

Evaluation of Photo Contents of Conversation Support System with Protocol Analysis Method

Zhou Xiaochun, Miyuki Iwamoto, Noriaki Kuwahara
Graduate School of Engineering and Science
Kyoto Institute of Technology
Kyoto, Japan

Abstract—With the deepening of aging and low birth rate in China, the solitary elderly or old couple living alone is becoming more and more, who has a higher risk of senile dementia caused by disuse of cognitive function because of loneliness without communication. Due to the shortage of care workers, the young volunteer is expected becoming communication partner for them. But it is difficult for the young volunteer without the experience of communication with the elderly, and for two generations to find common topics. However, Conversation Support System was proposed so that the elderly and the young volunteer can talk smoothly with common photo contents. In order to evaluate the utility of photo contents of the system in China, we did the conversation experiment by photos in China, to analyze the expression and stress of subjects during the conversation. As a result, the photos which made the elderly and the young volunteer feel easy and difficult for the conversation were found. Then we analyzed utterance data of subjects with protocol analysis method to discuss the common features of these photos.

Keywords—Protocol analysis method; the elderly; the young volunteer; photo contents; conversation

I. INTRODUCTION

A. Situation of Aging and Low Birth Rate in China

In China, the number of the elderly aged 65 and over was 158.31 million in 2017, and the percentage of the elderly rises to 11.4%. Meanwhile, China is in the face of low birth rate as the result of “One Child Policy”. In 2017, China’s population increased by 172.3 million, a year-on-year decrease of 630,000, and the birth rate dropped from 1.295 percent in 2016 to 1.243 percent in 2017 [1].

B. Shortage of Nursing Care Staff in China

The National Medium and Long-Term Development Plan for Civilian Talents (2010-2020) proposed that the number of nursing care staff should grow from 30,000 in 2010 to 6 million in 2020 [2]. However, by the end of 2016, the total number of skillful nursing care staff was only 12,144 [3].

C. Related Works

With the deepening of aging and low birth rate in China, the solitary elderly or old couple living alone is more and more, who has a higher risk of senile dementia caused by disuse of cognitive function because of loneliness without communication. However, due to the severe shortage of nursing care staff, the young volunteer is expected becoming communication partner for the elderly. But it is difficult for the young volunteer without any care and communication

experience. Co-imagination as a communication support method for prevention of dementia was proposed by Otake in 2006, which should support daily interactive communication between people without care experience and the elderly with images [4]. However, Iwamoto Miyuki investigated the Conversation Support System, which was expected that the young volunteer would feel less stress when talking with the elderly by common contents of photographs and videos [5]. Iwamoto decided to use photos as the contents of this system, because from the experimental results so far, although the video made the subjects feel less pressure, the photo made the elderly and young volunteers talk longer [5].

D. Photo Contents of System

Referring to related works on photo contents for Intergenerational Conversation Support System in Japan, the photo categories of “Food” and “Events” are common to any generation [6]. Meanwhile, from the results of the preliminary experiment in China, Chinese elders were interested in photo categories of “School” and “Commodity” [7]. Therefore, we selected photo categories of “Food”, “Events”, “School” and “Commodity” as contents of the system in this study, and for each category, we used ten photos in China for the conversation.

E. Protocol Analysis

Protocol analysis is a method of extracting the problem of the system by analyzing the utterance data (protocol) of the user while using the system [8]. In this study, the subjects are the elderly and young beginners who speak slowly when looking at the photo not seen before. Protocol Analysis is an effective method because utterance data of subjects is easy to be obtained.

F. Research Objectives

The goal of this study is to construct a system making the Chinese elderly and young volunteer talk smoothly by photo contents. In order to find the common features of optimal photos for the system, by which the elderly should talk smoothly, and the young volunteer should feel no stress in conversation, we evaluated each photo for each category with protocol analysis method.

II. EXPERIMENT

A. Summary

In this experiment, the elderly and the young volunteer talked each other while looking at the common photos. We

examined the degree of stress for the young volunteer, and the expression of the elderly depending on each photo in conversation.

The purpose of this experiment is to evaluate the photo categories which make the conversation between two generations easy or difficult.

B. Evaluation Item

1) Stress Check for Young Volunteer for Photo

The young indicated the degree of stress on the stress check sheet every minute in conversation. The stress check sheet represented a 1-7-scale. 1 means the least stress and 7 means the most stress (Fig. 1).

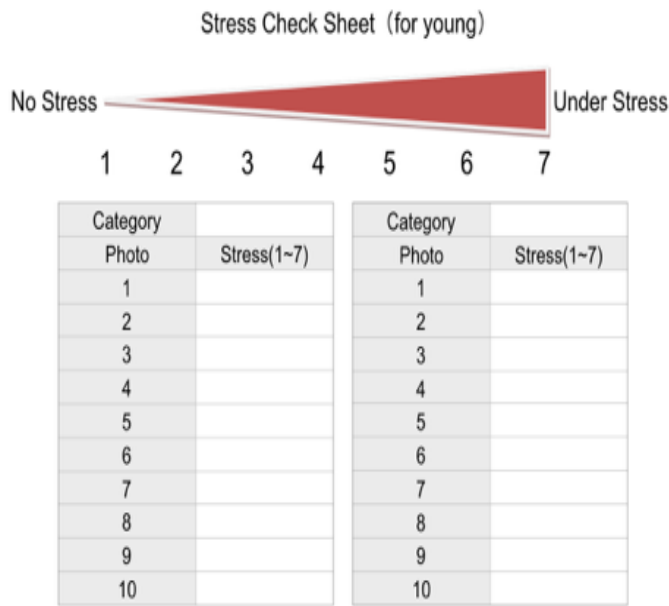


Fig. 1. Stress check sheet.

2) Questionnaire for Category

The elderly and the young answered ten questions by 5-stage subjective evaluation each time after each category was finished.

3) Expression Analysis for the Elderly for Photo

We analyzed the expressions of the elderly from the video recordings. We used the degree of happiness or unhappiness as a result of each photo.

C. Subject

The conversation participants were seven young female caregivers of Suzhou Social Welfare Home, and four senior women and two senior men without dementia living there. They don't know each other before.

D. Environment of Experiment

Because the function of eyes is often kept better than ears for the elderly, we brought a photo into close-up and displayed on the screen by the projector. The elderly and the young talked to each other while looking at the screen. The layout of the experiment was shown in Fig. 2 and 3.



Fig. 2. Experimental environment.

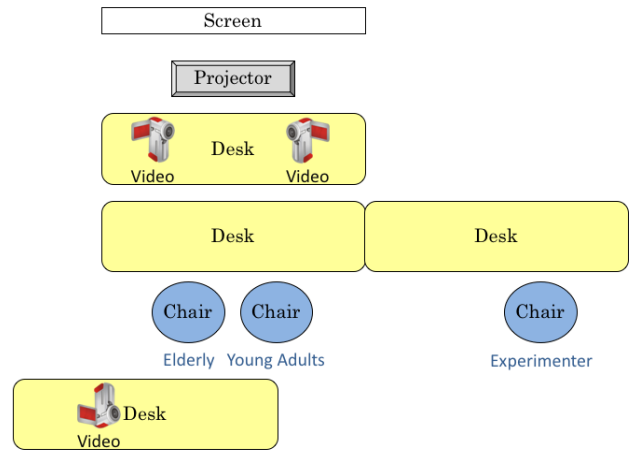


Fig. 3. The layout of the experimental environment.

- We used a meeting room, in which we placed desks and chairs side by side.
- We used a MacBook in which photos were uploaded for a 10-minutes conversation session.
- We used a camera to capture the expression of the elderly throughout the sessions.

E. Materials

This experiment used photo categories of “Food”, “Events”, “School”, and “Commodity”. Each category was with ten photos.

F. Methods

- We prepared four photo categories of “Food”, “Events”, “School” and “Commodity”, and ten photos for each category by Keynote.
- One elderly people and one young volunteer sat down side by side and faced to the screen. Then we displayed each photo on the screen for 1 minute. They were asked to talk about each photo for 1 minute.
- The young volunteer indicated the degree of stress on the stress check sheet every minute for each photo.
- Both the elderly and young volunteer answered the questionnaire for category each time after the conversation was over.
- A camera was used to capture the expression of the elderly throughout the sessions.

- The expression of the elderly was analyzed from the video recordings. We used the major literature “Expression Analysis” techniques to know which photo made the subject feel happy [9]. We framed the expression images of each category every second. Frame images for the number of conversation hours were classified by their respective pictures and analyzed by using the frame images. By making it into a frame, it is possible to delete facial expressions of parts other than the conversation in the video, and it can be thought that picture and emotion can be related more. We evaluated the degree of the smile on a frame-by-frame basis. We define 0% as a state of expressionlessness, and 100% as the state of the highest degree of smile. We used the highest degree of the smile as a result of each photo.

III. PROTOCOL ANALYSIS METHOD

Table I shows an example of protocol analysis method.

TABLE I. EXAMPLE OF PROTOCOL ANALYSIS

Speaker	Utterance Data		
	Utterance Contents	Utterance Amounts	Speaking Time
Young	It’s a blackboard.	1	1.3”
Elderly	Is it...?	1	0.9”
Young	Lei Feng	1	0.9”
Elderly	En, Lei Feng	2	1.4”
Young	Is it that Learn-from-Lei Feng a few years ago?	1	3.6”
Elderly	Yes yes.	1	1”
Young	We had to Learn from Lei Feng when we were children.	1	2”
Elderly	Yes yes.	1	1.3”

- Transcribed all utterances of subjects from the video recording, while “Speaker”, “Utterance Contents”, “Utterance Amounts”, “Speaking Time” are recorded for each photo. The speaking time used the time code of video recording.
- Extracted out photos that were supposed to be easy and difficult for the conversation on the results of stress check, expression analysis, and speaking time.
- Extracted out the key topic of photos above from the transcribed utterance protocol on the results of utterance amounts, which was classified to estimate their common features.

IV. RESULT

1) Questionnaire for Category

From the results of the questionnaire, we knew that all photo categories of “Food”, “Events”, “School”, and “Commodity” got the less stress for subjects under stage 3, which can help them talk easily [7].

2) Stress Check & Expression Analysis for Photo

Fig. 4 shows the relationship between the stress of young volunteer and expression of the elderly for each photo. The photos with both of no stress and happy are “Children’s Day”, “Duanwu Festival”, “Spring Festival Evening”, “Popcorn”, “Soy Milk”, “Fried Rice”, and “Television”. On the contrary, the extreme unhappy photo is “Dumpling”.

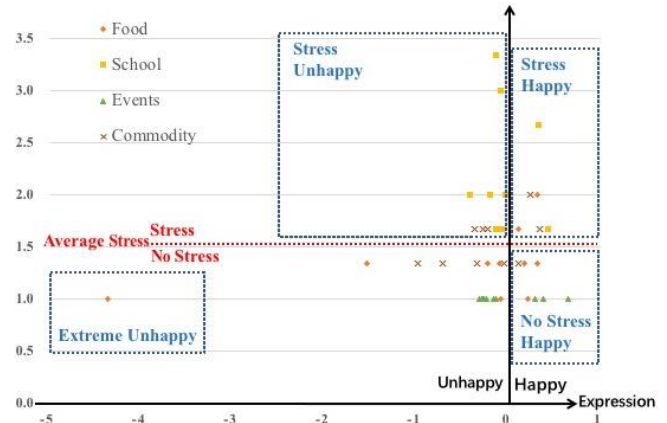


Fig. 4. Relationship between stress and expression.

3) Speaking Time for Photo

Fig. 5 shows the percentage of speaking time of the elderly for “Food” category. Because each conversation unit for each photo is the same 60 seconds, and the average speaking time of the elderly for each photo is 20.3 seconds. The percentage is over 34% (20.3 seconds) which is supposed to be good photo making conversation easy. From this figure, it is known that all photos in “Food” category are good.

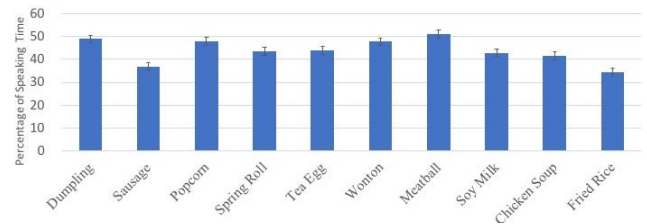


Fig. 5. Percentage of speaking time of the elderly for food.

Fig. 6 shows the percentage of speaking time of the elderly for “Events” category. From this figure, it is known that only photo of “Spring Festival Evening” is good.

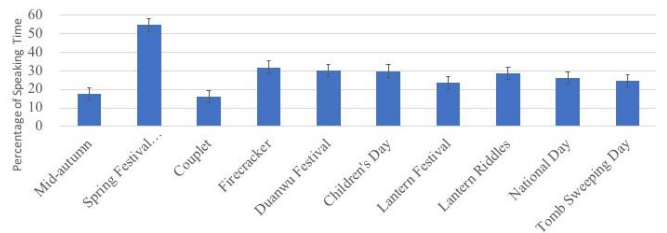


Fig. 6. Percentage of speaking time of the elderly for events.

Fig. 7 shows the percentage of speaking time of the elderly for “School” category. From this figure, it is known that only photo of “School Bag” is good. On the contrary, the extreme short speaking time is “Table-tennis”.

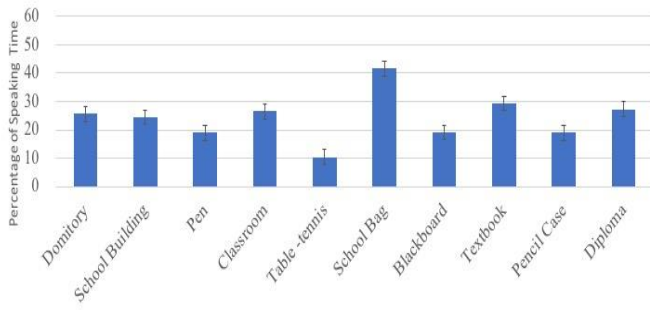


Fig. 7. Percentage of speaking time of the elderly for school.

Fig. 8 shows the percentage of speaking time of the elderly for “Commodity” category. From this figure, it is known that photos of “Sewing Machine”, “Lunch Box”, “Bicycle”, “Kerosene Lamp”, “Cold Cream”, “Quit”, “Hot Pot”, “Small Coal” are good.

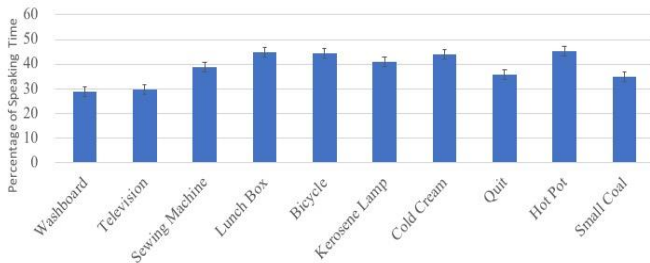


Fig. 8. Percentage of speaking time of the elderly for commodity.

Even if all categories are easy for conversation, the result of each photo is very different, from which we get the best photos with no stress, happy, and speaking time above 34%, that is “Spring Festival Evening” (Fig. 9), “Popcorn” (Fig. 10), “Soy Milk” (Fig. 11), and “Fried Rice” (Fig. 12). On the contrary, we also get the extreme unhappy photo of “Dumpling” (Fig. 13), and extreme short speaking time of “Table-tennis” (Fig. 14).



Fig. 10. Popcorn.



Fig. 11. Soy milk.



Fig. 9. Spring festival evening.



Fig. 12. Fried rice.



Fig. 13. Dumpling.



Fig. 14. Table-tennis.

4) Utterance Amounts for Photo

We analyzed the utterance contents of photos above and extracted out key topics on the result of utterance amounts.

Fig. 15 shows the percentage of utterance amounts of key topic for each photo.

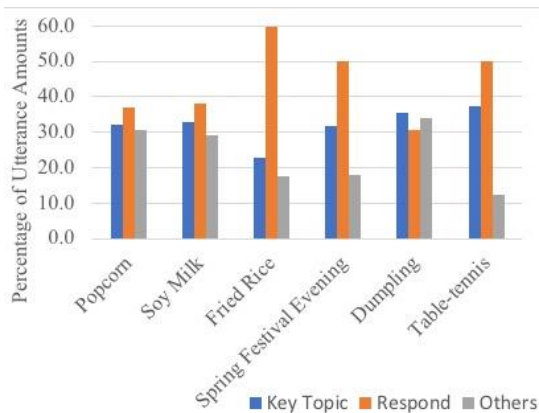


Fig. 15. Percentage of utterance amounts of key topic.

a) *Spring Festival Evening*: The most utterance amounts of the elderly are responding to the young partner. Except it, the key topic with the most utterance amounts concerns “Presenters of Spring Festival Evening”.

b) *Popcorn*: The most utterance amounts of the elderly are responding to the young partner. Except it, the key topic with the most utterance amounts concerns “How to Make Popcorn”.

c) *Soy Milk*: The most utterance amounts of the elderly are responding to the young partner. Except it, the key topic with the most utterance amounts concerns “How to Make Soy Milk”.

d) *Fried Rice*: The most utterance amounts of the elderly are responding to the young partner. Except it, the key topic with the most utterance amounts concerns “How to Make Fried Rice”.

e) *Dumpling*: The most utterance amounts of the elderly are key topic concerning “Northerners Like Dumpling” and “Southerners Don’t Like Dumpling”.

f) *Table-tennis*: The most utterance amounts of the elderly are responding to the young partner. Except it, the key topic with the most utterance amounts concerns “Countryside” and “Shabby”.

V. DISCUSSION

We classified the key topics of photos above:

A. Easy for the Conversation

1) Making Food Method

For the photo of “Popcorn”, “Soy Milk”, and “Fried Rice”, the key topic with the most utterance amounts of the elderly is that how to make the food and materials for making.

2) Recent Famous People

For the photo of “Spring Festival Evening”, the key topic with the most utterance amounts of the elderly is presenters of Spring Festival Evening, who are recent famous people in China.

From the result that the most utterance amounts of the elderly are responding to the young partner, we knew that these topics also made the most interaction between the elderly and the young volunteer.

B. Difficult for the Conversation

1) Different Area

For the extreme unhappy photo of “Dumpling”, the key topic with the most utterance amounts of the elderly is that northerners like the dumpling, not southerners. In China, different area has wholly different customs.

2) Shabby Things

For the photo of “Table-tennis”, the elderly talked few, only about what it looks like the countryside, and everything is shabby. Although photos of 1960’s are nostalgic, the elderly don’t like shabby things of that time.

VI. CONCLUSION

From this experiment, we got the common features of photos making the elderly feel easy to talk with the young volunteer as follows:

- Recent famous people.
- Diet can be done oneself.

This conclusion is similar to the related work on photo contents in Japan that regardless of age, what is familiar in the current environment can become a category of common

interest for the young people and the elderly. For example, when the young people see food they enjoy, the elderly can teach how to make it. This is a good topic that makes both the elderly and the young people feel happy [6].

Meanwhile, we have to avoid the features making conversation difficult:

- Different area.
- Shabby things.

This conclusion is also similar to the related work in Japan that in conversations, the burden of the young people will increase when there is a category with less commonality between the young people and the elderly [6].

For examining the common features of photo contents mentioned above, we should use more photos and do more experiments in China.

VII. FUTURE WORKS

In this study, we knew that the “Food”, “Events”, “School”, “Commodity” were suitable categories for conversation for the elderly in China, in which only four photos got the results with no stress, happy, and long speaking time. We should prepare more photos for the next experiment to examine more optimal photos and common features of them.

Because measure device is difficult for the elderly, in future research, we expect to evaluate the elderly and young volunteer’s emotional reaction to photos with sentiment analysis.

The young volunteers of this study were all the caregivers in this study. For the next experiment we should let the young

people without any care experience become the conversation partners.

ACKNOWLEDGMENT

This work was supported by JSPS KAKENHI Grant Number 15H01698.

REFERENCES

- [1] Statistical Communiqué of the People's Republic of China on the 2017 National Economic and Social Development, National Bureau of Statistics of China, 2018.
- [2] The National Medium and Long-Term Development Plan for Civilian Talents (2010-2020), Ministry of Civil Affairs of the People’s Republic of China, 2011.
- [3] China Civil Affairs’ Statistical Yearbook (Statistics of China Social Services), Ministry of Civil Affairs of the People’s Republic of China, 2017, p.134.
- [4] M. Otake, M. Kato, T. Takagi, and H. Asama, “Coimagination Method: Communication Support System with Collected Images and its Evaluation via Memory Task”, *Universal Access in Human-Computer Interaction Constantine Stephanidis (Ed.)*, LNCS 5614, Springer-Verlag, 2009.
- [5] M. Iwamoto, N. Kuwahara, K. Morimoto, “Comparison of Burden on Youth in Communicating with Elderly using Images Versus Photographs”, vol.6, No.10, IJACSA, 2015.
- [6] M. Iwamoto, N., Kuwahara, K., Morimoto, “A Study of Conversation Support System between the Elderly Person and Young Adults by Using Facial Expression Analysis”, Springer-Verlag, 2011.
- [7] X. Zhou, M. Iwamoto, N. Kuwahara, K. Morimoto, “A Study of Photographs as Communication Content for Intergenerational Conversation Support System,” *Lecture Notes in Computer Science*, vol 10287, Springer, Cham, 2017.
- [8] H. Kaioho, E. Harada, *Introduction to Protocol Analysis*. Shin-Yo-Sha, Japan, 1993.
- [9] T. Ota, M. Tamura, M. Arita, N. Kiso, Y. Saeki, “Expression Analysis-comparison of the Characteristics of the Facial Expression”, Ekman, 2010.