Practical Expectations of AI as Influenced by the Experience of Emotional Expression Technology Added to Synthetic Speech in Visually Impaired People

AKIRA KIMURA[1)]

Abstract: To clarify the effect on the change of expectation of AI through a series of experiments in which visually impaired people observe and comment on the emotions given to the synthetic voice.
Subjects: A group of visually impaired people who belong to the Gunma Prefecture Association for the Visually Impaired and who are able to travel to public facilities.
Methodology Research design: Cross-sectional survey and single-group experiment Participants were recruited by inviting members of the association to participate in a lecture meeting on the use of AI. expectations of practical use before and after listening to AI-enhanced synthetic voice (AI Inc.) were measured as outcomes. In the meantime, price changes were shown using a refrigerator as an example, and the effects of economic factors on the change in expectation were compared and analyzed.
Results Twenty-two participants, ranging in age from 27 to 80 years, were blind and lived in Gunma Prefecture, Japan. 12 (54%, 95% CI: 33%-75%, ns) responded yes to the expectations of AI, both before the intervention and with the ability to express emotions in a synthetic voice. 16 (72%, 95% CI: 54%-91%, $p<0.001$) responded no to the expectations after the AI experience. 54%-91%, $p<.05$), indicating a significant effect. Expectancy due to economic factors was 20 (90.9%, 95% CI: 78%-100%, $p<.05$).
Conclusion: The observation experiment of emotions attached to synthetic speech in the visually impaired showed that the expectation of AI is positive after examining the data of actual experience of emotional expression and economic estimation. These experiential opportunities are useful as a communication method for maintaining expectations.

Keywords: Practical Expectations, Experience of Emotional Expression Technology, Synthetic Speech, Visually Impaired People

1. Introduction

Information assurance is important for visually impaired people living in suburban areas in Japan. Research on information accessibility to contribute to information assurance is essential from the perspective of information assurance[1,2)]. This research activity requires the development of advanced technology and, at the same time, clinical trials of the users of that technology [3)]. In particular, Tachi et al.'s research on guide dog robots exists, but the protocols for subsequent clinical trials and the ongoing system of research on social support are not yet fully in place.[4)5)]. In the past, the development of behavioral assistive devices for the blind has also lagged behind clinical research in people with other disabilities, with no known clinical trial protocols or ethical issues to consider. One of the main reasons for this is that a visually impaired approach is also required in the so-called recruitment method, which is to invite the visually impaired subjects to participate in the research or study. In addition, it is important to cultivate the willingness of the people concerned to participate in the research, not as a matter of how to obtain their consent to participate in a single clinical trial, but as an active attitude and feeling that they are continuously playing a part in the research. However, there are few research reports on such issues in Japan.

Obtaining consent for participation in a clinical trial is not simply a matter of explaining the purpose of the trial and requesting the individual to suspend his or her participation for some reason.

Although it has been pointed out that the culture of participating in clinical trials has not been fully formed even among sighted people in Japan, it does not have to be the same for information accessibility, and there are many visually impaired people who expect improvements in real life through technological progress and implementation.

We have been developing information on safety support for mobility accessibility and have conducted many clinical trials, and the needs we have heard include expectations for guide dog robot-like technology such as AI suitcases, and a high level of interest in computer vision-based visual substitution devices that also use AI[2)].

However, their high expectations for these may give them a sense of disappointment if their individual expectation threshold is not exceeded the moment, they experience the actual technology. This situation, which we have named real-life experience shock, can hinder the development of assistive devices for the visually impaired developed from the perspective of people with clear vision, and hinders clinical trials during the development of useful technological products.

By introducing the health communication technique of step-by-step explanation of the gap between the expectation and the post-experience sensation as the cause of live-experience shock, we designed this study to confirm the effect of live-experience shock, so that future clinical trials and social implementation of advanced technologies can be conducted within a planned period of time. In this study, we will explain what AI is, and as an outcome, we will report on the effects of prior expectations on the gap between the actual experience and the post-experience experience by letting the participants experience emotional expressions on a synthetic voice. The purpose of this study was to determine whether the positive effects of real-life shocks emerge immediately or have a latency effect in visually impaired people who have participated in advanced technology trials living in suburban Japan.

---

1)Graduate School of Health Sciences, Guma Paz University, Gunma 370-0006.Japan.

2. Subjects and Methods

The subjects were blind people who belonged to the Gunma Prefecture Association for the Visually Impaired, and who attended a training session titled "AI and IT-based daily life support for the visually impaired" organized by the association, and whose consent was obtained before the session. A cross-sectional survey and a single-group experiment were conducted. The period of the study was conducted on January 28, 2021. The time was 60 minutes and was conducted with consent as part of a hands-on training session using artificial synthetic speech.

Method: I spoke to the group and said, "Let's enjoy the experiment using AIVOICE, and thank you for taking time out of your busy schedule to participate in the training session. Thank you for taking time out of your busy schedule to participate in our training session. 1.

I'm going to change the speed, so please raise your hand if you understand.

Which is your feeling, AI? I will ask you to listen and compare once and then answer.

Here I asked the following: 2-1.

2-1. Let's do a questionnaire and quiz first as a baseline for listening to the expectation of emotional expression. Do you believe that artificial voices have emotions? If you think yes, raise your hand. Now, those who think no, raise your hand. In this case, the artificial voice was set to standard speed, pitch 1.2 times, intonation 0.1, and the emotions to be given were anger 0, sadness 0, and joy 0.

As a baseline for listening to the expectation of the expressive ability of the artificial voice, do you think that the voice read by the artificial voice always has the same meaning if it is the same word? Raise your hand if you think yes, and raise your hand if you think no. At this time, the artificial voice was set to standard speed, pitch 1.2 times, intonation 0.1, and the emotions to be given were anger 0, sadness 0, and joy 0.

As an evaluation of practicality, can you hear AI-san's rapid speech? Question. Where is this person 1. grocery store/fruit shop 2. drugstore/pharmacist 3. coffee shop 4. soba/udon shop the researcher read this out loud.

2-4. Actual Test 1 Let's begin. Please listen to the words. Kitsune udon for two. The artificial voice at this time was set to Kansai dialect intonation, standard speed 4, pitch 1 times, intonation 1, and the emotions to be given were anger 0, sadness 0, and joy 0. Where is this person 1. grocery store/fruit shop 2. drug store/pharmacist 3. coffee shop 4. buckwheat noodle shop/udon shop This was read out by the researcher.

2-5. Actual Test 2 Let's begin. Please listen to the words. Kitsune udon for two. The artificial voice at this time was set to Kansai dialect intonation, standard speed 2, pitch 1 times, intonation 1, and the emotions to be given were anger 0, sadness 0, and joy 0. Where is this person? 1. grocery store/fruit shop 2. drug store/pharmacist 3. coffee shop 4. soba/udon shop. The researcher read this out loud.

2-6 Actual Test 3 Now let's begin. Please listen to the words. Kitsune udon for two. The artificial voice at this time was set to Kansai dialect intonation, standard speed 1, pitch 1 times, intonation 1, and the emotions to be given were anger 0, sadness 0, and joy 0. Where is this person? 1. grocery store/fruit shop 2. drug store/pharmacist 3. coffee shop 4. buckwheat noodle shop/udon shop the researcher read this out loud.

3-1 Evaluating the effect of emotional attachment, the researcher asked, "Is this proposal really, OK? Raise your hand when you think the answer is OK. This question is: AI, will you marry me after this training session?

Please listen to this voice and make your decision. Please listen to this voice and make your decision.

Raise your hand when you think the answer is OK.

3-2 We asked the participants how they felt about the phrase as an evaluation of the emotional attachment. The artificial voice at this time was set to Kanto dialect intonation, standard speed 1, pitch 1 times, intonation 1, and the emotions to be assigned were anger 0, sadness 1, and joy 0.

3-3 We asked the participants how they felt about the phrase as an evaluation of the emotion to be given. The artificial voice at this time was set to Kanto dialect intonation, standard speed 1, pitch 1, intonation 1, and the emotions to be assigned were anger 0, sadness 0, and joy 1. 4.

As a post-hoc evaluation of the listening experience of the expectation of emotional expression, we will ask you one more question about the same phrase as the baseline. Do you think that artificial voices have emotions? Please raise your hand if you think yes. 5.

As an error factor question to evaluate the experience, we asked the participants if they would definitely buy these guide dog robots with artificial intelligence if they were available. 6.

As an error-factor question to evaluate experience, we asked, "If it took 40 years for these artificially intelligent guide dog robots to become available for 100,000 yen from the time when refrigerators cost 3 million yen, would you be willing to buy a guide dog robot or AI suitcase for 100,000 yen? Would you be willing to buy a guide dog robot or an AI suitcase for 100,000 yen?

The analysis included descriptive statistics for each response and 95% confidence intervals for the proportion of yes responses, and the change in the proportion of the posterior assessment relative to the baseline was evaluated at a 5% level of significance.

For the evaluation of Ecological Momentary Assessment, moving averages of reactivity to stimuli were plotted, and the time series of negative and positive perceptions of expectancy were set from 0.5 points to 0.5 points, respectively, to indicate the time when the reversal of the measured ratio occurred. IBM SPSSv20 was used to draw the graphs.

This study was conducted with the approval of the Ethics Review Board of Gunma Paz University. This study was conducted after approval from the Ethical Review Board of Gunma Paz University, and was registered as a clinical trial.

3. Results

Twenty-two participants were included in the study, and the analysis was limited to the 17 participants who had once participated in the experiment with human proximity information. They ranged in age from 27 to 80 years old. All of them resided in

Gunma Prefecture and met by public transportation. 12 (54%, 95% CI: 33%-75%, ns) answered "yes" to their expectations of AI before the intervention, 16 (72%, 95% CI: 54%-91%, p<.05) after the experience, and 20 (out of 22) answered "no" to their expectations due to economic factors.) (90.9%, 95% CI: 78%-100%, p<.05).

The stimulus-response transition during the 60-minute period is shown in Figure 1.

The effect of time lag is shown in Figure 2, which is the result of cox regression analysis.

Momentary expectation of utility of AI technology For and Against(During 60 minutes).

FIG1 Momentaly Expectation of Utility of AI Technology with Affection For and Against (Real Dta and Moving Average) The stimulus-response transition during the 60-minute period

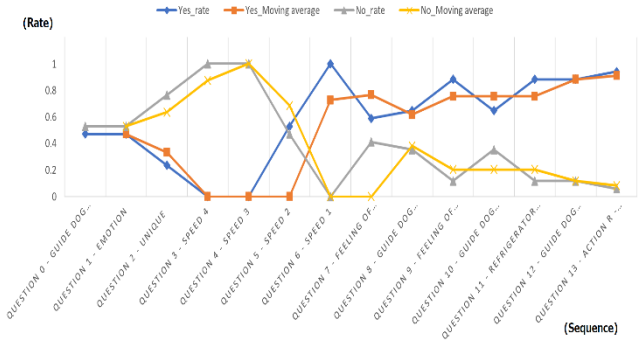
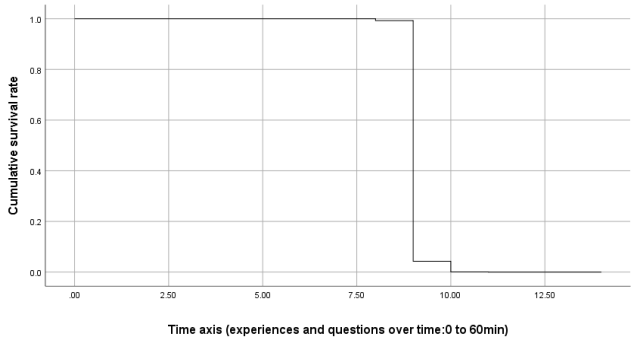


FIG2 The point at which the sense of negativity is converted to expectation by the survival function



4. Discussion

The change from negative expectations to positive expectations in the group occurred at time point 8. This indicates that the change occurred after one time point after listening to the artificial voice that gave the AI an emotion. In the bare data, the negative and positive emotions are reversed after six time points.

This was thought to be an increase in expectation based on the self-efficacy of being able to answer the question at the point where the speed of the artificial voice was slowed down and changed to a more natural speed. It is unclear whether the response to this stimulus had a greater effect than the expectation of the AI by adding emotion. In addition, although they would like to have a guide dog robot with AI technology, practical economic factors may have inhibited their expectations. It is likely that those who could visualize the price of the refrigerator responded.

Although order effect bias cannot be ruled out in this study, it can be confirmed that the staging effect of emotion impartation was in a situation close to their expectations, although the proportion of those who showed expectation after this changed significantly in the mother ratio.

Since such real-life shocks have traditionally been considered in terms of stimulus-response relationships, there have been few studies that have used such moving averages to discuss cumulative effects over a short period of time, and we believe that this is a novel way of confirming the events that can be shown in this study.

Few studies of visually impaired people participating in clinical trials have directly addressed the loss of expectancy or the positive change in negative feelings of expectancy. In terms of prior research, there are only scattered reports on stroke survivors who use wheelchairs[6]. Therefore, comparisons with previous studies can only be made surrogate in the general public. Toda et al. have conducted a background analysis of anxiety due to the evolution of technology[7]. They compared the distribution of anxiety about new technology in Japan and the United States. They reported that 17.0% (3.2+13.8) in Japan and 34.8% (12.2+22.6) in the U.S. tended to feel more income anxiety in the U.S. than in Japan. It is possible that they are answering in the sense that they cannot judge new technologies such as artificial intelligence and robots because they do not know anything about them, and the results of the cross-table analysis suggest that those who can flexibly respond to changes are more anxious. As a background, it is pointed out that people feel anxious because the impact of the introduction and diffusion of new technologies is uncertain and the impact cannot be grasped. If this is a possibility, it would be necessary for the government and others to have a deeper discussion and vision of the factors that will be affected by the introduction and diffusion of new technologies.

This view is also true for the visually impaired. Since the group in this study had been exposed to assistive technologies before, it is possible that the initial suppression of expectations was more of a pre-existing anxiety of not being able to imagine. Although it has been suggested that providing appropriate explanations in advance can reduce anxiety when introducing new technologies in healthy people, it is clear that among the visually impaired, especially the blind, it is difficult to form this image by words alone.

Therefore, when requesting participation in a clinical trial of such a new technology and conducting an RCT, some experiential component would be required to obtain prior consent. However, it will be necessary to do so carefully, estimating the expected effects of such real-life shocks. In this study, we are focusing on a group of visually impaired people living in a suburban area in Japan who are capable of navigating public transportation by themselves. Since they are already accustomed to assistive technology, it is possible that the real experience shock may be smaller than those without such experience, and therefore the results of this study are We showed that expectations of valuable artificial intelligence may change with a slight time lag after the actual experience, but the effect size may change if the order of technology presentation is changed. In the future, we hope that researchers who have such

opportunities will conduct tests to evaluate the order effect size of the effects of real experience shocks. It can be said that this research has found important findings in the field of information accessibility, where the need to spend explanatory time considering the time lag effect has not been pointed out in small-scale demonstrations of more practical AI-related technologies.

5. Conclusion

After examining the data of actual experiences and economic estimates, the utility expectations were shown to be positive. These experiential opportunities are useful as a communication method to maintain the sense of expectation.

Acknowledgments

We would like to express our sincere gratitude to MR. Isao Kimura, Chairman of the Gunma Prefecture Welfare Association for the Visually Impaired, and all the members for their cooperation in this research.

References

1. Guide Dog Robot. https://tachilab.org/jp/projects/guide-dog-robot.html (accessed 2021.6.21)
2.Suitcase becomes a guide dog... AI utilization, demonstration experiment opens. https://www.sankei.com/article/20201105-ROQLCDJMPBK4ROEI2TC24UUEFY/（accessed 2021.6.21）
3.https://pc.watch.impress.co.jp/docs/news/1259870.html#:~:text=%E4%BE%A1%E6%A0%BC%E3%81%AF%E7%B4%84800%E4%B8%87%E5%86%86,-%E5%AE%87%E9%83%BD%E5%AE%AE%20%E5%85%85&text=%E7%B1%B3Boston%20Dynamics%E3%81%AF%E3%80%814,%E3%81%A7%E7%B4%84796%E4%B8%87%E5%86%86)%E3%80%82（2021.6.21）
4. Susumu Tachi et al.: Electrocutaneous Communication in Seeing-eye Robot (MELDOG), Proceedings of the 4th Annual Conference on IEEE Engineering in Medicine and Biology Society, pp.356-361.1982.
5.Susumu Tachi, Kazuo Tanie, Kiyoshi Komoriya and Minoru Abe: Electrocutaneous Communication in a Guide Dog Robot (MELDOG), IEEE Transactions on Biomedical Engineering, Vol.BME-32, No.7, pp.461-469.1985.
6.Barker, D. J., Reid, D. & Cott, C. Acceptance and meanings of wheelchair use in senior stroke survivors. American Journal of Occupational Therapy, 58(2), 221- 230.2004.
7.Junji Toda,Hiroyuki Tuma,Susumu Hayashi,Koichi Kume, Analyzing the background of anxiety caused by the evolution of technology.IETI Discussion Paper Series 17J-049.pp1-17.2017.