

CoCoAJ: Supporting Online Correction of Hypermedia Documents for CALL

Hiroaki Ogata, Yoshiaki Hada, Cheng Feng and Yoneo Yano

Dept. of Information Science and Intelligent Systems,

Faculty of Engineering, Tokushima University

2-1, Minamijosanjima, Tokushima 770-8506, Japan

Tel: +81-886-56-7498 Fax: +81-886-23-2761

E-mail: {ogata, hada, feng, yano}@is.tokushima-u.ac.jp

URL: <http://www-yano.is.tokushima-u.ac.jp/ogata/>

This paper describes an online markup-based composition learning environment system called **CoCoAJ** (Communicative Collection Assisting System for Java). It allows students and teachers to exchange marked-up documents via Internet, and its environment is very similar to a real world one in which people use pen and paper. In order to record and exchange corrected compositions with marks and comments, this paper proposes XCCML (eXtensible Communicative Correction Mark-up Language), that is based on XML (eXtensible Mark-up Language). XCCML facilitates teachers to analyze and reuse the marked-up documents for the instruction.

Keywords: **Computer assisted language learning, Collaborative writing, XML, Online document correction, Hypermedia.**

1. Introduction

Recently, teacher-centered instructional approaches in traditional writing classrooms are replaced with more active and learner-centered learning approaches with collaborative writing tools[2]. These tools can (1) change the way students and teachers interact; (2) enhance collaborative learning opportunities; (3) facilitate class discussion; and (4) move writing from solitary to more active and social learning. Writing compositions includes various sub-processes such as planning, transcribing, and revising, which do not need to occur in any fixed order [19]. In particular, the review process assisted with computer-based writing tools, has recently received much interest (see as examples [4, 9]).

Many researchers developed online markup systems employing some markup models. However, it is very difficult to analyze and reuse the marked documents that are collected through the writing classroom because the documents do not have a common structure. Therefore, it is necessary to define the generalized format for encoding and exchanging the marked-up documents in order that online markup systems are used easily and widely.

CoCoA (Communicative Correction Assistant system) has been developed for supporting foreigners and teachers to exchange marked-up documents by e-mail [14]. Its environment is very similar to a real one in which people use paper and pen. CoCoA allows teachers not only to correct the compositions sent from foreigners by E-mail, but also foreigners to see where and why the teacher had corrected them. CoCoA improves the opportunities that foreigners have for writing Japanese compositions and for receiving instructions from teachers. **CCML** (Communicative Correction Mark-up Language) [15] has also been proposed for the representation of marked-up documents, which is based on **SGML** (Standard Generalized Mark-up Language) [8]. With CCML, teachers and students can exchange marked-up documents via e-mail [16, 17]. In the experimental use of CoCoA, most of users commented that CoCoA was easy for them to understand the mistakes in documents because of the use of marks, and that the optional view of the original, marked or revised text was very useful. However, CoCoA cannot show users a hypermedia document including figures, tables, movies and links because it deals with only text.

This paper tackles how to correct hypermedia documents by the extension of CoCoA. This paper proposes **CoCoAJ** (CoCoA for Java) to do so. Also this paper describes **XCCML** (eXtensible CCML) for correcting

hypermedia documents, that are based on XML (eXtensible Markup Language). XCCML is combined CCML with HTML (Hyper Text Markup Language) that can represent hypermedia documents including pictures, movies, audios and so on.

We have been investigating technological support for Japanese language learning among overseas students. For example, CAI systems called Kanji Laboratory [7], JUGAME [23], GRACILE[23] and JULLIET[1] were developed to support Japanese language learning. However, an on-line mark-up supporting system for Japanese language learning has not yet been proposed. Usually, in a Japanese writing classroom, teachers have to individually review learners' documents using pen and paper[18]. It takes a lot of time for teachers to do this. Therefore, we have implemented CoCoA for writing Japanese composition.

2. Online markup models

There are some editing systems that support teachers to review and correct the students' drafts with online mark-up. Farkas & Poltrock [5] classified the mark-up models as followings:

- (1) *Silent editing model*: This is the simplest model and it requires no special techniques. However, it is very difficult for the author to check the editor's work. This model is destructive because the editor cannot readily recover the original words once he/she has changed it.
- (2) *Comment model*: This model employs pop-up notes, temporary footnotes, hidden text, and special symbols placed within the text. This model can work for special groups and ad-hoc situations. A system called XyWrite[10] was proposed with this model.
- (3) *Edit trace model*: In this model, the editor works in the manner of an author, deleting, adding, and moving text as usual. The computer can compare the editor's new version with the original text, and allows the author to view the draft that contains the editor's changes. This model is apt to encourage heavier editing and less regard for the author's original text. Microsoft Word accepts this model.
- (4) *Traditional mark-up model*: This adapts the traditional paper mark-up model to the computer screen. The symbols are both familiar and intuitive for editors and authors; for example, deletion, insertion, and move. For instance, Red Pencil allows the editor to apply a complete set of traditional editing symbols directly to a document. The editor uses "digital ink" to mark a traditional editing symbol along with the words. Moreover, MATE[6] allows the editors to use both digital ink and voice command toward pen and voice computing. In this model, authors and editors can interpret the editor's markings much more readily than in the edit trace model.

There are many systems that employ traditional mark-up which allows multiple users to mark-up an electronic document as if they were marking up a printed copy of the document. However, such systems do not globally come into practical and wide use in composition writing classes because of their special format. Moreover, it is very difficult to analyze and reuse the marked documents because the marked documents are unstructured. Therefore, the system should provide a generalized and structured format for encoding and interchanging marked-up documents via the Internet.

3. XCCML

Based on the experimental results, we propose XCCML for exchanging marked-up documents. XCCML is an application of XML, and it supplies a formal notation for the definition of generalized mark-up languages. XML is a device- and system-independent method of representing texts in electronic form. That is to say, XML is a set of mark-up conventions used together for encoding texts. A mark-up language must specify what mark-up is allowed, what mark-up is required, how mark-up is to be distinguished from text and what the mark-up means.

3.1. Features of XCCML

The main characteristics of XCCML are:

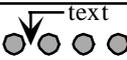
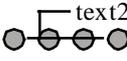
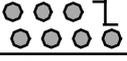
- (1) Based on the experiment, XCCML presents six marks and annotation XCCML tags.
- (2) The marks have three degrees of importance levels against respective corrections.
- (3) The original text is generated through removing all the XCCML tags.

- (4) The revised text is derived from the XCCML document.
 - (5) Because XCCML documents are text-formatted, it is easy to send them by e-mail.
 - (6) CCML documents easily make up full-text databases.
- Needless to say, XCCML inherits its features from XML.

3.2.XCCML structure

As shown in table 1, XCCML documents consist of three parts: header, body and close. “Header” represents additional information about the document. For instance, “next” tag denotes the next version of the document. The marks for review are shown in the “Body” as XCCML tags. “Close” shows the editor’s comments. In one sentence, “insert,” “replace” and “delete” marks were used, while “join,” “separate” and “move” marks were used over two sentences. The part between the start tag and the end tag denotes the learner’s mistakes. The “string” attribute represents the revised part of the document.

Table 1: Marks and XCCML tags.

Correction	Mark	Tag
1. Insert		<Insert string="text"/>
2. Replace		<Replace string="text2"> text </Replace>
3. Delete		<Delete> text </Delete>
4. Separate		<Separate/>
5. Join		<Join/>
6. Move		<Movefrom refid="id"/> <Moveto id="id"> text </Moveto>

(1) Root tags

Tag name	Explanation	Attribute	Attribute’s contents	End tag
XCCML	Beginning of XCCML tag	Version	Version number	</XCCML>
Head	Header information	None		</Head>
Body	Corrected document	None		</Body>
Close	Overall comments	None		</Close>

(2) Tags in header section

Tag name	Explanation	Attribute	Attribute’s contents	End tag
Title	Title of the document	String	Title name	None
Editor	People who corrected the document	Name	Name of the editor	None
		Email	Email address	
Author	People who write the original document	Name	Name of the author	None
		Email	Email address	

(3) Tags in body section

Tag name	Explanation	Attribute	Attribute’s contents	End tag
Insert	Insert words	String	Inserted words	</Insert>
		Level	Level of importance	
		Comment	Comment for the correction	
Replace	Change words	String	Corrected words	</Replace>
		Level	Level of importance	
		Comment	Comment for the correction	
Delete	Delete words	Level	Level of importance	None

		Comment	Comment for the correction	
Separate	Separate a paragraph	Level	Level of importance	None
		Comment	Comment for the correction	
Join	Join paragraphs	Level	Level of importance	None
		Comment	Comment for the correction	
Moveto	Moved position from the original spot	Fromid	ID	None
		Comment	Comment for the correction	
Movefrom	Original position of the words	Toid	ID	</Movefrom>
		Level	Level of importance	
Annotate	Comment for the correction	String	Annotation contents	None
		Level	Level of importance	

(4) Tags in close section

Tag name	Explanation	Attribute	Attribute's contents	End tag
Comment	Overall comment	None	None	Necessary

3.3.Level of marks

We found that the marks do not have the same level of importance. We identify corrections on the following levels:

- (1) Weak correction: The learner does not need to revise the document.
- (2) Normal correction: The learner should correct the document.
- (3) Strong correction: The learner must correct the document.

The strong corrections denote the important part of marks to be revised in the document. Using the importance level that the teacher had entered, the system provides the learner with the marks he/she wants to see. Therefore, the learner can avoid information overload from the reviewed documents. Every tag in table 1 has an attribute "level" that a teacher gives a number from one to three. Its default is two as normal correction.

3.4.Level of annotations

It is very important for a teacher to annotate the marked text for instruction in composition. For example, PREP Editor [12] is a word processor that allows writers and reviewers to create electronic margins, or columns, in which they can write and communicate through their annotations. We identify the following different kinds of annotations:

- (1) Explanation: This is used for explaining the reason of a correction.
- (2) Question: This is used for asking the learner a question; e.g., what do you want to write?
- (3) Comment: This shows the educational view of the teacher with respect to the document.

4.CoCoAJ

CoCoAJ consists of CoCoAJ-Editor and CoCoAJ-Viewer. A teacher corrects the document of overseas students with CoCoAJ-Editor and the system saves the marked document in XCCML format. CoCoAJ-Viewer interprets the XCCML document, and shows the learner the marked document.

4.1.Learning processes using CoCoAJ

By using CoCoAJ, a learner receives instruction about a Japanese composition from a teacher with the following processes:

- (1) The learner writes an original text with his/her familiar editor.
- (2) The learner sends the document to his/her teacher with his/her own e-mail tool.
- (3) CoCoAJ-Editor makes the document double-spaced. The teacher corrects the document with online marks

- and annotations. Then, the system allows the teacher to set the importance level to the marks in the document.
- (4) After CoCoAJ-Editor saves the marked text as a XCCML document, the teacher sends it to the learner by e-mail.
 - (5) CoCoAJ-Viewer provides the learner with the marked text after interpreting the XCCML document. Then, the system allows the learner to select the importance level to see the important part of the marked text.
 - (6) CoCoAJ-Viewer automatically generates both the original text and the revised one from the XCCML document. After editing the revised text, the learner can send it again to the teacher and continue refining the text.
 - (7) CoCoAJ maintains the version of the document, if the learner wants to revise the same document.

4.2. System configuration

Figure 1 depicts the learning environment of CoCoAJ.

- (1) XCCML parser: This module analyzes XCCML documents using the XCCML parser after reading them through the file management module. Then, it provides the results of correction according to the level of importance of marks.
- (2) Correction module: This module inserts XCCML tags into the learner's document, according to the revision of the teacher. After saving the marked text, the teacher sends it by e-mail to the learner.
- (3) Original text display module: This module generates the original text from the XCCML document by removing all the XCCML tags.
- (4) Revised text display module: This module generates the revised text by applying XCCML tags.
- (5) File management module: This module manages the versions of the documents. When the learner sends the teacher the revised document, the system creates a new XCCML document, inserts the "next" tag into the old XCCML document, and also enters the "previous" tag into the new XCCML document.

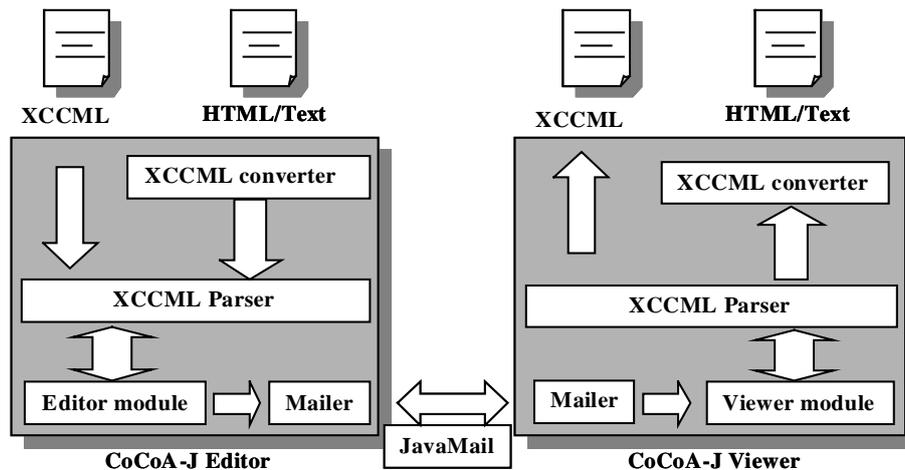


Figure 1: System configuration of CoCoAJ.

4.3. User interface

Figure 2 shows the screen snapshot of CoCoAJ-Editor. First, the learner writes a Japanese composition with a word processor and saves the document as HTML format. After that, the learner sends the document to the teacher by e-mail. By selecting a mark from the mark palette shown in the upper window, the teacher can revise the document. Moreover, the teacher can annotate the document using the annotation palette, and he/she can classify the marks according to the level of importance. The user can see the correcting document at the left side in the window and "*" means the user inserted the comment. The user can see the comments for the correction at the right side in the window. In this figure, the teacher substitutes "allow" with "allows" and gives a comment "*2". Also the teacher can see the original document and revised one by selecting window tag. After saving the marked document as a XCCML (see appendix A), the teacher can send it to the student by e-mail. Using CoCoAJ-Viewer, the learner obtains the same marked text that the teacher revised. By selecting the level of importance, CoCoAJ-Viewer provides only the marks over the level. The learner can reply to the teacher's comments and collaboratively write a composition with the teacher.

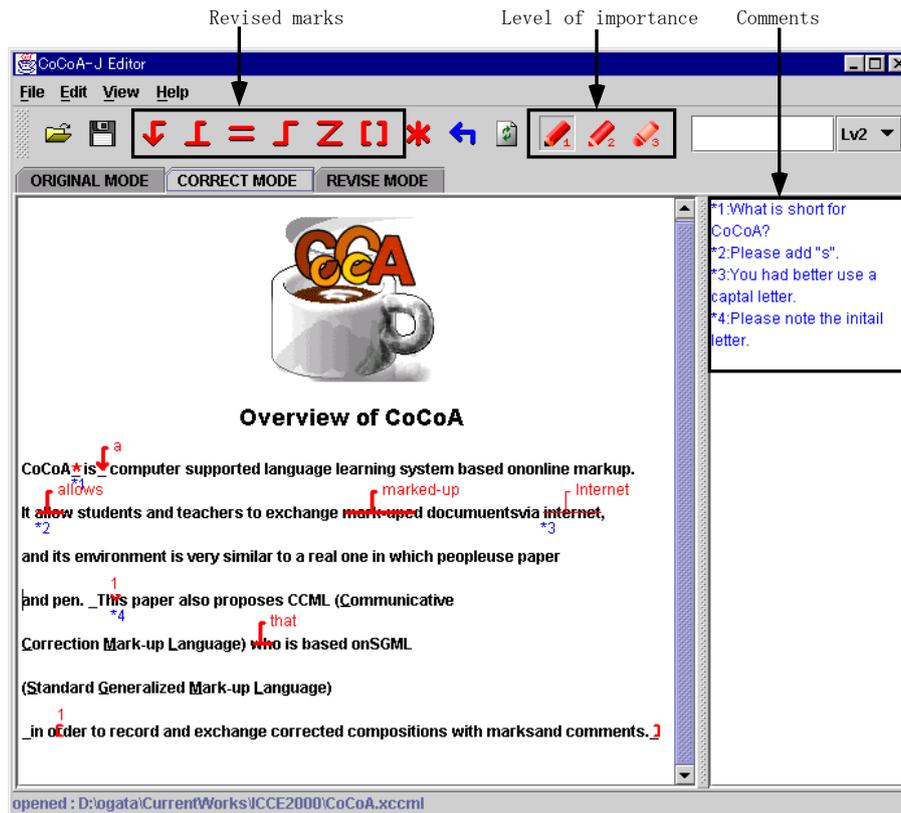


Figure 2: Screen snapshots of CoCoAJ-Editor.

5. Conclusions

This paper proposed a computer mediated language-learning system called CoCoAJ and XCCML for exchanging electronic marked-up documents. Now we are trying to propose XCCML to W3C (World Wide Web Consortium), and to show an XCCML document into Web browsers. After that, CoCoAJ will be able to be used for learning any language in an open-ended writing classroom. In our future research, we will investigate how to classify students' writing errors in their drafts, and how to assist a review process with AI technologies.

Acknowledgment

This research was supported in part by the Grant-in-Aid for Scientific Research (B) (2) No.09558017, No.09480036, and No.09230214 from the Ministry of Education, Science, Sports and Culture in Japan.

References

- [1] Ayala, G. & Yano, Y. (1995). "GRACILE: A framework for collaborative intelligent learning environments", *Journal of the Japanese Society of Artificial Intelligence*, vol.10. no. 6. pp.156-170.
- [2] Bonk, J. C. & King, S. K. (1995). "Computer conferencing and collaborative writing tools: Starting a dialogue about student dialogue," 1995 Proc. of Computer Supported Collaborative Learning.
- [3] Brock, N. M. (1995). "Computerised text analysis: Roots and research," *Computer Assisted Language Learning*, vol.8, no.2-3, pp.227-258.
- [4] Coniam, D. (1997) "A computerised English language proofing cloze program," *Computer Assisted Language Learning*, vol. 10, no. 1, pp.83-97.
- [5] Farkas K. D., & Poltrock, E. S. (1995). "Online editing mark-up models, and the workspace lives of editors and writers," *IEEE Trans. in Professional Communication*, 38(2), pp.110-117.

- [6] Hardock, G., Kurtebach, G., & Buxton, W. (1993). "A marking based interface for collaborative writing," *Proc. of the ACM Symposium on User Interface Software Techonogy*, pp.259-266.
- [7] Hayashi, T., & Yano, Y. (1994). "Kanji Laboratory: An environmental ICAI system for Kanji learning," *IEICE Trans. on Information and Systems*, E77-D(1), pp.80-88.
- [8] Herwijnen, E. (1990). *Practical SGML*, Kluwer Academic Publishers.
- [9] Kehagia, O. & Cox, M. (1997). "Revision changes when using word processors in an English as foreign languages context," *Computer Assisted Language Learning*, vol. 10, no. 3, pp.239-253.
- [10] Kincade, D. & Oppenheim, L (1994) "Marking it up as we go along: Into editorial production's electronic future," *Journal of Scholarly Publishing*, Vol. 25, No. 3, pp.233-242.
- [11] Matsumoto, Y, Kitauchi, A., Yamashita, T., Hirano, Y., Imaichi, O. & Imamura, T. (1997) "Japanese Morphological Analysis System ChaSen Manual," Nara Institute of Science and Technology Technical Report NAIST-IS-TR 97007. (in Japanese)
- [12] Neuwirth, M. C., & Wojahn, G. P. (1996). "Learning to write: Computer support for a cooperative process," *CSCL: Theory and practice of an emerging paradigm*, Koshmann, T. (Ed.), Lawrence Erlbaum Associates, Publishers, pp.147-170.
- [13] Ochi, Y., Yano, Y., & Hayashi, T. (1996). "JULLIET: Interactive learning environment for Kanji compounds learning," *Proc. of Educational Multimedia and Hypermedia 1996*, pp.539-544.
- [14] Ogata, H. and Yano, Y. (1997). CoCoA: Communicative Correction Assisting System for Composition Studies, *Proc. of ICCE 97*, pp.461-468.
- [15] Ogata, H., Yano, Y., and Wakita, R. (1998). CCML: Exchanging Marked-up Documents in a Networked Writing Classroom, *International Journal of Computer Assisted Language Learning*, Vol.11, No.2, pp.201-214.
- [16] Ogata, H., Feng, C., Hada, Y., and Yano, Y. (1999). Computer Supported Proofreading Exercise in a Networked Writing Classroom, *Proc. of ICCE 99*, Vol.1, pp.414-417.
- [17] Ogata, H., Feng, C., Hada, Y. and Yano, Y. (2000). Online Markup Based Language Learning Environment, *International Journal of Computers & Education*, Vol.34, No.1, pp.51-66.
- [18] Sato, M., Kano, C., Tanabe, K., & Nishimura, Y. (1994). *Practical Japanese Compositions*, Bonjin-sya. (in Japanese)
- [19] Scardamalia, M., & Bereiter, C. (1986). "Research on written composition," M. Wittrock (Ed.), *Handbook of research on teaching* (third ed.), Macmilian, pp.778-801.
- [20] Schwind, B. C. (1995) "Error analysis and explanation in knowledge based language tutoring," *Computer Assisted Language Learning*, vol.8, no.4, pp.295-324.
- [21] Shimano, H. (1986). *Handbook of proofreading*, Miki Publishers, Japan, in Japanese.
- [22] Yano, Y., Miyoshi, K., & Hayashi, T. (1993). "Development of a game style environmental ICAI system for Kanji idiom learning," *Proc. of International Conference on Computers in Education '93*, pp.411-413.
- [23] Yano, Y., Ogata, H., Sakakibara, R., & Wakita R. (1997). "CoCoA: Communicative correction assisting system for learning Japanese compositions", *Transactions of Japanese Society for Information and Systems in Education*, vol. 14, no.3, pp.21-28. (in Japanese)

Appendix A: XCCML document in figure 4.

```
<?xml version="1.0" encoding="Shift_JIS"?>
<!DOCTYPE XCCML SYSTEM "XCCML.dtd">
<XCCML>
<HEAD>
<Title string="Overview of CoCoA"/>
<Editor name="Hiroaki Ogata" email="ogata@is.tokushima-u.ac.jp"/>
<Author name="Yoshiaki Hada" email="hada@is.tokushima-u.ac.jp"/>
</HEAD>
<BODY>
<CENTER><IMG width="128" height="128"
src="image001.gif"/></CENTER><CENTER><H2>Overview of CoCoA</H2></CENTER> <H4>
CoCoA <Annotate level="3" comment="What is short for CoCoA?"/> is<Insert string="a" level="3"/>
computer supported language learning system based ononline markup.<BR/>
It<Replace string="allows" level="3" comment="Please add &#34;s&#34;."> allow </Replace> students
and teachers to exchange<Replace string="marked-up" level="3"> mark-uped</Replace> document via
<Replace string="Internet" level="1" comment="You had better use a captal letter.">
internet</Replace>,<BR/> and its environment is very similar to a real one in which people use paper<BR/>
and pen. <MoveTo fromid="1" level="2" comment="Please note the initail letter."/>This paper also proposes
CCML (<U>C</U>ommunicative<BR/> <U>C</U>orrection <U>M</U>ark-up
<U>L</U>angua<Replace string="that" level="3">ge</Replace> who is based on SGML<BR/>
(<U>S</U>tandard <U>G</U>eneralized <U>M</U>ark-up <U>L</U>anguage)<BR/>
<MoveFrom toid="1" level="2">in order to record and exchange corrected compositions with marks and
comments.</MoveFrom></H4></BODY>
<CLOSE>
<Comment/></CLOSE>
</XCCML>
```