Effects of Video Caption Modes on English Listening Comprehension and Vocabulary Acquisition Using Handheld Devices

Ching-Kun Hsu[1], Gwo-Jen Hwang[2*], Yu-Tzu Chang[3] and Chih-Kai Chang[3]

[1]Department of Technology Application and Human Resource Development, National Taiwan Normal University, Taiwan // [2]Graduate Institute of Digital Learning and Education, National Taiwan University of Science and Technology, Taiwan // [3]Department of Information and Learning Technology, National University of Tainan, Taiwan // chingkun.hsu@gmail.com // gjhwang.academic@gmail.com // t09955018@stumail.nutn.edu.tw // chihkai@mail.nutn.edu.tw
[*]Corresponding author

ABSTRACT

This study investigates the effects of different display modes of video captions on mobile devices, including non-caption, full-caption, and target-word modes, on the English comprehension and vocabulary acquisition of fifth graders. During the one-month experiment, the status of the students' English listening comprehension and vocabulary acquisition was evaluated on a weekly basis. From the experimental results, it was found that the learning achievement of the English target-word group was as good as that of the full-caption group in terms of vocabulary acquisition, while both groups outperformed the non-caption group. When the students' learning preferences were taken into consideration, the visual style students in both the English target-word group and the full-caption group showed significantly better learning effectiveness in terms of vocabulary acquisition than those in the control group. Furthermore, in terms of listening comprehension, the students in all three groups made remarkable progress.

Introduction

English has been recognized as being an important language for international communications. In the past decades, many non-English speaking countries have developed and utilized various computer systems to support English as Foreign Language (EFL) learners. Moreover, most of these countries intend to have their children learn English as early as possible; for example, the Ministry of Education in Taiwan has extended regular English instruction to the third grade of elementary school and is planning to initiate an English learning program for first graders within ten years.

Foreign language learning can be considered from the four aspects of listening, speaking, reading, and writing. Among these aspects, listening is an important capability of social interactions, and it has been found that people receive new messages more efficiently via listening than reading (Luo, 2008). The Ministry of Education in Taiwan even guides elementary schools to put emphasis on English listening first, followed by reading or writing. In the meantime, most forms of English certification include listening proficiency tests, implying the importance of fostering English listening competences.

With the advance of mobile technologies and multimedia, instructional materials for English listening training have been developed in a variety of forms. For example, MP3 players have become a new mobile device for learning via listening; moreover, several web2.0 media, such as YouTube, have become popular channels that provide audio and video materials for language learning (Godwin-Jones, 2007). In the meantime, the popularity of various mobile devices (e.g., smartphones and e-books) and wireless networks (e.g., Wi-Fi and Wi-Max) have enabled individual students to use those emerging learning materials or channels anywhere and at any time (Hwang, Shi, & Chu, 2011; Hwang, Wu, Tseng, & Huang, 2011; Wu et al., 2011). Therefore, it can be foreseen that students will eventually be equipped with a mobile device installed with proper learning tools, systems or materials so that they can make their own learning progress, and even learn with content of appropriate difficulty to match their level of proficiency (Hung et al., 2012; Hwang & Chang, 2011; Hwang, Chu, Lin, & Tsai, 2011; Norris, Hossain, & Soloway, 2011). In recent years, several studies concerning the use of mobile technology in language learning have revealed the benefits of this approach, such as the provision of opportunities for individual practice and a seamless learning environment (Ogata,

Matsuka, El-Bishouty, & Yano, 2009; Ogata & Yano, 2004; Ozcelik & Acarturk, 2011; Wong, Chin, Tan, & Liu, 2010).

On the other hand, previous studies have reported that videos embedded with captions are helpful for students in learning second language reading (Chun & Plass, 1997) and listening (Danan, 1992). Hsu and Chang (2010) have further indicated that hiding the easier vocabulary and showing only the relatively difficult words in the captions may contribute to undergraduates' listening comprehension. Consequently, this study developed different display modes of captions on mobile devices for improving English listening competence and promoting the vocabulary acquisition of students. The study aims at exploring whether different display modes of captions and subtitles result in different degrees of effectiveness in the listening comprehension and vocabulary acquisition of elementary school students. Moreover, the learning performance of students with different learner preferences is also compared. Furthermore, the students' perceived satisfaction with the approach, and their perceptions of the usefulness and ease-of-use of the mobile learning activity are reported as well.

Literature review

Video captions and English listening comprehension

Using videos or films as a learning resource has received a great deal of attention from researchers and has been successfully applied to various educational applications (Yang, Huang, Tsai, Chung, & Wu, 2009). Researchers have indicated that multimedia learning materials are more useful than traditional paper-based instruction (Mackey & Ho, 2008; Mayer & Moreno, 2002; Rose, 2003). Videos which provide visual, contextual, and non-verbal input provide foreign language learners with simultaneous visual and aural stimuli which can make up for any lack of comprehension resulting from listening alone (Brett, 1995; Hoven, 1999; Seo, 2002). Several previous studies have shown that such videos are highly accepted by learners during the learning process (Choi & Johnson, 2005; Choi & Johnson, 2007; Mackey & Ho, 2008). Moreover, captions have been perceived as the most useful and efficient auxiliary for watching videos (Hsu, 1994), while videos with authentic accents have been recognized as being a good learning resource for language learners (Dahbi, 2004).

Researchers have indicated that students need to receive a great quantity of comprehensible input so as to achieve the objective of language learning when they learn a foreign or second language (Krashen, 1981; 1985). When students watch videos in a foreign language, the contribution of comprehending and connecting the foreign language and its meaning is limited if they cannot understand what they have heard. Therefore, using captions and subtitles is useful in enhancing the effectiveness of listening comprehension. Scholars have defined subtitles as the on-screen text in the students' native language combined with a second language soundtrack in the video, while captions are the on-screen text in the original language combined with a soundtrack in the same language (Markham, Peter, & McCarthy, 2001). Therefore, in this study, the term subtitles refers to on-screen Chinese text combined with an English soundtrack, while captions refers to on-screen English text combined with an English soundtrack. In addition, bilingual subtitling has been defined as on-screen text in both the students' native language and the target language combined with a target language soundtrack (Katchen et al., 2001). That is, bilingual subtitling as used in this study refers to English audio with the simultaneous appearance of English and target-word Chinese texts on the screen. Furthermore, the target words represent the new or key vocabulary of the target lessons in the following discussions.

Captions and subtitles were initially used in foreign language instruction in the 1980s, and many studies have confirmed that combining captions with audio-visual materials is an effective instructional method to enhance the listening and reading comprehension of a second language (Borras & Lafayette, 1994; Danan, 2004; Garza, 1991; Markham & Peter, 2003) because learners can confirm the information they hear by way of the support provided by the captions (Froehlich, 1988; Grimmer, 1992; Vanderplank, 1988). In other words, captions visualize the auditory information of the foreign language which the learners hear in the video (Danan, 2004). Because of such additional cognitive processing, captions and subtitles promote language comprehension (Bird & Williams, 2002). Researchers have further indicated that videos with captions facilitate vocabulary acquisition (Plass, Chun, Mayer, & Leutner, 1998), reading (Chun & Plass, 1997), and listening comprehension (Danan, 1992; Hsu, 1994; Markham & McCarthy, 2001).

On the other hand, Vandergrift (2007) stated that providing native subtitles for learners will obstruct their familiarity with pronunciation. Among many different assisted methods for learners who use videos to train their foreign

language listening, Liou (2000) found that the highest proportion of students used foreign language captions as auxiliary support, especially the higher-achieving students. The replay function was the next, and the assistance of native subtitles was the third ranked strategy when learners needed support. Therefore, adopting an innovative approach, this study provided either full or partial foreign language captions, but did not provide full native subtitles for the learners in any of the three groups.

Mobile assisted language learning

Mobile learning creates diverse possibilities for innovative instructional methods to be carried out in the general classroom in more effective and efficient ways. Scholars have noted that mobile-assisted language learning (MALL) provides students with rich, real-time, convenient, and contextual learning opportunities, no matter whether they are inside or outside the classroom (Kukulska-Hulme, & Shield, 2008). At the present time, MALL is not only one of the main language learning resources for students, but it also contributes to the learning and utilization of new language skills (Hashemi, & Ghasemi, 2011). Scholars have found that learners pay most attention to captions, followed by video and audio, and acquire most words by associating them with visual images. They have therefore concluded that captioned video tends to aid recognition of written word forms and the learning of word meaning. On the other hand, non-captioned video tends to improve listening comprehension as it facilitates recognition of aural word forms (Winke, Gass, & Sydorenko, 2010).

Recently, an increasing number of MALL studies have been devoted to vocabulary acquisition and learning, and the learning outcomes have been significant (Chang & Hsu, 2011; Chen, & Chung, 2008; Lu, 2008; Kim, & Kim, in press). In one study, students were encouraged to use mobile phones to access video clips explaining English idioms (Thornton & Houser, 2005). This is part of an obvious trend that learners are spending more time using mobile devices rather than PCs to carry out learning activities (Stockwell, 2010). In another study, the language learners expressed positive attitudes toward using mobile phones to browse wireless application protocol sites, which was also found to be effective for developing listening skills and for student-centered learning (Nah, White, & Sussex, 2008). In virtue of the benefits of MALL, and as many students nowadays are so familiar with mobile devices and are used to applying them in learning, this study conducted a number of MALL instructional experiments, in which different modes of captions and subtitles were inset in the MALL video material for different groups, with the purpose of identifying better instruments for English as Foreign Language learners (i.e., EFL learners).

Learning style and the Technology Acceptance Model

Learning style refers to individual preferences regarding ways of learning, and affects how individuals accept stimuli, memories, thinking, and problem-solving. There are many different scholars who have proposed diverse categories of learning styles (Reichmann & Grasha, 1974; Fischer & Fischer, 1979; Keefe, 1979a; 1979b; Gregorc, 1979; Schmeck, 1982; Kolb, 1984; Myers & McCaulley, 1985; Honey & Mumford, 1986; Felder & Silverman, 1988). If teachers realize the difference in learning styles among learners, and design appropriate instructional methods or media, learners are likely to benefit. Taking the different learning styles of visual and verbal learners for example, visual learners remember best what they see, such as pictures, diagrams, flow charts, time lines, films, and demonstrations. Verbal learners on the other hand, get more out of words, such as written and spoken explanations. As this distinction between the visual/verbal dimensions is highly relevant to the use of videos in training the listening competence or vocabulary acquisition of a foreign language, this study employed the visual/verbal learning style dimensions to evaluate the learning preferences of the participants.

TAM is a well-known model proposed by Davis (1986) based on the Theory of Reasoned Action (TRA). This model provides a theoretical foundation for understanding how external variables influence the inner beliefs, attitudes and intentions of users, and then affect their use of technology. It also provides a way to explain users' behavior when accepting new information technology (IT) and to analyze the factors that influence their attitudes toward its use (Davis, 1989; Venkatesh & Davis, 1996). Perceived usefulness (PU) refers to the degree to which people believe that using a particular system would enhance their performance, while perceived ease-of-use (PEOU) refers to the degree to which people believe that using a particular system would be free from effort (Davis, 1989). This study uses the TAM questionnaire to investigate learners' perceptions of the different caption and subtitle modes inset in the MALL materials (i.e., video).

Method

Participants

The experiments were conducted in an elementary school in Taiwan, where the students learn English as a foreign language. The school categorizes the English proficiency of the students into three levels, A, B, and C, based on an English proficiency test developed by a group of experienced English teachers. The participants of this study were three classes of fifth graders who were 11 years old on average and were categorized as having an English proficiency of level C; that is, they were low-achievement students in English. During the learning activity, one class of 27 students, including 16 males and 11 females, was assigned to be control group one, another class of 28 students, including 12 males and 16 females, was assigned to be experimental group one, and the other class of 26 students, including 15 males and 11 females, was experimental group two.

Research design

The experiment was conducted for a month, as shown in Figure 1. The participants used PDAs (Personal Digital Assistants) to play the instructional video related to the lesson they studied in class each week. After spending time watching the video, they immediately took a test to evaluate their listening comprehension proficiency and vocabulary acquisition. Although this study provided PDAs as dedicated devices, it should be noted that other mobile devices, such as smartphones or e-books, could also be used to conduct the same learning activity.

During the learning activity, a video without any English captions or Chinese subtitles was used for the students in the control group because previous studies have indicated that providing no captions or subtitles helps students adapt to various pronunciation features, such as reduced forms, assimilation, elision, and re-syllabification (Vandergrift, 2007).

On the other hand, a video with full English captions and Chinese subtitles of only the target vocabulary was used for the students in experimental group one because a previous study showed that such a setting is helpful in training listening proficiency and comprehension, and confirmed that full Chinese subtitles are not needed (Hsu & Chang, 2010). Another video with both English captions of the target vocabulary and Chinese subtitles of the target vocabulary was used for the students in experimental group two.
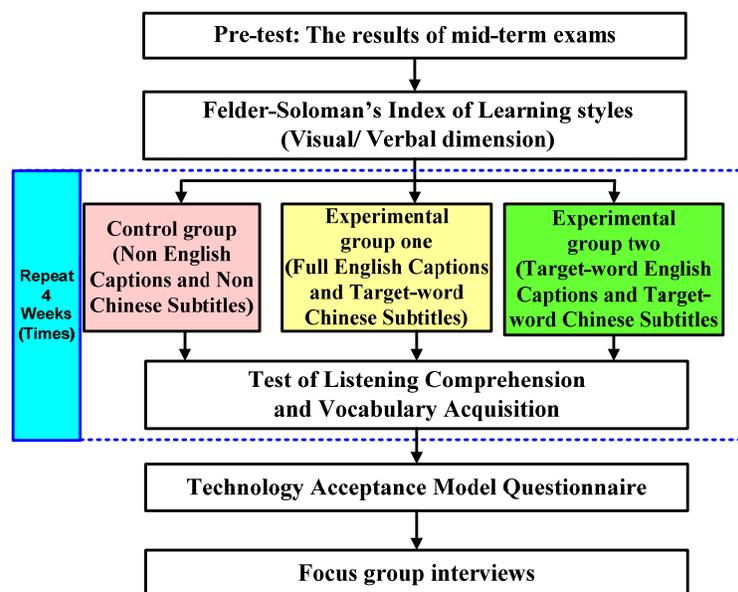

Figure 1. Experimental procedure

To summarize, the videos for the three groups consisted of the same content with different caption modes. No captions were provided for the control group, while full English captions and Chinese target words were provided for

experimental group one, and English and Chinese target words were provided for experimental group two. Figure 2 shows an example of the caption modes for experimental group one (left) and experimental group two (right).

Figure 2. The playing interfaces of experimental group one (left) and two (right)

Researchers have indicated that, for students to get used to the tempo of usual conversation, there is no need to provide fast forward or slow play functions; instead, only the functions of play, pause, and replay are necessary for listening training (Grgurovi & Hegelheimer, 2007; Winke, Gass, & Sydorenko, 2010). In order to take part in the mobile assisted listening training, the students in each group could use a stylus to select the functions of play, pause, and replay to listen in the limited time, as shown in Figure 3. As shown in the left photo of Figure 3, a timer was used to remind the students of the time limitation while watching the video. After conducting the post-test and administering the TAM questionnaire, this study further interviewed one third of the students (i.e., nine students from each group) to collect more detailed and in-depth feedback.

Figure 3. The students use PDAs to facilitate their English listening training

Research tool

The research tools in this study included the visual/verbal learning style measure, learning achievement tests, and the questionnaire for measuring the students' technology acceptance.

The test sheets were developed by two experienced teachers. The pre-test consisted of two groups of questions about the students' prior knowledge of the listening comprehension content and their vocabulary capability. It consisted of twenty-six multiple-choice items for examining listening comprehension and twenty-four multiple-choice items for testing vocabulary capability, with a perfect score of 100, and was also the school's midterm English exam. The weekly assessment and the post-test all consisted of five multiple-choice items for evaluating the students' listening comprehension, and five multiple-choice items for assessing the vocabulary acquisition of the lesson unit. The perfect score of the weekly tests and the post-test was 100. Each listening comprehension test was audio broadcast, and the students were asked to listen to the questions and fill out the answers on the answer sheet.

The Index of Learning Styles (ILS) Questionnaire was developed by Felder and Soloman (1991) based on the learning styles proposed by Felder and Silverman (1988). The original ILS measure consists of four dimensions (i.e., sensing/intuitive, visual/verbal, active/reflective and sequential/global), each of which contains 11 items. In this study, the "visual/verbal" dimension was adopted. The visual/verbal dimension consists of 11 items with a Cronbach's alpha value of 0.74 (Litzinger, Lee, & Wise, 2005).

In addition, this study employed the Technology Acceptance Model (TAM) questionnaire recently revised from the TAM questionnaire for mobile learning (Chu, Hwang, & Tsai, 2010; Chu, Hwang, Tsai, & Tseng, 2010). Three dimensions are explored in the questionnaire, that is, ease-of-use, perceived usefulness, and perceived satisfaction. The questionnaire uses a six-point Likert-scale, where 1 represents ''strongly disagree" and 6 represents ''strongly agree." The Cronbach's alpha values for the three dimensions, ease-of-use, perceived usefulness, and perceived satisfaction, are 0.88, 0.98 and 0.91, respectively.

Experimental results

Analysis of the pre-test and post-test

One of the objectives of this study was to examine the effectiveness of the different caption and subtitle modes with respect to the students' listening comprehension and vocabulary acquisition. It was therefore first necessary to ensure that the three groups had comparable listening comprehension and vocabulary levels before beginning the treatment. This was determined by way of the pre-test for which the mean values and standard deviations of the scores were 42.07 and 5.07 for the control group, 41.57 and 5.80 for experimental group one, and 42.15 and 8.27 for experimental group two. The ANOVA analysis results of the pre-test among the three groups do not show a significant difference ($p = 0.94 > .05$); that is, it was ascertained that the three groups of students had equivalent prior knowledge before the learning activity.

In addition, by employing ANCOVA on the post-test scores of the three groups, no significant difference was found between the listening comprehension scores of the three groups (F = 1.94, $p > .05$), while significant differences were found between the vocabulary acquisition scores of the experimental groups and those of the control group (F = 3.71, $p < .05$), as shown in Table 1. As the adjusted means of experimental group one (68.81) and experimental group two (67.17) were both significantly higher than that of the control group (51.10), the students who watched the videos with captions (no matter whether they were full English captions with Chinese target words or only English target words with Chinese target words) revealed significantly better learning achievements in English vocabulary than those who learned without captions. Therefore, the provision of Chinese target-word subtitles could be the crucial factor in promoting the effectiveness of vocabulary acquisition.

Table 1. ANCOVA analysis results of the vocabulary acquisition tests

Group	N	Mean	SD	Adjusted Mean	*F*	Pairwise comparisons
Control group (a)	27	48.15	24.34	51.10	3.71*	(b) > (a)*
Experimental group one (b)	28	71.43	26.35	68.81		(c) > (a)*
Experimental group two (c)	26	70.00	33.11	67.17		

*p < .05

Table 2. Paired sample t-test between the pre-test and the post-test

Group	N	Listening comprehension			Vocabulary acquisition		
		Pre-test Mean	Post-test Mean	*t*	Pre-test Mean	Post-test Mean	*t*
Control group (a)	27	42.07	65.93	-5.21**	23.90	48.15	-5.47**
Experimental group one (b)	28	41.57	77.14	-8.24**	26.61	71.43	-9.70**
Experimental group two (c)	26	42.15	71.54	-6.81**	26.56	70.00	-7.24**

**p<0.01

Table 2 shows the paired sample t-test results between the pre-test and the post-test. It can be seen that the three groups all made significant progress in listening comprehension and vocabulary acquisition during the four weeks

with p < .01. This finding complies with what has been reported by researchers, that children have more extensive opportunities to develop their listening proficiency if their instructors select appropriate multimedia materials and assistance (Van Scoter, Ellis & Railsback, 2001; Wright & Shade, 1994). In this study, the use of mobile devices with videos allows the elementary school students to have more extensive opportunities to practice in order to promote their listening comprehension and vocabulary acquisition.

In addition, this study also examined the learning progress of students' listening comprehension and vocabulary acquisition within the four weeks. Figure 4 shows the average test scores of the three groups over a period of a month. It is found that all of the three groups made progress each week; however, the two experimental groups had significantly better learning effectiveness than the control group, in particular, since the third week. In Figure 4, it can also be seen that the two experimental groups revealed similar learning performance during the month.
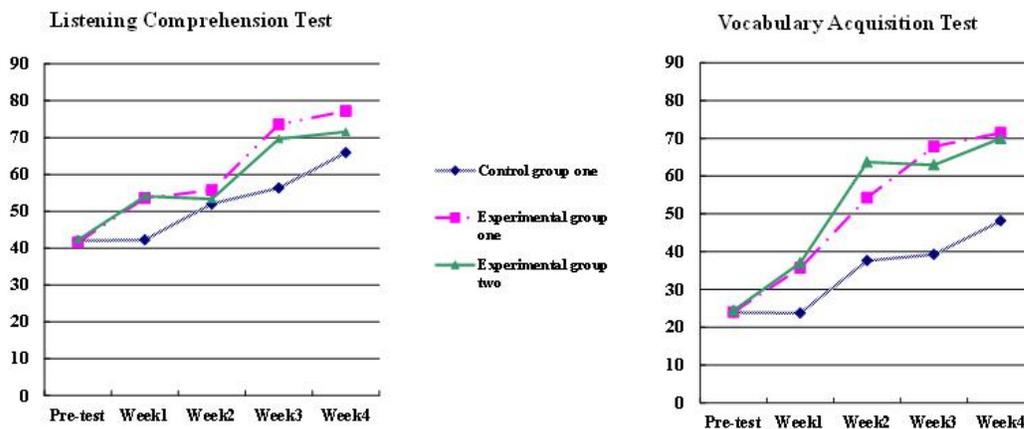
Figure 4. The improvement progress of listening comprehension (left) and vocabulary acquisition (right)

Difference in learning achievement of diverse learning preferences

This study further investigated the effect of different caption modes on the English listening comprehension and vocabulary acquisition of the students with different learner preferences. Table 3 displays the ANCOVA analysis of vocabulary acquisition which shows a significant difference for the visual learning preference students among the different groups. The students who prefer a visual style of learning in the target-word group performed as well as the students who prefer a visual style of learning in the full-caption group for vocabulary acquisition; moreover, the students who prefer a visual style of learning in both the target-word group and the full-caption group outperformed the students who prefer a visual style of learning in the non-caption group. Accordingly, for the students who prefer a visual style of learning, it is suggested to provide both English and Chinese target words to them, in particular, for those low-achievement students.

Table 3. ANCOVA analysis of vocabulary acquisition of the students preferring visual/verbal styles of learning

Learning style	Group	N	Mean	SD	Adjusted Mean	F	Pairwise Comparisons
Verbal (L)	Control group(L1)	7	51.43	27.95	51.17	0.84	
	Experimental group one(L2)	9	73.33	33.17	69.53		
	Experimental group two(L3)	9	62.22	38.01	68.42		
Visual (V)	Control group(V1)	18	47.37	24.23	51.55	3.23*	(V1) < (V3)*
	Experimental group one(V2)	19	70.53	23.45	68.71		(V1) < (V2)*
	Experimental group two(V3)	17	74.12	30.63	70.30		

*p<.05

Consequently, based on the findings of this study, it is recommended that learners who prefer a visual style of learning use the assistant mode applied in experimental group two for the purpose of promoting their learning comprehension, because they tend to pay attention to the image or video input rather than relying on listening to the

verbal input. In contrast, the students in the target-word group had similar performance in the listening comprehension to the students in the full-caption group. As a result, there is no need to provide full captions, as the use of target words only can provide students with more opportunities to practice listening to various pronunciation features, such as reduced forms, assimilation, elision, and re-syllabification.

Perceptions of the target word subtitles

Nowadays, there is a high level of acceptance of the use of information technology such as mobile devices as learning assistance. This study especially paid attention to the acceptance of experimental groups one and two. Therefore, Table 4 shows the descriptive statistics results of TAM for experimental groups one and two. On the six-point Likert-scale, where 1 represents ''strongly disagree'' and 6 represents ''strongly agree,'' the descriptive statistics of the technology acceptance model in experimental group two all attain scores higher than 5 on average. This means that the students agreed with and had positive perceptions of satisfaction, usefulness, and ease-of-use of the mobile learning system embedded with target-word Chinese subtitles and target-word English captions, as well as expressing high acceptance of the approach. There is no remarkable difference between the two experimental groups on the three scales of satisfaction, usefulness, and ease of use, as shown in Table 4. In addition, the students highly accepted the assistance of the target-word Chinese subtitles. The following discussion also presents the opinions of learners provided in the interviews for comparison with the results of the technology acceptance model investigation.

Table 4. The descriptive statistics of TAM for experimental groups one and two

Scale	Experimental group one			Experimental group two		
	N	Mean	SD	N	Mean	SD
Satisfaction	28	4.83	1.33	26	5.00	1.05
Usefulness	28	4.65	1.46	26	5.08	1.01
Ease of Use	28	4.58	1.49	26	5.07	1.04

Discussion and conclusions

This study found that the target-word strategies have better effects on vocabulary acquisition than on listening comprehension for low-achievement elementary school students. As students in non-English speaking countries generally lack English vocabulary, they especially need the assistance of the target words when watching videos for vocabulary acquisition. On the other hand, the results concerning listening comprehension are different from those of previous studies carried out in universities. One possible reason is that undergraduates have learned the frequently-used 2,200 English words, while elementary school students have only learned a few of these words. Consequently, the undergraduates were able to benefit from the practice, whereas the younger students could not (Hsu & Chang, 2010). In conclusion, this study suggests that the partial hidden caption mechanism can be used in an adaptive way that presents the selected vocabulary with different degrees of difficulty based on the learning level of the students.

From the interview results, there are several interesting findings. For example, the students in experimental group one (the full English caption group) indicated that it was not necessary to provide them with full English captions; moreover, they stated that showing full captions interfered with their listening to the learning materials. They believed that providing only target words was sufficient to assist them in improving their listening comprehension, which conforms to the results of the investigation of perceptions of using the system for learning English listening. The results showing the importance of target words and no need for full captions are similar to the research results of Guillory (1999) and Taylor (2005). In addition, some students said that they would like to learn from watching videos and playing computer games related to the topics in their textbooks.

Moreover, most of the students pointed out that the use of mobile devices was another important factor that motivated them to learn. They indicated that such a mobile technology-enhanced learning approach allowed them to practice anywhere and anytime, and hence learning the English listening content seemed to be easier and more interesting. The teacher who participated in this experiment further indicated that "*In the beginning, I wondered if the use of mobile devices could benefit the students. I worried that the screen of the PDA might be too small, and the*

students might spend lots of time playing games. Surprisingly, the students made significant progress each week, which has changed my perception in terms of using mobile technologies to learn." Such findings are consistent with the experiences reported by Wong et al. (2010).

In the future, we plan to conduct more experiments from several perspectives. The first is to use longer videos, since the present study only applied short videos of 20 seconds. This approach is suitable for elementary school students with low learning achievement, but may not be appropriate for advanced learners. Secondly, the study suggests that future researchers can determine which target words to show, and hide the other words in the captions of different video lengths for learners of different ages. Thirdly, although the results of this study seem to be desirable, it might be difficult to claim that the findings are significant since the test period was so short. It is possible that part of the results may be due to the novelty of viewing the videos with handheld devices. Therefore, it is worth conducting extended studies with a longer test period in the future. It is expected that the extended learning time may make contributions to listening proficiency, especially as there is a trend that mobile devices will become a common learning device in the near future (Norris, Hossain, & Soloway, 2011; Peng et al., 2009). Fourthly, as there are other handheld devices besides PDAs (e.g., e-books, notebooks, smartphones), it would be interesting to compare the learning performance of PDA learners with that of students who learn with other handheld devices. Fifthly, because English teachers may not have the necessary skills in information and computer technologies, we plan to establish caption processing tool freeware. When English teachers decide to employ this approach in their classes, they can use the freeware themselves to process the captions of their videos, or to select what words they want to show or hide in the captions, such as the target 1,200 words at the elementary and junior high school level, or the Dolch Sight words or keywords. Developing caption-filtering freeware contributes to making captions or subtitles reusable, diversiform and adaptive so that the application of videos in learning via mobile devices can be adapted to different users.

Acknowledgements

References

Bird, S. A., & Williams, J. N. (2002). The effect of bimodal input on implicit and explicit memory: An investigation into the benefits of within-language subtitling. *Applied Psycholinguistics, 23*(4), 509-533.

Borrás, I., & Lafayette, R. C. (1994). Effects of multimedia courseware subtitling on the speaking performance of college students of French. *Modern Language Journal, 78*(1), 61-75.

Brett, P. (1995). Multimedia for listening comprehension: The design of a multimedia-based resource for developing listening skills. *System, 23*(1), 77-85.

Chang, C. K., & Hsu, C. K. (2011). A mobile-assisted synchronously collaborative translation-annotation system for EFL reading comprehension. *Computer Assisted Language Learning, 24*(2), 155 - 180.

Chen, C.-M., & Chung, C.-J. (2008). Personalized mobile English vocabulary learning system based on item response theory and learning memory cycle. *Computers & Education, 51*(2), 624-645.

Choi, H. J., & Johnson, S. D. (2007). The effect of problem-based video instruction on learner satisfaction, comprehension and retention in college courses. *British Journal of Educational Technology, 38*(5), 885–895.

Choi, H. J., & Johnson S. D. (2005). The effect of context-based video instruction on learning and motivation in online courses. *The American Journal of Distance Education, 19*(4), 215–227.

Chu, H. C., Hwang, G. J., & Tsai, C. C. (2010). A knowledge engineering approach to developing Mindtools for context-aware ubiquitous learning. *Computers & Education*, 54(1), 289-297.

Chu, H. C., Hwang, G. J., Tsai, C. C., & Tseng, Judy C. R. (2010). A two-tier test approach to developing location-aware mobile learning systems for natural science courses. *Computers & Education, 55*(4), 1618-1627.

Chun, D. M., & Plass, J. L. (1997). Research on text comprehension in multimedia environments. *Language Learning & Technology, 1*(1), 60-81.

Dahbi, M. (2004). English and Arabic after 9/11. *The Modern Language Journal, 88*(4), 628-630.

Danan, M. (1992). Reversed subtitling and dual coding theory: New directions for foreign language instruction. *Language Learning, 42*(4), 497-527.

Danan, M. (2004). Captioning and subtitling: Undervalued language learning strategies. *Meta: Translator's Journal, 49*(1), 67-77.

Davis, F. D. (1986). *A technology acceptance model for empirically testing new end-user information systems: Theory and results* (Unpublished doctoral dissertation). MIT Sloan School of Management, Cambridge, MA.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319-340.

Felder, R. M., & Silverman, L. K. (1988). Learning and teaching styles in engineering education. *Journal of Engineering Education, 78*(7), 674-681.

Felder, R. M., & Soloman, B. A. (1991). *Index of learning styles.* Retrieved May 1, 2011, from http://www.ncsu.edu/felder-public/ILSpage.html

Fischer, B. B., & Fischer, L. (1979). Styles in teaching and learning. *Educational Leadership, 36*(4), 245-254.

Froehlich, J. (1988). German videos with German subtitles: A new approach to listening comprehension development. *Teaching German, 21*(2), 199-203.

Garza, T. J. (1991). Evaluating the use of captioned video materials in advanced foreign language learning. *Foreign Language Annals, 24*(3), 239-258.

Godwin-Jones, R. (2007). Emerging technologies digital video update: YouTube, flash, high-definition. *Language Learning & Technology, 11*(1), 16-21.

Gregorc, A. F. (1979). Learning/teaching styles: Their nature and effects. In J. W. Keefe (Ed.), *Student learning styles: Diagnosing and prescribing programs* (pp. 19-26). Reston, VA: National Association of Secondary School Principals.

Grgurovi, M., & Hegelheimer, V. (2007). Help options and multimedia listening: students' use of subtitles and the transcript. *Language learning & technology*, *11*(1), 45-66.

Grimmer, C. (1992). Supertext English language subtitles: A boon for English language learners. *Enrolled Agents Journal, 10*(1), 66-75.

Guillory, H. G. (1999). The effect of keyword captions to authentic French video on learner comprehension. *Computer Assisted Language Instruction Consortium Journal, 15*(1-3), 89-108.

Hashemi, M., & Ghasemi, B. (2011). Using mobile phones in language learning/teaching. *Procedia—Social and Behavioral Sciences, 15*, 2947-2951.

Honey, P., & Mumford, A. (1986). *The Manual of Learning Styles.* Maidenhead, England: Peter Honey Publications.

Hoven, D. (1999). A model for listening and viewing comprehension in multimedia environments. *Language Learning & Technology, 3*(1), 88-103.

Hsu, C.-K., & Chang, C.-K. (2010, Novermber). *Effects of automatic hidden caption classification on a content-based computer-assisted language learning system for foreign language listening.* Paper presented at the International Conference on Computers in Education, Putrajaya, Malaysia.

Hsu, J. (1994). *Computer assisted language learning (CALL): The effect of ESL students' use of interactional modifications on listening comprehension* (Unpublished doctoral dissertation). Department of Curriculum and Instruction, College of Education, Iowa State University, Ames, IA.

Hung, P. H., Hwang, G. J., Su, I. S., & Lin, I. H. (2012). A concept-map integrated dynamic assessment system for improving ecology observation competences in mobile learning activities. *Turkish Online Journal of Educational Technology, 11*(1), 10-19.

Hwang, G. J., & Chang, H. F. (2011). A formative assessment-based mobile learning approach to improving the learning attitudes and achievements of students. *Computers & Education, 56*(1), 1023-1031.

Hwang, G. J., Chu, H. C., Lin, Y. S., & Tsai, C. C. (2011). A knowledge acquisition approach to developing Mindtools for organizing and sharing differentiating knowledge in a ubiquitous learning environment. *Computers & Education, 57*(1), 1368-1377.

Hwang, G. J., Shi, Y. R., & Chu, H. C. (2011). A concept map approach to developing collaborative Mindtools for context-aware ubiquitous learning. *British Journal of Educational Technology, 42*(5), 778–789.

Hwang, G. J., Wu, C. H., Tseng, Judy C. R., & Huang, I (2011). Development of a ubiquitous learning platform based on a real-time help-seeking mechanism. *British Journal of Educational Technology*, 42(6), 992-1002. (SSCI)

Katchen, J. E., Fox, T., Lin, L., & Chun, V. (2001, November). *Developments in digital video.* Paper presented at the Third Pan-Asian Conference 2001, Kitakyushu, Japan.

Keefe, J. W. (Ed.). (1979b). *Student learning styles: Diagnosing and prescribing programs.* Reston, VA: National Association of Secondary School Principals.

Keefe, J. W. (1979a). Learning style: An overview. In J. W. Keefe (Ed.), *Student learning styles: Diagnosing and prescribing programs* (pp. 1-17). Reston, VA: National Association of Secondary School Principals.

Kim, D., & Kim, D.-J. (2012). Effect of screen size on multimedia vocabulary learning. *British Journal of Educational Technology, 43*(1), *62-70.* doi: 10.1111/j.1467-8535.2010.01145.x

Kolb, D. (1984). *Experiential learning: Experience as the source of learning and development.* Englewood Cliffs, NJ : Prentice-Hall.

Krashen, S. D. (1981). *Second language acquisition and second language learning.* New York, NY: Prentice Hall.

Krashen, S. D. (1985). *The input hypothesis: Issues and implications.* New York, NY: Laredo Publishing Company.

Kukulska-Hulme, A., & Shield, L. (2008). An overview of mobile assisted language learning: From content delivery to supported collaboration and interaction. *ReCALL, 20*(3), 271-289.

Liou, H. C. (2000). Assessing learner strategies using computers: New insights and limitations. *Computer Assisted Language Learning, 13*(1), 65-78.

Litzinger, T. A., Lee, S. H., & Wise, J. C. (2005). A study of the reliability and validity of the Felder-Soloman Index of Learning Styles©. *Education, 113*, 77-89.

Lu, M. (2008). Effectiveness of vocabulary learning via mobile phone. *Journal of Computer Assisted Learning, 24*(6), 515-525.

Luo, C. P. (2008). An action research plan for developing and implementing the students' listening comprehension skills. *English Language Teaching, 1*(1), 25-28

Mackey, T. P., & Ho, J. (2008). Exploring the relationships between web usability and students' perceived learning in web-based multimedia (WBMM) tutorials. *Computers & Education, 50*(1), 386–409.

Markham, P. L., & Peter, L. (2003). The influence of English language and Spanish language captions on foreign language listening/reading comprehension. *Journal of Educational Technology Systems, 31*(3), 331-341.

Markham, P., Peter, L., & McCarthy, T. (2001). The effects of native language vs. target language captions on foreign language students' DVD video comprehension. *Foreign Language Annals, 34(*5), 439–445.

Mayer, R. E., & Moreno, R. (2002). Animation as an aid to multimedia learning. *Educational Psychology Review, 14*(1), 87–99.

Myers, I. B., & McCaulley, M. H. (1985). *Manual: A guide to the development and use of the Myers-Briggs Type Indicator.* Palo Alto, CA: Consulting Psychologists Press.

Nah, K. C., White, P., & Sussex, R. (2008). The potential of using a mobile phone to access the Internet for learning EFL listening skills within a Korean context. *ReCALL, 20*(3), 331-347.

Norris, C., Hossain, A., & Soloway, E. (2011). Using smartphones as essential tools for learning: A call to place schools on the right side of the 21st century. *Educational Technology, 51*(3), 18-25.

Ogata, H., & Yano, Y. (2004, March). *Context-aware support for computer-supported ubiquitous learning.* Paper presented at the 2nd IEEE International Workshop on Wireless and Mobile Technologies in Education, JhongLi, Taiwan.

Ogata, H., Matsuka, Y., El-Bishouty, M. M., & Yano, Y. (2009). LORAMS: Linking physical objects and videos for capturing and sharing learning experiences towards ubiquitous learning. *International Journal of Mobile Learning and Organisation, 3*(4), 337-350.

Ozcelik, E., & Acarturk, C. (2011). Reducing the spatial distance between printed and online information sources by means of mobile technology enhances learning: Using 2D barcodes. *Computers & Education, 57*(3), 2077-2085.

Peng, H. Y., Chuang, P. Y., Hwang, G. J., Chu, H. C., Wu, T. T., & Huang, S. X. (2009). Ubiquitous performance-support system as Mindtool: A case study of instructional decision making and learning assistant. *Educational Technology & Society, 12*(1), 107-120.

Plass, J. L., Chun, D. M., Mayer, R. E., & Leutner, D. (1998). Supporting visual and verbal learning preferences in a second language multimedia learning environment. *Journal of Educational Psychology, 90*(1), 25-36.

Reichmann, S. W., & Grasha, A. F. (1974). A rational approach to developing and assessing the construct validity of a student learning style scale instrument. *Journal of Psychology, 87*, 213-223.

Rose, C. (2003). How to teach biology using the movie science of cloning people, resurrecting the dead, and combining flies and humans. *Public Understanding of Science, 12*(3), 289–296.

Schmeck, R. R. (1982). Teaching students through their individual learning styles: A research report. In J. W. Keefe (Ed.), *Student learning styles and brain behavior* (pp. 73-80). Reston, VA: National Association of Secondary School Principals.

Seo, K. (2002). The effect of visuals on listening comprehension: A study of Japanese learners' listening strategies. *International Journal of Listening, 16*(1), 57–81.

Stockwell, G. (2010). Using mobile phones for vocabulary activities: Examining the effect of the platform. *Language Learning & Technology, 14*(2), 95-110.

Taylor, G. (2005). Perceived processing strategies of students watching captioned video. *Foreign Language Annals, 38*(3), 422-427.

Thornton, P., & Houser, C. (2005). Using mobile phones in English education in Japan. *Journal of Computer Assisted Learning, 21*(3), 217-228.

van Scoter, J., Ellis, D., & Railsback, J. (2001). *Technology in early childhood education: Finding the balance*. Portland, OR: Northwest Educational Technology Consortium.

Vandergrift, L. (2007). Recent developments in second and foreign language listening comprehension research. *Language teaching*, *40*(03), 191-210.

Vanderplank, R. (1988). The value of teletext sub-titles in language learning. *English Language Teaching Journal, 42*(4), 272-281.

Venkatesh, V., & Davis, F. D. (1996). A model of the antecedents of perceived ease of use: Development and test. *Decision Sciences, 27*(3), 451–481.

Winke, P., Gass, S., & Sydorenko, T. (2010). The effects of captioning videos used for foreign language listening activities. *Language Learning & Technology*, *14*(1). 65-86.

Wong, L. H., Chin, C. K., Tan, C. L., & Liu, M. (2010). Students' personal and social meaning making in a Chinese idiom mobile learning environment. *Educational Technology & Society, 13*(4), 15-26.

Wright, J. L., & Shade, D. D. (Eds.) (1994). *Young children: Active learners in a technological age*. Washington, DC: National Association for the Education of Young Children.

Wu, P. H., Hwang, G. J., Tsai, C. C., Chen, Y. C., & Huang, Y. M. (2011). A pilot study on conducting mobile learning activities for clinical nursing courses based on the repertory grid approach. *Nurse Education Today*, *31(8),* e8-e15.

Yang, J. C., Huang, Y. T., Tsai, C. C., Chung, C. I., & Wu, Y. C. (2009). An automatic multimedia content summarization system for video recommendation. *Educational Technology & Society, 12*(1), 49–61.