ONLINE LANGUAGE GAMES FOR ENDANGERED LANGUAGES: jeux.tshakapesh.ca, www.eastcree.org/lessons

Marie-Odile Junker, Delasie Torkornoo

Carleton University (CANADA)
marie-odile_junker@carleton.ca, dtorkorn@connect.carleton.ca

Abstract

It is often a struggle to create a strong presence on the web for Aboriginal languages and to make use of Information and Communication Technologies to support language preservation and maintenance. One crucial aspect of Aboriginal language retention, at least in Canada, has been the development of literacy in Aboriginal languages. We report here on a series of projects with two Aboriginal linguistic groups in Canada: East Cree and Innu. For the past six years, using a collaborative (participatory action) research framework with partners involved in language teaching, we have been developing online language lessons and games aimed at bilingual Aboriginal speakers (Cree-English and Innu-French) who wish to become literate in their language. The first set of lessons and exercises, developed in 2006, was aimed at fluent adult speakers of East Cree who had been educated in English and wanted to learn basic syllabic orthography. The subsequent sets had to take into consideration multiple uses and users, including a parallel development for the Innu language, which does not use syllabics. The latter sets of games include vocabulary enrichment, the teaching of grammatical concepts, and the discovery of language structure.

Different interfaces allow for the ongoing creation of new lessons and exercises. Features include: sound and image, computer generated random variation, tractable results and progress, and specialized writing systems like Cree Syllabics.

Such online language games are database-driven web applications with an extensive collection of multimedia. They require a web and database server on the back-end, and a web browser for access on the user-end. These two requirements form the main technical challenges, but are compounded by the need to integrate the multimedia collection with other language applications such as online web dictionaries.

The project has provided training to Aboriginal educators. Team-based design of the tool as well as the games' content has allowed for flexibility and incremental development, as well as transfer of resources from one language group to another.

In this multimedia presentation, we report on the design process, and discuss the tools, technologies, and methods used to implement these language games.

Keywords: online games, language games, Aboriginal, endangered language, Canada, Cree, Innu

1 INTRODUCTION

It is often a struggle to create a strong presence on the web for Aboriginal languages and to make use of Information and Communication Technologies to support language preservation and maintenance. One crucial aspect of Aboriginal language retention, at least in Canada (Drapeau, 1992), has been the development of literacy in Aboriginal languages. We report here on a series of projects with two Aboriginal linguistic groups in Canada: East Cree and Innu. For the past six years, using a collaborative (participatory action) research framework with partners involved in language teaching, we have been developing online language lessons and games aimed at bilingual Aboriginal speakers (Cree-English and Innu-French) who wish to become literate in their language.

In this paper, we report on the design process and discuss the tools, technologies, and methods used to implement these language games.

2.1 Cree and Innu languages and dialects

The games are developed for Aboriginal speakers who wish to learn to read and write in their mother tongue: East Cree or Innu. East Cree and Innu belong to the Algonquian language family [1], [2]. East Cree and Innu are spoken in several Canadian provinces, as shown in Fig.1, but the languages with which we are concerned are both spoken in the French-speaking province of Quebec, in Eastern Canada. For historical and religious reasons, the East Cree people tend to have English as their second language (and French as their third), while the Innus in Quebec speak French as their second language, and their Aboriginal language as their first. There are about 13,000 East Cree speakers living in nine different communities and about 10,000 Innu speakers (formerly called Montagnais) living in eight different communities. Additionally, there is an Innu community in Quebec with a population of approximately 2,000 people, but very few Innu speakers remain, and there are two more Innu communities in Labrador, where people speak English as their second language.
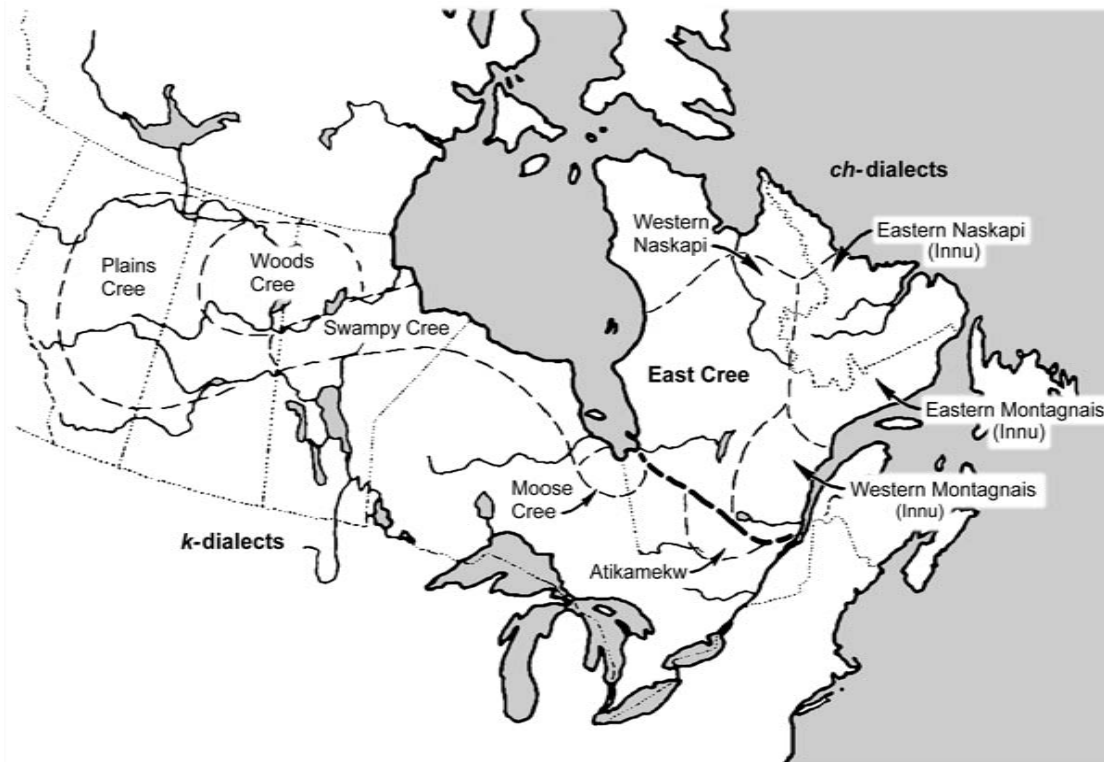

Fig. 1: Map of Cree-Innu languages in Canada

Within each language, there is some dialectal variation. East Cree can be divided into the Northern and Southern dialects, each with their respective orthographic conventions. Innu has one standardized spelling for all the dialects, but this results in spelling that is rather arbitrary when compared to the pronunciation. To address such variation, we chose to represent three Innu dialects and two East Cree dialects in the audio component of our games.

East Cree uses a syllabic orthography. The table below shows the correspondence between the syllabic characters and a standard roman orthography for a few of the 139 symbols.

▽		△	Ȧ	▷	⊳	◁	⊲		o	‖
e		i	ii	u	uu	a	aa		u	h
·▽		·△	·Ȧ	·▷	·⊳	·◁	·⊲			
we		wi	wii	wu	wuu	wa	waa			
V	·V	∧	Ȧ	>	⊃	<	⋖	·⋖	<	
pe	pwe	pi	pii	pu	puu	pa	paa	pwaa	p	
U	·U	∩	∩̇	⊃	⊃	C	Ċ	·Ċ	c	
te	twe	ti	tii	tu	tuu	ta	taa	twaa	t	

ᑫ	·ᑫ	ᑭ	ᑭ̇	ᑯ	ᑯ	ᑲ	ᑲ	·ᑳ	ᒃ	ᒃ
ke	*kwe*	*ki*	*kii*	*ku*	*kuu*	*ka*	*kaa*	*kwaa*	*k*	*kw*
ᒉ	·ᒉ	ᒋ	ᒋ̇	ᒍ	ᒍ	ᒐ	ᒐ̇	·ᒑ	ᒡ	
che	*chwe*	*chi*	*chii*	*chu*	*chuu*	*cha*	*chaa*	*chwaa*	*ch*	
ᒣ	·ᒣ	ᒥ	ᒥ̇	ᒧ	ᒧ̇	ᒪ	ᒪ̇	·ᒫ	ᒼ	
me	*mwe*	*mi*	*mii*	*mu*	*muu*	*ma*	*maa*	*mwaa*	*m*	
ᓀ	·ᓀ	ᓂ	ᓂ̇	ᓄ	ᓄ̇	ᓇ	ᓇ̇	·ᓈ	ᐣ	
ne	*nwe*	*ni*	*nii*	*nu*	*nuu*	*na*	*naa*	*nwaa*	*n*	

Innu uses a roman orthography, with one special character, the raised u: ^{u} in final position in some words.

2.2 Why language games and what for?

2.2.1 How the games and lessons came about

The first set of lessons and exercises, developed in 2006, was aimed at fluent adult speakers of East Cree who had been educated in English and wanted to learn basic syllabic orthography. As of 1995, the East Cree School Board had adopted Cree as the language of instruction in their schools, from Kindergarten to Grade 3, so children were coming home wanting to write notes in Cree to parents who spoke fluently, but could neither read nor write Cree. The demand for Cree literacy courses amongst adults was very hard to meet, however we hoped that online lessons and games designed to facilitate the learning of the syllabic orthography would help. We started developing the games with three Cree teachers and education consultants. The first series consisted of a set of 15 lessons per dialect, with hundreds of exercises, introducing the syllabic characters to adults, with written instructions in English and oral instructions in both dialects of East Cree (Fig. 2).
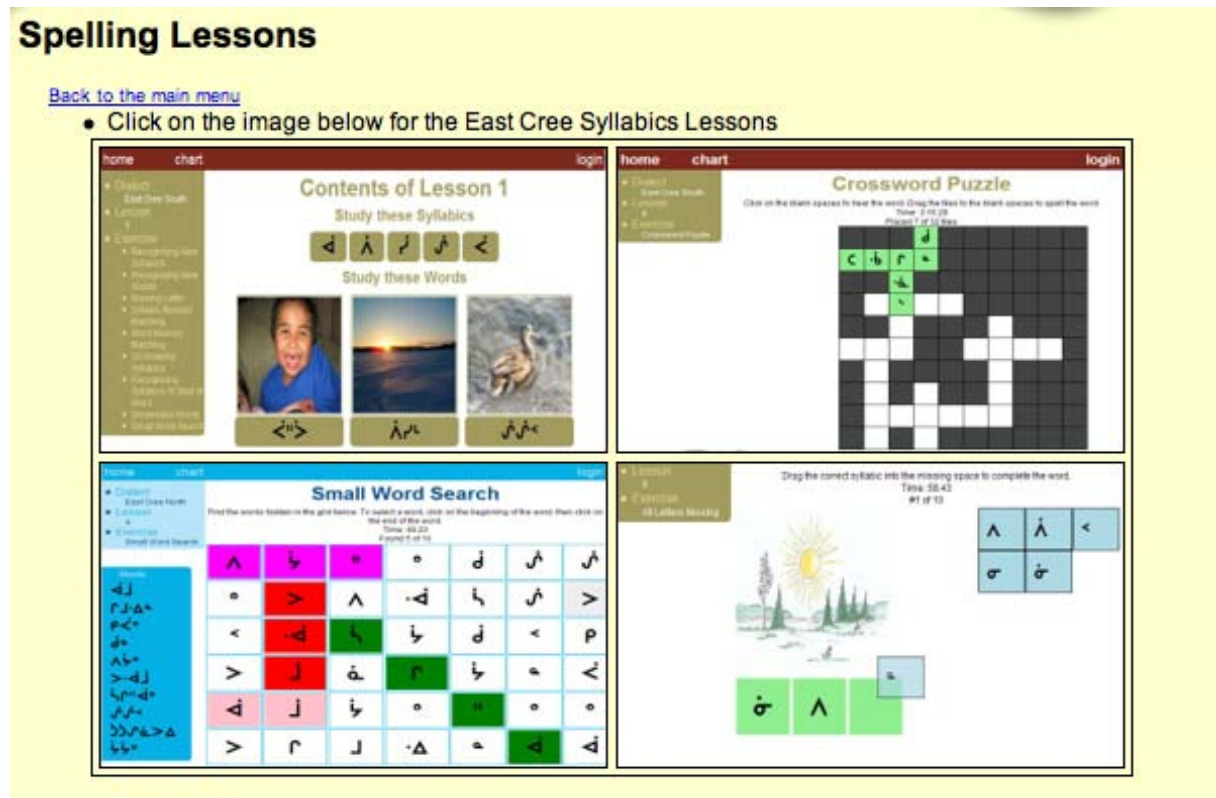

Fig.2: East Cree Syllabic Lessons and games in 2006.

The games were instantly popular with a much wider audience than anticipated. Second-language learners, even children and teenagers (most of whom were already proficient in reading) were using them. This success resulted in a demand for vocabulary games, and more advanced lessons targeting superior literacy skills, such as the mastery of grammatical orthography and the reading of texts. The

subsequent sets of lessons and games were aimed at more advanced learners, as well as second-language learners, children in school, or speakers who simply wanted to enjoy playing with their language online. The latter sets of games included vocabulary enrichment, the teaching of grammatical concepts, and the discovery of language structure. Therefore, a new type of software needed to be developed, with more emphasis on text and sound, rather than images, including the possibility of typing in answers, as we discuss in section 3.

By 2009 we were developing lessons and games for both Cree and Innu speakers. The first set of Cree lessons and exercises were adapted into Innu to teach basic literacy skills, and the advanced activities for both languages started being developed shortly thereafter, with the Cree and Innu language specialists working in parallel and inspiring each other. Our latest set of games (for Innu) was only just made public in June 2012.

2.2.2 Learners' profiles

Innu and East Cree learners' profiles are very similar, despite the fact that these learners speak different colonial languages. In general, most Cree or Innu speak their dialect but do not know how to write in the recently standardized orthography. Of the 10,000 Innu speakers targeted in this project, 200 at most have had training in the standard orthography, and less than 50 have really mastered it (Hélène St-Onge, p.c.). Similarly, of the 13,000 Cree speakers, while many can read and write in syllabics, few have really mastered the standard syllabic orthography. In both groups, some people over 60 years old are still unilingual speakers of Cree or Innu, while the others are bilingual speakers of Innu-French or Cree-English, and many others still are trilingual Cree-English-French.

All Innu communities in Quebec have a primary school, but only two have a secondary school. There is no course offered in Innu at an advanced level (beyond primary school), and no Innu language in schools outside the Innu communities. Unlike for the Crees, Innu has never been the language of instruction. Cree operate their own schools all the way to high school, but since the recent decision by the board to cancel the Cree as a Language of Instruction Program, Cree is now only taught as a language course all levels.

There is a tendency, once a colonial language orthography has been mastered, to use this orthography to write the language phonetically. This is ineffective and does not allow for communication across dialects and idiosyncratic speakers. The main objective for both language groups is thus to promote the standard orthography through these online games. Literacy in the Aboriginal language is seen as a desirable goal in terms of Aboriginal language and culture preservation [3], [4].

Many Innu and Cree also live in large Canadian cities outside of their community. Language preservation is an issue for them and can be assisted by the use of online language games. The same is true for one Innu community that has lost the language and wishes to regain it; for them, the language has to be taught as a second language and our site can be a good complement.

2.2.3 Objectives

As noted above, the main objective for both language groups is to promote the use of standard orthography through these online games. There is also a desire to slow down, or reverse as much as possible, language attrition. Areas of vocabulary ignored by young people were identified. Promoting knowledge of grammar in general, not only for grammatical orthography but also for language retention and accuracy, was also identified as another goal of the games. Both languages are richly inflected, and the mastery of this inflection is crucial for writing and also for speaking at an advanced level. For example, Cree and Innu verbs indicate whether things happen in a dream, are a possibility or a fact, who exactly is involved in an event and in which ranked role (including pure witnesses of all kinds), and whether the source of information is hearsay or eye-witnessed (also called evidential marking), to give only a few examples of categories that are not salient nor obligatory in the colonial languages.

Support for language teachers in both groups was another desirable outcome. In addition to managing each learner's profile, we are currently adding a class management tool so that teachers may use the games as a part of a language course, to track their students' progress. Yet another goal is to allow teachers to create their own sets of exercises, either by selecting different existing lessons and exercises and creating a new progression, or by creating entirely new ones. This allows for multiple uses.

2.3   Game interfaces

Different interfaces allow for the users to play the language games, as well as allow for the ongoing creation of new lessons and exercises of various kinds. The bulk of the lessons and games are stored in a relational database, which is explained in section 3. For now, we will take a quick look at the main features available to our users and content developers.

Level one of the East Cree lessons consists of a series of 15 to 17 lessons per dialect, introducing the syllabic characters with written instructions in English and oral instructions in Cree (either dialect).

Each lesson is built as follows. After the user is presented with new syllabic characters with accompanying sound, and words with matching sound and images containing these characters (offering both a phonic and a global reading approach), he or she is asked to proceed with 5 to 15 different types of exercises that use sound, image, and spelling matching in different ways: missing letter(s), word searches, crossword puzzles, and so on (Fig. 3) The order of items in the exercises is randomly generated, so they are never the same if a user plays again.

Fig. 3 Exercise selection for Level One

An Innu lesson for Level One similarly presents the letters used in the standard Innu orthography, except this time they are roman letters, with words to which sound and images are attached, followed by a series of appropriate games and exercises. Written instructions are available in French and oral instructions in Innu.

Thematic vocabulary lessons using the same software include weather expressions, plants, animals, days of the week, months, etc. An example for Innu is given in Fig. 4.
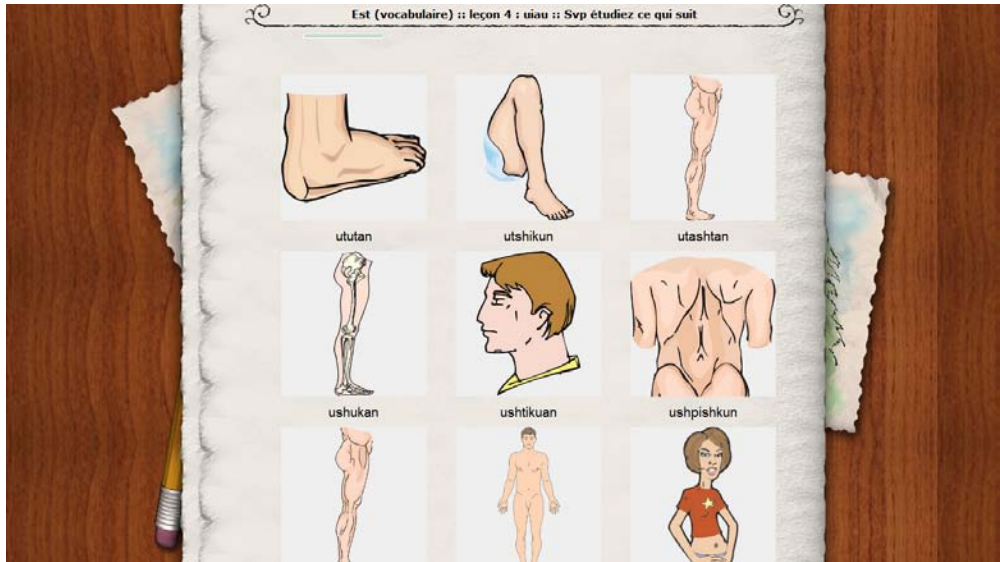
Fig. 4 Thematic Vocabulary for Innu- The body

Advanced levels for Cree and Innu feature text-oriented games such as:

- Dictation exercises: single words or sentences. After listening, the user types in the correct answer.

- Dictation in cloze texts: the user hears an entire sentence and has to type in the missing written parts.

- Cloze texts: The user has to fill in the missing words in a text by choosing from a selection of words, which can all be correct but need to go in specific places, or can include incorrect or incorrectly-inflected words.

Other types of activities for advanced users consist of categorisation games such as:

- Simple categorisation: the user has to match the given word or expression with one of the categories provided by clicking on it.

- Word-to-word matching: the user establishes a relationship between words shown in different columns, for example, antonyms or synonyms.

- Theme-word matching: the user has to group words that do not belong to a given category by dragging and dropping them in the appropriate area of the screen.

Cree and Innu are polysynthetic languages that use classifiers. For example, the family of words related to *caribou* can be identified by the morpheme *atikw*. The activity illustrated in Fig. 5 below is created by indicating in brackets words that do belong to the given category, and in curly brackets, some words that do not. The markup language (explained in section 3) has the following form: **atikw** [**atiku**ss, **atiku**-uiash, nau**atiku**an] {atikamekw, atimunish, ushakapishu}. This form has proven easy to learn and use for our content developers.

Fig.5: A theme-word matching activity (caribou words) with its markup language

3   TOOLS AND TECHNOLOGY

3.1   Development history

99% of our online language games program is based on Open Source Software (OSS). The major benefit of OSS is low development and maintenance costs, as the main components are available for free. The development of the lessons has undergone four generations of evolution. There were three driving forces behind these evolutions: the pedagogical requirements, the dynamic evolution of web technologies, and a commitment to making this technology as inexpensive as possible.

When the games were initially developed in 2006, their main pedagogical goal, as discussed above, was to teach basic literacy skills for fluent speakers. The program was given a set of syllabic characters and a set of words containing them, for which it would generate exercises. The program's codename was "level 1". A new set of game types - needed to cater to more advanced learners - was built as a separate software in 2009, and given the codename "level 2".

By 2011, web technologies had advanced to such a degree that a major re-haul of the underlying base code of both pieces of software became necessary. This presented an opportunity to move both programs (levels 1 and 2) to lower-cost technologies and to develop new modules to allow for full control of game creation by non-technical content providers. Previous versions of the games required a content provider and a technical support provider working together to update the content of the game. The current version (2012) has tools that allow the content provider to add or update game content with very little training and no technical support.

3.2   Architecture

The architectural pattern used for developing the online language games is the Model-View-Controller (MVC) paradigm. This allows us to separate the data (the Model) from the data manipulation (the Controller), and create different types of presentation for the data (the Views). Whenever we need to create a new game type, we simply add a new view to our system and associated controller. This allows the program to be easily expanded with very few side effects. The architecture is being maintained throughout the evolution of the program.

Currently the programs have three modules for dealing with various aspects of the system. One module is responsible for generating the navigational aspects of the program. This module creates paths along which different activities and exercises can be found. The second module is responsible for user and content management. It keeps track of registered users' scores and helps content providers manage the content and game media. The last module, called the content module, is at the core of the system. It is responsible for generating the actual games available on the system.

As shown in Fig. 6, the content module has two main models: the Exercise and the Activity models. The Exercise model drives the "level 1" type of exercises. Its main purpose is to teach vocabulary and the writing systems. This model has access to a vocabulary list from which it generates exercises automatically. Each vocabulary item (in text form) is expected to have at least two media items (audio and image). The Activity model drives the "level 2" type of exercises. Its purpose is to teach grammar and improve reading and writing skills. This model requires markup language to generate lessons. A markup language is a set of symbols that determine the way text is presented. Each type of activity has a special markup language.
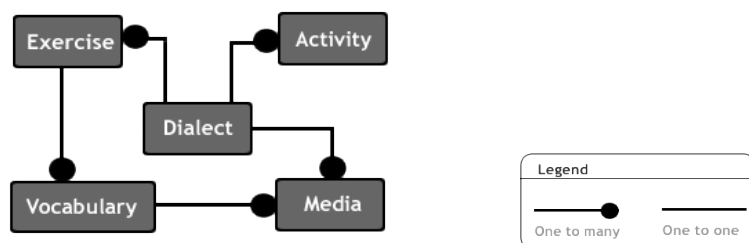

Fig. 6 The Content module

3.3   Open source programming languages

The online language games program is built on a web framework. This framework provides database access, session management and template tools that maintain a consistent look and feel across

interfaces. The first version of the program was built on a TurboGears Web Framework based on Python programming language. This framework used a MySQL database server as the Model storage, Kid template and Mochikit Javascript Library for the View(s). This remained through the first evolution of the program. However, CakePHP Web Framework based on PHP programming language was chosen as a replacement for TurboGears in the current generation of the program. CakePHP became an obvious replacement for two reasons. From a conceptual standpoint, TurboGears and CakePHP are very similar, despite their difference in programming language. This allowed for a smooth technical transition. The second and most important reason is the low cost of maintenance. CakePHP can run on most webservers (even on cheap webhost plans) and it is easy to learn for future updates.

3.4   Content provider and user interfaces

The game has two types of users: the general public and the content providers. The general public can either play games as anonymous or registered users. Registered users benefit from having their scores recorded. For every exercise or activity they have played, their best score is always available.

A content provider's job is to input data for the creation of all the various exercises and activities. Content providers in the earlier generations of the program worked with a system administrator to add or update content. However, in the current generation, the content provider does not need an administrator in order to proceed.

3.4.1   Vocabulary type games (code name 'Level 1')

The vocabulary type has two parts. First, a lesson consisting of vocabulary items and individual letters or syllabic characters is taught, and is then followed by exercises based on the words or characters in the lesson. Every lesson requires media. The exercises are based on six categories of games. The recognition category presents media and waits for the user to choose the associated orthographical representation. The matching category requires the user to match the given media with its associated orthographical representation. The missing category requires the user to choose the correct missing letter or letters. The memory category requires the user to match pairs of words or characters after a brief display. The word search category requires the user to find words hidden in a table of letters. Finally, the crossword category creates crossword puzzles, using sound or image as definitions.

There are three steps for the content provider to create the games. The first step is usually performed once, after installation. In this step, instances of the exercises from each category of game are created. The Cree language online game that teaches syllabics has a combined 21 instances of exercises available while the Innu version, which teaches roman orthography, has 13. This step is very technical because the required variables for the categories are set at this point to create an instance. A single variable value change creates a different exercise even within the same category. For example, one variable that needs to be provided is the source of words used in the exercises. Although every exercise has a lesson as a source, it is capable of sourcing from previous lessons as way to provide progressive learning. This variable can make an activity more difficult if turned on in one exercise and not in another. The next step is to create the navigation path for the games and choose the media for the lessons. In the final step, the vocabulary words or syllables for each lesson are chosen and the exercise instances are turned on. On the public user end, this generates dynamic exercises based on selected options (Fig. 7-8).
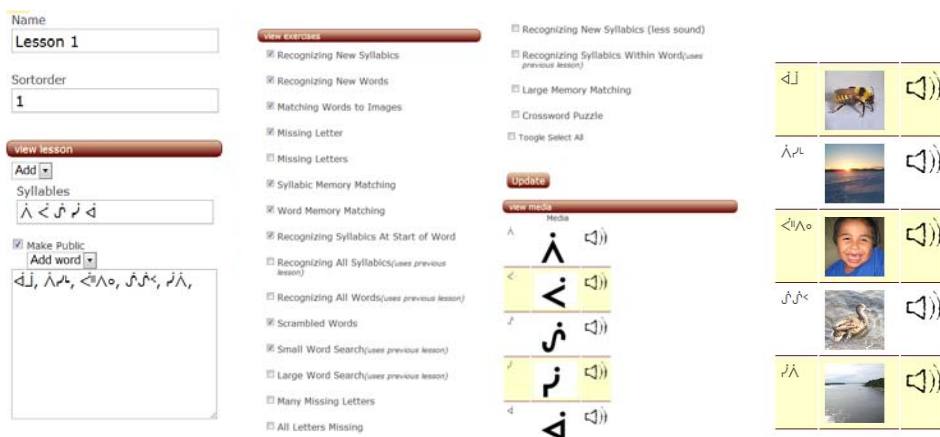

Fig. 7 Content provider interface for creating a lesson and activating exercises with media preview

Fig. 8 User interface with lesson and two corresponding exercises

3.4.2 Grammar and reading type games (code name 'Level 2')

This type of game is designed to enhance reading and grammar. There are dictation activities, categorisation activities, cloze text activities and different kinds of matching activities. Not all activities require media and there is capacity for long text. In this type of game the content provider has total control over each activity by the use of a specialized markup language. For example, in cloze text activities, the content provider can specify the missing parts of the sentence and the options made available to the player. The program, guided by markup language, generates the activity as specified. Fig. 9 shows data with markup and the generated activity:
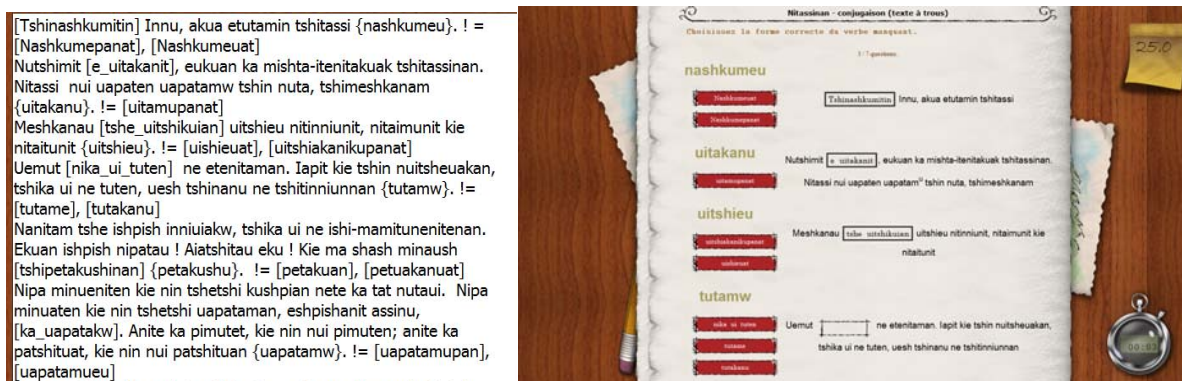

Fig. 9 Creating a cloze text exercise: markup by the content provider and corresponding activity

3.4.3 The Media Manager

The exercises and activities need to be associated with some media (sound or image or both). The system requires access to the URL of the media during game play. This could only be achieved by creating a system for mapping a piece of text onto a URL. The online games run on a webserver, making access to the underlying file system difficult, especially on servers without graphical interface. Our initial intent was to have the media filename match the text. This introduced three problems. The first problem was with non-roman orthography and the fact that most file systems do not support Unicode file naming. The second problem was the issue of homographs. The third and final problem was human error. The initial approach was to use a directory and file naming conventions that dealt with the first two problems since the human error factor is trivial. However, despite it being the most trivial, human error became the most expensive and the most time-consuming problem, because for every incorrectly spelled file, technical support was required to access the file system to perform the change.

In order to free the content provider from needing constant technical support, a Media Manager module was created for the current version of the game. The Media Manager module provides a graphical interface for content providers to upload, delete and rename media files. The Media Manager also provides the appropriate mapping between the text and URL by maintaining meta-data about the media files, for example: a dialect, a different orthography, a date, and so on. The meta-data is available to the content module of the game and this deals successfully with all three previous mapping problems.

The Media Manager module also allows us to integrate the multimedia collection with other language applications such as our online web dictionaries [3], [4].

3.5   Challenges: Compatibility issues

There are a few challenges with working with web browsers. The first challenge has been achieving cross-browser compatibility. Unfortunately, in the bid to win the ongoing browser war, the standards set by the World Wide Web Consortium (W3C) have become a casualty in that browsers would rather promote their own standards than conform to the W3C's. As such, making the games look and feel the same across all browsers is a distant reality. The heavy reliance of the program on Javascript for interactivity makes it vulnerable to browser updates that affect the Javascript engine. Finally, the most difficult problem to overcome and the reason why this program is not 100% OSS, is media playback in browsers. It is very difficult to achieve cross-browser media playback without a commercial proprietary media player. At this point we are using a customised flash-based media player. When W3C's HTML5 standards become the 'de facto' ones of all browsers, these problems will be easily remedied.

4   CONCLUSION

As is the case for most minority and Aboriginal languages, the general lack of resources forces one tool to be used for many purposes. Our language games are no exception. Originally developed for the literacy training of fluent speakers, they quickly became used for second-language training, vocabulary acquisition, and grammar learning. There was also a need to integrate the games with other resources such as an online grammar and dictionaries [5], [6], as well as with social media such as blogs and Facebook. Our next steps will consist of a fuller integration of these tools, as well as the creation of teachers' modules for class management and online learning. Since its inception, the project has provided training to Aboriginal educators in the areas of computer technology, sound editing and web content management. Team-based design of the tool as well as the games' contents has allowed for flexibility and incremental development, as well as the transfer of resources from one language group to another.

REFERENCES

[1]   Junker, M.-O. & Stewart, T. (2011). A linguistic Atlas for endangered languages: www.atlas-ling.ca. *Edulearn11 Proceedings*.

[2]   Junker, M.-O. (ed.) (2012). *The Cree-Innu-Ojibwe linguistic atlas*. www.atlas-ling.ca

[3]   Drapeau, L. (1992). Bilan de l'instrumentalisation et de la modernisation dans les langues autochtones. In Jacques Maurais (ed.) *Les langues autochtones du Québec.* Québec: Publications du Québec.

[4]   Burnaby, B. (2005). How Have Aboriginal North Americans Responded to Writing Systems in Their Own Languages? Unpublished. Memorial University of Newfoundland.

[5]   The Eastern James Bay Cree Dictionary on the Web: www.eastcree.org/cree/en/dictionary/

[6]   Dictionnaire Innu Dictionary Aimun-Mashinaikan: www.innu-aimun/dictionnaire