Abstract—Recent years have seen widespread adoption of the Internet for language teaching and learning. Interactive systems on the World-Wide Web provide useful alternatives to face-to-face tuition, and both teachers and learners can benefit from the exercises available. However, although there is a wealth of suitable material, it is hard to find because it is scattered around the web. Moreover, teachers are restricted by the material that is available, and cannot provide their own.

To tackle these problems we have constructed a digital library of language learning exercises that presents students with different kinds of exercise, and also lets teachers contribute new material. We first reviewed existing language learning systems on the web in order to develop a taxonomy of exercise types used for language activity. A prototype, ELLE, based on this taxonomy, provides various kinds of interactive exercises using material that teachers submit. The system has been evaluated by practicing language teachers.

Index Terms—Computer Assisted Language Learning, Digital Libraries, Online Learning, Teacher Participation

I. INTRODUCTION

Although computers began to be used for language teaching during the 1960s [1], the unprecedented growth of the Internet in the early 1990s altered their role in two far-teaching ways. First, the World-Wide Web provided an opportunity to augment, or even replace, face-to-face teaching by learning activities in the form of online exercises that were readily accessible to learners outside the classroom. Second, a plethora of language learning material became available for use by teachers and learners: newspapers, articles, stories, pictures, audio, and video.

There are countless language learning resources on the web: news sites like the BBC; commercial learning software; and dedicated language learning websites supported by educational establishments, community organizations, and individual teachers. Websites specifically designed for language learning commonly provide the learner with interactive exercises, and take advantage of multimedia technology to add interest. However, all of them contain teaching material that is fixed, determined by the system’s operator, and provide no facilities for teachers to participate in the learning activity by supplying their own material. This means, for example, that they are restricted to teaching one particular language.

We surveyed existing systems to discover what is available, determine how they are organized, and assess their strengths and weaknesses. Then we developed a taxonomy of seven exercise types that are used for language activity. Taking advantage of this, we built an extensible digital library environment, ELLE, which provides diverse kinds of interactive practice exercises, and draws teachers into the learning experience by allowing them to submit materials for exercises. It supports activities in the areas of vocabulary, grammar, reading and listening; and contains several exercise types for each one.

Section 2 introduces language learning systems that are already available on the web. Section 3 introduces several different aspects of language learning, and briefly describes the seven exercise types that we have identified. Section 4 describes ELLE in detail, including the method that teachers use to submit new exercise material. Section 5 describes an evaluation of the system by language teachers, and Section 6 discusses how the work fits in with current developments in educational digital libraries.

II. LANGUAGE LEARNING SYSTEMS ON THE WEB

The web has become a popular and effective place to learn and teach foreign languages. It transcends limitations of space and time: learners are able to learn languages in their spare time anywhere the Internet is available. It gives learners free choice of difficulty level and exercise type [2]. It provides privacy: learners do not lose face when they make mistakes.

There are a huge number of websites dedicated to language learning and teaching—Google returns about a million hits for the term “language learning”—including both online exercises and offline software. The former are generally open access (though they may require registration); the latter are often commercial.
Online resources
Online exercises allow learners to practice language constructions that involve vocabulary and grammar, and improve their skills in listening, reading and writing. They often take the form of quizzes, puzzles or games, but the boundaries between these categories—especially puzzles and games—blur. Learners can select from a given set of options, or construct their own answer. Text, graphics, audio and video can be used as the medium for presenting questions and accepting answers. Macromedia Flash and Java applets are popular ways of providing attractive and helpful feedback. Quizzes are the most popular type of resource. Puzzles and games add fun and help motivate learners; many reflect classic games such as pinball, hangman, crossword, and memorization. Some come with various difficulty levels—beginner, intermediate, and advanced. Some word puzzles minimize the difficulty by making the definition of the word, or the word itself, available to the learner.

Online courses are more extended and formal than individual exercises. They resemble those taught face-to-face in a classroom, except that they are given by a virtual teacher rather than a real one. Teachers organize material into sections that give a detailed explanation of some aspect of language (for example, a particular point of grammar), along with sample constructions and illustrative exercises. Answers are usually assessed and corrected by the system itself rather than by human markers.

Offline resources
Many language learning systems can be purchased over the web, and some websites provide non-commercial language software. Commercial systems usually have an abundance of learning materials and a sophisticated learning environment that integrates video, audio, graphics and animation into language teaching and learning design. Learners can watch animated pictures, listen to audio clips, record their own speech, and so on. Examples include the Language Adventure system, Rosetta Stone’s English Learning software, WordZap (which is freeware), and Spell It (shareware).

An online quiz generator
Some websites are hosted by professional language teachers who organize linguistic exercises and produce material for them. However, websites that allow teachers from elsewhere to contribute language material are rare. We found just one,¹ which provides an online multi-choice and bilingual quiz generator that teachers use to produce online quizzes or printable ones for classroom use. Teachers can contribute their quiz data to the central website.

III. LANGUAGE ACTIVITIES AND EXERCISE TYPE

Language learning has several different aspects [3]. Those supported by existing web-based systems can be classified into seven categories: vocabulary, grammar, reading, listening, writing, pronunciation, and speaking. Vocabulary activities help learners build a stock of words and master their meanings. This is the activity with which most beginners start their language learning. Grammar activities help learners study a language’s structural rules. Reading activities helps learners understand written materials. Listening activities help learners understand what is heard and improve their listening skills. Writing, pronunciation, and speaking activities assist learners in developing these skills.

To support each of these language learning activities, existing language learning systems use a variety of different types of exercise. We have analyzed these types to come up with a taxonomy to use as a basis for implementing a more generic system that gives a cleaner distinction between the general structure of an exercise and the specific topical material that it presents to learners. Most exercise types in our taxonomy could be used for all of the language learning activities identified above, although some are specific to one particular activity (and in this respect the taxonomy is not completely satisfactory).

Our taxonomy distinguishes seven types of exercise.

Multi-choice exercises take the form of a question and a set from which the correct answer must be selected. This is a traditional language learning exercise that is widely used for grammar, listening, and reading. Questions and answers can be phrased as text or images.

Matching exercises require the learner to find two matching items that have the same dictionary definition, or are opposites—for example, antonyms. They can easily cater for learners of different levels, and are widely used for vocabulary study. Examples include

- matching picture and word
- matching text and pronunciation
- matching word and definition
- matching words together.

Permutation exercises require the learner to sort items into some order. Items can be letters, words, events (perhaps represented by video clips), and phrases or sentences.

Fill-in-the-blank exercises, in which the learner must fill in gaps in a question, are widely used for learning grammar constructions and for developing reading and listening skills.

Type-the-answer exercises require the learner to construct the answers and type them in. For example, an audio clip might be presented, and learners must type in what they heard—words, sentences, a conversation, song or story.
Spelling exercises require learners to construct a word correctly. These are used for vocabulary learning, or for mastering easily misspelled words.

Category exercises require the learner to place things (e.g. words) into groups according to their type. They are widely used to help master vocabularies that are easily misused (e.g. countable and uncountable words).

IV. ELLE: AN EXTENSIBLE DIGITAL LIBRARY FOR LANGUAGE LEARNING

ELLE, an extensible digital library for language learning, helps learners improve their language skills through different kinds of interactive exercises, and also provides an interface for language teachers to contribute exercise materials. To help learners maintain a high level motivation it aims to provide exercises that are attractive, interesting and playful.

Activities and exercise types

Of the seven language activities in the previous section, ELLE supports the first four: Writing, pronunciation, and speaking are not covered because in these domains it is difficult to automatically determine the correctness of answers. Writing is well supported by online discussion forums and chat rooms; the other two require advanced audio technologies such as speech recognition.

Within each activity several of the exercise types identified above are implemented. For vocabulary learning ELLE provides permutation exercises involving scrambled words, several kinds of matching exercise (words against words, words with their definitions, and words against pictures), category exercises, multi-choice exercises, and spelling exercises. For grammar, teachers can explain a rule and provide practice exercises of several types: multi-choice, fill-in-the-blanks, permutation exercises in the form of scrambled sentences, and matching exercises that match the beginning and end of a sentence. For reading, ELLE provides fill-in-the-blanks and type-the-answer exercises that test whether learners understand an article by asking questions related to its content, or to vocabulary and grammar rules that appear in the article. For listening, type-the-answer exercises require learners to enter the words or sentences that they have heard, and fill-in-the-blanks and permutation exercises ask learners to insert words that are missing in an article, song, story, or conversation.

Figure 1 shows the learner’s home page, which displays individual exercise types grouped under by activity. The layout is plain because effort was concentrated on making individual exercises as attractive as possible. Furthermore, the exercise types that were implemented are by no means exhausted; more could easily be generated.

Figure 2 shows the Exercise List page for the Matching Words Together exercise type.

Levels of difficulty

Learners differ in their linguistic ability and benefit most from exercises whose difficulty matches the learner’s own level [4]. ELLE provides learners with the opportunity to select exercises and materials that are appropriate to them. Each exercise focuses on a particular language feature, and the exercise materials are classified into beginner, intermediate and advanced levels.

Examples

When a particular exercise type is selected from the home page shown in Figure 1, individual exercises are listed. Figure 2 shows the exercises under Vocabulary study. Matching words together. At this stage there are only two exercises (“antonyms” and “irregular plurals”), but teachers can add more to the list as described below.

Figure 3 shows one of the exercises, Matching antonyms. At the top right the user selects the difficulty level. The large central working panel contains mixed pairs of antonyms. When the user correctly matches two of these, they move to the results panel on the left. The navigation buttons at the lower right allow the learner to restart the exercise, and move on to the next one.

Submitting exercise material

ELLE’s exercise material is intended to be contributed by language teachers. When submitting material, teachers need to determine the target exercise and difficulty level. They access the system through a “teacher’s home page” that is structured just like the learner’s home page in Figure 1 except that each menu item has two links, one to view a sample exercise and the other to submit material. Suppose a teacher is submitting material for the Matching Words Together exercise type. First, they select an existing exercise name (“antonyms” or “irregular plurals” in this case) or types a new one. Then
they choose a difficulty level: beginner, intermediate, or advanced. Finally they enter the material itself into the large box. In this case paired words are entered in the format “#word1 : word2”—for example, “#narrow : wide”. It is not necessary to specify in advance how many pairs will be entered. Instructions for entry are given above the box. Exercise material can be submitted in the form of plain text, HTML, audio, and images. Teachers can type text or HTML into text boxes, and upload audio and images from pre-existing files. Exercises can have five up to different elements: instructions, level, hint, check and feedback. The implementation of these elements varies from one exercise to another.

**Instructions** explain how to do the exercise, for example “Click the matching words” in Figure 3. Also, tool tips are provided throughout ELLE to explain to the user the functionality of interface components.

**Level** elements specify a difficulty level (beginner, intermediate or advanced) for each exercise, determined by the teacher who submits the material.

**Hints** are provided for many exercises by a “I need a hint” button, in a way that varies from exercise to exercise—for example, *Matching Words* exercises highlight the next letter.

**Check** elements resemble judge correctness according to the answers provided by the teacher. ELLE provides both active and passive checking mechanisms; the former rejects wrong answers immediately, while the latter marks the answers only when a “check answers” button is clicked.

**Feedback** elements present statistical information about the learner’s performance and encourage them to proceed to the next exercise.

![Figure 3 The Matching Words Together exercise (Matching antonyms)](image)

**Exercise types**

Section 3’s taxonomy of seven exercise types is the key to the exercise submission system. Teachers can supply new material for existing exercises, and create new ones, but they are constrained by the structure of the exercise types that ELLE provides: multi-choice, matching, permutation, fill-in-the-blank, type-the-answer, spelling, and category. Spelling exercises are for vocabulary study; the other six can be used for all four activities. Some exercise types have sub-types. For example, *Matching Word and Picture* and *Matching Words Together* are sub-types of the *Matching* exercise type. A different submission page is provided for each combination of activity and exercise type, in other words, for each link in Figure 1.

There is insufficient space to describe in details how ELLE presents each of the exercises that have been implemented. Instead we briefly review the types that are provided.

Multi-choice exercises allow teachers to use different media—text, image, audio, video—to present questions and accept answers. Teachers must define several candidate answers, typically two to five. For text questions, teachers can provide explanations for each answer, which will be seen by learners.

Multi-choice exercises allow related items to be paired—for example a verb and its irregular present tense, or the beginning and end of a sentence, or, for reading activities, language expressions that have the same meanings in an article, or for listening activities, words with the same pronunciation. Different media can be used, for example, matching words and pictures.

Permutation exercises can involve rearranging letters, words, sentences, paragraphs of text, images, audio clips, or events in a movie. For example one might ask learners to sort images into the order in which they are described in an article, or reorder audio clips according to what is heard.

Fill-in-the-blank exercises are widely used for teaching language grammar. ELLE presents a description of a particular grammar rule and allows learners to drag answers from a closed list and slot them into the answer, or type their own words into the blanks.

Type-the-answer exercises are widely used in classroom teaching. Learners construct their own answer—in which case judging their correctness presents automated language learning systems with a serious challenge. ELLE incorporates both versions, and in the latter case allows the teacher to define alternative answers and uses whitespace normalization and case folding when matching.

ELLE’s spelling exercises are game-like and resemble classic games such as pinball, hangman, crossword, Tetris, and memorization. Collaboration and competition can be built into such exercises, challenging learners to outdo each other or forcing them to discover answers cooperatively.

Category exercises can be used for grammar, listening, and reading activities, although most examples on the web are for vocabulary study. Apart from words, items that can be categorized include phrases, sentences, articles, audio clips and video clips. For example, learners might sort audio clips into categories such as sport, weather, business, or categorize words whose vowels have the same pronunciation, or sentences according to their tense.
A Digital Library of Language Learning Exercises

Implementation
ELLE is written in Java and follows the client-server model. The client uses HTML pages to interact with learners and teachers, and sends requests to the server. The server uses PHP scripts to handle the requests, and sends responses to the client. Exercises are implemented as Java applets. For each exercise type and language activity there is an applet that implements all interaction with the learner. Only by using specially-written applets were we able to provide an interesting variety of interaction styles.

Internal data is divided into metadata and materials, both of which are stored in XML documents. Metadata describes the exercise types and sub-types, and contains exercise instructions, submission instructions, exercise lists and the pathname of the Java applet that implements the exercise. Materials define the exercise content, which can be entered online by teachers, and includes questions, answers, and related resources (e.g. image or audio files). Each exercise type or sub-type has one XML document that contains the metadata and a set of XML documents that contain the exercise materials, each pertaining to a particular topic.

The metadata document for the Scrambled Words exercise type is shown in Figure 4. Name names the exercise type, instructions inform learners how to do the exercises, index is used to uniquely name uploaded files, submissionInstruction guides teachers in making submissions, applet gives the path and size of the exercise applet, and exerciseList lists exercise names and their corresponding exercise material file.

The exercise material document for a Multi-choice exercise is shown in Figure 5. It specifies the author and creation time of the exercise. Topic elements tell what the material is about; level gives the level; and question elements describe the individual questions that make up the exercise. These vary from exercise to exercise. In this case there are four sub-elements: description gives the text of a question, picture is the file name of an image; answer is the answer, and correctAnswer is the sequence number of the correct answer.

V. EVALUATING ELLE

Four language teachers participated in an evaluation of ELLE. One experimented with the system in a usability lab, two were interviewed using screenshots to convey the operation of the system, and the fourth tested ELLE remotely and supplied feedback via email. Participants filled out a detailed questionnaire and answered 15 verbal questions.

The teachers found it reasonable to organize the language learning activities into vocabulary, grammar, reading and listening categories, once we explained why pronunciation, speaking, and writing activities were omitted, but suggested that there should be exercises that combined these activities. They confirmed the taxonomy of exercise types developed in Section 3 and did not propose any further types, although they did suggest some variants. They agreed that the three difficulty levels are appropriate for public use on the web.

The teachers felt capable of making submissions themselves, and one used the system to submit material for some exercises. However, the instructions could be clarified in both content and presentation. They thought it was attractive to be able to use HTML tags to format material, but questioned whether many language teachers could do so without assistance. Although the system can accept audio clips and image files as exercise materials, two teachers mentioned that it is not easy to produce image and audio files, and using existing web material could violate copyright. They wanted the system to provide tools for generating media files.

Teachers greatly preferred the passive method of checking answers, because active checking may encourage guessing and makes exercises too easy, even for beginners. They understood the difficulty of checking questions with many possible answers, and preferred
exercises with closed answer sets. They strongly criticized ELLE’s “hint” feature on the grounds that learners may use it to complete exercises without thinking about them first. They thought the system makes insufficient use of feedback for motivating language learning, and that it should provide more information about the learner’s performance and progress. Overall, the evaluation suggested that ELLE provides a suitable environment for learners to improve their language skills. The teachers liked the interface and the various features that exercises provide. They particularly liked how instructions are given in multi-choice exercises, and that learners can control the number of missing words in certain exercises. The teachers’ submission system is easy to use, but the instructions it presents and the way it presents them should be reviewed, and more information should be provided on how HTML tags can be used to format question text.

![Exercise materials document for a multi-choice exercise](image)

### VI. IS ELLE REALLY A DIGITAL LIBRARY

ELLE provides an integrated collection of structured learning exercises, with a submission system that makes it easy to extend the collection by adding new exercise material. In this sense it seems clear that it does indeed share the important characteristics of digital libraries. For example, it comfortably satisfies a common definition of a digital library as a focused collection of digital objects, including text, video, and audio, along with methods for access and retrieval, and for selection, organization, and maintenance of the collection [5].

A key feature of digital libraries, and one that distinguishes them from ordinary websites, is the fundamental role that metadata plays in the organization of the documents. It is metadata that allows libraries to be augmented with new documents that immediately enjoy first-class status in the collection—as distinct from ordinary websites, in which links to new documents must be manually woven into the existing structure. ELLE shares this important feature. When teachers define new exercises they are entering metadata that allows the new material to be integrated into the collection fully and immediately.

On the other hand, in its present form ELLE makes no use of existing digital library infrastructure or standards. The IEEE LOM standard for Learning Object Metadata [6] could certainly be used to describe a particular instantiation of ELLE so that it could be catalogued as a resource and made available from other educational digital libraries. However, LOM could play no useful role within ELLE itself. Although ELLE exercises are defined in terms of metadata, these are far finer-grained, and far more tightly interwoven with the operation of the exercise, than LOM would allow.

SCORM, the Sharable Content Object Reference Model, aims to foster the creation of reusable learning content as “instructional objects” within a common technical framework for computer and web-based learning. SCORM lessons can embed ShockWave or Java Applet files, and it is likely that ELLE applets could be embedded in an overall SCORM framework, although this would require further study. However, it is not clear that doing so would serve any useful practical purpose in simplifying or standardizing the implementation, because the necessary communication between exercise metadata documents and exercise materials documents would be hard to arrange except in an ad hoc way.

These reflections may indicate that our present notion of digital library is perhaps too narrow and conservative, and needs broadening to accommodate this kind of structured information system. In any case, there is no question that conventional digital libraries have an important role to play in language learning, both as a source of exercise materials and as a basis for novel kinds of comprehension and structural exercises. Further developments are planned, beyond the scope of this project, to integrate ELLE with digital library technologies in new ways.

### VII. CONCLUSION

ELLE shows how different exercise types can be placed within a common framework. Two features distinguish it from thousands of language learning systems on the web: the wide variety of exercises it supports and the source of
the material it presents. Exercise types are characterized by (a) a Java applet that presents the exercise to the learner, and (b) a metadata document that describes the structure of the information required to present the exercise. Language teachers can submit their own exercise materials, which are captured in a second metadata document. This means, for example, that teachers can devise exercises in different languages. Admittedly the potential for multi-lingual use is restricted by the fact that the interface to the system is presented in a single language, but we plan to capitalize on earlier work [7] to extend it so that translators can easily produce different language versions, and keep them updated as the system evolves.

There are several practical issues intrinsic to a language learning system that ELLE does not address. Different learners access the web using a variety of computers with a variety of browsers and connection speeds, and this variability should really be taken into account in designing online exercises. For the teacher’s online submission system, security is an important question—not only the technical integrity of the hosting system, which must withstand any attacks hidden in the material submitted, but more importantly the need to ensure that all material submitted is appropriate for student use. For a fully open system an workflow would have to be established that submitted new exercise material to a trusted authority for checking, before it was released. Finally, it would be nice for ELLE to be able to share exercises and exercise material, allow individual teachers to download and offer it locally or print it out for classroom use, or adapt it to the local environment.

Language learners and teachers already benefit from the wealth of resources available on the web. However, current language learning systems are closed and static; only their developers can add new exercises and new material. This project has devised a flexible structure for a dynamic and open language learning environment, a digital library that is able to grow organically as teachers submit new material, in different languages, targeted towards fresh audiences of language learners.

REFERENCES


AUTHORS

Shaoqun Wu completed her MSc on a digital library of language exercises at the University of Waikato, New Zealand, and is now a research assistant working on the Greenstone Digital Library Project.

Ian H. Witten is Professor of Computer Science at Waikato. He leads the Greenstone project. His research interests include information retrieval, machine learning, text compression, and programming by demonstration. He has published widely, including several books.

Arthur Edwards is senior lecturer at the College of Telematics of the Universidad de Colima, where he collaborates in research in the area of Computer Aided Language Learning, Virtual Reality, Online Collaborative Learning and Ad-Hoc Networks.

David M. Nichols is a senior lecturer in the Department of Computer Science at the University of Waikato. His research interests include digital libraries, human computer interaction and open source software.

Raúl Aquino obtained a PhD. in Electric and Electronic Engineering from the University of Sheffield, United Kingdom. Currently he is a Lecturer at School of Telematics in the University of Colima, México. He has Interests on Vehicular Ad-Hoc Networks, Wireless Sensor Networks and Mobile Computing Networks.

Manuscript received 02 May 2006.