

# Whether, When, and How to Utilize the Printed Script of a Film: A Good Substitute for L2 Subtitles?

紀要第 9 号抜刷

1995



財団法人 語学教育研究所

The Institute for Research in Language Teaching

# Whether, When, and How to Utilize the Printed Script of a Film: A Good Substitute for L2 Subtitles?

SHIZUKA Tetsuhito

## ABSTRACT

Despite the growing popularity of L2 captioned films in listening classes, they seem to have certain limitations as tools for proficiency development. This paper explored several ways of presenting the printed script of a film as a substitute for L2 subtitles. The subjects, forty Japanese EFL students (age 15-16), received three sets of different treatments: (1) watching a subtitled scene versus simultaneously listening to the soundtrack and following the script of the same scene, (2) simultaneously listening to the soundtrack and following the script versus reading the script of each sentence prior to listening to the soundtrack, and (3) when listening to the soundtrack after reading the script, following the script with the eyes open versus closed. The scores of the partial dictation tests and the subjects' verbal responses indicated the following:

- (1) A film, its script, and the soundtrack could serve as more than an excellent substitute for a film with closed captions.
- (2) First presenting the script and then the sound is more effective than simultaneous presentation.
- (3) When listening to the soundtrack, the script should be followed for better assimilation, and should be out of sight for better comprehension.

Based on these findings, a procedure of utilizing a film in a listening class was suggested.

## 1. INTRODUCTION

In accordance with the development of multi-media hardware, more and more EFL teachers seem to be utilizing ELT videos and/or commercial films in listening classes. In particular, films with closed captions seem to be drawing increasing attention from teachers and researchers. There is already a considerable body of research pertaining to the effects of multi-media presentations on learners' comprehension of the material and the development of their language proficiency. Although there exists an almost unanimous agreement that films, with or without captions, are powerful motivational tools (Edasawa et al. 1989, Takeuchi et al. 1990, Obari et al. 1992, etc), it seems premature to paint too rosy a picture about their effectiveness as language learning resources. This study explores the possibility of utilizing printed script instead of L2 subtitles for listening proficiency development.

## 2. REVIEW

Let us briefly review studies on the effects of adding pictures to audio and those concerning adding captions to audio. Numerous studies have been carried out to investigate the effects of pictures on students' comprehension of the subject matter and on their increase in listening proficiency over time. The results could be generalized as follows:

- (1) Adding pictures to audio generally increases learners' comprehension of the material content whether the language is their L1 or L2. It is especially effective when the picture and the audio carry redundant information (Nugent 1982, Levie and Lentz 1982, Reese 1984) or when the learners' proficiency is not high (Kamei 1994, Kamei & Hirose 1994).
- (2) It is rather indecisive or even doubtful, however, whether films are superior to audio-only materials in developing learners' listening proficiency (Edasawa et al. 1989, Takeuchi et al. 1990, Obari et al. 1992, Takai 1991, 1993).

One reason why visual presentation accompanying audio does not always contribute to learners' proficiency development may be that comprehending the situation through the visual channel does not guarantee paying attention to the words transmitted through the audio channel. Itakura (1986), who conducted an experiment on the effects of still-pictures on learners' listening comprehension, reported that adding pictures to audio indeed improved learners' comprehension of the material. However, she contended, it should not be interpreted that visuals enhanced the learner's listening comprehension ability, but rather, the improved comprehension was due to the amount of information received through visuals. Her findings are in line with those by Obari et al (1992), who compared video material with sound-only material on learners' comprehension of the story line. Although the learners who had viewed a video got higher scores in general comprehension tests than those who had only listened to the sound track of the same video, the two groups of subjects did not show any significant score differences in dictation tests or in fill-in-the-blank tests. Their findings substantiate Itakura's contention that visual materials do not affect the score when the learner tries to analyze the structures of verbal messages (Itakura 1986). Indirect support for this argument can be found in the research on the effects of pictures on learning to read in L1, which generally showed that the presence of pictures could be harmful when an instructional task such as the acquisition of sight vocabulary was the objective (Samuels 1967, Braun 1969, Harzem et al. 1976, Willows 1978, all cited in Levie and Lentz 1982).

With regard to the research which investigated the effects of closed captions on comprehension, the results were generally positive. Lambert et al (1981), Holobow et al (1984), and Danan (1992) examined various monolingual or bilingual conditions in which no dialogue, or dialogue in the first or second language was combined with no subtitles or subtitles in either language. All through their studies, the two most promising

combinations turned out to be “reversed subtitling” (L1 soundtrack plus L2 subtitles) and “bimodal L2 input” (L2 soundtrack plus L2 subtitling), which showed that students clearly benefited from processing the written representation of the dialogue. Hirose and Kamei (1993) explored the effects of English captions on learners’ comprehension in relation to learners’ proficiency level and to the type of information conveyed. The results suggested that any learner at any proficiency level could benefit from the captions and understood the film better and that the captions increased learner understanding of linguistic rather than emotional information. Kamei and Hirose (1994) again experimented on multi-media presentations, this time comparing four different conditions: English sound only, English sound with pictures, English captions with pictures, and English sound and captions with pictures. It was found that, in any proficiency group, the two conditions utilizing captions produced the highest comprehension scores.

Despite the promising results of these comprehension studies, results pertaining to proficiency gains are not too impressive, just as was with the case of motion picture studies. Many of them seem to cast doubt upon our expectation that exposing learners to subtitled films leads to their greater progress in listening proficiency than when films without subtitles are shown. Miyamoto (1990) compared L1 with L2 captions about their instructional impact on learner progress in proficiency over six months and reported no significant difference between the two in listening proficiency increase. Obari et al. (1993) conducted a longitudinal study on the effect of L2 captions by teaching one group with a video material with captions and the other group with the same video without captions. Out of the four sets of comparisons between two homogeneous groups, only one produced significant differences favoring the group which had used the captioned material. As for the other three pairs, the caption group gained only as much as the non-caption group.

Researchers have noted that one limitation of captioned films as listen-

ing comprehension material is that words on the screen are such a powerful attention catcher that learners cannot pay enough attention to the soundtrack (Miyamoto 1990, Obari et al. 1993, Kamei and Hirose 1994). In fact, among the four groups Kamei and Hirose experimented on, the picture+caption group did not differ significantly from the picture+caption+sound group in comprehension score, which means it didn't make any difference whether or not there was sound because the subjects were only reading the captions. This phenomenon could be accounted for by the more dominant processing of the visual modality than audition (d'Ydewalle 1991), and should not necessarily be attributed to the subjects' proficiency level.

There is no doubt that watching films is a lot of fun and that L2 captions are a blessing that makes it possible for language teachers to utilize authentic commercial films, which would be otherwise impossible or at least very difficult for learners to follow (Vanderplank 1990). However, if the purpose is to increase listening proficiency, the problem is that the enhancement of comprehension, which is brought about more by watching the motion picture or by reading captions than by listening to the audio is not likely to lead to improvement in listening ability. Obviously, just letting learners view captioned films will not achieve much. Then, what we must do if we intend to use captioned films for listening rather than reading or picture-watching training is to devise a way or ways of presenting, or not presenting, picture and/or captions in which learners can eventually learn to perceive and decode the sounds themselves.

Takahashi (1994) explored this area by comparing the effects of five ways of film presentation, four of which utilizing L2 captions, on subsequent dictation using the soundtrack of the same film. The five treatments were as follows. In T1, learners viewed the film without captions. In T2, learners viewed the film with captions on the screen. Both T1 and T2 took 2 minutes. In T3, learners read the caption for each sentence for

5 seconds before they viewed the corresponding part of the film. This treatment took 8 minutes. In T4, learners listened to the soundtrack two or three times to answer fill-in-the-blanks type exercises, before they were presented with answer keys for self-correction. This procedure took 20 minutes.

Computer-controlled T5 was rather complicated. First, learners viewed the caption of a sentence with two blanks in it. Then they viewed the corresponding part of the film and tried to fill in the blanks. Ten seconds later, the complete caption of the sentence was presented for learners' self correction. Finally the complete caption of the same sentence was presented again with the corresponding picture and audio, and learners tried to mouth the sentence without actually making the sound. This was done for all the sentences in the scene. When all the sentences were learned in this way, the whole procedure was repeated two more times using more difficult completion tasks. That is, in the second round, learners had to fill in not only two words but half of a sentence, and in the third round, they had to write down the whole sentence. All this was also conducted in 20 minutes.

Comparison of these five methods produced amazingly clear-cut results. The post treatment dictation scores in percentage were in the order of T1 (31.9) < T2 (54.9) < T3 (75.4) < T4 (86.1) < T5 (97.1), with a significant difference between every adjacent pair. Implications Takahashi mentioned included the following: (1) Learners learn little by simply viewing a captioned commercial film because the rate of speech is so fast that they cannot even read the captions. (2) Learners benefit from reading the captions of each sentence just before listening to the audio. (3) Learner gain is greatest in T5, in which they actively try to catch every word in a sentence.

### **3. THE STUDY**

The present study was primarily inspired by the results of Takahashi

(1994). Although the effect of T5 is truly impressive, it seems to involve much too complicated procedures to be implemented in a classroom which lacks CAI equipment. Since such classrooms are by no means rare even today, it would be justifiable to explore methods of presentation that are feasible in a conventional classroom with only a video cassette player and a caption decoder. Furthermore, since videotapes and scenarios of popular commercial films are widely available today, pursuing the possibility of utilizing printed text instead of captions would help teachers without closed caption decoders. Another potential difficulty with T5 for some teachers might be the length of time it takes in relation to the length of the film material. T5 took 20 minutes to familiarize learners with a 2-minute scene, which means it would need as long as 20 hours to deal with the whole of a typical 2-hour film. For motivational and practical reasons, it didn't seem worthless to explore alternative ways which would cover a longer part of a film in a shorter time, even if it meant a lower level of familiarization with each scene than in the case of T5. To summarize, the purpose of the present study was to explore alternative methods of presenting written representations of dialogues in commercial films feasible in a traditional teaching environment which lacks computerized equipment or even a video caption decoder. The research questions addressed were:

- (1) How does watching a subtitled scene compare with listening to the soundtrack while following the script of the same scene? Can the script be a substitute for L2 captions?
- (2) Is reading the script of each sentence prior to listening to the soundtrack more effective than simultaneously reading the caption and listening to the soundtrack?
- (3) When listening to the soundtrack after reading the script, should the script be followed with eyes or be out of sight?

As for the first question, no formal hypothesis was formed since, to the



knowledge of the present researcher, no research was carried out concerning the differences between the printed scripts and the captions on the screen. It was only informally predicted that the motion picture might distract learners. The second question was raised with an expected answer in mind based on Takahashi's study; it was hypothesized that the former style would be more effective than the latter since, in the case of a commercial film, the rate of speech would be too fast to process captions and audio at the same time. Research Question 3 was an extension of the preceding question. When listening to the audio after looking at the caption/script for some time to grasp the structure, should learners be presented only with the audio without captions or both audio and captions? Although no formal hypothesis was formed, our hunch was that captions would prevent learners from concentrating on the audio, thus resulting in lower scores than when only the sound was presented.

In this study, "effectiveness" was operationally defined as "the extent to which learners could perceive language form by listening to the soundtrack of the same film as was used in the treatment."

## **4. METHOD**

### **4.1 Subjects**

Forty first-year students (age 15-16) learning English at the Department of Communication Information Science, Fukushima National College of Technology participated in this study. At the beginning of the academic year 1994, the 18 male and the 22 female students in the department had been matched for sex and English proficiency, which was assessed by a part of a Pre-PET Test by Cambridge Examination Development Unit. Then each member of the matched pairs was randomly assigned to either of the two groups of twenty, each consisting of 9 males and 11 females. In these two groups, they were attending "English Conversation" class separately. These quasi-homogenous intact groups served as two groups of subjects to receive different treatments in this

study, which was conducted in the fall of the same academic year.

## **4.2 Materials**

### **4.2.1 Film**

Three consecutive scenes (6-8 minutes each) from *Punky Finds a Home* (CC-Study, Gakken) were used. This 50-minute-long ELT film, which is originally an episode in a serial produced by NBC, is subtitled verbatim specifically for language learning purposes, and comes with a script in print and the soundtrack in an audio tape.

### **4.2.2 Pre/Post/Retention tests**

In order to measure the extent to which subjects perceive the words they hear in the video, a partial dictation test was developed for each scene (see Appendix for an example). Subjects were to listen to an audio tape recording ten utterances taken from each scene and fill in the blanks. The first and the last word of each utterance were already provided. The mean length of the utterances was 7.6 words. Only thirty blanks per test, designated by underlines, were scored. When deciding which words to underline, care was taken so that the test results would be affected not by subjects' knowledge of vocabulary but by their listening perception skills. Words were chosen which were believed to be familiar to all the subjects but which were pronounced quite naturally. The words which came immediately after or immediately before the printed words were not underlined in order to prevent the subjects from trying to catch the sound without thinking about the meaning of the whole utterance. One point was given to one underlined blank correctly filled. If a correct word was provided in a blank next to the appropriate one, 0.5 point was given. Spelling mistakes were disregarded. The identical test was used as the pretest, the posttest, and the retention test for each scene.

### 4.2.3 Reflection Survey

This was a questionnaire to be conducted after the experiments. Its purpose was to look into the subjects' perception of different treatments. The questions (originally in L1) were:

- (1) Comparing viewing a captioned film with listening to the soundtrack of the film with the script at hand, which do you think is more effective for listening proficiency development and why?
- (2) If listening to the audio after looking at the script for some time produces better results in posttests than viewing a captioned film, what do you think is the reason?
- (3) Comparing, after looking at the script for some time, listening to the audio with your eyes closed and listening to the soundtrack with your eyes still fixed on the script, which do you think is more effective for developing your listening proficiency and why?

## 4.3 Procedure

### 4.3.1 Outline

Three comparisons were made in three consecutive class sessions in October 1994. Comparison 1 was carried out with Scene 1 on October 18th, Comparison 2 with Scene 2 on the 19th, and Comparison 3 with Scene 3 on the 25th. The reflection questionnaire was conducted on November 2nd. The retention test was given without notice on November 8th (see Table 1).

Table 1 : Procedure Outline

| Oct. 18th                 | Oct. 19th                 | Oct. 25th                 | Nov. 2th             | Nov. 8th          |
|---------------------------|---------------------------|---------------------------|----------------------|-------------------|
| Comparison 1<br>(Scene 1) | Comparison 2<br>(Scene 2) | Comparison 3<br>(Scene 3) | Reflection<br>Survey | Retention<br>Test |
| Pretest                   | Pretest                   | Pretest                   |                      |                   |
| Treatments                | Treatments                | Treatments                |                      |                   |
| Posttest                  | Posttest                  | Posttest                  |                      |                   |

### **4.3.2 Experiment Procedure in Detail**

On the day before each experiment, a Japanese translation of the dialogue and the glossary for the scene to be viewed in the next class session were distributed to the students. They were instructed to study the material in preparation for the actual viewing of the video. The purpose of giving the translation and the glossary was twofold. First, it was to facilitate the subjects' comprehension of the authentic material, which would discourage them by its difficulty if it were presented without any aid. Also, they were expected to level off, to a certain extent, the subjects' comprehension of the scene, thus letting us concentrate more on measuring their aural grasp of the language form.

At the beginning of each experiment, the pretest was given in 5-10 minutes. The subjects listened to each utterance twice to fill in the blanks. After the pretest was collected, the subjects watched the scene without the captions once. This viewing was intended to promote the subjects' general comprehension of the scene. After the primary viewing, the subjects received different treatments, which will be described in the next section, according to the group they belonged to. After the treatment, the subjects in either group watched the same scene without subtitles once again. Finally, the posttest was administered in the same manner as the pretest. All this was done within the 50 minutes of one class period. Please note that each experiment was meant as much to be a usual class session as an experiment for data collection.

### **4.3.3 Treatment**

The two groups of subjects received different treatments from each other in each class session. Note that the three comparisons were carried out using three different materials (scenes) and, therefore, the results were comparable only between the treatments in the same session, not between the treatments across sessions. Comparison 1 was designed in an attempt to answer Research Question 1, Comparison 2 to answer

Research Question 2, and Comparison 3 to answer Research Question 3.

#### **4.3.3.1 Comparison 1** (Using Scene 1)

Group A : Watch the scene with captions twice.

Group B : Listen to the soundtrack of the scene twice, following the printed script at the same time.

Thus, one difference was the presence/absence of the motion picture. The other was reading the captions on the screen versus reading the printed scripts on paper.

#### **4.3.3.2 Comparison 2** (Using Scene 2)

Group A : Read one sentence on paper silently for 5 seconds, and then listen to the soundtrack for the sentence. Repeat the procedure for all the sentences in the scene.

Group B : Watch the scene with captions twice.

The reason we let Group B watch twice was to keep the time length balanced between the two treatments. When we consider application in classrooms, it would be more practical to compare two treatments carried out in similar time lengths.

#### **4.3.3.3 Comparison 3** (Using Scene 3)

Group A : Read one sentence on paper silently for 5 seconds, and then listen to the corresponding part of the soundtrack while still looking at the print. Repeat this twice for every sentence.

Group B : Read one sentence on paper silently for 5 seconds, and then listen to the corresponding part of the soundtrack with eyes closed. Repeat this twice for every sentence.

Thus, the difference is whether or not to keep looking at the script while listening to the audio.

The three pairs of treatment are tabulated in Table 2.

Table 2 : The Pairs of Treatment

|              | Group A |         | Group B |         |
|--------------|---------|---------|---------|---------|
| Comparison 1 | picture | picture | $\phi$  | $\phi$  |
|              | audio   | → audio | audio   | → audio |
|              | caption | caption | script  | script  |
| Comparison 2 | $\phi$  | $\phi$  | picture | picture |
|              | $\phi$  | → audio | audio   | → audio |
|              | script  | $\phi$  | caption | caption |
| Comparison 3 | $\phi$  | $\phi$  | $\phi$  | $\phi$  |
|              | $\phi$  | → audio | $\phi$  | → audio |
|              | script  | script  | script  | $\phi$  |

Note: Media clustered vertically indicate simultaneous presentation.  
 $\phi$  indicates the lack of the medium.

## 5. RESULTS

### 5.1 Comparison 1

#### 5.1.1 Test Scores

The descriptive statistics and the results of paired t-tests for Comparison 1 are shown in Table 3. The change of means is illustrated in Figure 1. The pretest confirmed that the two groups were homogeneous at the outset of the experiment. Although both treatments significantly increased scores by 20.6% (Group A) and by 22.6% (Group B), neither posttest scores nor gains revealed any significant differences between groups. Retention scores and retained gains (R-gain: retention score minus pretest score) did not show any significant differences between the two groups, either. This means both viewing the captioned film and listening to the soundtrack while following the printed script produced significant gains, which were not significantly different from each other.

#### 5.1.2 Survey Results

To Question (1), which asked to compare caption+audio+picture

Figure 1 : Means for Comparison 1

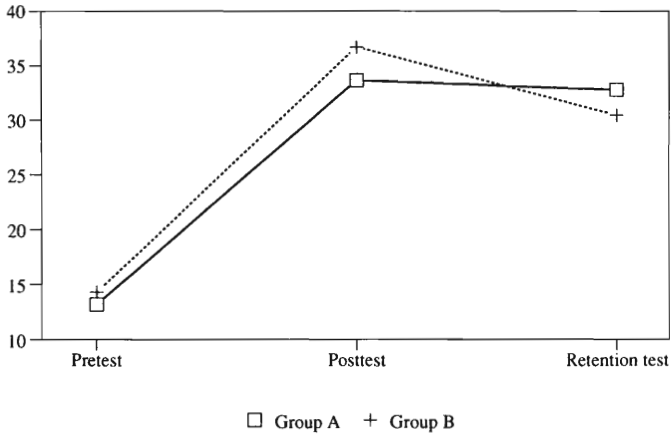


Table 3 : Means, SDs, and Results of Paired T-tests for Comparison 1

|         |      | Pre   | Post   | Gain   | t-value | Ret  | R-gain |
|---------|------|-------|--------|--------|---------|------|--------|
| Group A | Mean | 13.2  | 33.7   | 20.6   | 6.09**  | 32.6 | 19.4   |
|         | SD   | (6.9) | (14.7) | (14.7) |         |      |        |
| Group B | Mean | 14.1  | 36.7   | 22.6   | 11.43** | 30.6 | 16.5   |
|         | SD   | (6.7) | (11.4) | (8.6)  |         |      |        |
| t-value |      | 0.42  | 0.71   | 0.53   |         | 0.49 | 0.82   |

\*\*p&lt;0.01

with caption (script) + audio, 12 out of 40 subjects answered that they preferred the three-media presentation while the other 28 voted for the simpler caption(script) + audio mode. Reasons cited for supporting their choice are shown in Table 4.

Table 4 : Responses to Survey Question 1

| Reasoning   | # of subjects |
|---|---------------|
| ⟨For caption+audio+picture/Against script+audio⟩                  |               |
| The picture makes it easier to grasp the situation.               | 8             |
| The picture shows the movements of articulatory organs.           | 3             |
| Having fun facilitates progress.                                  | 1             |
| ⟨For script+audio/Against caption+audio+picture⟩                  |               |
| The picture makes it difficult to pay due attention to the audio. | 16            |
| The picture makes it possible to guess without listening.         | 2             |
| The captions are harder to read than scripts.                     | 2             |
| The captions are not exactly synchronized with the sound.         | 1             |
| The picture is superfluous for listening practice.                | 1             |

Table 5 : Means, SDs, and Results of Paired T-tests for Comparison 2

|         |      | Pre   | Post   | Gain   | t-value | Ret    | R-gain |
|---------|------|-------|--------|--------|---------|--------|--------|
| Group A | Mean | 17.2  | 34.6   | 17.4   | 7.76**  | 27.9   | 10.7   |
|         | SD   | (7.4) | (13.3) | (10.0) |         | (11.5) | (7.8)  |
| Group B | Mean | 17.8  | 27.9   | 10.1   | 4.33**  | 31.2   | 13.4   |
|         | SD   | (5.3) | (9.5)  | (9.9)  |         | (7.8)  | (8.0)  |
| t-value |      | 0.29  | 1.85   | 2.32*  |         | 1.06   | 1.07   |

\*p &lt; .05 \*\*p &lt; .01

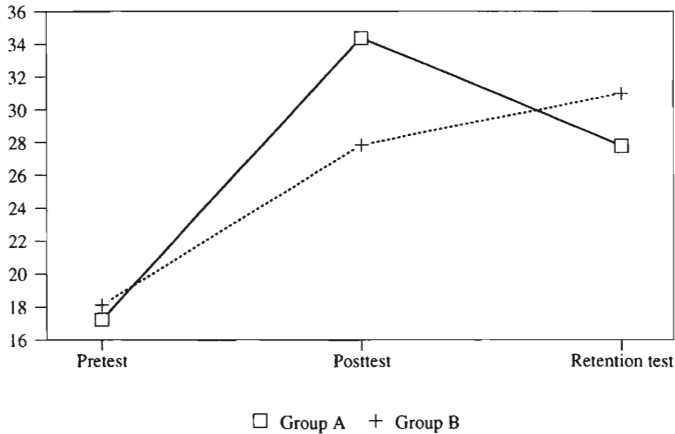
## 5.2 Comparison 2

### 5.2.1 Test Scores

The results of Comparison 2 are tabulated in Table 5 and illustrated in Figure 2. The pretest scores showed that there was no significant difference between groups before the treatments. However, the difference between the posttest scores approached a significance and the comparison of gains in fact revealed significant difference at  $p < .05$  level. Comparisons of pretest and posttest scores revealed a significant difference at  $p < .01$  level in either group. No significant difference was



Figure 2 : Means for Comparison 2



observed in retention scores or in retained gains. Thus, it was shown that looking at the script for some time and then listening to audio increased scores significantly more than watching captioned film, though the difference was not carried over until the retention test.

### 5.2.2 Survey Results

The answers to the second survey question, “If listening to the audio after looking at the script for some time produces better results in posttests than viewing a captioned film, what do you think is the reason?” are classified in Table 6.

## 5.3 Comparison 3

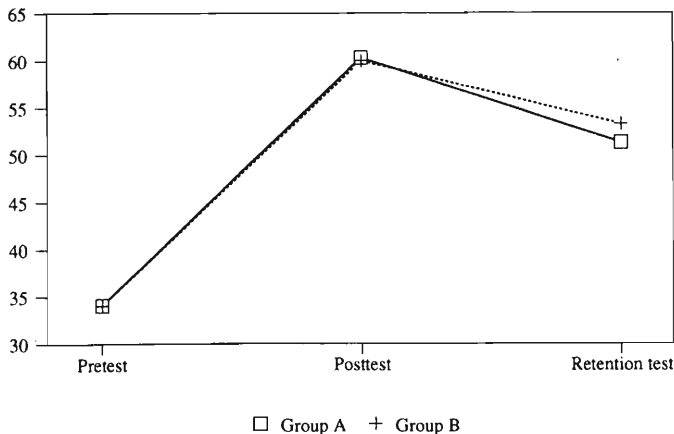
### 5.3.1 Test Scores

Table 7 shows and Figure 3 illustrates the test results for Comparison 3. Again the homogeneity of the two groups was confirmed before treatment. Just as in Comparison 1, though each group improved scores significantly after the treatment, no significant difference between

Table 6 : Responses to Survey Question 2

| Reasoning  | # of subjects |
|--|---------------|
| Looking at the script for some time enables/encourages me to memorize the line.  | 18            |
| The captions appear and disappear too quickly for the eyes or brains to catch up.  | 10            |
| Knowing in advance what will come helps me perceive words.   | 8             |
| Listening to the audio with the image of the script enables me to clearly perceive how certain words/phrases are pronounced. | 3             |
| When watching a subtitled video, all I can do is barely follow the captions. I can hardly pay attention to the audio.        | 3             |
| One thing at one time allows me to concentrate more.   | 3             |
| Following the captions lets me feel as if I were catching the sounds when actually I am not.                                 | 1             |

Figure 3 : Means for Comparison 3



groups was observed in posttest scores, gains, retention scores, or in retained gains.

Table 7 : Means, SDs, and Results of Paired T-tests for Comparison 3

|         |      | Pre   | Post   | Gain   | t-value | Ret    | R-gain |
|---------|------|-------|--------|--------|---------|--------|--------|
| Group A | Mean | 34.4  | 60.2   | 25.7   | 8.75**  | 51.8   | 17.4   |
|         | SD   | (7.3) | (13.3) | (13.1) |         | (12.5) | (13.7) |
| Group B | Mean | 34.1  | 60.1   | 26.0   | 10.19** | 53.7   | 19.7   |
|         | SD   | (9.6) | (14.5) | (11.1) |         | (12.8) | (11.9) |
| t-value |      | 0.14  | 0.02   | 0.07   |         | 0.48   | 0.57   |

\*\*p < .01

Table 8 : Responses to Survey Question 3

| Reasoning   | # of subjects |
|---|---------------|
| 〈For script+sound/Against sound-only〉   |               |
| The script allows me to better assimilate visual and acoustic images of the sentence.     | 8             |
| The script is necessary when the sentence is too long to retain in the short-term memory. | 4             |
| The script helps with comprehension even when the sound is not clear.                     | 2             |
| The script enables me to follow the audio.  | 2             |
| The script makes me feel secure.  | 2             |
| 〈For sound-only/Against script+sound〉   |               |
| The script prevents me from concentrating on the sound.                                   | 14            |
| The script makes me feel that I am catching the sound when I am not.                      | 3             |
| Looking at the script makes me pay too much attention to individual words.                | 1             |

### 5.3.2 Survey Results

To the third survey question concerning preference between listening to sounds with eyes following the script and without looking at anything, 25 out of 40 subjects answered that just listening to the sound is more effective and the other 15 answered to the contrary. Table 8 shows the elicited reasons.

## 6. DISCUSSION

### 6.1 Research Question 1

Research Question 1 was, “How does listening to the soundtrack while following the script compare with watching a subtitled film?” The test results of Comparison 1 showed no significant difference between the two, while the survey results favored script+audio rather than caption+audio+picture. It would be fair to say that script+audio was shown to be at least as effective as captioned films as far as assimilation of visual and audio images of utterances was concerned. This is good news to teachers without caption decoders since the possibility is that teaching with a script and non-captioned video could be more than a mere substitute for utilizing subtitled version.

To analyze students' preference for audio+script over captioned film, let us look more closely into perceived advantages and disadvantages of both modes. Looking at Table 4, it can be seen that the picture is perceived to do good by some students and to do harm by others. Most of the favorable opinions for the picture concern its role as a facilitator of overall comprehension, while the majority of the unfavorable views point out that the picture distracts their attention from the audio. Therefore, there are not really two opposing views about the functions of the motion picture but there is one contention made up of two parts. That is, the motion picture helps learners with overall comprehension but distracts them when they try to pay attention to the audio. This is a dilemma of which there is an easy way out. First watch the motion picture for schematic comprehension of the outline, and then, just listen to the soundtrack with the utmost attention to words. This sequential combination was precisely what Hasegawa and Lander (1985) advocated after comparing the “video-video group”, who saw a video tape twice, the “audio-audio” group, who heard the soundtrack of the tape twice, and the “video-audio group”, who saw the video tape once and then listened to the soundtrack once.

Unlike the motion picture, which was perceived to be a blessing and a curse at the same time, captions as opposed to printed scripts were not appreciated at all. None of the subjects who voted for captioned film mentioned merits of the words on the screen and some who opted for audio+script did so because of perceived demerits of closed captions. Two commented that captions were harder to read than scripts. This perception could be interpreted in several ways. It might simply mean that blinking captions on the screen are somewhat tiring to read than a script printed in black ink on white paper. Another explanation may be that learners are not used to processing captions, which do not always appear in sentences or in grammatical sense groups but in chunks synchronized with sometimes arbitrary breath groups. Also it is possible that learners feel pressed and insecure to keep processing words that keep disappearing in a few seconds. Although it could be an effective way to develop faster reading skills, this will surely take away more of the learners' attention which, in a listening class, should primarily be paid to the sound. One subject felt that the captions were not exactly synchronized with the sound, which could be true in scenes where exchanges were rapid. Although in the film used in this study, which was subtitled specifically for language learning, the captions were 100% accurate, in the case of ordinary subtitled films, as much as 20-40% (Kubota et al. 1990) or even 50% (Takahashi 1994) of the captions could differ from the sentences actually uttered. All these possibilities considered, it could be argued that, in the case of learners at this level, advantages of captions over script are rather difficult to find.

To summarize, in response to Question 1, the test scores and students' perception seem to indicate that using a film without captions for overall comprehension and the soundtrack plus printed script for assimilation of visual and audio images of words can produce results as good as, or better than, those obtained by utilizing the captioned version of that film.

## 6.2 Research Question 2

Research Question 2 was, “Is reading the script of each sentence prior to listening to the soundtrack more effective than simultaneously reading the caption and listening to the soundtrack?” Based on a finding by Takahashi (1994), the answer was predicted to be in the affirmative, which in fact turned out to be the case. Group A, who read the script prior to listening, showed a larger gain than Group B, who watched the scene with subtitles. It should be noted that this result was obtained even though Group B subjects viewed the video twice. Takahashi’s finding that listening after reading is more effective than listening and reading at the same time was confirmed.

Let us first discuss the limitations of simultaneous presentation. Takahashi argues that the average rate of delivery in commercial films is too fast for students. He reports that, even in the case of films which university students claim are relatively easy to listen to, the average rate of speech far exceeds the reading speed of average Japanese university students, which is almost always below 100 wpm (Ando 1979, cited in Takahashi 1994). They cannot even follow the captions by sight, much less by listening to the soundtrack. The responses by the subjects of this study to Survey Question 2 in fact support this argument. Thirteen subjects referred to the difficulty in following the captions.

However, fast speech may not be the only cause of low performance by the caption group. Even if the speed was not so fast, reading the captions and listening to the soundtrack could still be too much for learners because reading tends to take away most of their attention, as was pointed out by many researchers (Miyamoto 1990, Obari et al. 1993, Kamei and Hirose 1994). Some of the subjects in this study even employed the term “instinctive” to explain their attention automatically paid to the written words. In fact, d’Ydewalle et al. (1991), who investigated the eye-movement patterns of American and Dutch subjects watching films subtitled in either L1 or L2, reported that reading subtitles is more

or less obligatory because of the predominance of the visual information and its processing over the processing the soundtrack.

In contrast, the read-first-listen-later method has two advantages. First, when reading, learners can fully comprehend the sentence because ample time is allowed. Second, when listening, they can concentrate on aural perception because they already know what to expect.

Thus, Comparison 2 also downgraded captioned videos in favor of the simpler method of script plus ordinary videos or script plus the soundtrack because the read-first-listen-later method is impossible with subtitled films, unless they are controlled by computers. If repeating the procedure of looking at one sentence for a while and listening to the soundtrack is felt to be somewhat troublesome, even studying and getting familiar with the whole of the script to be covered in the next class could be recommended. Comprehending and memorizing the script as much as possible prior to a class session and devoting the whole period to pure listening activity could be considered one way of spending time efficiently in a listening class.

### **6.3 Research Question 3**

Research Question 3 was, "When listening to the soundtrack after reading the script, should the script be followed with eyes or be out of sight?" To put it another way, when listening with an expectation of what will come, which is more effective, the single-medium processing of audio only or the double-media processing of audio+script? Comparison 3 did not produce definitive results. The test scores revealed no significant difference between the two, while the survey results indicated there were more subjects who preferred audio-only than those who opted for audio+script.

Here again, looking into the subjects' verbal responses seems to help with examining the characteristics of the two treatments. Many of the subjects who voted for the double-media processing seemed to be aware

of the gap between their expectation of what a sentence would sound like and what it actually sounded like. As we all know too well, speech in the real world abounds in assimilation, weakening, omission, etc., which makes it really difficult for learners to associate the sounds with their written representation. The subjects who preferred audio+script probably wanted to see with their own eyes how the familiar sentences were pronounced, and that with good reason. It can be maintained that comparing the sound and its written representation is crucial, particularly when the gap between the two is felt to be wide. On the other hand, there were also many subjects who felt looking at the script prevented them from concentrating on the sound.

The question, then, is whether to look or not to look. A plausible answer may be, as in the case of the motion picture, that it depends. When the perceived gap between the sound and script is wide, following the script will surely help with assimilating the two. Also when the sentence is not yet stored in one's short-term memory, looking at the script will benefit learners. On the other hand, once the sentence gets into the short-term memory storage, depending on the script will become a hindrance to perceiving the sound itself. Therefore, a possible solution is looking at the script until the assimilation between the expectation and the reality is achieved and then concentrating on the soundtrack, or the reality.

#### **6.4 Suggested Teaching Procedure**

In the case of authentic commercial films, which are difficult even for advanced learners, simply developing the ability to catch the utterances in the film shown in class is challenging, motivating, and encouraging. Having such an ability could be considered a prerequisite to overtime proficiency development since at least part of such skills is expected to be transferred to when viewing other films (Takahashi 1994). Combining the insights drawn from the three comparisons above, we suggest the



following procedure for utilizing a film, its soundtrack, and the script with the aim of making learners be able to perceive, decode, and comprehend the soundtrack of the film.

- (1) Distribute the scenario to the students in advance. Instruct them to study the language and to comprehend the story line.
- (2) Show them the film to enhance their understanding of the setting and the situation.
- (3) Make them look at the script of each sentence/utterance for a while, and then have them listen to the corresponding soundtrack. Repeat this procedure for an appropriate number of times with each sentence. When listening to the soundtrack for the first few times, they should be allowed to look at the script and be encouraged to bridge the gap between the expected sound image and the actual sound. When learners feel they have confirmed how each utterance is pronounced, they are encouraged to concentrate on sounds with the script out of sight.
- (4) Have them listen to the soundtrack all through the scene.
- (5) When they feel they have become able to catch the sounds, they are shown the film once again.

## **7. CONCLUSION**

The present study compared several modes of script presentation in listening classes. Before summarizing the findings and the implications, it would be fair to point out some of the limitations of this study. One was the small number of the subjects. Since they were the only students whom the researcher was teaching in the academic year 1994, the number of the subjects could not be made larger. Another is that the reliability of the pre/post/retention tests was not confirmed mainly because the researcher was not familiar enough with the statistical procedures to determine how reliable the partial dictation tests were. Due to these limitations, the

findings of this study should not be interpreted as conclusive but rather as hypothesis-generating. Some of the hypotheses which could be formulated based on the test scores and the subjects' comments in this study were as follows:

- (1) A film, its script, and the soundtrack could serve as more than an excellent substitute for a film with closed captions.
- (2) First presenting the script and then the sound is more effective than simultaneous presentation.
- (3) When listening to the soundtrack, the script should be followed for better assimilation, and should be out of sight for better concentration.

Based on these implications, a procedure of utilizing a film in a conventional classroom setting was suggested. Further research should replicate this experiment, using pre/post tests with confirmed reliability measuring the performance of larger groups of subjects.

### ACKNOWLEDGEMENT

I would like to express my sincere gratitude to Ms. Ellen Eri Kawaguchi, one of my former classmates at Teachers College Columbia University, for her helpful advice and editing suggestions.

### REFERENCES

- Ando, Shoichi. "Sokudoku no hoho." *Yomu eigo*. Tokyo: Kenkyusha, 1979.
- Braun, C. "Interest-loading and Modality Effects on Textual Response Acquisition." *Reading Research Quarterly* 4 (1969): 428-444.
- Danan, Martine. "Reversed Subtitling and Dual Coding Theory: New Directions for Foreign Language Instruction." *Language Learning* 42.4 (1992): 497-527.
- d'Ydewalle, Gery, Caroline Praet, Karl Verfaillie, and Johan Van Rensbergen. "Watching Subtitled Television: Automatic Reading Behavior." *Communication Research* 18.5 (1991): 650-666.

- Edasawa, Yasuo, Osamu Takeuchi, and Kazuo Nishizaki. "Use of Films in Listening Comprehension Practice." *Language Laboratory* 26 (1989): 19-40.
- Harzem, P., I. Lee, and T.R. Miles. "The Effects of Learning to Read." *British Journal of Educational Psychology* 46 (1976): 318-322.
- Hasegawa, Kiyoshi and John Lander. "Visual Cues and Clues in Listening Comprehension." *Yokohama Kokuritsu Daigaku Kyoiku Gakubu Kyoiku Jissen Kenkyu Center Kiyo* 1 (1985): 41-55.
- Hirose, Keiko, and Setsuko Kamei. "Effects of English Captions in Relation to Learner Proficiency Level and Type of Information." *Language Laboratory* 30 (1993): 1-16.
- Holobow, N.E., W.E. Lambert, and L. Sayegh. "Pairing Script and Dialogue: Combinations That Show Promise for Second or Foreign Language Learning." *Language Learning* 34.4 (1984): 59-76.
- Itakura, Takeko. "Visual Aids and Listening Comprehension —Are They Really Beneficial?—" *Journal of Fukuoka Jogakuin Junior College* 22 (1986): 29-44.
- Kamei, Setsuko. "Eigo gakushu ni okeru maruchi media no koka: teiji naiyo tono kankei de." Paper read at the JACET 33rd Annual Convention. Nagoya, 10 September, 1994.
- \_\_\_\_\_, and Keiko Hirose. "Multimedia Effects on L2 Comprehension in Relation to English Learner Proficiency Level." *Language Laboratory* 31 (1994): 1-17.
- Kubota, Akira, Hiroshi Otake, and Hiroyuki Obari. "Tekisuto bunseki kara miru bideo kyapushon no kanosei." *LLA dai 30 kai zenkoku kenkyu taikai happyo yoko* (1990): 46-48.
- Lambert, W.E., I. Boehler, and N. Sidoti. "Choosing the Languages of Subtitles and Spoken Dialogues for Media Presentations: Implications for Second Language Education." *Applied Psycholinguistics* 2 (1981): 133-148.
- Levie, W. Howard, and Richard Lentz. "Effects of Text Illustrations: A Review of Research." *Educational Communication and Technology Journal* 30.4 (1982): 195-232.
- Miyamoto, Setsuko. "Eiga bideo kyoza no koka: nihongo jimaku to eigo jimaku o hikaku shite." *LLA dai 30 kai zenkoku kenkyu taikai happyo yoko*. (1990): 49-51.
- Nugent, Gwen C. "Pictures, Audio, and Print: Symbolic Representation and Effect on Learning." *Educational Communication and Technology Journal* 30.3 (1982): 163-174.
- Obari, Hiroyuki, Takako Sugiyama, John Lander. "Comparative Study of the Use of Video in the Teaching of Listening Comprehension." *Tsukuba Daigaku Gaikokugo Center Gaikokugo Kyoiku Ronshu* 14 (1992): 21-42.
- \_\_\_\_\_, Akira Kubota, and John Lander. "Eigo jimaku tuki bideo kyoza no choshu

- rikai ni oyobosu koka no kenkyu." *Tsukuba Daigaku Gaikokugo Center Gaikokugo Kyoiku Ronshu* 15 (1993): 45-67.
- Reese, Stephen D. "Visual-Verbal Redundancy Effects on Television News Learning." *Journal of Broadcasting* 28.1 (1984): 79-87.
- Samuels, S.J. "Attentional Processes in Reading: The Effect of Pictures in the Acquisition of Reading Responses." *Journal of Educational Psychology* 58 (1967): 337-342.
- Takahashi, Hideo. "Eigo hearing gakushu ni okeru moji joho no koukateki teijihou ni kansuru kenkyu." *Working Papers in Language and Speech Science* 4 (1994): 125-136.
- Takai, Osamu. "A Comparative Study of the Effectiveness between Audio Medium and Video Medium in Teaching English." *Otaru Shoka Daigaku Jinbun Kenkyu* 82 (1991): 17-53.
- \_\_\_\_\_. "Are Movies Really Effective for Japanese Students to Improve Their Listening Comprehension Ability? : A Pilot Study." *Otaru Shoka Daigaku Jinbun Kenkyu* 86 (1993): 117-137.
- Takeuchi, Osamu, Yasuo Edasawa, and Kazuo Nishizaki. "Do Films Improve EFL Students' Listening Comprehension Ability?" *Language Laboratory* 27 (1990): 81-98.
- Vanderplank, Robert. "Paying Attention to the Words: Practical and Theoretical Problems in Watching Television Programmes with Uni-lingual (CEEFAX) Subtitles." *System* 18. 2 (1990): 221-234.
- Willows, D.M. "A Picture Is Not Always Worth a Thousand Words: Pictures as Distractors in Reading." *Journal of Educational Psychology* 70(1978): 255-262.

## APPENDIX

Pre/ Post/ Retention Test Items (Sample)

Instruction: Listen to the tape and fill in the underlined blanks.

1. Maybe ( I ) ( need ) ( them ) ( more ) ( than ) ( you ) do.
2. So I ( can ) ( stay ) ( here ) ( until ) ( they ) ( find ) ( my ) mother ?
3. You ( must ) ( be ) ( the ) ( lady ) ( of ) ( the ) house.

(福島工業高等専門学校)