The Effect of Caption Modes on EFL Students' Video Comprehension

Jiangqiao Dong
School of Foreign Studies, Jiangnan University, China

Yajing Zhou
School of Foreign Studies, Jiangnan University, China

Guiru Liu
School of Foreign Studies, Jiangnan University, China

Abstract—Based on perception load theory and cognitive theory with respect to multimedia learning, an experiment was conducted to examine whether caption modes or the amount of captions (full captions, keyword captions and no captions) affected EFL students' distribution of attention resource in their visual channels and accordingly video comprehension. 147 Chinese students of English majors participated in the experiment. Results indicated that there were no significant differences between the three groups regarding the effect of video materials on participants' overall comprehension, which did not conform to Guillory's result. There might be three reasons for it. Apart from that, for the comprehension of picture information, there were no significant differences between the group with keyword captions and the group with no captions but both were obviously better than the group with full captions. As for the scores of language information, there were no significant differences between the group with full caption and the group with keyword captions, but both were better than the group with no captions. Two explanations for the result were proposed.

Index Terms—video comprehension, caption, multimodal, attention distribution

I. INTRODUCTION

In L1 learning environments, learners' auditory language ability is usually higher than their visual one in that their listening comprehension is better than reading comprehension. Hence video programs with captions mainly aim at young children for their reading training and adults for their literacy education (Goldman, 1993, quoted in Dai, 2005). In foreign language learning environments, however, learners' auditory language ability is usually lower than their visual one in that their listening comprehension is not as good as reading comprehension (Dai, 2005). Accordingly captions may make up for listeners' insufficiency of auditory language processing ability. The point is that this sort of compensation mechanism is rather complicated, because comprehension of multimedia videos with captions is closely associated with the material's organizational forms, learners' cognitive strategies and knowledge skills, objectives and requirements of cognitive processing tasks and their interactive relations. Previous researches with respect to the effect of captions focused on the overall effect of captions on video comprehension or the achievement effect of a single task and more or less neglected the discussion about the process itself. During a specific "video" processing, the effect of captions usually depends on the video watcher's level of perceptual load and selective attention.

II. THEORETICAL BASIS AND RELEVANT RESEARCHES

A. Theoretical Basis

Working memory (WM) refers to the resource-limited system that is used for temporary information storage and processing (Baddley, 1992) while selective attention refers to the cognitive system which selects the rich external information so as to ensure the efficient use of the limited cognitive resource (Down, 2000). Compared with the single linear text reading and sequential discourse listening comprehension, multimedia video watching need consume more attention resource. The perceptual processing of static pictures, dynamic videos and language information (including audio and visual ones) roughly corresponds to visual-spatial sketchpad, episodic buffer and phonological loop in Baddeley's model of working memory (Baddeley, 2000, see figure 1) but their occupation of attention resources is not "balance-loaded". Based on the researches on working memory, Lavie (1995; 2004) proposed the theory of perceptual load, believing that perception is a resource-limited process-cycle. The level of perceptual load for the current task determines the resource allocation in the selective attention. If the level is high, the limited attention resource is consumed, the other objects or information that is irrelevant to the current task can not be processed perceptually so that interfering effect will not be achieved. Otherwise the target stimulus can not consume the current attention resource, the redundant resource will automatically overflow, the processing will have nothing to do with the interfering stimulus.

The point is that the observer can not restrain the overflow of the redundant resource (Wei, 2005). Lavie's theory is deemed to have settled the psychological disputes on pre-selection and post-selection. Nevertheless, researches by Johnson et al. (quoted in Ren, 2009) further indicated that the decisive factor that affected selective attention was the resource limitation and that the load level of current task directly influenced the distribution of attention resources.
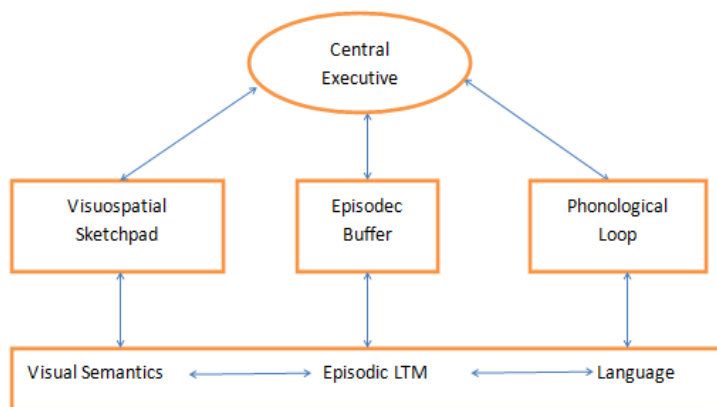

Figure 1 Baddeley's Model of Working Memory

Cognitive load is composed of the causal dimension that reflects the interactivity between tasks and learners' characteristics and the evaluative dimension that reflects the measurable notions such as psychological load, effort and performance, etc. To promote effective learning, appropriate materials should be selected so as to control the internal cognitive load, decrease the external cognitive load via effective teaching design, increase the relevant cognitive load and ensure that the overall cognitive load does not exceed the level that learners are able to bear.

Based on dual coding theory, Mayer (2001) proposed the three hypotheses of multimedia cognitive learning theory: dual channel hypothesis (visual channel and audio channel), limited capacity hypothesis (the processing capacity of each channel is limited) and active learning hypothesis. In multimedia learning, learners' processing comes from the information through two channels and the limited capacity merely allows part of text and picture information to be noticed by learners, enter their working memory and be actively processed, created into coherent structure and integrated with the previous knowledge so that meaningful learning may occur.

The above theories reveal the cognitive processing mechanism of multi-mode information in working memory from different perspectives. What they have in common is that they all involve the distribution of attention resource in the perceptual process with different focuses. Compared with single-channel linear text reading or sequential discourse listening comprehension, multimedia cognitive learning is multi-mode processing which processes the discrete visual information such as linear text, videos, sounds, pictures, etc. as well as the sequential audio information. In the audio-visual training based on acoustic visual material, Chinese students have obvious drawbacks in both visual competence of language processing (rapid reading of captions) and audio one (authentic sound listening), in particular in listening comprehension which is in most cases not as good as reading comprehension. Acoustic multimedia video watching in general belongs to cognitive processing of high perceptual load and different information in the same channel competes for watchers' attention resource. Videos with captions have the effect of cognitive compensation in language learning mainly in that information on target stimulus surpasses EFL learners' competence in a sense. The point is that on the similar conditions of high cognitive load, whether selective attention is self-conscious or automatic and on what conditions there will be participation of cognitive control process.

B. Literature Review of Relevant Researches

Scholars outside China made numerous researches on the positive effect of captions on video comprehension, for example, effect of L2 captions on content comprehension and memory (Price, 1983; Vanderplank, 1988; Garza, 1991), effect of captions on vocabulary acquisition (Neuman et al 1992), interactive effect between L2 proficiency and captions (Price, 1983; Vanderplank, 1988; Guillory, 1998; Danan, 1992) and effect of different captions and dub modes on L2 acquisition (Danan, 1992; Baltova, 1999; Markham et al., 2001, see Wang, 2005).

In China there have been comparatively less researches on application of captions in foreign language teaching and people began to make relevant researches mainly after 2000. Wang (2005) and Wang (2009) made some empirical studies, believing that captions were beneficial for Chinese L2 learners' accidental vocabulary acquisition and they advocated that teachers should make every effort to find the appropriate ways for students so as to maximize the effect of captions. Dai (2005) theoretically summarized the origins, development, categorization and the educational effect of captions, pointed out that captions might play some role in listening, speaking and reading and advocated the appropriate use of captions to assist listening comprehension. Ren (2009) also suggested that if designed and applied properly, the flexible captions based on SRT format might become good assistant in listening teaching and training.

Nevertheless, there have been few empirical studies on the mechanism of caption effect and design in foreign language video watching. Apart from that, amount of caption and combination forms of different categories of videos have not been given sufficient attention.

Barron & Atkins (1994) made researches on the different roles of key word caption, full caption and no caption videos in learners' comprehension of knowledge. The differences of the three groups lied in the amount of captions. The learning process was autonomously controlled by participants and repetition or playback was allowed. Results indicated that the amount of captions did not affect learners' learning effect. Nevertheless, the experiment was not a language learning task and the language used for the experiment was participants' native language (L1). Apart from that, participants were allowed to watched the videos repeatedly. Hence the role of text was weakened, captions had no obvious effect upon participants' listening comprehension and the change of caption modes had no effect on learners' amount of comprehension in terms of knowledge points.

Guilory (1998) put keyword captions into L2 teaching, respectively compared the differences between French full captions, keyword captions and no captions for adult French learners whose L1 was English in terms of videos comprehension. Results revealed that for the mean score of the three groups, there were no significant differences between keyword captions and full captions with both significantly higher than no captions. Guilory's study confirmed the positive effect of captions on comprehension and there were some points to be noted. Firstly, the researcher examined participants' comprehension of video content but failed to define what the content was. For example, was it information about language or picture? The researcher did not seem to investigate them respectively. Secondly, the researcher did not clearly pointed out the difficulty level of videos compared with that of learners' language proficiency. Thirdly, in the research L1 was English and L2 was French. The two languages were of the same origin and captions of such kind were beneficial for the prediction of word meaning. Guillory explained this and pointed out that the language of the same origin as keywords were more likely to attract learners' attention. English captions should not have similar effect on EFL teaching in the context of Chinese since English and Chinese are not the languages of the same origin. Hence whether Guillory's result can be applied to Chinese EFL teaching is to be validated.

III. RESEARCH DESIGN

A. Method

According to perceptual load theory, learners' capacity of working memory is limited. Multiple tasks will follow the principle of no change in the total amount with one more and the other less. The objectives of tasks will more or less affect learners' selective attention. When confronted with the same task, learners of different language proficiency and with different objectives usually make different psychological efforts so that different cognitive loads arise. Therefore the load of current task directly affects the distribution of attention resource. The researcher of this study believes that when there is no specific task, the attention resource for videos watching with high perceptual load is automatically distributed. But when specific learning tasks are given, it is selective for the distribution of attention resource for videos watching with both high and low perceptual loads. The watching effect depends on the resource competing level of single channel processing and the efficient use of channel characteristics of attention resource (Interfering information from audio channel is not affected by the level of visual perceptual load). Hence when there are merely pictures and texts, and attention resource is limited, the change of caption modes or amount of captions will absolutely change learners' selective attention and do not affect their performance.

This experiment intends to examine whether the amount of captions or caption modes (full captions, keyword captions and no captions) affect learners' distribution of attention resource in terms of picture and caption cognitive processing when video difficulty level is controlled within i +1 by Krashen so as to make a difference in their video comprehension of texts and pictures. In other words, will the differences of amount of information in learners' visual channel make differences in their comprehension of videos? The following three questions are to be answered: (1) Does the change of amount of captions affect learners' total comprehension achievement of videos? (2) Are there any differences between the comprehension achievement of language information and that of non-language due to the change of amount of captions? (3) Does the change of amount of captions exert the same influence upon videos in different contexts?

B. Participants

The participants of this experiment were 180 English majors from Hubei Engineering University, China. 33 of them took the pretest and the rest 147 were randomly divided into 6 groups who participated in the video test. For the experiment groups and the caption modes, see table 1. Before the experiment, the 6 groups took a simulation test of TEM4 (Test for English Majors-Band 4). ANOVA analysis indicated that there were no significant differences between the 6 groups in terms of English proficiency (F=.160, p=.997).

C. Materials

Experiment materials included 6 videos which lasted 25 to 35 seconds and a test paper whose reliability coefficient was 0.91. The difficulty level of videos was about the same as CET4 (College English Test-Band 4 and the mean speed was 130 wpm. Of the scripts, 14 words were not included in *College English Curriculum Requirements*, 2.7% of the

total amount. The topics of the 6 videos covered family life, natural disasters, social problems and future dreams, etc. that learners were familiar with.

According to Bejar's categorization of videos and Hymes' study of context variables, the 6 videos used for this experiment were divided into three kinds, namely videos with strong contexts, videos with weak contexts and videos with no contexts (Du, 2010) and each of them included 2 videos.

Key word captions were drawn from the scripts. Five senior English teachers and two native speakers of English (teachers from English-speaking countries) respectively read the scripts of the videos and picked up the key words from them. Words that were chosen four times or above were determined as key words which were 38% of the scripts. The full captions and keyword captions matched the dubs and appeared or disappeared as the audios played.

The test paper adopted multiple-choice and covered all the information points of the videos, including language and non-language ones. The researcher asked five of the students to watch the pictures with no audios or captions. The other five students read the scripts of the videos and wrote down the information points obtained. The points appeared three times and above were chosen as the items in the multiple-choice. Apart from that, in order to ensure the relative independence of language and pictures, keys to all items were gained from scripts or pictures only. In other words, the information that could not only be gained from the scripts but also from the pictures were excluded. The items were arranged in order of the appearance of information points themselves. Each time when the participants finished watching a video, they were demanded to finish the correspondent items.

D. Procedures and Data Collection

33 students took the pretest and based on the result of the test, the researcher adjusted the time of experiment and the contents of test items and became familiar with the procedures of the experiment. 147 participants were divided into 6 groups and finished the experiment in the sound lab. Each of the student used a computer and a pair of earphones. Each video was played once, then items about the videos appeared in the screen and students wrote down their keys on the answer sheets.

After the experiment, the researcher collected the answer sheets and used SPSS19.0 to analyze it. The data included the total scores, scores for language information, pictures, videos with strong contexts, language with strong context, pictures with strong context, videos with no context, language with no context and pictures with no context respectively. For the results and scores for each of the components, see table 2.

IV. RESULTS AND DISCUSSION

A. Effect of Video Materials on Participants' Overall Comprehension

From table 2 it can be seen that there were no differences between the three groups with captions in terms of the total mean scores (F=0.509; P=0.603). In other words, the three caption modes or amount of captions did not significantly affect learners' total amount of video information. This result did not conform to that of Guillory's who concluded that there were no differences between full captions and keyword captions with respect to video comprehension but both of them were better than no captions. There might be three explanations for the differences between the results of current research and that of Guillory's. *Firstly*, the testing methods were different. In Guillory's study, items included recall questions and inferencing questions and participants were demanded to write down their answers on their own accord, which involved the use of learning strategies. In the current research, however, all the questions were recall ones and the participants did not need to infer, which authentically reflected the amount of information that participants could obtain in current working memory and decreased the effect of differences between learning strategies and inference ability on the result of the experiment. *Secondly*, the research methods were different. In Guillory's study, participants' L1 was English and their L2 was French which was of the same origin as English. The appearance of captions in French might lead to the fact that participants used learning strategies of association and inference. In this research English and Chinese were not of the same origin and captions in Chinese would not produce similar effect. *Thirdly*, there were no significant differences between the three groups in total scores, which could neither explain the fact that different amount of captions had the same effect upon the videos watching effect for the three groups, nor indicate that in case of different combinations of videos, audios and captions, effects of videos watching for the three groups were simply the same.

In this experiment, in view of the measurability of captions' cognitive compensation effect, the videos chosen were somewhat difficult (Being too difficult or easy, the compensation effect would accordingly be decreased and even disappear and become pure reading comprehension). Videos watching with full captions easily caused high perceptual load defined by Lavie, due to the large amount of target visual information (video pictures and caption text). Hence when the group with full captions watched the videos, both the picture information and the caption text information caused the distribution competition of attention resource for the visual channel. Participants could merely process part of the visual information and neglected the rest, which affected the total scores.

In contrast with videos with full captions, videos with keyword captions presented much less language information, and amount of information for watchers' decreased, which would cause the loss of part of language information but videos with less captions meant less perceptual load and as well as less need of attention resource. According to Lavie's perceptual load theory, if the current task could not consume attention resource, the extra attention resource would

automatically overflow and process the other information. The group with keyword captions processed less amount of text information than the previous group in their visual channels and the attention resource that was released when watching would automatically overflow and process the non-language information in the videos (pictures were the greatest stimulus next to captions in visual route). In this case, the information that participants obtained would compensate each other with one more and the other less.

As far as the group with no captions was concerned, in their visual channels there was nothing but pictures. Accordingly they obtained more picture information which compensated the language information performance that had been lost for the lack of captions in terms of amount and there were no significant differences in their total scores.

B.   Differences between Pictures and Language Information in Terms of Caption Modes Effect

For participants' scores of pictures and language information and their analysis, see table 3 and 4 (F=10.389; P=0.000). For the comprehension of picture information, there were no significant differences between group two (with keyword captions) and group three (with no captions) but both of them were better than group one with full captions; for the scores of language information there were no significant differences between group one and two (see table 5), but both were better than group three (F=10.248; P=0.000). The results indicated that for the comprehension of picture information, group with keyword captions and group with no captions had overflowed resource to process picture information, and both of them obtained more scores than the group with full captions and for the comprehension of language information, groups with captions were better than the group with no captions. There might be the following possible explanations.

When participants from group with full captions watched the picture information and caption information in their visual channels, the amount of information in it (pictures + text) exceeded the attention resource that could be distributed. The two categories of information in the same channel competed for attention resource and hence neither could be effectively processed. This group gained least scores for picture information, indicating that participants' attention focused on caption reading (so as to make up for their insufficient EFL listening) so that large amount of picture information was lost and that the participants got low scores. It could also be speculated from the above that the limitation of resource was the key factor that affected selective attention and that learners would unconsciously focus on caption reading to effectively understand what they were watching.

Group with keyword caption read comparatively less amount of information, the attention resource needed by their visual channels was accordingly less and the automatically overflowed resource would be used to process other stimulus such as audios and pictures. As the limitation of attention resource had the characteristics of channels, the interference information from the auditory channel was not affected by the visual and perceptual load (Tellinghuisen et al. 2003). Hence the relatively rich visual resource might process the video information, which might lead to the effect of context clue (For the details, see 4.3) and the focus of attention resource on the location of the target stimulus so that interference from outside could be avoided and interference effect could be decreased on the conditions of low load. Therefore this group could obtain both picture information and language one and learnt effectively in a sense.

Group three with no captions had merely pictures in their visual channels and should obtain the highest scores for picture information, but results revealed that there were no differences between group three and group two (group with keyword captions). The possible reason might be that in this experiment, the processing of picture information could not consume participants' attention resource and the extra attention resource overflowed and processed the other interference stimulus (object or information that had nothing to do with videos) so as to form interference and in turn affect the scores for picture information. The language cognitive performance of this group was also the worst of the three due to the lack of compensation effect of captions on listening.

Via comparison between the participants from the three groups in terms of picture information and language information, it could be seen that when the difficulty level of videos was under control and the condition of i + 1 was satisfied, participants' attention resource was gradually released so that they had free time to process picture information as the amount of caption reading decreased. However, without captions, the attention resource that had overflowed exceeded the amount needed for picture processing, would again overflow automatically, process the other irrelevant information and eventually lead to the interference effect on the conditions of low perceptual load.

C.   Effect of Context on Video Comprehension of Materials with Captions

For the analysis of scores obtained by participants from the six groups in their comprehension of language information, see table 4. There were no significant differences between group one and four (P= .967). The differences between the two groups lied in the videos watched with or without pictures. The result indicated that the existence of pictures did not affect participants' performance. Apart from that, when participants watched the videos with full captions, their attention resource were completely occupied by the captions and no extra resource overflowed to process the picture information. Group four processed the double modes of audios and texts with no extra visual (picture) interference and effectively made use of the channel characteristics of attention resource so that their language performance was about the same as that of group one. From the analysis in 3.2, it could be seen that group with keyword captions obtained significantly more scores than group with full captions with respect to pictures, indicating that the videos were not that difficult. It could be concluded from the above analysis that the difficulty level of videos used in this experiment was appropriate and could meet the requirements of the experiment.

From table 4 it could be seen that the existence of pictures seemed to exert no effect on participants' performance and that there were no significant differences between group one and four, group two and five, group three and six, which did not conform to Mayer's (2001) multimedia effect. Nevertheless, as the above analysis was made on the basis of various videos, no obvious multimedia effect of video pictures was found. Scores for the various videos revealed their specific effects. The six videos might be categorized into videos with strong context, videos with weak context and videos with zero context.

For the analysis of scores for the first three groups in videos watching with strong context, see table 5. Results indicated that there were no significant differences between the three groups in terms of total scores, meaning that the total amount of information the three groups obtained was about the same. See table 6 for the analysis of scores for the six groups in videos watching with strong context with respect to language information, and table 7 for the analysis of scores for the first three groups in picture information. Although the three groups were not equal to each other in terms of the amount of captions presented, there were no significant differences in their scores for language information, indicating that when pictures were closely associated with video contents, the language information that participants had lost in their visual channels could be compensated via the discourse information and clues with strong context in their auditory channels so as to produce the multimedia effect called by Mayer, which could be also seen from the comparisons between group four, five and six who watched videos without pictures. Without the compensation of pictures, the three groups mainly obtained language information from audios and captions, which led to the significant differences between the three groups (Group four was obviously better than group five and six). Simultaneously group two gained more scores than group five and group three got more than group six.

V. CONCLUSION AND IMPLICATIONS

Based on perceptual load theory, cognitive load theory and multimedia learning theory, this research examined learners' distribution of attention on language and picture information in different cases via control of caption modes in their visual channels and there were some implications for the results with respect to how to make better use of video materials.

From the above results it could be seen that when learners watched videos with relatively high level of difficulty, different amount of captions would affect their distribution of attention. Their insufficiency of videos watching led to the fact that the learners passively turned to caption reading for help and the mode with full captions occupied most of their attention resources. Although learners could obtain the maximum amount of language information, they had no extra attention resource to be distributed to the processing of picture information and caused the loss of picture information. The drawbacks of videos with no captions were obvious and accordingly large amount of language information was lost. Particularly when there was lack of context clues (videos with weak and zero contexts), the effect of multimedia decreased and the irrelevant interference increased and the watching effect was poor. Hence the use of keyword captions not only provided the main language information, but also released part of attention resource to be used in processing the picture information via the decrease of reading amount of visual language information so as to obtain both categories of information.

In this experiment the researcher also found the interaction between captions and different kinds of videos. In videos with strong context, context clue might activate multimedia learning effect and accordingly decrease the use of captions as much as possible so as to decrease cognitive load, start selective attention and improve the processing effect of information on the target stimulus. In videos with zero/no captions, picture information could not provide context clue, easily caused extra effect and interference on language processing. It should be noted that videos watching of full captions with no videos or keyword captions proved to be a good choice for sheer listening training in languages, because the limitation of resources had the characteristics of channels and the interference information in the auditory channel was not affected by the high or low visual and perceptual load. Accordingly in the use of auditory and visual materials, EFL teachers are supposed to discriminate different categories of videos and use different caption modes so as to maximize the teaching and learning effect.

APPENDIX.   TABLES

TABLE 1
CAPTION MODES FOR THE SIX GROUPS

Groups	Audios	Pictures	Full Captions	Keyword Captions	No Captions	
1	√	√	√			
2	√	√		√		Audios+pictures+captions
3	√	√			√	
4	√		√			
5	√			√		Audios+captions
6	√				√	

TABLE 2
SCORES FOR EACH OF THE COMPONENTS FOR THE SIX GROUPS

	Pretest	Total scores	Scores for pictures	Scores for language
Group 1 M (SD)	72.29(6.58)	54.71(9.57)	49.71(13.91)	58.20(10.19)
Group 2 M (SD)	73.57(6.24)	55.53(9.04)	60.00(11.66)	52.40(11.00)
Group 3 M (SD)	73.23(6.40)	53.06(7.80)	65.14(10.78)	44.60(9.46)
Group 4 M (SD)	73.44(4.03)			58.33(13.16)
Group 5 M (SD)	72.94(4.85)			50.83(11.76)
Group 6 M (SD)	73.06(5.43)			39.79(12.55)
F	0.160	0.509	10.389	10.249
Sig.	0.977	0.603	0.000	0.000
P<0.05				

TABLE 3
COMPARISON BETWEEN SCORES FOR PICTURE INFORMATION

Group	Group	Mean Difference (I - J)	Std. Error	Sig.	95% Confidence interval	
					Lower limit	Upper limit
Full caption	Keyword caption	-10.28520*	3.44704	.004	-17.1568	-3.4136
Full caption	No caption	-15.43000*	3.44704	.004	-22.3016	-8.5584
Keyword caption	No caption	-5.14480	3.44704	.004	-12.0164	-1.7268
P<0.05						

TABLE 4
ANALYSIS OF SCORES FOR LANGUAGE INFORMATION

(I) Group	(J) Group	Mean Difference (I - J)	Std. Error	Sig.	95% Confidence interval	
					Lower limit	Upper limit
1	2	5.80000	3.22549	.074	-.5766	12.1766
	3	13.60000*	3.22549	.000	7.2234	19.9766
	4	-1.3333	3.25891	.967	-6.5760	6.3093
	5	7.36667	3.25891	.025	.9240	13.8093
	6	18.40833*	3.25891	.000	11.9657	24.8510
2	3	7.80000*	3.22549	.017	1.4234	14.1766
	4	-5.93333	3.25891	.071	-12.3760	.5093
	5	1.56667	3.25891	.631	-4.8760	8.0093
	6	12.60833*	3.25891	.000	6.1657	19.0510
3	4	-13.73333*	3.25891	.000	-20.1760	-7.2907
	5	-6.23333	3.25891	.058	-12.6760	.2093
	6	4.80833	3.25891	.142	-1.6343	11.2510
4	5	7.50000*	3.29200	.024	.9919	14.0081
	6	18.54167*	3.29200	.000	12.0336	5.0497
17.5497	5	6	11.04167	3.29200	.001	4.5336
P<0.05						

TABLE 5
ANALYSIS OF SCORES FOR EACH GROUP IN VIDEOS WITH STRONG CONTEXT

(I) Group	(J) Group	Mean Difference (I - J)	Std. Error	Sig.	95% Confidence interval	
					Lower limit	Upper limit
1	2	-3.27200	4.51789	.470	-12.2036	5.6596
	3	-5.81840	4.51789	.200	-14.7500	3.1132
2	3	-2.54640	4.51789	.574	-11.4780	6.3852
P<0.05						

TABLE 6
ANALYSIS OF SCORES FOR LANGUAGE INFORMATION IN VIDEOS WITH STRONG CONTEXT

(I) Group	(J) Group	Mean Difference (I - J)	Std. Error	Sig.	95% Confidence interval	
					Lower limit	Upper limit
	2	1.71440	5.22043	.743	-8.6060	12.0348
1	3	2.85720	5.22043	.585	-7.4632	13.1776
	4	3.45252	5.27453	.514	-6.9749	13.8799
	5	15.35627*	5.27453	.004	4.9289	25.7836
	6	14.16585*	5.27453	.008	3.7385	24.5932
2	3	1.14280	5.22043	.827	-9.1776	11.4632
	4	1.73812	5.27453	.742	-8.6893	12.1655
	5	13.64187*	5.27453	.011	3.2145	24.0692
	6	12.45145*	5.27453	.020	2.0241	22.8788
3	4	.59532	5.27453	.910	-9.8321	11.0227
	5	12.49907*	5.27453	.091	2.0717	22.9264
	6	11.30865*	5.27453	.034	.8813	21.7360
4	5	11.90375*	5.32808	.027	1.3705	22.4370
	6	10.71333*	5.32808	.046	.1801	21.2466
5	6	-1.19042	5.32808	.824	-11.7237	9.3428
P<0.05						

TABLE 7
ANALYSIS OF SCORES FOR PICTURE INFORMATION IN VIDEOS WITH STRONG CONTEXT

(I) Group	(J) Group Difference	Mean (I - J)	Std. Error	Sig.	95% Confidence interval	
					Lower limit	Upper limit
1	2	-12.00000	6.04612	.051	-24.0527	.0527
	3	-21.00000*	6.04612	.001	-33.0527	-8.9473
2	3	-9.00000	6.04612	.141	-21.0527	3.0527
P<0.05						

REFERENCES

[1] Baddeley, A.D. (1992). Working memory. *Science*, 55 (31), 556-559.
[2] Baddeley, A.D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Science*, (4), 417-423.
[3] Baltova, I. (1999). The effects of subtitled and staged video input on the learning and retention of Content and vocabulary in a second language. Unpublished doctoral dissertation. University of Toronto.
[4] Barron & Atkins. (1994). Audio instruction in multimedia education: Is textual redundancy important? *Journal of Education Multimedia and Hypermedia*, (3/4), 295-306.
[5] Dai, J. (2005). Captions in movies and foreign language teaching. *Computer-assisted Foreign Language Education*, 137(3), 47-53.
[6] Danan, M. (1992). Reversed subtitling and dual coding theory: new directions for foreign language instruction. *Language Learning*, 42(4), 54-65.
[7] Downing, P. E. (2000). Interaction between visual working memory and selective attention. *Psychological Science*, 11(6), 467-473.
[8] Du, B. (2010). Effect of situational videos on EFL listening comprehension in the context of CALL. Unpublished master dissertation, Jiangnan University, China.
[9] Garza, T.J. (1991). Evaluating the use of captioned video materials in advanced foreign language learning. *Foreign Language Annals*, 24(3), 81-94.
[10] Guillory, H.G. (1998). The effects of keyword captions to authentic French video on learner comprehension. *CALICO Journal*, 15(1-3), 89-108.
[11] Krashen, S. D. (1985). The Input Hypothesis: Issues and Implications. Torrance, CA: Laredo Publishing Co.
[12] Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology*, 21(3), 41-55.
[13] Lavie, N. et al. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology*, 133(3), 74-86.
[14] Markham, P., Peter, L & McCarthy, T. (2001). The effects of native language vs. target language captions on foreign language students' DVD videos comprehension. *Foreign Language Annals*, 34(5), 63-78.
[15] Mayer, R.E. (2001). Multimedia Learning. New York: Cambridge University Press.
[16] Neuman, S. B. & Koskinen, P. (1992). Captioned television as comprehensible input: effects of incidental word learning in context for language minority students. *Reading Research Quarterly*, 27(1), 23-39.
[17] Price, K. (1983). Closed-captioned TV: An untapped resource. *MATSOL Newsletter*, (2), 12-25 .
[18] Ren, K.X. (2009). EFL listening teaching based on flexible captions. *Modern Education Technology*, 54(10), 69-74.
[19] Tellinghuisen, D. J. & Nowak, E.J. (2003). The inability to ignore auditory distractors as a function of visual task perceptual load. *Perception & Psychophysics*, 65(5), 817-828.
[20] Vanderplank, R. (1988). The value of teletext subtitles in language learning. *ELT Journal,* 42(4), 112-125.
[21] Wang, H. (2005). Effect of captions on incidental vocabulary acquisition. *Computer-Assisted Foreign Language Education*, 137(2), 56-60.
[22] Wei, P. (2005). Understanding selective attention from the perspective of perceptual load theory. *Advances in Psychological Science*, 13(4), 413-420.

Jiangqiao Dong is a professor in the School of Foreign Studies, Jiangnan University. His research interests include English applied linguistics and study of intercultural communication, with a focus on computer-aided language teaching.

Yajing Zhou is a lecturer in the School of Foreign Studies, Jiangnan University. Her research interests include computer-aided language teaching.

Guiru Liu is a master student in the School of Foreign Studies, Jiangnan University. Her research interests include English applied linguistics.