, the main effect of skill was not significant but that of session and the two-way interaction did show a significant tendency: $F(1, 26) = 2.81, p = .07, \eta_p^2 = .10$ (session), and $F(1, 26) = 2.83, p = .06, \eta_p^2 = .10$ (skill \times session) (Fig. 6).

In the result of the simple main effect, there was a significant difference of experienced radio duo through three sessions ($p = 0.03, d = 0.24$). A post-hoc t -test with Bonferroni's correction showed that experienced radio duos expressed longer frequency of representational gesture per speech in the audience-absent session than in the audience-absent/post-talk session ($p = 0.03, d = 0.96$) and in the audience-present/post-talk session than in the audience-absent/post-talk session ($p = 0.08, d = 0.87$).

3.3 Post-experiment Questionnaire

Figure 7 shows the results of a post-experiment questionnaire on the attention of the performer to the audience.

The authors performed two-way analysis of variance to examine the effects of skill (i.e., experienced versus inexperienced) and session (i.e., audience-present, audience-absent).

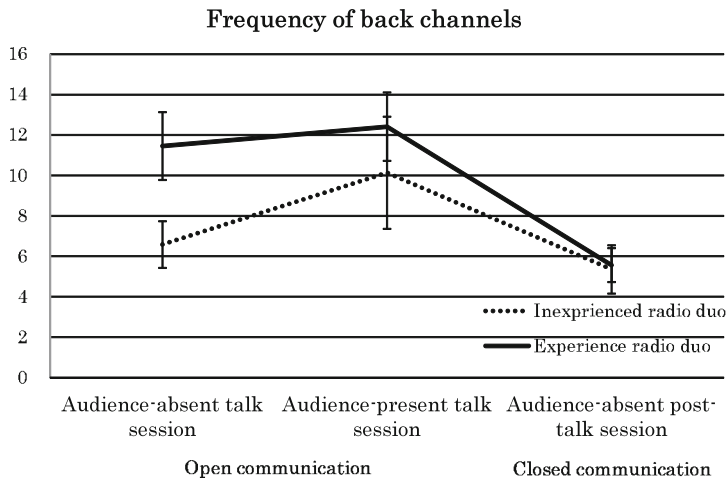


Fig. 4. Results of frequency of back channels per speech

For the frequency of beat gesture, the main effect of session was significant: $F(1, 26) = 5.93, p = .02, \eta_p^2 = .16$). This result suggested that the partners in both experienced and inexperienced radio duos talked to each other while paying more attention to the audience in the audience-present talk session than in the audience-absent talk session.

3.4 Correlation Between Verbal/Nonverbal Behaviors and Attention to the Audience

The correlation analysis between the post-experiment questionnaire and verbal and nonverbal behaviors showed that there were both a negative correlation between the attention to the audience and the turn duration ($r = -0.728, p < .01$) and a positive correlation between the attention and the frequency of back channel ($r = .560, p < .05$) in the audience-absent session in inexperienced radio duos (Table 1). In other words, the more attentive the radio duo was, the shorter the turn duration in an inexperienced radio duo was, or the larger the back channel was.

On the other hand, there was a negative correlation between attention and the frequency of beat gesture ($r = -0.659, p < .05$) in the audience-present session in experienced radio duos. In other words, the more attentive the radio duo was in experienced radio duos, the smaller the frequency of beat gesture was.

4 Discussion

In the present study, we found differences in verbal and nonverbal behaviors between experienced and inexperienced radio-duo talk through three kinds of

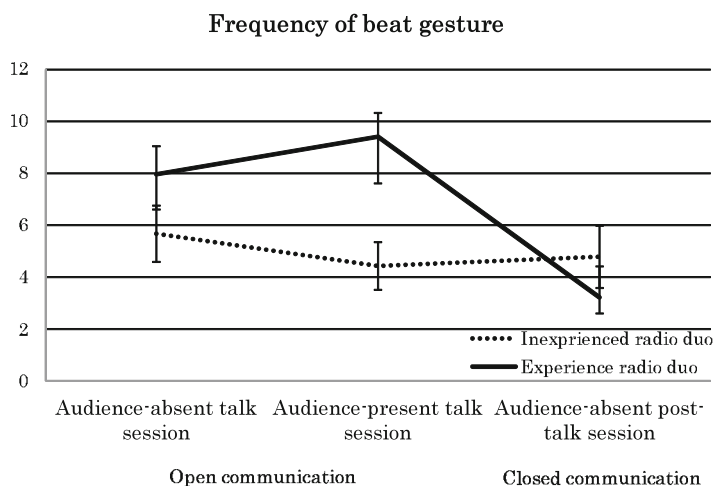


Fig. 5. Results of frequency of beat gesture

Table 1. Correlation between attention to the audience and verbal/nonverbal behaviors

	Attention to the audience			
	Experienced radio duo		Inexperienced radio duo	
	Audience-absent session	Audience-present session	Audience-absent session	Audience-present session
Turn duration	-.075	-.497	-.728 * *	-.065
Speech interval	.109	-.028	-.318	.372
Back channel	.301	.140	.560*	.243
BG	-.253	-.659*	.234	.133
RG	.031	.029	-.079	.287

BG: Beat gesture, RG: Representational gesture

*: $p < 0.05$, **: $p < 0.01$.

sessions, i.e., audience-present, audience-absent as open communication and post-talk as closed communication, in terms of “turn duration,” “speech interval,” “frequency of back channel,” “frequency of beat gesture,” and “frequency of representational gesture.” We also found correlations between the attention to the audience and the radio duo’s behaviors.

In their verbal behavior, experienced radio duos did not change so much in turn duration. Inexperienced radio duos used a longer turn duration in the audience-present talk session than in the audience-absent sessions. On the other hand, both radio duos seemed to show the similar tendency on the speech interval or the frequency of back channels. Although the amount of speech with the audience present increased more than that without an audience in the case of

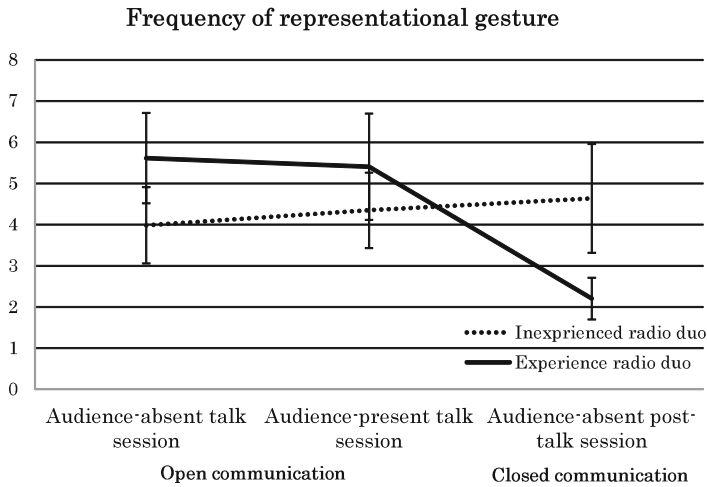


Fig. 6. Results of frequency of representational gesture

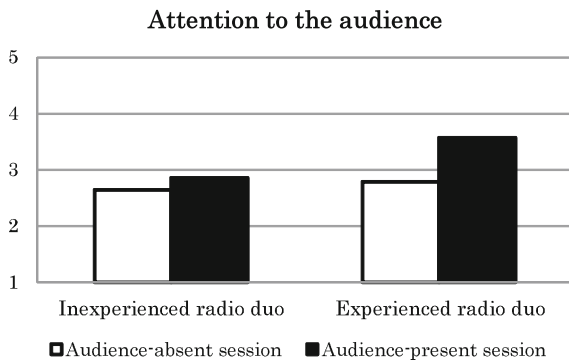


Fig. 7. Results of post questionnaire: attention to the audience

comic duos [9], our results suggested that the communication style, open or closed, affected the verbal behaviors. The difference in radio talk skill seemed not to affect the verbal behaviors so much except the speech interval. From the results of the correlation with attention to the audience, the turn duration has a negative correlation while the back channel has a positive correlation in inexperienced radio duos in the audience-absent session. The turn duration and the number of back channel might have a potential of participant-directed behaviors in inexperienced radio duos.

In the nonverbal behavior, inexperienced radio duos did not change so much their frequency of either beat or representational gestures. Experienced radio duos used a larger number of both beat and representational gestures in audience-absent and audience-present sessions than in the post-talk session. Although the presence of audience affects the frequency of both beat and

representational gestures in the conventional studies in gesture production [1,6], our results suggest that both the communication style, open or closed, and the difference in radio talk skill seemed to affect the frequency of gesture production. From the results of the correlation with attention to the audience, the frequency of beat gestures has a negative correlation in experienced radio duos in the audience-present session. The frequency of beat gesture might have a potential of addressee-directed behaviors in experienced radio duos.

In conclusion, different levels of radio talk skills show different type of nonverbal expressions. The open communication style also affects the radio duo's speech and gesture production. Besides, it might be possible that the verbal behaviors indicate participant-directed acts, although the nonverbal behaviors indicate addressee-directed acts. In our future work, we will search for explicit speech and gesture orientation by analyzing more detailed aspects of the verbal and nonverbal behaviors of radio duos based on Clark and Carlson, 1982 [2]. These results may lead to several implications for constructing a narrative-strategy model for communication robots that can alleviate the sense of alienation felt by the users.

Acknowledgment. The contents of this study are based on the second author's master thesis. We thank 28 students at Doshisha University for their participation in the experiment. We also thank Haru Nitta, Kana Shirai, Koshi Nishimoto, and Tomoki Yao for their fruitful discussions on this study.

References

1. Alibali, M.W., Heath, D.C., Myers, H.J.: Effects of visibility between speaker and listener on gesture production: some gestures are meant to be seen. *J. Mem. Lang.* **44**, 169–188 (2001)
2. Clark, H.H., Carlson, T.B.: Hearers and speech acts. *Language* **58**(2), 332–373 (1982)
3. ELAN. <https://tla.mpi.nl/tools/tla-tools/elan/>
4. Matsuyama, Y., Taniyama, H., Fujie, S. and Kobayashi, T.: Framework of communication activation robot participating in multiparty conversation. In: *Proceedings of AAAI Fall Symposium, Dialog with Robots*, pp. 68–73 (2010)
5. McNeill, D.: *Hand and Mind: What Gestures Reveal about Thought*. University of Chicago Press, Chicago (1992)
6. Mol, L., Krahmer, E., Maes, A., Swerts, M.: Seeing and being seen: the effects on gesture production. *J. Comput.-Mediated Commun.* **17**(1), 77–100 (2011)
7. Ohba, M., Okamoto, M., Enomoto, M., Iida, H.: Tittering and laughing: a case of manzai audience. In: *Proceedings of LIBM 2008: First International Workshop on Laughter in Interaction and Body Movement*, pp. 34–39 (2008)
8. Ozyurek, A.: The influence of addressee location on spatial language and representational gestures of direction. In: McNeill, D. (ed.) *Language and Gesture: Window into Thought and Action*, p. 6483. Cambridge University Press, Cambridge (2000)
9. Sakata, M., Tanno, M.: Multimodal interactions in duo-comic acts manzai. In: *Proceedings of the 8th International Conference on Humanized Systems 2012 (ICHHS 2012)*, pp. 63–66 (2012)