

Communicative Language-Learning Environment Using Video Annotation System

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 a communicative language-learning environment between a native speaker and a learner using a videoconference system. This paper describes Viclle (Video-based Communicative Longuage Loarning System) that supports both a teacher and a learner. A teacher (a native speaker) can correct a recorded video with Viclle. This paper also describes VCML (Video-based Correction Markup Longuage) based on XML to show corrected information. The VCML file can use the video files on network. Using this system, both a teacher and a learner can communicate by e-mail without sending video files.

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

1. Introduction

In the present, foreign language learning bases on systematic theories by written language mainly. Therefore, it is difficult to acquire communicative skills like a pronunciation, an intonation and a gesture. If a class is formed by many people, it is not easy to coach. In order to acquire skills to communicate with a foreign language, it is important for a learner to hear the language with a learner's ear and to understand. So, the environment to communicate with a foreign language is needed.

The approach of the language learning through communication is called "communicative approach", and attracts attention. This theory uses a language as a means. In this theory, a learner practices to tell what a learner wants to tell with a foreign language. The gap of the knowledge between a teacher and a learner is filled up by using a learning language. Therefore, CALL (Computer Assisted Longuage Loarning) with communicating between a teacher and a learner is paid attention.

In the other side, the technology of network and computer is progress. Smoother communication can be performed using not only a text but a video (LeeTiernan, et al., 2001). For example, in a school of conversational English, the environment to communicate with a teacher (a native speaker) using a video conference system is realized. In learning with a videoconference system, it is difficult to record and use conversational situation for educational instruction.

In this paper, we propose a system that can add what a teacher wants to revise and to comment. Usually, it is difficult to understand where a teacher corrected in an original video because a fixed video is created. Therefore, a videoconference system is extended to record a video talking between a learner and a teacher. This paper proposes video correcting system, Viclle (Video-based Communicative Longuage Loarning System) that is a language learning system to support correcting a conversational video between a teacher and a learner. Video correction in this paper is to add the corrected information, various media and unnecessary parts where a teacher wants to delete. It is easy for a learner to understand where and how a teacher corrects. Viclle has the features as followings.

- (1) A learner can correct a video talked between a teacher and a learner.
- (2) A learner can learn by interacting with this system based on marks corrected by a teacher.

- (3) Errors in an original video are related with correction marks. This system can show where and how a learner made mistakes.
- (4) A revised conversational video is generated if VCML tags are applied to an original video.
- (5) Describing corrections of a conversational video uses VCML (Video-based Correction Markup Language) based on XML (eXtensible Markup Language).

Some systems that can insert comments into a video are developed as related researches. For example, there is a system to insert the comments that are video or text into a static image decomposing a machine (Lieberman, 1997). To increase recognizable and communicative skills of children, a system is proposed to communicate using other children's comment after inserting things a child noticed while playing video as a teaching material (Bouras, et al., 2000). Visual Language shows the mean of the point on a video by using icons (Davis, 1993). These systems can edit or create but can not correct video.

This paper is organized as the following. First, Video Correction is proposed in section 2. Playing of revised video is proposed in section 3. Section 4 illustrates the implementation. Finally, the concluding remarks are given in section 5.

2. Language learning by correcting video

2.1 Video Correction

The correction on paper is written with a red pen. The corrected text can be understood what is wrong somewhere because the revised contents are shown in a red part. So we developed CoCoA (Communicative Correction Assisting System) (Ogata, et al., 1997) and CoCoA-J (Ogata, et al., 2000) that environment is very similar to real one with a pen and paper. They can generate the original text and the revised one from the corrected text. It can show corrected situation by a viewer. But CoCoA can not deal with video and sound.

Correction in video on computer by comparing with correction on paper. The constructing smallest unit on a video is one frame, and the constructing smallest unit on a text is one letter. Text correction shown as Figure 1 (A) operates deleting or inserting letters that are the smallest units. Similarly, video correction shown as Figure 1 (B) operates deleting or inserting frames that are the smallest units. It is called Video correction. Figure 2 is shown comparing video editing with video correction.

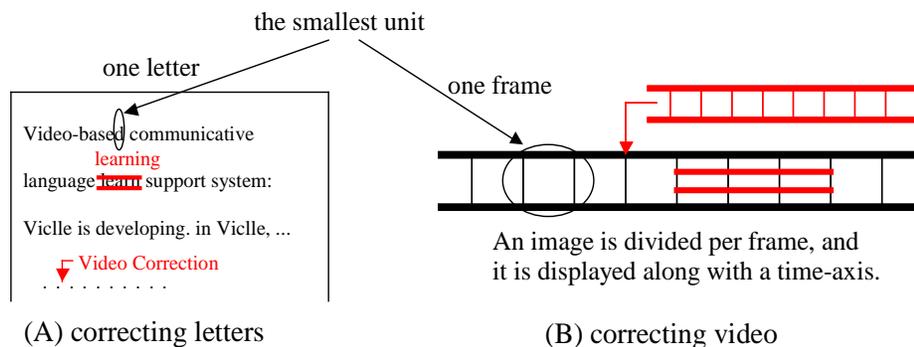


Figure 1 Correcting text and video.

2.2 The definition of correction information

After correcting a video, there is a lot of information inserted by a teacher. Theirs definitions are shown as followings.(refer Figure 2)

- (1) **Original video:** The conversational video talking between a teacher and a learner. A target to correct.
- (2) **Mistaken video:** The mistaken part in an original video.

- (3) **Unnecessary video:** The disused part in an original video.
- (4) **Explained information:** The information inserted into an original video to explain.
- (5) **Revised information:** The information replaced a mistaken video.
- (6) **Commentated information:** The additional information to comment like an advice.
- (7) **Index information:** The information to describe correction. Revised processes are described.

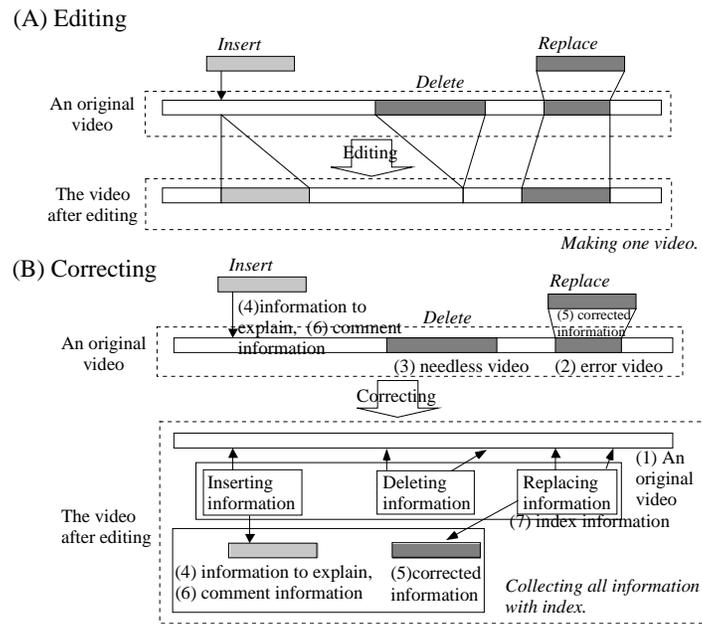


Figure 2 Compare editing and correcting a video.

2.3 Conversational learning by correcting video.

The learning flow shown as Figure 3 is as followings.

- (1) This system records conversational video where both a teacher and a learner learn to talk.
 - (2) A learner and a teacher put recorded video in (1) on file server. A learner sends the address to a teacher. After this, a teacher corrects the video.
 - (3) After correcting, a teacher sends VCML file to a learner. A learner learns by playing the revised video with a viewer.
 - (4) A learner answers questions from a teacher and inserts question where a learner can not understand. After this, this system generates VCML file that describes questions and learning situations and sends to a teacher.
 - (5) A teacher grasps a learner's learning situations from VCML file and refers to next time.
- The above is repeated in learning.

3. Playing revised video

In this chapter, an interactive learning way between a learner and a teacher to play is proposed.

3.1 Kinds of correction

If a teacher explains something, it is difficult to explain with only speech and a gesture. So, writing is used to explain. An important part where a teacher wants to stress is repeated. In practicing to pronounce, a teacher is

understood by speaking each other. A teacher not only explains but also asks a learner. From the above, the correction on Vicle is classified as followings.

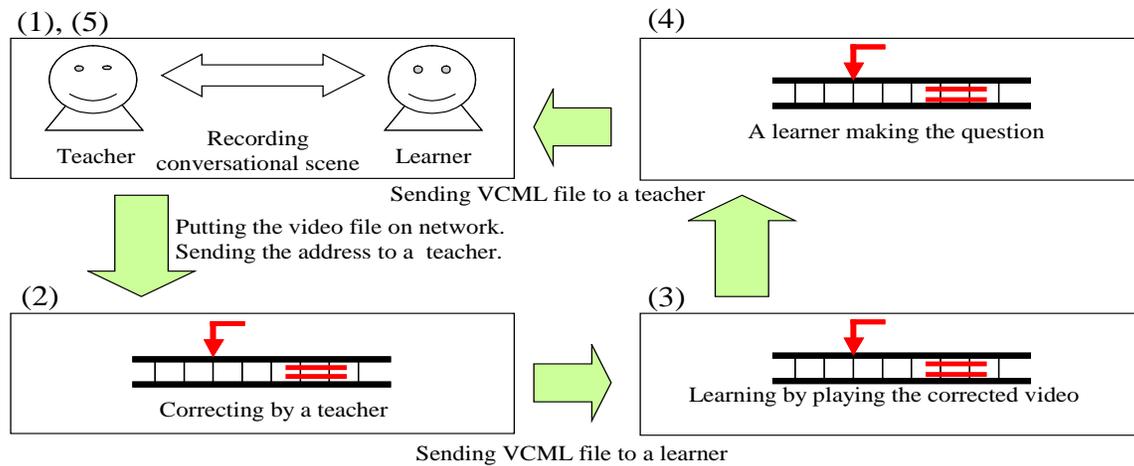


Figure 3 Learning flow in Vicle.

- (1) **Insert:** Inserting explained information into specified frame on an original video. Playing an original video is paused while explained information is playing.
- (2) **Overlap:** Playing letters and figure as explained information on specified frames in an original video. Explained information is shown at that time when an original video is played.
- (3) **Delete:** Specifying unnecessary video in an original video. The frames of unnecessary video are not played.
- (4) **Replace:** Replacing the part of a mistaken video with revised information. The way to play a revised video is translated by classified correction.
- (5) **Annotation:** Inserting commented information on specified frames. Commentated information is only appeared on another windows while playing an original video. An original video is not stopped.
- (6) **Repeat:** Repeating how many times a teacher wants emphasis.
- (7) **Pause:** Showing a dialog to ask whether a student understands after pausing on the specified frame.

Moreover, these corrections except (5)Annotation are classified as followings. This information set the playing way against a learner by a teacher. Each correction mark is set.

- (1) **Normal:** It is a normal correction. Showing a dialog to ask whether a student understands after playing. In replace correction, revised information instead of a mistaken video is played.
- (2) **Comparison:** This can select in only replace correction. A comparison between a revised video and a mistaken video is played. This has two modes. The one is playing alternate a revised video and a mistaken video, the other one is synchronous playing a revised video and a mistaken video.
- (3) **Question:** Asking a question after playing a correction. A dialog to ask a question is shown after playing.

In addition, the level to show importance is prepared to tell whether its correction is important as followings.

- (Level 1) Proposing other turn of phrase or making a suggestion.
- (Level 2) It is preferable to correct.
- (Level 3) It should correct.

The initial value is two. The value is changed if a teacher judges to change.

3.2 Situated information of a learner

If a learner does not understand what a teacher explains while a learner studying, a learner may think again to understand. But if a learner really do not understand, a learner may give up. Therefore, it is important for a teacher to grasp the situation of a learner after studying. So, situated information that is returned to a teacher is classified as followings.

- (1) **Opinion:** Describing the question or the opinion a learner asked.
- (2) **Answer:** Describing a reply and what a learner wants to say.
- (3) **Repeat:** Recording what a learner repeat. This is recorded automatically.
- (4) **Playing Situation:** describing the situation understood by a learner or whether the correction is played.

3.3 VCML: A markup language to describe video correction

There are various media as corrected information after correcting video. There are a learner's situations after a learner studying. So, we proposes VCML (Video-based Correction Markup Language) that is used in language learning to describe index information. This markup language is based on XML. The advantage described correction information with using XML is shown as followings.

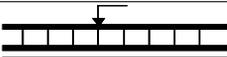
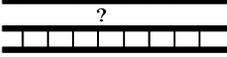
- (1) Making a new video file by processing old video files is not needed.
- (2) It is possible to use files on network.

A communication between a teacher and a learner from (2) is supported by only a VCML file. The features of VCML is shown as followings.

- (1) It contains an original video and corrected information.
- (2) It becomes a revised video by applying the tags shown revised information.
- (3) It becomes an explanatory video by applying the tags shown explained information.
- (4) It is possible to describe the situation of a learner after studying.
- (5) It is not depended on kinds of computers and software.
- (6) It is possible to send and to get with an e-mail because it is text format.

The structure of VCML is shown as Figure 4, an example of a VCML file is shown as Figure 5. So, the explanation of each part is shown as followings.

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

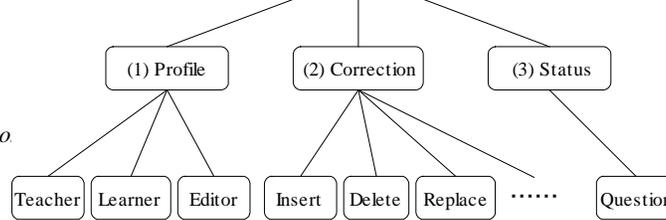


Figure 4 Structure of VCML

Table 2 Media used by VCML in correction part.

Correction	Letters	figure	sound	movie	link
Insert	false	false	true	true	true
Overlap	true	true	false	false	false
Replace	false	false	true	true	false
Delete	false	false	false	false	false
Repeat	false	false	false	false	false
Pause	true	false	false	false	true
Annotation	true	true	true	true	true

Table 3 VCML tags in status part

Response	Contents
Opinion	Describing a learner's opinion
Question	Describing the answer from a teacher
Repeat	Recording a learner's repeating point

- (1) **Profile part:** It is described the profile of a learner and a teacher and it is specified original videos.
- (2) **Correction part:** Corrected information is described. Corrected information is shown by correction mark shown as Table 1. In this correction mark, the media shown as Table 2 are used. Those correction marks are added after studying.
- (3) **Status part:** the situation of a learner except playing situation is described. It is made after a learner studying. The tags shown as Table 3 that is described situation information are used.

3.4 Playing video correction

The playing mode of corrected information is shown as followings.

- (1) Playing an original video: Only an original video is played.
- (2) Playing correction: All corrections are played.
- (3) Playing revised correction: an original video and explained information are played. Disuse videos are not played.
- (4) Playing revised correction: Only an original video and its revised information are played. Disuse videos are not played.
- (5) Playing to practice: Only the part that is played the comparison.

4.Implementation

We have developed Viclle on Windows NT. The development language is Java2. The system configuration is shown in Figure 5. Viclle consists of Editor for a teacher and Viewer for a learner. The snapshot of Viclle Editor is shown in Figure 6. In the Editor, a teacher inserts corrections in the original video.

The flow of correction is explained by Figure 6. The video to correct is shown as (A)Correction Editor. Frames decomposing the video are shown as (C)Time Line Panel. A correction mark is inserted by pushing (B)Correction Palette after a teacher finds where to correct. In this interface, replace correction is inserted. When replace button is pushed, (D)Correction Dialogue is appeared, and the correction is inserted. Similarly, other corrections are inserted. (F)Correction Status Bar shows inserted corrections. It is possible to check playing situation in (E)Play Mode after correcting.

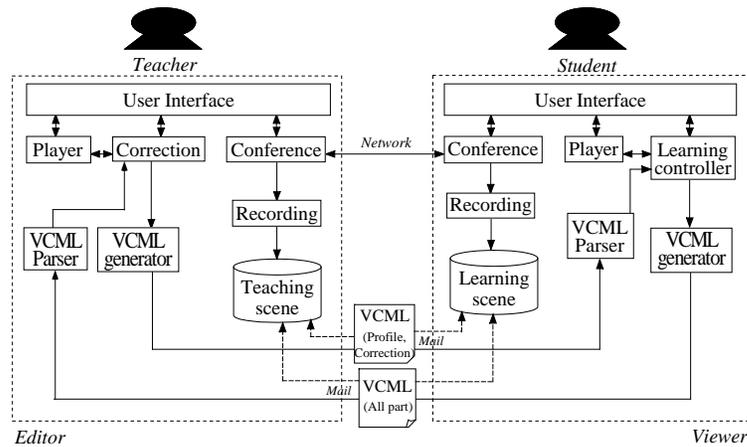


Figure 5 System Configuration

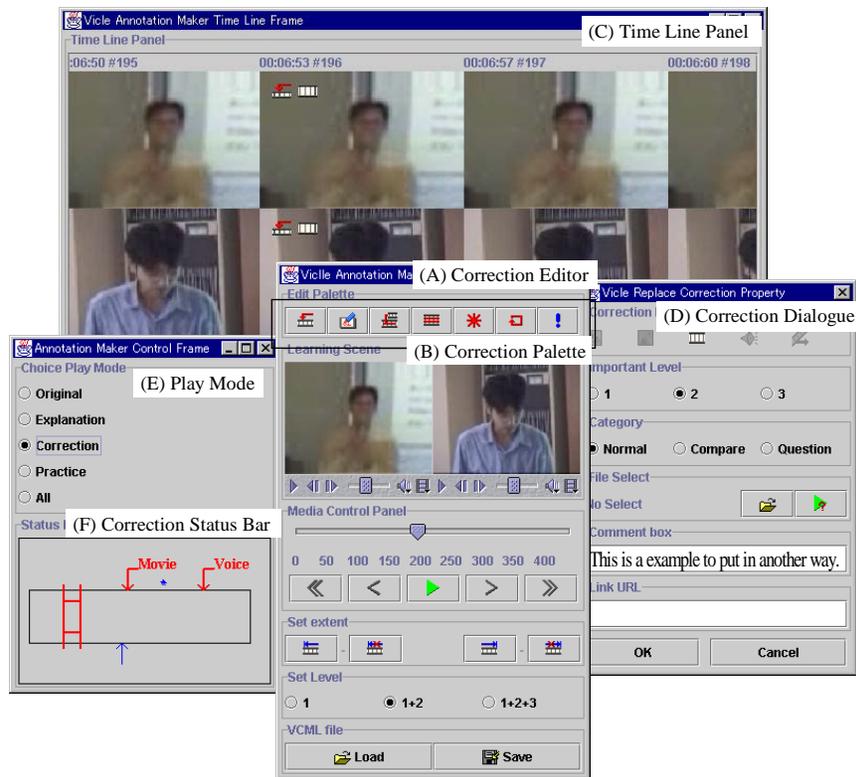


Figure 6 Snapshots of Editor on Vicle.

5. Conclusion

In this paper, we proposed *Viclle* that can correct video and VCML that is a markup language to exchange information between a learner and a teacher. In the future work, we will evaluate *Viclle* and investigate how to support video-based language learning in a cyber space, e.g. using speech recognition, searching static images to help correction work.

Acknowledgement

This work was partly supported by the grant to the research project at Doshisha University named "Intelligent Information Science and Its Applications to Problem Solving in Engineering Fields" from the Ministry of Education, Science, Sports and Culture, Japan, and also supported in part by the Grant-in-Aid for Scientific Research No. 13780121, No. 12558011, and No. 11878032 from the Ministry of Education, Science, Sports and Culture in Japan.

References

- 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).
- Davis, M.: Media Streams: An Iconic Visual Language for Video Representation, available at <http://www.interval.com/papers/mediastreams/index.html>(1993).
- LeeTiernan, S. and Grudin, J. Fostering Engagement in Asynchronous Learning through Collaborative Multimedia Annotation, INTERACT 2001 (To appear in Proc.), 2001.
- Lieberman, H.: A User Interface for Knowledge Representation from Video (1997). available at <http://lieber.www.media.mit.edu/people/lieber/Lieberary/Mondrian/Knowacq/Knowacq.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).