

Time Spent Viewing Captions on Television Programs

**Carl J. Jensema
Ramalinga Sarma Danturthi
Robert Burch**

Institute for Disabilities Research and Training, Inc.
2424 University Boulevard West
Silver Spring, MD 20902
301-942-4326

June, 1999

Abstract

The eye movements of 23 deaf subjects, ages 14 to 61, were recorded 30 times per second while the subjects watched four 2.5-minute captioned television programs. The eye movement data were analyzed to determine the percentage of time the subject actually looked at the captions on the screen. It was found that subjects gazed at the captions 84% of the time, at the video picture 14% of the time, and off the video 2% of the time.

Age, sex, and educational level achieved appeared to have little influence on time spent viewing captions. When caption speed increased from the slowest speed (100 wpm) to the fastest speed (180 wpm), mean percent of time gazing at captions increased only from 82% to 86%. A distinctive characteristic of the data was the considerable variation from subject to subject and also within subjects (from video to video) on the percent of time spent gazing at captions.

Introduction

The technology for captioning television programs for the benefit of people who are deaf or hard of hearing has been around since the early 1970s, and the first national closed captioned television broadcast was done on March 16, 1980. Since 1993, closed captioned television decoders have been built into every television set with a screen larger than 13 inches. More and more captioning is being done, and in accordance with federal law, within a few years, most programs will have closed captions.

Captioning is similar to subtitling and is the process of converting program audio to text and displaying that text on the screen. This is a much more complicated process than it may seem and requires many decisions concerning timing and screen placement. Captioning television programming is a whole new profession, one that has grown very rapidly in recent years.

The eye movement of caption viewers is an important issue within the captioning industry because it influences many decisions on how captioning is done. Unfortunately, almost no research has been conducted to examine how people move their eyes when viewing captioned television. The only recent research publication seems to be an article by Jensema, El Sharkawy, Danturthi, Burch, and Hsu (1999) that gave a general overview of eye movement among captioned television viewers.

The study presented here takes this a step further and examines the percentage of viewing time actually spent focusing on captions. The general research questions to be addressed are:

1. What percent of viewing time do deaf people actually spend looking at captions?
2. Does the percentage vary with age, sex, and education?
3. Does the percentage vary from program to program and subject to subject?
4. Does caption speed influence the percent of time spent viewing captions?

Procedure

This research study utilized a modified version of the Eyegaze Development System designed by LC Technologies, Inc. This system, as modified by the Institute for Disabilities Research and Training, Inc., (IDRT) was described in detail by Jensema, et al. 1999. Basically, the subject sits in front of a computer screen and watches a television program shown on the screen. A camera with an infrared light is mounted beneath the screen and the light is shined into the subject's eye. The camera photographs the reflection of the light off the subject's eye and uses this information to calculate where the subject's eye is looking on the screen. These photographs and the calculations made from them are done 30 times a second. The coordinates on the screen where the subject is looking are recorded on a computer file for later analysis.

Each subject is seated so that their eyes are between 18 and 22 inches from the computer screen. The video picture on the computer screen is 6 inches by 4.5 inches, a size that, when seen from 22 inches away, is roughly similar to watching a 27" television set from 6 or 7 feet away.

The stimulus materials for this study were four short captioned videos on the topics of Pyramids, Salmon, Sailing, and Space. Each video was exactly 150 seconds long (2.5 minutes) and was divided into five 30-second parts. Each part was captioned at a different speed. Jensema, McCann, and Ramsey (1996) had examined a sample of 205 captioned television programs and found that the captioning rates had a mean of 141 wpm and a standard deviation of 21 wpm. Using this as a guide, the speeds used for the four videos in the present study were 100 wpm, 120 wpm, 140 wpm, 160 wpm, and 180 wpm. The videos were used in a previous study (Jensema, 1998) and were custom-made for the research applications. There was no audio and the caption wording was carefully created to be consistent in reading difficulty across the different video segments.

Each of the four videos was 2.5 minutes long, giving a total of 10 minutes of stimulus material. Since eye position was measured 30 times a second, each subject's eye position was measured a total of 18,000 times (10 x 60 x 30).

The four videos were shown to 25 deaf people randomly selected from the Washington, D.C. metropolitan area. The data from two of these subjects was discarded because the eye movement data indicated they were clearly not paying attention to the videos. The remaining 23 subjects consisted of 11 females and 12 males. They varied in age from 14 to 61 years, with a mean of 29 years and a standard deviation of 13.7 years

The data collection procedure was very straightforward. All testing was done in the IDRT research lab in Silver Spring, Maryland. Each subject was given a brief introduction to the project, the equipment was explained, and any questions the subject had were answered. After filling out a short demographic questionnaire, the subject was seated in front of the computer screen. The equipment was calibrated to the subject's eye, and the presentation of stimulus material was begun. Each 2.5-minute video was presented without a break. There was a brief pause between videos. After the four videos were shown, the subject was questioned and any feedback comments were noted. The entire procedure from start to finish took less than an hour. Each subject was given an honorarium of \$25.

All eye movement data for each subject was recorded into a computer file in three columns. The first column represented the sequence number among the 18,000 observations for each subject. The second and third columns represented the X and Y coordinates for the eye gaze location on the computer screen. Since it was known exactly how long each video segment was, it was possible to divide the data according to these segments.

Each of the 18,000 data points for each subject was categorized by location as being a) on captions, b) on picture, or c) off screen. All captions were known to appear

in a specific area near the bottom of the screen. Figure 1 graphically illustrates the coordinates used for categorization, and Figure 2 shows a plot of the gaze points of a typical subject for a 2.5-minute video. This particular subject watched the captions 85% of the time, the video picture 12% of the time, and was gazing off the video 3% of the time. Note that the "caption area" is arbitrary. In this study the "caption area" was defined as the actual space occupied by two lines of captions plus one line above and one line below these two caption lines. If some other definition of the caption area was used, the statistics in this study would obviously change somewhat.

Results and Discussion

Most of the analysis in this paper focuses on the percent of time subjects watched the caption area of the screen. Over all subjects and all video material, subjects gazed at the captions 84% of the time, at the video picture 14% of the time, and off the video 2% of the time. The off-video gaze points are thought to be mostly due to eye blinks and normal saccadic eye movement.

The fact that viewers in this study appear to spend about 84% of their television viewing time looking at a program's captions suggests that viewing captioned television is primarily a reading task. This has extremely important implications for the development of reading skills. For example, it is generally assumed that people watch about 30 hours of television programming per week. If those programs are captioned, a typical viewer may spend about 25 hours a week looking at printed material. That much exposure to print is bound to influence reading skills.

There were 12 males and 11 females in the sample. No significant difference in mean percent of time spent looking at captions was found between males and females. ($t = -.361$, $df = 21$, $p = .72$). Sex does not appear to be a significant factor in percent of time spent viewing captions.

The age range of the sample was 14 years to 61 years. The correlation between age and percent of time spent viewing captions was .253. Age was not an important factor in percent of time spent viewing captions in this study. It had been thought that age might be a factor and that older people, with their generally slower reflexes and less acute eyesight, might need more time to read captions. The .253 correlation indicates that age and caption viewing time were not closely linked in this study. However, the mean age in this study was 29 years and the oldest subject was 61 years. A group of older people, such as those aged 50 to 80 years, might have different results.

Subjects were asked the highest educational level they had completed. The list below gives the mean percentage of time spent viewing captions by the educational levels reported. Percentage of time spent viewing captions is significantly different for different education groups at the .03 level of confidence ($F = 3.422$, $df = 4, 18$, $p = .03$), but it is not clear from the data in this study exactly what this means in practical terms. Further research is needed to investigate the relationship between education and percent of time spent viewing captions.

	<u>N</u>	<u>Mean</u>
Junior High School	2	81
High School	10	86
Some College	3	86
College Graduate	7	79
Graduate School	<u>1</u>	<u>92</u>
Total	23	84

Table 1 gives statistics for each topic and for each words-per-minute category over all subjects. The various topics and categories have means between 82% and 86%, very close to the 84% mean for all video clips in this project. Table 1 has two other particularly interesting results. First, it shows that the overall percent of time an audience spends looking at the captions seems to remain fairly constant across different programs. Second, the overall mean percent of time spent watching captions does not increase greatly as caption speed increases. This second finding is particularly interesting. It had been expected that, as the number of words shown on the screen increased, the percent of time spent looking at them would also increase. There is an increase, but it is not great. As caption speed increased from 100 wpm to 180 wpm, the percentage of time spent looking at captions only increased from 82% to 86%.

Probably the most distinctive characteristic of the data is the wide variation in time that subjects watch captions on different videos. For example, Table 2 gives statistics on the percent of time each subject watched captions during each of the 20 video clips in this project. The extreme is Subject 13, who watched captions 44% of the time on one video clip and 99% of the time on another video clip, a range of 55 points. Obviously, many factors come into play here, including how interested the subject is in the topic, how interesting the video is, how difficult the captioned material is to read, and how much prior knowledge the subject has of the topic.

The eye movement processes which deaf and hard of hearing people use to view captioned television are very complex and, as yet, little understood. This study has been a small step toward understanding these processes, but much remains to be explored.

Reference

Jensema, C.J., El Sharkawy, S., Danturthi, R. S., Burch, R., and Hsu, D. Eye Movement Patterns of Captioned Television Viewers. Paper submitted to the American Annals of the Deaf, 1999.

Figure 1
Computer Screen Layout

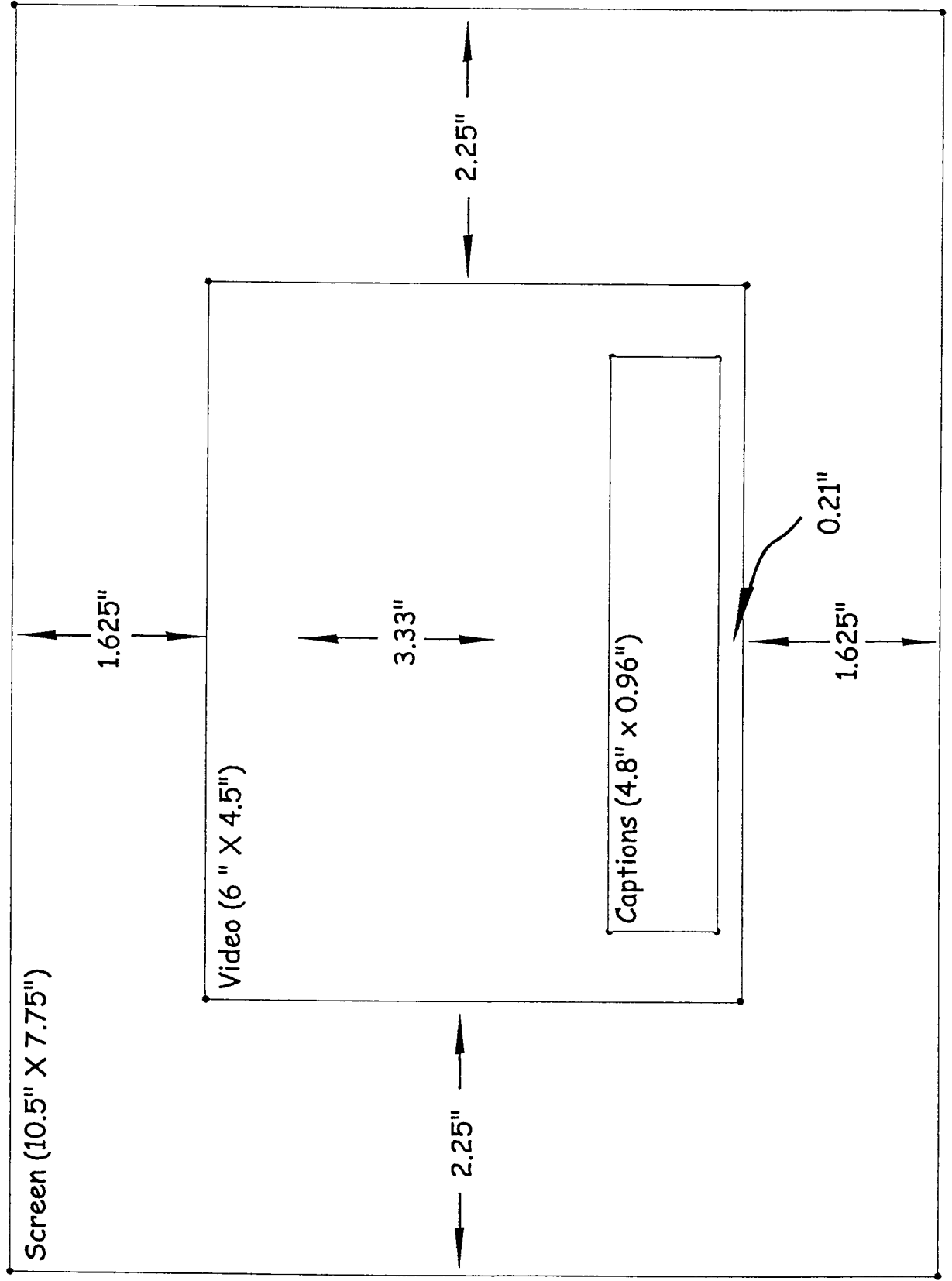


Figure 2

Plot of gaze points for a typical subject (4500 points over 2.5 Mins)

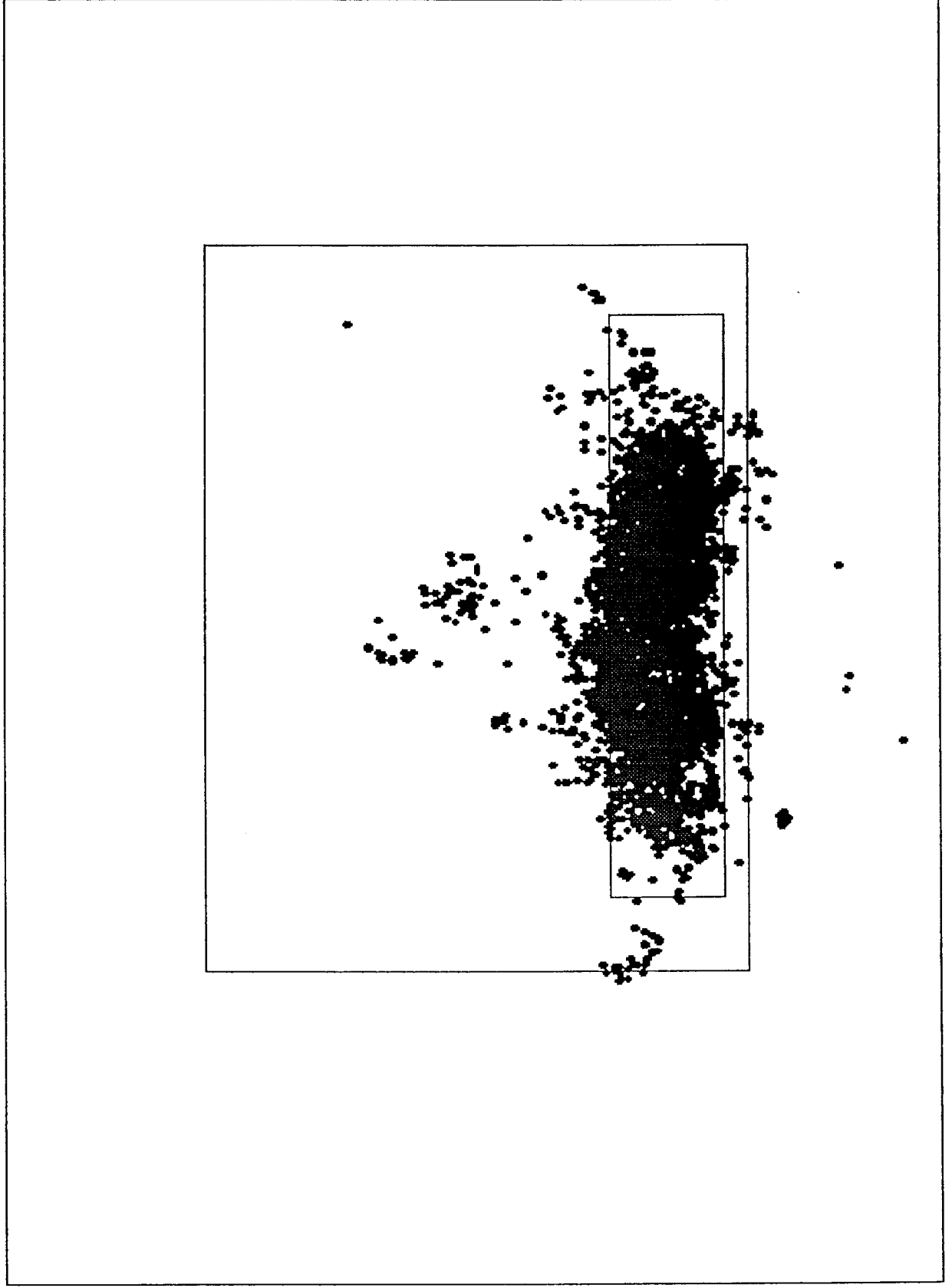


Table 1
Percentage of Viewing Time Spent Looking at Captions
By Topic and Speed
 (Based on 23 subjects.)

	Topic				Words Per Minute					All Video Clips
	Pyramids	Salmon	Sailing	Space	100	120	140	160	180	
Mean	83	86	82	84	82	82	83	85	86	84
S.D.	11	10	11	11	11	10	12	11	9	11
Min	54	54	44	51	52	62	44	44	54	44
Max	100	100	100	100	99	100	99	100	100	100
Range	46	46	56	49	47	38	55	56	46	56

Table 2
Percentage of Viewing Time Spent Looking At Captions
By Subject

(Based on 20 video clips for each subject.)

Subject No.	Mean	S.D.	Min	Max	Range
18	72	9	52	84	32
3	73	9	51	89	38
13	76	18	44	99	55
14	77	10	62	92	30
8	79	13	56	97	41
11	80	12	63	99	36
9	81	7	67	90	23
1	82	9	64	95	31
5	82	6	68	92	24
7	82	7	69	94	24
2	84	10	65	96	31
4	85	7	70	95	24
21	85	10	64	96	32
23	85	11	62	99	37
10	86	8	64	95	31
15	86	9	54	95	42
17	87	8	72	100	28
6	88	9	62	99	37
22	88	11	61	98	37
16	90	6	81	97	16
12	91	8	71	99	28
19	92	6	82	100	18
20	93	5	82	99	17