

Endangered Language Documentation and Transmission

D. Victoria Rau

Department of English Language, Literature and Linguistics
Providence University, Taiwan

dhrau@pu.edu.tw

Meng-Chien Yang

Department of Computer and Communication Engineering
Providence University, Taiwan

mcyang2@pu.edu.tw

Maa-Neu Dong

Department of Anthropology
National Museum of Natural Science, Taiwan

tong@mail.nmns.edu.tw

Endangered Language Documentation and Transmission

Abstract

This paper describes an on-going project on digital archiving Yami language documentation (<http://www.hrelp.org/grants/projects/index.php?projid=60>). We present a cross-disciplinary approach, involving computer science and applied linguistics, to document the Yami language and prepare teaching materials.

Our discussion begins with an introduction to an integrated framework for archiving, processing and developing learning materials for Yami (Yang and Rau 2005), followed by a historical account of Yami language teaching, from a grammatical syllabus (Dong and Rau 2000b) to a communicative syllabus using a multimedia CD as a resource (Rau et al. 2005), to the development of interactive on-line learning based on the digital archiving project. We discuss the methods used and challenges of each stage of preparing Yami teaching materials, and present a proposal for rethinking pedagogical models for e-learning.

1 Introduction

Computer applications in second language acquisition (CASLA) have become indispensable in language teaching, assessment and research (Chapelle 2001). Applied linguists no longer question whether it is necessary to use computers to assist in language teaching and learning but debate how to most effectively use the educational technology.

The technology of e-learning has dramatically changed the genre of the language textbook, the context of language teaching and learning, and the traditional roles of language teachers and learners. The key benefits of e-learning, such as the ability to meet individual learning needs, access a wider range of resources, and be exposed to multimedia materials, including both images and sounds, are particularly useful for the design of teaching materials for less commonly taught languages and/or endangered languages, since such textbooks are not readily available and are extremely costly to produce. However, little research has been conducted to explore the potential benefits of using information and communication technology (ICT) in the promotion of minority language teaching and research.

Many projects around the world are now seeking to preserve endangered languages (e.g., Lublinskaya and Sherstnova 2002; Psutka et al. 2002). The attempt to preserve an endangered language includes several steps: documenting and recording the oral and written literature, compiling the grammar and a dictionary of the language, and annotating the documentation related to this language. It is important that an endangered language preservation and documentation project should be comprehensive and carefully planned. This project needs not only to take advantage of state-of-the-art technologies and establish an environment for learning, but also to find an effective approach to teach the endangered language to the ethnic group using the language, particularly to members of the younger generation who often live in urban areas without any connection to their place of origin.

With a two-year grant (2005-2007) from the Endangered Language Documentation Project (ELDP, MDP0114, <http://www.hrelp.org/grants/projects/index.php?projid=60>) from the School of Oriental and African Studies (SOAS), University of London, our research team, represented by the current authors, has begun to prepare digital archiving materials to document the Yami language, a Philippine Batanic language spoken on Orchid Island by less than 4,000 people, and facilitate preservation and dissemination of the Yami materials collected since 1994. According to Rau's (1995) sociolinguistic survey on Orchid Island, Iraralay is the only community of the six villages on the island where children still use some Yami for daily interaction. Although Yami has been offered as an elective in elementary school since 1998, Yami is gradually being

replaced by Mandarin Chinese. Among the junior high school students on the island, 60% either believed Yami would die eventually or were uncertain about the fate of the language.

The approach proposes a comprehensive series of steps to collect and record the Yami language. In addition, the work includes development of a learning method that will be effective with Yami youngsters who live in urban areas away from Orchid Island. Although the complete work of documentation will take many years, the Yami language is in danger of being lost due to rapid urban migration. Therefore, we have developed a strategy to make language items available in learning materials as soon as they have been collected, taking advantage of information technology and computer networking. Using these technologies we have developed an integrated platform for documenting, processing and learning that will help both Yami youngsters and other students taking Yami as a second language.

The integrated platform is built on a main web server with several supporting servers. The main server is designed as the server for resource management and the supporting servers are designed for different purposes. The purpose of this design is to effectively edit the oral recording of the Yami language and to make the language learning materials available. The proposed platform includes three subsystems:

1. a subsystem to manage and edit the digital archiving of the Yami language,
2. a subsystem to handle the workflow of collecting oral recordings of the Yami language,
3. a subsystem to create and manage the Yami language learning materials.

Each subsystem is installed on one or two servers. All these subsystems will be described in detail in Section 3.

Although most ideas in the proposed integrated framework have been used for other language documentation and learning, the proposed framework is an initiative for archiving and teaching an endangered language¹. The attempt of our study is not only to use technologies to preserve an endangered language but also to develop a well-accepted platform for this language. Hence, people can and will learn and appreciate this language and its cultural heritage.

The aim of this paper is to provide a progress report of the current development of the integrated framework, and discuss the various stages of Yami language teaching leading toward e-learning and rethinking pedagogical models. The remainder of this paper is organized as follows. We begin by describing the background of this project and the progress we have made during the first six months and future activities for the second year. Section 3 is an introduction of the integrated framework. Section 4 illustrates the current development of the system, followed by Section 5, a historical account of the developmental process of the Yami language

¹ The integrated framework was first presented at the 2005 ALR Workshop (Yang and Rau 2005).

teaching materials as we move from a traditional grammar translation model towards meeting new challenges in e-learning.

2 Digital Archiving Yami Language Documentation

The Yami digital archiving project began on August 1, 2005. The first two authors serve as the PI and co-PI, respectively. We have also hired four computer technicians, two graduate students, and three Yami staff members for this project. Among the three Yami staff members, our senior language consultant, Maa-neu Dong, has been responsible for transcribing and translating all the materials (tape recording, videotaping, and pictures) in Yami. She has also served as a go-between for our research team and the Yami staff members on Orchid Island.

The other two Yami staff members were asked to provide audio and/or video recordings and pictures at least once a month. It took two months at the beginning of the project to purchase equipment, write up a detailed user's manual, communicate with the Yami staff members to ensure they understand how to use the equipment properly to collect data, and wait for them to sign a paper on Intellectual property rights. One of the staff members has been able to contribute materials regularly since September. However, the other member who originally agreed to participate in the project decided to take another job and thus was replaced by someone else at the last minute. It took the new staff member a much longer time to understand how to collaborate with us.

In addition to digitally archiving the 20 narratives, reference grammar, trilingual dictionary with 2000 entries (Rau and Dong, in press), and multimedia pedagogical materials (Rau et al. 2005), we also collaborated with local consultants to document daily conversations, business transactions, festivals, and ceremonies.

The topics were selected based on consultation of previous research on Yami ethnography, and were designed to meet the standards stipulated by the R.O.C. Ministry of Education for developing Austronesian teaching materials in Taiwan. The topics are closely related to those selected for inclusion in the four volumes of Yami multimedia teaching materials (Rau et al. 2005).

Over the past six months, we have developed the metadata following the structure of the OLAC format. The web site for recording the progress of this project has been established at Providence University with the IP address: <http://yamiproject.cs.pu.edu.tw/yami/>. So far, we have built the e-learning platform and put the four volumes of Yami teaching materials there for online language courses. The activities to be undertaken in the second year are as follows:

- The query interface and the query language for the archiving Yami language database will be designed using the techniques of intelligent searching and common sense indexing.
- An on-line Yami New Testament dictionary will be created.
- E-learning materials with different levels of difficulty will be created and installed into the e-learning platform.
- The server and accessing entrance of the Yami language digital archiving will be procured and installed.
- The preliminary corpus of the Yami language including the language learning materials, the dictionary and the Yami narratives translated into English and Chinese will be established.
- The e-learning materials will be field tested in the PI's graduate seminar, "Austronesian Linguistics: Yami structure, function, and teaching" in Spring 2006.

3 Integrated Framework

In this section, we describe our design and the theoretical framework behind it. The project is divided into four major steps:

- (1) field recording: recording the oral sound data of the Yami language,
- (2) archiving: editing the sound data and annotating the data using the metadata,
- (3) multimedia transformation: analyzing the original data and creating a multimedia Yami dictionary and text description,
- (4) e-learning: creating online Yami language learning materials.

The framework is designed to meet two requirements of our Yami language archiving project:

- (1) to build a complete and original archiving database for Yami language including speech of various genres, grammar, vocabulary and cultural artifacts,
- (2) to create learning materials in an easy-to-learn environment via internet and computer.

3.1 Field Recording

First of all, the existing records collected by the research team since 1994 were organized and digitalized, along with new field recordings. In our project, we have developed an oral speech archiving database to store these oral recordings. Each recording is scanned to find the basic sound characteristics and transferred into digital data. The sound characteristics are used for

comparing and tracking these recordings. The field recordings are arranged by segments, ranging from words in isolation to “idea units” or “tone units” (Chafe 1979) in continuous speech.

Once a segment of the field recording has been completed, the original data is stored in the computer and two different types of digital data are created. These include MP3 data that are used for creating the learning materials and the annotated digital data in which the recordings are separated into sentences with Chinese and English translations. All these data are stored in a relational database with the recording date used as the searching key.

The processing of field recordings is considered to be the preparation and preprocessing stage of the Yami language documentation project. The voice database is used to create the archived data and learning materials.

3.2 Archiving

The archiving step begins with editing the voice database and construction of the Open Language Archives Community (OLAC) metadata for each entity in the voice database. The original sound tracks in the field recording database are edited to improve clarity of the sound by using sampling techniques (Kientzle 1998). The edited sounds are stored as the new sound records in the voice database.

The metadata used for describing Yami language is the OLAC metadata, an extended Dublin Core set with basic elements of language resources. To meet the requirement of the linguistic community, certain new extension elements are put in the OLAC set following the Dublin Core Metadata Initiative (DCMI) guidelines (DCMI 2000). To build a proper OLAC metadata for the Yami language, we have chosen to adopt the OLAC set proposed by Bird and Simons (Bird et al. 2001, Bird and Simons 2003) for this project. Because Yami is primarily an oral language, we use a subset of this OLAC set. The OLAC elements used in this project are: {Title, Creator, Subject, Subject language, Description, Publisher, Contributor, Date, Type, Format, Identifier, Source, Language, Relation, Rights}. The reason for selecting these elements is to create a common description of the Yami language. Furthermore, after reviewing the field study materials, we can show that the above OLAC subset can meet the basic requirement for describing the Yami language. The rules to apply these OLAC elements to each recording of the Yami language are:

- (1) Each OLAC element can be optional and repeatable;
- (2) Each OLAC element can describe only one single identification or one single range;
- (3) Data format of each OLAC element follows the rules in DCMI (DCMI 2002).

Each OLAC element used in describing the Yami language is given following the OLAC and ELDP guidelines. Suppose there is a Yami language sound track to be described, the OLAC element set of this sound track is shown as follows:

Title: the Chinese name of the Yami language sound track. A second Title element is used to store the English translation.

Creator: the Yami speaker who uttered this speech. A second Creator element is used to store his/her Chinese name.

Subject: the keyword used to classify the content of the Yami language sound track. The keywords and controlled vocabularies are being collected.

Subject language: the Chinese linguistic description of the Yami language. A second element is the corresponding English description.

Description: the usage and the multimedia data related to this Yami language sound track. Some multimedia data are collected using the Multimedia Transformation step described in Section 3.3.

Publisher: the research teams and the sponsoring institutions.

Contributor: the research teams and the person who recorded this sound track.

Date: the date this sound track was recorded and the date the archiving process was completed.

Type: the genre of the content of the Yami language sound track. We are transferring many Yami-language linguistic and anthropological terms into DC-type. These DC-type terms are used as the Type element.

Format: the digital data type of the Yami language sound track.

Identifier: the ELDP identifier for this Yami language sound track. We have followed ELDP guidelines to create identifiers for the archived sound track.

Source: the location of the archiving database and the location for storing the field study draft.

Language: English and Chinese (traditional and simplified characters)

Relation: the related Yami language sound tracks.

Rights: copyright information of this sound track.

In the archiving step we also consider how to build a database of the controlled vocabularies for the Yami language. We use three sources for the controlled vocabulary in this project: lexicon, primary text and language description.

The table of OLAC metadata is created in two forms, one XML text table format and one relational table format. The voice database from the first step is edited and connected to the metadata table.

Another goal of this step is to build a Yami language online dictionary. The OLAC metadata are used for parsing and editing the voice database to create a Yami language online dictionary. We are developing an auto dictionary-generating program that can process the OLAC metadata and find suitable terms. In addition, we will use the grammar and course materials of the Yami language multimedia courseware created by Rau et al. (2005) to build our on-line multimedia Yami language dictionary.

When the metadata of a set of the Yami language sound tracks is completed, the results will be published online on our web site. So far our focus has been aligning the OLAC metadata of the Yami language sound tracks with the multimedia courseware by Rau et al. (2005). Beginning in 2006, we will try to use ontology to determine rules for creating metadata automatically and to develop an automatic metadata generator for the Yami language.

3.3 Multimedia Transformation

To preserve the Yami language, we use an image database to annotate the language. In addition, each word in Yami is annotated with its orthography stored in a sound database. The purpose of this transformation is to build an image for each Yami word. Therefore, the meaning of the word can be related directly to a picture. The reasons why we have chosen to use this approach to annotate the Yami language are as follows:

- (1) The annotated pictures help learners understand the traditional lifestyle on Orchid Island and give them more incentive to learn the language.
- (2) The pictures include many Yami cultural artifacts. The annotated pictures can thus preserve descriptions of their cultural heritage.

The steps for multimedia transformation of the Yami language are as follows:

- (1) Collect suitable images for building the annotated image database. In addition to the materials being provided monthly by our Yami consultants, we consulted other research teams to borrow Yami images and video recordings. Jian-Xiang Lin's sample materials can be retrieved from the following websites:

<http://dlm.ncnu.edu.tw/Lanyu/DigitalReservation/script/scheme/preview.asp?xml=415.xml&dxsl=Acul04.xsl>,
<http://www.sinica.edu.tw/~video>.

- (2) Design criteria to choose the images. We will select appropriate images and develop possible connections between Yami expressions and a set of pictures.

- (3) Build a special annotated database and use the Yami language to annotate the image data. The annotated algorithms will be based on the fuzzy logic style (Kecman 2001) or the Coherent Language Model (Jin, Chai and Si. 2004).
- (4) Build a corresponding mapping relation between a Yami expression and a set of annotated images. The mapping relations will be a set of contexts and symbolic tables similar to a set of induction rules.
- (5) Build a sound connection between each Yami word and its phonetic symbols by using the fuzzy logic learning algorithm.

The results of multimedia transformation will be able to be used as a foundation for creating online learning material. The results are stored in a relational multimedia database as well as the XML pages.

3.4 e-Learning

The final task of our project is to find an effective way to teach the Yami language to urban Yami youngsters and other learners of Yami as a second language. To create an open environment conducive to self-study, our choice is computer-based or Internet-based learning. There have been various discussions about how to use information technologies and the web to learn a different language. Gerbault (2002) showed that it is viable to set up a multimedia environment for learning a language without a teacher's participation. Fujii et al. (2000) demonstrated a project using the Internet as a tool for the teacher to post course materials and create an online learning environment.

E-learning generally consists of self-access, reference sources, discussion forum, and virtual learning classrooms. The main motives for introducing e-learning include improving student multimedia learning experience, enhancing learner autonomy and broadening participation. Finally, e-learning can be controlled primarily by tutors or students, depending on the objectives, contents, learning tasks, length/time/place of study, or choice of assessment activities.

A computer-based learning environment is a very effective way to help students learn. In order to provide an effective learning environment, Leung (2003) suggested that four contextual issues should be considered in design and implementation of computer-based learning. These issues are topic selection, authenticity, complexity, and multiple perspectives. The design of the web-based computer-assisted learning program for the Yami language takes these four issues into consideration. We outline our design as follows.

The proposed e-learning environment in this project is designed to function as a virtual classroom. Students can select the Yami language learning materials based on their interest and

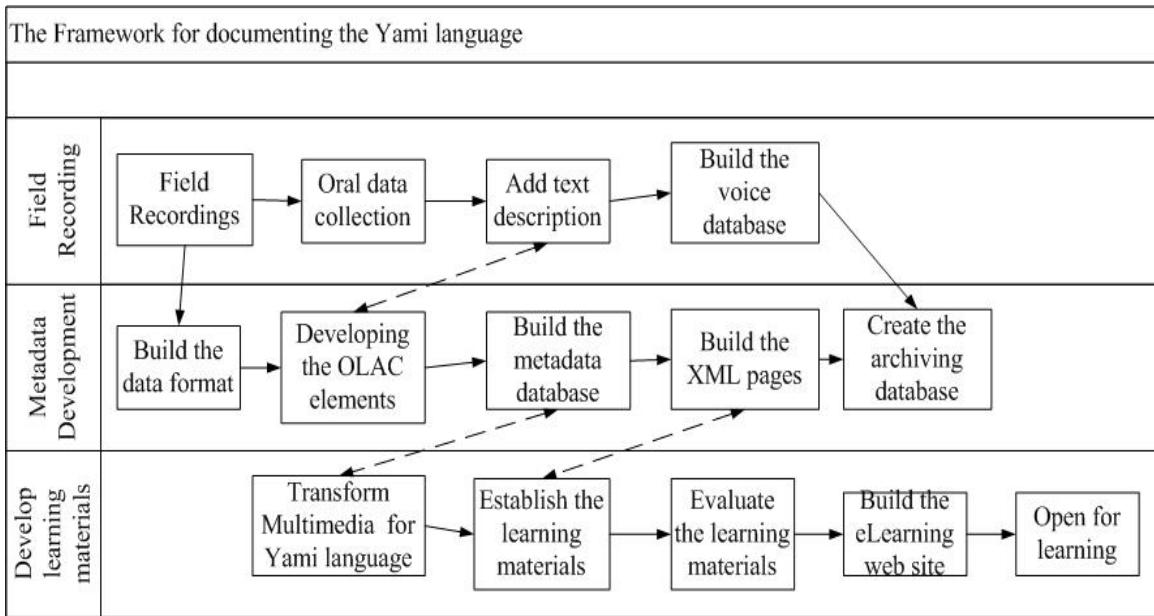
capability. If a student asks for clues or explanation of a specific Yami word or expression, a suitable image or video clip is retrieved from the multimedia database. If a student is not familiar with a specific Yami sound, a similar phonetic symbol is provided to him/her. The learning materials are arranged in three different settings: scenario setting, easy-to-difficult condition setting, and learner's choice setting. The scenario setting uses related scenes in Yami society such as the flying fish festival as a main theme of the learning materials. The easy-to-difficult condition setting allows the learner to select different levels of the Yami language materials. The levels are based on word frequencies and complexity of grammar. The learner can arrange the learning materials in the chosen setting. The learning system will give detailed guidelines to explain how to choose the learning materials. If a student wants to learn the Yami language, s/he can choose different learning materials based on his/her interest. The learning materials are designed as theme units with exercises and rubrics for self-assessment. The design of these Yami language exercises is based on a study about the reactions of students to using a web-based system for learning Chinese in Taiwan (Yang, 2001).

We also plan to use multimedia to help students learn the Yami language. We will build an annotated image database as a tool to help the learners understand the meaning of Yami words or expressions. To promote clearer understanding, the database will include both a video clip and several still images to help learners visualize Yami culture.

The interface of the proposed learning environment is built on a web server with a dynamic web page. To establish a more efficient learning environment, all the learning materials are edited into reusable learning objects. The user interface is developed as an adaptive style following Mich et al.'s (2004) PARLING system.

The proposed framework is illustrated in Figure 1.

Figure 1: The integrated framework for the Yami Language preservation project



4 Current progress in development of the e-learning platform

Based on the above-mentioned framework, we have developed the first version of the multimedia e-learning platform for the Yami language. The e-learning platform is developed following the guidelines:

1. the e-learning system should provide an interface simple enough for the beginner;
2. the e-learning system should include different levels of the course materials with the possibility of adding new materials;
3. the e-learning system should provide many visual effects to stimulate interest in learning the Yami language.
4. the e-learning system should provide a comprehensive dictionary and an effective searching tool.

Currently, the Yami language materials collected to build the e-learning platform include:

1. the Yami language course materials by Rau et al. (2005),
2. the sound tracks of each utterance in the forty lessons,
3. images and video clips collected by the three Yami staff members.

The e-learning system is developed as a group of dynamic web pages which are stored in a web page server with a database engine for manipulating the course materials and the learning records.

The current e-learning platform for the Yami language includes the following components:

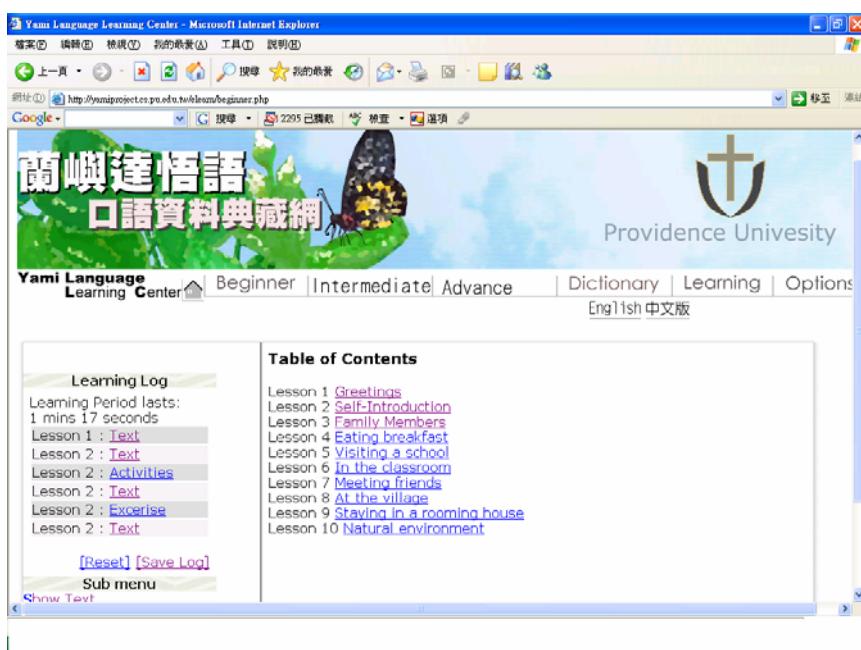
1. the Yami course materials, classified into three levels: beginner, intermediate and advanced;

in each lesson, the materials include the Yami text, Chinese and English translation, word analysis, grammar, learning activities, and exercises,

2. the Yami dictionary, organized in alphabetic order,
3. the system setting tool which includes the interface setting, the learning log setting, and the web display setting,
4. the online learning activities which students can use to collaboratively practice with their partners,
5. the virtual learning group which allows students to email their learning logs to their own mailbox and the teacher's mailbox.

The e-Learning platform for the Yami language is shown in Figure 2.

Figure 2. The e-learning platform for the Yami language



5 Development of Yami language materials: a historical account

The process of building a platform for the e-learning materials needs to be informed by an in-depth investigation of the contexts of language teaching and learning, so that the design of ICT materials can meet the learners' needs. The present researchers discovered quickly that it would not be sufficient to simply make the four volumes of prepared language materials available on the Internet without making drastic modifications to the format. What are the benefits and constraints of e-learning? What on-line support is required? In other words, we need to rethink pedagogical models for e-learning.

Before we propose how to explore the task of developing Yami e-learning courseware from the viewpoints of applied linguistics and computer science, we provide a historical account of the development of Yami language materials and the teaching and learning contexts since 2000.

5.1 From audiolingual to communicative competence

The first Yami textbook (Dong and Rau 2000b) for adult learners was compiled when Rau was teaching a graduate course on Austronesian Linguistics at Providence University. Three students participated in that course, learning the structure of Yami as an introduction to Austronesian languages before they could read and understand journal articles in the field. They were also employed at the same time to help with Yami data collection for a project funded by the Council of Aboriginal Affairs (Dong and Rau 1999, 2000a). The reason that a course on aboriginal languages in Taiwan was offered in an English graduate program is also related to the wider social context where ethnic languages were given opportunities to reverse their language vitality in Taiwan.

After the announcement of the policy of the Ministry of Education (MOE) to offer ethnic languages (including aboriginal languages) as elective courses in grade one beginning on June 4, 1999, many Austronesian language communities were provided with grants, but lacked sufficient training to create teaching materials for primary schools to meet the immediate needs. As a result, the quality of the textbooks created by non-professionals left much room for improvement. Native speakers could provide constructed dialogues and stories in their languages, but had no knowledge of the structure and pedagogy necessary for writing teacher's manuals. Since that gap can only be filled by a unique combination of manpower, the first author, who learned her trade from John Wolff while compiling a Pilipino textbook (Wolff, Centeno and Rau, 1991) collaborated with the third co-author, a Yami language worker, who received training to transcribe and translate Yami speech data at the Institute of Ethnology at

Academia Sinica, to embark on this mission.

The first Yami textbook was heavily influenced by Wolff's audiolingual, self-instruction design but organized by a "functional" syllabus, advocated by Huang (1992). The book begins with a general introduction including principles of pronunciation, followed by ten lessons and a reference grammar. The last lesson is a short story, but the other nine are constructed dialogues based on a plot of a Chinese woman visiting her Yami friend on Orchid Island. Each lesson consists of four sections: (1) dialogues, (2) word analysis, (3) sentence patterns with drills and grammatical explanations, and (4) culture notes. The curriculum design was modeled after the late twentieth century's "Cornell experience" with native speakers leading the drills in L2 and linguists explaining the grammar in L1.

However, with the rapid development of computer technology, any new language teaching materials in the 2000s have been expected to add multimedia components. During 2003-2004, Dong was hired to teach Yami as a required course in the Department of Language Communication at National Dong Hua University on the east coast of Taiwan. She was approached by the department head to plan a new series of textbooks with four volumes to meet the standards for the Austronesian language curriculum, as stipulated by the MOE. During the same time Rau was a visiting scholar at Center for Advanced Research on Language Acquisition (CARLA), University of Minnesota, conducting research on second language acquisition. We rose to the challenge and planned our second Yami textbook (Rau et al 2005).

We laid out three principles in preparing the multimedia materials: (1) selecting interesting topics covering a wide range of genres and structures, (2) adopting a communicative task-based approach closely tied with standards and assessment, and (3) using ICT (information and communication technology) to enrich the contents and modes of presentation. The topics and sequence of grammatical structures followed the MOE standards closely. Similar to our first textbook, the first two volumes comprise constructed dialogues based on a story line of a Chinese student touring Orchid Island. The main character flies back to Taiwan at the end of Volume Two. The third volume includes many dialogues constructed based on the folklore and narratives we collected previously (Rau and Dong, in press). Volume four is a collection of short stories and descriptions of cultural activities ranging from traditional to contemporary events. The second textbook is enriched with extensive comments on grammar, classroom activities, student exercises, and a CD.

We explicitly stated how to use this textbook in the preface. The four volumes are designed as a four-semester "Yami as a second language" course for college level students, meeting four hours per week. The language instructor is foregrounded, leading classroom activities, whereas the linguist is invisible because she has provided all explanations in writing. The students are required to listen to the CD and preview the lesson (including the dialogues,

vocabulary and grammar) before they attend the class. During the first two hours of each lesson, the native speaker can use the classroom activities and assign exercises as homework. When they meet in the third hour, the students may present the results of their assigned homework. In the fourth hour, the instructor can choose to provide feedback on the students' homework or entertain questions.

The two textbooks were field tested with undergraduate students at Dong Hua University and graduate students at Providence University. Dong was invited to teach Yami at Dong Hua University during 2003-2005. She began with seven undergraduate students, two of whom were Yami native speakers. Whereas the five Chinese students were taking it to fulfill language requirements, the two Yami students were learning literacy skills and grammatical structure of their native language. The students met once every week on Saturdays for four hours. It was later changed to once every two weeks for six hours due to the long transportation time for the instructor. She started by focusing on dictation, pronunciation, conversations, and sentence patterns without using any textbook the first semester. Not until the second semester did she use the first Yami textbook (Dong and Rau 2000b) as the major textbook for the course. The grammar description in the book is simple enough that was easily received by the students.

In the second year, five students continued with the course (4 Chinese and one Yami). They began to use the third and fourth volumes of the multimedia materials (Rau et al. 2005). The learning activities and exercises caught their attention. They were especially interested in drawing comics for story retelling as a suggested activity. But the grammar explanation seemed to be too technical and beyond their comprehension.

In the Spring 2005, Dong and Rau co-taught Yami lessons to five graduate students at Providence one hour per week as part of the Austronesian Linguistics seminar, using the first two volumes of the multimedia materials. Dong was responsible for providing Yami language input by modeling pronunciation and supplying vocabulary and sentences to mediate communication in Yami, a modified Community Language Learning approach (CLL, Curran 1976). Rau, on the other hand, played the role of a linguist, explaining grammatical structures in L1. As expected, the graduate students in linguistics picked up the grammar much more quickly than the undergraduate students without such training. They also exceeded the undergraduate students in classroom activities and assigned homework. However, the undergraduate students outperformed the graduate students in oral proficiency and pronunciation. The different course objectives certainly contributed partially to the different results.

While trying to prepare e-learning materials, transforming the traditional language teaching from print materials to on-line multimedia learning, we quickly noticed that the linguist could not simply upload all the word files and let the computer technicians worry about the rest. The

real challenges are really beginning. Thus we need to rethink pedagogical models and redesign our materials for e-learning.

5.2 Proposal for rethinking pedagogical models in e-learning

While a new genre is emerging (Swales 1990), we propose to explore (1) the learning experience, learners' strategies, autonomy, identity, and needs, (2) the effectiveness of the adaptive materials and e-portfolio, and (3) the major genre differences between the traditional language teaching materials and the newly designed e-learning materials. We plan to use nexus analysis (Scollon and Scollon 2004), an ethnographic discourse analysis, as our guide for a qualitative study. In the area of computer technology, we will develop online Yami language course materials and authentic assessment with a rich variety of cultural items, focusing on techniques to adapt the Yami language learning platform. The online platform is designed for both second language acquisition and first language literacy education. The system will automatically adjust the online learning materials presented based on the user's progress. In addition, we will also explore the techniques necessary to develop handheld portable Yami language e-learning materials.

6. Conclusion

This paper describes the ongoing project on Yami language documentation and transmission. Our collaborative work might serve as an example of preserving and promoting endangered language in Taiwan. The integrated framework for archiving, processing, and teaching the Yami language has provided a good model for language preservation. Our experience has also pointed to future studies in rethinking pedagogical models in designing new materials to face the challenges of e-learning.

References

Bird, S., G. Simons and C-R Huang. 2001. The open language archives community and Asian language resources, *NLPRS 2001*: 31-38.

Bird, S. and G. Simons. 2003. Extending Dublin core metadata to support description and discovery of language resources, *Computers and Humanities*, No. 37: 378-388.

Chafe, W. L. 1979. The flow of thought and the flow of language. In *Syntax and Semantics 12: Discourse and Syntax*, ed. by T. Givón, 159-181. New York: Academic Press.

Chapelle, C. A. 2001. *Computer Applications in Second Language Acquisition: Foundations for Teaching and Research*. Cambridge: Cambridge University Press.

Curran, C. A. 1976. *Counseling-learning in Second Languages*. Apple River, IL: Apple River Press.

DCMI 2000. Dublin Core Qualifiers.
[\[http://dublinecore.org/documents/2000/07/11/dememsqualifiers/\]](http://dublinecore.org/documents/2000/07/11/dememsqualifiers/)

DCMI 2002. DCMI Elements and Element Refinements – a current list.
[\[http://dublinecore.org/usage/terms/dc/current-elements/\]](http://dublinecore.org/usage/terms/dc/current-elements/)

Dong, M-N and D. V. Rau. 1999. *Yami Texts and Vocabulary* (in Chinese). First Year Progress Report, Council of Aboriginal Affairs, Executive Yuan, R.O.C.

Dong, M-N and D. V. Rau. 2000a. *Yami Texts and Vocabulary* (in Chinese). Second Year Progress Report, Council of Aboriginal Affairs, Executive Yuan, R.O.C.

Dong, M-N. and D. V. Rau. 2000b. *Yami textbook* (in Chinese). Taipei: Crane.

Fujii, S. Iwata, J. Hattori, M. Iijima and T. Mizuno. 2000. “Web-Call”: a language learning support system using Internet. In *Seventh International Conference on Parallel and Distributed Systems*, 326-331.

Gerbault, J. 2002. Information technology and foreign language learning: what happens when no teacher is around? In *International Conference on Computers in Education*, 394-398.

Jin, R., J Chai and L. Si. 2004. Effective automatic image annotation via a coherent language model and active learning, In *The 12th ACM Annual Conference on Multimedia (ACM-MM)*, 892-899. New York, USA.

Huang, L. 1992. *Beginning Atayal* (in Chinese). Taipei: Crane.

Kecman, V. 2001. *Learning and Soft Computing: Support Vector Machines, Neural Networks, and Fuzzy Logic Model*, MIT press.

Kientzle, T. 1998. A Programmer’s Guide to Sound, Addison-Wesley.

Leung, A.C.K. 2003. Contextual issues in the construction of computer-based learning programs, *J. Computer Assisted Learning*, Vol. 19: 501-516.

Lublinskaya, M. and T. Sherstnova. 2002. Audio collections of endangered arctic languages in the Russian federation, In *Proceedings of the 5th International Conference on Text, Speech and Dialogue*, ed. by P. Sojka, I. Kopeck, and K. Pala, 347-353. Lecture Notes in Computer Science; Vol. 2448. Springer-Verlag, London, UK.

Mich, O., E. Betta and D. Giuliani. 2004. PARLING: e-literature for supporting children learning English as a second language, In *Proceedings of the 9th International Conference on Intelligent User Interface*, ed. by J. Vanderdonckt, N. J. Nunes and C. Rich, 283-285. ACM Press, New York, NY, USA.

Psutka, J., P. Ircing, J. V. Pusta, V. Radova, W. J. Byrne, J. Hajic, S. Gustman, and S. Gustman, 2002. Automatic transcription of Czech language oral history in the MALACH project: resources and initial experiments. In *Proceedings of the 5th International Conference on Text, Speech and Dialogue*, ed. by P. Sojka, I. Kopeck and K. Pala, 253-260. Lecture Notes in Computer Science; Vol. 2448. Springer-Verlag, London, UK.

Rau, D. V., M-N. Dong, M-Y Lin , H-H. Chang, and Y-C Hsu, 2005. Multimedia Materials of Yami Language, Technical Report, Department of English Language, Literature and Linguistics, Providence University.

Rau, D. V. and M-N. Dong. In press.. *Yami Texts with Reference Grammar and Vocabulary*, Language and Linguistics. A-10.

Rau, D. V. 1995. Yami Vitality. NSC report NSC84-2411-H-126-001. Presented at the Symposium on Language Use and Ethnic Identity, Institute of Ethnology, Academia Sinica 1995/5/16..

Scollon, R. and S. W. Scollon. 2004. *Nexus analysis: Discourse and the Emerging Internet*. Routledge, London and New York.

Swales, J. M 1990. *Genre analysis: English in Academic and Research Settings*. Cambridge University Press, Cambridge.

Wolff, J., T. Centeno and D-H. Rau 1991. *Pilipino through Self-Instruction*. Four volumes. SEAP, Cornell University.

Yang, M-C. and D. V. Rau 2005. An integrated framework for archiving, processing and developing learning materials for an endangered aboriginal language in Taiwan. Paper presented at the ALR-2005, October 14, Jeju Island, Korea.

Yang, S. C. 2001. Integrating computer-mediated tools into the language curriculum, *J. Computer Assisted Learning*, Vol. 17: 85-93.