

- seminar: Focusing on international contexts of multimedia and information technologies in higher education.* Tokyo: NIME.
- van Lier, L. 1991. Inside the classroom: Learning processes and teaching procedures. *Applied Language Learning*, 1 (2), 29-69.
- van Lier, L. 1996. *Interaction in the language curriculum: Awareness, autonomy, and authenticity.* London: Longman.
- Vygotsky, L.S. 1978. *Mind in society: The development of higher psychological processes.* Cambridge: Harvard Univ. Press.
- Vygotsky, L.S. 1986. *Thought and language.* Cambridge: MIT Press.
- Warschauer, M. (ed.) 1995. *Virtual connections* Honolulu: SLTCC, Univ. of Hawaii.
- Warschauer, M. 1997. Computer-mediated collaborative learning: Theory and practice. *Modern Language Journal*, 81(4), 470-481.
- Warschauer, M., L. Turbee, and B. Roberts 1996. Computer learning networks and student empowerment. *System*, 24 (1), 1-14.
- Wood, D., J. S. Bruner, and G. Ross 1976. The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, 17, 89-100.
- Xiaomin, K. 1994. Foreign language education and media: From a situated learning's perspective. *Monograph Series* Vol.5, 81-89. Kobe: Kansai Chapter, the Language Laboratory Association of Japan.
- Zhao, Y. 1997. Language learning on the world wide web: Toward a framework of network based CALL. *CALICO Journal*, 14 (1), 37-51.

Modalities of Subtitling and Foreign Language Learning

Haruyo Yoshida	Setsunan University
Shigeo Uematsu	Setsunan University
Shinsuke Yoshida	Setsunan University
Osamu Takeuchi	Kansai University

要約

本論文は、(1)映像と英語音声+日本語字幕、(2)映像と英語音声+英語字幕、(3)映像と日本語音声+英語字幕の三つの異なる音声と字幕のコンビネーションが、言語獲得の前提となる内容理解にどのように影響するかを実証的に考察したものである。被験者は、日本人中級英語学習者 156 名で、英語聴解力テストにより 3 つの等質なグループに分けられた。このそれぞれのグループに、上記の 3 条件で題材を提示し、その内容理解を調べた。題材にはアメリカ映画の一部を利用した。その結果、(2)→(1)→(3)の順に内容理解が高いことがわかった。この結果を被験者のアンケートなどから得られた結果と総合すると、中級学習者にとっては、多様な情報源(映像、音声、文字)を利用することが困難であり、特に情報が外国語で提示された場合には、外国語字幕→外国語音声の順で困難さが増すことがわかった。本論文では、この結果に基づき、効果的な音声と字幕のコンビネーションの在り方、今後の研究の方向性などについても言及していく。

1. Literature and Background

It has been more than two decades since video was first introduced to EFL (English as a Foreign Language) classrooms in Japan with the aim of facilitating instruction. Video provides learners with real, authentic target language information through the use of media-combinations such as pictures and soundtrack or pictures, soundtrack and subtitles (also known as captions). For example, in watching foreign commercial videos with subtitles, we receive information from three sources of information at one time; pictures (visual), soundtrack in L2 (second language, English) and subtitles in L1 (first language, Japanese) translated

from L2 scripts.

Paivio (1986) states that in the case of second or foreign language learning, when there are richer and more meaningful non-verbal referents, such as objects, events or emotions, that are taken in by learners in the form of non-verbal information, better language recall and appropriate production can occur. One can easily understand the beneficial effects of multi-presentation when recalling his or her own experience of the use of mnemonic associations in acquiring new information, such as an association between a vocabulary item (*e.g.*, apple) and the visual image of the object referred to (red, round-shaped). These mnemonic associations can be also reinforced by the auditory information channel in the form of utterance of the target vocabulary on the soundtrack.

In watching captioned video, there might be a further implementation in mnemonic associations when subtitled textual information which appears in captions is sufficiently absorbed by learners. Cohen (1987) contends that learners use some associative memory aids automatically, and it is critical for language teachers to examine ways to maximize the benefits of mnemonic help for learners. Though it is crucially important for educators to know in depth, the nature and the process of these beneficiary interactions of multi-presentations as d'Yedewalle and Gielen (1992) points out, the distribution of attention between different channels of information turns out to be a diverse and complicated process among different individuals, since viewers seem to develop strategies that allow them to process these channels at their own choice according to their own interest.

The advantage of multi-presentation with picture and soundtrack based on dual coding theory (Paivio, 1986) has been supported by research (*e.g.*, Levie and Lenz, 1982 ; Bagget and Eherenfeucht, 1983). Concerning the "dual coding theory", Paivio states:

"There are two classes of phenomena handled cognitively by separate subsystems, one specialized for the representation and processing of information concerning nonverbal

objects and events, the other specialized for dealing with language." (Paivio, 1986:53)

However, the effect of adding subtitles has not been thoroughly explored yet in terms of its possible combinations with other channels of information, or in the choice between L1 and L2 as the means of input of messages.

Regarding the adding of subtitles in the EFL video instruction, there are three possible ways of combining subtitles and soundtracks; The first combination is "standard" which consists of L1 subtitles with L2 soundtrack. The second type is "bi-modal", with L2 (English) subtitles and L2 (English) soundtrack. The third is "reversed", with L2 (English) subtitles and L1 (Japanese) soundtrack.

Holobow, Lambert, and Sayegh (1984) tested the comprehension and memory in L2 of English speaking pupils in French immersion programs using two different combinations of channels. One was "reversed", a combination of dialogues in L1 (English) and coordinated with printed scripts in L2 (French). The other was "bi-modal", consisting of coordinated dialog and scripts both in L2. They found that "reversed" combination showed impressive effects on L2 comprehension and L2 contextual meaning comprehension, and observed that "bi-modal" facilitated advanced learners' comprehension and memory of L2 text given.

Danan (1992) examined the L2 comprehension of a total of 102 U.S. college students in beginning and intermediate French classes using French videos under the three viewing modes: standard, reversed, and bi-modal. She found advantages in both reversed and bi-modal subtitles, while standard subtitles contributed little to learners' retrieval of L2 information. She also advocates that, based on her findings, a model which integrates both reversed and bi-modal inputs into a complete curriculum should be explored.

Hirose and Kamei (1993) examined the effects of subtitles using bi-modal subtitled video for 275 Japanese EFL students. They found that students understood better when seeing the film with captions, and that

L2 English subtitles facilitated understanding through the learner's linguistic knowledge such as grammar and vocabulary. They claim:

"Captions help learners, especially those who have grammatical knowledge and vocabulary, but are relatively poor at listening, such as the intermediate group in this study, to understand the dialogue. Captions can supplement lack of listening ability with other abilities students are more proficient in." (Hirose and Kamei, 1993: 11)

The aim of this study is to investigate the effectiveness of the multi-presentation of video under three different combinations of L1/L2 inputs in subtitles and in soundtracks, i.e., standard, bi-modal, and reversed. The focus is especially on the effectiveness of "reversed" mode which has not been investigated for Japanese EFL learners. Due mainly to the technical difficulties, reversed videotape has neither been used in the classroom nor investigated. The replacement of English (L2) dialogue with Japanese (L1) dialogue requires a special technique to coordinate the Japanese dialogue with what is going on in the pictures. Sometimes translation does not work without appropriate modifications and manipulations.

The research questions based on the literature review are as follows; (1) when an English video is presented to three groups with different viewing conditions (modalities) of standard, bi-modal, and reversed, is there any significant difference to be observed among respective mean test score for their comprehension of the story? In addition, (2) which of these three combinations do the learners feel is most helpful in understanding the context?

2. The Study

2.1 Subjects

The subjects of this study were 156 university freshmen. They were from three different sections of the same English course in a private university in Osaka. Each class consisted of about fifty members. They had

studied English more than 6 years in junior high school and high school. Subjects who had experiences in living in English speaking countries were excluded from the study group.

2.2 Method

After the instructor explained the goals and objectives of this study to the subjects of these classes, a listening section of the Comprehensive English Language Test (CELT ; Form A) was administrated as a pre-test in order to determine if any significant proficiency difference existed. (See Harris and Palmer, 1986 for the details of the CELT test.)

Table 1 Descriptive Statistics for Pre-Test

	n	Means	SDs
Standard	52	71.37	9.69
Bi-modal	51	71.41	13.56
Reversed	53	72.28	9.29

The test took approximately 30 minutes to complete. Table1 shows the descriptive statistics for pre-test. A one-way ANOVA shows that there is no significant differences among mean scores ($F(2,153) = .12, p < .89$). Furthermore, Hartley's test rejected the homogeneity of variance among three groups ($F_{max} = 2.13, p < .05$). As a result, those three groups were almost the same in terms of their mean scores for the pre-test.

Then after the pre-test, each of these three groups was shown a ten minute scene excerpt from an American video, *The Purple Rose of Cairo*, directed by Woody Allen, with the combinations of dialogue in soundtrack subtitles as follows;

- (1) Group one (standard mode): English soundtrack and Japanese subtitles
- (2) Group two (bi-modal mode): English soundtrack and English subtitles
- (3) Group three (reversed mode): Japanese-edited soundtrack and English subtitles

They were then asked to answer ten comprehension questions (See Appendix 1) based on the video excerpt. All of these questions were multiple choice type questions. Each question was written so it could be answered correctly if a learner was able to use the intended information source from picture, subtitle and soundtrack, or use a combination of more than one of those. In addition, standard and reversed groups were asked to give written comments concerning their experiences.

The study took up about a half of two different 90 minute lecture periods. In the first class, subjects were informed of the goals and objectives of this study and they were also informed that their participation in this study was optional, that is, if they did not want to take part in the study, they had a choice not to do so. In the second class, they sat for the CELT listening test. They were then informed that their participation in the test would have nothing to do with grading.

2.3 Data Analysis

The answers were checked and scored respectively by the researchers. For each question, there was only one appropriate answer and the total number of correct answers was counted up for each subject. Mean scores for each question item and category were compared across the three groups. For the entire mean score comparison among three groups, Sheffe tests were administered. For question level analysis, the FREC procedure of SAS package of statistical program was used to create the three-way frequency and crosstabulation tables.

3. Results

The outcome of the comprehension test indicates that the mean score of reversed group turned out to be higher than that of the standard group, and both of these were significantly higher than that of bi-modal group. ($p < .05$ see tables 2 & 3)

As for the micro-level analysis, the FREC procedure revealed that bi-modal group was found to be significantly lower ($p < .05$) than other groups in terms of its scores for Questions 1,2,6,7,8, and 9. (table 4)

Table 2 Descriptive Statistics for three groups

	n	Means	SDs
Standard	52	8.21	1.87
Bi-modal	51	5.90	1.56
Reversed	53	8.74	1.55

Table 3 The Result of Sheffé Test

	Reversed	Bi-modal
Standard	.278	.000*
Bi-modal	.000*	

($p < .05$)

Table 4 The Crosstabulation of Comprehension Questions

		Standard	Bi-modal	Reversed	Total
Q1	Incorrect	5	23	6	34
	Correct	47	28	47	122
Q2	Incorrect	6	7	3	16
	Correct	46	44	50	140
Q3	Incorrect	8	11	3	22
	Correct	44	40	50	134
Q4	Incorrect	13	13	11	37
	Correct	39	38	42	119
Q5	Incorrect	12	32	14	58
	Correct	40	19	39	98
Q6	Incorrect	7	31	4	42
	Correct	45	20	49	114
Q7	Incorrect	14	33	8	55
	Correct	38	18	45	101
Q8	Incorrect	8	26	7	41
	Correct	44	25	46	115
Q9	Incorrect	10	22	4	36
	Correct	42	29	49	120
Q10	Incorrect	10	11	7	28
	Correct	42	40	46	128

4. Discussion

The results could be interpreted as follows: in order to better answer the comprehension questions, subjects had to understand the context to a sufficient level. When information was given only in L2 (bi-modal), it was difficult for these intermediate EFL learners to understand what was happening on the screen, which contrasts with what Hirose and Kamei's findings (1993) in which intermediate EFL students were able to utilize even L2 caption for better comprehension. As Holobow *et. al* (1984) pointed out, bi-modal seems to work only when subjects are advanced learners. It might be difficult for intermediate EFL learners to gain substantial amount of contextual information from pure L2 linguistic environment. When context information was given in learners' L1(Japanese), however, it was relatively easier for them to answer the content questions. Also, when soundtrack was provided in learners' L1 (reversed mode), it might have been a better source of information than captions in L1 (standard) because L2 sound information was processed chronologically with picture, hence better associated and understood.

Why was the bi-modal group consistently outperformed by other two groups? Although the information was presented through three sources, *i.e.*, picture, soundtrack and subtitle, if there was no L1 contained, it could not be utilized by learners possibly because of the lack of sufficient skills to understand information conveyed solely through L2. As for questions which could draw answers from pictures such as Q's 2,3, and 10, there were no significant differences among the three groups. Without information presented either in the form of picture or through L1, it seems, that they could not follow the context of the story.

The written comments made by both standard and reversed groups show that 29% subjects of standard group felt safe and comfortable because they got information in L1 through subtitles, while 72% of the subjects of the reversed group felt the same because they got information in L1 through the reversed soundtrack. In the reversed group, some subjects commented that they did not feel like they were improving L2 listening comprehension ability because all the aural information was

given only in L1(Japanese). However, they felt secured and could enjoy the video content because they did not suffer interference from the anxiety often caused by the difficulties in understanding English language. Regarding the combinations of L1/L2 in subtitles and soundtrack, 21% of the standard group subjects thought that the combination pattern of their group would improve listening skills. 32% of the standard group subjects pointed out that the combination of their group contributed not only to the improvement of listening skills, but also to reading skills that might contain some English/Japanese translation work.

On the other hand, as many as 54% of the subjects in the standard group pointed out that they could not do anything other than read the Japanese captions because they felt unable to do two things at the same time. 25% of the reversed group subjects pointed out that information conveyed through the soundtrack was so dominant that they did not pay any attention to reading subtitles. 8% of the standard group subjects claimed that even if one could read subtitles, there was no room available for utilizing pictures. 11% of the reversed group subjects pointed out that when they concentrated on reading English captions, they were unable to listen to the Japanese soundtrack, hence were unable to understand the context.

5. Conclusion

Before concluding, some shortcomings of the present study should be pointed out. First, this study tried to find a one-time effect on the EFL learners' comprehension of the three modes of soundtrack-subtitle combination. Long-term effects, which might be different from that of a one-time experience need to be investigated in a separate study. Second, the number of the subjects, a total of 156, is relatively small when divided into the three treatment groups. This small number might have caused some anomaly in the data analyzed. In connection to this, our subjects were from one university, and, thus, any findings of this study need to be replicated in other populations. Third, only a portion of one movie was used in our study. Consequently, the idiosyncrasies of the movie could

have exerted some influence on the results obtained.

With these shortcomings in mind, the following summary is in order: Our study reveals that, in facilitating intermediate EFL students' comprehension, reversed mode is the most effective combination of soundtrack and subtitles (followed by standard mode). Contrary to the Hirose and Kamei's study (1993), however, our subjects were unable to utilize L2 information sources, and, thus, bi-modal mode turned out to be the least effective one. In this connection, our micro-level analysis of the comprehension questions also indicates that if there is no L1 information available, it might be extremely difficult for intermediate EFL students to obtain necessary information from authentic movies.

Using L1 information, be it subtitle or soundtrack, can be an effective option, if properly used, for promoting intermediate EFL students' comprehension. As Danan (1992) implies, investigating the proper use of L1 information in language teaching is the next agenda that needs to be explored.

References

- Bagget, P. and Eherenfeucht, A. (1983). Encoding and Retaining Information in the Visuals and Verbals of an Educational Movie. *Educational Communication and Technology Journal*, 31(1), 23-32.
- Cohen, A. (1987). The Use of Verbal and Imagery Mnemonics in Second Language Vocabulary Learning. *Studies in Second Language Acquisition*, 9, 43-62.
- Danan, M. (1992). Reversed Subtitling and Dual Coding Theory: New Directions for Foreign Language Instruction. *Language Learning* 42(4), 497-527.
- d'Yedewalle, G. and Gielen, I. (1992). Attention Allocation with Overlapping Sound, Image, and Text. In Rayner, K. (Ed.), *Eye Movements and Visual Cognition: Scene Perception and Reading*. New York: Springer-Verlag.
- Harris, D.P. and Palmer, L.A. (1986). *CELT: Examiners' Instructions and Technical Manual*. New York: McGraw-Hill.
- Hirose, K. and Kamei, S. (1993). Effects of English Captions in Relation to Learner Proficiency Level and Type of Information. *Language Laboratory* 30, 1-16.

- Holobow, N. E., Lambert, W.E., and Sayegh, L. (1984). Pairing Script and Dialogue: Combinations That Show Promise for Second or Foreign Language Learning. *Language Learning*, 34(4), 59-76.
- Levie, W. H. and Lenz, R. (1982). Effects of Text Illustrations: A Review of Research. *Educational Communication and Technology Journal*, 30(4), 195-232.
- Paivio, A. (1986). *Mental Representation: A Dual-Coding Approach*. New York, NY: Oxford University Press.

【Appendix 1. Comprehension Questions】

The followings are ten comprehension questions on the video, *The Purple Rose of Cairo, Scene One*. Each question is followed by the bracket which shows the information source(s) from which subjects gain the information necessary to answer the question. The abbreviations P, C, S respectively refer to Picture, Caption (subtitle), Soundtrack. The asterisk before number shows the correct answer for each question.

- Q1. What does the manager tell Cecilia about the new movie? (C, S)
1. More exciting. 2. More interesting.
3* More romantic. 4. More depressing.
- Q2. Where does Cecilia work? (P)
1* At a diner. 2. At an ice factory.
3. At a theater. 4. At the Copacabana.
- Q3. What is Monk doing when Cecilia goes to meet him? (P)
1. Looking for a job. 2. Trying to forget his troubles.
3. Going to the ice factory. 4* Playing a game.
- Q4. Where does Cecilia tell Monk there might be some jobs? (C, S)
1. At the movie theater. 2. At the restaurant.
3. At the laundry. 4* At the ice factory.
- Q5. What does Monk ask Cecilia for? (C, S)
1. Some tips. 2. His laundry.
3. The dice. 4* Some money.
- Q6. What city do the characters in the movie decide not to visit? (C, S)
1. Casablanca. 2. Cairo.
3. Tangiers. 4* Paris.
- Q7. How long have Henry, Jason, and Rita been traveling when they meet

Tom? (S)

1. Since yesterday.
2. A week.
3. For years.
- 4*Two weeks.

Q8. Why is Tom in Egypt? (C,S)

1. Because he is meeting the king.
- 2*Because he is looking for the Purple Rose of Cairo.
3. Because he is buying a new pith helmet.
4. Because he is digging in an old tomb.

Q9. Where did a fortune teller predict Tom would fall in love? (C,S)

1. Cairo.
2. the Copacabana.
- 3*New York.
4. Morocco.

Q10. How many people go the Copacabana? (P,C,S)

1. Four.
2. Five.
- 3*Six.
4. Seven.

【Appendix 2. Comments】

Translated into English from Japanese originals. Each comment is followed by the bracket which shows the total number of the subject who gave similar comments.

< Standard group >

Supporting Opinions

- *This method is good for me to practice an English listening skill.(9)
- *I often check out a video film at a rental video shop, so I am used to this method.(9)
- *I can make sure the content both by reading Japanese subtitles and listening to English dialogues.(7)
- *By repeating listening to the English dialogues, I feel I will be able to understand the content without reading the Japanese subtitles.(2)
- *If I can't understand the English dialogues, I try to supplement the meaning of the context with the Japanese subtitles.(2)
- *I can get a lot of information through both English and Japanese channels.(2)
- *This method is dynamic and makes me feel as if I were in a movie

theatre.(2)

- *I can identify the easy English expression only by listening to the English dialogues, as well as I can make sure of the meaning in the Japanese subtitles.(1)
- *This method is a good way to train the translation ability from English into Japanese.(1)

Reluctant Opinions

- *I only read Japanese subtitles without listening to the English dialogues.(11)
- *This method is inevitable because I don't have enough ability to understand the English dialogues provided only with English sound.(7)
- *I can't understand the context in details.(3)
- *I devote myself to reading the Japanese subtitles too much, which prevents me from watching and enjoying the picture part.(3)
- *I strongly depend on the Japanese subtitles, which fixes me to the screen.(3)
- *I can't do two things such as reading the Japanese subtitles and listening to the English dialogues at once.(2)
- *I tend to skip the unknown words because of concentrating on reading the Japanese subtitles.(2)

< Reversed group >

Supporting Opinions

- *Because of the Japanese dialogues, I can understand the meaning of the context easily.(20)
- *I can understand the content and the meaning of the vocabulary by both the Japanese dialogues through the ears and the English subtitles through the eyes.(7)
- *I can make sure the meaning of various expressions by reading the Japanese subtitles.(5)
- *This method is a good way to practice the English composition.(2)

- *I don't know if it's good or not, but I can follow the story without reading the English subtitles because the Japanese dialogues are provided at once. (2)
- *I don't have to concentrate on listening, so I can devote myself to read the English subtitles.(1)
- *This method will improve my English reading ability.(1)
- *I realized there were many expressions of translating the English.(1)
- *I feel easy, because I can read more than hear the English.(1)
- *This method will improve my English reading speed.(1)
- *First I read, then I hear, which helps me understand the difficult English expression.(1)
- *I can understand the grammar through both English and Japanese channels.(1)
- *I can study English by comparing with English and Japanese.(1)
- *I am used to the doubted foreign movies on TV, so I feel comfortable.(1)

Reluctant Opinions

- *I don't read the English subtitles because the Japanese dialogues are too strongly come into my mind.(13)
- *Without the English sound, I don't think I learn anything in the English class.(9)
- *I devote myself to read the English subtitles too much, which prevents me from watching and enjoying the picture part.(4)
- *When I concentrate on reading the English subtitles, the Japanese dialogue sounds tend to be ignored, as a result I ended up not grasping the meaning of the context.(4)
- *My English reading speed is not fast enough to follow the English subtitles on the screen, so I got frustrated. (2)
- *This method lacks in dynamics.(2)
- *Foreigners speak Japanese (doubted) fluently is funny and unnatural.(1)
- *I feel pain of my eyes.(1)

- *I feel frustrated with the slight timing difference of the Japanese dialogues and the English subtitles.(1)