Title	Communi-gate : a visual speech facilitator to enhance empathic communication in cross-lingual online discussion
Sub Title	
Author	堀井, 玲耶(Horii, Reiya) 南澤, 孝太(Minamizawa, Kōta)
Publisher	慶應義塾大学大学院メディアデザイン研究科
Publication year	2021
Jtitle	
JaLC DOI	
Abstract	
Notes	修士学位論文. 2021年度メディアデザイン学 第862号
Genre	Thesis or Dissertation
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO40001001-00002021-0862

Master's Thesis
Academic Year 2021

Communi-gate: A Visual Speech Facilitator to Enhance Empathic Communication in Cross-Lingual Online Discussion

Keio University
Graduate School of Media Design

Reiya Horii

A Master's Thesis
submitted to Keio University Graduate School of Media Design
in partial fulfillment of the requirements for the degree of
Master of Media Design

Reiya Horii

Master's Thesis Advisory Committee:

Professor Kouta Minamizawa	(Main Research Supervisor)
Professor Kai Kunze	(Sub Research Supervisor)

Master's Thesis Review Committee:

Professor Kouta Minamizawa	(Chair)
Professor Kai Kunze	(Co-Reviewer)
Professor Hideki Sunahara	(Co-Reviewer)

Abstract of Master's Thesis of Academic Year 2021

Communi-gate: A Visual Speech Facilitator to Enhance Empathic Communication in Cross-Lingual Online Discussion

Category: Design

Summary

Online cross-lingual discussion is not only challenging for non-native speakers (NNS) but also for native speakers (NS). Lack of social signals and nonverbal cues in remote communication make NS hard to adapt their speech to NNS' needs. In this paper, we propose Communi-gate, a concept of visual speech facilitator that acts as a gateway to enhance online communication between NS and NNS. Communi-gate aims to represent the communication experience of the listener that is difficult to be judged online through visual effects to encourage the speaker to adjust their speech behaviors in real-time.

We implemented and investigated the feasibility of two visual effects (Blur and Slow-motion) to provoke speakers' awareness of listeners' comprehension. The results from our validation study suggest that the Blur effect is the most effective measure to induce speakers' accommodative changes without causing extra perceived workload, which ultimately led to a psychological comfort in some listeners to try to understand the conversation better.

Keywords:

computer-mediated communication, cross-lingual communication, speech accommodation, psychophysiology, visualization

Keio University Graduate School of Media Design

Reiya Horii

Contents

List of Figures

List of Tables

Acknowledgements

To my mother, professors, friends, and partner who have supported me and my projects over the past 2.5 years by giving me the freedom to express creative ideas and the confidence to act upon them.

To my mother, thank you for respecting my wishes when I said I wanted to go to graduate school after college and for giving me the opportunity to study abroad twice. Every time I made a decision that required a lot of money, you always told me not to worry about financial matters but to enjoy what I wanted to do first, so I was able to spend these fruitful 2.5 years. I understand it's not easy to make such a decision to spend so much of your hard-earned money on me. I respect you as a woman and a mother.

To my advisor, Professor Kouta Minamizawa. Thanks to your precise advice, my research has been guided in the right direction. I think of you as a kind of "salt". Because your advice is sometimes salty in ways by pointing at which I didn't realize or didn't want to realize, and it comes with pains. However, if I still carefully improve on those areas, the flavor of my research becomes more profound. Also, I notice that you care about your students behind the scenes, which gives me a sense of security. All in all, I think that I was able to cook my research well because I chose Embodied Media.

To my sub-supervisor, Professor Kai Kunze. Thank you for connecting me to the excellent Geist members and for expanding my project. Whenever I tell you about my research progress, you honestly say "interesting". Do you know how much it boosted my confidence? I can't thank you enough for staying up late the night before I submitted my conference papers - doing final checks over and over again. You were not only a "checker" but also a "player" of my research. This fact was the most reassuring and thankful thing for me.

To my co-researcher, Professor Naomi Yamashita. Your presence has been a great help to my research. First of all, thanks to your research I was able to find meaning in my research, which was inspired by my personal experience. And I am very grateful for the opportunity to talk to you and have your help in my research. Thank you for always being kind to give me detailed feedback despite your busy schedule. When I told you that I finished submitting my thesis, you told me “If you don’t like your job, come back to our research.” I was happy to hear your words. I am very proud to have such a wonderful woman like you working in the research world, and I hope we can do research together again someday. I am very grateful to have met you through my research!

I would like to thank my senior assistant, Dr. Yun Suen Pai, for his logical and constructive comments throughout. Your responsiveness and efficient facilitation during the busy days of my Master’s thesis period was much appreciated. I am also very grateful to you for carefully reviewing and revising my long paper late into the night, despite your other duties. Your thoughtful and accurate feedback gave me the energy to go the extra mile and do my best. Thank you very much for your dedication and support.

To my collaborators, Karen Han, George Chernyshov, Zhuoqi Fu(Taku), Chi-Lan Yang, and Yurike Chandra. To Karen, I am glad that we have been able to work together on this research project as a tripod. Your generous and warm support has enabled me to proceed with this research at 100%. To George, I would like to thank you for your help, especially with the technical aspects of my research. Thank you for taking the time out of your busy schedule to work with me to figure out how to implement my ideas. To Taku, thank you for your kindness and willingness to help with my experiment so many times. To Chi-Lan, thank you for your objective and accurate advice on experimental design. I really benefited from your constructive and straightforward feedback. When you said to me, “I’m very impressed by what you’ve done for your research as a master student,” I was very happy to know that someone really appreciated my efforts so far. I am looking forward to the day when we can do research together again. To Yurike, you have built the fundamental part of my research. When I just joined EM, I didn’t know how to proceed with my research. You gave me a lot of helpful advice, which lead to my project as it is now. Your advice is always right to the point. Thank you for sticking with me through all these years and supporting me in giving shape to my thoughts.

To my friends and professors whom I met through KMD and GID journey. I want to thank all of you for giving me good advice on my research. I express my gratitude for your willingness to work with me on my ideas and to participate in my experiments. Without your help, I could not have completed all the experiments. Especially, I would like to express my sincere gratitude to my doctoral students, Takayoshi Hagiwara and Harry Krekoukiotis. Thanks to the help of Mr. Hagiwara, I was able to process the data smoothly in the analysis and discussion of my master's thesis. And to Harry, thank you for being my mentor and helping me with my research since my first year. You have always guided me in the right direction. I really appreciate all your consideration and time.

Also, I would like to express my gratitude to Professor Akira Kato for his help during my thesis writing. I would like to express my sincere gratitude to you for allowing me to use your paid overleaf account and for helping me reduce the data in my files. I was very happy with your dedicated support.

Lastly, to my partner and our dear hamster, Taraco, who supported me emotionally during the pandemic. Firstly to my partner, thank you for being the person who understands me the best at all times and for supporting my growth. Whenever I start panicking due to my impatience, you always teach me to take a breath and look at the big picture of the situation. When I feel overwhelmed with anxiety, you are always on my side to reassure me. You are the source of my strength and this research. Thank you for all your love and support. Also, I admire you, Taraco, for not caring about human stress. Your optimism always relieves my tension. Taraco, you are absolutely adorable and I'm lucky to have you. (PS (2021/07/26): When I was writing the sentences above, I thought it was very natural to see you, and I didn't expect you to cross the rainbow bridge such soon. Now I'm in deep sadness and I wish many times this was a dream. I really hope I could have met you in person and stroked your fluffy fur this summer. Although I'm just waiting for time to shed this sadness for now, I am truly grateful that Taraco came to us for the past six months. You have given us irreplaceable time and smiles. Thank you and I love you.)

All of you have been a great help to me,
and without you, I would not have made it this far in my research.

Chapter 1
Introduction

1.1. The Relationship with Foreign Language Users

With the changing demographics of the world, attention to intercultural communication and the process of communicating across language barriers continue to grow. In the education field, there were major changes to the network of international student mobility in higher education over a 10-year period between 1999 to 2008 [11]. The U.S., a country with the highest number of international students, experienced an all-time high in international student enrollment in the 2018/19 academic year, marking the fourth consecutive year that the number of international students exceeded one million [12]. In Japan, the total number of international students at higher education institutions and Japanese language institutions was reported to be 312,214, according to the 2019 Survey of International Student Enrollment submitted by the Japan Student Services Organization [13, 14].

In light of all this, it is clear that there are increasing opportunities for people with different mother tongues to communicate both around the world and in Japan. Therefore, the ability to navigate cross-lingual communication has become more essential for individuals living in the same society than ever.

1.2. The Challenge in Cross-Lingual Communication

International students often have difficulty in communicating with others because of language differences when living abroad. For example, when I was an exchange student in the U.S., I experienced the challenge and loneliness of communication due to cultural and language barriers. Although having prepared well for the study abroad, I could not keep up with the conversations of native speakers(NS), and I sensed a lack of my own presence in class. Especially when the school I studied at only had a small number of international students, it was mentally tough not being able to receive support from the local students and instructors who understood what I was going through. After returning to Japan, I found that my international student friends had also gone through the difficulty of making friends with Japanese students and understanding ambiguous expressions unique to the Japanese language. These findings made me become aware of the problems of the communication barriers that international students, as non-native speakers(NNS), go through due to the language and cultural differences.

With this event as a turning point, I would like to create a method of communication that would allow everyone to have an inclusive dialogue regardless of the language differences. In order to achieve this goal, it is necessary to create more understanding and empathy in people who are close to foreign language learners.

There are two main factors that drive this communication problem.

- NS's lack of understanding of NNS
- NNS's hesitation to ask for help or unwillingness to interrupt

First, NS need to be aware in order to consider NNS's comprehension level. Since NS hardly ever imagine how difficult it is for NNS to use their language, they often speak the same way as how they would to other NS - without knowing there is an increased burden on the NNS. Second is the strong sense of insecurity that NNS face when using a foreign language. As a result, NNS tend to underestimate their language abilities and become reluctant to speak. They avoid speaking out of fear of being judged by others especially if they are not confident in their language abilities.

Despite their inner struggles, this attitude of NNS is often misread by NS as an individual personality and the difficulties on the part of NNS are often not recognized. In addition, this communication problem is even more serious in online communication, where it is difficult to see the other person's response and the emphasis is on talking. With the increase of online classes and meetings due to the Coronavirus pandemic, this communication problem between NS and NNS has become a critical issue that needs to be resolved.

1.3. Current Focus on Communication Support

Unfortunately, this foreign language communication disorder suffered by NNS is only considered to be a temporary impairment of the individual's language skills and is correctable. Therefore, many studies aim to mitigate this problem by helping NNS to learn themselves. Based on this understanding, not many studies have investigated the willingness of NS to respond in multilingual communication with the aim of facilitating their efforts to respond to NNS. This is untoward, because NS may be able to allocate more cognitive resources to additional tasks (e.g., helping NS understand how to best respond) while communicating in their native language.

Additionally, this problem is especially more acute in online situations. This is because the difficulty in interpreting non-verbal cues makes it more difficult for the NS to interpret the NNS's struggles in comprehension. In response to this problem, however, the research area of discussion facilitation systems has not considered the issue of communication between NS and NNS, although there are studies that aim to support the views of minority or passive participants. Therefore, in online cross-lingual discussions, NS are often unaware that NNS are struggling to understand their conversations, and this language-induced communication problem has not been solved. However, it is very important to take into account the difficulties of NNS in online communication.

1.4. Contribution

Building on previous research, my objective is to bridge a gap in communication between NS and NNS in online settings. In this study, NS is defined as the facilitator to mitigate the communication gap and my aim is to trigger NS's awareness to NNS's understanding and behavioral change in their speech. To bridge the gap in understanding NNS's listening comprehension, in this paper, I am designing Communi-gate, a visual speech facilitator that helps NS to adjust their speaking behavior to enhance NNS's comprehension. The purpose of this visual speech facilitator is to give live feedback to the NS prompting them to change the speaking style (slow down, use shorter utterances, simpler language, etc.) when NNS has a difficulty understanding during an online conversation.

To sum it up, this thesis provides the following contributions:

- I proposed Communi-gate, a visual speech facilitator that assists online communication in multicultural communication.
- I found that the Blur effect is overall the most preferred visual effect, and from our evaluation (N=21), the results suggest that the Blur effect to be the most effective in terms of inducing behavioral changes in speakers without causing extra perceived workload.
- I presented design guidelines and potential applications for Communi-gate to be used not just for current remote video platforms, but also for future platforms across a variety of languages.

1.5. Thesis Structure

This paper consists of 5 chapters. Chapter 1 provides the background which explains the need for this thesis. Following the Chapter 2 describes the problems in foreign language communication, their causes, and the current research, concepts, and applications that are being done to alleviate the communication issue in online settings. Chapter 3 discusses the problem and goal setting of our project, fieldwork and user research for concept design, and the concept validation process. Chapter 4 presents a series of verification experiments to evaluate the system developed in Chapter 3. Finally, Chapter 5 provides a summary and conclusion of this research.

Chapter 2
Literature Reviews

2.1. Challenges that NNS have faced in Multilingual Communication

Understanding a discussion in a second language comes with its challenges for NNS. In particular, there is a big difference between NNS and NS in terms of the difficulty of language operation. In general, global teams use a common language such as English to communicate, but for NNS, using the common language as a foreign language is a burden because it reduces their ability to think and process information [15]. For example, when there are a large number of NS, conversations are tend to be carried on among themselves at a faster rate than when talking to NNS one-on-one [16, 17], especially when the discussion is heated [1, 18].

Why is there so much difference in mutual-understanding between NS and NNS? It is due to the unique anxieties associated with learning a second language. According to Khan and Al-Mahrooqi [19], foreign language learners often feel frustrated and anxious when communicating in a foreign language due to their lack of vocabulary and ability to express themselves fully in the language. When language skills are limited, they tend to monitor themselves too much for fear of being judged by others who are more proficient in the language. This particular anxiety that foreign language learners experience when communicating in a foreign language is called foreign language communication anxiety (FLCA).

Another study has also confirmed that speaking is one of the most anxiety-producing activities in the use of a foreign language. Learners of a foreign language may feel insecure about their language ability and avoid communicative situations where they have to speak in a foreign language in groups or in public, or where they have to listen to and understand spoken language [20]. Hence, understanding the learners' emotional state and reasons that trigger anxiety is crucial for helping them to succeed in language learning.

2.2. Misunderstanding of NS toward NNS's FLCA

Previous research on collaboration between NS and NNS has suggested that NS may not be aware of the extent of the linguistic challenges faced by NNS and how such challenges interfere with their interaction [21]. This conjecture may be particularly salient in video conferencing, where non-verbal cues are limited and audio information is predominant [22]. This is because in online meetings, NNS tend to avoid speaking out of concern for the attention of others, making it difficult for NS to notice NNS' linguistic difficulties. Also, compared to in-person, online conversations provide limited visual information to capture the listener's comprehension level. For that reason, if NNS say very little during a meeting, NS may attribute their low level of participation to temperamental factors (e.g. shyness) rather than language barriers [23].

Although some studies admit that NS are motivated to respond in a way that is understandable to NNS, for example by reducing their own speaking speed, it has not practically succeeded [24–27]. Because in practice, speech characteristics such as speech rate are often beyond the conscious control of the speaker, especially when he or she is excited or trying to persuade [28, 29]. Also, even if the NS consciously tries to respond, speech modification may not be appropriate or effective because the NNS may not have requested it or the NS may not be aware of the level of response required by the NNS [30].

2.3. Multilingual Support Tools

2.3.1 Improve NNS's English Conversation Skills

While many past works have aimed to mitigate this communication problem in foreign language usage, most of them focus on helping NNS to improve their English conversation skills.

Previous research has focused on supporting NNS in multi-person collaborative work, and self-learning tools have been developed [31]. The most widely adopted method is to provide NNS with an automatic recording. This aims to improve NNS understanding in audio and video conversations, and to enable NNS to catch up with conversations they have missed. However, such an approach has both benefits and costs for NNS, since reading an automated transcript while following an ongoing conversation imposes a higher cognitive load. Yamashita et al. [1] responded to this problem by creating artificial silent gaps in real-time conversations, which helped NNS to understand and participate better, but also had some negative effects, such as speech overlap (See Fig.2.1).

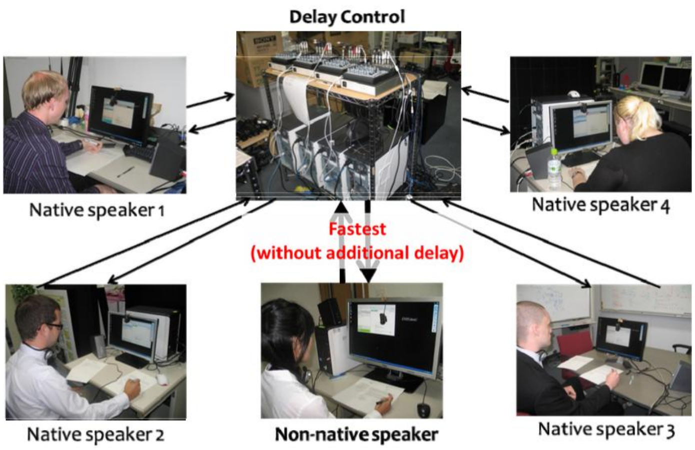


Figure 2.1 The Insertion of Silent Gaps between Conversational Turns [1]

Another research of Nishida [2] is also being done on an accent modification system that combines a speech recognition engine with a speech synthesis engine to convert native speech into a foreign language accent. The purpose of the system is to make the pronunciation of the NS resemble the pronunciation of the language learner's native language, as an auxiliary function to make it easier for the NNS to understand the conversation (See Fig.2.2). However, due to the implementation limitations of the system, the user needs to press a button before starting a conversation, and the accent-transformed speech is played back after the user has finished speaking, which takes time to process and makes it difficult to have a fast-paced conversation with the NS.

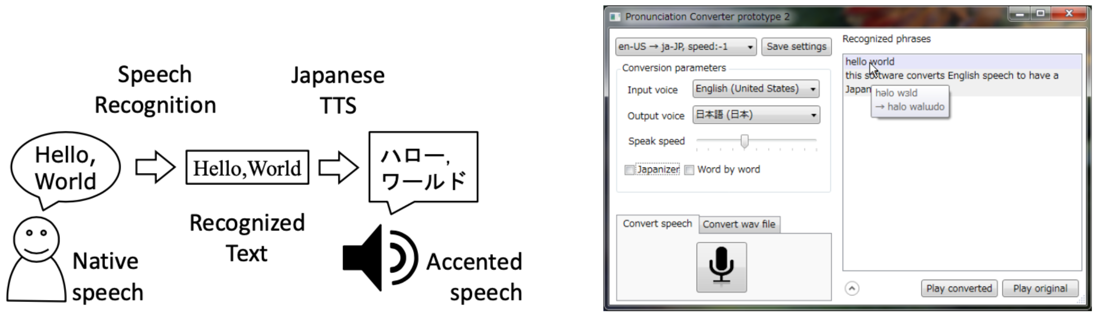


Figure 2.2 Native-to-Foreign Speech Accent Conversion [2]

2.3.2 The Role of NS to Mitigate the Communication Gap

Previous research have focused on exploring tools to support NNS, but only a few studies have taken into account the role of the NS when designing tools to support NNS.

In one of those few examples, Duan et al. [3] investigated the use of a real-time speech speedometer to allow NS to monitor their own speaking speed and adjust their NNS accordingly. However, their statistical analysis suggests that the speech speedometer causes NS to lose track of their thought processes, making communication more laborious (See Fig.2.3).

In the study of Kitayama and Nishimoto [4], they proposed “DAFlingual”, a method to assist native English speakers in conversations with non-native English speakers by applying delayed auditory feedback (DAF). DAF is a phenomenon known as speech suppression without physical discomfort when the speaker’s voice is delayed by about 200 ms and fed back to the speaker. The results of their experiments confirm that DAF has the effect of reducing the speed of speech, inducing word jams and suppressing long sentences for the speaker(NS) (See Fig.2.4).

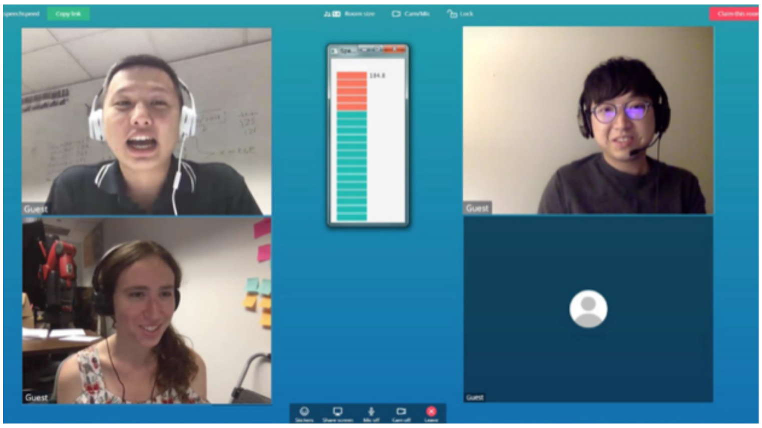


Figure 2.3 The Real-Time Speech Speedometer [3]

Figure 2.4 DAFlingual: Applying Delayed Auditory Feedback [4]

However, such forced behavioral change of NNS by the intervention of speech information increases the cognitive load on NS at the same time, making it difficult for them to hold a conversation smoothly. Even worse, they may become tired of speaking, and lose their motivation to care about the NNS's comprehension. Therefore, it is essential to strike a good balance between interventions in speech and improving the NS's motivation to consider NNS's comprehension.

2.4. Virtual Communication Assistance Tools

Many of the current online facilitation tools are aiming (1) to support group work online (such as equal participation, decision-making, and productivity), or (2) to improve speakers' speaking skills and mutual understanding with listeners by integrating with listeners' feedback. In order to achieve these objectives, three main methods are used.

2.4.1 Capturing Group Work Dynamics for Self-reflection

In recent years, there has been an increasing amount of research on "mirroring" displays, which attempt to enhance group work by visualizing specific elements of the group work [32–35]. Typically, the display visualizes the percentage of each participant's participation in the conversation to help balance the conversation share across participants.

For instance, the study of Pooja and Karrie [5] is developing conversation mapping programs in order to get an overview complete picture of communication in online groups by showing the amount of time each participant has been speaking. Analysis of the studies has shown that the visualization shows the personality traits of each participant expressed in explicit behavior (See Fig.2.5). For example, the those who were quiet during the conversation or the those who were dominated the conversation. More interestingly, the use of visualization also led to some discoveries, such as a decrease in the amount of conversations for users who had many conversations, and an increase in the amount of conversations for users who had few conversations. Research by the MIT Media Lab [6] has also created a visualization system to identify patterns of turn-taking, overlapped speech, and conversation timeline in face-to-face meetings (See Fig.2.6). This research shows that speakers can reflect on their social interactions by visualizing their speech patterns.

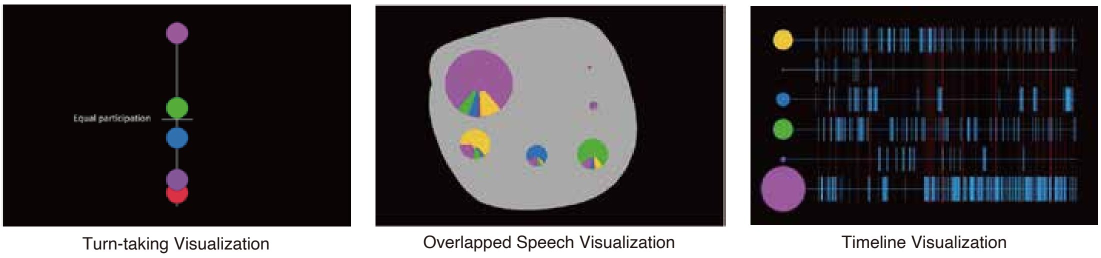


Figure 2.5 The Visualization of Group's Interaction Dynamics [5]

Figure 2.6 The Visualization Interface of Speaker Participation [6]

These studies show that a visualization system with explicit behavioral representations of each participant's personality traits can encourage speakers to self-reflect on their own speech activity. Also, it may lead to changes in their attitudes and internal motivations to consider other participants during the conversation. This can be a valid basis for encouraging NS to deepen their understanding toward NNS and to change their behaviors in conversations between NS and NNS.

2.4.2 Inducing Changes in Perception through Visual Modification

There has also been research into the use of visual illusions in online group work to generate behavioral change and decision-making in users. For example, Seta et al. [7] conducted a study using a novel video-chat system, "Divided Presence," which aims to reduce the bias of majority rule by equalizing the apparent number of majority and minority participants using a pseudo-population increase. Computer graphics avatars are assigned to each member of an arbitrary minority group, and these avatars speak on behalf of the other members, apparently increasing the number of members in the discussion. With this feature, there was increase in the level of agreement when the majority participant finally agreed with the minority opinion (See Fig.2.7). In other words, by visually amplifying the number of participants, a behavioral change occurred in which the majority listeners listened carefully to the minority side.

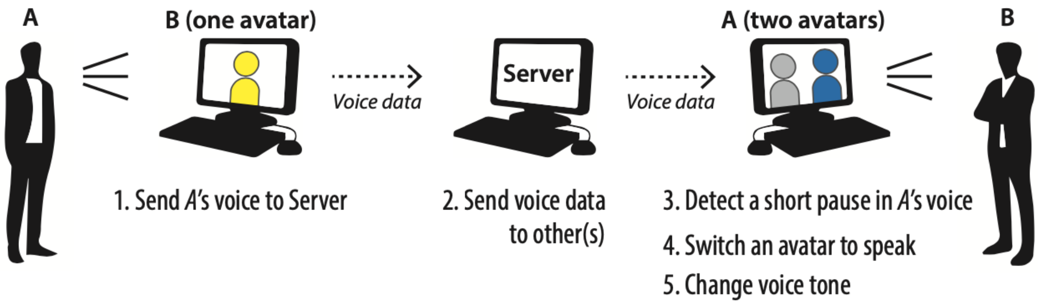


Figure 2.7 Divided Presence: Pseudo-Population Increase [7]

On the other hand, there are research that aim to create a collaborative video conferencing environment by using computers to subtly change the facial expressions of participants, thereby changing the mood of the entire group [36, 37]. In general, psychological research has shown that people's facial expressions affect their creativity, i.e. that emotions affect creativity and that the facial expressions of others affect their own emotions. Based on these findings, Yoshida et al. [8] proposed "Smart Face", a method to improve creativity by changing people's facial expressions during video conferencing. They developed a method to influence creativity by changing facial appearance in real time, including facial expressions and facial similarity. The results of their user study showed that modifying the participants' faces to smile had a significant effect on creativity during the conversation (See Fig.2.8).

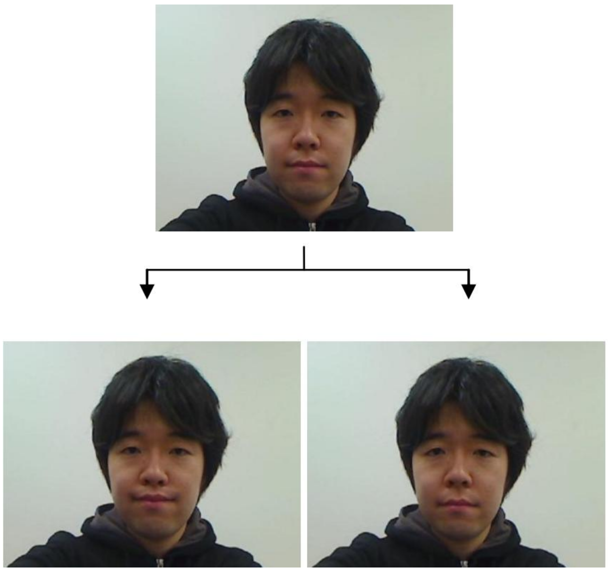


Figure 2.8 Smart Face: Real-time Facial Deformation [8]

From these studies, we can say that the intervention of a computer as a third party can change the attitude, stance and behavior of group members towards other members. In other words, by intervening with visual illusions on behalf of NNS, who are too anxious to interrupt a conversation to ask for help, the computer may be able to bring about behavioral changes in the way NS speak. Thus, visual intervention could be one way of bridging the communication gap between NS and NNS without placing an unnecessary psychological burden on the NNS.

2.4.3 Capturing Audience Feedback

In contrast to face-to-face communication, online communication provides limited feedback from the audience. Thereby, it is very difficult for the speaker to establish an intimate relationship with the listener and effectively adapt the content of the conversation. Unfortunately, current video conferencing platforms are limited in the solutions they can support. As a result, in order to understand audience responses better during online communication, there has been a great deal of interest in research on emotion sensing to facilitate online communication by capturing audience responses.

To capture the audience's responses, behavioral signals (facial reactions and head gestures) and physiological signals (skin conductance and heart rate variability (HRV)) are commonly used. For example, AffectiveSpotlight was implemented in Microsoft Teams [9], which analyzes the facial reactions and head gestures of the audience and dynamically spotlights the most expressive users (See Fig.2.9). Wei et al. [10] also developed a visualization system that predicts the audience's psychological state (anxiety, excitement, boredom, etc.) in real time based on the audience's physiological signals (See Fig.2.10). In their experiments, several speakers stated that they achieved communication feedback that allowed them to adjust their presentations as needed (e.g., to provide additional explanations) based on the audience's responses. Some participants in the experiment said that the system was useful in providing timely support to confused audiences. Whereas, constant display of the audience's biological responses and behavioral changes on the screen increased the cognitive load for the speaker in some aspects. From this finding, we can conclude that it is necessary to present the speaker with information that is effective in building a relationship with the listener at the right time.

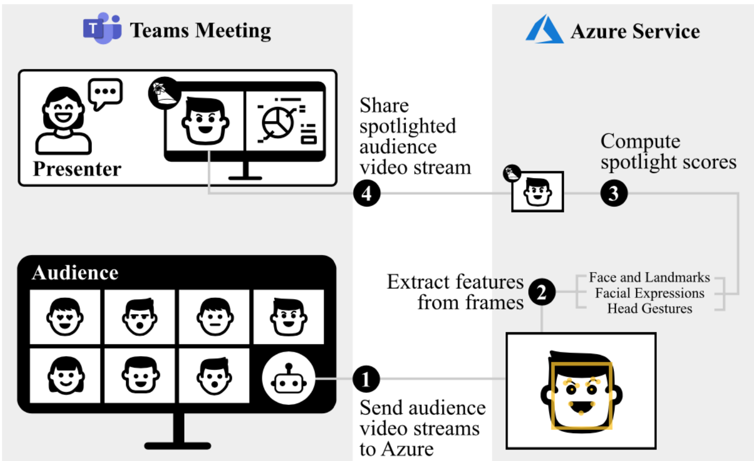


Figure 2.9 AffectiveSpotlight: A Real-time Feedback Support System for Online Presenters [9]

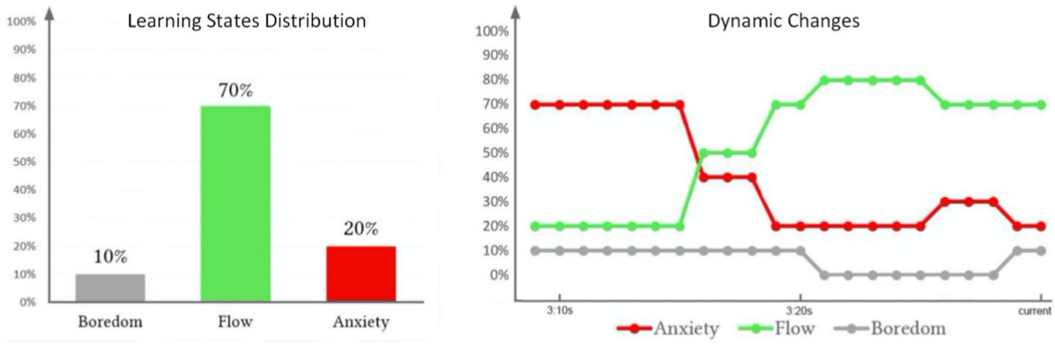


Figure 2.10 The Visualization of Audience's Psychological States [10]

Therefore, if we relate these results to the context of our study, we may be able to motivate NNS to talk in accordance with their NNS by indirectly communicating the sense of anxiety that NNS perceive from their biological responses to the NS. In order to prevent distraction of the speaker, the method of informing the NNS when they are not sure is considered to be effective rather than always sharing the psychological situation of the NNS.

2.5. Summary

Based on findings from those related works, it is found that NNS encounter a large mental load (FLCA) in foreign language conversation. Contrary to this situation, the suffering they are experiencing is often overlooked as something that can be remedied in time by self-learning, and is almost never understood, especially by NS. In addition, online conversations make it even more difficult for people around NNS to be responsive to the reactions of NNS. However, with the support of those around them, these communication gaps can be greatly improved, and have a significant impact on NNS' enthusiasm and achievement in learning a foreign language.

Many of the current online facilitation tools are designed for (1) visualization of conversational activities for self-reflection, (2) use of visual information to generate perception and behavioral changes, and (3) audience feedback to improve mutual understanding between speaker and listener. However, there are only a small number of studies that assume the use case of a foreign language. In addition, these few foreign language support tools increase the cognitive load on NS by forcing them to intervene with audio information, making it difficult for them to carry out smooth conversations. Therefore, it is necessary to come up with a method that does not place a stressful load on the NS and naturally promotes behavioral change.

Chapter 3
Empathic Communication Design

3.1. Problem Definition and Our Mission

Following an ongoing online discussion conducted in a second language is challenging for NNS. In particular, there are often communication problems between NS and NNS, especially when there is a difference in the level of acquisition of the language used as a common language. For instance, when NS speak too fast [15], the NNS can have a hard time following. During group discussions, some of them would hesitate to ask questions - afraid of interfering. It is also harder to signal difficulties in comprehension as non-verbal cues are harder to interpret in online meetings [38,39]. Therefore, it is crucial for system designers to consider the plight of NNS in online communication.

3.2. Concept Design Analysis

For this communication problem, I define international students as the target users. In order to understand the issues international students faced, I participated in the Global Innovation Design (GID) course, an inter-university study abroad program (See Fig.3.1). In the GID course, students studied at the Royal College of Art (RCA) in the UK and the Pratt Institute in the U.S. The purpose of this participation was to experience the difficulties of communication due to the language differences as an international student myself and to deepen my understanding of the factors that brought the issues.

To clarify the issues faced by international students, these three kinds of research below were mainly conducted during the study abroad program.

- Survey for those who had studied abroad in the past

- Observations and interviews with international students and local students in RCA
- Fieldwork at a local language school in London

From these fieldwork and interview experiences, we drew out factors that driven into the problem from the analysis of user behaviors and designed solution concepts.

Figure 3.1 Global Innovation Design Course

3.2.1 Survey for Students have experienced Study Abroad

To investigate what percentage of international students experienced the above language communication problems, a survey was conducted with 30 students who had studied abroad (16 female and 14 male). The questions included the place of study, number of times studied abroad, and initial feelings about studying abroad. The distribution of students' nationalities includes 5 Americans, 2 British, 1 Canadian, 2 Dutch, 1 French, 1 Polish, 1 Indian, 7 Chinese, 1 Taiwanese, 4 Indonesians, 3 Japanese, 2 Singaporeans.

According to the results of the survey, 12 out of 30 students answered that they had negative feelings at the beginning of studying abroad. Most of the students who responded to this answer, regardless of the number of times they had studied abroad, had studied in a country where the language was different from their native language. In other words, they had to use a foreign language to communicate with the locals. In fact, 3 of them said the reason was because of the language barrier. For instance, "I didn't know how to start a conversation with the local people," "I was worried if I could make friends," "I was frustrated because I couldn't make my presence felt in the conversation," and "The nervousness just stops when I actually get into it and start talking to people."

One other student mentioned, "I think the problem with meeting people while studying abroad is that everyone goes back to a different country right away. In other words, if you don't make friends with them from the beginning, there is no chance of them becoming your friends." Most international students spend the first few months of their stay trying to overcome their language problems. It means that only in the latter of their stay they were able to start making friends and joining classes actively to enjoy their time abroad. Therefore, if I can shorten the time spent on language problems, international students can have a more satisfying experience with their study abroad.

3.2.2 Observations and Interviews with International Students and Local Students in RCA

RCA is a school where people from very different cultural backgrounds come together, and much of the classwork was done in groups. However, from the field interviews, many students reported that language barriers and differences in verbal and non-verbal communication styles caused tension in groups. More importantly, they mentioned that it was difficult to talk to each other about this tricky and sensitive topic, especially in tense group work situations where there are time pressure, conflicting ideas, and other factors.

For example, there was a division in the level of participation and contribution depending on the student's English proficiency. Students with higher English proficiency were easily involved in the work while those with lower proficiency only listened and contributed less to the group work (See Fig.3.2). The less English proficient students also tend to avoid interacting with other students after class. Although the students themselves were aware of this problem, they did not have a chance to discuss it because their priority was to complete the group work. As a result, some of the local students assumed that the lack of interactions was due to cultural differences.

From these observations, it is evident that within group work this foreign lan-

Figure 3.2 A Scene of Group Work in RCA

guage communication problem frequently occurred. At present, however, although the students themselves are aware of the communication problems between group members, they do not know whether they are caused by a language barrier, cultural background, or individual personalities; sadly no concrete solutions have been taken. Moreover, this seems to affect the social relationships among students. For all of these reasons, I believe that students need to be aware of this significant problem and that it needs concrete solutions otherwise it can reduce the overall quality of education.

3.2.3 Fieldwork at a Local Language School

In order to investigate how international students feel when using English as a foreign language, a fieldwork study was conducted at a local language school in London, the Edward Language School. Surprisingly, during this fieldwork, there were no students who strayed from the conversation in group conversations, as was seen in the RCA. On the contrary to my expectation, all members of the group were actively participating in the conversation. This gave me an insight into the fact that international students of the same language ability are not afraid to express themselves in a foreign language. Comparing the two situations at the RCA and the Edward Language School, the reason for this difference can be attributed to the basic language skills in communicating among the students (See Fig.3.3).

Figure 3.3 Fieldwork at the Edward Language School

3.3. Goal Setting

The results of all the researche showed that most of the international students experienced a sense of isolation when talking with English native/fluent speakers. It is because there is a gap in basic language skills when communicating in English. This problem seems to be a common occurrence, especially in classes where group work is emphasized. This greatly affects the international students' participation in group work and their relationships with the team members.

Based on these findings, this study aims to develop a communication support tool for international students who often face language barriers with local students. The goal is to bridge the communication gap between NS/fluent English speakers and NNS and ultimately to reduce the anxiety NNS have about language barriers. Our objective is to create an experience in which NS or fluent English speakers show understanding of the linguistic constraints of NNS and naturally change the way they speak to communicate effectively with each other.

3.4. Concept Building Process

3.4.1 Iteration 1: Speed-Up Speaker

Figure 3.4 Speed-up Speaker

There is a major difference in the cognitive load of communication between those who have different basic language skills. For example, it requires more work for NNS to catch up in normal conversations compared to NS. As the initial discourse on this issue, I organized a hands-on exhibition in the final presentation of my individual project at RCA to examine this assumption: “How can we break down language barriers to create a greater understanding of linguistic differences and affect behavior change?” The objective of this exhibition was to help students at RCA understand and empathize with the difficulties international students experience when speaking in English as a foreign language, utilizing the perspective of NNS.

Hypothesis

By having the same experience, people can put themselves in others' shoes to understand how other people feel and act. Based on the information, I hypothesized *if people experience conversations in the perspective of a NNS, they will feel sympathy for NNS and be motivated to modify their own speech activity.*

To foster a greater understanding of linguistic differences and to break down language barriers, I decided to introduce the concept of simulating the experience of NNS listening to a conversation in an unfamiliar foreign language.

Conceptual Idea

From design analysis, I found that international students experienced a sense of alienation when they talk with native or fluent speakers. This is due to the fast tempo of the conversation and the unwillingness to interrupt. With this viewpoint in mind, I designed a hands-on exhibit using "Sped-up Speaker" that mimics the difficulties that foreign speakers face when having a conversation with native speakers (Fig.3.4). The purpose of this exhibition was to create a mechanism to visualize the perceptual experience of the NNS (the NS' speaking speed sounds twice as fast as NNS') and to make the NS aware of the problem that the NNS had been facing.

Prototyping

I used a node function in MAX/MSP 2[1], which interpolates audio data graphically. The function of the audio system was to record audio input from the microphone and play it back at twice the speed (Fig.3.5). An external speaker was connected to a laptop over Bluetooth.

Test (Hands-on Exhibition)

The purpose of this exhibit was to test our concept and to see if it could help NS understand the communication problems better faced by NNS. In the exhibition, I set a task for participants to play 'Face Guessing Game' with a sped-up speaker

1 https://cycling74.com

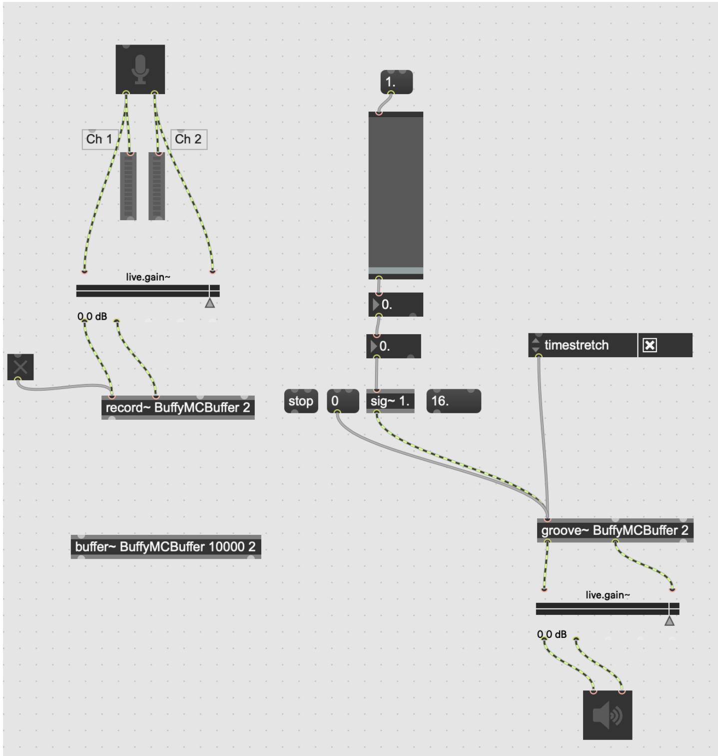


Figure 3.5 A Sample of Audio Play Back Speed-up Code

who would deliver information well to the receiving end. To achieve this task, mutual understanding was important for both parties to share the information correctly. Therefore, the speaker would be required to change the way they spoke in order to convey information to the other person efficiently under the intervention of the sped-up speaker (Fig.3.6). In total, 27 students of RCA participated in the exhibition.

Procedure

The participant (the speaker) had to make a face from the pieces provided without showing it to their partner. After recording their description, it was played aloud at twice the speed for both participants to hear. The other participant (the listener), who did not make the face, must try to build the face in how NS responded to communication issues with NNS. Therefore, it is important to demonstrate the

Figure 3.6 A Scene of the Exhibition in RCA

next goal of establishing awareness and behavior change.

Results and Insights

According to results from the survey, some people mentioned that the experience through this exhibition remind them of a similar experience in communicating with people that do not share the same language such as group work in school. Luckily, some opinions received from the participants agreed with our concept, such as "I was happy that I was able to share my problems that I usually have through the experiment (from NNS)" and "I would like to consider my way of speaking in group work considering the level of understanding of foreign students (from NS)". This proved that the installation could create empathy in local students to understand language barrier that International students have faced.

For the next step, it requires me to think of methods to implement it in a more practical way, in everyday usage scenarios. This is because an exhibition usually ends up being a one-time experience. It means even though it gives NS awareness of the issues, participants will easily forget about this afterward. In addition to raising awareness to the communication issue, it is also necessary to provoke behavioral changes in how NS actually respond to communication issues with NNS. Therefore, it is important to think of ways to demonstrate the next goal of establishing awareness and behavior change.

3.4.2 Iteration 2: Bubble Visualization Overlay

Figure 3.7 Speech Speed Visualization System

Due to the impact of the pandemic, the structure of classes was moved from offline to online. Therefore, our concept scenario changed to solve the communication gap between NS and NNS in online settings. Because the communication problem would be more complicated online, a dialogue style that relies on voice conversation would discourage NNS from asking questions or asking for help. It's also hard for NS to notice NNS' struggles due to the lack of non-verbal cues to capture NNS' comprehension level. On this iteration especially in technical implementation, I gained massive help from my senior student, Yurike Chandra.

Hypothesis

Speed is one of the most common factors that hinders NNS to comprehend what NS say. Furthermore, due to NS' familiarity with the way English works, they may unconsciously speak faster. Hence, we hypothesized *if a speaker sees their own speech activity, it encourages the speaker to modify their own speech.*

Conceptual Idea

With the transformation of the education environment to online, we decided to implement our idea in online meeting situations. Hence, we developed the concept is a visualization system that would help NES realize their speaking speed (Fig.3.7).

In terms of designing speaking speed, in general, it can be calculated by dividing the total number of words by the number of minutes the speech took. In other words, a faster speed will have more information conveyed in a shorter amount of time. Reflecting such a relationship between speed and information, we decided to visualize information density as an indicator of NS's speaking speed.

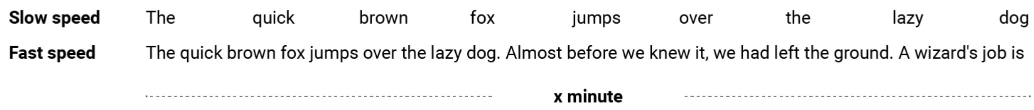


Figure 3.8 Correlation between Speed and Information Density

We then used a sequence of bubbles to represent the information density. When NES speaks at a fast speed, the sequence of bubbles will become denser, and vice versa (Fig.3.9).

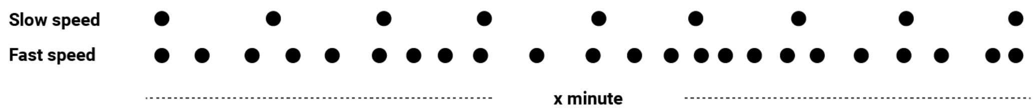


Figure 3.9 A Sequence of Bubbles Represents Information Density

Prototyping

We used TouchDesigner, a visual programming language [2], to create our visualization system. TouchDesigner would take in the sound input from the computer's microphone and calculate the frequency. When the sound reached a threshold, it would generate a bubble through a particle node. We then connected TouchDesigner output to NDI Virtual Input [3], which streamed the visuals to an online video conference software, Zoom [4] (Fig.3.10).

2 https://derivative.ca/product

3 https://ndi.tv/tools/

4 https://zoom.us/

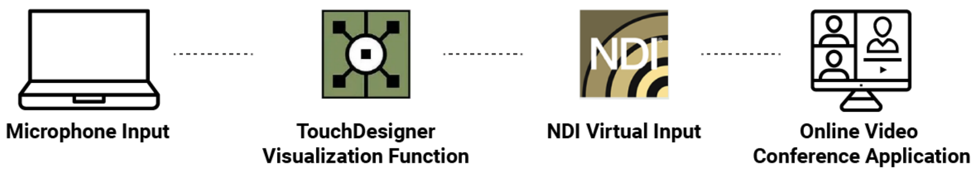


Figure 3.10 Components of Speech Speed Visualization System

Test

The study aimed to test our concept and see how visualizing speech speed affects NS' actual speaking speed. We used a within-subjects study design to compare an online discussion experience with two experimental conditions: (1) with visualization and (2) without visualization. We hypothesized that *if the visualization helps NS understand their speaking speed, they would speak slower during the session with visualization than the session without visualization.*

In total, 13 adult participants (4 Japanese - NNS, 9 English Speakers - NS) took part in this study. Participants were briefed on the experiment and signed a consent form. We divided our participants into 4 groups with 1 Japanese participant each and 2 to 3 English-speaking participants. We invited them to join a Zoom meeting with the speech bubble visualization installed on the English-speaking participants' video (See Fig.3.11).

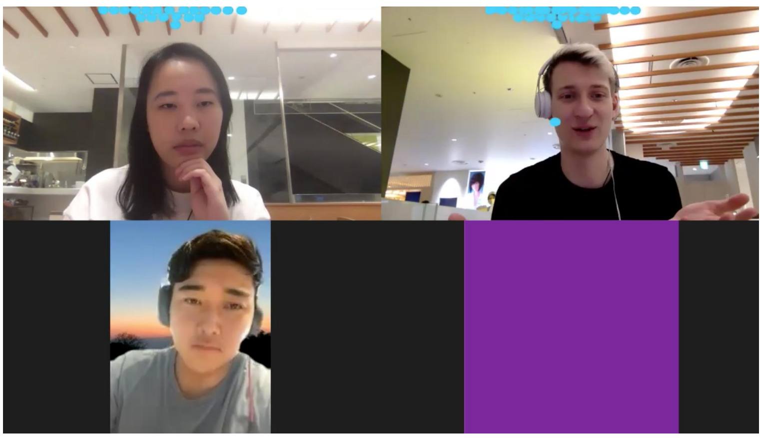


Figure 3.11 Online Discussion Setting

Then, they went through 2 discussions with the two previously mentioned conditions (10 minutes each). A 5-minute break was given in between each discussion (See Fig.3.12).

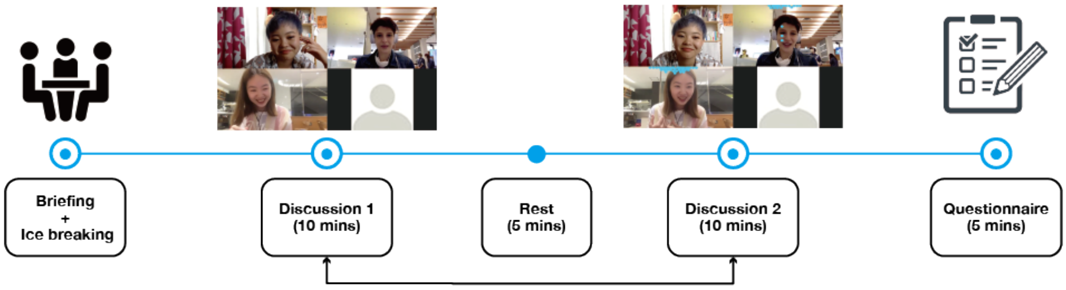


Figure 3.12 Experiment Procedure

The appearance of the conditions in the discussion was randomized. Before the study began, participants were told that the generated bubble represented their speaking speed. They were also given time to familiarize themselves with one another through a 5-minute ice-breaking session. The selected topics were open-ended questions that allowed participants to speak easily without the need to have prior knowledge. At the end of the experiment, participants were asked to answer some questionnaires regarding the experience.

Insights from Results

The qualitative evaluation suggested that the visualization helped most NS to understand their speech speed. However, they were still unable to modify their speaking rate, which caused the Friedman test to show no significant difference ($X^2(2) = 0.1111$, p $= 0.73888$) in the speech speed for both conditions (See Fig.3.13).

A similar observation was found in the previous research when using a speech speedometer [3], which implied that the NS required cues (e.g., timing to slow down, ideal speech rate) to be able to adjust their speaking behavior. Moreover, in the interview, many NS mentioned that when they talk, they tend to look at the listeners' screen to see if they understand. This behavior further highlights the importance of the listeners' feedback (e.g., head-nodding, confused face) in

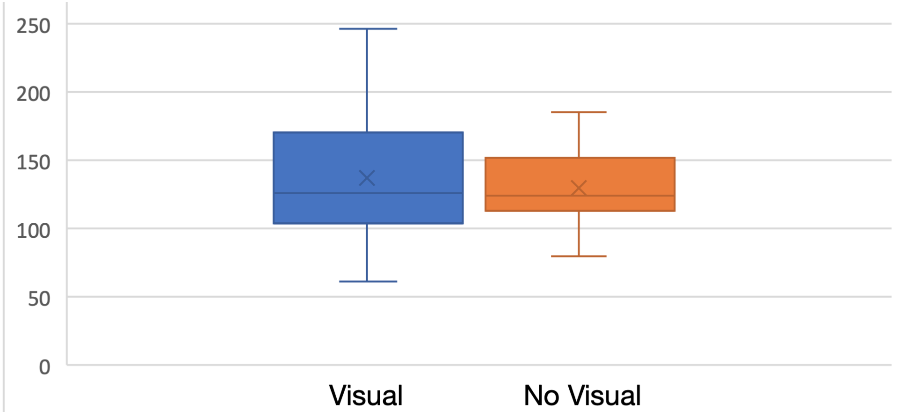


Figure 3.13 Mean and Standard Deviation of Speech Speed

aiding speakers to change their speech rate.

Based on these findings, we hope to integrate listeners' responses in our next iteration of the speech visualization system. Additionally, we would like to implement psycho-physiological measurements such as head-nodding or eye-blink rate to gain a comprehensive assessment of the participants' levels of understanding during the discussion.

3.5. Communi-gate: Visual Speech Facilitator

Image by LooseDrawings[5]

Figure 3.14 Overview of Communi-gate: Visual Speech Facilitator

To bridge the gap in understanding NNS' listening comprehension, we are proposing Communi-gate, a visual speech facilitator that helps NS to adjust their speaking behavior to enhance NNS' comprehension (See Fig.3.14)[5]. The purpose of our concept is to give live feedback to the NS prompting them to change the speaking style (slow down, use shorter utterances, simpler language, etc.) when NNS show difficulty in understanding during an online conversation.

Final Goal

From the concept building process, it appeared that by creating a system to visualize the perceptual experience of NNS, it evoked understanding in NS and brought changes in their behavior. Moreover, in response to the shift to online classes due to the pandemic, we developed the Bubble Visualization Overlay, a system that visualizes the dialogue characteristics of speakers and projects them on the Zoom screen. However, the results of the experiment proved that feedback from listeners was more important in motivating speakers to change their speaking style than the visualization of their own speaking style. This showed that it is important to integrate the listeners' feedback in order to encourage psychological and behavioral changes in the speaker.

5 https://loosedrawing.com/

From these findings, the integration of listener feedback and visual filter effects is the next step to this online communication support facilitator. Thus, the ultimate goal is to design visualizations that allow NS to empathize with how NNS are able to understand their conversations and adjust their speaking style. In my study, I aim to use visual information to intervene in people's speaking behavior by triggering psychological changes. The design of visual effects is based on the following two concepts.

Principle Phenomena

Concept.1: Generate Psychological Change in Speakers by Hiding the Listeners' Facial Expressions

Facial expressions can express an individual's emotions or indicate their intentions in social situations [40]. For example, in the change detection task of Ro et al. [41], participants were sensitive to facial changes. From facial expressions, we can understand the other person's emotions and how much the other person is concentrating on the conversation. Hence, the interpretation of the other person's face is very important for building a relationship between the listener and the speaker in a conversation.

Based on this theory, it can be said that facial expression is an important determinant for the speaker to distinguish how much the listener understands the conversation. From this finding, I hypothesized that *it could create an opportunity for the speaker to pay attention to the listener's level of understanding by hindering the other person's facial expressions.* For example, by adding changes to the brightness, resolution, and visibility of facial expressions, it may be able to change the listener's emotional state as interpreted by the speaker (See Fig.3.15). Therefore, I aim to provoke awareness in the speaker to consider the listener's level of understanding by making it difficult to grasp the other person's facial expressions with visual effects.

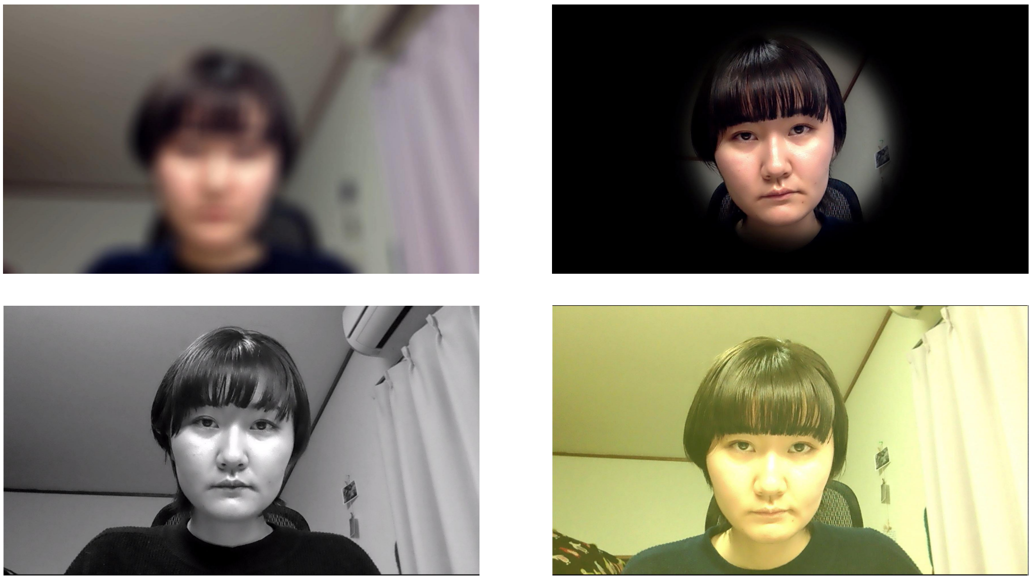


Figure 3.15 Visual Effect Samples of Facial Expression Modification

Concept.2: Generating Artificial Network Disruptions to Interference with Speech

Delayed Auditory Feedback (DAF) is a well-known method of external intervention in speech in which the speaker's voice is delayed by about 200ms and fed back to the speaker. Inhibiting speech without distraction or physical discomfort, the conversational inhibition effect of DAF has the advantage that it does not stop the speech completely. Nishida [4] conducted an experiment on the effects of DAF function in conversations between English proficient students and English learners. This study showed that the use of DAF caused a decrease in speech rate and interruptible pauses.

There are studies that show that the speed of another person's movement can automatically affect the timing of the observer's execution of the action [42]. The study of Ichikawa [43] also reported that the reaction time to visual stimuli is affected by the speed of the image observed before the measurement. Depending on the playback speed of the image to be observed, the observer's behavior immediately afterward becomes faster or slower. Another study also reported that delayed visual feedback, when combined with delayed auditory feedback, is highly effective in influencing speech fluency [44]. These studies suggest that the tempo of perception and action is influenced by the environment and by stimuli from other people.

From this point of view, I hypothesized that *by slowing down the playback speed of the listener's screen, it would be possible to slow down the speaker's speaking speed.* For example, when you are talking to someone via video chat, if the other person's internet connection goes down, you may stop talking once or start wondering if your voice is reaching them properly. Using this daily practice as a reference, we wondered if it would be possible to intervene in the speaker's speech by intentionally causing this phenomenon (See Fig.3.16).

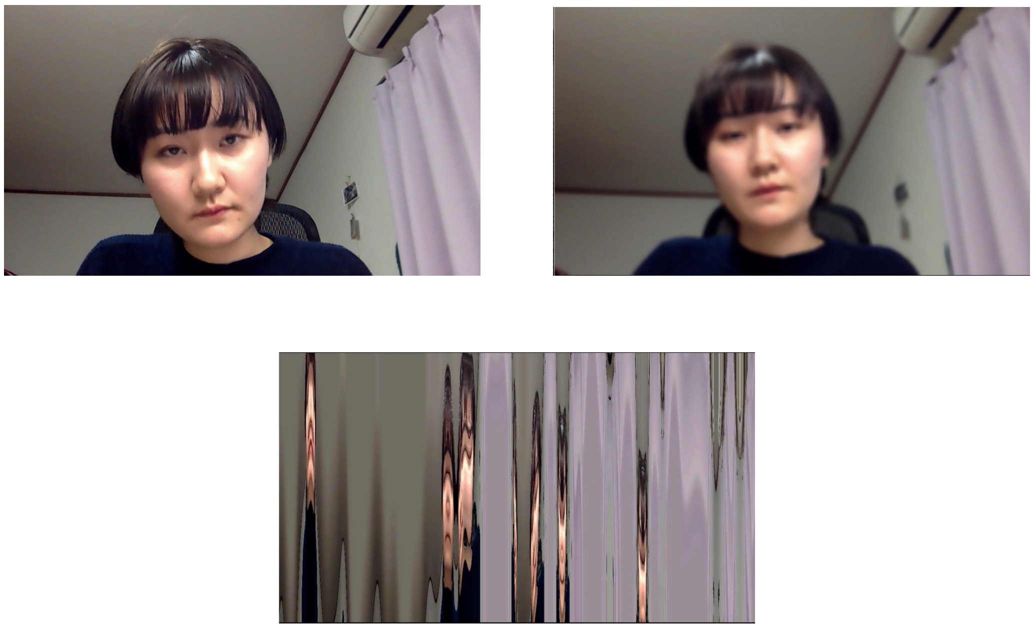


Figure 3.16 Visual Effect Samples of Artificial Network Disruptions

System

This visual speech facilitator aims to visualize the understanding level of the listener through visual effects, which are difficult to judge online, and to encourage the speaker to improve their speech in real-time. The system is configured to intervene as a third party on the network in a conversation between NNS and NS to represent the NNS' requests. As a result, it will help NNS to reduce their stress by repeatedly checking the understanding of the conversation.

NNS wear a device on their hands that measures their stress level. When the stress level exceeds a threshold, a visual effect is projected on the screens of all listeners. When the stress level goes down, the visual effect disappears from the screen (See Fig.3.17).

The following two methods are being considered for visual effects on NS: Blur effect and Slow-motion effect. Depending on the features of the visual effect, it can be selected or combined. With the blur effect, it aims to draw NS' attention to the fact that NNS' comprehension is declining, and with the slow-motion effect, it aims to demand improvement in the way NS speak.

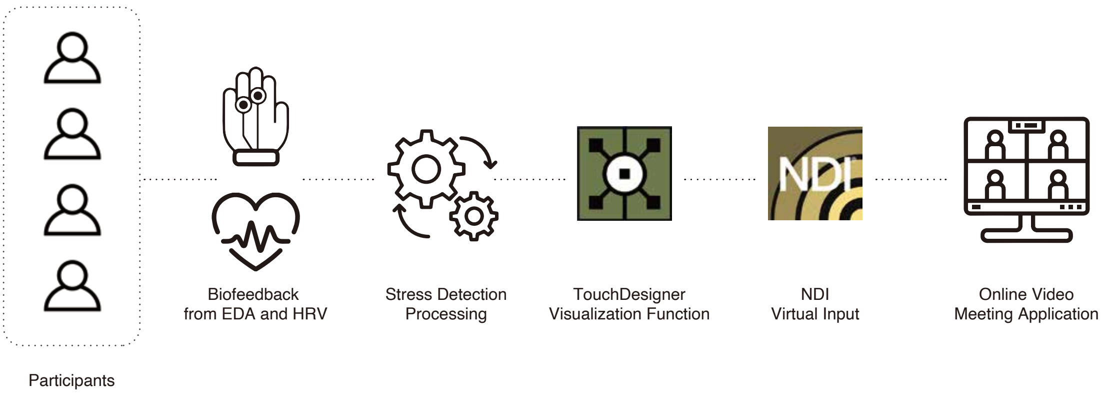


Figure 3.17 The Overview Structure of the System

Anticipated Experience

When the NS notice these visual effects, behavioral changes can be expected, such as checking NNS' comprehension, making a pause, slowing down the speaking speed, asking a question, repeating the word with rephrasing it to a simpler term, and summarizing the story more briefly.

3.6. The Potential Application

Communi-gate is intended to be used during online group discussions and presentations when international students are present. Facilitating remote communication between people of different languages is expected to promote co-creation, smoother communication, and forming stronger relationships in remote environments. Therefore, it will be important to have such a facilitation tool, especially for universities and departments with a large number of international students. In addition, the introduction of online facilitation tools can lead to the effectiveness of online lectures, especially in foreign language classes and classes with a mix of international students. It allows for high-quality instructions tailored to each student's level of understanding.

It can also be used in other situations, not limited to the educational field. For example, it is useful in international academic conferences where discussions are held in different fields and languages. In fact, Yamashita [3] mentions that Japanese researchers' performance at international conferences is limited by language barriers. It may create an opportunity for researchers who have not been able to speak out due to language barriers to share their knowledge. In daily life, it can also be recommended for online medical services for foreigners. According to Teraoka et al.'s research [45], many foreign residents in Japan have experienced difficulties in fully understanding doctors' instructions or in communicating their symptoms with doctors. Hence, the introduction of Communi-gate will enable smoother communication and sharing of information between foreign patients and doctors (See Fig.3.18) [6][7].

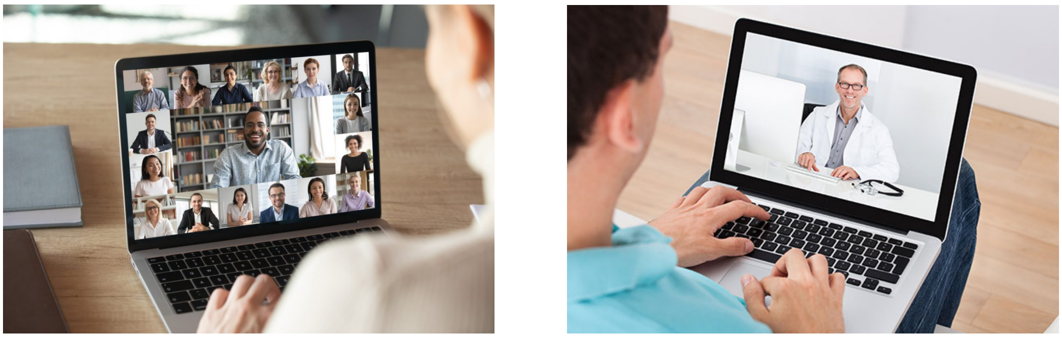


Photo by MoneyCashers[6] and MEDPAGETODAY[7]

Figure 3.18 Other Application Scenarios

6 https://www.moneycrashers.com/online-virtual-conferences-pros-cons/

7 https://www.medpagetoday.com/resource-centers/osteoporosis/assessing-telemedicines-pros-and-cons/697

Chapter 4
Proof of Concept

4.1. Overview of the Concept

To evaluate Communi-gate, we conducted a 5-part study with the following 3 objectives.

- Observing biological responses during human performance of task cues in listening comprehension (2 settings: pre-recorded, real-time)
- Evaluating of visual effects that induce psychological stimuli and regulate conversational behavior (2 settings: pre-recorded, real-time).
- Assessing the effects of inducing psychological stimuli and regulating conversational behavior in conversation

The experiments from Chapter 4 onwards were conducted as a joint project with members of Geist (Han Jiawen, George Chernyshov, Zhuoqi Fu, and Kai Kunze). I also received the cooperation of Prof. Naomi Yamashita (NTT Communication Science Labs) and Chi Lan Yang (Ph.D., The University of Tokyo) for the experiment design.

4.2. Technical Implementation

4.2.1 Visual Effect Design

Prerequisites in TouchDesigner

We used TouchDesigner, a visual programming software, to create our visual effects. TouchDesigner would take video from Webcam through Video Device

TOP and output through NDI out TOP. Then we use NDI Virtual Input, which streamed the visuals to an online video conference software, Zoom.

For activation of the visual effect, Serial CHOP is used to connect the Arduino and set it up to be controlled by an external button. For the implementation of the final experiment, a change was made to introduce an internal switch in TouchDesigner instead of an external switch by Arduino to activate the visual effects (See Fig.4.1). For the conversion of function values, we used Math CHOP to output as ON(1)/OFF(0). Also, for function values in Math CHOP, Bind CHOP was used to set each function area.

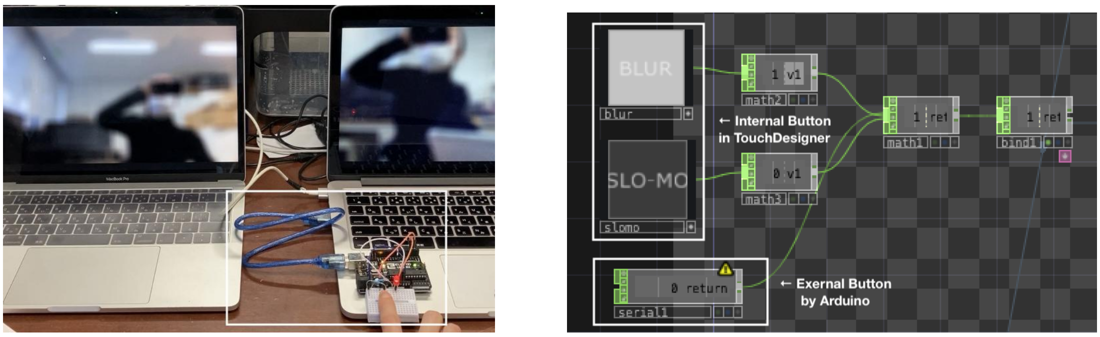


Figure 4.1 Activation Setting of the Visual Effect

TouchDesigner Patch Setting per Visual Effect

TouchDesigner PATCH details for each effect are attached in the Appdendix.

- **Frame Drop, Freeze**: Adjust the playback speed of the image using the FPS (Frame per second) function of the NDI output TOP by setting Freeze to 30 (0) and 0 (1) and Frame Drop to 30 (0) and 2 (1) in Math CHOP, and reflect it in the FPS value in Blind CHOP.

- **Blur**: Blur TOP is used, and Math functions 1(0) and 30(1) are assigned to the Filter Size of Blur TOP in the Blind CHOP.

- **Slow-motion**: Slow-motion integrates the above Frame Drop and Blur configurations, and assigns the Math CHOP function 30(0), 2(1) to the FPS value of the NDI input TOP using the Blind CHOP. Similarly, the Math

CHOP functions 0(0) and 5(1) are assigned to the Filter Size of the Blur TOP using Blind CHOP.

- **Glitch**: Assign the function 0(0), 7(1) in Math CHOP to LFO and Null as "frequency" using Blind CHOP. The frequency value is then assigned to the Amplitude value in the Noise CHOP and converted to the noise of the video. Then, using CHOP to TOP, the Amplitude value in Noise CHOP is processed as TOP, and projected on the screen using Display TOP.

- **Black and White**: Using Monochrome TOP, the function 0(0), 1(1) in Math CHOP was integrated into Monochrome TOP using Blind COP.

- **Narrow FOV**: Generate a circle of scale 1×1 with Circle TOP. (4) Narrow FOV: Generate a circle with a scale of 1×1 using Circle TOP, set the softness to 0.2 and the softness Type to Hybrid Cubit to slightly Blur the circular lines, and use Transform TOP to expand the circle from 1 to 3 (enough to cover the screen) depending on the output of an external switch. Use Composite TOP (operation type: Multiple) to project the input video from Video Device In TOP onto the generated circle. In order to project the full screen (radius of circle = 3) when the switch is off and a circle of radius 1 when the switch is off, the Math CHOP function is set to 3(0), 1(1).

- **Flash**: Project a color (yellow: 0.92618, 1, 0.141) on the screen using Circle CHOP, and assign the functions 0(0) and 2(1) in Math CHOP to LFO CHOP and Null CHOP as "projection frequency" using Bind CHOP. This reproduces the Flash effect by projecting the yellow image in small increments.

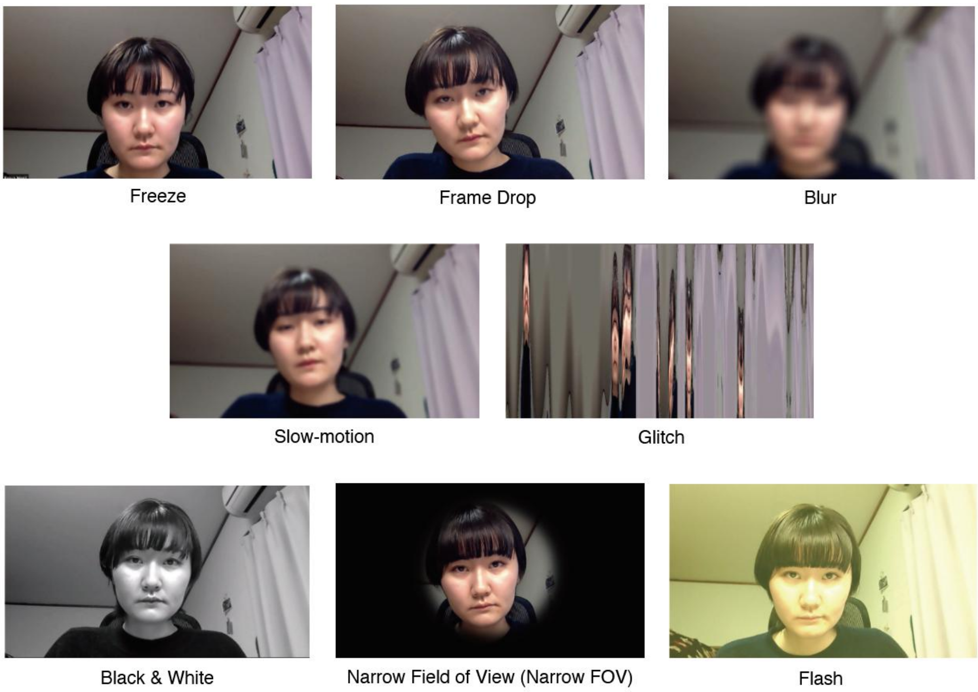


Figure 4.2 Visual Effect Examples

4.2.2 Effect Streaming Setup

The configuration and data flow of this system is shown in Fig.4.3. Each PC opens the TouchDesigner file with the IP address of the computer on the server side. The server re-streams the visual effect activation (ON/OFF) to the other computers it controls under a common IP address. Likewise, the participant's physiological data will also be sent to the server of another computer via WiFi. The participant's physiological data is stored locally on the server-side computer.

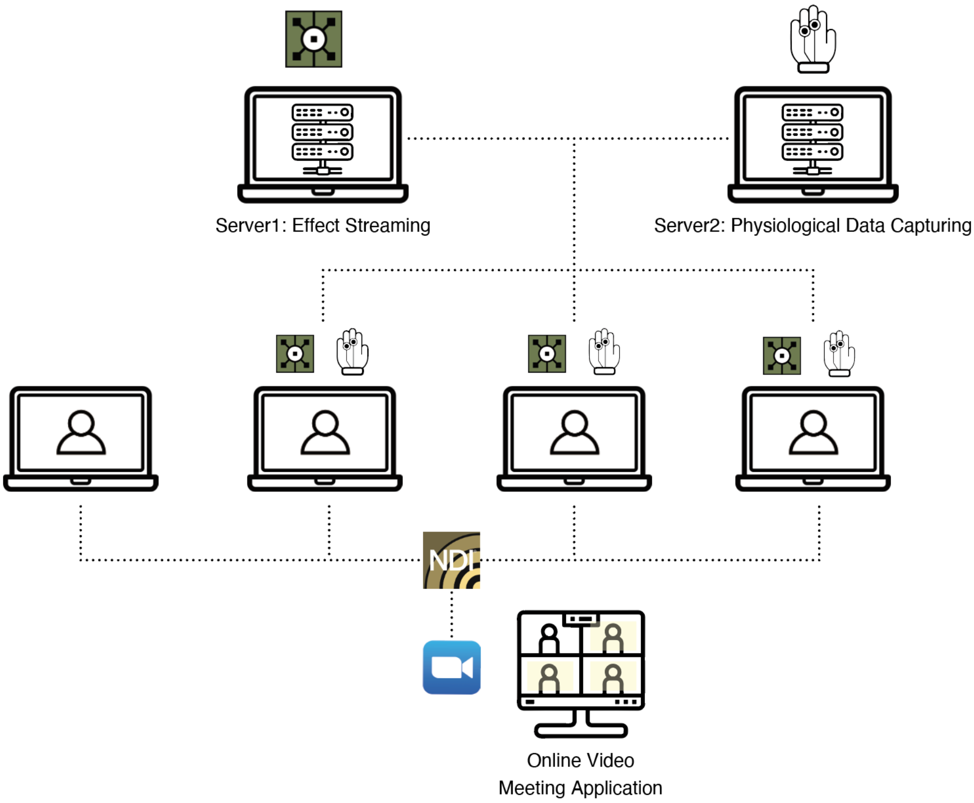


Figure 4.3 The Configuration and Data Flow

In order to connect Touchdesigner to the server and control multiple computers remotely, we used Data OPerators (or DATs). DATs are used to hold text data and we include multiple lines of text as in a script that allows patch information of TouchDesigner to be exchanged between computers connected to a common IP address (See Fig.4.4). (The script is attached to Appendix.) The activated visual effect is projected onto the screen of the Online Video Meeting Application, ZOOM, via the NDI Virtual Input.

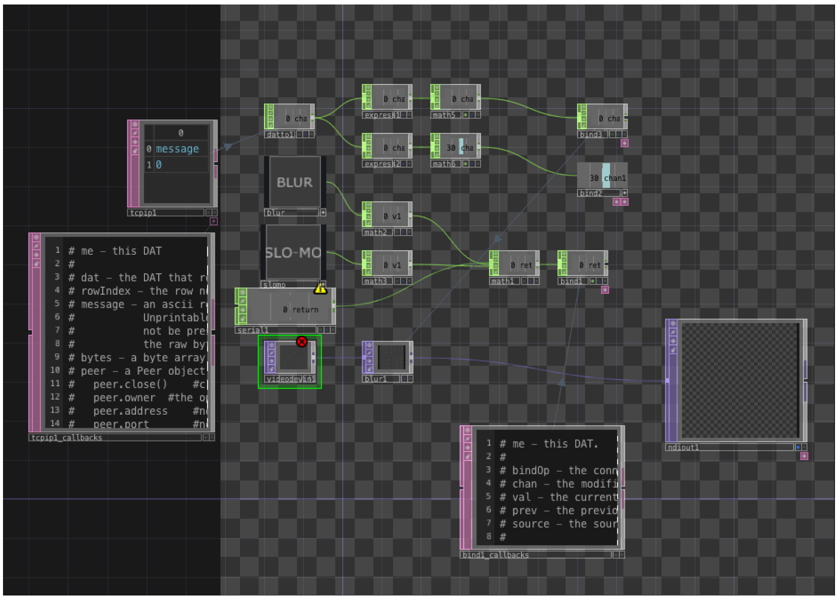


Figure 4.4 TouchDesigner Patch for Effect Streaming

The physiological data of participants was captured using custom built smart wristbands that introduced by Shugawa et al [46] (See Fig.4.5). The physiological data includes heartbeats, EDA, and the LF/HF ratio related to the heart rate variability. The device was designed to be worn on the wrist, measuring EDA from two electrodes on the finger and heart rate using an optical blood volume pulse sensor.

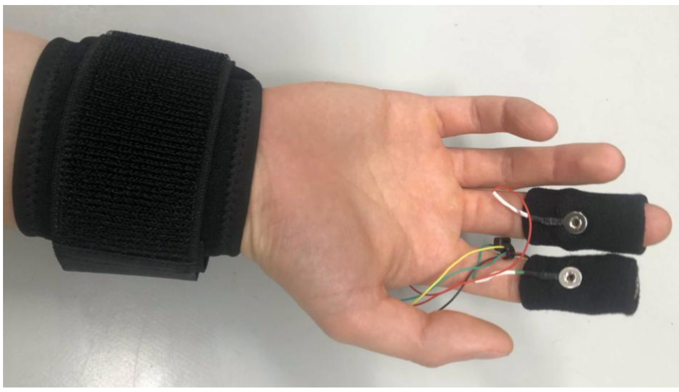


Figure 4.5 Wristband with Electrodermal and Heart Activity Sensors

4.3. Preliminary Study: Biological Indicators of Listening Comprehension

4.3.1 Study Design

As a basis for this implementation, we conducted an initial study to investigate the relationship between psychophysiological measurements and the listener's comprehension. In recent years, the use of multiple physiological information (such as heart rate and skin conductance) to quantify listening effort during spoken communication has received much attention [47–50]. In this study, we conducted experiments in two different settings (Phase.1: pre-record, Phase.2: real-time) to compare and evaluate the stress and increased listening effort associated with communication in a foreign language using heart rate and skin conductance. We used a within-subjects study design to compare participants' physiological signals with listening comprehension.

We will focus on these 3 points below:

- Investigation of the changes in physiological signals during a series of English listening comprehension tasks
- Investigating the relationship between the psychophysiological readings and listener's comprehension level rather than just task-induced stress estimation
- Collecting data labeled with the objective comprehension level and subjective mental demand for the implementation of the next stage of the envisioned assistance tool

Our Hypotheses are

- *"The higher the difficulty of the listening test, the higher the sweating and heart rate values."*
- *"If the correct answer rate of the question is low, the sweating/heart rate values increase."*
- *"People with a low English proficiency level have significantly higher psychophysiological values than those with a high level."*

Method: Phase.1 (pre-recorded)

In phase.1, as the first step for our system implementation, we need to explore the listeners' physiological response when they are performing English listening exercises of varying difficulty. The listening materials to elicit physiological changes were from British Council and represented CEFR level A1(Beginner), B1(Intermediate), and C1(Advanced). These levels are defined in the Common European Framework of Reference (CEFR) by the Council of Europe[1].

Method: Phase.2 (real-time)

In phase.2, to recognize listening effort and stress in real-time settings, we plan to collect our data in a longer time for a more precise prediction. Participants in the experiment will be asked to record the timing of their failure to understand as a time stamp using an external button while watching a video in a foreign language. The listening material used was a TED talk[2]. From a list of three of each of the most viewed in 5 representative categories (Technology, Design, Business, Science, and Global Issue), each subject was asked to select 3 videos of his or her choice to watch.

4.3.2 Procedure: Phase.1 (pre-recorded)

In total, 14 adult participants (7 Japanese, 6 Chinese, 1 Indian - NNS, 8 males 6 females, MODE= 24-26) took part in this study. The experimental setup was in accordance with the Ethics-board and rules of blinded experiment.

Participants were asked to wear and use equipment provided by the researcher. We were collecting different types of data from them through sensing devices such as skin conductance, heart activity, and facial expressions. Heart rate and skin response was assessed from wearable wireless sensory device placed on the fingers of non-dominant hand. Facial expressions was assessed using web camera of the laptop (See Fig. 4.6).

1 https://learnenglishteens.britishcouncil.org/skills/listening

2 https://www.ted.com/talks

Figure 4.6 Preliminary Study (Phase.1): Experimental Setup

Participants were asked to complete a self-reported online linguistic background questionnaire to assess their level of English proficiency and language usage. The questionnaire was designed by adapting questions from these sources: the Language Experience and Proficiency Questionnaire [51]. The study typically took 8 minutes (per listening task) with 1 minute for a short questionnaire of NASA Task Load Index2 between the tasks. In total, 3 listening comprehension tasks took part (See Fig.4.2).

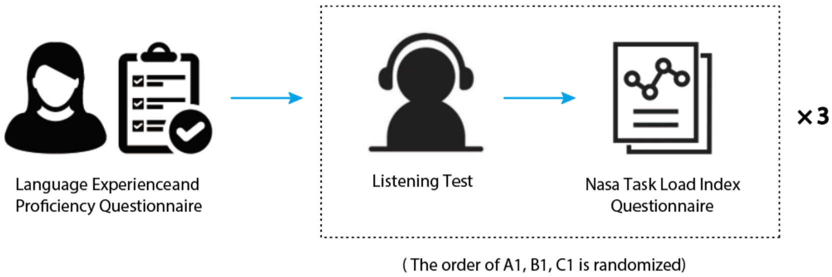


Figure 4.7 Experiment Procedure

4.3.3 Results: Phase.1 (pre-recorded)

Data Analysis

Results1: Analysis of Mental Demand Differences over Listening Tasks

We compared paticipants' mental demand differences from Nasa Task Load Index and answer accuracy differences among A1, B1, and C1 sessions. According to the pairwise comparison with the Bonferroni correction, significant mental demand differences exist between A1($M = 2.23$,$SD = 1.09$) and C1($M = 5.31$, $SD = 1.38$) ($p <.001$), B1($M = 3.54$, $SD = 1.81$) and C1 ($p = .003$). We used the answer accuracy as an objective measurement of paticipants' understanding level. We assigned accuracy rate above 85% as high accuracy for A1, 80% for B1, and 60% for C1, which are average scores across all paticipants for each session. From this analysis, it can be seen that a significant difference in mental demand and percentage of correct answers was generated by C1 compared to the other A1 and B1.

Results2: The Relationship between the Psycho-physiological Indications and Listener's Comprehension Level

Our experimental findings suggest the potential of using HR features to predict listeners' objective listening accuracy as well as differences of EDA dynamics across listening sessions in varying levels of difficulty.

Figure.4.8 shows HR features' changes within A1, B1, and C1 sessions. Although there are no significant differences observed among three trials overall, obvious changes occur near the end of each session when paticipants were answering questions after each listening material. It might be because of the fact that answering questions requires paticipants to actively think and connect what they had heard before, in contrast to the listening period. And this could result in more obvious physiological differences between mental states when paticipants can understand the materials and when they can not.

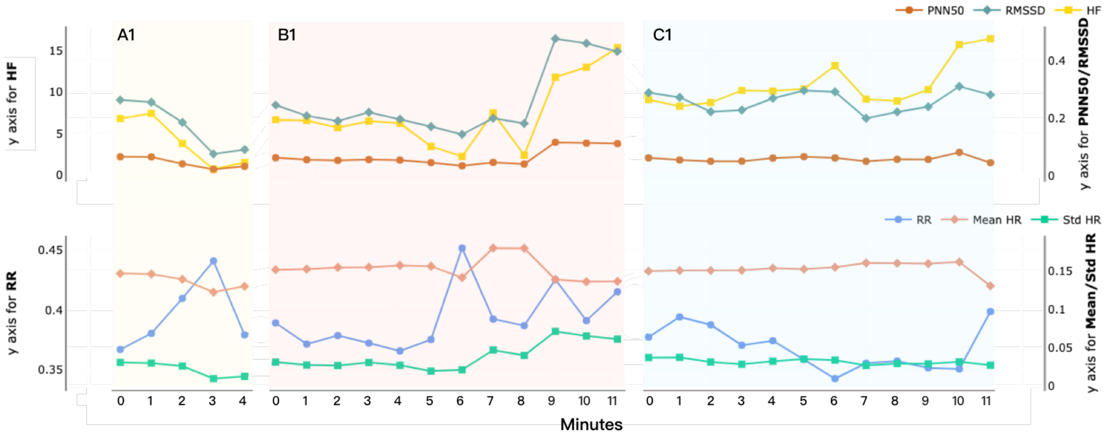


Figure 4.8 Changes of key HR and HRV features used in this study across three trials (A1 session:yellow highlight; B1 session: red highlight; C1 session: blue highlight).

EDA Data is presented in Fig.4.9. Each bar represents the total number of local maxima registered in each minute aggregated over all the participants. Comparing the numbers of spikes per bin, ANOVA test showed p=0.005 for all three levels. Which leads us to think that the distribution of spikes is not random. For A1 vs C1 p=0.002; for B vs C p=0.0645; for A vs B p=0.145, which shows that the EDA activity patterns are significantly different between the easiest and most difficult levels. Differences between adjacent A1 and B1 levels as well as B1 and C1 levels are noticeable, but short from significant.

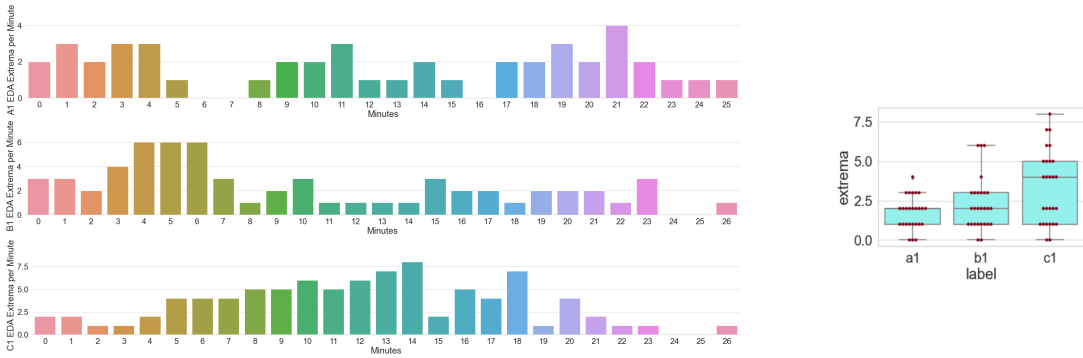


Figure 4.9 Distribution of detected spikes binned by minute for each of the difficulty levels.

4.3.4 Discussion: Phase.1 (pre-recorded)

The data collected in this study was recorded in the laboratory over a relatively short period of time. In addition, the listener's comprehension is derived from the percentage of correct answers to the listening questions, and does not define the moment when the listener does not understand. Therefore, in the following work, we plan to collect data in a way that extracts the moments when listeners do not understand in real time by using time stamps and other methods over a longer period of time in order to make more accurate predictions.

4.3.5 Procedure: Phase.2 (real-time)

In total, 17 adult participants (12 Japanese, 5 Chinese - NNS, 10 males 7 females, MODE= 24-26) took part in this study. The wristband device used in this experiment is the same as the one used in Phase.1.

Participants were asked to complete a self-reported online linguistic background questionnaire that we used in Phase.1. The study typically took 12 minutes (per listening task) with 1 minute for a short questionnaire of NASA Task Load Index2 between the tasks. In total, 3 listening comprehension tasks(3 Ted talks) took part. Participants were instructed to keep pressing the button until they understood if they started feeling difficult in following the speech while watching the video (See Fig.4.10).

Figure 4.10 Preliminary Study (Phase.2): Experimental Setup

4.3.6 Results: Phase.2 (real-time)

Data Analysis

Results1: The Relationship between the Psycho-physiological Indications and Listener's Comprehension Level

The following is an example of EDA and HR features for all paticipants over 3 listening comprehension tasks (See Fig.4.11, 4.12) The area marked with grey marker in both EDA and HRV is the time period when the paticipants pressed the button, which suggests when paticipants could not understand the talk or had difficulty in following the talk. Data of some paticipants with lots noises was removed.

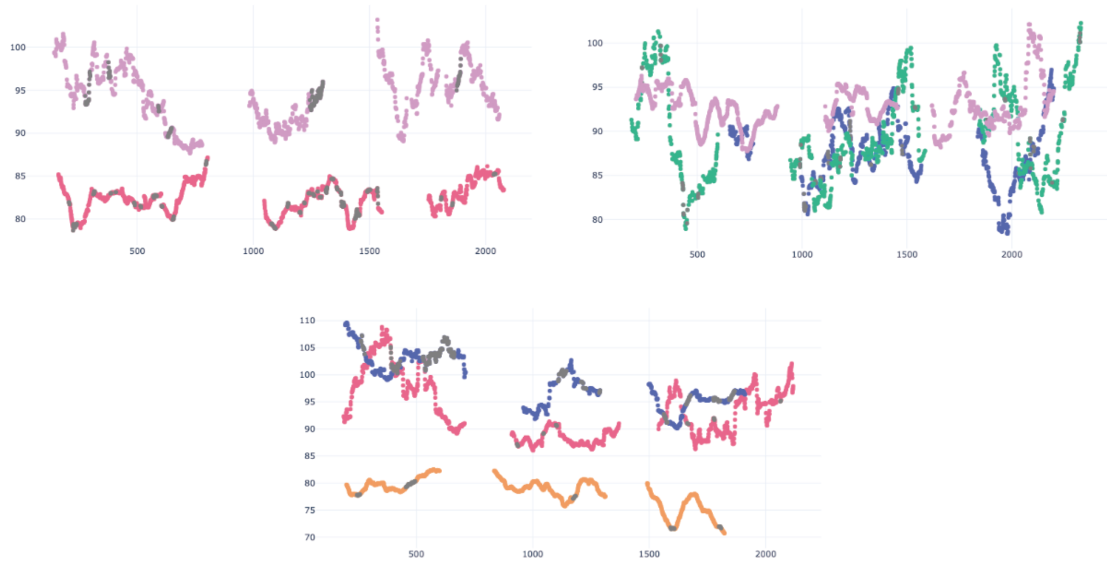


Figure 4.11 Changes of HR features

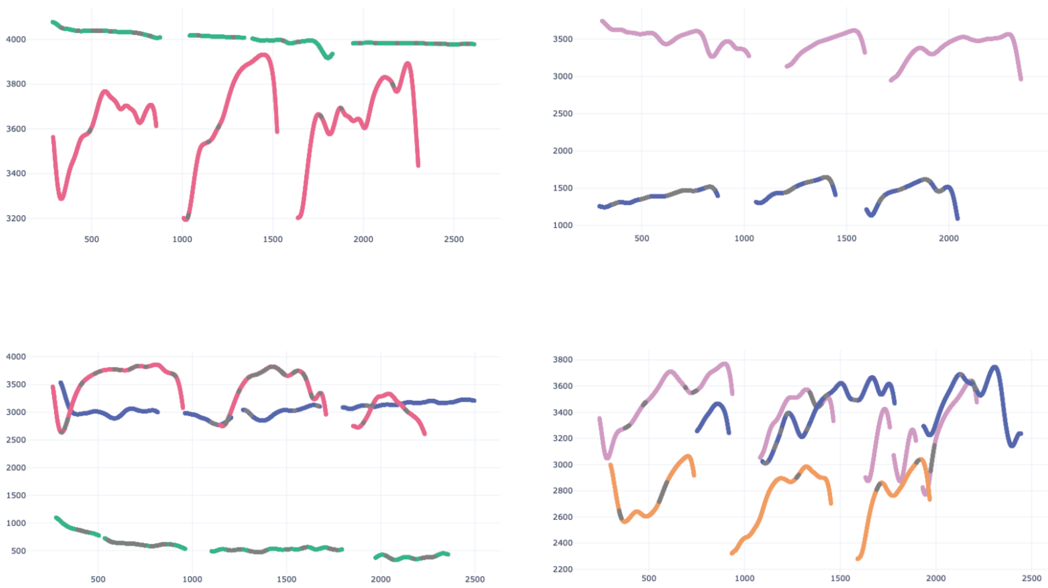


Figure 4.12 Changes of EDA features

Qualitative Analysis

Results2: Judgment of Listening Comprehension

From responses of participants during the experiment, we found that the degree to which listeners judged their own "understanding" varied from person to person, depending on the degree to which they considered themselves not to have understood. For example, there were several patterns: pressing when they did not understand a word, pressing when they could not follow the context, and pressing when they could not understand the whole story after listening to it. In addition, there was a lag between the time when they realized that they did not understand and the time when they expressed their intention and pressed the button.

4.3.7 Discussion: Phase.2 (real-time)

As a result of the experiment, we found that it is difficult to prove the degree of understanding of the listener based on paticipantive judgment from biological responses. Moreover, responses of participants during the experiment, we found that the degree to which listeners judged their own "understanding" varied from person to person, depending on the degree to which they considered themselves not to have understood.

Based on these results, we realized that it is difficult to define a certain threshold, because the paticipantive criterion of "comprehension level" would produce separate individual differences and variations in the results. Therefore, in this final experiment, we decided not to employ integration with biological responses.

4.4. Study.1: Visual Effect Evaluation

4.4.1 Study Design

In this study, we conducted an evaluation experiment of 8 visual effects. The visual effects were designed based on 2 concepts in Chapter 3.5. These are visual effect filters that will be projected on the user's screen during online meetings such as Zoom, with a total of 8 types: Freeze, Frame Drop, Blur, Slow-motion, Glitch, Black & White, Narrow Field of View(NarrowFOV), Flash (See Fig.4.13).

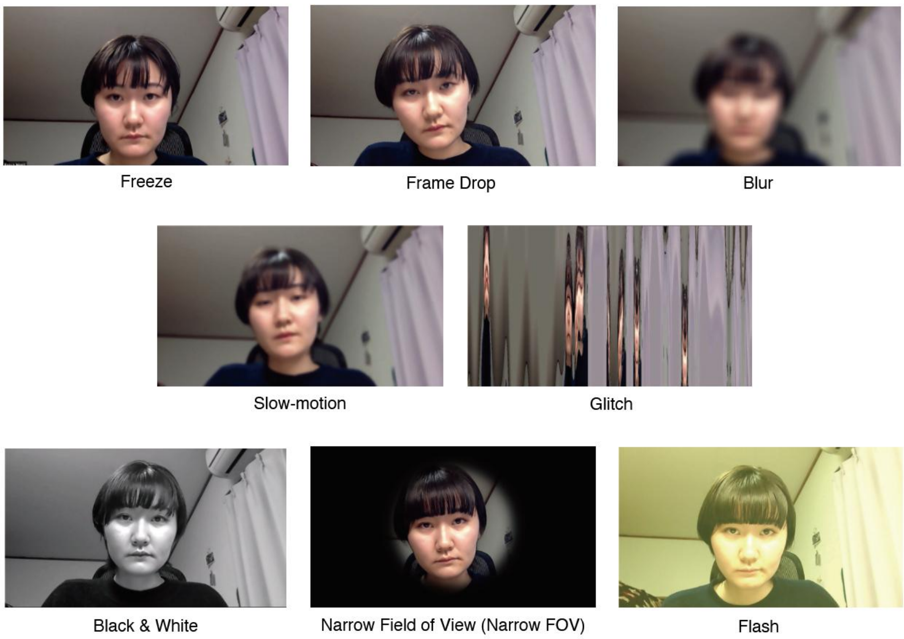


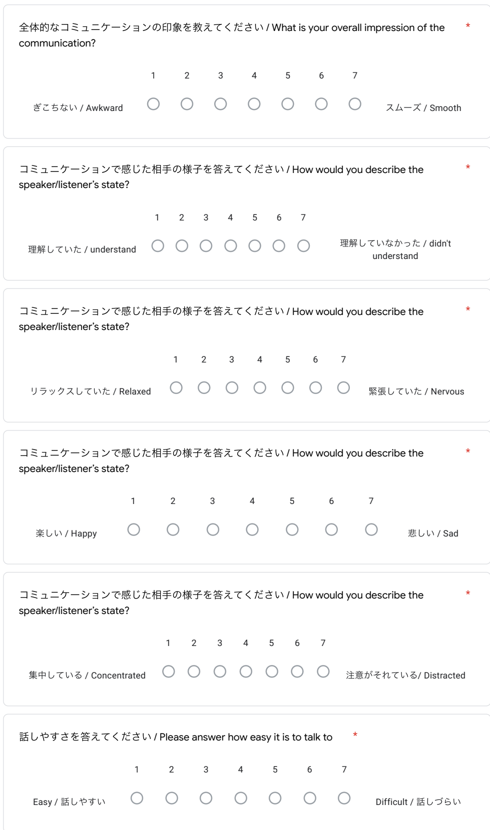
Figure 4.13 Study.1: Visual Effect Examples

Our Hypothesises are

- *"When the listener's face is obscured by the Blur filter, the speaker feels psychologically distant from the listener and becomes anxious."*
- *"If the listener's screen movement slows down (Frame-drop, Slow-motion, Freeze), the speaker feels less comfortable to keep talking."*

Method

In order to evaluate visual effects, we followed the the Semantic Differential (SD) method as done with subjective assessment of the mental state of a participant through a questionnaire that is composed of pairs of opposing adjectives, each of which is evaluated on a 7-point scale [52]. In our study, the 6 adjective pairs "smooth – awkward," "understand – didn't understand," "relaxed – nervous," "happy-sad," "concentrated - distracted," and "easy to talk - difficult to talk" were used to evaluate the extent of feeling from a speaker over the conversation with visual effects. These questions aim to assess the following four criteria. (1) the overall impression of the conversation, (2) the interpretation of listener's comprehension, (3) the interpretation of listener's emotion (4) the impact on ease of speaking (See Fig.4.14).



全体的なコミュニケーションの印象を教えてください / What is your overall impression of the communication? *

1 2 3 4 5 6 7

ぎこちない / Awkward ○ ○ ○ ○ ○ ○ ○ スムーズ / Smooth

コミュニケーションで感じた相手の様子を答えてください / How would you describe the speaker/listener's state? *

1 2 3 4 5 6 7

理解していた / understand ○ ○ ○ ○ ○ ○ ○ 理解していなかった / didn't understand

コミュニケーションで感じた相手の様子を答えてください / How would you describe the speaker/listener's state? *

1 2 3 4 5 6 7

リラックスしていた / Relaxed ○ ○ ○ ○ ○ ○ ○ 緊張していた / Nervous

コミュニケーションで感じた相手の様子を答えてください / How would you describe the speaker/listener's state? *

1 2 3 4 5 6 7

楽しい / Happy ○ ○ ○ ○ ○ ○ ○ 悲しい / Sad

コミュニケーションで感じた相手の様子を答えてください / How would you describe the speaker/listener's state? *

1 2 3 4 5 6 7

集中している / Concentrated ○ ○ ○ ○ ○ ○ ○ 注意がそれている/ Distracted

話しやすさを答えてください / Please answer how easy it is to talk to *

1 2 3 4 5 6 7

Easy / 話しやすい ○ ○ ○ ○ ○ ○ ○ Difficult / 話しづらい

Figure 4.14 Study.1: Questions in the Questionnaire

4.4.2 Procedure

In total, 20 adult participants (13 female and 7 male) took part in the study. The native language of the paticipants was a mixture of both Japanese and English.

In the experiment, we used a within-subjects study design to compare 9 effects in total with including a session of no effect. No effect is the beginning as a baseline and the order of the rest effects appearance is counterbalanced in order to minimize ordering effects. A speaker was instructed to speak to the screen of a pre-recorded video for 30 seconds. The study typically took 30 minutes in total with 1 minute for a short questionnaire between sessions. After all the sessions were completed, additional interviews were conducted with the paticipants about "in which session you had the most difficulty in continuing to speak" for 5 minutes. The overview of the process is as shown in the following diagram (See Fig.4.15).

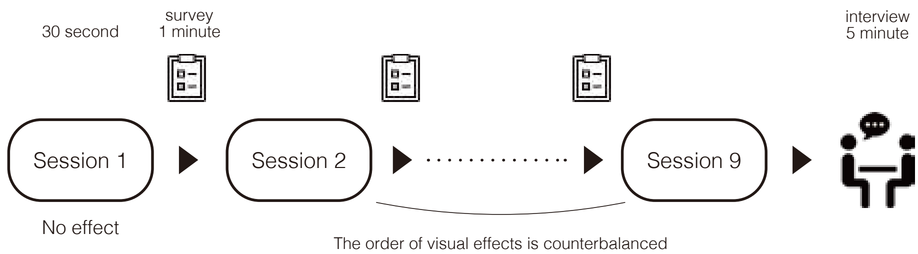


Figure 4.15 Study.1: Experiment Procedure

4.4.3 Results

Data Analysis

Results.1: The Extent of Feeling from a Speaker over an Online Speech Experience with Visual Effects

A one-way repeated measures ANOVA across the 8 effects (Freeze, Frame Drop, Blur, Slow-motion, Glitch, Black & White, NarrowFOV, Flash) was conducted. The results of SD method are shown below (See Fig.4.20). The following are the criteria for the t-test by ANOVA: (*p $<$ .05, **p $<$ .01, ***p $<$ .001.)

- **Overall Impression of the Conversation (Fig.4.16)**

Participants responded highly with "awkward" to the Glitch effect. We found a significant main effect (F(7, 133) = 3.890, p <0.001, η_p^2 = 0.170). Holm's post hoc test indicated that Glitch was significantly smaller than NarrowFOV and Black and White (NarrowFOV: t(19) = 4.420, adj.p = 0.008, Black and White: t(19) = 3.873, adj.p = 0.027)

- **The Interpretation of Listener's Comprehension (Fig.4.17)**

  Participants responded highly with "not understand" to the Glitch effect. (F(7, 133) = 3.890, p <0.001, η_p^2 = 0.218). Holm's post hoc test indicated that Glitch was significantly larger than Freeze (t(19) = 3.200, adj.p = 0.031)

- **The Interpretation of Listener's Emotion (Fig.4.18)**

  Blur, Slow-motion, and Glitch were relatively answered "sad" more than other effects. (F(7, 133) = 5.782, p <0.001, η_p^2 = 0.233). Holm's post hoc test indicated that Blur, Slow-motion, and Glitch were significantly larger than Freeze (Blur: t(6) = 6.475, adj.p = 0.000, Slow-motion: t(19) = 5.287, adj.p = 0.001, Glitch: t(19) = 4.781, adj.p = 0.003)

- **The Impact on Ease of Speaking (Fig.4.19)**

  In the case of Freeze, Slow-motion effects, it was reported that it was difficult to speak. (F(7, 133) = 3.890, p = 0.001, η_p^2 = 0.157). Holm's post hoc test indicated that Glitch and Blur were significantly larger than NarrowFOV (Glitch: t(19) = 4.125, adj.p = 0.016, Blur: t(19) = 4.032, adj.p = 0.019)

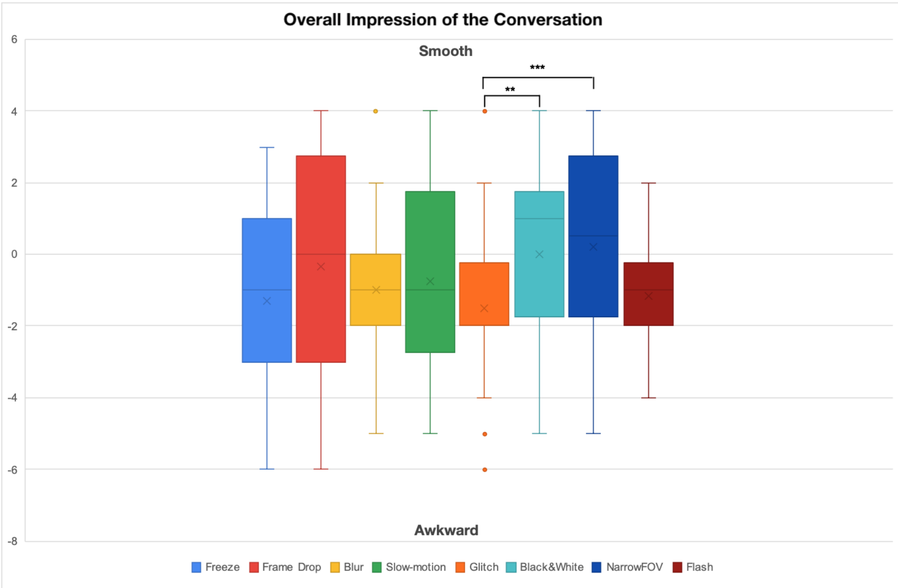


Figure 4.16 Overall Impression of the Conversation (Smooth-Awkward)

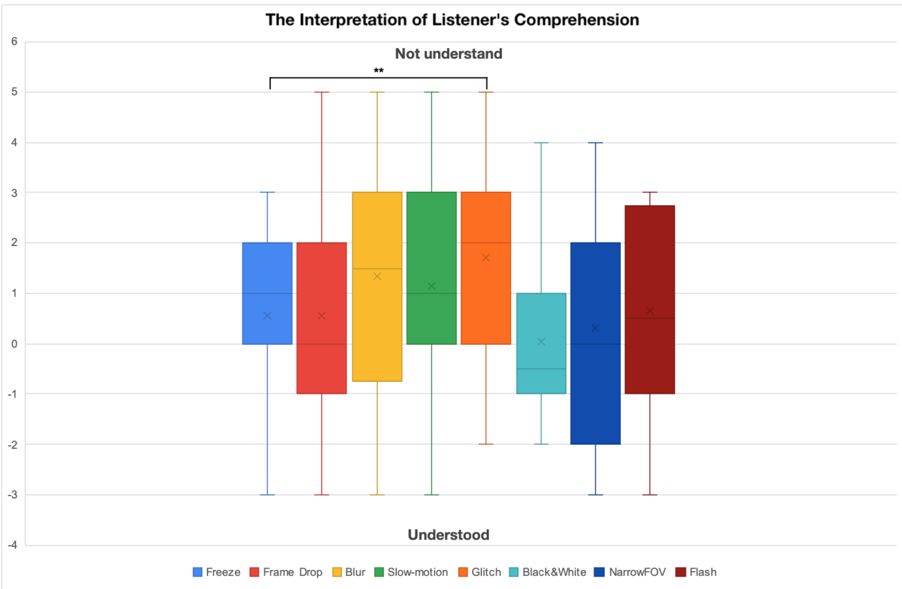


Figure 4.17 The Interpretation of Listener's Comprehension (Not understand-Understood)

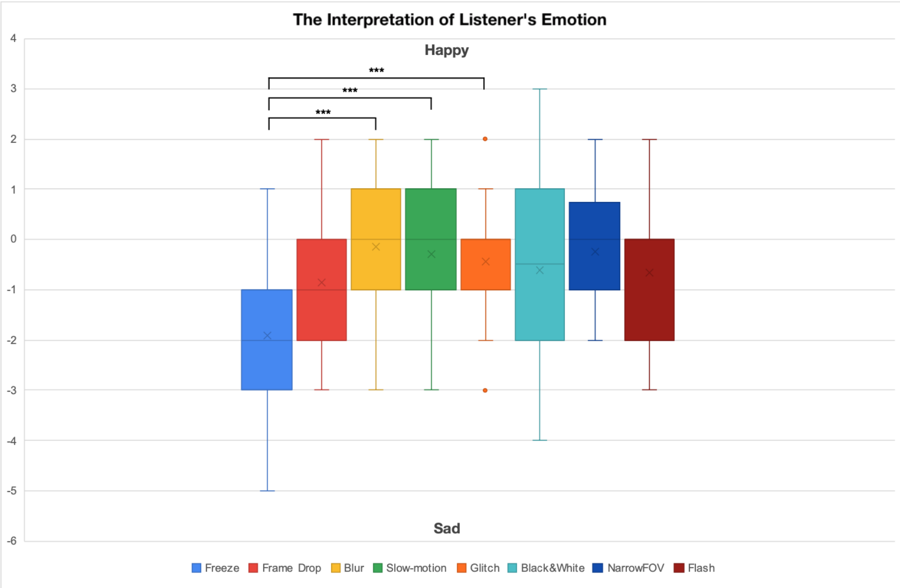


Figure 4.18 The Interpretation of Listener's Comprehension (Happy-Sad)

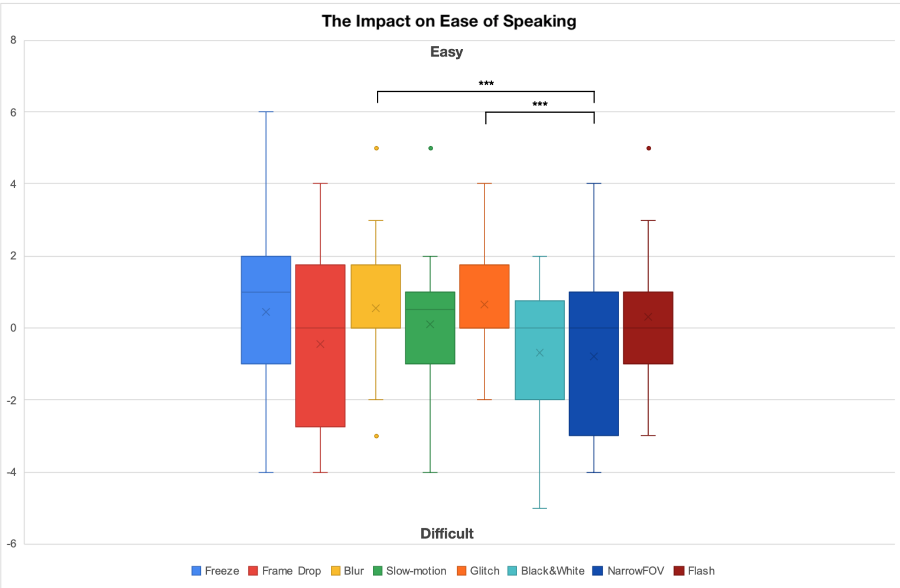


Figure 4.19 The Impact of Ease of Speaking (Easy-Difficult)

Figure 4.20 Study.1: Visual Evaluation Feedback Analysis

Qualitative Analysis

Results.2: Insights from Interviews

From the responses of the additional individual interviews, 7 of the respondents felt uneasy with Blur effect because they could not see the other person's expression. In addition, many participants responded that they thought it was an error in the Internet connection for Freeze, Slow-motion, and Glitch effects, and they were worried about whether the listener was hearing what they were saying well. In particular, those respondents also mentioned that they were used to seeing Freeze and Slow-motion in their daily online meetings, while some were surprised by Glitch because it was an effect they did not usually see.

4.4.4 Discussion

Results from this study indicate that Blur has the effect of making the speaker feel more distant from the listener. Freeze, Slow-motion, and Glitch were also able to arouse the speaker's attention through visual delays, as we intended. However, there was no noticeable change in behavioral change due to the manipulation of these visual delays. This was because the speaker was talking to the recorded video, so there was little incentive to change the way he/she spoke. From this finding, it is our next step to investigate what kind of psychological and behavioral changes the speaker will make in each of the Blur, Freeze, Slow-motion, and Glitch effect. Therefore, we decided to conduct another in-depth experiment to evaluate these four effects in a situation where the speaker is talking to the listener in real-time.

4.5. Study.2: Behavioral Change Induced by Effects

4.5.1 Study Design

In this study, we aim to investigate a speaker's actual reaction in real-time toward effects. The visual effects were selected: Blur, Freeze, Slow-motion, and Glitch effect (See Fig.4.21). In particular, our objective is to elucidate the characteristics of the speaker's behavioral change in Slow-motion, Freeze, and Glitch.

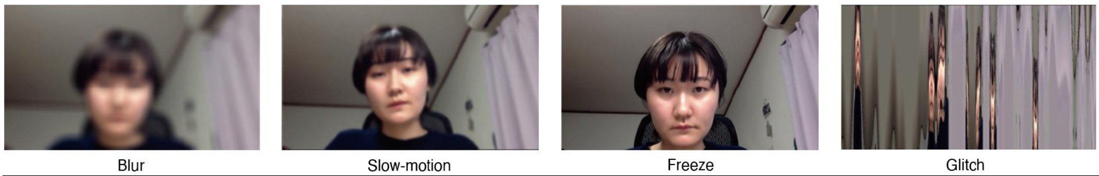


Figure 4.21 Study.2: Visual Effect Samples

Our hypothesises are

- *"When the listener's screen is still or changes due to Freeze or Glitch effect, the speaker stops thinking and several pauses are created."*
- *"When the listener's screen movement is delayed by slow motion, the speaker feels difficulty speaking and slows down."*

Method

We collected subjective survey measures after each session. First, manipulation check of speaker's awareness of visual effects. Speaker's awareness of each effect was assessed by a two-option question asking if they identify any visual effect displayed on any screen during the speech ("I wasn't aware" "I was aware of visual effect feedback") [3]. Second, subjective assessment of the extent feeling from a speaker over the conversation with visual effects. This time, the 3 adjective pairs: "understand – didn't understand," "concentrated - distracted," and easy to talk - difficult to talk" were used. Third, for the speaker's behavioral change, those who noticed the effect were asked a descriptive descriptive open-end question, "How did your consciousness or behavior change when the effect occurred?" We prepared a descriptive question for those who noticed the effects (See Fig.4.22.

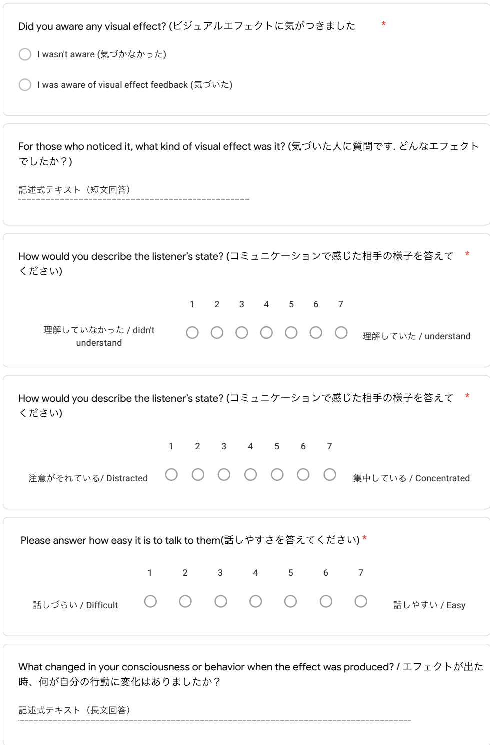


Figure 4.22 Study.2: Questions in the Questionnaire

4.5.2 Procedure

In total, 11 adult participants (8 female, 3 male) took part in the test. The native language of the paticipants was a mixture of both Japanese and English.

We used a within-subjects study design to compare an online speech experience with 5 experimental conditions: No effect, Blur effect, Slow-motion effect, Freeze effect, and Glitch effect. The order of visual effects is counterbalanced in order to minimize ordering effects. No effect was used at the beginning as a baseline.

The topics were chosen to be of general interest. For example, the usage of social media, online shopping, and Kindle. A template of three slides containing the pros, cons, and a potential personal verdict of the topic that presenters had to modify. Presenters had 1 minute for preparation in the beginning and then delivered 1 minute talk with looking at the screen of the confederate where a different visual effect feedback appears per session. The study typically took 30 minutes in total with 1 minute for a short questionnaire between sessions. After all the sessions were completed, additional interviews were conducted with the paticipants ("which session was the most difficult session to continue speaking," "what cues do they rely on to observe a listener's understanding," and "which session is the easiest and hardest to see the non-verbal cues from the listener") for 10-15 minutes. (See Fig.4.23, 4.24).

The Criteria to the Confederate

A confederate acted as a listener. The confederate was instructed to mute their microphones during the presentation and to not interrupt the presenter via other means such as chat messages. The confederate always saw the participant with an average-rated face.

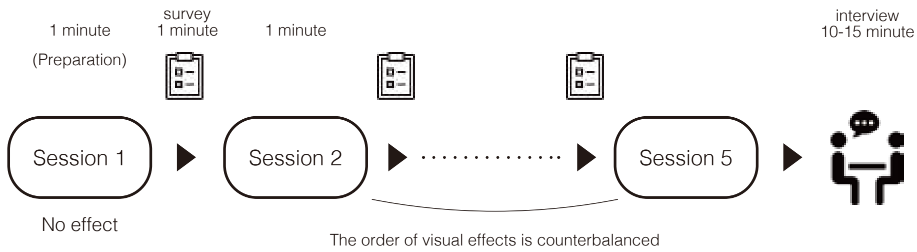


Figure 4.23 Study.2: Experiment Procedure

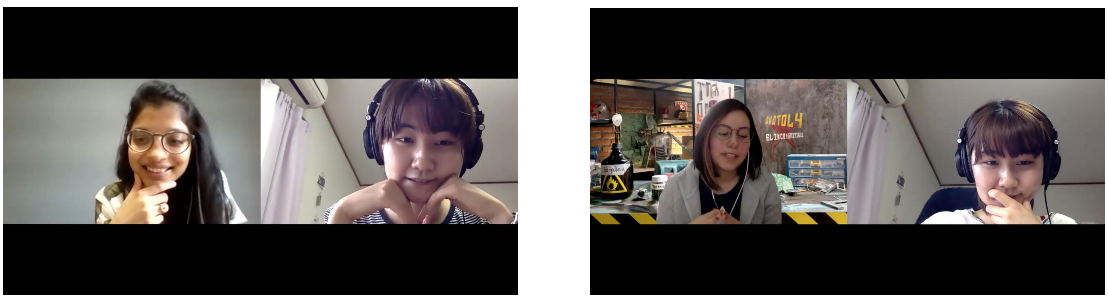


Figure 4.24 Study.2: Experimental Setup

4.5.3 Results

Data Analysis

Results.1: Speaker's Awareness of the Effects

Regarding the prominence of each visual effect from the speaker's point of view, all effects except "Freeze" were noticed 80% of the time (See Fig.4.25). From the follow-up interview, we found that the changes in "glitch" and "Blur" were easy to notice because they reflected images that are not usually seen. as for Slow-motion, there were comments such as "the timing of the other person's gesture was off" and "the other person's screen was jerky and there was a time lag. The results showed that it was easy to recognize the listener's movement. On the other hand, although the elements of Slow-motion and visual effects were similar, Freeze stopped for a moment but returned to the normal state again, making it difficult for the listener to recognize. Some of the interviewees who answered that they did not notice the effect seemed to feel that they had noticed that the image stopped several times, but did not think it was an effect.

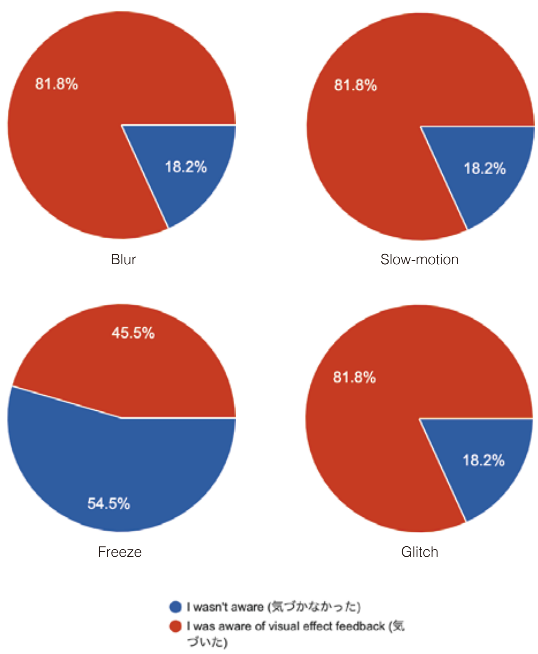


Figure 4.25 Manipulation Check of Speaker's Awareness of Visual Effects

Results.2: The Extent of Feeling from a Speaker over an Online Speech Experience with Visual Effects

The results of SD method are shown below (See Fig.4.29). Unfortunately, ANOVA analysis did not show any significant changes. There are two possible reasons for this result. The first reason is that in this experiment, we compared the effect with a similar effect created with the same concept(Artificial Network Disruptions). Therefore, it is presumed that the design of the effect was only slightly different and that there was not that much difference in the level of recognition. Another reason is the small number of paticipants (11). Therefore, we think it is necessary to increase the sample size to capture the changes for further generalization of the results.

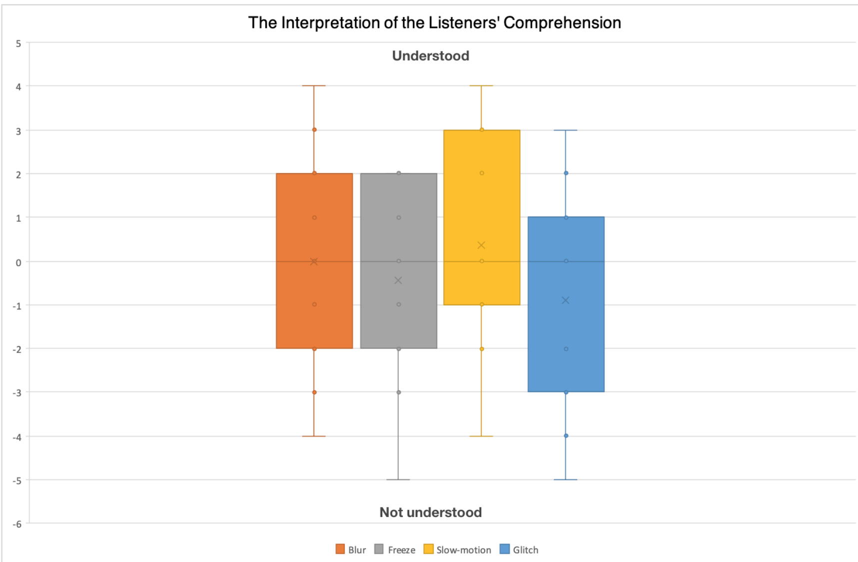


Figure 4.26 The Interpretation of the Listeners' Comprehension (Understood-Not understand)

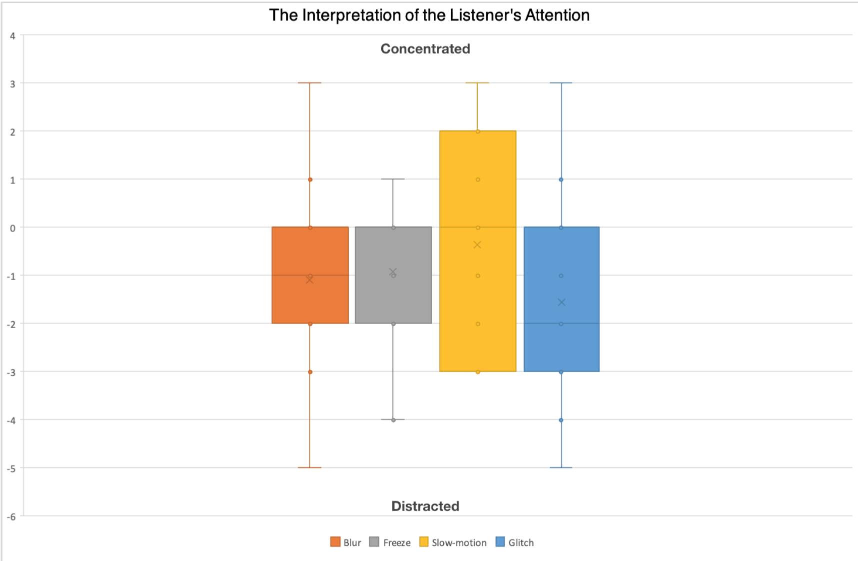


Figure 4.27 The Interpretation of the Listener's Attention to the Speaker (Concentrated-Distracted)

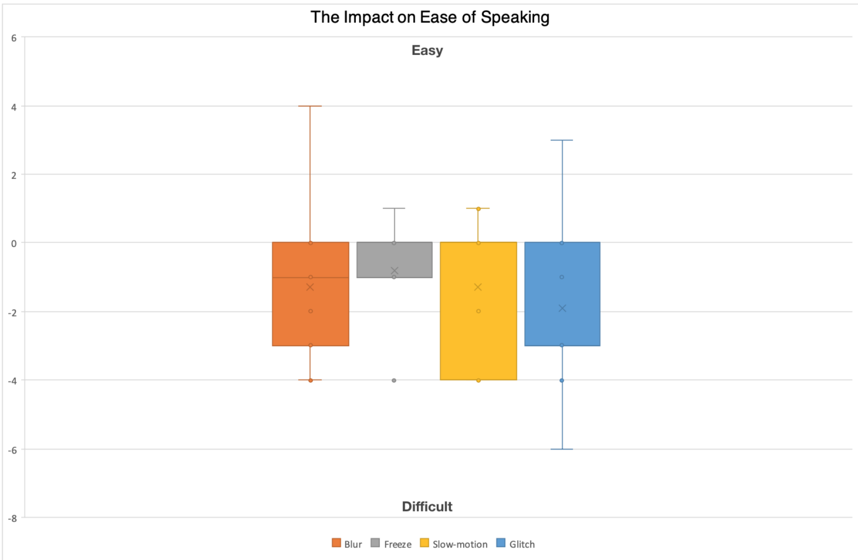


Figure 4.28 The Ease of Speaking (Easy-Difficult)

Figure 4.29 Study.2: Visual Evaluation Feedback Analysis

Qualitative Analysis

Results.3: Insights from Interviews

As a result of further individual interviews, 5 participants commented that the Glitch and Blur effects were unnatural and deliberate, and that they felt the listener was not paying attention to what the speaker was saying. Regarding Blur effect, one respondent said, “it is easy to notice, but it is not that intrusive”. In the case of Freeze, the screen stops for a moment, but then returns to normal, so it doesn’t bother participants, or they don’t notice it. Perhaps for this reason, Freeze is the least noticeable effect over 4 effects as Fig.4.25 showed. As for Slow-motion effect, some speakers said that they adjusted the timing of their speech based on the other person’s reaction or slowed down their own speaking speed.

4.5.4 Discussion

From the SD method questionnaire responses in Fig.4.29, the results of the effective ranking of effects can be interpreted as follows.

- Regarding the interpretation of the listeners’ comprehension, more participants answered that they did not understand with effects of Glitch and Freeze.
- Regarding the interpretation of the listener’s attention to the speaker, participants said that a listener “looked distracted” more with the effects of Glitch.
- As for the ease of speaking, the respondents answered that the Glitch and Slow-motion effects were “difficult to speak”. However, the difference over visual effects is quite subtle.

Given the results from the interviews, the results show that the features of each effect on the speaker in Slow-motion, Freeze, and Glitch are as follows: Glitch interferes with the conversation as noise, Freeze is difficult to notice, and when noticed, the speaker stops thinking, and Slow-motion motivates speakers to adjust their speaking style to respect the understanding of the listener. Therefore, since the goal of this study is to “improve speech”, we can see that Slow-motion has the feature that make it an ideal intervention for speakers.

However, in this study, we did not see any significant change in the actual speaker’s speech itself. For this outcome, participants responded, in a setting where one person gives a one-sided presentation, even if they noticed an effect, they were not sure whether if they should talk to the listener, then they just keep talking. Therefore, for the next experimental setting, it is necessary to change the setting to the experimental setting in which both the speaker and the listener would interact together, rather than that in which only one speaker speaks.

4.6. Validation Study

4.6.1 Study Design

Our goal is to design visualizations that allow NS to empathize with how NNS understand their conversations and adjust their own speaking style. Therefore, we aim to use visual effects to intervene in speaker's speaking behavior by triggering psychological changes. The design of visual effects are Blur effect and Slow-motion effect.

In this study, the language of the native speakers is set as English. In the experiment in the environment of Keio Media Design, we define that the NS is the September batch students who take classes in English, and the NNS is the April batch students who take classes in Japanese.

Building on previous research, we will focus on these 3 points below:

- Psychological change in a speaker to understand listeners' comprehension
- Changes in speaker's speaking style during a conversation
- Effect on the listeners' comprehension during a conversation

To achieve the goal, we hypothesized that

1. *"Slow-motion effect would be possible to slow down the speaker's speaking by slowing down the playback speed of the listener's screen."*
2. *"Blur effect could create an opportunity for a speaker to pay attention to the listener's level of understanding by hindering the other person's facial expressions."*
3. *"Comparing conversation with those effects, the stress level of NNS will decrease with Blur effect."*

Method

In terms of experimental design, we will follow the study method: Wizard of Oz, which is a method in which participants are led to believe they are interacting with a working prototype of a system, but in reality, a researcher is acting as a proxy for the system from behind the scenes. Thus, in this experiment, the activation of the effect is manipulated by the researcher in this experiment. The visual effect will be activated three times per session (once per minute). In order to measure the stress level of the listeners, the two listeners were equipped with wristbands that measure EDA and heart rate as described in 4.2.2.

4.6.2 Procedure

In total, 14 adult participants(11 female, 10 female) took part in the experiment. We will conduct our experiment with 3 settings of No effect, Blur effect, and Slow-motion effect. The order of effect appearance is counterbalanced. At the beginning of the experiment, participants will be given a practice session to get familiar with the task and other members. Participants will be asked to fill the consent form and the questionnaire to check English proficiency level before the experiment [51], which is the used in the previous experiment in Preliminary Study(Phase.2). The study typically takes 45 minutes in total with 1 minutes for a short questionnaire between sessions. Each session will take 3-5 minutes. A speaker will be given 1 minute for one target word and participants will guess 3 words per session. After the experiment, a speaker will be asked for an extra interview by a researcher (See Fig.4.30, 4.31).

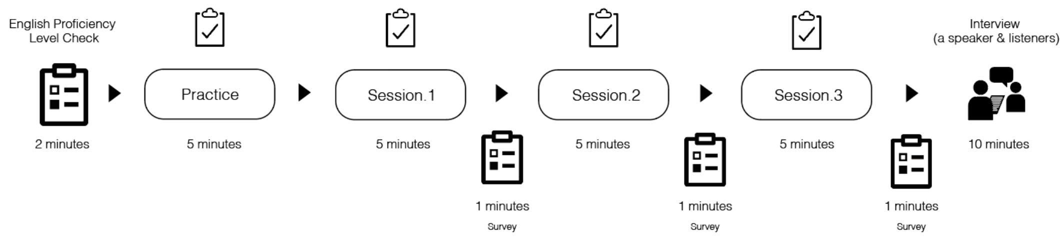


Figure 4.30 Validation Study: Experiment Procedure

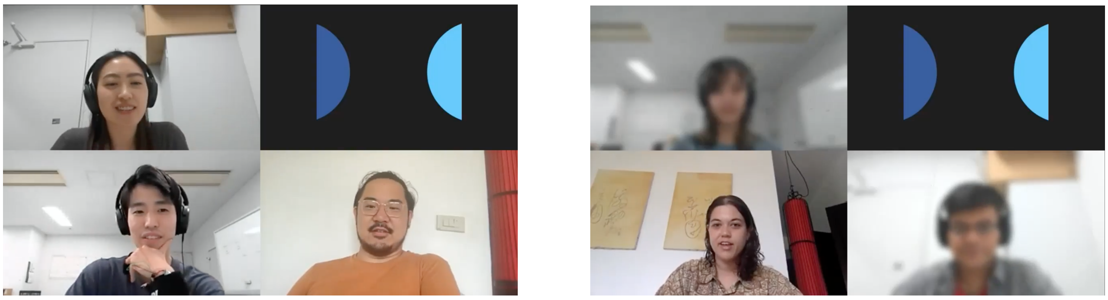


Figure 4.31 Validation Study: Experimental Setup.1

Fig.4.32 shows the overview of the experiment setup. In the experiment, participants play the taboo game [3] in a group of 3 people in Zoom. 1 person is assigned for a cue giver(=a speaker), who will deliver a guess-word to your team members without using taboo words listed. And the rest of 2 people are assigned for guessers(=listeners), who will get the guess-word from listening explanation by a cue giver. A set of one international student (=L1) and one Japanese student (=L2) per group. The goal of this game is the percentage of correct answers for the entire team. To achieve the task, a speaker will be required to change the way one speaks in order to convey information to the other person efficiently under the intervention of visual effects.

Listeners are allowed to ask a speaker questions and provide additional explanations to guide their answers. However, they cannot directly ask the subject word. The question must be in a format that can be answered with “yes” or “no”. (e.g., “Is it a living thing?”) During the study, all answers by both a speaker and listeners will be kept secret as we need to measure the speaker’s perceived listener’s understanding and the listener’s confidence in his or her own understanding.

3 https://playtaboo.com/playpage

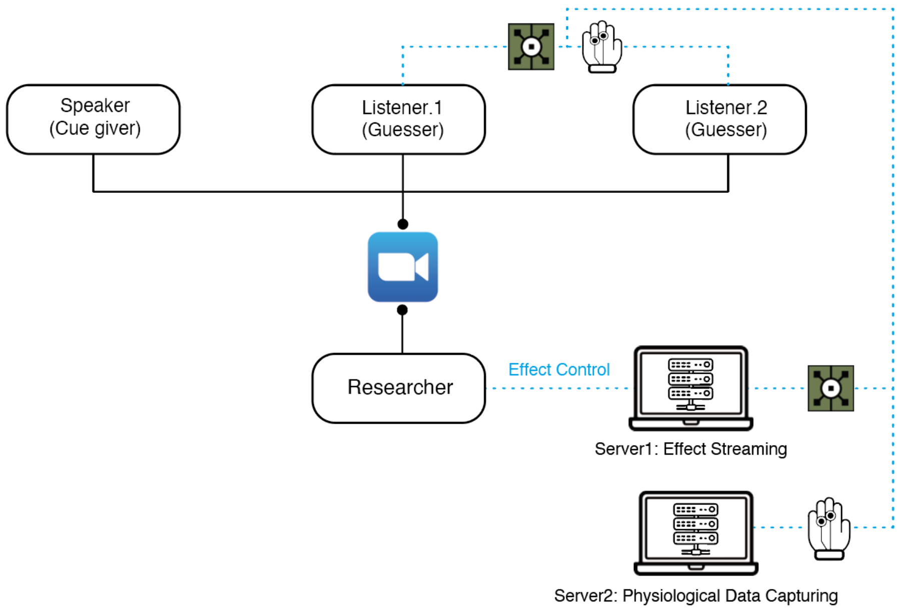


Figure 4.32 Validation Study: Experimental Setup.2

Measures

In order to collect subjective survey measurements after each session, we refer to the measurements of Duan at el. [3] for the questionnaire contents.

- **Manipulation Check of Speaker's Awareness of Visual Effects**

  As in Study.2, speaker's awareness of the visual effect was assessed by a multiple choice question asking them to identify what was displayed on their screen during the experiment. "I was aware of visual effect feedback" "I was sometimes aware of visual effect feedback" "I wasn't aware"). The type of effect was not given in advance, so those who answered that they noticed it were asked to write down what kind of effect it was.

- **Subjective Assessment of Speaker's Perception of the Visual Effects**

  In order to assess the speaker's perception of the visual effects used in the conversation, we used questionnaire with SD method. In this case, 4 adjective pairs were used. “smooth-awkward,” “understood-did not understand,” “concentrated-distracted,” and “easy-difficult to talk”.

- **Workload**

  After each trial, both a speaker and listeners completed the NASA-TLX, a widely used tool to measure subjective task workload, using a 7-point Likert scale (1=extremely low, 7=extremely high). We dropped the “Temporal Demand” dimension out of the 6 dimensions (mental, physical and temporal demand, performance, effort and frustration) to have participants not to feel pressured with the time limit.

- **Physiological Responses of Listeners in Conversation**

  All listeners were asked to provide and wear the wristband introduced in Fig.4.5 by the researcher. We were collecting skin conductance and heart activity through sensing devices. Heart rate and skin response was assessed from wearable wireless sensory device placed on the fingers of non-dominant hand.

- **Listeners Perceived Clarity of Speaker’s Speech Behavior**

  This was measured using one item. “A speaker spoke clearly enough for me to understand” and listeners rated on a 7-point (1=strongly disagree, 7=strongly agree) Likert scale.

- **Listeners’ Perceived Speaker’s Accommodation**

  We measured listeners perception of each speaker’s accommodation using a 4-item scale adapted from Gasiorek [17]. All listeners responded on a 7-point Likert scale (1=strongly disagree, 7=strongly agree) about whether they felt each speaker “adjusted his/her ways to communicate for me”, “did not adjusted his/her communication appropriately”, “was intending to be helpful” and “was having good intention”. To prevent bias when answering, the questionnaire items were mixed with positive and negative statements.

4.7. Results

Unfortunately, the ANOVA analysis did not reveal any significant changes in the results of all the questionnaires except for the physiological responses. We believe this was due to the small sample size, which allowed individual differences between groups to show up in the results and prevented us from producing standardized results for generalization.

Data Analysis

Results1: Percentage of Correct Responses to the Task

Table.4.1 and Fig.4.35 show the percentage of correct responses to the experimental task (taboo game) in 3 settings: No effect, Blur effect, and Slow-motion effect, for the whole team and for each listener. The figure.4.33 shows the correct response rate of the team, which is the sum of the correct response rates of L1 and L2, and the figure.4.34 shows the correct response rate of each listener separately. There are seven teams with three members per team (one speaker and two listeners).

This figure represents the number of questions that the two listeners (L1&L2) were able to guess for the word of the theme explained by the speaker. If the listener's answer is exactly the same as the word, 1 point is given, and if the word is different but the meaning is the same, 0.5 point is given. (Full score = 3) Table.4.1 shows the SD and Mean of the percentage of correct responses to the task in terms of teams(L1&L2), L1 and L2 in the 3 settings above.

Table 4.1 SD and Mean of the percentage of correct responses to the task

	No effect		**Blur**		**Slow-motion**	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**
Team Accuracy(L1&L2)	1.92	1.01	1.57	1.10	1.92	1.08
L1 Accuracy	2.35	0.94	2.21	0.69	2.64	0.47
L2 Accuracy	1.50	0.95	0.92	1.09	1.21	1.07

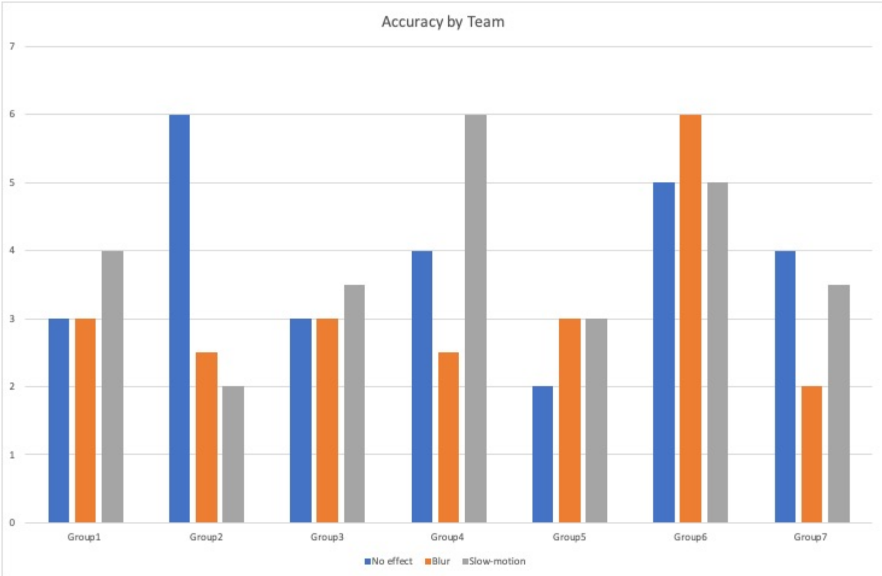


Figure 4.33 The Percentage of Correct Answers to the Tasks as Team(L1&L2)

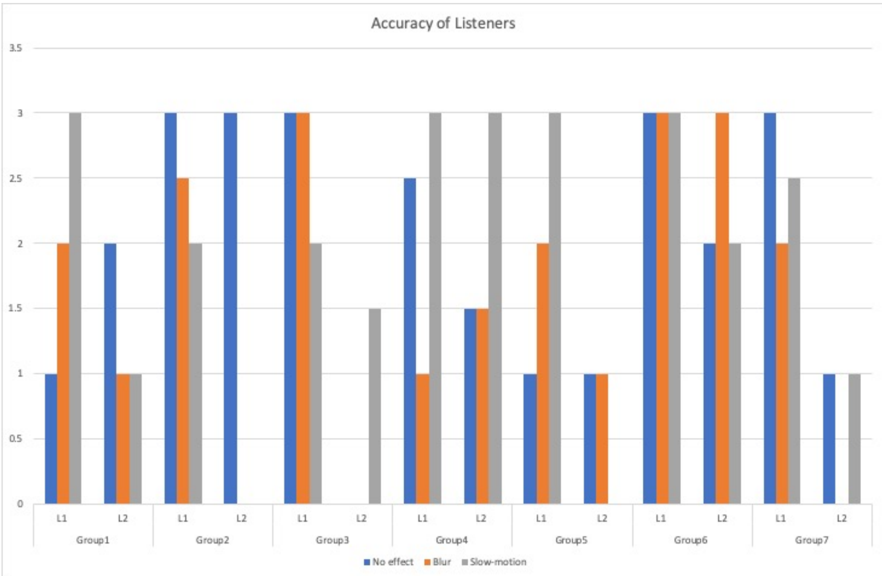


Figure 4.34 The Percentage of Correct Answers to the Tasks per Listener

Figure 4.35 The Percentage of Correct Answers to the Tasks in the 3 settings: No effect(blue), Blur effect(orange), and Slow-motion effect(grey)

Results2: Manipulation Check of Speaker's Awareness of Visual Effects

In order to check whether they were aware of the visual effects, a three-choice question was asked asking whether they were aware of the visual effects displayed on the screen during the conversation. The Fig.4.39 are the responses of the 7 speakers to the 3 effects.

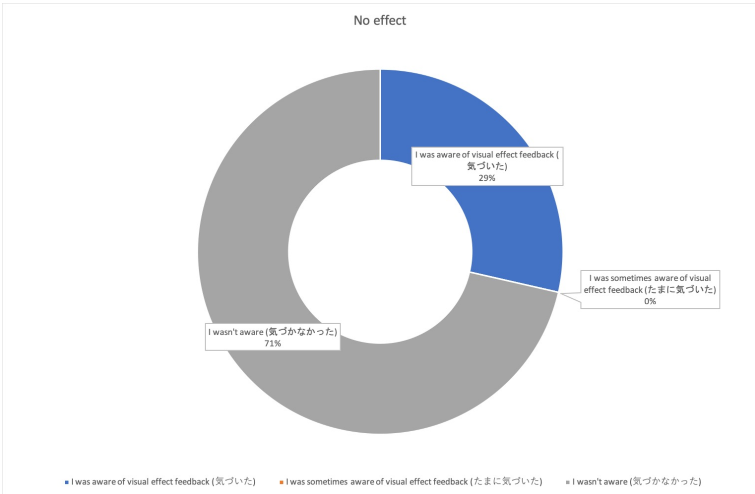


Figure 4.36 No effect

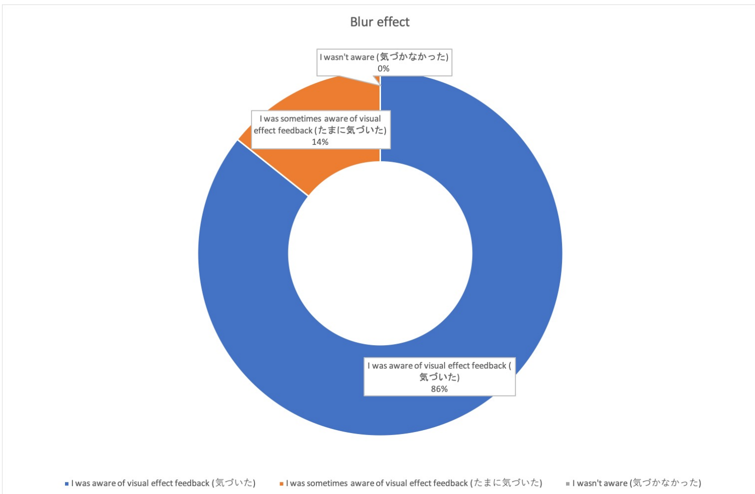


Figure 4.37 Blur effect

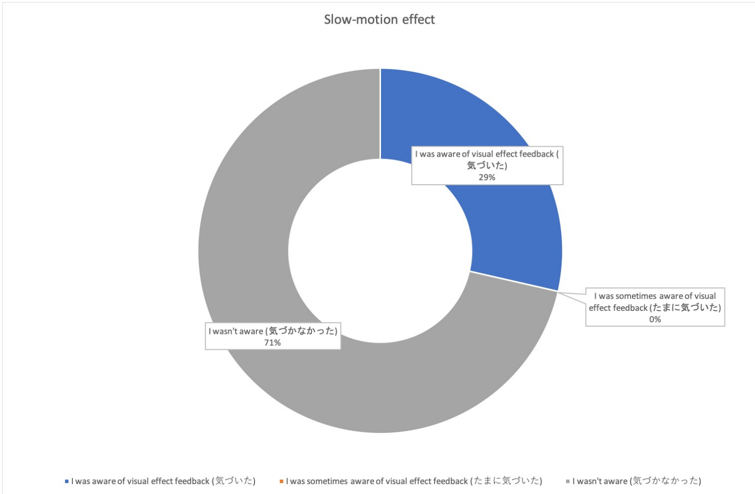


Figure 4.38 Slow-motion effect

Figure 4.39 Manipulation Check of 3 effects: Aware (blue), Sometimes aware effect(orange), Didn't Aware (gray)

For the No effect and Slow-motion effect, 2 out of 7 respondents (29%) answered that they "I was aware of the visual effect feedback" them, while the other 5 (71%) answered that they "I wasn't aware" them. On the other hand, 2 out of 7 respondents (29%) answered that they "I was sometimes aware of the visual effect feedback" the blur effect, while the other 5 (71%) answered that they "I was aware of the visual effect feedback".

Results3: Subjective Assessment of Speaker's Perception of the Visual Effects

In order to assess the speaker's perception of the visual effects used in the conversation under 3 different visual effect conditions, the SD method with 4 assessments was used: (1) Overall impression of the communication, (2) The interpretation of the listeners' comprehension, (3) The interpretation of the listeners' attention, (4) The impact on the ease of speaking. The Fig.4.44 shows the results of SD method regarding 3 different conditions: No effect (blue), Blur effect(orange), and Slow-motion effect (gray).

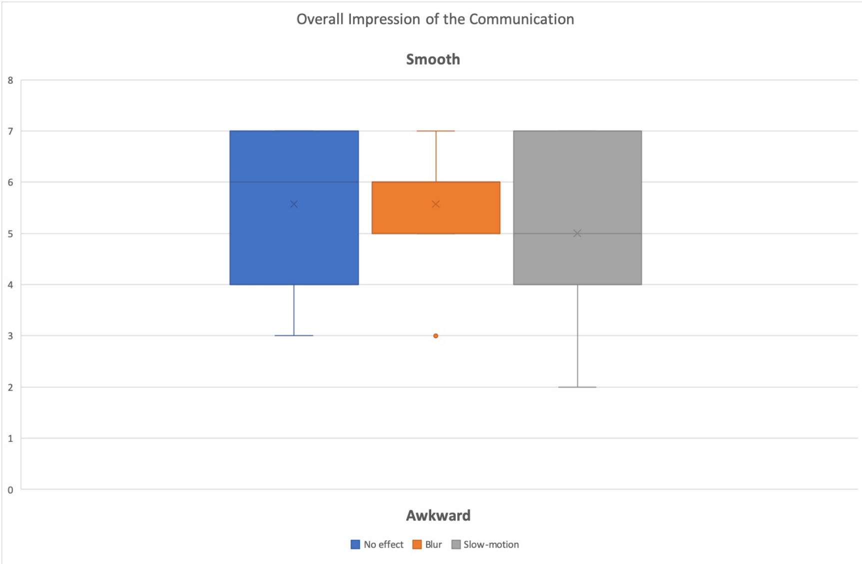


Figure 4.40 Overall Impression of the cCommunication

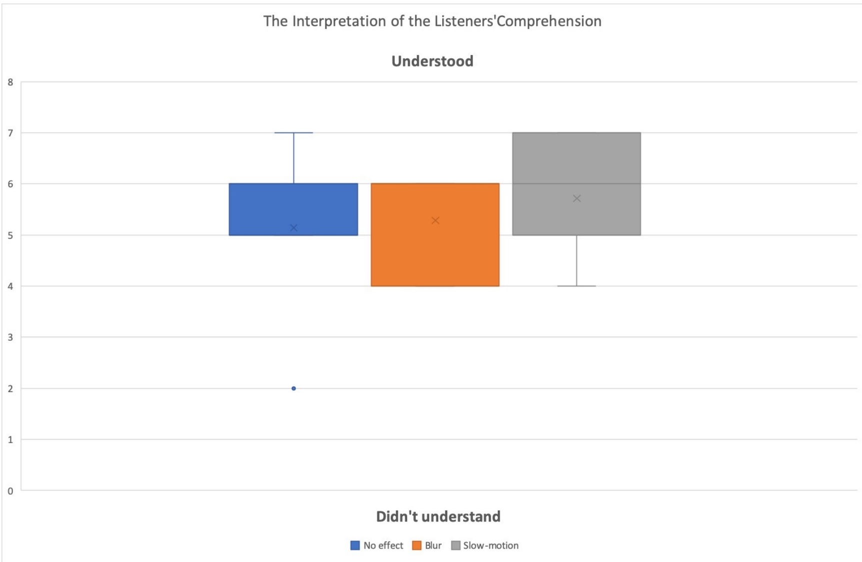


Figure 4.41 The Interpretation of the Listeners' Comprehension

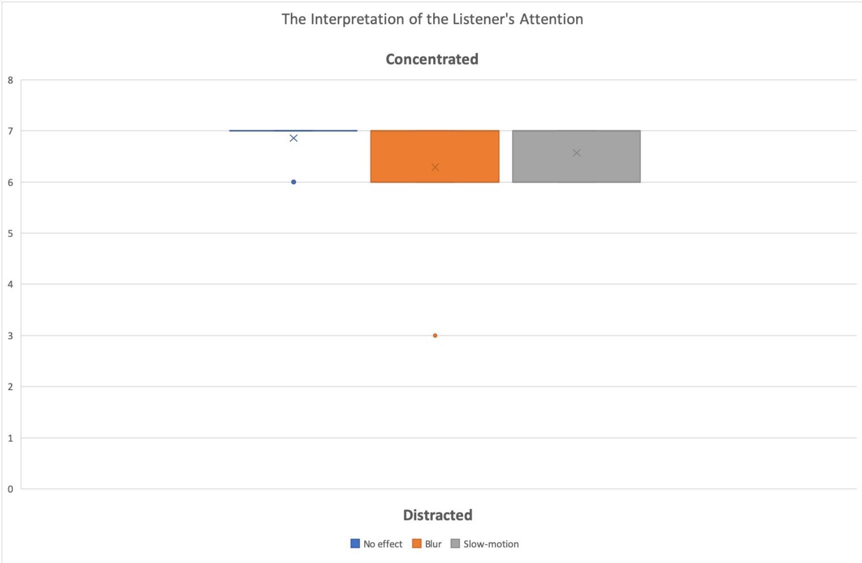


Figure 4.42 The Interpretation of the Listeners' Attention

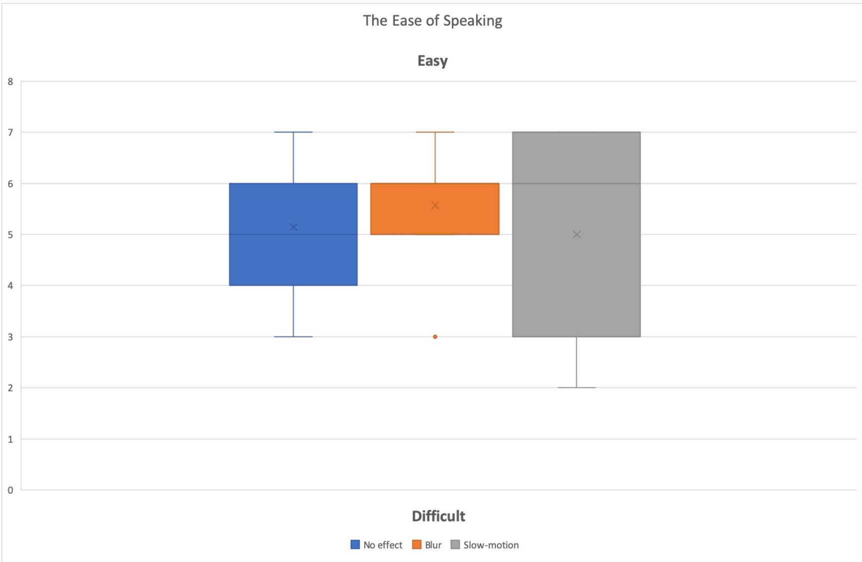


Figure 4.43 The Impact on the Ease of Speaking

Figure 4.44 Validation Study: Visual Evaluation Feedback Analysis

Results4: Workload

The following tables show the results of NASA-TLX for the participants: Table.4.2 shows the speakers, Table.4.3 shows all listeners (L1&L2), and Table.4.4 shows only L2. A one-way repeated measures showed no significant difference in workload from both speakers and listeners over 3 conditions: No effect, Blur effect and Slow-motion effect.

Table 4.2 Answers for NASA Load Index from Speakers

	No effect		**Blur**		**Slow-motion**	
Speakers	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Mental Demand	1.23	0.48	1.57	0.78	1.57	0.78
Physical Demand	4.85	1.46	4.57	1.39	4.85	1.46
Performance	5.28	0.95	5.42	0.78	5.14	1.34
Effort	5.00	1.68	5.28	1.11	5.71	0.95
Frustration	2.71	1.97	2.42	1.39	3.00	1.57

Table 4.3 Answers for NASA Load Index from Listeners(L1&L2)

	No effect		**Blur**		**Slow-motion**	
Listeners(L1&L2)	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Mental Demand	4.71	1.89	4.71	1.26	5.00	1.70
Physical Demand	2.57	1.91	2.78	1.84	3.50	1.60
Performance	3.85	2.24	4.00	2.18	3.85	1.84
Effort	4.50	2.06	4.85	1.40	5.07	1.59
Frustration	3.92	2.12	4.28	1.93	4.14	2.07

Table 4.4 Answers for NASA Load Index from L2

	No effect		Blur		Slow-motion	
Listeners(L2)	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Mental Demand	6.14	0.69	5.42	0.53	6.28	0.75
Physical Demand	3.14	2.41	3.14	2.11	4.00	1.82
Performance	2.28	1.70	2.42	1.71	2.28	0.48
Effort	6.14	0.69	5.85	0.89	6.42	0.53
Frustration	5.57	1.27	5.71	0.75	5.58	0.89

Results5: Physiological Responses of Listeners in Conversation

The following are the physiological responses of the listeners in the 3 conditions: No effect(blue), Blur effect(orange) and Slow-motion(grey) (See Fig.4.48). We mainly analyzed the results of three types of skin conductance response peaks (SCR_peaks), BPM (normalized), and HRV SDNN (normalized). The results of pairwise comparisons for each item are as follows.

- **SCR_peaks**: Significant difference between Blur effect and No effect. ($F(1.17, 9.33) = 6.66$, $p = .026$ ($p < .05$))
- **BPM (normalized)**: No significant difference among Blur, No effect, and Slow-motion session. ($F(2, 16) = .902$, $p = .426$)
- **HRV_SDNN (normalized)**: No significant difference among Blur, No effect, and Slow-motion session. ($F(1.25, 9.97) = 3.97$, $p = .068$)

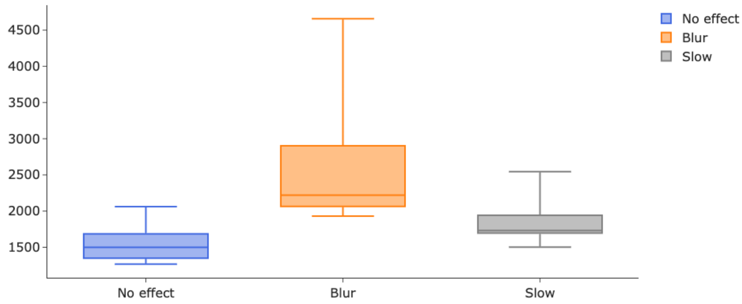


Figure 4.45 SCR_peaks (skin conductance response peaks) of all listeners

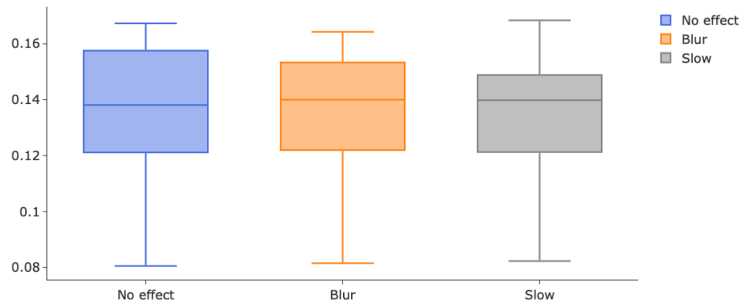


Figure 4.46 Normalized_BPM of all listeners

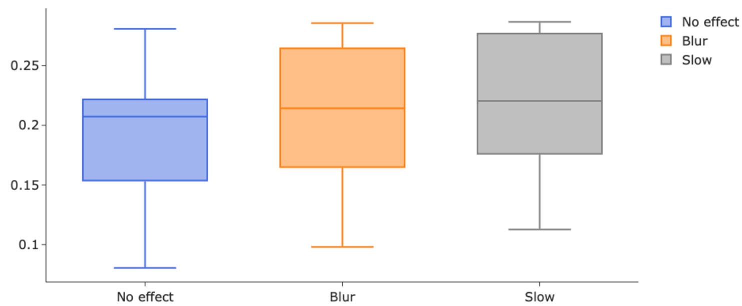


Figure 4.47 Normalized_HRV_SDNN of all listeners

Figure 4.48 The Physiological Responses of the Listeners

Results6: Listeners' Perceived Clarity of Speaker's Speech Behavior

Fig.4.51 shows the responses of L1 and L2 to the question about how listeners perceive the clarity of the speaker's speech under 3 different visual effect conditions. There are two figures, one for L1 and L2 summed responses and one for L2 alone.

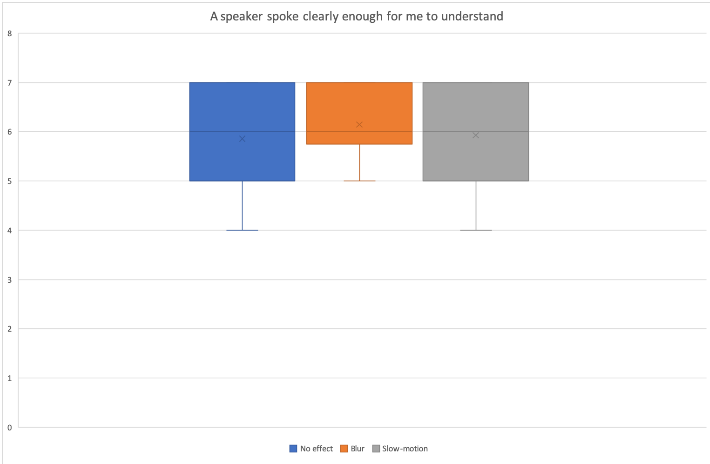


Figure 4.49 The Perceived Clarity of Speaker's Speech from L1&L2

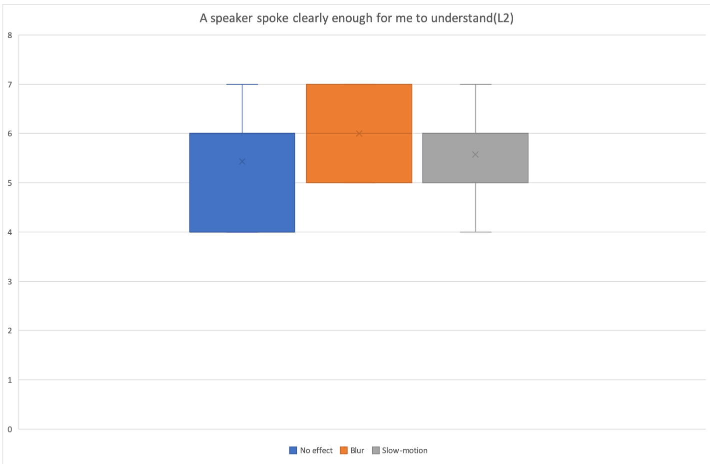


Figure 4.50 The Perceived Clarity of Speaker's Speech from L2

Figure 4.51 Listeners' Perceived Clarity of NS Speech Behavior on a scale of 1 to 7 in 3 settings: No effect(blue), Blur effect(orange), and Slow-motion effect(grey)

Results7: Listeners' Perceived Speaker's Accommodation

These figures below show the answers to four questions assessing whether the listener perceived the speaker's consideration under 3 different conditions of visual effect: the sum of L1 and L2 answers (Fig.4.56), only L2 answers (Fig.4.61).

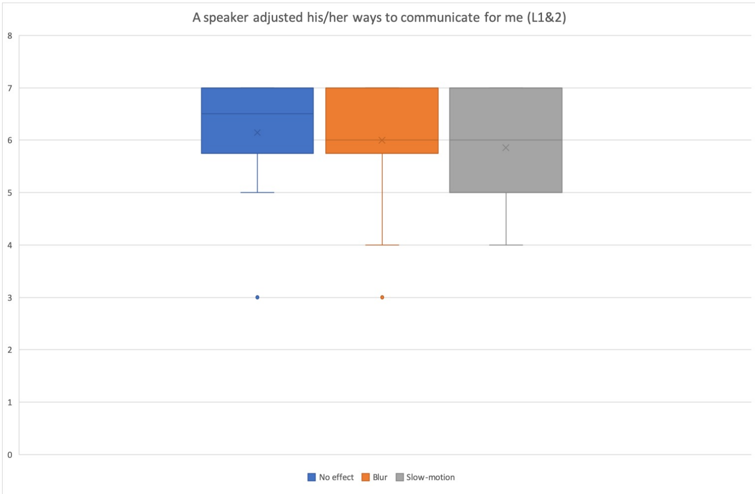


Figure 4.52 Q1: A speaker adjusted his/her ways to communicate for me (L1&L2)

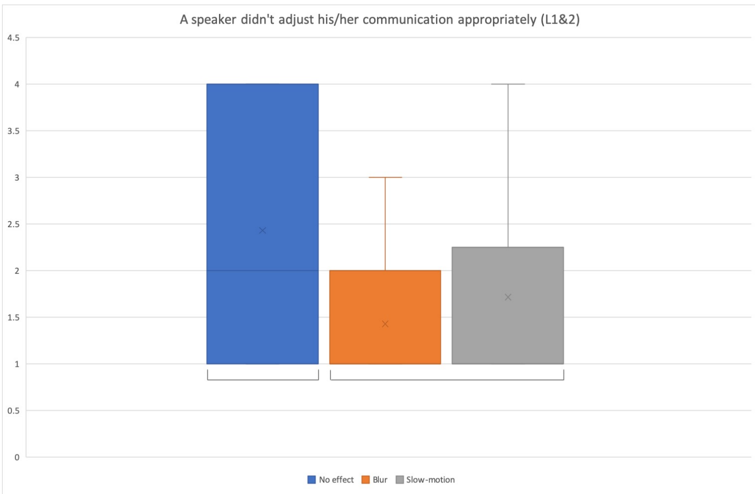


Figure 4.53 Q2: A speaker didn't adjusted his/her communication appropriately (L1&L2)

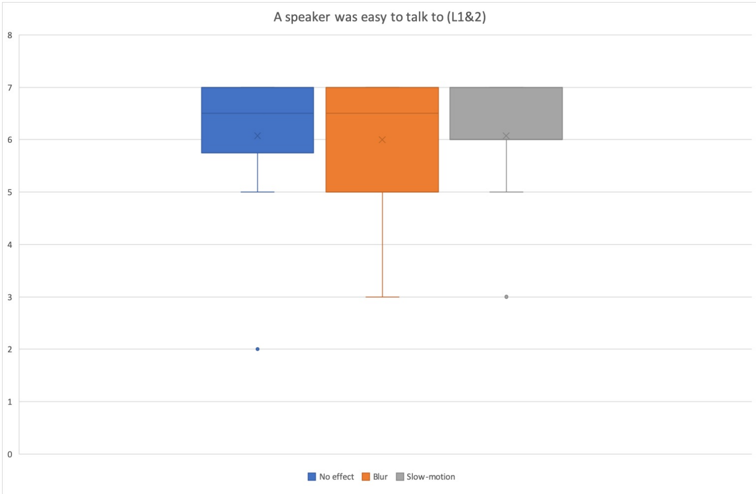


Figure 4.54 Q3: A speaker was easy to talk to (L1&L2)

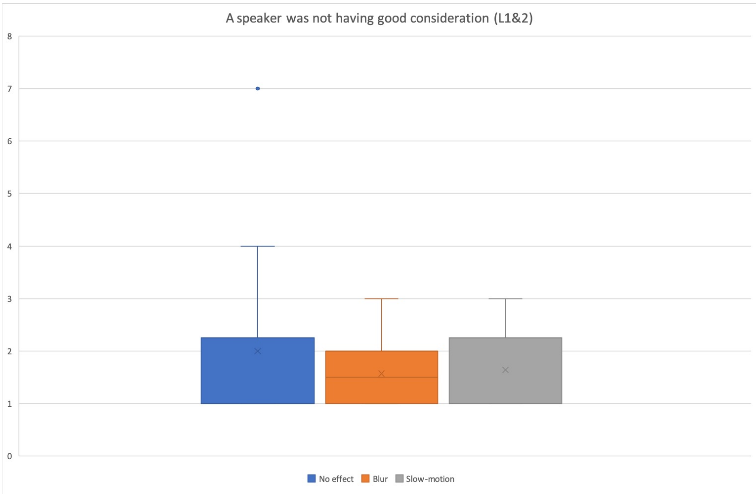


Figure 4.55 Q4: A speaker was not having good consideration (L1&L2)

Figure 4.56 Questions for listeners' perceived speakers' accommodation on a scale of 1 to 7 in 3 settings: No effect(blue), Blur effect(orange), and Slow-motion effect(grey)

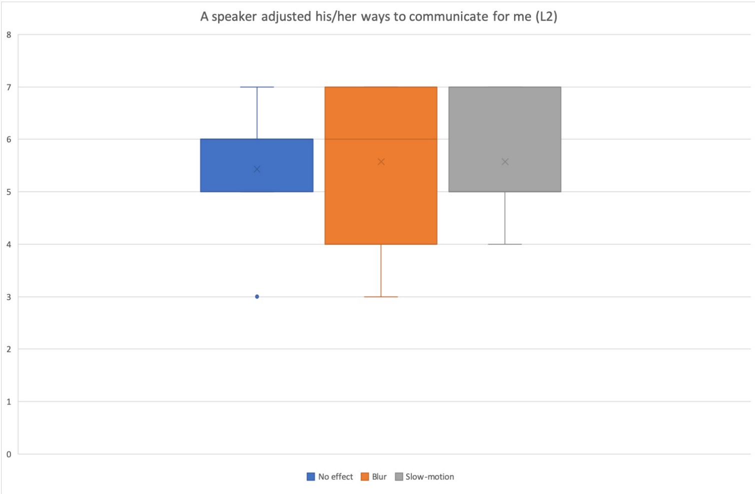


Figure 4.57 Q1: A speaker adjusted his/her ways to communicate for me (L2)

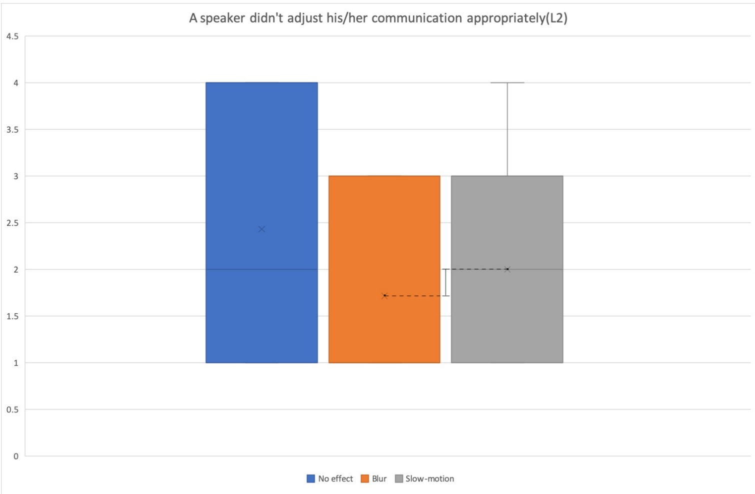


Figure 4.58 Q2: A speaker didn't adjusted his/her communication appropriately (L2)

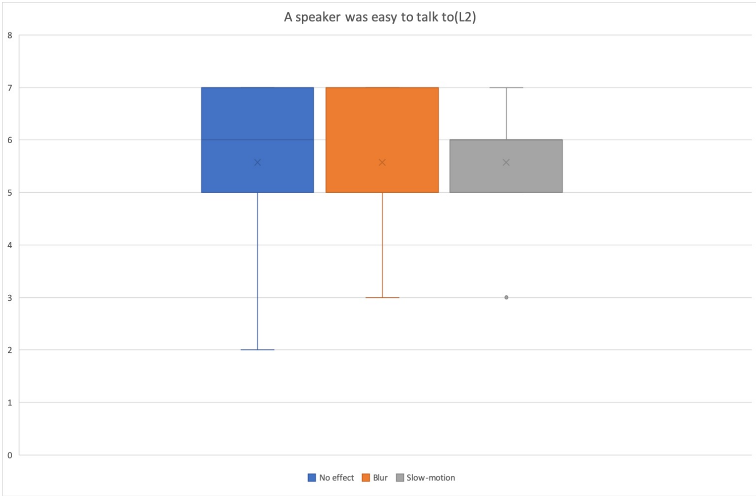


Figure 4.59 Q3: A speaker was easy to talk to (L2)

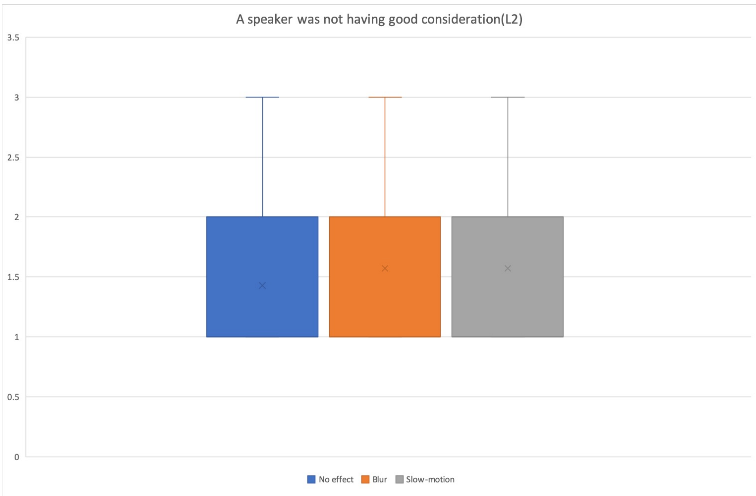


Figure 4.60 Q4: A speaker was not having good consideration (L2)

Figure 4.61 Questions for L2' perceived speakers' accommodation on a scale of 1 to 7 in 3 settings: No effect(blue), Blur effect(orange), and Slow-motion effect(grey)

Results8: Insights from Interviews

All interviews with speakers and L1 were conducted in English and audio-recorded and transcribed by the researcher; all interviews with L2 were conducted in

Japanese and audio-recorded and transcribed by the researcher, a native speaker of Japanese. All Japanese transcripts were translated into English using Deep Translate[4]. Using an inductive and interpretive approach [53], we first extracted themes from the quotations, then identified high-level themes and relationships between themes. This analysis led to the emergence of several interesting themes.

Interviews with Speakers

The following 5 questions were asked to the speakers.

Q1: How many target words do you think listeners guessed correctly? Q2: What cues do you rely on to make adjustments on your instruction to make you understood by the listener? Q3: How do you interpret the visual effect? Q4: How do you think is the cause of the effect? Q5: (After explain the function of effect trigger was the stress level from listeners) Which listener do you think became the trigger of the effect overall as impression level?

As a result of analyzing those questions, the following 3 themes were extracted.

The first theme that ran through interviews with all speakers concerned **factors that contribute to the speaker's confidence in his or her explanation**. Many speakers answered to grasp their listeners' level of understanding from their engagement in the conversation (questions, talking back as checking), facial expressions, and head nods/tilt.

- *They(listeners) don't say anything at all, they may look away from the screen, they may laugh, but they don't say anything. -S2*
- *Verbal and physical cues, facial expressions, and then the more questions they(listeners) asked, the more confident they seemed to be. -S4*
- *When they (listeners) were talking back, nodding heads, say things to get react. Then, I checked whether they were following. -S7*

The second theme that ran through interviews with all speakers concerned **speaker's interpretation of the visual effects**. 5 of 7 speakers interpret the Blur effect as internet or Zoom connection issue. However, as for the cause of the

4 https://www.deepl.com/ja/translator

visual effect, 2 responded that the Blur effect was triggered due to technical causes and 3 responded that it was due to personal causes by the listener (biological response, excitement level, and understanding level). The other 2 responded that it was due to the researcher's operation or random.

- *I wondered if it was a problem with the internet or wifi. I kept talking, worrying if they could hear me, but they didn't show any signs of confusion. It made me hesitate. (About the cause of the visual effect) Maybe reactions to their physiological signal because they were wearing the wristbands. -S3*
- *The screens going blurry so I thought my zoom connection was bad. Maybe the operators of the experiment were causing the visual effect as a distraction. -S5*
- *With blur of the screen in the 2nd session or something, I can't see their nodding. It might have distracted me from watching them. But I mainly watched on the list of words. (About the cause of the visual effect) part of their reaction, maybe represent their understanding? Not technical issues because when my wifi connection was worse, it did not show like that -S7*

The third theme that ran through interviews with all speakers concerned **the speaker's understanding of the listeners' understanding.** The judgments of the 7 speakers as to which listeners caused the most visual effect were as follows: 4 guessed L2, 2 guessed L1, and 1 answered don't know. The reasons given were that listeners talked less, asked more questions, and the listeners' facial expressions showed that they looked stressed.

- *L1. She looked bored. -S2*
- *L2. She seemed most confused, least engaged. -S4*
- *L2. Because he communicated less, but often asked questions. -S6*

Interviews with Listeners

The following 5 questions were asked to the listeners.

Q1: Which session was the easiest/most difficult to understand? Q2: How many target words do you think listeners guessed correctly? Q3: Did you notice any changes in the speaker's reaction/speaking style during 3 sessions? Q4: How did the speaker adjust her/his speaking style? Q5: How did the behavioral change influence your understanding their speech? As a result of analyzing those questions, the following 4 themes were extracted.

The first theme that ran through interviews with all listeners concerned **elements of understanding for the listeners**. The listeners reported familiarity with the target word including cultural background and the speaker's speaking speed are related to their understanding.

- *Some of the words related to some cultural background. -L1.1*
- *Depends on how you are familiar with the words. -L1.2*
- *Session.2 was easy to understand because the speed was just right and many of the words were familiar. Session.1 was more difficult than sessions 2 and 3 because there were more words I didn't understand. -L2.2*

The second theme that ran through interviews with all listeners concerned **the listener's confidence in one's own comprehension**. Regarding their own percentage of correct answers out of 9 questions, 6 out of 7 questions in L1 were 7 or more, and all the L2 students answered that it would be 4 or less. Many of the listeners, once they had some idea of the answer, asked the listener questions to confirm the accuracy of their answers.

- *When she understood the word, to confirm whether she was right or wrong, she asked questions about it and confirmed, and she had the confidence that she was right. -L1.6*
- *When I got it halfway through, and as he(speaker) explained, it fit my answers, and I also asked questions to confirm. -L1.7*

However, some of the L2 were unable to ask questions due to peer pressure, time pressure, and hesitation to ask questions in English. In some cases, L2 were able to confirm the content of the questions asked by L1, but most of them were not able to ask questions make sure of their answers.

- *Even if I understood the nuance of the word, I wasn't sure if it was correct, and the time ended without me being able to be sure. I didn't ask questions properly then I ended up being vague. In some cases, I finished answering by reading the flow around me. -L2.2*

- *To be honest, I could not understand any of them. I didn't understand some of the words in the explanation. When I could see the other listener already understood, I had to follow their pace, and there was also the pressure of the time limit. On the other hand, the more information the other listener got out of the questions, the easier it was to understand for me as well. I could compare my guesses with the other. -L2.4*

- *Sometimes I thought the speed was too fast, then I couldn't catch the words. But since I didn't know how to ask questions in English, I felt discourage in asking the speaker. -L2.7*

The third theme that ran through interviews with all listeners concerned **listener's interpretation of changes in speaker's behavior**. 10 out of 14 listeners noticed the visual effect. In sessions where there was a visual effect, listeners noted behavioral changes in speakers such as repeating explanations, emphasizing, stammering, asking questions, or speaking slowly.

- *During the 2nd session, there was a bit of blur, so that kind of threw her(speaker) off. There was a time lag, she couldn't explain properly. -L1.5*

- *When there were effects on the screen appeared, he(speaker) was taken by surprise, he repeated the words he were trying to say. And he slowed down after he started repeating words, he opened the floor up for questions. I wasn't sure if the other participant got it. -L1.7*

- *She(speaker) was trying to communicate very well from the beginning, but I wonder if she slowed down in Session.3 because he was worried about blurring and getting the information across to the listeners. -L2.5*

The forth theme that ran through interviews with all listeners concerned **effect of speaker's behavior change on listener's understanding**. 8 out of 14 listeners (L1:3, L2:5) responded that the change in the speaker's behavior helped with their understanding.

- *Yeah,it helped. I felt like a conversation rather than a lecture so it was easy to communicate (with the speaker). -L1.4*
- *When he(speaker) was repeating himself and slowing down, I started trying to understand him more and listening more closely. -L1.7*
- *It was easy to understand; she(speaker) said the word twice because it was hard to hear at once. When I was blurry, he said it twice. -L2.5*
- *(After the speaker asked me if I was okay,) Rather than not saying anything, I felt reassured that it was okay to ask. Also, since the speed of speech slowed down, it was easier to hear than when it was faster. -L2.7*

On the other hand, it seemed that some participants felt frustration from mismatch between effect appearance and their actual understanding.

- *I had to listen more closely to him(speaker) because his speaking style was becoming a bit erratic. -L1.7*
- *"I was afraid of my blurred face misleads the speaker to think the listener don't understand.". -L2.2*

4.8. Discussions

4.8.1 Discussions for Each Result

Results1: Percentage of Correct Responses to the Task

The mean of the percentage of correct responses for the entire listener(L1&L2) is lower for the Blur effect than for the No effect and Slow-motion effect. Next, analysis by listeners shows that the percentage of correct responses of international students (L1) who are accustomed to using English has a higher score in the Slow-motion effect and a lower score in the Blur effect. For Japanese students (L2) who are not used to communicating in English, the correct response rate was highest in the No effect, followed by Slow-motion, and lowest in the Blur effect. In other words, the correct response rate was higher in the session using Slow-motion than in the Blur effect.

Data Analysis

Discussion2: Manipulation Check of Speaker's Awareness of Visual Effects

The responses in Fig.4.39 clearly show that the Blur effect is more noticeable than the Slow-motion effect. From interview analysis, it is found that 5 of 7 speakers thought it was a problem with the internet or wifi connection and were worried that the listeners was not getting through. In addition, the Blur effect made it difficult for speakers to see the listeners' facial expressions and nods, which further led to their anxiety. For this reason, we also observed the speaker asking the listener, "Are you okay?"

Slow-motion, on the other hand, went unnoticed by most participants. This is probably because people do not move their bodies during online conversations in general, so it is difficult to see changes in their movements by Slow-motion effect. Especially the Japanese participants(L2) did not change their facial expressions or body language much, so there was not much difference between when the effect was activated and when it was not activated. Moreover, when multiple people are talking, it is generally hard to focus on each person's actions, so it may have been more difficult to notice small changes in the other person's behavior than when talking 1-on-1.

Discussion3: Subjective Assessment of Speaker's Perception of the Visual Effects

From the SD method questionnaire responses in Fig.4.44, the results of the effective ranking of effects can be interpreted as follows.

- Regarding the overall impression of the communication, more speakers felt "awkward" in the session with Slow-motion.
- As for the ease of speaking, the speakers answered that the Blur effect was "easy to speak".

Referring to the results of this study and the Manipulation Check, we can see that the Blur effect is easily noticeable, but not so intrusive as to interfere with speech. In other words, it can function as a visual effect for ambient notifications. As for the Slow-motion effect, although most of the speakers did not notice the effect in the Manipulation Check, they felt the awkwardness of the conversation on unconscious level.

Discussion4: Workload

Workload in Speakers

According to Table.4.2, the value of frustration for Blur effect is the lowest among the 3 conditions, even though all the speakers noticed the appearance of the effect in the Manipulation Check. On the contrary, the Slow-motion session seems to have required more effort and frustration from the speakers than the no-effect and blur sessions.

From this perspective, the following characteristic points can be made about each effect. The Blur effect is moderately noticeable, but not so much as to interfere with the speaker's conversation. The Slow-motion effect is not perceived as an intentional effect, but is thought to have some change in the impression of the conversation on an unconscious level.

Workload in Listeners

A comparison of workload with and without effects for all listeners(L1&L2), referring to Table.4.3, shows that the frustration value was highest in the Blur session. Next, a comparison between effects shows that the session using Slow-motion effect requires relatively higher mental demands and effort than those using Blur effect, and considerably higher physical demands.

Based on the NASA-TLX results for L2 shown in Table.4.4, the impression of L2's workload from the comparative observations between L1 and L2 is as follows: the evaluation for performance is low, and physical, mental, effort, and frustration are high. From this finding, it can be seen that L2 has the highest psychological and physical workload during the experiment and has the lowest self-confidence and accomplishment in the task. In fact, the responses from L2 in the interviews about the L's confidence in their own comprehension showed that many L2 felt time pressure, peer pressure, and pressure to ask questions in English. In addition, it was also found many L2 were not confident about their answers since they could not share their understanding with other members. With these reasons, the workload of L2 was higher than L1.

Discussion5: Physiological Responses of Listeners' in Conversation

According to Fig.4.45, SCR peaks are significantly higher in the Blur session than No effect session, which is known as the SCR peak, for quantifying stressful feelings [54].

From the NASA-TLX results, we can see that the values of frustration and effort of the listeners during the Blur session are the highest than the other 2 conditions. According to the interview, some listeners mentioned that they felt anxious when the timing of the effect and the actual level of understanding of the listeners did not match. Therefore, we can attribute these results into the finding that some listeners felt frustration from mismatch between effect appearance and their actual understanding.

Discussion6: Listeners' Perceived Clarity of Speaker's Speech Behavior

According to listeners' responses in Fig.4.51, it can be said that listeners assumed that most speakers generally spoke clearly. Among the 3 conditions, scores in the Blur effect session were subtly higher, especially in responses from L2. Observations during the experiments and the interview responses also showed that the speakers were attentive to their listeners' responses, such as nods, questions, and facial expressions, as they gave their explanations.

Discussion7: Listeners' Perceived Speaker's Accommodation

As with the responses to the listeners' perceived clarity of speaker's speech behavior in Fig.4.56, we can tell listeners' impressions of the speaker's accommodation were highly positive over 3 sessions in general. The questions "A speaker adjusted his/her ways to communicate for me", "A speaker was intending to be helpful A speaker was intending to be helpful", "A speaker adjusted his/hers ways to communicate for me", and "A speaker was having good intention" are relatively similar in all 3 conditions.

However, the response that "A speaker did not adjust his/her communication appropriately" was higher in the sessions where effects (Blur and Slow-motion) were used than in the sessions where no effects were used from both L1&L2 and L2 alone (See Fig.4.64). This indicates that the speaker's adjustment was more appropriate for the listeners in the session with effect. In terms of comparison between effects, specially for L2, many of them answered that the speaker's adjustment was more optimal in the session with Blur effect than Slow-motion.

In the interviews, we found listeners identified changes in the speaker's behavior such as repeating explanations, emphasizing, stammering, asking questions, and speaking slowly in sessions with visual effects. For those changes in the speaker's behavior, 8 out of 14 listeners (L1: 3, L2: 5) indicated that it helped their understanding.

This result indicates that the speaker tends to make adjustments more appropriately for the listener when the effect is present. In addition, for listeners with a relatively low proficiency level in English, the speaker's adjustment was most effective during the Blur effect.

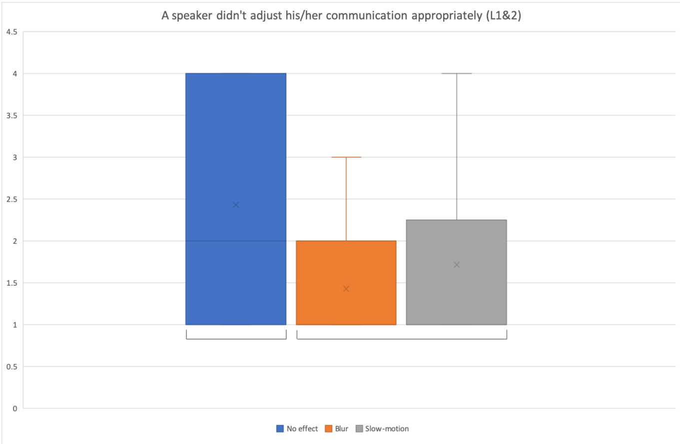


Figure 4.62 Answers from all Listeners(L1&L2)

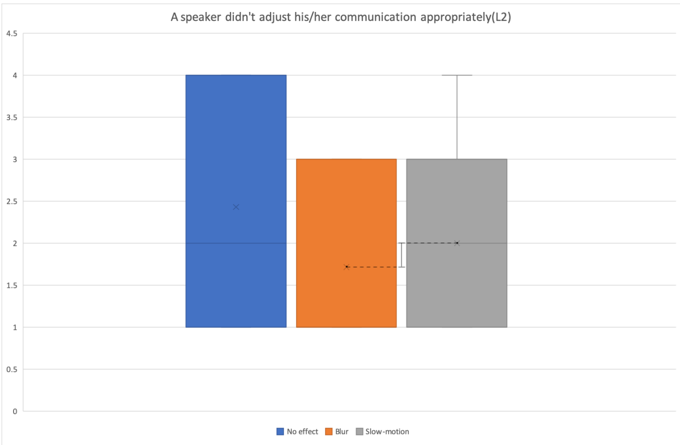


Figure 4.63 Answers Japanese students(L1)

Figure 4.64 Answers for a question on A speaker's speech adjustment from all listeners(L1&L2) and Japanese students(L1) in 3 settings: no effect(blue), Blur effect(orange), and Slow-motion effect(grey)

4.8.2 Discussion Summary

Overall, our analysis and discussion of the results suggest the following 3 findings.

Finding.1: Impact of Visual Effects on Behavioral Change of Speakers

From an interview, we found that 5 of 7 speakers who noticed the visual effect interpreted the cause of blurring the screen as a delay in the Internet or Zoom connection issue. This resulted in changes in the speakers' behaviors, such as asking questions, reciting, and speaking slowly in concern for the listener. In particular, the blur effect was found to be effective in triggering behavioral changes in speakers without causing extra cognitive load on speakers as the value of frustration NASA-TLX results for speakers during the blur session is the lowest among the three conditions.

Finding.2: Impact of Visual Effects on the Listener

From the interview, 8 of 14 listeners said that such behavioral changes in speakers helped them understand the conversation better. The results suggest that visual effects contribute to accommodative changes in the speaker's speech, which ultimately affect the listener's understanding of the conversation. Especially for Japanese students(L2), who had were unable to ask questions due to time and peer pressure, and hesitation to ask questions in English, seemed to be more comfortable to asking questions after seeing behavioral changes in speakers. From those reactions, we can assume our visual effects led to a psychological comfort in listeners to try to understand the conversation better.

However, the visual effect was not effective for all listeners. The number of SCR peaks in the blur session is significantly higher than that in the no effect session. Based on listeners' NASA-TLX responses, the frustration value was highest as well in the blur session. According to the interview, some listeners mentioned that they felt anxious when the timing of the effect and the actual level of understanding of the listeners did not match.

Finding.3: Characteristics and Evaluation of Visual Effects

The speaker's awareness of the blur effect was 100%, while that of the slow-motion effect was 29%. The analysis with including results from NASA-TLX revealed that each visual effect has the following characteristics: Blur effect is moderately noticeable, but does not interfere with the speaker's conversation. Slow-motion effect is not perceived as an intentional effect, but gives the speaker a sense of discomfort on an unconscious level. Therefore, while the Blur effect could make the speaker think that a technical problem has occurred and create an opportunity to pay attention to the listener's understanding, the Slow-motion effect is difficult for the speaker to see, and as a result, no noticeable behavioral change occurred.

4.9. Limitation and Future Work

Overall, we believe that the Blur effect has the potential for modulating speaker's accommodative speech behavior without affecting the perceived task load. However, the current design has its limitations, and together with the insights gained from the user interviews, there is room for further improvement and development. In the next section, we will elaborate on the current limitations and future challenges.

Integrating Listener's Feedback to Trigger Visual Effects

Finding.2 states that some listeners felt frustrated more because of mismatch in timing between effect appearance and their actual understanding. From the interview, most speakers judge their listeners' level of understanding based on their level of participation in the conversation, such as asking questions or talking back for confirmation, their facial expressions, and their head nods/tilt. Hence, the next step is to integrate the above-mentioned physical movements of the listener with the triggering the visual effect, so that the visual effect can be applied when the listener has trouble understanding the conversation. We believe that this integration will improve the speaker's awareness and consideration to the listener's comprehension and avoid extra frustration on listeners.

Improving the Experimental Setup

There was a slight discrepancy between the conditions of the experimental setup and the conditions of the actual conversation as the assumption of use. As a result, the behavior of the speaker and listener was somewhat limited. The two major reasons are as follows.

First of all, the ease of explanation and understanding varied greatly depending on the game's target words. For proper nouns with a strong cultural background, the percentage of correct answers depended on whether or not the participants understood the culture of the country. Therefore, it is highly likely that each session was not conducted under the same conditions of difficulty. Secondly, the one-minute explanation time was too short for everyone to perform the task. Due to the pressure of the time limit, they often ended up with one-sided explanations by speaker's explanation, and the number of conversations from the listeners was often low. As a result, some listeners said that they did not know when to ask questions because the speaker was speaking in one direction. Thus, it is necessary to set a longer time so that both the speaker and the listener can interact.

Hence, we need to improve the design of the experiment in the future in order to make the environment closer to the online discussion we envisioned. Eventually, we hope to use Communi-gate for online lectures, group work, and workshops. In this way, professors and other members of the group can facilitate discussions according to students' level of understanding, which we believe will contribute to the improvement of online collaborative activities.

4.10. Summary

In preliminary study, we explored the relationship between listening comprehension and biological responses in 2 conditions: the pre-recorded condition using listening materials (phase.1), and the real-time condition using time stamps (phase.2). The results showed that the EDA data captured the most subtle changes in the participants' biological responses during listening task. On the other hand, in the experiment in the real-time situation of phase 2, we found that the standard of "being able to understand" varied greatly among individuals, and it was difficult to standardize it to a certain standard. Therefore, in the final validation study, we decided not to use biological responses to trigger the visual effect, but to have the researcher activate it randomly.

In study 1 and 2, we conducted experiments to evaluate the visual effects, including visual effects designed based on the two concepts introduced in Chapter 3.5 (1. Hiding Facial Expression, 2. Artificial Network Disruptions) and other sample visual effects. We conducted a comparative evaluation using a questionnaire with the SD method along 3 main evaluation criteria: impression of the conversation, impression of the speaker to the listener, and the speaker's speaking style. As a result, the Blur effect and the Slow-motion effect were selected as as its scores on the above 3 points matched the concept.

In the final validation study, we investigated the effectiveness of visual effects on the speaker's behavior and the listener's behavior under the three conditions: No effect, Blur effect, and Slow-motion effect. The results showed that the Blur effect is the most effective measure to induce speakers' accommodative changes without causing extra perceived workload, and those changes in speakers led the listeners have to psychological comfort in trying to understand the conversation better.

Chapter 5
Conclusion

NNS has faced challenges and felt alienated when speaking with NS/fluent speakers due to the anxiety they feel from using a foreign language. On the contrary, the suffering they are experiencing is often overlooked as something that can be met by their self-learning and is rarely understood, especially by NS. Therefore, in order to bridge the gap in multicultural communication between people with different language abilities, it is important to design assistant tools that take into account the plight of NNS in foreign language communication.

In this context, many studies have attempted to mitigate this communication problem, but most of them have focused only on helping NNS improve their language skills. Therefore, not many studies have investigated the willingness of NS to respond in multilingual communication with the aim of facilitating their efforts to respond to NNS. In addition, online conversations make it more difficult for those around the NNS to respond to the NNS's responses. However, so far, there has not been much research and development of online communication tools to support the communication gap for foreign language users.

Given all these factors, we determined to design Communi-gate, a visual speech facilitator to help NS adjust their speaking style to improve their understanding of NNS in online cross-lingual communication. The goal is to create an experience in which NS and fluent speakers of English understand the linguistic constraints of NNS, and naturally change their speaking style for delivering communication more effectively. The results of interviews and fieldwork during the GID course have shown that most international students have experienced a sense of isolation in conversations with native/fluent speakers due to differences in their language proficiency. This problem was especially prevalent in group work classes and seemed to have a significant impact on international students' participation in group work and their relationships with team members. These factors led us to believe that it is critically important to develop a tool that would help NS adjust their speaking behavior in order to improve their understanding of NNS and thus reduce their stress.

To implement this concept, we designed visual effects (Blur effect and Slow-motion effect) as feedback to induce psychological changes and intervention in conversational behavior for a speaker. The results of our research experiments showed that visual effects contribute to behavioral changes in speakers, and that behavioral changes in speakers ultimately influence the listener's level of understanding in the conversation. In particular, the Blur effect was found to be effective in triggering behavioral changes in speakers without causing extra cognitive load on speakers. However, the timing of the appearance of the visual effect did not necessarily match with the listener's actual comprehension, which cause frustration on some listeners. Thus, our next step would be to integrate these listener's physical movements with the visual effect triggers so that the visual effect can be applied when the listener is unable to understand the conversation.

The future potential use of Communi-gate is for online classes with a mix of international students in educational institutions, international academic conference, and online medical service for foreign residents or travelers as mentioned in Chapter.3.6. But moreover, we believe this tool is not limited for foreign language users but can also be used for other users. For example, it could be useful as a classroom support tool for students with hearing impairments. With the recent increase in online classes, students with hearing impairments, who used to converse through physical movements and detailed facial expressions, are unable to read the conversations of professors and other students well and cannot keep up with the class [55–57]. Thus, by facilitating remote communication between people of different language and physical capabilities with our Communi-gate, we can expect to see even more active co-creation across those language and physical barriers.

References

[1] Toru Ishida Naomi Yamashita, Andy Echenique and Ari Hautasaari. Lost in transmittance: How transmission lag enhances and deteriorates multilingual collaboration. pages 923–934. ACM Conference on Computer Supported Cooperative Work Social Computing Social Computing, 2013. URL: https://doi.org/10.1145/2441776.2441881, doi:doi.org/10.1145/2441776.2441881.

[2] Takeshi Nishida. Promoting intercultural awareness through native-to-foreign speech accent conversion. CABS '14, page 83–86, New York, NY, USA, 2014. Association for Computing Machinery. URL: https://doi.org/10.1145/2631488.2634058, doi:10.1145/2631488.2634058.

[3] Wen Duan, Naomi Yamashita, and Susan R. Fussell. Increasing native speakers' awareness of the need to slow down in multilingual conversations using a real-time speech speedometer. *Proc. ACM Hum.-Comput. Interact.*, 3(CSCW), November 2019. URL: https://doi.org/10.1145/3359273, doi:10.1145/3359273.

[4] Shiro Kitayama and Kazushi Nishimoto. Daflingual:a supporting method to improve english conver-sation skill by applying delayed auditory feedback. *IPSJInteraction*, pages 894–896, 2017. URL: http://hdl.handle.net/10119/15113, arXiv:http://hdl.handle.net/10119/15113.

[5] Pooja Mathur and Karrie Karahalios. Visualizing remote voice conversations. In *CHI '09 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '09, page 4675–4680, New York, NY, USA, 2009. Association for Computing Machinery. URL: https://doi.org/10.1145/1520340.1520719, doi:10.1145/1520340.1520719.

[6] Joan Morris DiMicco, Katherine J. Hollenbach, and Walter Bender. Using visualizations to review a group's interaction dynamics. In *CHI '06 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '06, page 706–711, New York, NY, USA, 2006. Association for Computing Machinery. URL: `https://doi.org/10.1145/1125451.1125594`, `doi:10.1145/1125451.1125594`.

[7] Keisuke Seta, Masanori Yokoyama, Shigeo Yoshida, Takuji Narumi, Tomohiro Tanikawa, and Michitaka Hirose. Divided presence: Improving group decision-making via pseudo-population increase. In *Proceedings of the 6th International Conference on Human-Agent Interaction*, HAI '18, page 260–268, New York, NY, USA, 2018. Association for Computing Machinery. URL: `https://doi.org/10.1145/3284432.3284443`, `doi:10.1145/3284432.3284443`.

[8] Naoto Nakazato, Shigeo Yoshida, Sho Sakurai, Takuji Narumi, Tomohiro Tanikawa, and Michitaka Hirose. Smart face: Enhancing creativity during video conferences using real-time facial deformation. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work and Social Computing*, CSCW '14, page 75–83, New York, NY, USA, 2014. Association for Computing Machinery. URL: `https://doi.org/10.1145/2531602.2531637`, `doi:10.1145/2531602.2531637`.

[9] Prasanth Murali, Javier Hernandez, Daniel McDuff, Kael Rowan, Jina Suh, and Mary Czerwinski. Affectivespotlight: Facilitating the communication of affective responses from audience members during online presentations. *arXiv preprint arXiv:2101.12284*, 2021.

[10] Wei Sun, Yunzhi Li, Feng Tian, Xiangmin Fan, and Hongan Wang. How presenters perceive and react to audience flow prediction in-situ: An explorative study of live online lectures. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW):1–19, 2019.

[11] Robin Shields. Globalization and international student mobility: A network analysis. *Comparative Education Review*, 57(4):609–636, 2013.

[12] IIE Press Office. Number of International Students in the United States Hits All-Time High. https://www.iie.org/Why-IIE/Announcements/2019/11/Number-of-International-Students-in-the-United-States-Hits-All-Time-High, 2019. [Online; accessed 30-May-2021].

[13] The Yomiuri Shimbun. More international students to be admitted to Japan's national universities. https://the-japan-news.com/news/article/0006900930, 2020. [Online; accessed 30-May-2021].

[14] The Japan Student Services Organization (JASSO). Survey Results of Foreign Students' Status of Residence in 2019. https://www.studyinjapan.go.jp/ja/statistics/zaiseki/data/2019.html, 2020. [Online; accessed 30-May-2021].

[15] Yohtaro Takano and Akiko Noda. A temporary decline of thinking ability during foreign language processing. *Journal of Cross-Cultural Psychology*, 24(4):445–462, 1993.

[16] Ann R Bradlow and Tessa Bent. Perceptual adaptation to non-native speech. *Cognition*, 106(2):707–729, 2008.

[17] Jessica Gasiorek. Perspective-taking, inferred motive, and perceived accommodation in nonaccommodative conversations. *Journal of Language and Social Psychology*, 34(5):577–586, 2015.

[18] Willem J. Levelt. *Speaking: From intention to articulation (Vol. 1)*. MIT press, 1993.

[19] Asfia Khan and Rahma Al-Mahrooqi. Foreign language communication anxiety (flca) among tertiary level omani efl learners. *The Asian EFL Journal Quarterly March 2015 Volume 17, Issue*, 57, 2015.

[20] Lisa M Guntzviller, Robert N Yale, and Jakob D Jensen. Foreign language communication anxiety outside of a classroom: Scale validation and curvilinear relationship with foreign language use. *Journal of Cross-Cultural Psychology*, 47(4):605–625, 2016.

[21] Chien Wen Yuan, Leslie D. Setlock, Dan Cosley, and Susan R. Fussell. Understanding informal communication in multilingual contexts. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*, CSCW '13, page 909–922, New York, NY, USA, 2013. Association for Computing Machinery. URL: https://doi.org/10.1145/2441776.2441880, doi:10.1145/2441776.2441880.

[22] J. Storck and L. Sproull. Through a glass darkly what do people learn in videoconferences. *Human Communication Research*, 22:197–219, 1995.

[23] Helen Ai He, Naomi Yamashita, Ari Hautasaari, Xun Cao, and Elaine M. Huang. Why did they do that? exploring attribution mismatches between native and non-native speakers using videoconferencing. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, CSCW '17, page 297–309, New York, NY, USA, 2017. Association for Computing Machinery. URL: https://doi.org/10.1145/2998181.2998205, doi:10.1145/2998181.2998205.

[24] Leslie M. Beebe and Howard Giles. Speech-accommodation theories: A discussion in terms of second- language acquisition. 46, pages 5–32. International Journal of the Sociology of Language, 1984. URL: https://doi.org/10.1515/ijsl.1984.46.5, doi:doi.org/10.1515/ijsl.1984.46.5.

[25] Ge Gao, Naomi Yamashita, Ari MJ Hautasaari, Andy Echenique, and Susan R Fussell. Effects of public vs. private automated transcripts on multiparty communication between native and non-native english speakers. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 843–852, 2014.

[26] Helen Ai He, Naomi Yamashita, Chat Wacharamanotham, Andrea B. Horn, Jenny Schmid, and Elaine M. Huang. Two sides to every story: Mitigating intercultural conflict through automated feedback and shared self-reflections in global virtual teams. *Proc. ACM Hum.-Comput. Interact.*, 1(CSCW), December 2017. URL: https://doi.org/10.1145/3134686, doi:10.1145/3134686.

[27] Mei-Hua Pan, Naomi Yamashita, and Hao-Chuan Wang. Task rebalancing: Improving multilingual communication with native speakers-generated highlights on automated transcripts. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, CSCW '17, page 310–321, New York, NY, USA, 2017. Association for Computing Machinery. URL: `https://doi.org/10.1145/2998181.2998304`, `doi:10.1145/2998181.2998304`.

[28] Simon Arnfield, Peter Roach, Jane Setter, Peter Greasley, and Dave Horton. Emotional stress and speech tempo variation. In *Speech under Stress*, 1995.

[29] Albert Mehrabian and Martin Williams. Nonverbal concomitants of perceived and intended persuasiveness. *Journal of Personality and Social psychology*, 13(1):37, 1969.

[30] Emma Sweeney and Zhu Hua. Accommodating toward your audience: Do native speakers of english know how to accommodate their communication strategies toward nonnative speakers of english? *The Journal of Business Communication (1973)*, 47(4):477–504, 2010.

[31] Ge Gao, Naomi Yamashita, Ari MJ Hautasaari, Andy Echenique, and Susan R. Fussell. Effects of public vs. private automated transcripts on multiparty communication between native and non-native english speakers. CHI '14, page 843–852, New York, NY, USA, 2014. Association for Computing Machinery. URL: `https://doi.org/10.1145/2556288.2557303`, `doi:10.1145/2556288.2557303`.

[32] Amy Soller, Alejandra Martínez, Patrick Jermann, and Martin Muehlenbrock. From mirroring to guiding: A review of state of the art technology for supporting collaborative learning. *Int. J. Artif. Intell. Ed.*, 15(4):261–290, December 2005.

[33] Taemie Jung Kim. *Enhancing distributed collaboration using sociometric feedback*. PhD thesis, Massachusetts Institute of Technology, 2011.

[34] Daniel Olguín Olguín. *Sociometric badges: wearable technology for measuring human behavior*. PhD thesis, Massachusetts Institute of Technology, 2007.

[35] Janienke Sturm, Olga Houben-van Herwijnen, Anke Eyck, and Jacques Terken. Influencing social dynamics in meetings through a peripheral display. In *Proceedings of the 9th International Conference on Multimodal Interfaces*, ICMI '07, page 263–270, New York, NY, USA, 2007. Association for Computing Machinery. URL: `https://doi.org/10.1145/1322192.1322238`, `doi:10.1145/1322192.1322238`.

[36] Teresa M Amabile, Sigal G Barsade, Jennifer S Mueller, and Barry M Staw. Affect and creativity at work. *Administrative science quarterly*, 50(3):367–403, 2005.

[37] Alice M Isen, Kimberly A Daubman, and Gary P Nowicki. Positive affect facilitates creative problem solving. *Journal of personality and social psychology*, 52(6):1122, 1987.

[38] Roser Cañigueral and Antonia F de C Hamilton. The role of eye gaze during natural social interactions in typical and autistic people. *Frontiers in Psychology*, 10:560, 2019.

[39] Jinni A Harrigan. Listeners'body movements and speaking turns. *Communication Research*, 12(2):233–250, 1985.

[40] Qiang Xu, Yaping Yang, Qun Tan, and Lin Zhang. Facial expressions in context: Electrophysiological correlates of the emotional congruency of facial expressions and background scenes. *Frontiers in Psychology*, 8:2175, 2017.

[41] Tony Ro, Charlotte Russell, and Nilli Lavie. Changing faces: A detection advantage in the flicker paradigm. *Psychological science*, 12(1):94–99, 2001.

[42] Katsumi Watanabe. Behavioral speed contagion: automatic modulation of movement timing by observation of body movements. *Cognition*, 106(3):1514–1524, 2008.

[43] Makoto Ichikawa. Temporal illusions in visual perception. *Kougaku*, 39(2):82–88, 2010.

[44] Jennifer Chesters, Ladan Baghai-Ravary, and Riikka Möttönen. The effects of delayed auditory and visual feedback on speech production. *The Journal of the Acoustical Society of America*, 137(2):873–883, 2015.

[45] M Teraoka and Y Muranaka. Aspects of cross-cultural experience perceived by foreigners living in japan when using its healthcare services. *J Jpn Acad Nurs Sci*, 37:35–44, 2017.

[46] Moe Sugawa, Taichi Furukawa, George Chernyshov, Danny Hynds, Jiawen Han, Marcelo Padovani, Dingding Zheng, Karola Marky, Kai Kunze, and Kouta Minamizawa. Boiling mind: Amplifying the audience-performer connection through sonification and visualization of heart and electrodermal activities. In *Proceedings of the Fifteenth International Conference on Tangible, Embedded, and Embodied Interaction*, pages 1–10, 2021.

[47] Jackson Beatty. Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological bulletin*, 91(2):276, 1982.

[48] Sophia E Kramer, Theo S Kapteyn, Joost M Festen, and Dirk J Kuik. Assessing aspects of auditory handicap by means of pupil dilatation. *Audiology*, 36(3):155–164, 1997.

[49] Carol L Mackersie and Heather Cones. Subjective and psychophysiological indexes of listening effort in a competing-talker task. *Journal of the American Academy of Audiology*, 22(2):113–122, 2011.

[50] Nagahama Toru, Miyanishi Ukako, and Yosuke Morita. A trial on measurement of stress using finger plethysmogram on learning activity. 41(4):149–152, 2018.

[51] Viorica Marian, Henrike K Blumenfeld, and Margarita Kaushanskaya. The language experience and proficiency questionnaire (leap-q): Assessing language profiles in bilinguals and multilinguals. *Journal of speech, language, and hearing research*, 2007.

[52] Masashi Nakamura, Harumi Ikei, and Yoshifumi Miyazaki. Physiological effects of visual stimulation with full-scale wall images composed of verti-

cally and horizontally arranged wooden elements. *Journal of Wood Science*, 65(1):1–11, 2019.

[53] Barney G Glaser, Anselm L Strauss, and Elizabeth Strutzel. The discovery of grounded theory; strategies for qualitative research. *Nursing research*, 17(4):364, 1968.

[54] Jorn Bakker, Mykola Pechenizkiy, and Natalia Sidorova. What's your current stress level? detection of stress patterns from gsr sensor data. In *2011 IEEE 11th international conference on data mining workshops*, pages 573–580. IEEE, 2011.

[55] NHK. What is important in communicating with the hearing impaired? `https://www3.nhk.or.jp/news/html/20210203/k10012846461000.html`, 2021. [Online; accessed 28-June-2021].

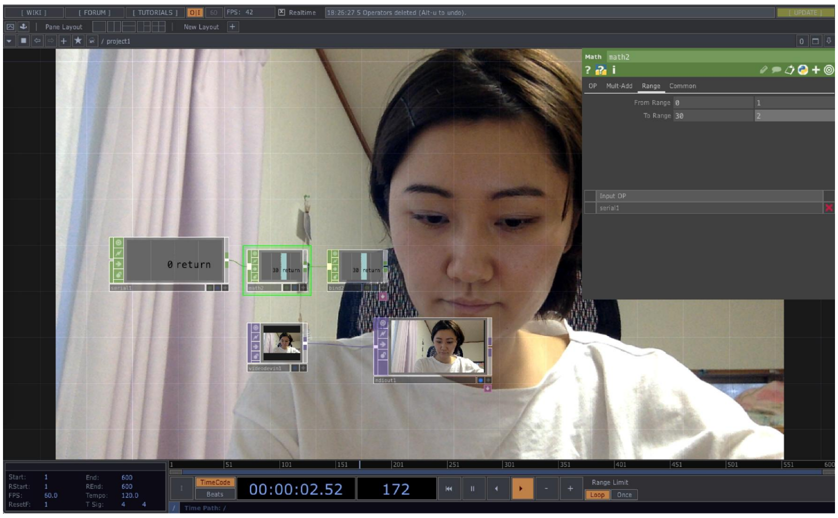
[56] The Mainichi. New Corona: Distance Learning for the Hearing Impaired. `https://mainichi.jp/articles/20200806/ddm/013/040/021000c`, 2020. [Online; accessed 28-June-2021].

[57] Waseda Weekly. Online classes for students with disabilities: Remote support from Yamagata and other locations. `https://www.waseda.jp/inst/weekly/feature/2020/11/30/80785/`, 2020. [Online; accessed 28-June-2021].

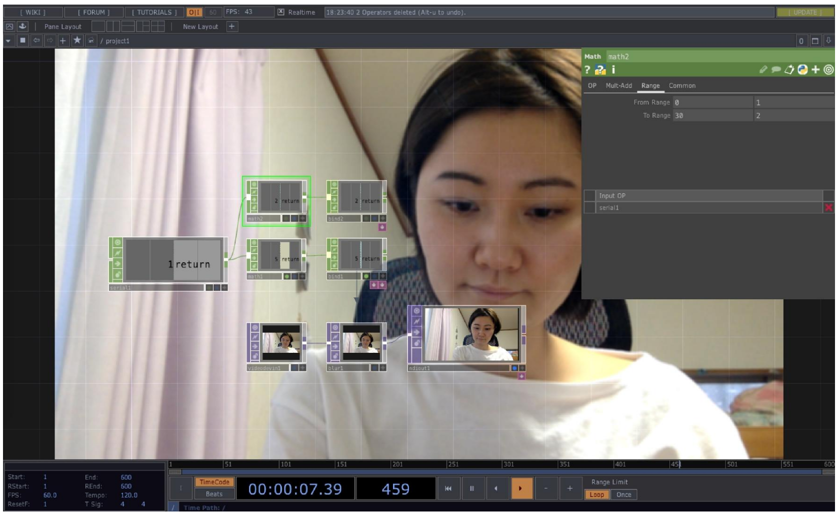
Appendices

A. TouchDesigner Patch of Visual Effects

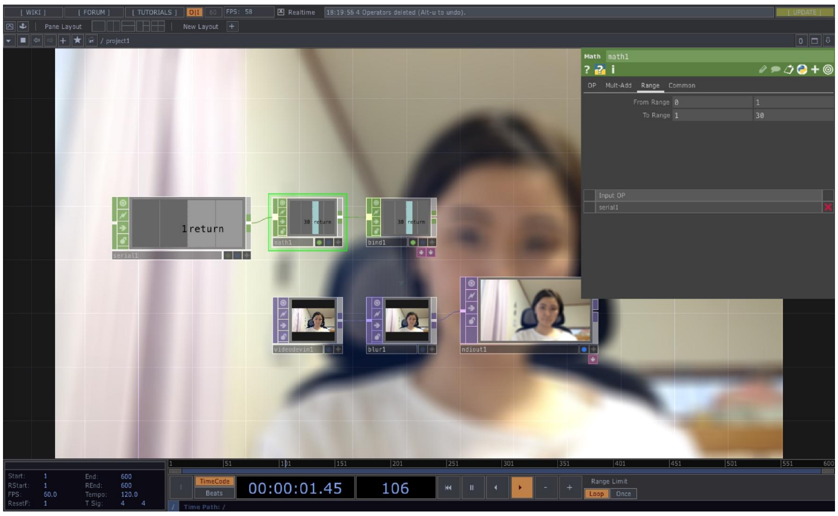
A.1 TouchDesigner Patch of Freeze and Frame Drop effect



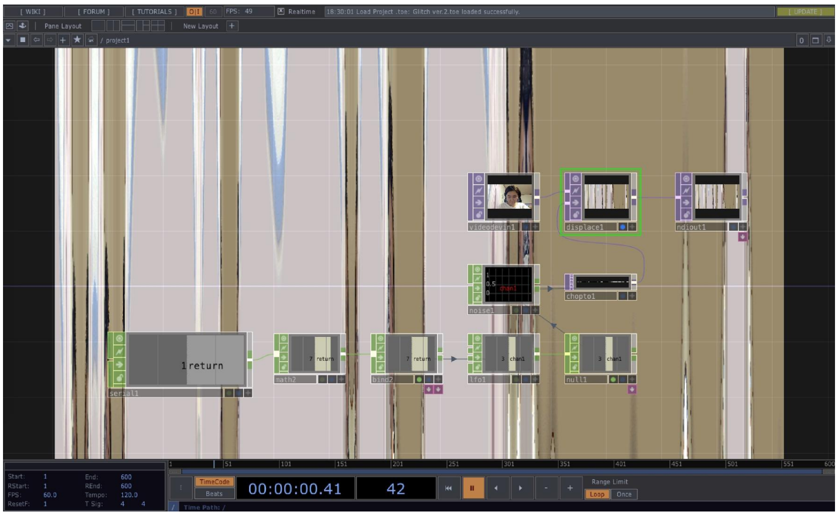
A.2 TouchDesigner Patch of Slow-motion effect



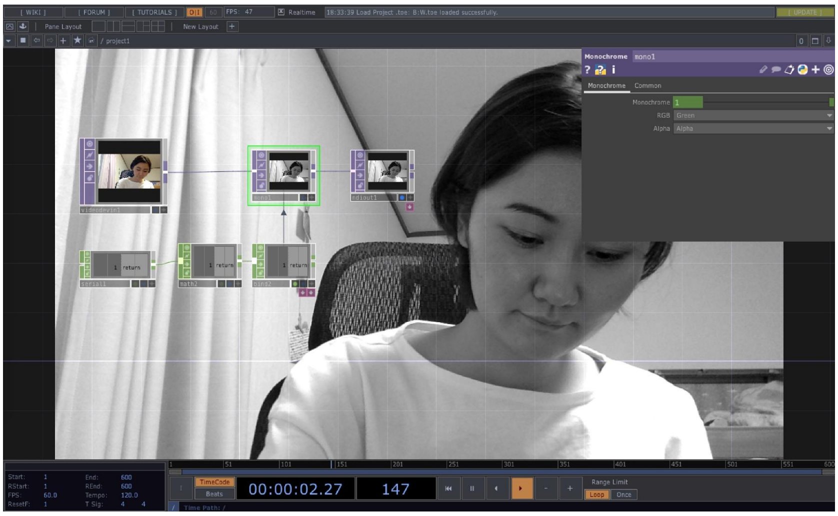
A.3 TouchDesigner Patch of Blur effect



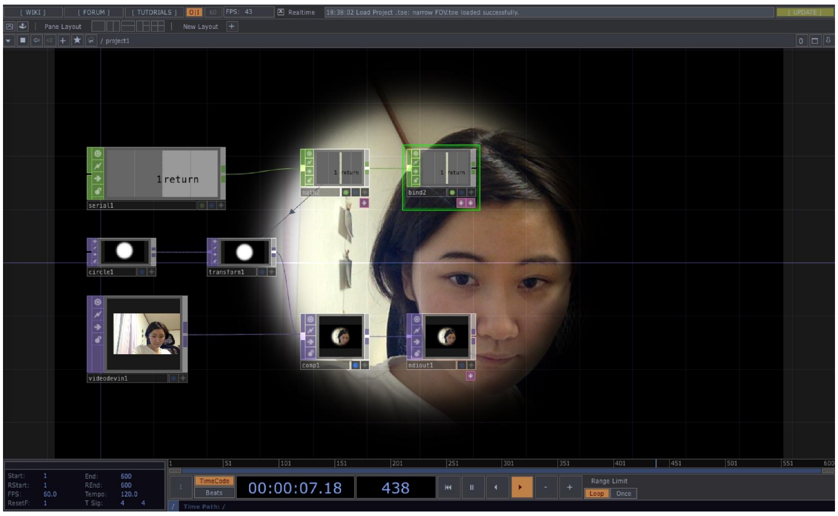
A.4 TouchDesigner Patch of Glitch effect



A.5 TouchDesigner Patch of Black and White effect



A.6 TouchDesigner Patch of NarrowFOV effect



A.7 TouchDesigner Patch of Flash effect

B. Code of TouchDesigner in DAF

B.1 Code of tcpip1_callbacks

```
# me - this DAT
#
# dat - the DAT that received the data
# rowIndex - the row number the data was placed into
# message - an ascii representation of the data
#                       Unprintable characters and unicode characters will
#                       not be preserved. Use the 'bytes' parameter to get
#                       the raw bytes that were sent.
# bytes - a byte array of the data received
# peer - a Peer object describing the originating data
#   peer.close()        #close the connection
#       peer.owner      #the operator to whom the peer belongs
#       peer.address    #network address associated with the peer
#       peer.port               #network port associated with the peer

def onConnect(dat, peer):
        peer.sendBytes("0.0")

        return

def onReceive(dat, rowIndex, message, bytes, peer):

        return

def onClose(dat, peer):
        return
```

B.2 Code of bind1_callbacks

```
# me - this DAT.
#
# bindOp - the connected Bind CHOP
```

```
# chan - the modified Channel
# val - the current value
# prev - the previous value
# source - the source of the change (input Channel or reference parameter)
#

def onValueChange(bindOp, chan, val, prev, source):
        op('tcpip1').send(val)
        return
```