

Tentative model of integrating authentic captioned video to facilitate ESL learning

WANG Yan-dong, SHEN Cai-fen

(Basic Courses Department, PLA University of Foreign Languages, Luoyang 471003, China)

Abstract: Guided by relevant learning theories and based on the analysis of the attributes of captioned video, the Chinese ESL learners' characteristics and the characteristics of ESL learning itself, the paper tries to propose a model of integrating captioned video in ESL acquisition process. The model focuses on the feasibility of using captioned video to facilitate or support ESL learning.

Key words: authentic captioned video; ESL; information processing theory; dual coding theory

1. Introduction

With the ever increasing popularity and availability of English language feature films, sitcoms, and documentaries in China, using captioned video to facilitate ESL learning becomes widely welcomed. Many Chinese English learners frequently express their enjoyment of watching movies in English, both for gaining exposure to the authentic foreign culture as well as for improvement of their English language skill in authentic cultural context. Therefore, it is necessary to conduct some researches in this area.

This paper is based on the hypothesis that integrating captioned video into the ESL acquisition process can greatly improve the performance of ESL learning. Previous research on learning and on language learning in particular has yielded many constructive fruits, which can serve as the theoretical basis for integrating captioned video in ESL learning. Based on the analysis of the attributes of captioned video, the Chinese ESL learners' characteristics and the characteristics of ESL learning itself, the paper tries to propose a model of integrating captioned video in ESL acquisition process.

The paper starts with a brief review of the previous researches on using captioned video to enhance learning and introduces the background for language learning nowadays. Following the background introduction, the attributes of captioned video are analyzed. With characteristics of ESL acquisition being described in the third part, the paper tries to list some theoretical basis to support the notion of using captioned video to facilitate ESL learning in the fourth part. In the fifth part of the paper, a model for integrating captioned video in ESL learning is proposed and the specific processes in the model are explained to show its reasonability and feasibility. In the concluding part, some caution is given for better integration of captioned video in ESL learning.

2. Background

2.1 Urgency of ESL learning

WANG Yan-dong (1984-), male, graduate student of Educational Technology, PLA University of Foreign Languages; research fields: CALL, educational technology and foreign language teaching.

SHEN Cai-fen (1957-), female, associate professor of Basic Courses Department, PLA University of Foreign Languages; research fields: educational technology and foreign language teaching.

With the wheels of human history rolling into the 21st century, the world has changed greatly. The world of today, characterized by enormous flows of various kinds of information, has been an age of “global village,” when communications between people have expanded way beyond their nationalities, culture or mother tongues. On the international scene as a whole, an increased openness and extended cooperation between nations are apparent signs of the times. All these trends have sparked great interests of second language learning all over the world. It seems that in this era, the mastery of a second language, especially ESL (English as a Second or Foreign Language), which is regarded as an international language in common use, is extraordinarily necessary. Based on this reality, a lot of researches, aiming at finding the most efficient and effective way of learning English, have been conducted. Among these studies, exploring the possible benefits from using video for ESL acquisition has been a widely explored subject. Although as a new subject in China as it is, the apparent benefits have been displayed in numbers of cases. Unfortunately, however, little attention has been given to the overall effects of captioned video on ESL acquisition and no models that can be used to integrate captioned video to facilitate ESL learning. Therefore, it is interesting and necessary to examine what captioned video can contribute to ESL acquisition and propose models useful.

2.2 Review of researches on using captioned video to facilitate SL learning

In some western countries, captioned video are not only seen as a charity contribution to help the deaf or the hearing impaired to understand information in videos, but also examined, exploited, and used by lots of educators and language teachers as important instructional and teaching resources.

The great popularity of captioned videos in L2 (second language) among second language learners is a strong impetus for researchers trying to determine what benefits they can offer to the learner of a foreign or second language. The history of research in this area is quite long, or rather, talking and writing about captioned video’s educational value is almost as old as the medium itself.

With the advent of captioned video in 1970s, many studies on using captioned video to enhance language learning have been conducted. Early in 1980s, there had been researches on this area, such as using closed caption TV to teach reading to adults, especially immigrants in America. A lot of researchers in this area have involved in investigating what priorities can be provide by captioned video and how to use them to enhance language learning. Researchers such as Price, K. (1983), Vanderplank, R. (1988), Koskinen, P. & Neuman, S. (1992), Markham, P., Peter, L. & McCarthy, T. (2001), Anne R. Friedman (2001) have contributed a lot to the literature of the area. However, little attention has been paid to this area by Chinese researchers. Up to now, some Taiwan and Hong Kong scholars have conducted a few researches, which is still far from enough. Moreover, it is hardly to find any research conducted from the instructional technology perspective, which is an incentive for me to have a try from the perspective of instructional technology.

3. Attributes of Authentic Captioned Video as a Learning Medium

When any of the media is selected as an educational or instructional medium, it has its own superior features to outstand itself. When it comes to authentic captioned video, there is no exception. Captioned video not only shares the common attributes with non-captioned video materials when it is viewed as an educational means, but also possesses some particular attributes which enable it to be an ideal language learning (especially) and teaching medium.

3.1 An introduction to captioned video

To integrate captioned video into ESL learning process, we should at first know what captioned video is. Captions are textual versions of the spoken words such as dialogue, monologue, narration, as well as other unspoken information on the screen. They usually appear in two or three lines at the bottom of the screen. Captions are usually seen in TV programs such as news, documentaries and weather reporting, as well as in other video media formats such as VCDs or DVDs. Captions are also used when intelligibility is reduced by poor voice quality, dialecticisms, colloquialism or other features of speech. What's more, commercials make extensive use of captioning, sometimes with calligraphic expression. Self-evidently, when a video is added with captions (either English or Chinese captions) it becomes captioned video.

Captioned video was originally invented in the United States to allow both audio and video information to be perceivable to those who do not have access to audio and to be understandable to a wider audience. The history of captioned video can be traced back to the 1970s. It was developed on the basis of captioned TV. In the year of 1971 and 1972 respectively, the ABC (American Broadcast Company) successfully held two exhibitions of captioned TV in succession, which symbolized the primary success in the technology of captioned TV. It was originally intended for hearing lost or hearing impaired audience. With its apparent characteristics of charity, it gained popularity as well as the support of US federal government. Later on, the National Captioning Committee, aiming at spreading and popularizing of TV captions, was established in 1979. The day of March, 16th 1981 was regarded as "the day of breaking silence" by the American media, because for the first time captioned TV series were broadcast on the air. From then on, the deaf or hearing impaired audience can appreciate TV programs through captions. During the 1990s, the development of the captioning technology bound a great leap. As of July 1993, all TV sets (13 inches or larger) sold in the United States by law must have captions decoder circuitry built in.

While captioned video allows millions of deaf and hard-of-hearing citizens' access to television, it also serves as a useful teaching aid in second language learning, child or adult acquisition of literacy, and in most types of educational programming. With the development of captioning technology, all kinds of captions are available at the convenience of video audience's choice. These captions include L1 (first language) captions (sometimes called subtitles), L2 captions, closed captions¹, full-text captions, key-word captions, reversed captions² and etc.

3.2 Attributes of captioned video

Captioned video and non-captioned video share a lot attributes possessed by all video types. Video refers to the electronic storage of moving images together with soundtrack (videotapes, videodisc, DVD, etc.) (Robert Heinich & Michael Molenda & James D. Russell & Sharon E. Smaldino, 2002). Generally speaking, there are four types of video formats: videotapes or videocassette, DVD, videodisc, and Internet streaming video. Actually, it is too hard to list all the video version formats and all its features here. However, when used as an instruction medium, the common attributes possessed by the authentic foreign-language videos (both captioned and non-captioned) can be listed as followings:

- (1) They tend to be cheaper and more efficient to transmit information than the traditional ways of transmitting information through books or people.
- (2) They can record the live scene or the authentic scene of lectures, seminars, conference, which can serve to the time-shift or space-shift lesson schedule. It can be very convenient for learners who have time or distance

¹ Closed captions refer to the kind of captions that can be turned on or off at the user's will.

² Reversed captions consist of the original video lines (or L2 lines) displayed on the screen but meanwhile the videos are dubbed with L1 audio.

limitation without lessen the authenticity of the live scene.

(3) It has the ability to manipulate temporal and spatial perspectives. Video can compress the time it takes to observe an event, such as American daily life, through a technique known as time lapse. Time can be also expanded in motion media through a technique called slow motion for students to see the events that occur too fast to be observed by naked eyes. Moreover, video and live television allow us and our students to observe two events occurring simultaneously but many miles apart using split screen. In this way special effects can be obtained during presentation.

(4) Motion and animation. Moving images have an obvious advantage over still visuals in portraying concepts in which motion is essential to understanding. With this attribute, we can take advantage of persistence of vision to give motion to otherwise inanimate picture to illustrate how things work out, such as how pronunciation occurs.

(5) Readily available and simple to use. Nowadays, foreign language movies and other videos (especially English language videos) have become widely spread in China and are available on almost any topic and for all types of learner in all the domains of instruction—cognitive, affective, psychomotor, and interpersonal. Meanwhile, it is easy to control. It can be played at various speeds in both forward and reverse. What's more, usually foreign videos issued in China have two audio tracks, for example two different languages can be added to aim at two different levels of difficulty or for foreign language learners.

(6) Stimulating and engrossing. It can presents information in a more vivid way than only text can. Studies have shown that audience usually is attentive to motion images projected onto a large screen. Video has the power to hold the learners' spellbound as a human drama unfold before their eyes.

(7) Ideal material to learn foreign languages. It is an excellent media for learning foreign languages. It allows the ESL learner to hear word or expressions pronounced by native speaker and to experience the authentic context or environment in which these idiomatic expressions are used. Moreover, cultural understanding can be developed through viewing video depicting the country of the native speakers. It can take students to places they might not be able to go otherwise, for example to a foreign country, which is very beneficial to foreign language learners.

(8) Individualized learning. Because of the ease of operating, video lends itself to individual study. For example, the foreign language learner learns the language.

(9) Oral practice. An important language skill is speaking in ESL learning. It needs as much practice as possible to develop this skill. Students can use excerpts from video as a part of oral reports, turning the sound off and using their own narration. Furthermore, role playing is another convenience provided by captioned video.

(10) Documenting actual events and bringing them into classroom, such as a pageant where an interpreter is needed. By viewing the live scene the foreign language learner can develop their own skill of interpreting. Moreover, if one has no access to a high-degree academic conference, he or she can experience the live situation by viewing a video.

Aside from the aforementioned ten attributes shared by captioned authentic video and non-captioned video, some special attributes lends captioned video being an excellent medium for ESL learning.

Firstly, captions of captioned video possess the following three characteristics:

A. Synchronized—the text content usually appears at approximately the same time that audio would be available.

B. Equivalent—content provided in captions usually is equivalent to that of the spoken word.

C. Accessible—caption content usually is readily accessible and available to those who need it.

Secondly, captioned video combines visual, auditory, and lexical stimuli. Therefore, it is motivating, entertaining, and educational. Thirdly, captioned video provide English learners with real language and cultural information with accuracy by adding the more reliable textual information. Fourthly, captioned video provides a basis for producing and practicing language for learners, as well as developing their language skills, for it provides verbal message for learners. Thus, language learners not only have access to auditory language experiences, but also to visual and text information.

Moreover, these characteristics have been proven to be beneficial to language learning. A number of closed captioned video studies have revealed that viewing L2 videos with L2 captions added to them method that (a) compels students to do more language-specific processing; (b) improves listening comprehension (Price, 1983, pp. 7-8; Vanderplank, 1988, pp. 221-234); (c) enhances vocabulary learning (Danan, 1995, pp. 253-282; Neuman & Koskinen, 1992, pp. 94-106); and (d) ameliorates verbatim recall and memory of content (Vanderplank, 1990, pp. 221-234). Last but not least, there is evidence that students who are exposed to L2 subtitled video relate positively to it and are confident that subtitles help them learn (Chung, 1996, pp. 321-330). Furthermore, according to some authoritative surveys, students commented that the subtitles enhanced their ability to notice, comprehend, spell, and recall new L2 material (Iva Baltova, 1999, pp. 26-38). In a word, these exclusive characteristics of captioned video make it of great value for second language learners.

4. Characteristics of ESL Learning

4.1 Characteristics of ESL learning in general

To understand the general characteristics of second language acquisition is important to understand the Chinese ESL learning in particular. There are some general characteristics of second language acquisition being listed as followings.

Firstly, the general communicative objectives of the second language learning require the teaching of L2 comprehension and production in a meaningful and realistic context, and attaining its general experiential objective entails the development of knowledge through various “real life” language experiences (Tremblay, 1990, pp. 18-20). As demonstrated by a body of recent research, L2 acquisition is enhanced in environments where language is the medium of communication rather than merely the object of analysis (Larsen-Freeman & Long, 1991, p. 258).

Secondly, according to Rod Ellis (2001), two main factors affecting the second language acquisition are external and internal factors. The external factors are the social milieu (or social context) and the input. The former refers to the influence of the social conditions on the opportunity of exposure to the authentic and original L1 language context. The latter refers to the samples of language to which a learner is exposed. Language cannot occur without some input. So what type of input facilitates learning? It is found that learners benefit less from input that has been simplified from the original authentic L2 language or deprived of the original authentic context. That is to say we need the indigenous language that is used in the authentic context to facilitate SLA. From this perspective of view, using video, especially captioned video is an ideal choice for SLA. Internal factors, on the other hand, are cognitive mechanism and the extent to which L1 is drawn on when L2 is learnt. Thus, SLA can be much easier and more efficient if the L1 and L2 are integrated into the same video clip. Two obvious examples of this are the model of native-language captions and the model of reversed captions.

Thirdly, accidental learning is more often to happen in second language acquisition.

It is not possible to learn a complete language intentionally—there is too much to learn.

Intentional learning, which is more likely to result in explicit than implicit knowledge and therefore, will not contribute to the kind of proficiency needed to use language easily and fluently.

In other words, incidental learning plays a very important role in achieving L2 proficiency. What is opportune is that captioned video can bridge the gap between intentional learning and accidental learning by tying the two ways of learning together through captions. On the one hand, visual clues and soundtrack can create an authentic culture and language environment in which accidental learning can easily take place. On the other hand, captions can be used as a resource for intentional learning, which can compensate the drawbacks such as the missing of some information, of the accidental learning.

In light of these characteristics, using captioned video in EFL classroom will compensate the drawbacks of the ordinary textbook-based language teaching.

4.2. Characteristics of ESL learning in china

4.2.1 The characteristics of English Language itself

(1) Alphabetic language. Unlike Chinese characters, English is a kind of alphabetic language, thus, English words usually don't look like the thing they stand for. Usually, they are remembered by its echoic characteristic. In other words, one needs to learn it by reading or listening.

(2) Cultural background. According to Krashen's (1985) input hypothesis "Human acquire language in only one way—by understanding messages, or by receiving 'comprehensible input'..." In most of the cases, learner always failed to understand the materials provided for them due to their unfamiliarity with the English culture, let alone to acquire the language.

(3) Enormous vocabulary and idiomatic expression. For most Chinese ESL learners, a main impede is their small vocabulary and lack of idiomatic expression. They always encounter the situation in which they just cannot express themselves clearly because they do not know how to say.

4.2.2 The characteristics of Chinese ESL learners

(1) Most of the ESL learners have passed the critical period for language learning. According to Chomsky, human are innately equipped for language. Two decades ago, some linguistics has proposed a Critical Period Hypothesis theory for language learning. The theory points out that human possess a special capacity for language learning. This period usually ranges from the age of three to the age of seven and after the age of seven, the older the age on arrival, the lower performance in language learning.

(2) Influenced by first language. For most of Chinese ESL learners, Chinese is their mother tongue. When they learn English, they hardly can refrain the interference of their mother language. Therefore, although after many years' English learning and training, still a lot of learners cannot make themselves understood to native English speakers due to their excessive reliability on mother tongue. To find a way to avoid Chinglish is very urgent.

(3) Lack of interest. A lot of ESL learners take English learning as a tedious work of memorizing words, grammatical rule, and reciting texts. They are passive learners and language learning is just an endless work to them. They often complain that they seem to be stuck in a situation in which they see no progress in their English learning no matter how hard they have tried.

(4) Test-oriented trainees with low language competence. In China, ESL is a compulsory course and nearly all the students need to pass the national English degree test such as CET-4 or CET-6 to get their bachelor or master degree. This results in a phenomenon known as "high marks with low competence", which is fatal to

language. Language is learned to communicate not to deal with tests.

4.2.3 The characteristics of environment

To learn English well, an ideal way is to immerse in the native environment in which English is spoken. However, it is improbable for all of the Chinese ESL learners to go abroad to learn that language. Being in an environment, with people around the learner speaking Chinese only, it is really hard to learn authentic and real English. Moreover, a lot of Chinese ESL learners have no chance to communicate in English. On the one hand, they have no partners in such an environment. On the other hand, learners complain they have nothing to talk about because they always forget most of things they have learned in class. Just as Hymes (1972) has put it: “the key to understanding language in context is to start not with language but with context.” Without ideal context to understand and practice language, the speed of language acquisition and language skill development will be greatly slowed down.

5. Theoretical Basis of Using Captioned Video in ESL Learning

5.1 Krashen’s second language acquisition theory

According to American linguist S.D. Krashen’s (quoted in NI Xiu-ying, 2005) second language acquisition theory, human acquire language only through understanding messages or through receiving “comprehensible input”. Krashen believes “adults have two distinctive ways of developing competences in second language...acquisition, that is, by using language for real communication...learning...‘knowing about’ language.” In his Input Hypothesis, which consists of five linked “hypotheses”, input, acquisition/learning, monitor, natural order, and effective filter, Krashen proposes that language is acquired through comprehensible language input. Thus, language learners should be provided with optimal input to facilitate their language acquisition. According to him, it needs four conditions to achieve the optimal input: comprehensible, interesting and relevant, not form-focused, and quantitative. All these four aforementioned conditions can be found in captioned authentic video.

5.2 Theory of how learning occurs

Learning is an old topic and is not simple to define concisely. However, there are some definitions as followings:

“Learning is a relatively permanent change in behavior due to experience. Learning is a relatively permanent change in mental representations or associations as a result of experience” (Ormrod, J. E., 2004).

“Learning refers to lasting changes in the learner’s knowledge, where such changes due to experience” (Mayer,R.E.,2003).

Learning is “a persistence change in performance or performance potential that results from experience and interaction with the world” (Driscoll,2005).

From the definition of learning cited above, we can see that experience plays a key role. This indicates that the quality of the experience potentially enhances what is actually learned, and language learning is no exception, or rather needs more emphasis on immersion of the experience—in language context. Captioned authentic video as one of educational technology can enhance learning experiences through the meaningful integration of it into ESL learning.

5.3 Cognitive theory of multimedia learning

Mayer’ research (2003) on Cognitive of Multimedia Learning indicates, for example, that individuals have a limited capacity to attend to and process incoming information through a single sense modality (e.g., auditory).

That is, only a certain amount of information can be heard, understood, and processed by a learner at any time. More than the restricted amount will not be retained because a form of “cognitive overload” (in some books also called as working memory) may occur. However, if a second channel (e.g., visual) is utilized to convey additional information, then the limited capacity for processing information by learners can be expanded to some degree.

One way to ease the “over-load of working memory” is to activate prior knowledge and mental automation. The importance of prior knowledge is concisely summarized by David Ausubel (epigraph): “The most important single factor influencing learning is what the learner already knows. Ascertain this and teach him accordingly” (quoted in Haido Samars et al., 2006). Therefore, in language learning, new information should be connected in a meaningful way to information previously learned. Integrating captioned video will help learners activate existing schemata making the information provided to them more meaningful and reducing cognitive load.

Through captioned video, learners are actively engaged in the learning process; learners create language knowledge by making connections with previously learned material, for learning environments created by captioned video allow and encourage students to make these connections.

5.4 Information processing theory

One of the more influential cognitive theories used in instructional design is the Information Processing Theory. It focuses on how the human memory system acquires, transforms, compacts, elaborates, encodes, retrieves, and uses information. Information Processing Theory approaches learning primarily through the study of memory. The memory system is divided into three main storage structures: sensory registers, Short-Term Memory (STM), and Long-Term Memory (LTM).

The first stage of processing is registering stimuli in the sensory memory system. Individuals receive information from the environment through a sensory receptor; ears, eyes, nose, mouth or sense of touch. The information is a sensation as it enters the body. The sensory registers (one for each sense) briefly hold raw information until the stimulus pattern is recognized or lost. The sensation that is not discarded is transformed into a message and sent to working memory. The message is sent when the individuals are paying attention to a stimulus or if the message activates a known pattern. STM holds information in its recognized form instead of its raw sensory form (e.g., visual—icon, auditory—echo, textual—alphabet). It is critical that the learner concentrates to the information at this stage and further stages in order for information to move on to the next. An important characteristic of STM, for our purpose, is that both visual information and phonemic coding are the important modalities of learning, especially learning a language.

Information that we do remember for a short amount of time goes into long-term memory. Long-term memory has an unlimited capacity to store knowledge. This is the most essential part of the Information Processing Theory. Elaboration and distributed practice are the two processes that move information into long-term memory. Thus, by using captioned video during instruction can help the students to encode information.

5.5 Dual coding theory

Paivio (1971, 1986) developed the dual coding theory, which stated that the two types of information (verbal and imaginal) are encoded by separate subsystems. One specialized for sensory images and the other specialized for verbal language. He distinguished two separate representational systems: the verbal system and the imagery system composed of nonverbal objects and events. According to Paivio’s model, the two systems are functionally independent, yet representation in one system can also activate those in the other because dually code items (coded verbally and nonverbally) are linked by referential connections. Visual traces are remembered better than verbal components and also have an additive effect when items or subjects are encoded dually. Consequently,

according to Paivio (1986: chapter 11), in the case of second or foreign language acquisition, as more foreign words and expressions are learned in direct association with appropriate nonverbal referents (objects, but also experiential elements such as events and emotions), the richer and more meaningful are the referential interconnections, thus resulting in better language recall and appropriate use.

In captioned video, motion pictures, verbal captions, and spoken words are presented to the audience simultaneously. In this way, a word or a concept represented as an image can also find the verbal label in video captions or soundtrack, or vice versa. Moreover, Paivio (1986) argued that pictures were often remembered better than verbal information. Those remembered images can be served as prior knowledge of later more advanced language learning. Naturally, if appropriately arranged and implemented, captioned video can facilitate ESL learning.

6. A Tentative Model of Integrating Captioned Video in ESL Learning

Based on the analysis above, a model of integrating captioned authentic video in ESL learning can be proposed. The model consists of four sub-models, which are divided deliberately to elaborate this tentative model. The model is based on several hypotheses.

Firstly, ESL learning through captioned authentic video is a cognitive process. This process is composed of three steps according to Vanpatten (1996: quoted in Peter Skehan, 1998). In his model, input comes first, intake second and developing language structure finally. On the basis of Vanpatten and supplemented with Krashen's second language acquisition theory, there comes the first hypothesis in the paper. A sub-model is proposed in line with this hypothesis, as shown by Figure 1. In the sub-model, captioned authentic video is the source of language in-put. Based on its special attributes of authentic captioned video, it can be viewed as an optimal ESL language in-put. Language acquisition requires meaningful interaction in the target language (Krashen, 1982) and language out-put is initiated by language in-put. Although authentic or original English videos can provide a real communication environment, yet this might be too difficult to be understandable to the learners if they only rely on their listening comprehension, let alone acquire anything. To facilitate this in-put, textual captions can help a lot. Moreover, Authentic or original captioned can provide multiple-channel information, which can ease the "load" of learner's working memory. In this way, learners can receive adequate input and develop their sense of English language and meanwhile, prepare them for the right output. English video, affordable and available to nearly all ESL learners, is the right choice.

Secondly, a key factor that influences ESL learning is memory. According to Cognitivists' Information Processing Theory, learning something must follow three stages of memory: the sensory memory, the STM, and the LTM. Based on this theory, the second sub-model is proposed, as shown by Figure 2. In this model, captioned video can meet the required conditions by the optimal memory. During language learning, not all input has equal value and only input that is noticed then becomes available for intake and effective processing (Schmidt, 1990: quoted in Peter Skehan, 1998, p.40). Noticing or attention, in other words, serves as a necessary but not a sufficient condition for effective processing to take place. Since video itself can catch and hold learners' interest, it is easy to be noticed by learners. Moreover, captions in video channel learner audience's attention and brings into awareness what otherwise would have been missed (in the case of taking listening as a primary learning task; it is vice versa when taking reading as a primary learning task). When the input language is stored in long-term memory, it is divided into two parts and stored in the form of semantic memory and episodic memory separately. Since video

material is usually episode-oriented, the content of the video can be easily remembered in LTM. This can serve to activate the content of semantic memory and ease the “load” of working memory. This satisfies the principle of “in language learning, new information should be connected in a meaningful way to information previously learned.” Moreover, authentic captioned video can be used to train or activate language skills such as fast reading, listening comprehension, writing inspiration, speaking, translating and interpreting.

Thirdly, language is constructed by learners through interaction with authentic scenes in which that language is used. This hypothesis is based on the Constructivists’ view. The third sub-model (as shown in Figure 3) is trying to elaborate the importance of authentic or real context in which language is used. This distinguishes authentic English video from other video. Moreover, in this model, soundtrack is depicted as echoic information and motion images as iconic information. Emphasizing on this distinction is significant to ESL learning because unlike Chinese characters, English is an echoic language. Finally, when captioned video is integrated into ESL learning process, it mingles with the process and forms a circulating cycle.

Based on these three hypotheses and sub-models elaborated above, a model of integrating captioned video into ESL learning is tentatively built up (as shown by Figure 4). In this model, three channels of language information is presented to learners by the authentic captioned video, auditory information serving as auditory stimuli, motion images serving as contextual clue, captions serving as noticing and textual assistance. Auditory stimuli can be used to train listening comprehension skill, synchronized interpretation skill, and pronunciation and intonation rectification. Motion images can be used to train imitation and recitation skill. Captions provide learners with rich vocabulary, idiomatic expression with accuracy. Captions can also be used as interpretation training and listening checking. In a word, authentic captioned video provides an immersive environment for language learning, which can facilitate ESL acquisition and competence in every sense.

7. Conclusion

Using any media to help improve learning should take into account a broader range of factors such as the attributes of the media, the specific learning tasks, and the learner characteristics. Thus, to integrate captioned video in ESL learning needs to talk about the attributes of captioned video and the characteristics of ESL learner in China. With the guide provided by the theory in educational technology and second language acquisition research, to integrate captioned video to facilitate ESL learning is possible to be realized. Therefore, the model proposed in this paper can be used in practice. However, just as the AECT 2004 draft definition of Instructional Technology has pointed out: technology and media used for educational or instructional purpose should be “appropriate”. That is to say, the selection of methods and media should be made on the basis of “best practices” applicable to a given situation, denoting suitability for and compatibility with their intended purposes³. Therefore, captioned authentic video used to facilitate ESL learning should be and must be task-appropriate and content-appropriate. Of course, there is still a long way to go to make the model more mature and feasible to practice.

References:

- Anne R. Friedman. *Using closed captioned video in the ESL reading class*. Retrieved April 25, 2006, from <http://www.umkc.edu/cad/nade/nadedocs/96conpap/afcpap96.htm>.
- Chung, J. M. 1999. The effects of using video texts supported with advance organizers and captions on Chinese college students’ listening comprehension: An empirical study. *Foreign Language Annals*, 32(3), 295-308.

³ AECT definition and terminology committee, *The Meanings of Educational Technology*, June 1, 2004.

- D.-Li.n. 1999. Processing a spoken text in a subtitled video program. In: *The proceedings of the English international symposium in English teaching*. Taipei: The Crane Publishing, 321-330.
- DAI Jin. 2005. Captioned video and foreign language learning. *CAFLE*, (103).
- Danan, M. 1995. Reversed subtitling and dual coding theory: New directions for foreign language instruction. In: B. Harley (Ed.), *Lexical issues in language learning*. Ann Arbor, MI: Language Learning, John Benjamins, 253-282.
- David M. Moore, John K. Burton & Robert J. Myers. 2004. Multiple-channel communication: The theoretical and research foundations of multimedia. In: David H. Jonassen. (Ed.), *Handbook of research on educational communications and technology* (2nd ed.). New Jersey: Lawrence Erlbaum Associates, 979-1005.
- Driscoll, M. P. 2005. *Psychology of learning for instruction*. (3rd ed.). Boston: Allyn & Bacon.
- Haido Samars, Thanasis Giouvanakis, Despina Bousiou & Konstantions Tarabanis. 2006. Toward a new generation of multimedia learning research. *AACE Journal*, 14(1), 3-30.
- HE Gao-da. 2002. *Modern educational technology and modern foreign language teaching*. Guilin: Guangxi Normal University Press.
- Helen Gant Guillory. 2001. The effects of keyword captions to authentic French video on learner comprehension. *CALICO Journal*, 15, 1-3.
- Hymes, D. 1972. On communicative competence. In: J.B. Pride & J. Holmes (Eds.). *Sociolinguistics*. Harmondsworth, England: Penguin Books.
- Iva Baltova. 1999. Multisensory language teaching in a multidimensional curriculum: The use of authentic bimodal video in core French. *Canadian Modern Language Review*, 56(1).
- Jonas Borell. 2000. SUBTITLING OR DUBBING? An investigation of the effects from reading subtitles on understanding audiovisual material. *Cognitive Science*, 61-80.
- Koskinen, P. S., Wilson, R. M., Gambrell, L. B. & Neuman, S. B. 1993. Captioned video and vocabulary learning: An innovative practice in literacy instruction. *The Reading Teacher*, 47(1), 36-43.
- Krashen, S. 1985. *The input hypothesis*. New York: Longman Group.
- Larsen-Freeman, D. & Long, M. H. 1991. *An introduction to second language acquisition research*. London: Longman.
- Markham, P. L., L. A. & McCarthy T. J. 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. 2003. *Learning and instruction*. Upper Saddle River, NJ: Merrill/Prentice Hall
- Neuman, S. B. & Koskinen, P. 1992. Captioned television as comprehensible input: Effects of incidental word learning from context for language minority students. *Reading Research Quarterly*, 27, 94-106.
- NI Xiu-ying. 2005. Krashen's second language acquisition theory and the teaching of spoken English. *US-China Foreign Language*, 3(4).
- Ormrod, J. E. 2004. *Human learning*. Upper Saddle River, NJ: Merrill/Prentice Hall.
- Paivo, A. 1971. *Imagery and Verbal Processes*. New York: Holt, Rinhart & Winston.
- Paivio, A, Clark, J. & Lambert, W. 1988. Bilingual dual-coding theory and semantics repetition effects on recall. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 14, 163-172.
- Paivio, A. 1986. *Mental representation: A dual-coding approach*. New York: Oxford University Press.
- Peter Skehan. 1998. *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Price, K. 1983. Closed-captioned TV: An untapped resource. *MATSOL Newsletter*, 12, 7-8.
- Robert Heinich, Michael Molenda, James D. Russell & Sharon E. Smaldino. 2002. *Instructional media and technologies for learning* (7th ed.). Pearson Education.
- Rod Ellis. 1997. *Second language acquisition*. Oxford Introductions to Language Study.
- Rod Ellis. 2001. *Second language acquisition: The role of consciousness*. Lecture given at National Chengchi University, Taiwan.
- SHU Ding-fang. 2004. *Foreign language teaching reform: Problems and solutions*. Shanghai: Shanghai Foreign Language Education Press.
- Timothy J. Newby, Donald A. Stepich, James D. Lehman & James D. Russell. 2006. *Educational technology for teaching and learning* (3rd ed.). New Jersey: Pearson Prentice Hall.
- Tremblay, R., Duplantie, M., & Huot, D. 1990. National core French study: The communicative/experiential syllabus. Winnipeg: CASLT.
- Vanderplank, R. 1988. The value of teletext sub-titles in language learning. *ELT Journal*, 42(4), 272-281.
- Vanderplank, R. 1990. Paying attention to the words: Practical and theoretical problems in watching television programs with uni-lingual (CEEFAQ) sub-titles. *System*, 18, 221-234.

WANG Hui. 2005. The effects of captions on Chinese EFL students' incidental vocabulary acquisition. *CAFLE*, 102.
 ZHANG Dong-zheng. 1999. *Foreign language teaching pedagogy*. Scientific Press.

(Edited by Robert, Doris and Stella)

Appendix A: Model Graphics

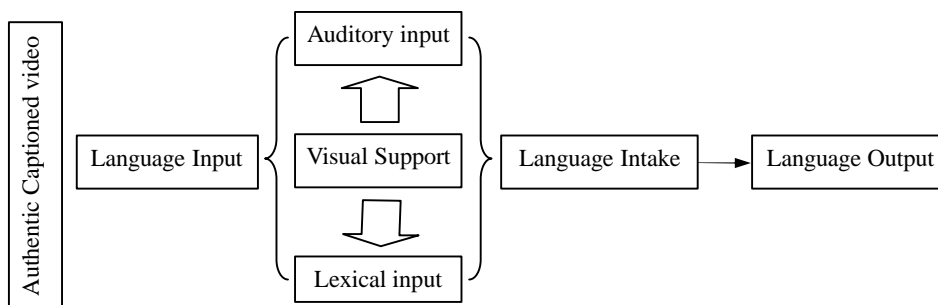


Fig. 1 A cognitive process of language learning through authentic captioned video

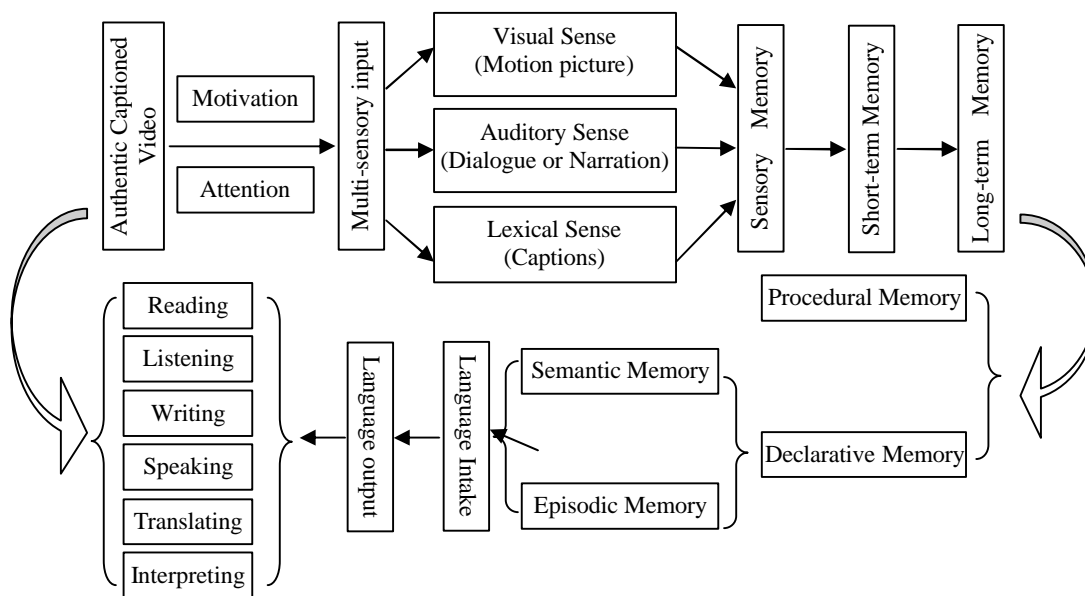


Fig. 2 Information processing during language learning through captioned video

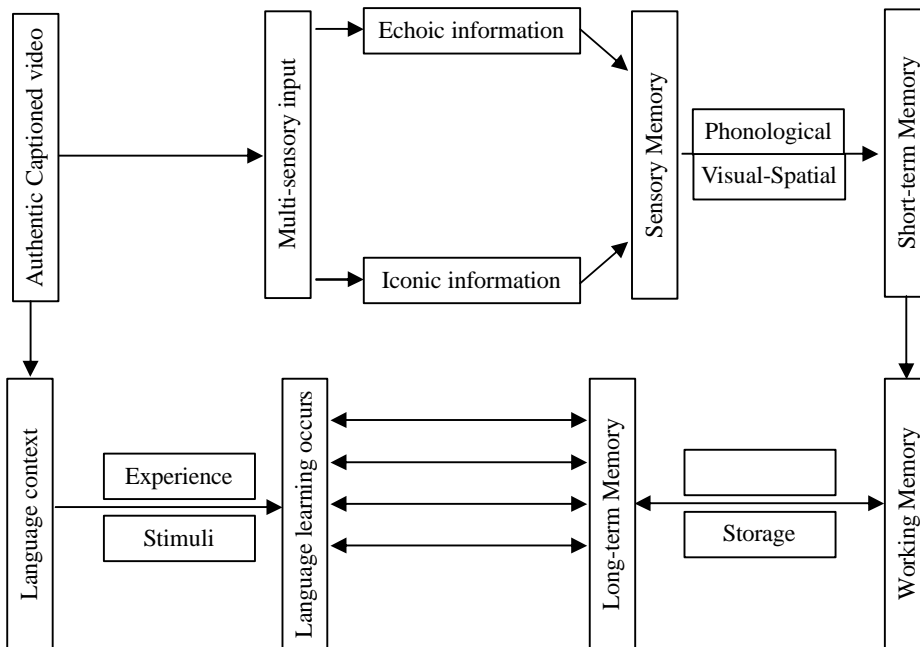


Fig. 3 Learning language in an interactive way through captioned video

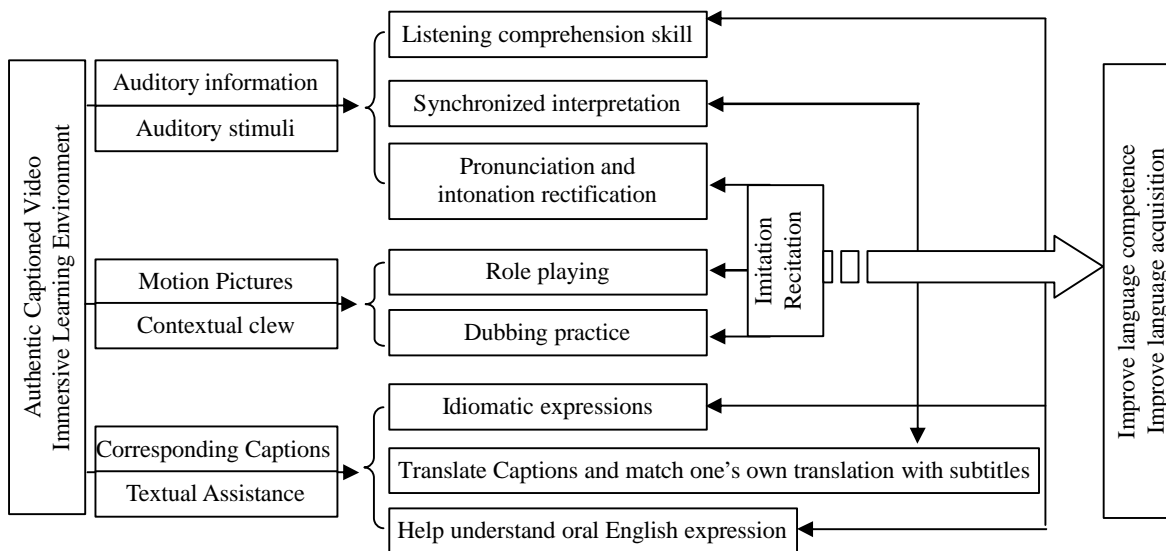


Fig. 4 A tentative model of integrating captioned video into ESL learning