Supporting Online Video-based Correction for Language Learning through Markup-based Video Editing

Yoshiaki Hada, Hiroaki Ogata, Yoneo Yano

Department of Information Science and Intelligent Systems, Tokushima University
2-1, Minami-Josanjima, Tokushima 770-8506, Japan
E-mail: {hada, ogata, yano}@is.tokushima-u.ac.jp
URL: http://www-yano.is.tokushima-u.ac.jp/

Abstract: This paper focuses on online video based correction for language learning. The environment supports between a native speaker as a teacher and a non-native learner using videoconference system. We extend a videoconference system so that the system can record the video based conversation as a learning scene. If teacher edits the video to explain, it is very useful as teaching materials. However, ordinal video editing system creates quite new video file. With this system, learner can hardly understand teacher's task. Therefore, this paper proposes video correction and the online video-editing model. Our prototype system called Viclle (Video-based Communicative Language Learning System) supports both teacher (native speaker) and learner. Using this system, both teacher and learner can communicate by e-mail without sending video files.

Keywords: XML, Language learning, Video-based correction, Videoconference, multimedia.

1. Introduction

With the fast development of the Internet, people around the world have more chances to communicate directly with CMC (Computer Mediated Communication) tools, such as e-mail, chat and bulletin board system. By applying CMC tool to language learning, it is possible for learners to communicate with native speakers in foreign languages. In fact, many approaches, which are intended to improve exchange among different cultures and language learning with CMC tool, have been proposed. Hanson et al. (1998) used CMC tool for improving international collaboration and comprehension among university learners around the world. Through the experiment, learner's ability to comprehend and express themselves in foreign languages was highly improved. They also pointed out that to communicate directly with native speaker was very important for foreign language learning. The approach that emphasizes the foreign language study through communication is known as "communicative approach" (Johnson & Morrow, 1981). This approach is a foreign language learning method, which emphasizes communication primarily. Grammatical rules and vocabularies are of secondary importance. Recently, it has attracted much interest in CALL (Computer Assisted Language Learning).

From the technical point of view, the high-speed and broadband network has realized rich, smooth and synchronous communication and collaboration using not only text but also audio and video (LeeTiernan et al, 2001). For example, in a school of English conversation in Japan, Japanese learners can master English with distant teacher (native speaker) using videoconference system. The interactive video makes learning more effective (Dalton, 1990). What seems to be lacking, however, is to reuse the conversational video for language learning by reviewing and editing video of the past lesson.

This paper proposes online video based correction system for language learning, which allows teacher and learner to revise and edit their conversational video. First, the videoconference system in the system records their conversation into video file. If teacher edits video using video-editing tool, it may very difficult for learner to understand where teacher has corrected in original video because the ordinal tools creates the quite new video. Therefore, we propose markup-based video-editing system called Viclle (Video-based Communicative Language Learning Support System), which provides the similar environment to the real one where we correct text using red-pen and paper. Using this system, teacher and learner can correct their conversational video, e.g., user can insert new video segment and annotations with correction mark. That is easy to understand teacher's task for correction. The features of Viclle are as follows:

(1) Teacher and Learner can correct conversational video.
(2) Learner can learn by interacting with this system based on correction marks.
(3) Teacher relates learner's errors in original video with correction marks. The marks show what learner has made mistakes.

(4) The system use VCML (Video-based Correction Markup Language) to describe the contents of video correction. The language is based on XML (eXtensible Markup Language).

(5) It generates revised video if VCML tags are applied to original video.

(6) Teachers and learners share video as teaching materials.

This paper is organized as the following. First, Video Correction is proposed in the next section. After that, the paper describes playing of revised and the implementation of Viclle. Finally, concluding remarks are given in section 5.

2.Related Work

Recently, a wide range of authoring systems has been developed, which is allowing both text and media annotations. Multimedia annotation environment for young children using Java has been developed (Bouras, et al., 2000). It is used to provide children with the opportunity to reflect upon and annotate episodes from their everyday life. This system can lead to argue with many learners. However, it has no functions to learn with the annotations.

The approach followed in the project of ACTS082, called DIANE was to allow the recording what an arbitrary application outputs as the basic content of multimedia document and to annotate it with all kind of media available to the user. This system defines the annotation framework but the target for usage is not defined clearly. Therefore, it can cope with the ordinal video but cannot contend with two videos recorded by video conference system. Because it has no players to show these synchronously.

Classroom 2000 project focuses on the teaching material. There is a hypothesis; tools to aid in the capture and subsequent access of classroom information will enhance both the learning and teaching experience. Therefore, teacher can make teaching materials with various media and show it to learner. Learner can annotate on the teaching material. However, learner cannot study with learner's own video.

Lieberman et al (1997) has developed a system to use video as a tool for procedural knowledge representation. However, this system cannot annotate for the invisible knowledge like language learning. Because the comment is drawn on the static image from original video.

In order to search and retrieve from large video archive, Davis (1995) has proposed an iconic visual language called Visual language. The prototype system called Media Stream produces representations in a frame language. Visual language can classify the meaning of the frame mechanically. This system uses this idea to show the contents of correction.

As well known, SMIL (Synchronized Multimedia Integration Language) is also defined to enable simple authoring of interactive audiovisual presentations. SMIL is typically used for "rich media"/multimedia presentations which integrate streaming audio and video with images, text or any other media type. Therefore, SMIL makes a linear presentation stream. VCML is different from SMIL because VCML can give a non-linear and hyperlinked presentation stream. For example, learner can see only the original video, and compare it with the revised one. Moreover, SMIL links video with timestamp but VCML does it with a frame for an easy-video-editing like video-editing tool, e.g., MediaStudio (Ulead Systems Inc.) and Premiere (Adobe Systems Inc.).

3.Language Learning using Video Correction

3.1 Video Correction

Text-based correction shows wrong part with red pen on the same paper. It is easy for learner to grasp teacher's task. According to such a traditional correction manner, we have developed an online markup-based collaborative writing support system, called CoCoA (Communicative Correction Assisting System) (Ogata et al, 1997) and CoCoA-J (CoCoA for Java) (Ogata et al, 2000), which is very similar to real correction environment with pen and paper. The system has been used for language learning, and it can automatically generate original text and revised one from corrected text. However, the system cannot deal with multimedia contents such as video and audio. Therefore, it cannot correct video like text.

We focus on the smallest unit. Text consists of many characters. Text-based correction operates for one character. Similarly, video consists of many frames. Video can operate like text if teacher operates frame as character. Video-based

correction operates for one frame. Figure 1 shows text correction and video correction. Text correction in (A) shows deletion and insertion for characters. Video correction in (B) shows them for frames.
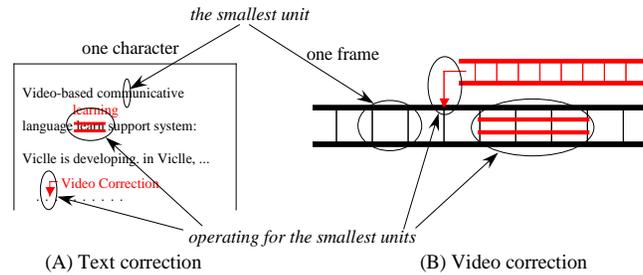

(A) Text correction (B) Video correction
Figure 1   Text correction and Video correction.

3.2 Online video-editing model

There are several editing systems that support teachers to review and correct learners' drafts. Farkas & Poltrock (1995) classified online text-editing model; silent editing model, comment model, edit trace model, and traditional markup model. In silent editing model, teacher overwrites and corrects learner's document. In comment model, annotations are inserted in original document. In edit trace model, teacher works in the manner of author, deleting, adding, and moving text as usual. The computer can compare editor's new version with original one, and allows learner to view draft that contains teacher's changes. Microsoft Word accepts this model. Traditional markup model employs correction marks to the computer screen, for example, deletion, insertion, and move. According to the classification, we propose online video-editing model as follows: (See figure 2)

(A) *Silent model:* After correcting wrong parts of original video with general video-editing tool, teacher creates new video file. This is the simplest model but it is very difficult for learner to check teacher's work. This model is destructive because teacher cannot readily recover original video once he/she has made revised video.

(B) *Comment model:* This model employs pop-up notes, temporary footnotes, hidden text, and special symbols placed within original video. For example, system based on this model is developed with SMIL. It often takes much time for teachers and learners to obtain revised video from comments in original video.

(C) *Mark-up model:* This adapts the traditional marks with pen and paper to video editing on the computer screen. The symbols are both familiar and intuitive for teachers and learners; for example, deletion, insertion, and move. For instance, MATE (Hardock et al, 1993) that is text-based collaborative writing tool, allows teacher to use both digital ink and voice command toward pen and voice computing. In this model, both learners and teachers can interpret teacher's modification with markings much more readily than in the other models. Moreover, they can easily derive both original video and revised one from marked-up video.
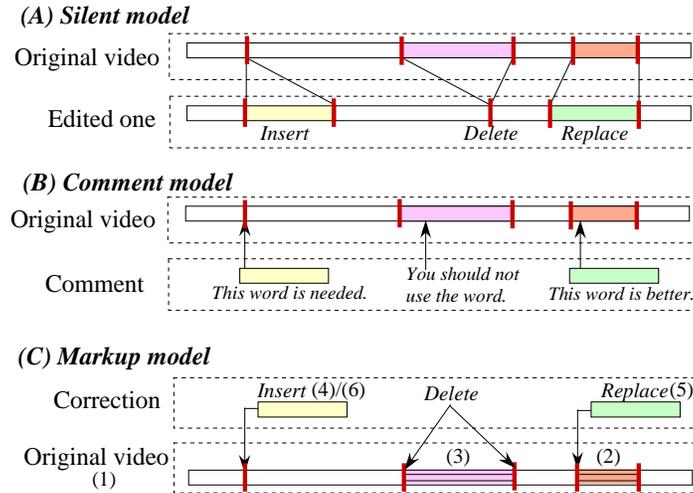
Figure 2: Taxonomy of online video editing models.

Many systems employ traditional mark-up that allows multiple users to mark-up an electronic document as if they were marking up a printed copy of document. However, no studies have ever tried to adapt markup model into video-editing system. Therefore, this paper proposes video markup model for supporting language learning.

3.3 Correction information

After correcting video, there is a lot of information. This chapter classifies the definition as follows: (See the item (C) in Figure 2)

(1) **Original videos:**   The original conversational videos of both teacher and learner.

(2) **Error:**   The error part in the original videos.

(3) **Delete:**   The deleted part in the original video.

(4) **Explanation:**   The information to explain.

(5) **Revision:**   The information to replace the mistaken part.

(6) **Comment:**   The information to comment like an advice.

(7) **Index:**   Representing the corrected contents of video.

3.4 Learning process using Viclle

The learning flow shown as Figure 3 is as follows.

(1) **Recording with video conference system:**   Viclle records the videos of both a teacher and a learner. The system stores the file-database with the video files after recording. The learner and the teacher share the files.

(2) **Correcting the videos:**   A teacher corrects the recorded videos with Viclle Editor. A teacher sends VCML file (made by Viclle Editor) to a learner after correcting.

(3) **Learning by the correction:**   A learner studies by the corrected videos with Viclle Viewer.

(4) **Entering questions:**   A learner answers questions from a teacher and enters a learner's questions. Viclle Viewer generates VCML file (questions and learning situation added) and sends to a teacher.

(5) **Grasping the learner situation:**   A teacher grasps the learner situation with Viclle Editor and refers to work in next time.
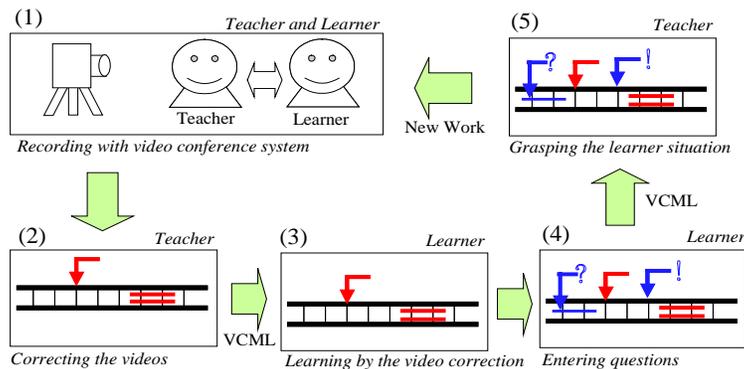
The above is repeated in learning.

Figure 3 Learning flow using Viclle.

4. VCML: Markup Language for Video Correction

There are various media as corrected information after correcting video. There is the learner situation after a learner learns. We propose VCML (Video-based Correction Markup Language) to represent these information. This markup language is based on XML. The advantages are shown as follows:

(1) The corrected video is shown by the link between the original video and the corrected information.

(2) It is possible to use files on network.

The features of VCML are shown as follows:

(1) It does not edit the original video.

(2) It becomes the revised video with applying the tags of the revised information.

(3) It becomes the explained video with applying the tags of the explained information.

(4) It is possible to represent the learner situation.

(5) It is independent of kinds of computers and software.

(6) It is possible to send and to get by e-mail because it is text format.

Table 1 VCML tags for video-based correction.

Correction	Status Mark	Contents
Insert		Inserting media
Overlap		Composing comments
Replace		Replacing with other media
Delete		Deleting the part
Repeat		Repeating the part
Pause		Pausing the point
Annotation		Inserting media as a comment

Table 2 Media by VCML in correction.

Correction	Text	Picture	Audio	Video	Web
Insert			✓	✓	✓
Overlap	✓	✓			
Replace			✓	✓	
Delete					
Repeat					
Pause	✓				✓
Annotation	✓	✓	✓	✓	✓

4.Implementation

We have developed Viclle on Windows 98/NT/2000. The development language is Java SDK1.3.1. Java Media Framework (JMF) 2.1.1a as a framework to deal with media, JAXP (Java API for XML Parsing) as a XML Parser, RELAX (Regular Language description for XML) to analyze XML documents, and Relaxer as a parser of RELAX are used.

The snapshot of Viclle Editor is shown in Figure 7. In the Editor, teacher edits original video with correction mark. (A) Correction Editor shows original video for correcting. (B) Timeline shows frames of original video. (C) Timeline Controller manages it Teacher can select correction mark with (D) Correction Palette. Correction mark is inserted by clicking (C). An example that inserts replace correction is shown. When replace button is clicked, (E)Correction Dialogue appears. Then, user selects media type; text, audio, video, and so on. In this case, user selects and inputs text. After that, correction of replace is completed. Similarly, user can insert other corrections. (F) Correction Status Bar shows inserted corrections. The contents for playback can check with (G) Play Mode.
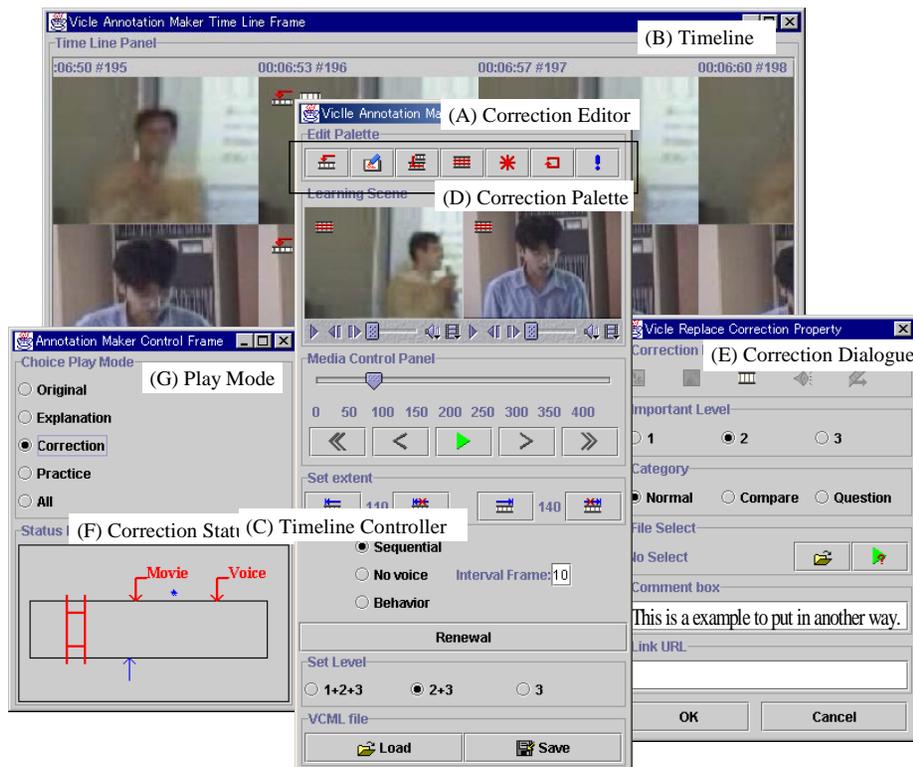

Figure 7   Snapshots of Editor on Viclle.

5.Conclusion

This paper proposed online video-based correction system for language learning. The prototype called Viclle can correct video and VCML can exchange corrected information. Although we tried to use Viclle for language learning, the system is independent of the domain. For instance, it can use sports training for the form; acquire the gesture in the different culture like between Japanese and USA. In our future work, we will use VCML document to retrieve from video database that are stored VCML documents. The data can use as error database. In another point, we examine to use speech recognition system to help the correction work and train learner's pronunciation or intonation.

Acknowledgement

References

Lieberman, H.: A User Interface for Knowledge Representation from Video (1997). Conference of the American Association for Artificial Intelligence, Seattle, August 1994.

Abowd G. D., Chris Atkeson, Ami Feinstein, Cindy Hmelo, Rob Kooper, Sue Long, Nitin "Nick" Sawhney and Mikiya Tani. Teaching and Learning as Multimedia Authoring: The Classroom 2000 Project. In the Proceedings of the ACM Multimedia'96 Conference, pp. 187-198, November 1996.

Jason A. Brotherton, Janak R. Bhalodia, and Gregory D. Abowd. Automated Capture, Integration, and Visualization of Multiple Media Streams. In the Proceedings of IEEE Multimedia '98, July, 1998. To appear.

Davis, M.: Media Streams: An Iconic Visual Language for Video Representation, In: Readings in Human-Computer Interaction: Toward the Year 2000, ed. Ronald M. Baecker, Jonathan Grudin, William A. S. Buxton, and Saul Greenberg., pp. 854-866. 2nd ed., San Francisco: Morgan Kaufmann Publishers, Inc., 1995. (Also available at: http://www.interval.com/papers/mediastreams/index.html)

Ogata, H. and Yano, Y.: CoCoA: Communicative Correction Assisting System for Composition Studies, Proc. of ICCE 97, pp.461-468, Malaysia, Dec., 1997.

Ogata, H., Hada, Y. and Yano, Y.: CoCoAJ: Supporting Online Correction of Hypermedia Documents for CALL, ICCE 2000. Vol.1, pp.323- 329, Taipei, Taiwan, Nov. 2000.

Ogata, H., Feng, C., Hada, Y. and Yano, Y.: Online Markup Based Language Learning Environment, International Journal of Computers & Education, Vol.34, No.1, pp.51-66 (2000).

Benz, H., Bessler, S., Fischer, S., Hager, M., and Mecklenburg, R. DIANE: A Multimedia Annotation System in Proceedings of the Second European Conference on Multimedia Applications, Services and Techniques (ECMAST'97) (Milan IT, May 1997), Fdida, S., and Morganti, M. (Eds.), Lecture Notes in Computer Science 1242, Springer Verlag, Berlin, 183-198.

Benz, H., Fischer, S., Mecklenburg, R., and Dermler, G: DIANE - Hypermedia Documents in a Distributed Annotation Environment in Proceedings of the Conference on Hypertext - Information Retrieval - Multimedia (HIM'97) (Dortmund DE, September 1997), Norbert, F., Gisbert, D., and Tochtermann, K.(Eds.), Schriften zur Informatik, UVK Universitaetsverlag, Konstanz, 293-306.

Bouras, C., Kapoulas, V., Konidaris, A., Ramahlo, M., Sevasti A., Van de Velde, W.: Using Multimedia to Support Reflection on Past Events for Young Children, ED-MEDIA2000 Proceedings, pp.105-110 (2000).

Sevasti, A. V. and Bouras C. J.: Using Java to implement a multimedia annotation environment for young children, Proc. of the 8th ACM international conference on multimedia, Oct. 30-Nov. 3, 2000, Los Angels, CA USA, pp. 187-194.