A Mobile Vocabulary Acquisition Application for Health Science Students: a Proposed Study

Ilana Wilken\textsuperscript{1}, Febe de Wet\textsuperscript{1} & Elsabé Taljard\textsuperscript{2}

\textsuperscript{1}Human Language Technology Research Group, Meraka Institute, CSIR, Pretoria, South Africa
\textsuperscript{2}Department of African Languages, University of Pretoria, Pretoria, South Africa

IWilken@csir.co.za, FDWet@csir.co.za, Elsabe.Taljard@up.ac.za

Abstract

Communication plays a vital role in everyday life and in some situations multilingual communication is a necessity. The South African constitution recognises 11 official languages. Multilingual communication therefore occurs almost everywhere, like at hospitals and in clinics. When the need arises for someone (who does not speak a global language like English) to seek medical attention in South Africa, it becomes a challenge to find a health practitioner that speaks their first language. This study aims to develop and evaluate the use of a mobile application designed for supplementary Northern Sotho language learning by health science students.

Index Terms: mobile-assisted language learning, human-computer interaction, user experience

1. Introduction

South Africa is home to eleven official languages and although it is expected of medical professionals to communicate in the languages of the area where they work few of them do. Communication barriers and the associated challenges experienced by health professionals are widely documented in South Africa’s health care sector [3, 4, 5, 6]. Health professionals hardly ever speak their patients’ first languages. This communication barrier is especially problematic in the rural communities of South Africa [7]. Medical students often start working in areas other than those they grew up in, so they have to learn another language to function in professionally, and this is sometimes challenging [1].

Language acquisition consists of three language subsystems: grammar, phonology and vocabulary [8]. Even though vocabulary is of vital importance in language acquisition, teachers often neglect it and rather focus on grammar and phonology [8]. It has been claimed that there is an increased chance of communicating in a foreign or second language if the speaker knows well-chosen basic vocabulary in the target language than when they have mastered the language’s grammar [8]. This is especially true of communication in a specific domain. The domain in which a medical practitioner and a patient find themselves in, adjusts the size of the vocabulary to that which is relevant for the situation. For example, when discussing a patient’s broken leg, the only vocabulary the health professional will need is vocabulary relating to the patient’s leg and the procedures around it. The language learning issue is addressed to some extent by South African universities through teaching health science students an additional language.

Mobile applications can be seen as extensions for learning in new and different environments and this in turn allows language acquisition not to be limited to the classroom [9]. This can be a big advantage for busy students who can use portable devices to learn whenever they have free time available [9]. Mobile applications combine different elements to achieve the goal of supplementing traditional learning methods. Two of these elements are addressed in this study: 1) mobile assisted language learning and 2) human language technologies.

1.1. Mobile assisted language learning

Mobile assisted language learning (MALL) is a subarea of mobile learning (mLearning) [10]. Even though scholars are not in agreement about whether mobile means the mobility of the technologies or the mobility of the learner, it is often stressed that it refers to the mobility of the learning content [11]. Computer assisted language learning (CALL) on the other hand takes place when students or learners use computers during a language course [12].

Although many commonalities between MALL and CALL are observed, it has to be noted that there is a difference between the two. MALL differs from CALL because it can be used in personal and portable devices and it enables students or learners to learn in new and spontaneous ways [11].

1.2. Human Language Technologies

Human language technologies (HLT’s) have the goal of getting computers to perform useful tasks involving human language [13]. Two of these tasks include improving human-human communication and the processing of text or speech [13]. Text-to-speech synthesis (TTS) and automatic speech recognition (ASR) are two elements of HLT that are relevant to this study. TTS systems automatically generate speech from text input [14] and ASR systems recognise the user’s speech and responds accordingly [15]. HLT is used within MALL programmes to enhance the technology and offers learners an environment in which they can use multi-media to learn a new language.

South Africa’s languages can all be regarded as being under-resourced in terms of developing HLT’s [16]. Even South African English can be classified as an under-resourced accent of English [17]. The situation is much worse for the indigenous languages that cannot leverage on existing resources in other languages [17]. A special effort has therefore been made to collect resources and to develop basic HLTs for all 11 languages [18, 19]. The ASR and TTS systems developed for Northern Sotho are used in this study.
2. Previous work

There is a large number of MALL applications on the internet. These applications range from basic to advanced language learning (like Duolingo [20] and Babbel [21]), to domain focused applications (like Mobile Xhosa [22]). Duolingo is a mobile application for European languages. Users get to learn a new language in an entertaining way through grammar, writing, listening and speaking exercises. Babble is an application similar to Duolingo, but only the first lesson of each category is free. Users need to purchase new lessons if they want to continue using the application. Mobile Xhosa is also a mobile application, but its focus is mainly on teaching Xhosa to medical professionals in South Africa to assist them when they communicate with patients who can only speak Xhosa. The application provides users with basic medical phrases with translations in English and Xhosa, as well as a translator to translate other words or phrases into English or Xhosa.

3. The proposed study

Northern Sotho is one of South Africa’s official languages and the University of Pretoria offers a Northern Sotho language learning course to students in the health care sciences. As mentioned previously, language learning applications with the same goal in mind already exist, but to the authors’ best knowledge, none exists for Northern Sotho in the health domain. This was the biggest motivation behind this study. Because of the need for a set of specialised vocabulary items suited for specific purposes and situations in which the students will function, a MALL application was developed for the health domain. The application focuses only on the vocabulary subsystem of language acquisition and health science students can use the application to practice their Northern Sotho skills that involve vocabulary. A content management system (CMS) was also developed which is used to register users, upload lessons, assign specific lessons to specific users, and to store any data recorded by the user.

3.1. The application

One application was developed, but with three different modes. The application was developed for tablets using the Android operating system. The modes are:

- Silent (vocabulary acquisition without TTS or ASR)
- Listen (vocabulary acquisition with TTS)
- Speak (vocabulary acquisition with TTS and ASR)

The application is able to work with or without internet access. Internet access is only needed when a user logs in for the very first time in order to download the lessons to the device. Once a lesson has been completed and an internet connection is detected, the data on the device is sent to a database.

3.1.1. Functional description of the application

The basic functionality of all three modes of the application are the same, as described under Silent mode.

Silent mode

This mode only has the basic functionality of the application. As can be seen in Figure 1, the user sees a question (a word or a phrase) in English and this is supported by an image. Three Northern Sotho translations of the English word or phrase are given as possible answers. The user will then have to choose an answer. The user gets two chances to answer correctly. If they fail to do this, the correct answer is shown and the application proceeds to the next question. Figures 2 and 3 show the feedback given to a user when they get the answer incorrect and correct, respectively.

Figure 1: An example of a question.

Figure 2: The feedback given to a user when the incorrect answer was chosen.
Listen mode
This mode has an extra functionality where the pronunciation of the correct answer is automatically played by means of pre-rendered Northern Sotho TTS, as shown in Figure 4, as soon as a user gets an answer correct. When a user has unsuccessfully attempted to answer the question two times, the user will still get the opportunity to listen to the audio of the question. This forces the user to listen to the audio at least once, after which they will also have the opportunity to listen to the pronunciation as many times as they like by pressing the ‘Listen’ button.

Speak mode
This mode has a second extra feature in the form of ASR, along with the TTS pronunciation. As indicated by Figure 5, with the ASR functionality, the user has an opportunity to practice their pronunciation of the answer. After getting a question correct, the user gets three attempts to correctly pronounce the word or phrase. If the user gets the pronunciation correct, the application moves on to the next question. If after three attempts the user has not pronounced the word or phrase correctly, the application automatically moves on to the next question. Again, even when a user has unsuccessfully attempted to answer the question two times, the user will still get the opportunity to listen to the audio of the question and to pronounce the word or phrase.

Text-to-speech
The TTS ‘voice’ that was used was developed by the HLT Research Group at the Meraka Institute of the CSIR in Pretoria, South Africa. Because the application is functional without internet access, it was decided to use pre-rendered TTS audio files that will be downloaded with the lessons and stored on the device. TTS was also chosen because it offers flexibility when new pronunciations have to be added to the system, as a voice artist doesn’t have to go to a recording studio to record new words. The TTS is also used to provide the users with a target pronunciation and to practice their listening skills, because first language speakers are not always readily available to listen and speak to.

Automatic speech recognition
The Northern Sotho ASR system was developed by the HLT Research Group at the Meraka Institute of the CSIR in Pretoria, South Africa. The system is fully functional and is used in the application. The main aim of the ASR is to elicit spoken

---

1http://www.qfrency.com/
output. A global form of feedback is given in terms of judging whether an utterance is correct or incorrect in its entirety, but no feedback is given on the pronunciation itself. The application’s purpose is thus to provide the students with an opportunity to practice their oral skills, which is not always possible in class due to time constraints and students are often too shy to speak in front of their peers.

The Northern Sotho ASR system that is used in this study was developed using speech produced by native speakers of the language [23]. It is well-known that non-native speech is more difficult to process accurately using ASR than native speech [24]. The recognition library will thus be restricted in order to make the ASR in this application more sufficiently robust. Very few resources are available for Northern Sotho itself and there is currently no speech data available for non-native speakers or learners of the language. It was therefore decided to capture all the speech input that will be produced by the users during application usage. The data can be used to improve the automatic recognition of the non-native speech for subsequent versions of the application and to start investigating the possibility to enhance the speak mode of the application by providing feedback on the users’ pronunciation.

3.1.4. Scoring

It was decided to include scoring in the application as a method of keeping track of a user’s progress, as well as to motivate a user to keep using the application. There are two scores in the application: a score for the number of times the user correctly selected the correct Northern Sotho answer for the English question; and a score for the number of times the user correctly pronounced the word or phrase on their first attempt.

3.2. Content management system

The CMS provides a platform for lecturers, teachers, researchers, etc. (known as administrators) to upload content that will be used in the application. An administrator can register users, create groups for each class, upload lessons (including graphics and the pre-rendered TTS audio files) and view all data related to the application. This data includes the usage statistics of the application as used by the users, as well as the utterances made by the users when using the speak mode. The statistics can be viewed in an easy-to-read format. With the statistics, the administrators can control which users have completed lessons, which lesson(s) users are struggling with, how many times a user has done a particular lesson, etc. An administrator can also listen to all the audio files that were recorded by the users.

3.2.1. Content

The content used for this application is the same content the students are given in their lessons during lectures. It was decided to only use the vocabulary and phrase list given in the course study material as it is suitable to transfer to the application environment, and because the main focus of the study is on vocabulary acquisition. The content was received in the form of a vocabulary list which was divided into categories matching the different medical courses the students are enrolled for (e.g. dietetics, nursing, etc.). The lists were combined into one list, which was then divided into different sections/lessons, namely body parts, internal organs, food, etc. The graphics are of a high quality and match all the words and phrases, in order to strengthen the impact of a lesson.

3.3. Connectivity

Broadband connectivity is not yet available in most places across South Africa, and thus it was decided to design the application in such a manner that it only uses intermittent internet access. Therefore TTS and ASR are able to function offline. Figure 6 is a representation of the connectivity between the CMS, the internet and the application on a device.

The CMS is accessible on the internet. This is where the administrators can upload new lessons, register new users, etc. When a device is connected to the internet (online), all the content allocated to a specific user is accessed directly from the CMS and is pulled onto the device. Once a lesson has been completed, all results and data are sent directly to the CMS where it can be viewed by the administrators.

A solution to working online when there is no formal internet connection, is by means of a laptop that is used as a server. All the data that is on the CMS is synchronised with the laptop. Together with a router, the laptop creates a closed wireless connection. The users can use the devices and complete lessons like they would if were they online. All data is sent to the server, which after being synchronised with the internet, sends all data to the CMS.

When a device is completely offline a user can still use the application and complete lessons, given that they have logged in at least once when connected to the internet. This is necessary because the lessons have to be downloaded onto the device. Once the user has completed a lesson, all the results and data are stored on the device. As soon as the device detects an internet connection when the user logs in again, the results and data will automatically be sent to the CMS.

Figure 6: A representation of the connectivity between the CMS, the internet and the application.

3.4. User tests

A successful efficiency evaluation is difficult to do in these circumstances because a multitude of factors could have an influence on the students’ development. Thus effectiveness will not be evaluated during this study. The students will however write vocabulary tests before and after the evaluation period, but this cannot be seen as rigorous evaluation on progress because of the many influencing factors present during such a study.

The students that will participate in this study will be first language Afrikaans and English speakers and all the students will be provided with the exact same hardware. The main focus of this study will be on evaluating user experience. At the end
of the evaluation the usage statistics collected by the system and the answers received through the questionnaires will be compared. In order to evaluate the user experience, the students will use the application once a week for twelve weeks in class as a supplementary tool to their formal studies. After twelve weeks, the students will complete a user experience questionnaire and in the following week, a focus group discussion will be held. The results will be used to not only determine whether or not the application has a place in the classroom as a supplementary learning tool, but also to establish if the students prefer one mode of the application above another.

The questionnaire consists of open- and close-ended questions, and are asked to collect feedback on the user friendliness of the application, as well as the preference the students have to one mode of the application above another. The user experience questions include questions about the users’ first impression of the application, whether or not the application was easy to use, what they liked most and least about the application, what they would change, whether or not they would recommend the application to other Northern Sotho students, etc. The user preference questions include questions about which mode do they think helped them most with their vocabulary, reading, speaking, and grammar skills, which mode they preferred overall, etc. All the answers to the questionnaire’s questions will also be used to improve future versions of the application.

4. Conclusion

The application’s main aim is to act as a supplementary tool for students learning a new language, and not to replace their current learning material and methods. The user study will be done in collaboration with the lecturer of the language course in order to maximise its impact.

5. References