



## **How Hearing Impaired People View Closed Captions of TV Commercials Measured By Eye-Tracking Device**

**Takahiro Fukushima, Otemon Gakuin University, Japan**

**Takashi Yasuda, Dai Nippon Printing Co., Ltd., Japan**

**Mika Oda, Dai Nippon Printing Co., Ltd., Japan**

### **Abstract**

We conduct experiments to measure effects of various types of closed captions for Japanese TV commercials using eye-tracking device. The experiment materials are real TV commercials, and 92 hearing impaired people participate in the experiments. We prepare six TV commercials, each of them captioned in four different ways: no closed captions and three types of captions, which are fully captioned (horizontal), partially captioned (horizontal), and fully captioned vertically. The participants watch them in silence using eye-tracking device, accompanied by a questionnaire asking about caption readability for two CMs. We define Areas of Interest (AOIs) as the caption lines and record fixation counts in AOIs and visit counts to AOIs. The results show that there are no statistically significant differences among the three types of captions on both fixation and visit counts, which means the participants watch the CMs similarly. However, further analyses of the data suggest some characteristics of the way the hearing impaired participants view the CM captions. The questionnaire reveals that there are significant differences among the three types of captions for a detergent CM, where partially captioned CM scores higher than the others in readability.

## Introduction

There are two kinds of captions for TV programs and commercial messages (CMs). One is open captions, characters or texts are a part of video and thus open to all the viewers. The other is closed captions which are hidden and the viewers can watch them with operations of TV remote control, usually by pressing CC button of the remote. Closed captions are one of the important sources of information for hearing impaired and hard-of-hearing people. Ministry of Internal Affairs and Communications (MIC) in Japan has been promoting closed-captioning TV programs, and as a result, most of TV programs, 84.8% for NHK (the public TV station) and 95.5% for main five private TV stations, were closed captioned, according to the latest survey conducted by MIC (MIC statistics, 2014).

However, although about 20% of the TV programs by the key five private stations are commercial messages, they are not closed captioned at all, except a few commercials. Thus in order to increase captioned CMs, MIC recently started to investigate how to produce and promote captioning of CMs with cooperation of major TV stations, advertising agencies, and commercial sponsors in Japan. There have been just a few studies on closed captioned CMs in Japan (Fukushima & Inoue, 2012; Inoue, 2012) but they did clearly show the need and importance of closed-captioned CMs not only for hearing impaired, hard-of-hearing people but also senior citizens in the society. However the studies did not specify how it should be done. As a first step, we conducted a research experiment to measure effects of several types of closed-captioned Japanese TV commercials using eye-tracking device for hearing impaired people. Below, we explain the details of the experiment, experiment results as well as our observations and analyses of the results.

## Experiment Methods and Experiment Materials

We describe below the details of the experiments, the materials, and the participants, and the procedures. We used six real TV commercials (CMs) provided by Kao Corporation. They include two for laundry detergent and four for cosmetics. All of them are 15 second CMs. The details of the CMs are as follows:

	Product name	Content
CM 1	Attack Neo Reset Power	Laundry detergent
CM 2	Attack Neo Ex Bio Power	Laundry detergent
CM 3	AUBE couture	Cosmetics
CM 4	SOFINA whitening	Cosmetics
CM 5	SOFINA Beaute (milky lotion)	Cosmetics
CM 6	SOFINA Beaute (milky lotion UV protection)	Cosmetics

Table 1 Details of TV commercials

We selected a variety of CMs in terms of the following characteristics:

- Load level of the viewers
- Number of scenes

- Degree of scenes changes while captions are shown
- Degree of movements
- Number of scenes in which human face or faces are shown
- Number of open captions

Table 2 shows the characteristics of the six CMs.

	CM 1	CM 2	CM3	CM4	CM5	CM6
Load level	High	Small	High	High	Middle	Small
No of Scenes	12	6	18	13	11	9
Degree of Changes of Scenes	High	Low	High	High	High	Middle
Degree of Movements	High	Low	High	High	Middle	Middle
Face appearance (No. of Scenes)	5	4	14	9	6	5
No of Open Captions	7	5	7	6	6	5

Table 2 Characteristics of CMs

We have prepared four different types of captions for each CMs. First, CMs are NOT captioned at all, i.e. such CMs have only open captions (type 0). Second, CMs are closed-captioned according to conventional standard (type 1). Third, CMs are closed captioned but some captions are omitted where open captions are shown (type 2). In some CMs, open captions are provided for speech in the CMs, thus we do not need closed captions for such scenes. If we add closed captions in such CMs, the viewers will see the same characters or texts twice as both open and closed captions. The last type is vertical captions. In Japanese language, we use both horizontally and vertically aligned text in our daily life. As the fourth type, CMs are closed captioned vertically (type 3).

We divided the participants into four groups (A to D) each having 21 to 24 of them. The assignment of the CMs for each group is as follows:

	CM 1	CM 2	CM 3	CM 4	CM 5	CM 6
Type 0 (open caption only)	A	B	C	D	A	B
Type 1 (closed caption, conventional)	D	A	B	C	D	A
Type 2 (closed caption, some omitted)	C	D	A	B	C	D
Type 3 (closed caption, vertical)	B	C	D	A	B	C

Table 3 Assignment of CMs for four participant groups

Images from captioned CMs are shown below.



Figure 1 CM image (type1)



Figure 2 CM image (type2)



Figure 3 CM image (type3)

These CM images are taken from the experiment material. Please note that CM type2 (Figure 2) has no captions because open caption (text in the CM video) is provided, and also note that the dots in the images correspond to eye gaze points, pink is those of women and blue stands for men.

### Experiment Participants

The participants of the experiments are all hearing impaired and the total number is 92. In terms of gender, 48 of them, 52.2% are male and 44 or 47.8% are female. The breakdown in terms of age is as follows:

Age	Number	Percentage
30-39	10	10.9%
40-49	14	15.2%
50-59	26	28.3%
60-69	31	33.7%
70-79	10	10.9%
80-89	1	1.1%

Table 4 Ages of Participants

We also asked them how they communicate with hearing people, and 88 of them answered. Out of 88, 51 (58%) use lip reading and gestures, 62 (70.5%) need written information, and 69 (74.8%) rely on sign language.

### Experiment Producers

First, a participant watched a short video to adjust the eye-tracking device, which located between the participant and TV screen. Then each of the participants watched six CMs, variously captioned, once per CM. This was done silently without any sound. Lastly, they were asked to recall what they remembered about the CMs. The types and order of CMs were decided according to the assignment (Table 3). We used written texts on paper board for experiment instructions, and asked them questions with help of sign language interpreters. We conducted the experiments at five different places in Japan, on five different dates in August and September 2014.

Figure 4 shows the experiment setting. The CMs were shown on a 32 inch high definition TV monitor (1920x1080), and there were an operator of the equipment (left in the figure) and a sign language interpreter (right in the figure).



Figure 4 Experiment setting

The eye tracking equipment was Tobii X120 Tobii Technology AB, Firmware: 2.0.7 Studio 3.2.3.336 (Enterprise edition), with sampling rate of 60 Hz. The Tobii system uses the Velocity-Threshold Identification (I-VT) fixation classification algorithm (Slavucci & Goldberg 2000) to filter and determine fixations among eye movements.

After the eye-tracking evaluation, we asked the following three questions for two CMs, one is CM1 (a detergent CM) and the other is CM3 (a cosmetics CM).

Question 1: whether you finishes reading all the captions or not

Question 2: which one is the easiest to read

Question 3: which one is the easiest to view both captions and videos (non-caption portion)

The participants were asked to choose one or more for question 1, and the only one answer for question 2 and 3.

## Results

We defined Areas of Interest (AOI) as the caption lines, where the closed captions are shown on the screen, and recorded fixation counts on AOI and visit counts to AOI. Fixation count is the total number of fixations on AOI (caption area) and visit count is the total number of visits to AOI. Thus if a participant visited AOI, and the gaze moved three times within AOI, its fixation count is three and visit count is one.

## Overall Results of Fixation and Visit Counts

The following two tables show the basic statistics of the results for fixation and visit counts. “Mean” signifies average counts per caption, and “SD” stands for standard deviation. Following the tables, graphs of both counts (mean values) are shown.

CM No	Type	Mean	SD	CM No	Type	Mean	SD
CM1	type1	3.71	2.15	CM4	type1	3.72	2.01
	type2	3.77	2.38		type2	2.86	1.50
	type3	3.01	1.73		type3	3.69	1.93
CM2	type1	4.42	3.00	CM5	type1	3.27	2.08
	type2	4.40	2.75		type2	3.48	2.35
	type3	4.40	2.31		type3	3.41	2.03
CM3	type1	4.21	2.66	CM6	type1	4.31	2.70
	type2	4.65	2.89		type2	4.15	2.40
	type3	4.18	2.10		type3	4.01	2.34

Table 5 Basic statistics of fixation counts

CM No	Type	Mean	SD	CM No	Type	Mean	SD
CM1	type1	1.45	0.57	CM4	type1	1.42	0.58
	type2	1.56	0.63		type2	1.24	0.43
	type3	1.29	0.53		type3	1.46	0.60
CM2	type1	1.42	0.62	CM5	type1	1.43	0.60
	type2	1.39	0.67		type2	1.38	0.62
	type3	1.59	0.69		type3	1.37	0.58
CM3	type1	1.62	0.78	CM6	type1	1.74	0.77
	type2	1.71	0.80		type2	1.77	0.70
	type3	1.58	0.66		type3	1.55	0.88

Table 6 Basic statistics of visit counts

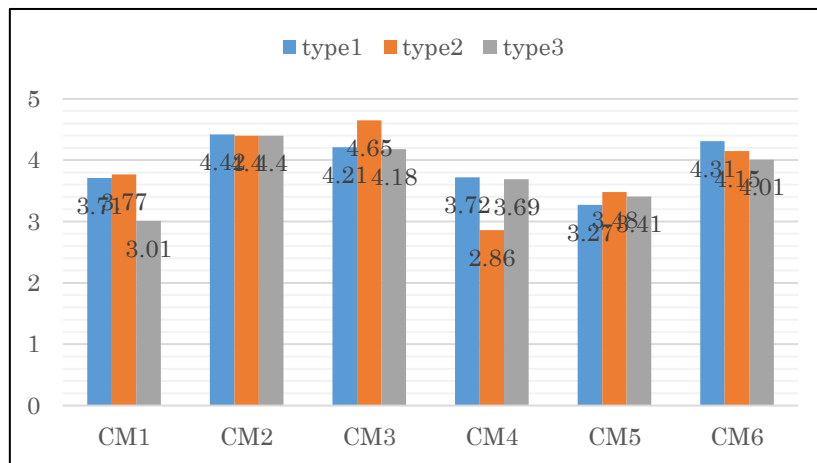


Figure 5 Fixation counts for 6 CMs

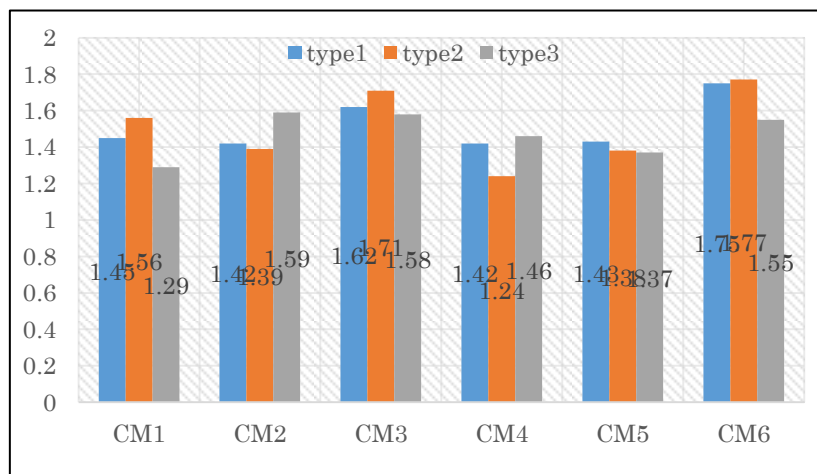


Figure 6 Visit counts for 6 CMs

### Statistical Significance among the Fixation and Visit Counts

We looked into if there are some statistically significant differences among the fixation counts and among visit counts for all the CMs. We used analyses of variance (ANOVA) and we found out no significant difference among the three sets of scores for type1 to type3, for both fixation and visit counts at both 0.05 and 0.01 level.

### Further Analyses of Fixation and Visit Count Data

We further analyzed the fixation and visit count data, we found the following:

- 1) One closed caption is read with 3 to 4 fixations
- 2) About half of the participants watch the caption only once.
- 3) Fixation counts per second are about between 1 and 2.
- 4) The number of characters (in the caption) which are read with one fixation is about 5 or 6, excluding special symbols such as period, exclamation mark, question mark, and parentheses.

## Questionnaire Evaluation Results

The next two figures show the results of questionnaire evaluation for three types of captions.

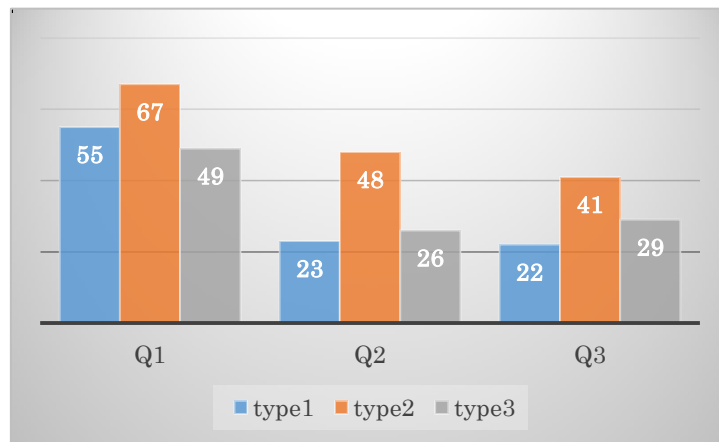


Figure 7 Result of questionnaire evaluation for CM1

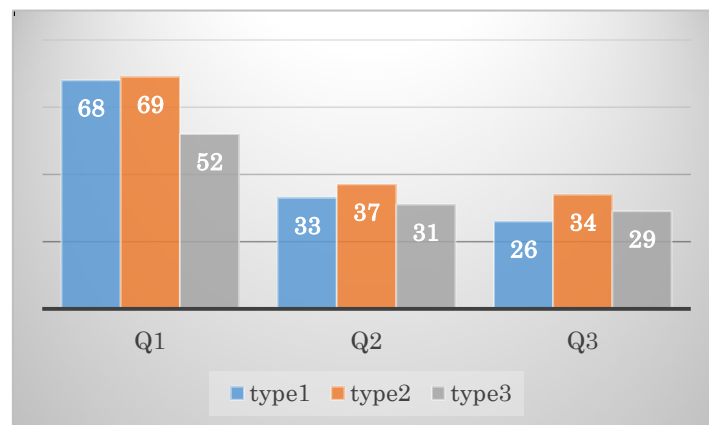


Figure 8 Result of questionnaire evaluation for CM3

We compared the scores for the three types of captions for CM1 and CM3 with ANOVA analyses, and found that there was at least one significant difference among the scores for CM1 for all the three questions at 0.05 level, however we found no significant difference for CM3 data. Please note that although the all participants (N=92) were asked to choose the best one for question 2 and 3, some of them could not choose the best, thus no answer, or a few of them chose two best ones, which led to different total numbers for question 2 and 3 in figure 7 and 8.

## Observations

As the results of the experiment and their statistical analyses show, we did not discover statistically significant differences among the fixation and visit counts. This means that the participants viewed the three differently captioned CMs similarly even though there are visible differences between horizontal and vertical captions. Further analyses suggest that since the number of characters which are read per fixation in general for Japanese captions is about 3.2, the result this time, 5 to 6 characters per fixation, is almost twice as many, and thus



the hearing impaired participants read more characters at a fixation.

The number of fixation per second is said to be 3 to 4 in general and our data showed the number is 1 to 2, thus we can say that about a half of the fixations are used for reading captions and the other half is used to read or view non-AOI area. The results of the questionnaire evaluation showed that type 2 caption was evaluated better than the other one or two types for all the questions and thus we can say that type 2 caption where captions are partially done had better readability and easiness to view. In order for us to find out relation between eye tracking data and characteristics of six CMs shown in table 2, we computed correlational values between them. We computed correlational scores between fixation counts and the number of scenes and number of face appearances for the six CMs, and we did the same for visit counts. The correlational result showed there were no statistically meaningful relations among them. Canadian researchers reported that hearing impaired people tended to pay more attention to human faces than hearing people (Chapdelain, Gouiller, Beaulieu and Gagnon 2007). Thus we expected appearance of human faces may influence the fixation or visit counts, but we did not find such influence this time.

### **Conclusions**

We have conducted experiments in order to find the effects of various types of caption for real TV commercial. The results show that there was no clear difference among the fixation and visit count data, however they also suggest that the hearing impaired participants read more characters with one fixation and about half of the fixations are spent for viewing non-AOI, i.e. other areas than the closed captions. The questionnaire result clearly shows that the participants prefer one type of caption, partially captioned CM to the other types. We would like to continue to research on CM captions in order to find what captions make good caption for TV commercial, and what type of captions help viewers to understand CMs better. Finally, we would like to thank Kao Corporation for providing us with their TV commercials as experiment materials. This research has been supported in part by Hosono Bunka Foundation grant (Hosono Bunka Foundation).

## References

- Chapdelaine C., Gouaillier V., Beaulieu M., Gagnon L. (2007). Improving Video Captioning for Deaf and Hearing-impaired People Based on Eye Movement and Attention Overload, *Human Vision & Electronic Imaging (SPIE #6492)*, San Jose.
- Fukushima T. & Inoue S. (2012). Research on The Usefulness of Closed Captions on TV Commercials (4) - Analyses of Closed-Captioned TV Commercials in terms of Caption Speed and Caption Readability -, in *Proceedings of the 4th International Conference for Universal Design in Fukuoka 2012*.
- Hoso Bunka Foundation. <http://www.hbf.or.jp/eng/>.
- Inoue S. (2012). Research on The Usefulness of Closed Captions on TV Commercials (1) An Opinion Survey of the Hearing-Impaired and Challenges It Reveals, in *Proceedings of the 4th International Conference for Universal Design in Fukuoka 2012*.
- MIC statistics. (2014). *Jimakuhousou-tou no jisseki*, Statistics of closed captioned TV broadcasting.  
[http://www.soumu.go.jp/menu\\_news/s-news/01ryutsu09\\_02000106.html](http://www.soumu.go.jp/menu_news/s-news/01ryutsu09_02000106.html) (in Japanese).
- Slavucci, D. D. & Goldberg, J. H. (2000). Identifying fixations and saccades in eye-tracking protocols, in *Proceedings of the symposium on Eye tracking research & applications, ETRA 2000*, Palm Beach Gardens, Florida, United States, pp. 71-78.