LiveTalk: Real-time Information Sharing between Hearing-impaired People and People with Normal Hearing

● Shinichi Ono ● Yasuaki Takamoto ● Yoshiki Matsuda

In recent years, more and more people have been recognizing that communication is very important in pursuing collaborations and innovations. Recent efforts to involve people from a variety of backgrounds in the development process, such as brainstorming, evaluation, and modification, have shown that such efforts can create new user experience (UX). Notably, there are projects where participants include people with certain difficulties. In April 2015, Fujitsu launched a new communication tool, FUJITSU Software LiveTalk, aiming to include hearing-impaired people in the circle of community. Throughout the developmental phases, we collaborated with contributors who were hearing-impaired. In the development of LiveTalk, we observed the participants to identify characteristic behaviors of hearing-impaired people in their workplaces as well as challenges they encounter when communicating with people who can hear normally. Their opinions were also shared with us to help create a prototype equipped with features that addressed them. Through the user evaluation and feedback on this prototype, we repeatedly improved the model to make it easier to use. It is an application that realizes smooth bilateral communications between hearing-impaired people and people with normal hearing, based on a new UX design. This paper explains the development of LiveTalk.

1. Introduction

ICT has become prevalent today. It is increasingly important to take accessibility into account in designing various products, and a growing number of devices and services cater to the needs of people with disabilities. In Japan, the law concerning the minimum ratio of employing people with disabilities was amended in April 2013, and it is now mandatory for private companies and national/regional public bodies to ensure that people with disabilities account for more than 2.0% and 2.3% of employees, respectively. Thus, job opportunities for these people are increasing. In 2016, the Disability Discrimination Act was introduced, which obligated national/regional public agents to make efforts to cater to the needs and requests of people with disabilities. In this climate, it is important that companies who employ people with disabilities implement measures to support them.

Fujitsu has been proactive in the promotion of universal design since the late 1990s. ATMs and smartphones equipped with accessibility features are some examples. We are also making efforts to improve the work environment around those who are hearing- or sight-impaired, trying to increase the future career opportunities for challenged people.

This paper describes our efforts to develop a communication tool that facilitates smooth information-sharing between hearing-impaired people and people with normal hearing, for example, in a meeting. During the development process, we created prototypes based on the user experience (UX), and piloted them with people with hearing difficulties.

2. Communication environment for hearing-impaired people

There are approximately 340,000 hearing-impaired or hard-of-hearing people in Japan. It is vital for them to have support in terms of communication. Some companies employ external interpreters (professional transcribers and/or sign language interpreters) for important occasions such as business presentations.[1] Elsewhere, support is given in forms such as the person sitting next to a hearing-impaired person taking notes so that the main points of the meeting can be shared, or a summary of it, being transcribed with a PC and shared with the person electronically.[2)]

The authors also co-work with hearing-impaired colleagues, with whom information is shared using the PC-based, note-sharing method. However, there were problems; even simple meetings needed a long time for preparation (PC, projector, etc.), note-taking was too slow to keep up with the meeting proceedings, and the note-taker would lose track of the meeting agendas. On the part of the hearing-impaired people, they tended to feel like a burden to the meeting, for all these efforts had to be made for them to access information.

To address these problems, Fujitsu launched a new communication tool in April 2015. FUJITSU Software LiveTalk aims to include hearing-impaired people in the circle of communication while reducing burdens on all parties participating in a meeting.[3)] With LiveTalk, a person's speech is transcribed immediately using a speech recognition engine, and the texts are displayed on multiple computer screens in the meeting room. LiveTalk is installed on each individual computer (**Figure 1**), using the AmiVoice SP2 engine made by Advanced Media, Inc.

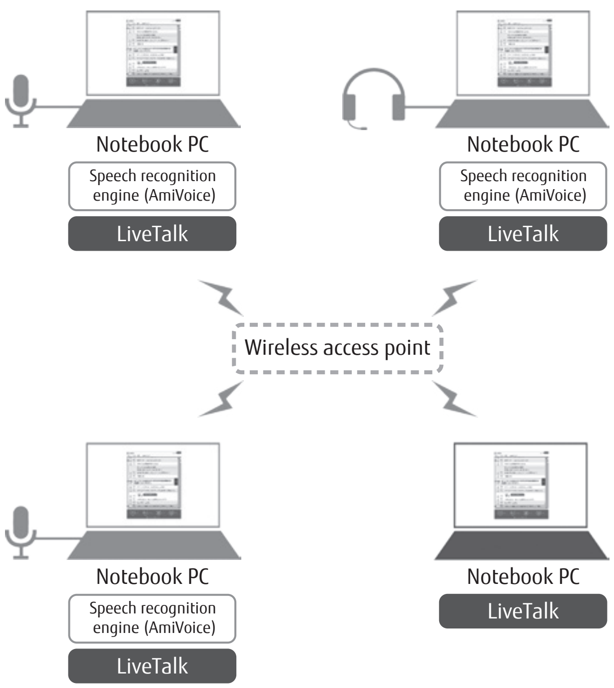


Figure 1
LiveTalk configuration.

We had hearing-impaired people participate in the planning and development phases of LiveTalk. As communication plays an important role in the UX design based on diverse user groups, in the development of Live Talk, we observed the participants to identify characteristic behaviors of hearing-impaired people in their workplaces as well as challenges they encounter when communicating with people who can hear normally. Their opinions showed that they could not keep up with the meeting, they had difficulties in joining in conversations, and they hesitated to say when they did not understand.[4)]

Taking these as a starting point, we set out to achieve real-time communication, comprehensive understanding of multi-directional conversations, and opportunities for people with disabilities to have their voices heard. We then developed a prototype based on the usability for these people, and improved it through trials conducted in real meetings.

3. User evaluations of prototypes and improvements

We developed LiveTalk through having the prototype tested by hearing-impaired people and their colleagues/managers who could hear normally, repeating evaluations and improvements.

3.1 First prototype

The first prototype aimed to develop and evaluate two functions. One was visual communication (**Figure 2**). The PCs of the meeting participants were connected to microphones, which captured the speech of each individual. A speech recognition engine converted the speech into texts. A screen that shows all participants' faces using a 360-degree camera displayed the texts in balloons over the face of the speaker. This feature enables participants to understand who said what in real time. The second function was text-based communication (**Figure 3**), shown as a list of speech texts in chronological order.

Through the pilot tests by users including hearing-impaired people, we identified the following issues.

1) The face images captured by the 360-degree camera were too small to recognize facial expressions.
2) The hearing-impaired people were too busy reading the text to look at the face images. For the same reason, they could not pay attention to the

Visual system using 360-degree camera and array microphone

360-degree camera and array microphone

Figure 2
Visual communication feature.

Figure 3
Text-based communication feature.

surrounding situation.[5)]

3) The hearing-impaired people experienced great visual strain when looking at the computer screen and the surroundings in the meeting room alternately.
4) Typological errors in the auto-transcription were distracting.

3.2 Second prototype

To address the above 1), we developed a system using a revolving camera (**Figure 4**). Installed in the middle of the meeting table, it captures participants' facial expressions, which are combined with each person's speech obtained through the aforementioned text-based communication feature. This system, however, had problems both with the camera and the text-based communication feature. The revolving camera needed to be adjusted in terms of the capture ranges. When multiple people spoke simultaneously, the camera could not rotate fast enough to capture all of them at once.

As for the text-based communication feature, improvements were needed in the readability of the on-screen text, real-time speech recognition, and visibility of texts when several people spoke at the same time. Transcription errors also proved difficult to correct.

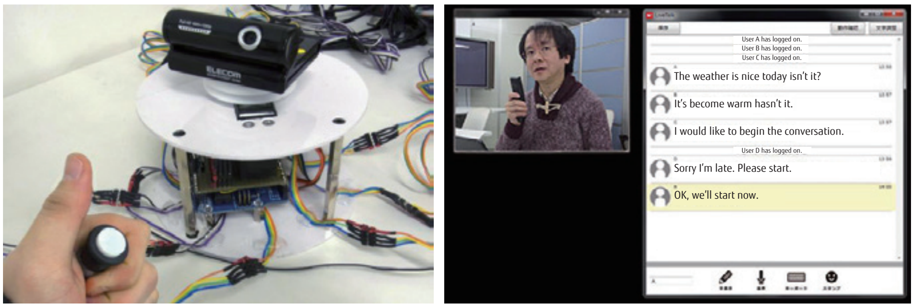
3.3 Third prototype

The first prototype had the aforementioned problem 2). The second model had the camera that could not keep up with the speed of alternating speakers. In addition to these, it has been known that, by looking at the speaker's face and his/her speech in text alternately on the computer screen, users spend much time moving their visual focus, and this interfered with their understanding.[6] To address this problem, we developed a system which could display texts near the image of the speaker by means of a wearable device, using MOVERIO BT-200 by Seiko Epson Corp.

The texts generated through the speech recognition engine are displayed in the view field of the wearable device, together with an indication of the direction from which the speech is made. On the top-right corner of the view field, speakers are represented by white circles that indicate their positions in relation to the device user and the volume of the voice. These circles expand as the respective person speaks [**Figure 5(a)**].

We also developed another system so that, when someone spoke, the side of the view field (top, bottom, left, or right) corresponding to the speaker's direction expanded [**Figure 5(b)**]. If the speaker is in front of the user, the top edge of the view field expands. Similarly, the speaker behind the user is indicated by the bottom side, and the right/left-hand side correspond to the respective side of the field.

People with a hearing impairment put on the wearable device equipped with these features, and evaluated it in terms of the color and size of the texts, their positions in the display area, the number of lines, the indication of the voice directions, and the history display method. The responses were generally good, because they did not have to look everywhere all the time, and they could notice the speech of those who were outside the view field. Apart from this, negative responses included that it was difficult to look at the speaker's face, surroundings, and texts all at once, and the wearable device was heavy. To make this prototype practical, these issues must be addressed.



Revolving camera

Visual and text representation

Figure 4
System with a revolving camera.

(a) The voices and their loudness are represented by circles that are arranged in six directional points

(b) Voices coming from the front and sides of the user are indicated by arches that respond to the loudness of the respective voices

Figure 5
Example of the wearable device's view display (evaluation of the text color and directional indication).

3.4 Characteristic behaviors of hearing-impaired people

Through the evaluations of these prototypes, we obtained some feedback from the hearing-impaired people that was contrary to our initial expectations. For example, some of them felt that it was unnecessary to convert all spoken words into texts using the speech recognition engine, or they felt more comfortable with conventional means such as sign language interpretation and note-taking.

Many hearing-impaired people use communication methods such as sign language, writing, lip-reading, and chat applications. Due to a lack of experience in voiced communications as they happen between people with normal hearing, they are not sufficiently exposed to such occasions as unintended errors, hesitations, grammatical errors, and pauses in speech. Because the speech recognition picks up all these erroneous sounds and converts them into texts, the resulting sentences are often perceived to be more onerous to understand than interpreters or note-sharing, to which people with a hearing impairment are more accustomed.

For this reason, the tool to be developed must complement the conventional information-sharing techniques, or be used selectively according to the situation. It is also important that the cognitive characteristics of hearing-impaired people, which are influenced by personal experiences, are taken into account in the development of the tool.

Meanwhile, with the first prototype, the texts converted from speech were displayed after mistranscription by the speech recognition engine was manually corrected. However, the feedback from the hearing-impaired people revealed that they would like to see the texts in real time, and the corrections could be made later.

Based on this, a new system was installed in the second prototype, so that the texts were displayed immediately when speech was made, before misconversions were corrected. In addition to this real-time display feature, we aimed to realize bilateral communication so that people with a hearing impairment could say, while keeping up with the pace, "I understand," "it's not clear," "wait a moment," and "repeat it, please."

4. Key features of LiveTalk user interface

As described above, we developed three prototypes. While the wearable device used with the third prototype had some advantages, we decided that issues such as the device weight needed more time to be addressed. Therefore, we abandoned this prototype for the current development.

We continued the development of LiveTalk based on the features of the second prototype and by enhancing them. In the following, we will describe the key features of LiveTalk.

1) Real-time display of speech

We made the data transfer of the speech recognition engine more efficient so that the text was output without much delay. This helps hearing-impaired people to understand conversations in a meeting in real time, and gives them more opportunities to express their opinions during the meeting.

2) Display of simultaneous speech

Considering a situation where several people may speak at the same time in a meeting, we have made improvements in terms of the visibility of each portion of speech. This makes it easier to manage unexpected remarks.

3) Bilateral communication

With the addition of stickers and fixed-phrases to the keyboard input, hearing-impaired people have more support in expressing themselves. This improvement makes communication between people with normal hearing and hearing-impaired people more fluent.

4) Correction of errors in transcriptions

It is inevitable that the speech recognition engine involves some transcription errors. LiveTalk is thus equipped with a feature that allows any participant to correct transcription errors at any time using their connected computers. This feature has helped to reduce the distraction to hearing-impaired participants due to misconversions.

5) Saving conversation logs

A new feature has been added for saving the texts (before and after text moderation) as conversation logs. These data will help the users to create bullet notes and meeting minutes, thus also realizing some benefits to participants with normal hearing.

Today, LiveTalk with these features is used in meetings where participants include people with a hearing impairment. We are anticipating further improvements to the tool's features and usability as well as coordination with other applications in the future.

5. Conclusion

In this paper, we described the development of a communication tool based on UX design, facilitating information-sharing with hearing-impaired people by converting conversations into texts in real time.

In the future, we will work to improve and expand the user interface and features, adapting LiveTalk to various situations. We would like to develop this tool that leverages visual communication as a technology not only for people with disabilities, but also for everyone.

This project "Meeting Support System for Hearing-impaired People" was conducted with the subsidy of technical research for the resolution of the digital divide (FY2013–FY2015), provided by the Ministry of Internal Affairs and Communications.

We would like to thank Prof. Nobuko Kato, Assistant Prof. Daisuke Wakatsuki, and Assistant Prof. Shigeki Miyoshi of the Tsukuba University of Technology, and Prof. Tadashi Uesugi of the Kawasaki Municipal School for the Deaf, for their invaluable advice. We are also indebted to all trial participants for their willing cooperation and feedback.

References

1) M. Iwayama: Observations on the comprehensive framework of initiatives to help hearing-impaired people to have lasting careers. Journal of the doctorate studies in social sciences, Vol. 10, pp. 1–24, 2013 (in Japanese).
2) E. Mizuno: Corporate attitude to employment of people with disabilities. Life Design Report, No. 178, pp. 24–31, 2007 (in Japanese).
3) Fujitsu: Fujitsu announces LiveTalk, a participatory communication tool for hearing-impaired people (in Japanese).
http://www.fujitsu.com/jp/group/ssl/resources/news/press-releases/2015/0414.html
4) E. Mizuno: Workplace communication for hearing-impaired people—Present states and challenges identified from surveys to people with hearing impairment and corporations. Life Design Report, No. 182, pp. 4–15, 2007 (in Japanese).
5) T. Suzuki: Study on deaf and hard-of-hearing students with their design practice support. Tsukuba University of Technology Techno Report, Vol. 18, No. 2, pp. 68–72, 2011 (in Japanese).
6) H. Kuroki, et. al.: Research on the real-time captioning system for the deaf and hard-of-hearing regarding the speakers' facial images and the timing of displaying texts with misrecognized words. Journal of Visual Information Media, Vol. 65, No. 12, pp. 1750–1757, 2011 (in Japanese).

Shinichi Ono
Fujitsu Ltd.
Mr. Ono is currently engaged in research and development in the fields of universal design and accessibility.

Yasuaki Takamoto
Fujitsu Ltd.
Mr. Takamoto is currently engaged in research and development in the fields of universal design and accessibility.

Yoshiki Matsuda
Fujitsu Ltd.
Mr. Matsuda is currently engaged in research and development in the fields of universal design and accessibility.

