

Musically Expressive Doll in Face-to-face Communication

Tomoko Yonezawa^{†, *1} and Kenji Mase^{‡, *2}

[†] ATR Media Integration & Communications Research Laboratories

[‡] ATR Media Information Science Laboratories

yonezawa.tomoko@lab.ntt.co.jp, mase@atr.co.jp

Abstract

We propose an application that uses music as a multi-modal expression to activate and support communication that runs parallel with traditional conversation. In this paper, we examine a personified doll-shaped interface designed for musical expression. To direct such gestures toward communication, we have adopted an augmented stuffed toy with tactile interaction as a musically expressive device. We constructed the doll with various sensors for user context recognition. This configuration enables translation of the interaction into melodic statements. We demonstrate the effect of the doll on face-to-face conversation by comparing the experimental results of different input interfaces and output sounds. Consequently, we have found that conversation with the doll was positively affected by the musical output, the doll interface, and their combination.

1. Introduction

In this research, we sought to adopt musical expression as a new form of communication that runs parallel to other verbal and nonverbal modalities. People communicate with each other by using multi-sensory expressions that exploit gestures, gaze and other detectable elements. These nuances make communication smooth, natural, redundant, unburdening, and robust. Recently, humans have developed new means for expression that employ technology and enhanced media, such as communicative acoustics. However, people have not yet acquired a musically expressive method to augment conversation.

We propose using musical expressions as a means of multi-modal communication that activates and supports message conveyance during conversation. We adopted a stuffed toy interface as a musically expressive device [14] and outfitted the doll with various sensors. An internal PC interprets the doll's present situation and the tactile interaction of the user, then translating this interaction into melodic utterances. In the preliminary experiment, we observed that people enjoyed the system as they displayed intimate interactions and accepted the musical sound output as responsive expressions of the doll interface. We now find this research needs systematic evaluation to study whether such an interface is useful for expression,

and if it may serve as a new channel of communication parallel to conversation. In this paper, we first introduce our musically expressive doll interface and then analyze experimental results on the effectiveness of the doll as a new communication device.

2. Motivation

When people converse using words, they instinctively add to their vocal communication with other expressions such as gestures or fingering. We assume that musical or sound expressions give several modalities to the communication and that the doll interface permits richer interactions.

There are several existing forms of musical expressions that are equivalent to the multi-modalities of communication. For example, there are interesting cultural media such as *Utagaki*, a traditional Japanese custom in which young boys and girls contend with each other in songs to communicate romantic emotions to members of the opposite gender. In that communication style, indirect musical expressions are used along with verbal expressions. Music is also often used in movies or musicals to make the scene dramatic, which is a one-way augmentation to the audience. Our work was originally motivated by the desire to integrate conversational communication with additional musical channels.

We now propose communication using the various music expressions that are performed by a doll interface and we examine its effects and the doll device itself. Juslin [5] introduced aspects of music modality, focusing on the peculiar nature of sound in relation to emotions. We examine the variety of musical expressions that are generated by various interfaces and situations, including communication between people.

We then consider the several functions of a doll. People treat dolls as another self or a partner. There is a function called "self-adapter" in nonverbal expressions, as seen in someone who is gesturing and fingering during a conversation. The doll's embodied physical nature would seem to elicit such a behavior. When a doll is used as another self it would be serving the function of "embodiment," as seen in ventriloquism and house-playing. This is probably caused by the doll functioning as a personified interface for controlling other media or interacting with it.

The importance of nonverbal communication, regardless of whether it's conscious expression or not, was emphasized by Vargas [12] and Raffler-Engel [13]. They

^{*1} Currently with NTT Cyber Space Laboratories

^{*2} Currently with Nagoya University / ATR-MIS

witnessed several functions of nonverbal communication shown by people using facial and body expressions. It has been observed that people who interact with computers behave as if they were communicating with another human [7]. We feel that a doll interface could provide some functions of these nonverbal expressions more explicitly with the aid of its personified look and role as a metaphor.

Following this trend, we propose the musical expression of the doll interface as a communication method. Focusing on the multi-modality of communication, we examine a new style that accompanies sympathy, conception, and creation. Furthermore, we expect the doll to express ambient or projected emotion, or the intentional expression that reflects a user who uses non-verbal communication.

3. Related Works

There have been several efforts to develop interactive systems using dolls. "Noobie" [3] was proposed as a sympathetic computer-human interface for children. The "Swamped!" testbed [4] was also built as a "sympathetic interface" using a sensor-doll. The doll is mainly used as the controller for a character's behavior in its own graphically presented virtual world. Cassell et al. [1] proposed a peer doll as a storytelling interface for children by exploiting the tendency of people to act with a doll as if it were alive. From those researches we see that doll interfaces have been used as both familiar and personified device. "ActiMates Barney" [10] is a commercialized sensor doll available as a playmate for learning through interaction. "My Real Baby" is a similar robotic device that emulates an infant with various facial expressions.

In contrast to these systems, we use sound effects and music, rather than the conventional media controls of toy dolls, to extend the expressive capability of the actuators and apply them in an unobtrusive manner. "The Familiar" [2] is an automatic diary system using a sensor-equipped doll (or a backpack sensor apparatus) and a context recognition system. Its purpose is to achieve automatically record the user's behavior and an outline of its associated context. We adopt this system's context sensing mechanism in our prototype.

Interesting work has been done with personified robots performing as human-to-human communication tools. A robot for elderly people has been introduced as a healing and communicating tool [6]. "RobotPHONE"[9] is a pair of simple shape-shared dolls with sensors and motion actuators. Suzuki et al. have introduced a "Mobile Robot Platform for Music and Dance Performance" [11], which is an interactive system built with a robot that uses musical expression. This work involves both musical expression as a communication method and as a personified interface. However, this system does not emulate communication between humans.

4. Context-aware Stuffed Toy Prototype System With Musical Expressions

We previously proposed Com-Music, a sensor-equipped doll for musical expression (Fig. 1) to facilitate communication among people [14]. We believe that a doll interface can become a kind of medium for musical communication by providing the user with several expression controls and harmonizing their combinations. The doll can also be a playmate toy for entertainment purposes, a part of a music edutainment system, and a human-to-human communications support device over a network. Thus the doll interface can be used as a familiar and flexible communication device.

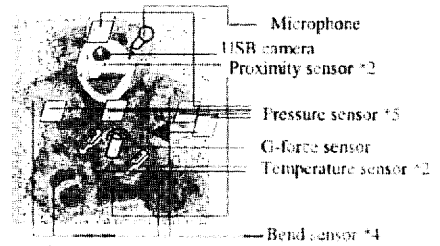


Figure 1: Sensor-doll mappings and setup

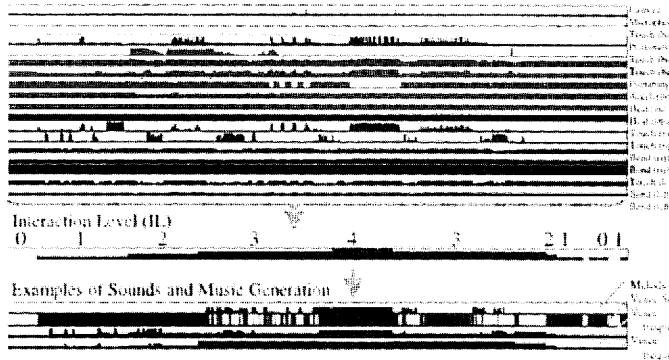


Figure 2: Sensor data, interaction levels, and music expressions

We first designed our context-aware sensor-doll as a reactive communicator for the user with the doll itself [14]. We also constructed a testbed of a networked musically expressive doll to enable musical communications [8]. The sensor-doll has several internal modes and accepts two kinds of reactive controls: (1) context-recognition for mode switching and (2) direct input translation in each mode. The internal modes of the doll are divided into five states that represent the doll's internal status, which resembles a particular mood. The transitions between states are controlled by the interaction with a finite state machine. Each state's design closely corresponds to the strength of activities.

The system generates music and sounds controlled through a context-based interpreter, which processes raw sensor data (Fig. 2). In the music performance state the doll performs as a musical controller while allowing its partner to play music cooperatively. The musical expressions have global or partial controls such as melody, rhythm, key, and chord.

The prototype system was provided to a number of visitors to our laboratory. As a preliminary experiment, we observed people's reaction to the doll's shape, musical expression, and mode change based on its context sensing. We observed that people enjoyed their interaction with the doll. For people expecting an intelligible voice, the muttering sound the doll makes is always regarded as unusual, but users come to enjoy making various sounds quite quickly. The mode change is recognizable from the sound transitions between the internal states [8].

5. Designed Experiment

5.1 Experiments on Face-to-face Communication with Musically Expressive Doll

For musical expression to be effective as a new communication channel, high-quality conversations must be generated and the doll should provide supplemental expressions to the modality of conversation. To discuss the effect of the musically expressive doll on communication between people, we observed human-to-human conversations using the musical sensor-doll.

Sensors	Mapping 1 (melodic)	Mapping 2 (voice-like)
Pressure sensors	Harmonics, Volume	Harmonics, Volume
Bend sensors	<i>Discrete</i> notes	<i>Indiscrete</i> notes

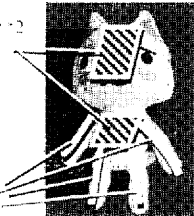


Figure 3: Setup of music-doll for experiment

We adopted a cat-like stuffed doll for this experiment (Fig. 3). The doll is equipped with four bend sensors in its legs and arms and two pressure sensors in its head and trunk. The sensors are connected to an A/D converter through cables, and the signals are sent to a Macintosh computer that generates music and sound.

For this experiment, we prepared special audio mappings that are very easy for musical novices to control in conformity with the Com-Music system. The basic controls include 1) volume and harmonics mapped at the head and trunk, and 2) melodic elements at the bend sensors. Each condition, described in the section on Conditions, has a different mapping to basic musical elements. For example, some music notes, actuated by the bend sensors change either non-discretely or discretely based on the signal of the sensors' values. A simple explanation of the mappings is given in Fig. 3. Mapping 1 of the figure is the musical sound mapping using melodic sound, and Mapping 2 is the voice-like sound made by continuous change in the melody, which is equivalent to the pitches of our speaking voice.

Method: We divided the subjects into two groups: Group **P** (Player) was comprised of the examinees who were given devices, and Group **L** (Listener) consisted of the examinees who were not given the device for use during the experiment. In each test, we formed 14 same-

gender pairs of one person from Group **P** and one person from Group **L**. The members of the pairs met each other for the first time and had never talked with each other before the experiment; this constraint was adopted to observe how the system *helps* their first communication. The partners conversed with each other under various conditions discussed later. The topics of the conversation were prepared on simple subjects such as "melons" or "tomatoes," and given to the pair before the experiment. The orders of conditions and topics were randomized.

Through headset microphones worn by the examinees, we recorded the speaking sounds of the Player and the Listener as the conversation's voice dialogue, and we also recorded the musical sound from the device (if any) performed by the Player. A video recording of each experiment was also made for our reference in the analysis.

Hypotheses: We made two predictions before conducting the experiment. First, that the conversation would be influenced differently in terms of the total amount of speaking and its balance between speakers. We anticipated this would be as a result of assigning each speaker a different musical device including a traditional musical instrument (piano), a sensor-doll, or no device at all. Secondly, we expected that the conversation would be different in terms of the total amount of speaking and its balance between speakers. This would develop as a result of assigning device different sound mappings to the speaker's device such as melodic, voice-like, or no sound at all.

Conditions: We conducted two experiments: the first studied the effect of different musical devices on conversation, and the second the effect of different musical reactions using the sensor-doll device. *Experiment #1* consisted of three conditions: I) the player must use the sensor-doll with a mapping from its tactile sensor input to a simple musical control (C_{doll} , which is the same as Mapping 1 of Fig. 3) during the conversation, II) the player must use a piano (C_{piano}) in the same way as C_{doll} and III) the player may not have a device (C_{no_device}). *Experiment #2* had three more liberal parameters: I) the player may play the musical control (melodic and harmonic sound) with the sensor-doll but it is not required (C_{melody} , Mapping 1 of Fig. 3), II) the player controls sound in the same way as the voice control (C_{voice} , not melodic, non-discrete frequency shown in Mapping 2 of Fig. 3), and III) the player can control nothing but may talk to Listener while touching the doll (C_{no_sound}).

Subjects: Twenty-eight undergraduate, graduate students, and researchers, from 18 to 35 years old (14 males and 14 females).

Procedure: The experiment was performed with the following procedure.

Before the experiments, (I) a subject (Player) was given the musically expressive device under each condition (piano, doll, nothing). (II) Player had a few minutes to become familiar with the new musical device. If the

condition was C_{no_device} , this step was skipped. (III) Both subjects, Player and Listener, learned of the topic. (IV) The subjects talked with each other under one of the three conditions (C_{doll} , C_{piano} , or C_{no_device}). During the conversation, Player touched or controlled the device (or nothing).

Afterwards, (V) *Sound Performance Change*: Experiment #1 repeats steps (I) to (IV) under each of the other two conditions. The order of the conditions was randomized. (VI) *Sound Feedback Change*: Experiment #2 is conducted in the same way as in step (V) under each condition (C_{melody} , C_{voice} , and C_{no_sound} , randomized). (VII) After conducting the experiments under all conditions, the subjects answered questionnaires.

Instructions: The experiment conductor asked the subjects to talk with one another about the given topic in each condition for 1-5 minutes. The conductor also explained the following conditions to the subjects: (A) the conversation must be more than one minute long, but if the conversation exceeds five minutes it will be stopped, (B) each subject has to express at least one opinion on the given topic or relate a narrative to the other, (C) each subject must listen carefully to the remarks of the other. They were also told that they could stop the conversation anytime if they wanted to, and (D) that they could freely talk with each other provided that the above conditions were satisfied.

The Players were also told to perform music or sound devices freely following the directions of the experimenter under each condition during the conversation.

Measures: For measuring the eagerness of the conversations, we detected utterance-intervals from each recorded sound (Listener, Speaker, and Sound of music performed by using the experimental system) by an ON/OFF judgment that uses a threshold of the sound volume. We then added the period of all intervals of each recorded sound under various conditions.

For convenience, we call the total time of the "volume ON" all combined intervals of each speaker or sounds from the device $L_{condition}$, $P_{condition}$, and $S_{condition}$ (the total time of Listener's voice, Player's voice, and Device's sound, respectively, for each condition of piano, doll, no_device, melody, voice and no_sound).

To measure whether the music and its device affect the conversation, we adopted the **L/P value**, which is the ratio of intervals between L and P and thus represents the balance of the conversation between the two people. The initiative of a conversation does not stay with one side but changes during the conversation. However, in general we can assume that either Player or Listener took the initiative, which can be determined by the dominance of speaking time.

We additionally adopted **P*** and **S*** values to express the total time while the Player concentrates in talking without playing the device (**P***) and the total time while the Player concentrates in performing some sound without

talking (**S***). These are different from the S and P values which are the simple totals of the utterance periods.

The questionnaire given to the examinees after the experiments used the SD method with a 1-7 scale. A higher score means that they feel the questionnaire statement matches their experiences.

5.2 Results

In this experiment, 13 out of the 14 pairs of examinees talked over the maximum length experiment (300 seconds) for every condition. The remaining pair talked for about 172.8 seconds (average). Both Listeners and Players sometimes lost the conversation about the given topic and tried to recover it. Several Players touched the doll or the piano when they lost the topic while talking with Listeners. Furthermore, some Listeners asked Players to perform the music when they lost the topic, apparently motivated in the same way as the Players. Thus, the five-minute conversations included both periods of lively mood and periods of subdued mood. Therefore, we concluded that a five-minute conversation was sufficient to provide statistically reliable data that includes diverse states of conversation.

Figure 4 shows two photos of this experiment. We could observe two unique interactions between Players and the doll regardless of music, sound, or no feedback. Some Players tried to show the doll to the Listener during the conversation (Fig. 4(a)) while making no reference to the content of the conversation. Other Players continued to look at the doll even during the conversation (Fig. 4(b)).

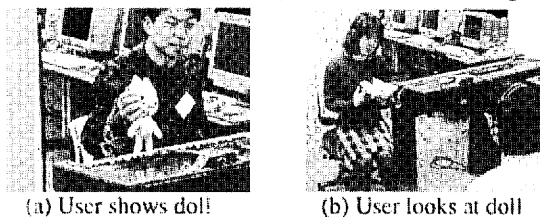


Figure 4: Talking experiment using doll

5.3 Analysis

[Analysis 1]: General analysis

Analysis 1-1: Significant Period of Performance

Under the condition C_{doll} of Experiment #1, we computed the normalized S_{doll} values, the totals of the Players making sound using the doll, by dividing S_{doll} by the sum of L_{doll} and P_{doll} , $S_{doll}/(L_{doll}+P_{doll})$. The {average, variance, max, minimum} of these normalized values were {0.576, 0.053, 1.44, 0.25}.

Analysis 1-2: Significant Balance of Performance

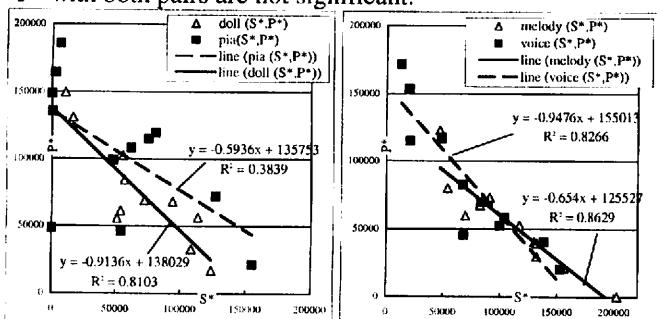
Figure 5 shows the plots of **S*** and **P*** values under each condition except the conditions C_{no_device} and C_{no_sound} , since these conditions do not have any sound output. Only 12 samples were taken into account in this analysis after excluding two samples: one pair's conversation was too short, and the other suffered from a failure to

record voice. The figure shows that the period of each experiment, $S^* + P^*$, was around 150 seconds.

In C_{piano} , the values of the ratio of S^* to P^* are scattered (Fig. 5(a), plotted by "■"). Compared with the R^2 values (fitness) of the other conditions, the R^2 value of S^* to P^* in C_{piano} was lower (0.3839, Fig. 5(a)) than the others (0.8103, 0.8629, 0.8266). These values can be interpreted as showing that the ratio of S^* to P^* in C_{piano} irregularly loses the balance that was maintained in C_{doll} .

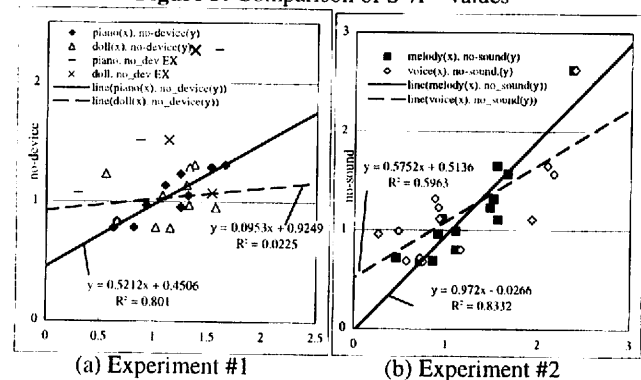
Analysis 1-3: Independent utterance of Sound and Player

We examined T-tests of both S^* and P^* values of the pairs (C_{doll} , C_{piano}) and (C_{melody} , C_{voice}). The T-value of S^* with (C_{melody} , C_{voice}) is significant, $T = 2.44 > 2.20$ (12, $p < 0.05$). The T-values of S^* with (C_{doll} , C_{piano}) and that of P^* with both pairs are not significant.



(Horizontal axis: S^* , vertical axis: P^* in both figures [millisecond])
(a) S^* , P^* Values (C_{doll} , C_{piano}) (b) S^* , P^* Values (C_{melody} , C_{voice})

Figure 5: Comparison of S^*/P^* values



(a) Experiment #1 (b) Experiment #2
(H: L/P_{doll} or L/P_{piano} , V: L/P_{no_device}) (H: L/P_{melody} or L/P_{voice} , V: L/P_{no_sound})

Figure 6: Comparison of L/P values

Analysis 2: Comparison of Total Performing Period

Analysis 2-1: Total Performing Period in Experiment #1

Next we focused on the total time of the performance with the doll and the piano by the Player. In Experiment #1, the average of S_{piano} was about 87.0 seconds, ranging from 0.8 to 188 seconds. We found two clustered groups in the samples. In the first group, consisting of nine of fourteen pairs, the average of S_{piano} values was 133.6 seconds, while in the other group, the values were all under 10 seconds, with an average of 3.8 seconds. On the other hand, the average of S_{doll} was about 108 seconds, generally dispersing from 5 to 250 seconds.

Analysis 2-2: Total Performing Period in Experiment #2

In Experiment #2, the averages of S_{melody} and S_{voice} were 174.14 and 124.75 seconds, respectively. For the difference between S_{melody} and S_{voice} , the T-value was $2.336 > 2.160$ (14, $p < 0.05$).

Analysis 3: Comparison of L/P values

We then analyzed data with an L/P value adopting the convenient description $L_{condition} / P_{condition} = L/P_{condition}$.

Analysis 3-1: L/P Values of Experiment #1

The 13 ratio values of (L/P_{piano} , L/P_{no_device}) and (L/P_{doll} , L/P_{no_device}) are plotted in Fig. 6(a). One sample was excluded because of voice recording failure. Linear relevance between L/P_{piano} and L/P_{no_device} can be seen except for the values represented with "-" marks, which are regarded as exceptions. Compared to this result, the plot of the ratio L/P_{doll} to L/P_{no_device} is scattered except for the data corresponding to "x," which are plotted by "-". The weak but similar tendency of a linear relation is observed even in the exceptional data above (Fig. 6(a)) with a different slant. The slant of resulting linear regression for L/P_{piano} is significantly larger than that of L/P_{doll} .

Analysis 3-2: L/P Values of Experiment #2

In Experiment #2, the relationship of L/P_{melody} to L/P_{no_sound} is roughly approximated as the expression $y \approx x$ (Fig. 6(b)). On the other hand, the relationship of L/P_{voice} to L/P_{no_sound} is scattered, as shown by the difference in the R^2 values, $\{(L/P_{melody}, L/P_{no_sound}), (L/P_{voice}, L/P_{no_sound})\} = \{0.8332, 0.5963\}$ (Fig. 6(b)). Moreover, the ratio of L_{voice} is higher than that of P_{voice} , as can be seen from their approximated lines with lower slant than ($L/P_{melody}, L/P_{no_sound}$).

We compared C_{no_device} with C_{no_sound} , and the correlation between L/P_{no_device} and L/P_{no_sound} was 0.04.

Analysis 4: Subjective evaluations

We found several remarkable results in the subjective evaluations. In the evaluations of the ease of conversation, the average of C_{piano} values by both Player and Listener are low (3.29, 3.71) compared to C_{doll} : 3.85, C_{no_device} : 4.64 for Player and C_{doll} : 3.93, C_{no_device} : 4.5 for Listener. The values are totally dispersed. In particular, the Players' evaluations are dispersed in C_{doll} and C_{piano} (variance: 3.82 and 3.14, respectively, when C_{no_device} is 2.40).

6. Discussion

We first confirmed that the musically expressive doll was utilized effectively during conversation. While Analysis 1-2 shows the total of the Player's expression by only speech or sound was stable, Analysis 1-1 shows that the Player sufficiently performed as much as he/she talked. We then concluded that Players were able to make the sensor-doll generate a sufficient quantity of expressions without hesitation.

The doll interface was easily played comparing with C_{piano} (subjective evaluations). Each of the examinees

may have different preferences for traditional musical instruments. It is shown in Analysis 2-1 that two clustered groups appeared in piano performance. Traditional instruments would be difficult to play during a conversation, especially for beginners. In contrast, the Player shows some tendency of playing a longer time with the doll interface in terms of average and dispersion. Therefore, we concluded that using the doll interface for musical expression could be an effective way to introduce a new modality of conversational communication, functioning in a different way from the piano.

Having the doll gives the conversation irregular and various effects. The Listener/Player utterance balance (L/P values) of Analysis 3-1 and 3-2 show the effects on the balance of conversation with playing the doll were irregularly changed, regardless of kinds of sound expressions, while the effect of playing piano is stable, and the balance was not changed. Players had different feelings than the Listener in regards to the subjective aspects of the experience.

Additionally, we observed a difference in conversation by the kinds of sound feedback. Players made the musical sound with the doll more than they made the voice-like sound in both the whole and independent utterances (observed in Analyses 2-2, 3-2). From this we conclude that voice-like sound is disturbing to conversation and that there is a positive role of melodic expression. It is supposed that this form of musical expression is attractive and effective for face-to-face conversation. As a supplement, Analysis 1-3 shows that *sound feedback* affects the balance of the expressive device's independent utterance, rather than *the type of input device*.

Finally, we conclude that the musical expression using the doll is easy, and it affects face-to-face communication behavior. From these results showing the effect on the conversation having this doll, we regard this new type of musical expression as a positive addition to multi-modal communication.

7. Summary

In this research, we aimed to adopt musical expressions as a new channel of communication that operates parallel to other verbal and nonverbal methods. The experiments conducted with a musically expressive doll, which included analyses of Player's talking-only time, Player's sound performing-only time, and the Listener/Player utterance balance led us to conclude that the doll provided a new form of communication that affects the balance of conversation.

Some remaining issues for further study include: 1) experiments on two-person communication when both individuals use the musically expressive doll, 2) investigation of the superiority of musically experienced people, and 3) a redesign of the doll and musical expressions for more suitable expression.

Acknowledgements

The authors would like to thank Brian Clarkson, Kazuyuki Saito, Yasuyuki Sumi, David Ventura, Ryohei Nakatsu, Norihiro Hagita and other ATR members for their help and discussion on this work. This research was supported in part by the Telecommunications Advancement Organization of Japan.

References

1. Cassell, J., Ananny, M., Basu, A., Bickmore, T., Chong, P., Mellis, D., Ryokai, K., Smith, J., Vilhjalmsson, H., and Yan, H., "Shared Reality: Physical Collaboration with a Virtual Peer," Proceedings of CHI2000, pp.259-260, 2000.
2. Clarkson, B., Mase, K., and Pentland, A., "The Familiar, a living diary and companion," CHI2001 Extended Abstracts, pp. 271-272, 2001.
3. Druin, A., "NOOBIE: The Animal Design Playstation," SIGCHI Bulletin, 20(1), pp. 45-53, 1988.
4. Johnson, M., Wilson, A., Kline, C., Blumberg, B., and Bobick, A., "Sympathetic Interfaces: Using Plush Toys to Direct Synthetic Characters," Proceedings of CHI 98, pp. 288-295, 1998.
5. Juslin, P. N., "Perceived Emotional Expression in Synthesized Performances of a Short Melody: Capturing the Listener's Judgment Policy," MUSICA SCIENTIAE, Vol. 1, No. 2, pp. 225-256, 1997b.
6. Lytle, M., "Robot care bears for the elderly", <http://www.globalaging.org/elderrights/world/teddybear.htm>, 2002
7. Reeves, B. and Nass, C., "The Media Equation: how people treat computers, television, and new media like real people and places," CSLI Publications, 1998.
8. Saito, K., Yonezawa, T., and Mase, K., "Awareness Communications by Entertaining Toy Doll Agent," International Workshop on Entertainment Computing 2002, to appear.
9. Sekiguchi, D., Inami, M., and Tachi, S., "RobotPHONE: RUI for Interpersonal Communication," CHI2001 Extended Abstracts, pp. 277-278, 2001.
10. Strommen, E., When the Interface is a Talking Dinosaur: Learning Across Media with ActiMates Barney. Proceedings of CHI 98, pp. 288-295, 1998.
11. Suzuki, K., Tabe, K., and Hashimoto, S., "A Mobile Robot Platform for Music and Dance Performance," Proceedings of International Computer Music Conference 2000, pp. 539-542, 2000.
12. Vargas, M., "Louder than Words -An Introduction to Nonverbal Communication-", Iowa State University Press, 1986.
13. von Raffler-Engel, W., "Aspects of Nonverbal Communication," Swets and Zeitlinger, 1980.
14. Yonezawa, T., Clarkson, B., Yasumura, M., and Mase, K., "Context-aware Sensor-Doll as a Music Expression Device," CHI2001 Extended Abstracts, pp. 307-308, 2001.