See-Through Captions in a Museum Guided Tour:

Exploring Museum Guided Tour for Deaf and Hard-of-Hearing People with Real-Time Captioning on Transparent Display

Ippei Suzuki*, Kenta Yamamoto*, Akihisa Shitara, Ryosuke Hyakuta, Ryo Iijima, Yoichi Ochiai

Research and Development Center for Digital Nature, University of Tsukuba, Japan

*These authors contributed equally to this research.
Background & Introduction
How can we update the accessibility of museum guided tours for deaf and hard-of-hearing people?
Background Accessibility of Museum Guided Tour for DHH People

Approaches to accessibility of audible information

Sign-language guided tours

Auditory information via mobile device
Background
Accessibility of Museum Guided Tour for DHH People

Sign-language guided tours
Difficult to recruit an interpreter

Auditory information via mobile device
One-way information
Cannot communicate with guide

Image (Left): Namatame et al. 2020. The Science Communication Tour with a Sign Language Interpreter
Automatic Speech Recognition (ASR)
Background  Automatic Speech Recognition (ASR)

Approaches to utilize automatic speech recognition

ASR on mobile devices  ASR on augmented reality devices
Background  Automatic Speech Recognition (ASR)

ASR on mobile devices

The facial expression and body language of the partner are overlooked

ASR on augmented reality devices

Speaker cannot confirm whether the speech has been correctly recognized.
Introduction  Our Previous Work

See-Through Captions: Real-Time Captioning on Transparent Display for Deaf and Hard-of-Hearing People.
Introduction  Transparent Display
Implementation  Transparent Display
**Implementation**

Transparent Display

---

<table>
<thead>
<tr>
<th>Specification</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Resolution</strong></td>
<td>320 × 360 pixels</td>
</tr>
<tr>
<td><strong>Number of Colors</strong></td>
<td>4,096 Colors</td>
</tr>
<tr>
<td><strong>Transmittance</strong></td>
<td>87%</td>
</tr>
<tr>
<td><strong>Weight</strong></td>
<td>Approx. 130 g</td>
</tr>
<tr>
<td><strong>Brightness (Center)</strong></td>
<td>270 cd/m²</td>
</tr>
<tr>
<td><strong>Contrast Ratio</strong></td>
<td>20:1</td>
</tr>
</tbody>
</table>

Japan Display Inc.
Implementation  Microphone

Shure; WH20XLR
Headset Microphone

Unidirectional cardioid directivity

Less surrounding noise
Implementation Backpack

- **Audio Interface**
- **Computer (Tablet PC)** *Inserted into backpack*
- **Battery**
- **Display Drive Board**
- **Mobile Wi-Fi Hotspot**

Weight: approx. 3.3 kg
Implementation Automatic Speech Recognition API

Web Speech API on Google Chrome

Image (Google Chrome Logo): https://www.google.com/chrome/
Implementation Overview

- Display
- Microphone

- Audio I/F
- Computer
- Drive Board
- Battery
- Mobile Wi-Fi Hotspot
Case Study: Guided Tour in Museum
Case Study: Guided Tour in Museum

Miraikan
THE NATIONAL MUSEUM OF EMERGING SCIENCE AND INNOVATION

Science Communicators

Bunsuke Kawasaki  Sakiko Tanaka

Image (Right): https://www.miraikan.jst.go.jp/aboutus/
Image (Science Communicators): https://www.miraikan.jst.go.jp/en/aboutus/communicators/
Case Study: Guided Tour in Museum

Tour theme: “The difference between humans and robots”
Case Study: Guided Tour in Museum

Communication Method
Tours were conducted in Japanese language

Guide Person ➔ DHH People
See-Through Captions

DHH People ➔ Guide Person
Speech or Writing
When ASR system stopped...

Guide express “wait” in gestures of sign language
Communication Protocol

When participants wanted to talk
They raise their hand or notepad

When someone talked one’s idea
“Applause” in gestures of sing language
Procedure

1. Participants were asked about the preferred position of display and asked about preferred infection-prevention methods (face shield or face mask)

2. The guide described the theme of the tour and conducted some quiz games about Miraikan

3. Guided tour

4. Participants were asked to fill out the questionnaires and be interviewed
Case Study: Guided Tour in Museum

Display Position: Basic
Case Study: Guided Tour in Museum

Display Position: Overlay
Case Study: Guided Tour in Museum

Display Position: Hands-Free
Case Study: Guided Tour in Museum

Display Positions
Case Study: Guided Tour in Museum

Participants

11 **DHH Participants** | 18-53 years old

4 **Hearing Participants** | 36-56 years old

+ 1 Hearing Participant without questionnaires

Tour Groups

9 Groups

Each tour group contained at least one DHH person; some groups contained a few hearing people.
Results & Discussion
## Results  Quantitative Evaluation

<table>
<thead>
<tr>
<th>Question</th>
<th>Group</th>
<th>Mean (M)</th>
<th>Standard Deviation (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1. Readability of the ASR results</td>
<td>DHH</td>
<td>4.45</td>
<td>.50</td>
</tr>
<tr>
<td></td>
<td>Hearing</td>
<td>4.00</td>
<td>.63</td>
</tr>
<tr>
<td>Q2. Noticeability of misrecognition</td>
<td>DHH</td>
<td>4.27</td>
<td>.86</td>
</tr>
<tr>
<td></td>
<td>Hearing</td>
<td>2.80</td>
<td>.75</td>
</tr>
<tr>
<td>Q3. Whether they wanted to continue utilizing this system</td>
<td>DHH</td>
<td>4.73</td>
<td>.45</td>
</tr>
<tr>
<td></td>
<td>Hearing</td>
<td>4.20</td>
<td>.40</td>
</tr>
</tbody>
</table>
ASR sometimes misrecognize the words

It was difficult to read when misrecognition occurred

Possible solution:
The speaker acquire utterances and speaking styles that were easy for the system to recognize correctly

Dictionary registration for technical terms / nouns
Results & Discussion  Readability of Captions

The readability affected by background and reflection

Difficult to see in some settings especially when there is a strong light in the background

Possible solution:
The guide pays attention to that Easily text design changeable system
Results & Discussion  How to Display Captions

Subtitle design is for a larger transparent display

The character flow was too fast

The screen was filled with rephrasing when misrecognition occurred

Possible solution:
Function to look back at the history
Little larger transparent display
Results & Discussion  Benefits of Transparency

Participants could see the subtitles while looking at the contents of the exhibition.

It was easy to communicate in both directions by being able to see the guide’s face and make eye contact.

Transparency made it possible to see the whole without obstructing the view, and that they did not feel any gap.
Results & Discussion  Display Position

Handheld setup makes us easy to change the position

We asked participants which position is preferred

“If the display is held near the face, it is easier because there is only one place to watch.”
Results & Discussion  Display Type and Size

As a future work, it is necessary to compare See-Through Captions with other methods in detail.

Participants mentioned:
- AR glasses was tiring but See-Through Captions was easier
- Display size was small

Example of other methods: Two-sided tablet
Results & Discussion  Challenges Specific to Guided Tours

See-Through Captions was originally developed as a 1:1 communication

When multiple people participated, their voice is NOT displayed

Possible solution:
Participants also wear microphones
Participants also hold displays
Future Work  How Can DHH People Communicate with Tour Guide?

The current system assumes that DHH people speak using voice

Some DHH people do not tend to speak by their voice

Possible solution: Additional input interface?
Future Work  How Can We Communicate with DHH People Who Prefer Sign Language than Text?

The current system assumes that DHH people read texts

Some DHH people prefer to read sign language

Possible solution: Text <-> Sign language Translator?
Summary of Contributions

1. Implementation of the smaller version of See-Through Captions
2. Case study: a guided tour in a museum
3. Discussion of findings based on the results
See-Through Captions in a Museum Guided Tour:
Exploring Museum Guided Tour for Deaf and Hard-of-Hearing People with Real-Time Captioning on Transparent Display

Research and Development Center for Digital Nature, University of Tsukuba, Japan
Acknowledgements

Transparent display provided by

User study support and assisted by

Miraikan
THE NATIONAL MUSEUM OF EMERGING SCIENCE AND INNOVATION
Bunsuke Kawasaki, Sakiko Tanaka, Chisa Mitsuhashi

This work was supported by

JST CREST Grant Number JPMJCR19F2, including the AIP Challenge Program, Japan.