White Paper
Language Learning App

Contents

1 Project Abstract

This project collaborates with speakers of the Mi'gmaq language in Listuguj, Quebec to make an app that helps learners of all ages speak Mi'gmaq. The app is also generalizable to other languages, and has the ability to sync content with LingSync, an app for linguists to store data corpora. The Learn Mi'gmaq app is customized for minority languages specifically, emphasizes verbal and aural communicative competence, and allows students to learn offline as well as online. Additionally, it allows students to make their own lessons, tailoring the material to their own interests and needs. It accomplishes these tasks by learning from the triumphs and mistakes of the language-teaching apps already on the market, and building off the programming and linguistic experience/expertise of its designers.

2 Statement of need

The Learn Mi'gmaq app is designed to fill a gap in the current collection of apps designed for language-teaching. This gap occurs where learners want to acquire their her-

itage language in a context where it is not the majority language. Our app is a response to the case of Mi'gmaq (Canadian Maritimes), where there are approximately 5,000 speakers, the vast majority of whom are over 50 years old. This is a common pattern in Aboriginal languages of North America: grandparents and great-grandparents can speak the language, but their children and grandchildren cannot. There is now a strong desire in those latter generations to reconnect with their heritage and consequently their heritage languages.

Community members in Listuguj have expressed the desire for mobile language-teaching software that will support the existing language revitalization efforts there, and the apps already on the market were found wanting. Though there are many existing apps for language-teaching, most of them match the profile below and fall under some of these limitations:

- **Text-based** Despite the immense importance of speaking competence, most apps on the market (with some notable exceptions like *Listen and Speak* and *Babbel*) focus on text-based language-learning and never require the learner to speak the target language. This quite obviously impairs the learner's own fluency when communicating in that language.

- **Expensive** Many established apps (such as *Rosetta Stone*, *Fluenz*, and the somewhat lesser-known *Tell Me More*) are prohibitively expensive, costing $500 or more for an individual learner to have access to the full range of materials in their target language. This financial aspect limits the access of potential learners and shrinks the pool of people who may use and learn from the app.

- **Majority language** Existing apps appeal to the largest markets by teaching popular business or travel languages. These have the advantages of a well-advanced technical and operational side. However, teaching a majority language requires different resources for the student than teaching a minority language.

  - **Different needs** Majority language-learners are often travelling to a new place, where they will be immersed in that language. Their learning requires preparation in areas like "reading signs," but absolutely no preparation in areas like "how to convince your grandmother to stay in Mi'gmaq when you can converse easily in English instead."

  - **Help for minority language learners** It is this latter area that minority language-learners often need assistance in the form of tips, guidance, and specific phrases and vocabulary to steer conversation. Considering this clear difference in context, simply creating a Mi'gmaq unit of, for instance, *Rosetta Stone*, would be a misguided endeavour.

  - **Existing minority language apps** There are existing apps that teach minority languages (for instance, those made by *Thornton Media Inc.*) but these apps

tend to take the form of phrasebooks or talking dictionaries, with little guiding structure to introduce the student to the language for more complicated communicative purposes.

- **Difficult balance of didacticism** Children learn language in an immersion atmosphere apparently effortlessly. However, when adults learn a language, they often seek out the guidance of explicit grammatical rules. Striking a balance between full immersion and over-explaining concepts can be difficult. Some apps that stand out as being well-constructed in this regard are *Babbel* and *Busuu*.

- **Read-only and impersonal** These two aspects go hand-in-hand, and are a large problem with existing programs. Language-learning apps are designed to be consumed: they deliver a specific amount of content, and the learner learns only that content. Even the best app cannot promise learners that they will find material tailored to themselves–the lessons that are relevant to them may be buried amongst others, or simply missing. Some programs (such as *Memrise*) allow learners to create playlists of pre-existing materials, but none that we know of allow learners to work with speakers to create and upload their own unique content.

Our team is taking these advantages and limitations into consideration while designing this app. Additionally, the principles of The National Foreign Language Center's STARTALK program (http://startalk.umd.edu/principles/) will inform the solutions created for the pitfalls of other language-teaching apps.

3 App description

Our app guides learners through the process of language acquisition in a way that allows them to learn at their own pace. The linguistic information is grouped into thematic units (i.e.: *Food*, *Family*, *Hockey*) that students may select according to their individual interests and needs. Within these units, the student encounters lessons that increase gradually in difficulty–the community has already recorded over 100 lessons that, pending speaker permissions, may be input directly to the app. Each lesson opens with a "heads-up", a short introduction to what the student should focus on within the lesson's material. The lesson material then comes in **three** parts,

- First, a **dialogue set** consisting of a short video dialogue and comprehension questions. The comprehension questions are delivered aurally as well as with time-synced text. The student responds to the questions by recording themselves speaking and comparing their answer to the audio of a sample answer.

- Second, a collection of **vocabulary** items. These include a picture/video showing the definition, a recording of a native speaker saying the word/phrase (with

time-synced text), and a space for the student to record themselves saying the word/phrase. Included in this portion is a speech analyzing tool that informs the student how accurate their repetition was.

- Third, a varied set of **skill-developing exercises and games**. Some games will simply be a flashcard-style review of the material the student just learned. Other exercises (such as Hangman) will be aimed at improving the students' skills with orthography and general familiarity with words. Others, (such as Build-A-Word) will highlight some fundamental aspects of Mi'gmaq, specifically the fact that many of its words are long and made up of smaller parts; playing with these words in the context of a game will help the students overcome the initial uneasiness many students feel when confronted with very long words.

Students also have the option of adding a **tutor**, a person whose grasp of the language is greater than theirs. This tutor is able to check up on their learner's progress to give them support and feedback, and respond to questions the learner might have. Tutors are a way of integrating the current speakers of the language into the learning process without forcing them all into the role of "teacher"; the creation of a decentralized support network for the students

In addition to overcoming certain shortcomings of other apps, we aim to innovate with respect to how the material for the app is created in two ways. Firstly, we are **crowd-sourcing** some data creation by allowing students to work with speakers in their community to build their own language lessons. The app evolves with the learner and tailors itself to their needs, with an easy-to-use interface to create personalized lessons. These lessons will have multiple privacy settings so the student may keep the lesson private, or publish it for other learners to see (after it has been verified by another Mi'gmaq speaker). Being able to produce content means that students can use this app in the context of a class assignment, or as a tool in one-on-one study with a speaker in the Master-Apprentice Program.

Secondly, this project is **collaborative** with LingSync, an app for field linguists to collect, store, and share linguistic data. Considering the large overlap between data that linguists use to theorize about a language and lessons learners require to learn the language itself, we plan to make it simple and easy for linguists to import their data to this app as language lessons.

We want these lessons to not only teach and revitalize, but also preserve and document– this entails another aspect of the collaboration with LingSync. In addition to linguist-created materials becoming language lessons, learner-created lessons can optionally be synced with the linguists' app. In LingSync, an interested linguist may then examine the data created by the learner, and the observations of the linguist may help gain some insight into the foundations of the Mi'gmaq language.

Lastly, it is very important to the team that this app be fun for students to use. To that end, we will make it easy for students to track their progress, and give achievements

for specific milestones that the learner has reached. It will also be easy to share these progress reports and achievements using a social networking site such as *Facebook*; this will help learners stay in touch with each other, and demonstrate to the speakers in their lives just how seriously they are taking their commitment to the language.

3.1 Functionality

- **Offline** The app is basically offline (with online sharability), users do not have to sit in front of a computer to do lessons. Packaged for Android, users can download and use the app on their mobile devices (phones, tablets) with or without wireless connection. (accessible)

- **Listen-Speak oriented** Lesson materials are presented in audio format with accompanying pictures or videos.

- **Check progress**

- **Individualized/Customizable** Using Android's camera and audio capacity, users can take their own pictures and audio recordings and create their own lessons according to their needs and interests. This functionality opens the possibility of using the app for classroom activities and encourages users to talk to native speakers in their community. Helps immersion. (applicable to classroom, fun to use, connect to community)

- **Sharable** Possible to share lessons with other users (students) and teachers. Teachers can check students progress and correct mistakes. (applicable to classroom, connect to community)

- **Connected to LingSync** Minority languages tend to have little learning/teaching materials while the data collected for linguistic research purposes can be large. The learning app is connected to LingSync, a database app, which allows teachers to extract data relevant for language learning from the LingSync database and convert them into lessons.

- **Organization** A unit consists of lessons, a lesson consists of dialog, vocal and exercise sections. Units can be organized by themes (food, animal, weather etc.) or target grammatical constructions (question, past tense, etc.).

4 Goals and Objectives

- Upload all existing lessons and expand on them (exercise development)

- Fully implement and facilitate the make-your-own-lesson tool (include sharability)

- Implement sharability (integrate with Facebook etc)

- Implement architecture for "game" atmosphere (gamify)

- Become fully syncable with any linguist's data from LingSync

5 Budget and timeline

The prototype of the language learning module was began on September 13th 2012 and finished on October 3rd, 2012.

Iteration	Hours	Technology
– Prototype		
Software Architecture Design	55	Software Engineering
Lesson Datum Model	15	Javascript
Lesson Datum Listen and Repeat View	20	Javascript
XML import logic	6	Javascript
Corpus to Language Learning logic	10	Map Reduce
Main Manu Dashboard	6	HTML
Student LIsten and Repeat Dashboard	10	HTML
Student Instructions View	6	Javascript
Audio Play/Record View	10	Javascript
Audio Text Time Alignment logic	20	Javascript
Audio Record Android logic	6	Java
Android Packaging	15	Java
Software Architecture Design	40	Software Engineering
Teacher User Model	10	Javascript
Teacher User View	40	Javascript
Language Learning Corpus	10	Javascript
Language Learning Corpus View	40	Javascript
Language Learning Corpus Dashboard	10	HTML
Student User Model	30	Javascript
Student User View	60	Javascript
Datum Shadowing Lesson Model	10	Javascript
Datum Shadowing Lesson View	10	Javascript
Student Shadowing Lesson Dashboard	40	HTML
Datum Quiz Multiple Choice Model	20	Javascript
Datum Quiz Multiple Choice View	20	Javascript
Student Quiz Dashboard	120	HTML
Datum Prompted Production Model	20	Javascript
Datum Prompted Production View	20	Javascript
Student Prompted Productions Dashboard	40	HTML
Feedback Comment Model	10	Javascript
Feedback Comment View	40	Javascript
Comment Teacher Feedback logic	6	Map Reduce
Audio import to Multiple Utterances logic	60	Praat & Javascript
Audio File Server	140	Node

Table 1: The Language Learning Module enables to use data in the database to create lessons for language learners.