Regular Papers

Development of Universal Communication Aid and Its Design Concept—for Use by Hearing-impaired People and Foreign Travelers

Kaoru Nakazono†, Mari Kakuta, Yuji Nagashima, and Naotsune Hosono

Abstract

We are developing a communication aid called VUTE (visualized universal talking environment) that operates on portable information devices. VUTE is aimed at foreign travelers or hearing/speaking impaired people with a disability for speaking Japanese (or the local language). It displays motion pictograms that give users better comprehension. Furthermore, it is unique in incorporating the expressive manner of sign language in the design of the motion pictograms. We created a prototype of VUTE (VUTE 2009), which can be used to call an ambulance and/or fire engine in an emergency such as a sudden illness or fire. This paper describes the basic concept of our research project and describes the design concept and an overview of VUTE 2009. Finally, results of an evaluation test of VUTE 2009 are presented.

1. Introduction

People who are hard of hearing as a result of old age, deaf people, and people traveling abroad experience similar communication barriers because they cannot communicate freely in the spoken language. In an emergency such as a natural disaster, accident, or sudden illness, it is crucial that people have access to minimum essential communication functions to enable them to call for help. Our goal is to establish a communication technique for such people and situations.

To achieve this goal, we are currently designing a communication aid called VUTE (visualized universal talking environment) that does not rely on a specific language and can be used without prior training. As a preliminary stage in this project, we created a prototype of VUTE (VUTE 2009), which can be used to call an ambulance and/or fire engine in an emergency. VUTE is implemented on a portable electronic device and displays motion pictograms in contrast to most of the current communication aids, which use still pictograms.

In the study described in this paper, our aim was to confirm that a communication aid using motion pictograms is a practical means of achieving our goals. This paper first describes the basic concept of our research project. It then describes the design concept and gives an overview of VUTE 2009. Finally, it presents the results of an evaluation test of VUTE 2009. This evaluation was a preliminary version of the full-fledged test to follow, but it surely proved the system's efficiency.

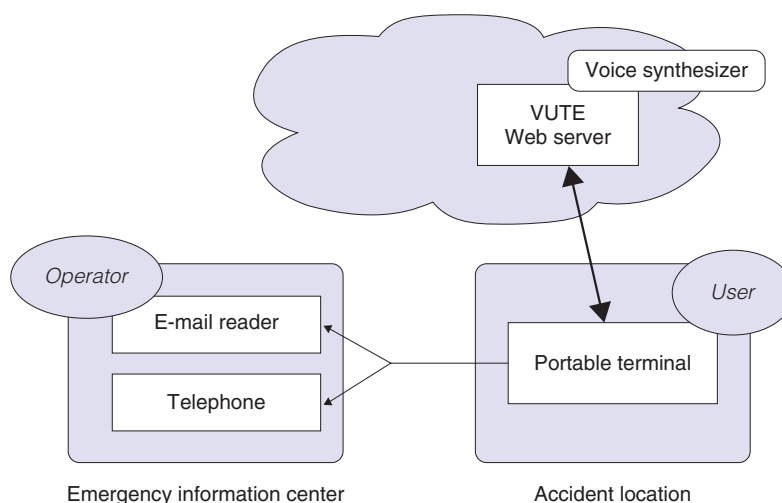† NTT Network Innovation Laboratories
  Yokosuka-shi, 239-0847 Japan

Fig. 1. Block diagram of VUTE in practical use.

2. Features of VUTE

When people with different backgrounds try to communicate with each other, they may encounter a variety of communication barriers, such as differing language, culture, sense of values, knowledge, experience, and physical disability. The attempt to use information and communications technology (ICT) to overcome such barriers is called *universal communication*. Among its various forms, what is called augmentative and alternative communications (AAC) is used by many people as an alternative communication means [1], [2]. AAC includes low-technology techniques, such as the user pointing at letters on a board to make a sentence or showing picture cards. In contrast to such techniques, methods of assisting communication by using electronic devices are often called *communication aids*.

VUTE is a technology intended to achieve universal communication. It targets (1) people who cannot speak aloud due to sudden illness or injury, (2) people who have a hearing disability or articulation disorder, and (3) people, such as foreigners, who cannot speak or read the local language. It uses motion pictograms instead of characters to navigate through a conversation in an emergency situation. A block diagram of VUTE in practical use is shown in **Fig. 1** and the technical features of VUTE are described below.

2.1 Character-free design

Most AACs based on conventional pictograms use text prompts as a supplementary means of communi-

cation. Although text does considerably enhance understanding and convenience for those who can read it, characters are of no help to those who cannot read them.

Many people say that hearing-impaired people can communicate adequately using written messages or emails, and hence do not require pictograms. However, it should be noted that, although this is not generally well known, people who lost their hearing ability before acquiring a language usually find it difficult to learn a spoken language, such as Japanese or English. For example, for those hearing-impaired people who first learned Japanese sign language (JSL), which is a completely different language from Japanese, JSL is their native language and Japanese is a second language. This is why many of them find it difficult to read and write Japanese. This has been confirmed through experiments [3].

For this reason, we think that the use of text is exclusive, so we decided not to use text at all in our pursuit of universal communication. As an alternative, we chose motion pictograms, which are expected to serve as an effective means of assisting the communication of hearing-impaired persons.

2.2 Use of sign-language-like expressions by hearing-impaired persons

As with spoken languages, in sign languages the vocabulary, word order, and grammar differ from one country or region to another. However, it is known that people using different sign languages can begin to make minimal communication in a matter of one or

two hours after meeting for the first time. We believe that this suggests that the means of expression used by hearing-impaired persons, such as sign language expressions, contains some elements that make universal communication possible.

When hearing-impaired persons communicate in a sign language, they use more than their hands and fingers. For example, the direction of gaze and facial expressions (non-manual signals) convey important meanings, mainly providing grammatical meaning, such as indicating a question, imperative statement, or conditional clause. Moreover, people often use non-verbal expressions that cannot be considered part of a sign language, such as gestures and mimicking of certain actions. In this paper, we call all such expressions used by hearing-impaired persons (including nonlinguistic expressions) *sign-language-like expressions*.

We aim to identify elements that enhance universality in communication and apply them to the design of VUTE. Although we are still at a preparatory stage in applying sign-language-like expressions to the design of individual pictograms, we have started to introduce them from a broader perspective, as discussed in section 4.2. Aiming to broaden the application area of VUTE in the future, we are analyzing sign-language-like expressions in the belief that they will prove useful as the vocabulary of VUTE grows and pictogram expressions become more and more complex.

2.3 Use of motion pictograms

Most existing AACs assume the use of paper, and thus use static pictograms. In contrast, VUTE assumes the use of personal digital assistants (PDAs), so it can use motion pictograms. Motion pictograms enable the viewer to understand movements at a glance, so they can convey the concept of a verb or other function words—concepts that are difficult to express with static pictograms. This raises the viewer's level of understanding. Moreover, since motion pictograms can show movement, they raise the level of understanding of sign-language-like expressions.

We introduced manga-like expressions into the design of motion pictograms. This is a sharp contrast to the very simple, abstract, highly designed styles are used in the design of ordinary pictograms. Excessively simple pictograms are expected to have difficulty expressing various things and events in the real world. Fujisawa et al. made still pictograms that introduced manga-like expression and confirmed their comprehensibility. We introduced manga-like expressions in the design of VUTE motion picto-

grams in the expectation that manga-like pictograms would contain rich information and could express abstract feelings or complicated body movements.

3. Design of a universal communication aid for emergency situations

As the first step in developing a communication aid based on motion pictograms, we chose to focus on the navigation of an emergency conversation and developed VUTE 2009, a prototype that can be used for both evaluation and demonstration. Since VUTE 2009 was built as a Flash program, it can be used on any terminal that supports a Flash-compatible Internet browser. The main features of the design of VUTE 2009 and its process flow are described below.

3.1 Analysis of emergency conversation

In this paper, we define an *emergency conversation* as one used in the event of an emergency by an involved person or his/her family member (hereinafter, *rescue seeker*) to call the emergency number or an emergency response center (hereinafter, center) to seek help. We also define *emergency conversation navigation* as the process of supporting an emergency conversation so that the rescue seeker can achieve his/her objective.

To design the emergency conversation navigation flow, we sought the cooperation of Kasuga, Onojo, and Nakagawa Fire Departments in Fukuoka Prefecture and investigated their procedures for emergency voice conversations. By examining their Voice Response Manual (not disclosed to the public), we extracted typical conversation patterns, identified the minimum necessary information that the center needs to obtain when they are contacted by a rescue seeker, and studied the most efficient sequence in which different items of information can be sought from the rescue seeker. As a result, we were able to obtain the minimum necessary information even if the emergency conversation is fixed to a pattern in which the rescue seeker simply answers questions posed by the center. We also found that in most cases the main information can be obtained in up to ten Q&A (questions and answers) sets.

Of course, a conversation would be more natural if both sides were able to pose questions. For example, a rescue seeker might ask "Is there anything I should do while waiting for the ambulance?" or "How soon will the ambulance arrive?" If the rescue seeker were to ask these questions, he/she would be able to treat the injured person appropriately and not have to wait
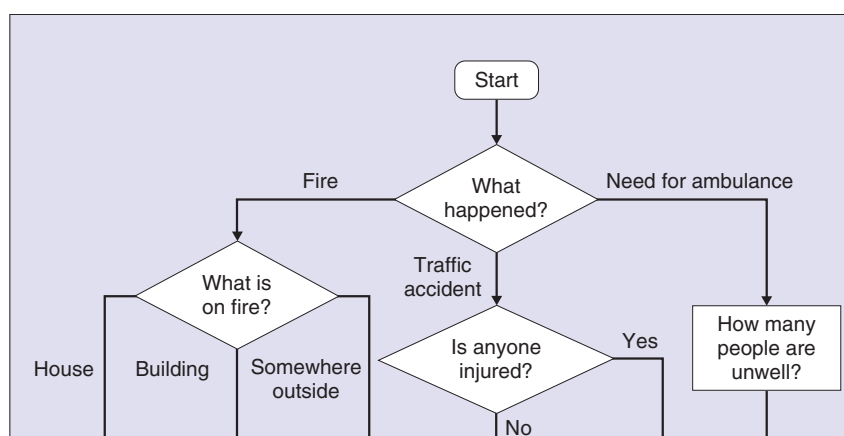
Fig. 2.   Part of a flowchart of a dialog for requesting emergency services.

impatiently. However, if we considered cases where the rescue seeker is allowed to ask questions, then the number of conversation flow variations would increase so dramatically that it would be impossible to identify a manageable number of conversational patterns. Besides, it would be extremely difficult to express such questions and answers with pictograms.

For these reasons, in the prototype system, only the center asks questions.

3.2   Development of emergency conversation flowcharts and multiple-choice question standardization

To standardize the emergency conversation procedures, we described emergency conversations in flowcharts. We simplified the procedures to reduce the number of exchanges in a conversation. In addition, we limited the question types to multiple-choice questions, in which the user selects one answer from a set of options provided. Part of one of the developed flowcharts is shown in **Fig. 2**. The flowcharts of emergency conversations implemented in VUTE 2009 consist of *branching nodes*, where the flow branches depending on the answer, *non-branching nodes*, and links connecting these nodes.

3.3   Expressions used in multiple-choice questions

Each conversation used in VUTE 2009 is a series of multiple-choice questions, posed and answered. Each answer option is expressed by a pictogram. The user answers questions by touching the appropriate option. Therefore, in the course of operating the system the user must understand that he/she is expected to select one of the multiple answer options. A cartoon-like character in the center of the screen, called the agent, plays an important role in helping the user to understand what he/she is expected to do.

How do hearing-impaired persons express multiple-choice questions? When they want the other person to select one of several options, they associate each option with a finger. They touch a finger with the other hand and explain the option associated with that finger. For example, if they want to ask whether the other person wants Chinese, Japanese, or Italian food, they point at the index, middle, and ring fingers in sequence and associate one type of food with one of these fingers. By associating an option with a finger, they detach the options from the conversing people and treat them as something independent of the people concerned. These gestures enable the viewer of the sign language to understand that he/she is expected to choose one of several alternatives.

Similarly, in VUTE 2009, an agent is placed in the center of the screen as the talker, and answer options are placed around the agent. The agent points at options one by one and tilts his/her head to indicate that he/she is asking "which one?"

3.4   Processing flow of VUTE 2009
1)   Entry of the user name and location

First, the user (rescue seeker) enters his/her name and location. (In a practical implementation, user names are pre-registered and the user's current location can be set automatically using GPS (global positioning system).
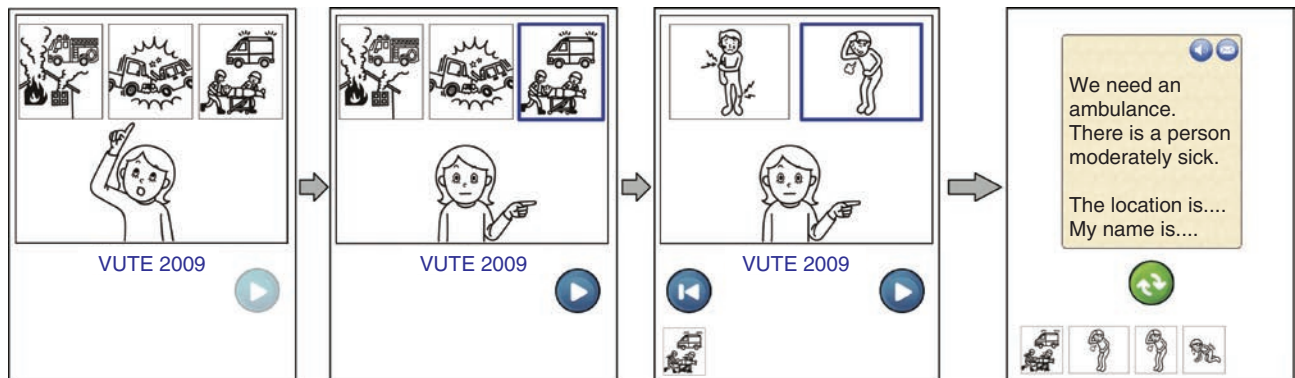
2)   Q&A procedure

Fig. 3.   Screenshots of VUTE 2009 conversation process.

In accordance with the emergency conversation flowcharts developed in section 3.3, the system presents questions and the user answers them all. The agent is displayed at the center of the screen, and a number of pictograms representing answer options are displayed at the top of the screen. The agent points at the options one by one. This is to indicate that the user is expected to select one option. Each motion pictogram consists of just two or three frames, which are repeated.

The user selects one pictogram by clicking or touching it. Depending on the selected answer, the system selects the next question and presents it. Thumbnails of the pictograms selected by the user are displayed at the bottom of the screen in their selection order.

3)   Output of results

When all the questions have been asked and answered, text in the local language is generated on the basis of the information obtained from the Q&A session. The text can either be displayed or read aloud as appropriate.

The above conversation procedures are shown in **Fig. 3**.

4.   Experiments to evaluate the effectiveness of VUTE 2009

We carried out an evaluation test by asking a number of subjects to operate VUTE 2009 to determine whether the system does obtain the minimum necessary information from an emergency conversation. The task achievement times with VUTE 2009 were compared with achievement times with a text-based navigation system.

4.1   Experimental method

For comparison, we developed text-based navigation systems, one using the Japanese language and the other using the English language. In these systems, each VUTE 2009 pictogram was replaced by a Japanese or English sentence.

The mission of the subjects was to seek rescue by operating, one by one, the three navigation systems: VUTE 2009, the Japanese text-based system, and the English text-based system. After each experiment, they were asked to provide opinion scores, and the time it took to complete each task was recorded.

The three navigation systems were given to subjects in random order. This was done twice so that each subject conducted a total of six experiments. In each experiment, a specific situation was defined, such as whether the problem was an accident, sudden illness, or fire; how serious the state of the victims was; and how many people were involved. Sentences describing the situation were also created.

(1)   Experimental system. The English and Japanese language programs were written using Super-Card. An example of a user's answers in the English system is shown in **Fig. 4**. These programs ran on a MacBook G4. Since the user interfaces used by VUTE 2009 and those used by the text-based navigation systems were different, differences in their operability might affect the subjects' evaluation. To hide such differences, we asked the subjects only to point at the pictogram they had selected and a separate operator actually operated the personal computer running VUTE 2009.

(2)   Subjects. The subjects were 13 Japanese people without hearing disability, 8 hearing-impaired Japanese, and 5 non-Japanese without hearing disability.

Details are given in **Table 1**. We first asked the subjects to rate their own reading and writing ability in English and Japanese on a scale from 1 (cannot read or write the language at all) to 5 (native tongue) in increments of 0.5.

(3) Scenario presentation. We developed six scenarios in Japanese and English. Each scenario was printed on a sheet of A4-size paper. Before the start of an experiment, the subject was given a randomly selected scenario description and asked to make himself/herself fully familiar with the scenario. An example of a scenario description is shown in **Fig. 5**.

(4) Achievement time measurement. Once the subject had understood the situation, he/she was asked to operate each navigation system without any explanation about how to use it. The period from the time when the subject started the operation to the time when he/she gave an answer to the last question was defined as the achievement time. In the case of VUTE 2009, the number of answers ranged from three to five, and the achievement time varied accordingly. In addition, some subjects did not give the expected answers, which resulted in a number of answers being different from what was expected. Considering that such variations may be balanced out by taking the average, we simply used the average achievement time.

(5) Intelligibility assessment. After each experiment, the subject was asked to assess how much he/she understood the sentences or pictograms presented during the experiment (intelligibility) on a scale from 1 (did not understand them at all) to 5 (completely understood them).
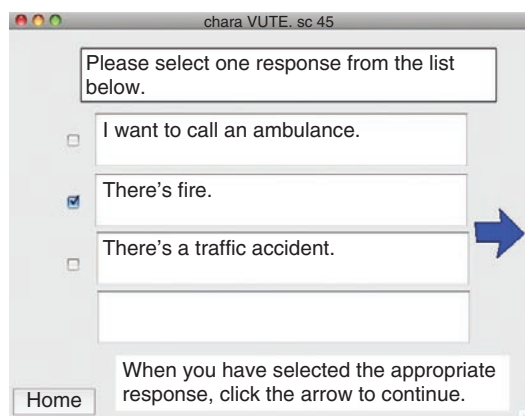

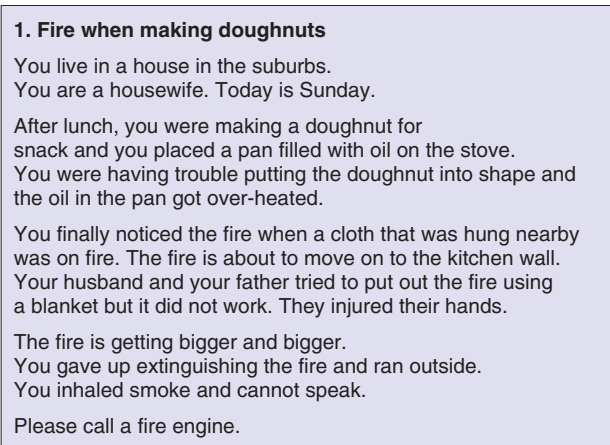Fig. 4. Example of English emergency conversation guidance during the experiment.

Fig. 5. Example of English scenario description used in the experiment.

Table 1. Breakdown of subjects.

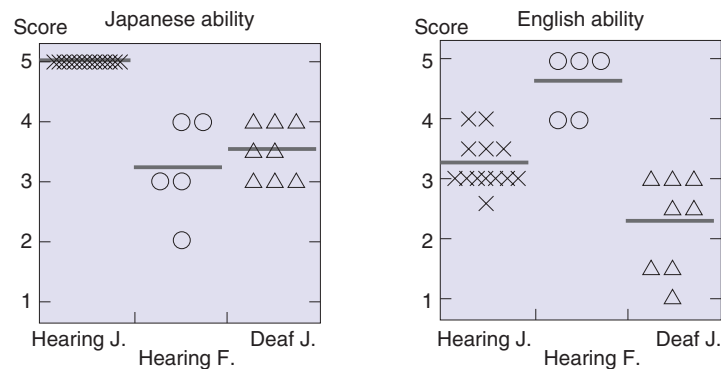|		Hearing Japanese	Hearing foreigner	Japanese hearing-impaired
Gender	Male	10	3	4
	Female	3	2	4
Age	20s	8	4	4
	30s	2	1	1
	40s	3	0	1
	50s	0	0	0
	60s	0	0	2
Nationality	Japanese	13	0	8
	English or American	0	3	0
	Thai	0	2	0

Fig. 6. Self-estimation of reading and writing ability in Japanese and English.
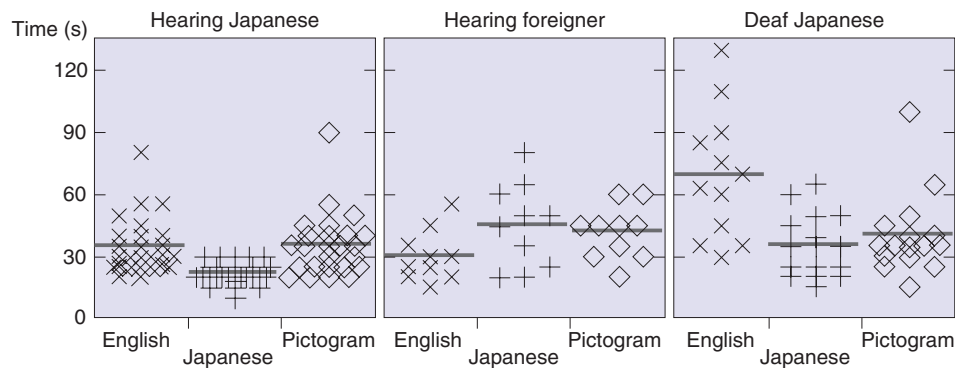

Fig. 7. Distribution of achievement times for different subject groups and data types.

4.2 Experimental results

The subjects' self-assessments of their Japanese and English abilities are shown on scatter diagrams in **Fig. 6**. The score of a Japanese without hearing disability (Hearing J.) is indicated by ×, that of a foreigner without hearing disability (Hearing F.) by ○, and that of a Japanese with a hearing disability (Deaf J.) by Δ. The average of each subject group is shown by a horizontal bar.

In the experiments with pictogram navigations, we confirmed, by watching the experiments, that all subjects were able to proceed to the last question and provide appropriate outputs. Therefore, the percentage of subjects accomplishing their given mission was 100% and the validity of pictograms used in VUTE 2009 was confirmed.

The achievement time distributions are shown in scatter diagrams in **Fig. 7** for each subject group. The achievement time for an experiment with English navigation (English) is indicated by ×, that for one with Japanese navigation (Japanese) by +, and that for one with motion pictograms (Pictogram) by ◇. The average of each experiment group is shown by a horizontal bar. Similarly, the intelligibility levels are shown in **Fig. 8**.

Furthermore, the distributions of the difference in achievement times between Japanese and English text navigation and motion pictogram navigation are shown in scatter diagrams in **Fig. 9** for each subject group.

From Fig. 7, we determined the following. When the Japanese and non-Japanese without hearing disability used motion pictograms, their achievement times for motion pictograms were equivalent to those for their foreign language. For the Japanese with a hearing disability, their achievement times for motion pictograms were equivalent to those when using Japanese texts. However, when they used English (their foreign language), their achievement times were longer than for Japanese or pictograms. These
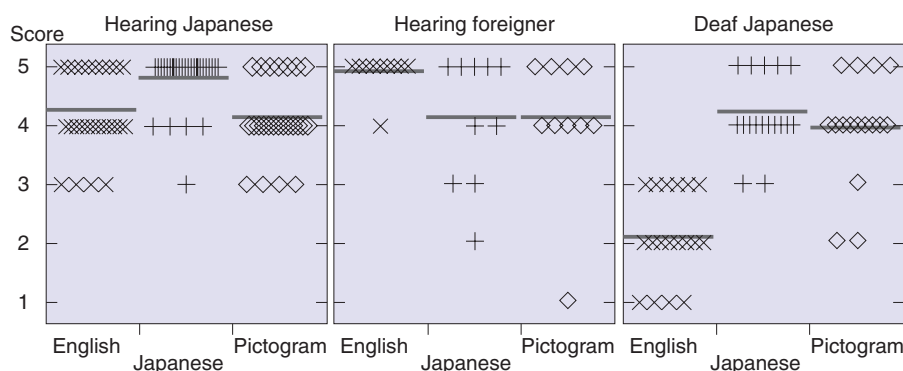
Score

Fig. 8.   Distribution of intelligibility scores for different subject groups and data types.
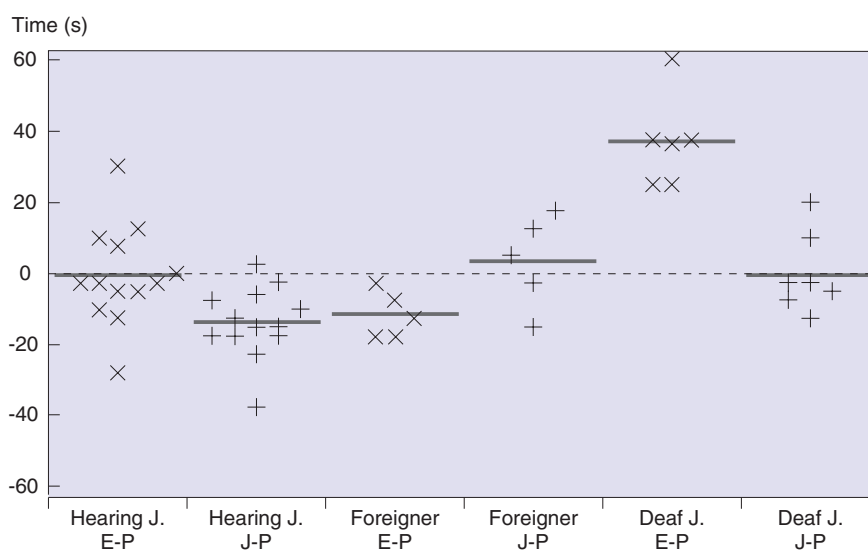

Fig. 9.   Differences in achievement times for English text, Japanese text, and pictograms.

findings became clear when we inspected Fig. 8 because, if the marks (× or +) are plotted around the zero line (dotted line), the achievement times for English or Japanese navigation are equal to those for pictogram navigation and if they are plotted over the zero line, they are longer than for pictogram navigation, which means that text-based navigation is more time-consuming than pictogram-based navigation.

A similar inspection of Fig. 8 revealed the intelligibility levels.

4.3   Discussion

The above experiment results show that it is almost certain that the communication aid using motion pic-

tograms enables the user to accomplish the goal of emergency conversation. In addition, it enables people without any hearing disability to accomplish emergency conversation within almost the same length of time as the case where they deal with a foreign language that they are not particularly good at.

The experiments confirmed some issues that were only abstractedly guessed beforehand. They are summarized below. They will be solved through future experiments.

(1)   Explanation of defined situation

In these experiments, the predefined situation, such as a disaster or illness, was described in Japanese or English text. As we expand the types of people acting

as subjects, some may not be able to understand the situation well if the explanation is given only in Japanese or English. For Japanese with a hearing disability, a description of the situation in Japanese is not necessarily the ideal means of explanation. It is necessary to study methods of explanation that do not use written characters but use pictures or cartoons.

(2) Degree of mission accomplishment

In these experiments, we did not examine the degree of mission accomplishment precisely because there was not a single absolutely correct answer in these experiments, so it was difficult to assess the degree of accomplishment objectively and numerically. For example, in the case of a traffic accident, the original mission was accomplished even if the user selected a pictogram of an ambulance instead of a traffic accident. When a user needs to indicate how serious an injury is, he/she may not be able to tell whether the injury is slight or somewhat more serious. It is necessary to study whether we can improve the evaluation with regard to these aspects.

5. Conclusions

This paper presented the basic concept and design policy for VUTE, a universal communication aid based on motion pictograms. It gave an overview of the functions of VUTE 2009, a prototype communication support system we have developed. VUTE 2009 focuses on emergency conversations. We carried out preliminary experiments and confirmed that the minimum information necessary for the dispatch of a rescue team can be conveyed by motion pictograms. Although motion pictograms are less efficient in conveying information than using the native tongue, they promise to be useful in situations in which the user does not understand the language spoken at his/her location.

Using the experimental results, we will conduct more precise experiments with a wider range (in age, nationality, severity of the disability, etc.) of subjects. Later full-fledged experiments, examining the subjects' actions closely (for example, measuring the time required to choose appropriate answers for alternative pictogram sets), are expected to lead to improved designs of pictograms and user interfaces.

We aim to develop the system into a communication aid that is closer to natural conversation and applicable to a wider range of purposes. Specifically, we will apply it to conversations taking place in railway stations, where there are likely to be travelers who cannot understand the local language. For this purpose, we will continue to analyze sign languages and study how to implement the system on PDAs and mobile phones.

VUTE 2009 is currently available on the Internet [4]. Anyone can access this site to experiment with the motion-pictogram-based communication aid.

References

[1] K. Fujisawa and T. Inoue, "Young Children's Comprehension of the Visual Symbols of PIC-J (Japanese Version of Pictogram Ideogram Communication)," Proc. ISAAC 1998, Taejon, Korea.
[2] T. Kojima, "Augmentative and Alternative Communication (AAC) for Individuals with Severe Language Disabilities: Topics from Recent AAC Practices and Research," Technical Report of IEICE, Vol. 101, No. 36, pp. 27–34, 2001 (in Japanese).
[3] K. Nakazono and S. Kim, "Availability of Textual Information for Hearing Impaired People," Human Interface, Vol. 10, No. 4, pp. 395–402, 2008 (in Japanese).
[4] http://vute.ilab.ntt.co.jp/vute (in Japanese).

Kaoru Nakazono

Senior Research Engineer, NTT Network Innovation Laboratories and Research Professor at Kyoto Institute of Technology.

He received the B.E. and M.E. degrees in biological engineering from Osaka University in 1980 and 1982, respectively, and the Ph.D. degree in engineering from Chiba University in 2006. He joined the Electrical Communication Laboratories of Nippon Telegraph and Telephone Public Corporation (now NTT) in 1982. From 1990 to 1993, he worked for the Advanced Telecommunications Research Institute International (ATR). His research interests include assistive technology, human cognition and communication, and basic computer science. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE) of Japan and the Human Interface Society of Japan (HISJ).

Mari Kakuta

Ph.D. student and teaching assistant at International Christian University and part-time lecturer at Yokohama National University.

Yuji Nagashima

Professor, Kogakuin University.

He received the B.E., M.E., and Ph.D. degrees in electronic engineering from Kogakuin University, Tokyo, in 1978, 1980, and 1993, respectively. He joined the Faculty of Engineering at Kogakuin University as an assistant and then became a lecturer and then an assistant professor. He has been a professor there since 2003. His research interests include human interface technology, sign linguistics engineering, assistive and rehabilitation engineering, and developmental disorders. He is a member of IEICE, HISJ, IEEE, and the Association for Computing Machinery.

Naotsune Hosono

Senior Managing Consultant, Oki Consultant Solutions Co., Ltd., Lecturer at Seikei University, Interdisciplinary Researcher at Keio University.

He received the M.Sc. degree from Keio University, Kanagawa, in 1974, the M.Sc. degree from the University of Manchester, UK, in 1981, and the Ph.D. degree from Keio University in 2003. He is a member of the Academic Society, HISJ, and the Human Centered Design Organization (HCD-Net) and a Fellow of the Japan Ergonomics Society. He is a Certified Professional Ergonomist. He is an author of several books in Japanese: "Usability Testing" (Kyoritsu Shuppan Co., Ltd.), "Dictionary of Ergonomics" (Maruzen Co. Ltd.), and "Universal Design with ICT" (Maruzen Co. Ltd.).