Development and Utilization of TUFS Language Modules

Chunchen Lin* Yuji Kawaguchi**

Abstract: This paper describes the developmental procedures and theoretical backgrounds of the “TUFS Language Modules” created by the Tokyo University of Foreign Studies in the framework of the 21st Century Center of Excellence (COE) Program. The most salient feature of the TUFS Language Modules lies in the fact that language materials cover 19 different countries and regions of the world and are designed in the same theoretical frame of reference. Each language unit contains pronunciation, dialog, grammar, and vocabulary modules. Most of the materials are XML-tagged, which allows us to easily revise present materials and effectively create new language units. At present, the materials of the TUFS Language Modules have been further divided into smaller elements according to the SCORM 2004 standard, so that by using a SCORM-compliant e-Learning system, we can develop, for instance, teaching materials of exchanging greetings in 18 different languages. We will also review how Japanese modules have been utilized so far in the classroom.

Keywords: TUFS Language Modules, XML, SCORM, COE program

1. What are the TUFS Language Modules?

The language structure is divided into modules that have several actions, and while the modules are independent to some extent, they also function together on the whole. This is the concept of a language module. Generally, the language structure is comprised of four modules: pronunciation, grammar, vocabulary, and dialog. This module hypothesis was adopted by Fodor and has a certain level of support in generative grammar and second language acquisition studies that are influenced by it.[1]

The development of the TUFS Language Modules covering 18 different languages would have been impossible without the foundations of linguistic theory and second language acquisition. In fact, the Pronunciation Modules have been designed based on phonetic and phonological theory, the Dialogue Modules are built on notional and functional syllabus, and the Vocabulary Modules are classified in terms of their semantic categories based on Japanese lexical taxonomy as elaborated by the National Institute for Japanese Language.

* Tokyo University of Foreign Studies
** Tokyo University of Foreign Studies

[1] Fodor, J.A. 1975. *The modularity of mind*. MIT Press. Chomsky, N. 1980. *Rules and representations*. Columbia University Press.

The research field that implies the systematic integration of computer science, linguistics, and language education is called Linguistic Informatics. It was so named by the COE Program (2002–2006) and later taken over by the Global COE Program (2007–2011).[2]

In the TUFS Language Modules[3] [1], all the data on teaching material[4] besides the pronunciation module practice exercises are defined using XML; therefore, it is possible to create new teaching materials by creating and adding teaching materials data matched to the XML standard. This is advantageous in that it does not warrant the creation of a separate web design. We commenced the development of the TUFS Language Modules in 2003, beginning with the pronunciation module practice exercises, and have gradually released the modules in the form of a web page. At the same time, we have developed the e-Learning system for the operation of the respective learning courses.

SCORM, a standard for guaranteeing the interoperability of the teaching materials and the Learning Management System (LMS), was released by a Japanese non-profit organization, e-Learning Consortium Japan. We have been gradually making the TUFS Language Modules compliant with the SCORM 2004 standard since fiscal year 2007. Currently, we are operating them in Moodle, an LMS system, which is SCORM 2004 standard compliant. Below, we will first discuss the design of the TUFS Language Modules and their compliance with the SCORM 2004 standard; then, we will explain examples of operation of the Japanese dialog module; and finally, we will talk about the future development of the modules.

2. Design of TUFS Language Modules and Compliance with the SCORM 2004 Standard

The practice exercises in the pronunciation modules have been developed as a web page; however, for the other modules—the dialog module, grammar module, and vocabulary module—the data is defined by XML, and a mechanism to generate web teaching materials was adopted as required. Here, we explain the methods for the development of each of the teaching materials as well as the composition of the developed teaching materials and the application program

[2] Usage-Based Linguistic Informatics, URL: http://www.coelang.tufs.ac.jp/english/index.html; Corpus-based Linguistics and Language Education, URL: http://cblle. tufs.ac.jp/index.php?id=21.
[3] URL: http://www.coelang.tufs.ac.jp/modules/
[4] This refers to the elements comprising the teaching materials. For example, the text of example sentences from dialogs and its translation, audio, and video are elements, and when they are combined, they constitute teaching materials. It corresponds to the term "asset" in SCORM; however, since it is not a web page or assessment object, the definitions of the two terms are not the same.

interface (API) used to achieve compliance with the SCORM 2004 standard, and introduce the course formulation using Moodle.

2.1 Practice Exercises in the Pronunciation Modules

In the practice exercises in the pronunciation modules, we avoided specialized phonetic terms as much as possible and took care to develop easily comprehensible teaching materials that use simple explanations [2, 3]. For all 17 languages, we established three goal stages—"for survival," "for fluent communication," and "for pronunciation as good as native speakers"—and prepared several teaching topics for each of the goals. Each of these topics consists of several web pages featuring the explanation, practice problems, and answers to the practice problems, which together constitute one learning unit.[5] In order to comply with the SCORM 2004 standard, we implemented processing which calls an Initialize and Terminate API, and made the modules compliant with the SCO[6] teaching materials. Figure 1 shows the French pronunciation module being operated in Moodle.

2.2 Dialog Modules

For the development of the dialog modules, we adopted a "functional syllabus" that provides learning content that is common to each language. First, we identified example sentences presented in the existing Japanese language teaching materials and conducted a similar survey of Korean and Chinese teaching materials to verify whether these example sentences were effective in the acquisition of other languages as well [4]. Based on the results of these surveys, we created a functional syllabus consisting of 40 language functions. The dialog modules include not only text information such as dialog sentences and their translations but also multimedia information such as audio and video of the text. The formats of the dialogs were largely the same regardless of language, and we adopted the method of creating the teaching materials data in the XML format and automatically generating the teaching materials page in script language [5]. The teaching materials data of each dialog is composed of the head and the body. The head is the part that contains information other than the dialog, including the creator and informant of the dialog, the status of the dialog, etc. The body includes the text of the dialog itself. The body also contains

[5] This refers to a collection of several learning pages, which constitute one learning unit.
[6] This is the lowest level of granularity of learning resources that can be tracked by an LMS using the SCORM run-time environment.

information such as the line, sentence, target language, translation, part of speech, grammar, vocabulary, etc. Figure 2 shows the schema expression of XML that composes the teaching materials data of the dialog module.

Currently, we have released two types of teaching materials: teaching materials for use in the classroom that teachers present to learners using projectors in classrooms and teaching materials to use for learning with which learners teach themselves at home.

We have prepared four screen patterns for teaching materials for use in the classroom. The teachers can use these screens for different purposes. For example, the students can use pattern 1 to listen to and remember the individual speeches by the participants in the dialog. Next, with pattern 2, they can look at the individual speeches while practicing pronouncing and writing them. Then, they can go on to pattern 3 to practice listening to and remembering the overall dialog, and finally use pattern 4 to do role-plays, etc.

On the other hand, the teaching materials for the use of learners stipulate clear learning goals, display teaching materials data specialized for those goals on the learning pages, and are designed so that one learning unit is composed of several pages. The currently prepared pages for learners consist of four types of teaching materials that combine the following skills essential for language learning: the information acquisition skills of “seeing and listening” and the information transmission skills of “speaking and writing” [6]. Currently, pattern 3 for use in the classroom is compliant with the SCORM 2004 standard and is being operated in Moodle as shown in Figure 3. The dialog module is also an SCO teaching material, which implemented the Initialize and Terminate API, so that the management status, such as the page usage time of the learner, can be checked on the Moodle management screen as shown in Figure 4.

2.3 Grammar Modules

The structure of grammar is very different for each language, and the languages covered by the grammar modules are from regions throughout Asia and Europe and have different grammatical structures. Therefore, the content of the grammar modules also inevitably differs. For example, in Japanese grammar, the explanation of the sentence pattern is central, as follows:

Taro-wa gakko ni kayoimasu (*N1 + N2 + V*).

On the other hand, the Chinese module contains the following explanation:

An abbreviated question equivalent to asking “… wa?” can be created by attaching the tone particle “呢” to a noun phrase.

In this way, an explanation of vocabulary, which is a distinctive characteristic of Chinese grammar, is central to the grammar. For this reason, it is difficult to develop teaching materials by establishing a common grammatical framework for different languages [7]. Therefore, when creating the grammar modules, we defined “pattern” taking into consideration the reusability of the teaching materials data and constructed individual grammar items—the smallest unit composing the grammar modules—from the following four elements.

- Cards: Cards summarizing the minimum amount necessary for the grammar items
- Explanation: Supplementing the cards to give detailed explanations of exceptions to grammar rules and more detailed grammar rules
- Example sentences: Collection of example sentences linked to audio
- Practice problems: Collection of practice problems related to the grammar items

We made the grammar items, comprising the above four elements, the smallest unit and wrote the teaching materials data in XML. Compared to the dialog module, the section explaining grammar has more data in a written format, and so we decided to write each element in rich text format. For example, the content of the card elements and that of the para-tag showing the XML data structure of the explanatory sentences in Figure 5 are written in rich text format, as shown in Figure 6, and the instructor is able to write anything he/she likes using accent characters. On the other hand, we defined the example sentence section in great detail, including example sentences and their translations, audio files, etc. In the card section, explanation section, and example sentence section, we implemented the Initialize and Terminate API, making the sections compliant with SCO teaching materials, just as we did in the dialog module. The practice problems of the grammar module include single-answer problems and multiple-choice problems, matching problems, fill-in-the-blank problems, and sorting problems. These problems not only implement the Initialize and

Terminate API but also comply with the SetValueAPI of SCORM 2004. For this reason, after the learner has finished studying, he/she is able to not only confirm his/her usage time but also record the grade and answers using the Moodle report function, as shown in Figure 7.

2.4 Vocabulary Modules

The objective of the vocabulary modules is to understand and learn basic vocabulary. For this reason, for all languages, we have included approximately 700 to 900 words based on the basic vocabulary of 800 words in the Japanese Language Proficiency Test Level 4 in addition to the distinctive vocabulary of each language. As shown in Figure 7, in addition to the meaning and pronunciation of the vocabulary, we established tag sets so that learners can create example sentences. Furthermore, we assigned a classification code based on the National Institute for Japanese Language's "Bunrui Goihyo" (Word List by Semantic Principles) to each vocabulary item and divided all vocabulary items into meaning classifications and scene classifications so that learners could teach themselves basic vocabulary. A scene classification is a vocabulary classification based on the scene, for example, overseas vacations, sports, at the company, the library, etc. A meaning classification is a classification based on meanings, such as pronouns, adjectives, space, color, etc. Moreover, we have prepared practice problems so that the learner can check his/her knowledge after studying the basic vocabulary with meaning classifications and scene classifications. Figure 8 is the Moodle screen. Each vocabulary item is the smallest unit comprising the teaching materials, and it implements the Initialize and Terminate API of SCORM 2004; the practice problems comply with the SCO teaching materials, which are compliant with the SetValueAPI.

3. Examples of Use of the Japanese Dialog Modules

We anticipate that overseas learners of the Japanese language will use the Japanese modules. For that reason, we translated them into English, French, Chinese (traditional and simplified), Korean, Turkish, and Mongolian in order to prepare dialog modules in multiple languages.[7] Only over 20 learners of Japanese at universities in Taiwan have actually used our "listening and speaking" model in the pages for learners of the Japanese dialog modules, which were created in traditional Chinese script and implemented in the LMS.

[7] URL: http://www.coelang.tufs.ac.jp/english/modules/index.html

Regarding the number of times used, we performed a cluster analysis, and the results showed that the average number of times used could be divided into two groups: 45 times and 100 times. Moreover, as shown in Table 1, based on the comments provided by the supervising instructors, we were able to detect a correlation between the two groups defined by the number of times used and the scores of the learners. We decided to call the group who used the modules a lesser number of times the “moderate users” and the group who used the modules a higher number of times the “active users.” Table 2 shows the average hours of use of the pages for learners by the moderate users and the active users.

	成績	
	上位 (名)	下位 (名)
グループ 1	1	6
グループ 2	13	5

Table 1: Comments of the Instructor and Number of Times Used

** 1%有意

	適当組(分)	積極組(分)
Page1	1.1	3.6**
Page2	3.6	3.6
Page3	1.4	1.5
Page4	4.9	3.4**
Page5	2.3	3.9**

Table 2: Hours of Use of the “Listening and Speaking” Pages

The “listening and speaking” model of the pages for learners is composed of the following five learning processes: (1) “understanding the overall flow of the dialog using video,” (2) “confirming the meaning of each utterance,” (3) “practicing by repeating each utterance,” (4) “confirming the dialog sentences in written form,” and (5) “role-play.” As shown in Table 2, for Pages 1 and 5, the hours of use of active users are significantly longer (at 1% significance); however, the hours of use of moderate users are significantly longer (at 1% significance) for Page 4. The content of Page 1 is listening through entire dialogs, so we can see that active users are more determined to take time to understand the dialog. The content of Page 2 consists of confirming the meaning of each utterance. The average hours of use of Page 2 by active users is the same as their average hours of use of Page 1; thus, we can see that they are setting aside approximately the same amount of time. The hours of use were low in both groups for Page 3; however, we wonder if this might be because the content of Page 3 is similar to the content of Page 2 for the learners. On Page 4, the learners confirm the meaning of each utterance with written sentences and audio. The hours of use of moderate users were higher than the hours of use of active users only for this page. Page 5 is the role-play page, and the results showed that moderate users spent more time using this page. By using an LMS in this way, we can not only get the learners to simply answer practice problems and evaluate their scores but also investigate the usage status of the learners

and subject it to analysis and evaluation.

4. Conclusion

The TUFS Language Modules also include an interface for creating teaching materials. Therefore, in order to simply create teaching materials, general language teachers can also input teaching materials data from the web interface without having to worry about tags, etc. We believe that now that we have made them compliant with the SCORM 2004 standard, teachers will be able to utilize the teaching materials data even more freely than before. In the future, we plan to enhance the content of the TUFS Language Modules in two directions. The first direction is to introduce more languages so that languages apart from the current 18 languages will be included. The other direction is diversification. For example, this includes the PuTongHua (standard Mandarin) used in Beijing, Suzhou, and Taiwan, and the French used in the province of Quebec in Canada.

References

[1] Kawaguchi, Yuji. “Foundations of Center of Usage-Based Linguistic Informatics (UBLI),” in Yuji Kawaguchi et al. (eds.) *Spoken Language Corpus and Linguistic Informatics*, John Benjamins, pp. 9–33 (2006).

[2] Kigoshi, Tsutomu. “The Creation of the TUFS Pronunciation Module,” in Yuji Kawaguchi et al. (eds.) *Linguistics Informatics State of the Art and the Future*, John Benjamins, pp. 316–332 (2005).

[3] Nakata, Shunsuke, Shin Abe, and Hajime Mochizuki. “Design and Development of Multilingual E-learning Materials, TUFS Language Modules – Pronunciation,” Proceedings of IASTED International Conference on Computers and Advanced Technology in Education, pp. 591–596 (2003).

[4] Yuki, Kentaro, Kazuya Abe, and Chunchen Lin, “Development and Assessment of TUFS Dialogue Module – Multilingual and Functional Syllabus,” in Yuji Kawaguchi et al. (eds.) *Linguistics Informatics State of the Art and the Future*, John Benjamins, pp. 333–357 (2005).

[5] Lin, Chunchen, Kazuya Abe, and Kentaro Yuki. “A Method for Developing Multilingual E-Learning Material Based on Functional Syllabus and XML Scripting: Dialogue Module in TUFS Language Modules,” National Conference of the Institute of Electronics, Information and Communication Engineers (2002).

[6] Naganuma, Naoyuki and Chunchen Lin. “Development of Learning Module

for the Web-based Language Learning," World Conference of SITE 2004, pp. 4003–4004 (2004).
[7] Minegishi, Makoto. "Developing Grammatical Modules Based on Linguistic Typology," in Yuji Kawaguchi et al. (eds.) *Spoken Language Corpus and Linguistic Informatics*, John Benjamins, pp 331–348 (2006).

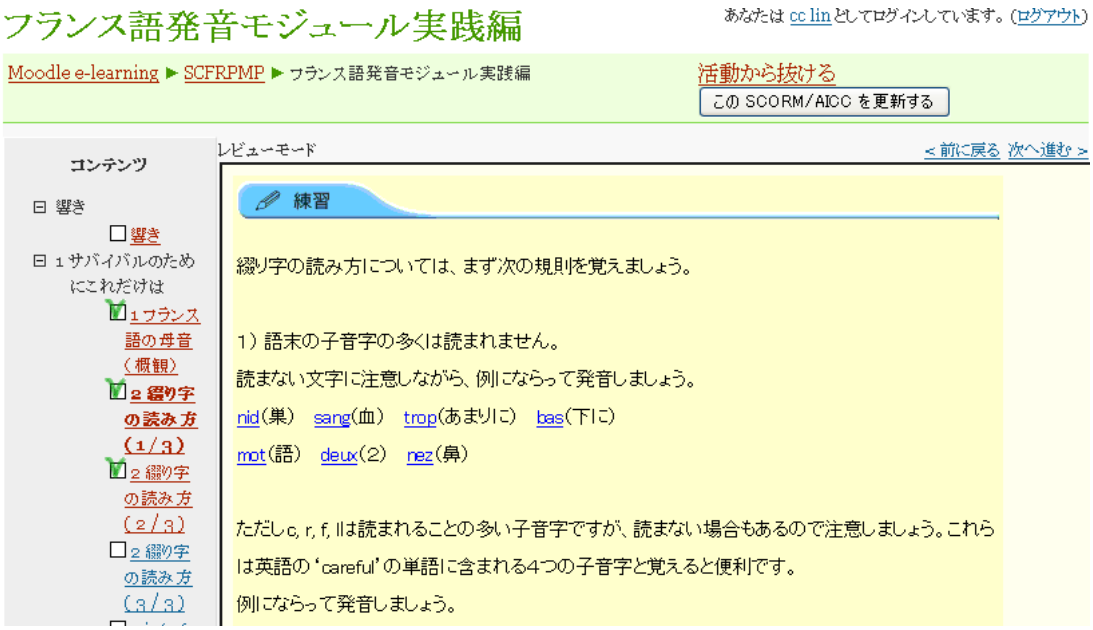

Figure 1: Operating the Pronunciation Module Practice Exercises in Moodle

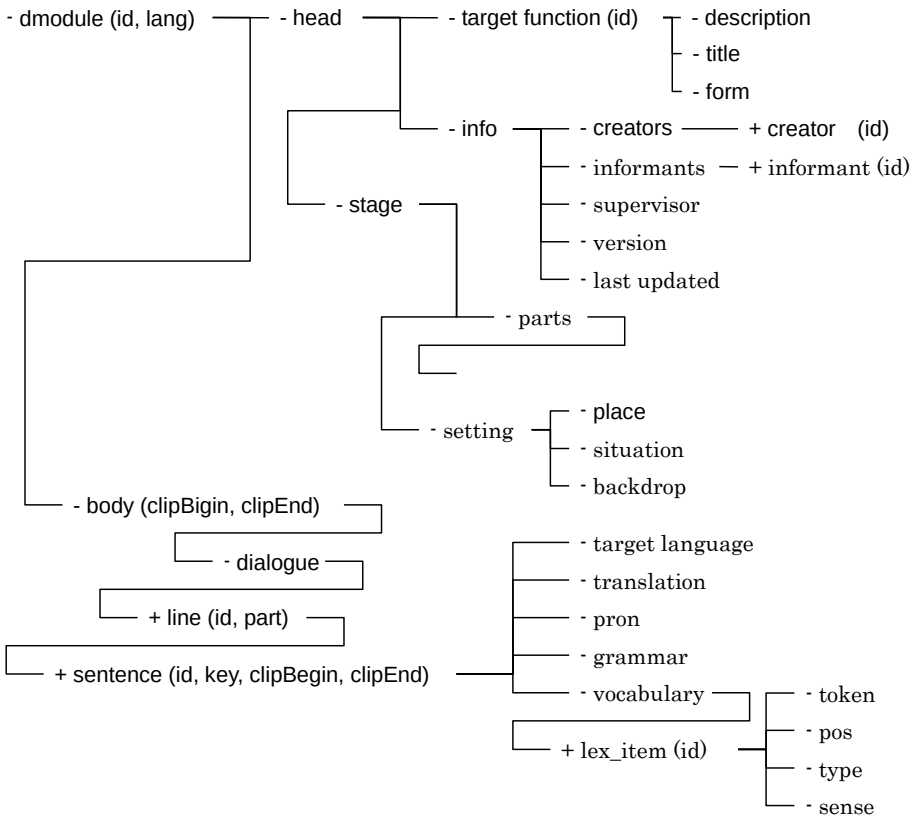

Figure 2: XML Schema for Dialog

Figure 3: Dialog Module Teacher

挨拶する	完了	PT15H15M51S		トラック詳細
感謝する	完了	PT3H3M14S		トラック詳細
注意をひく	完了	PT2H2M38S		トラック詳細
自己紹介する	完了	PT2H2M2S		トラック詳細
謝る	完了	PT2H2M18S		トラック詳細

Figure 4: Information in Moodle about Use of Dialog Teaching Materials by the User

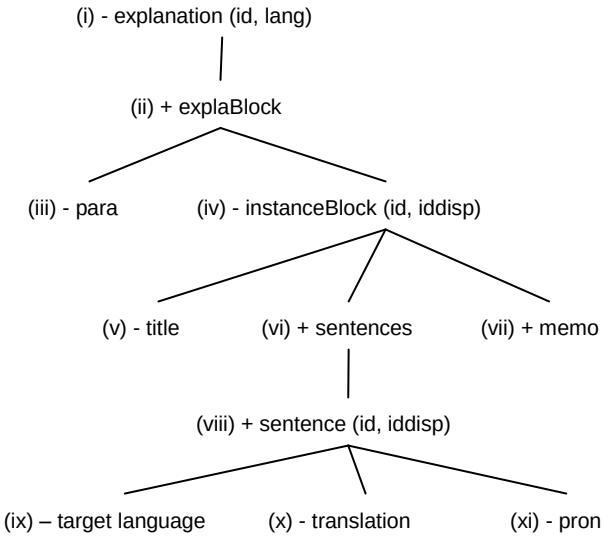

Figure 5: Data Structure of the “Explanation” in the Grammar Modules

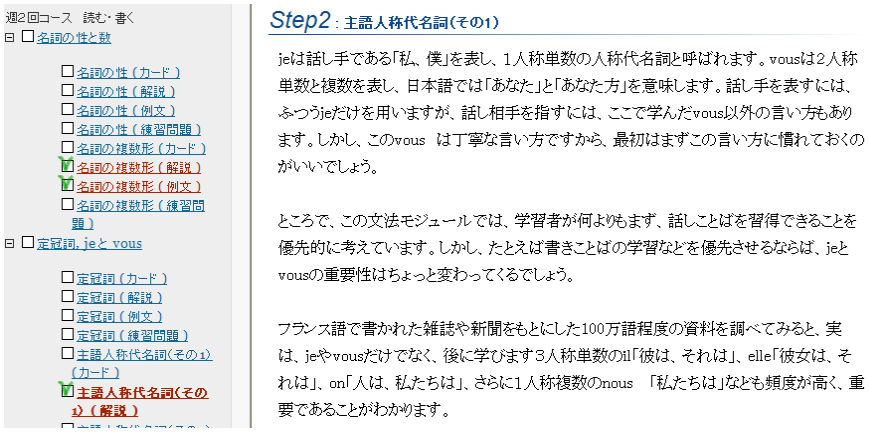

Figure 6: Para-Tag Explaining French Grammar Teaching Materials

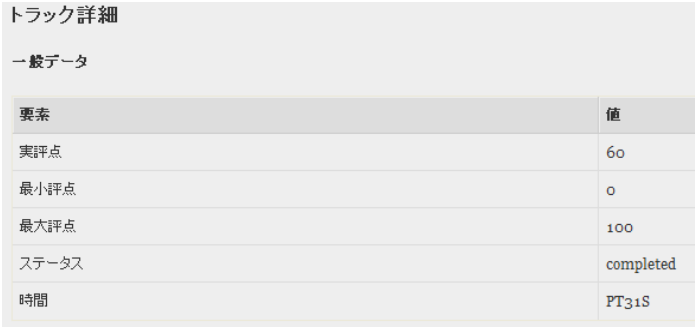

Figure 7: Track Details of Grammar Practice Problems

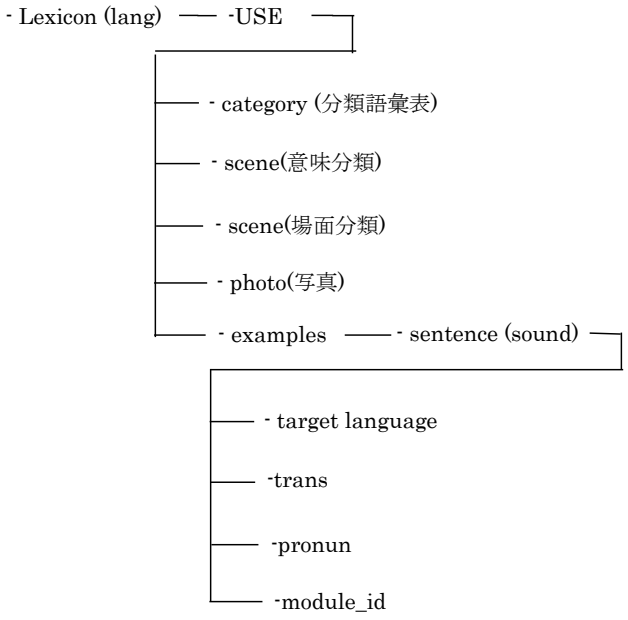

Figure 7: XML Schema for Vocabulary Module

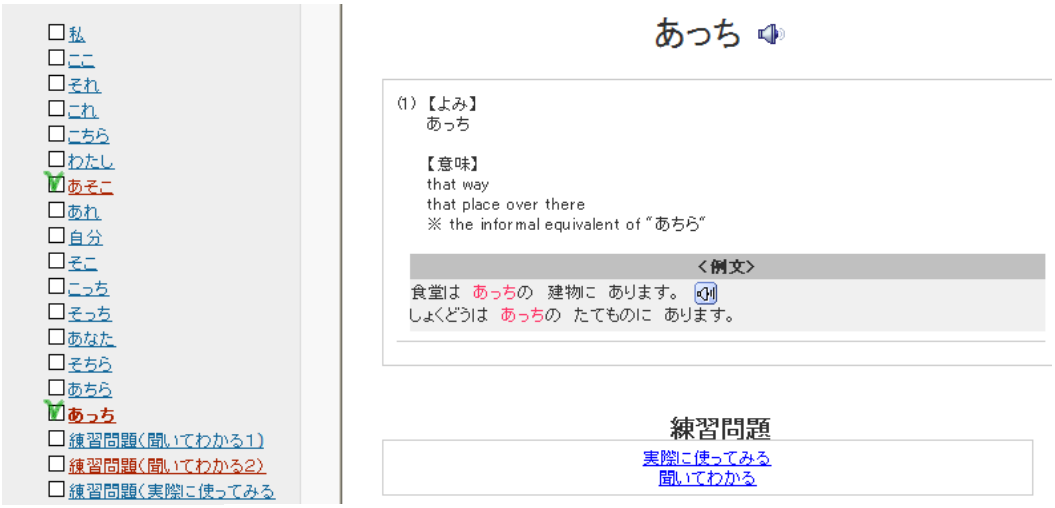

Figure 8: Vocabulary Module Moodle Screen