Design of a Transcription Tool for the Kelabit Community of Bario, Sarawak

Emmy Dahliana Hossain, Sarah Flora Samson Juan, Jane Labadin and Priscillica Agas
Faculty of Computer Science and Information Technology,
Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak.
hedahliana@unimas.my

Abstract—**This paper describes the design of a transcription tool developed for the Kelabit community of Bario, Sarawak based on the community's requirements and feedbacks. Transcribing is the process of making a full written copy of spoken or dictated material. Using this transcription tool, users can listen to audio files of recordings and proceed to do the transcribing within the tool itself. Afterwards, they can save the transcription into text files for future uses. Users can control the audio file while it is playing, and do updating or editing to the transcription. This project was carried out on the motivation to assist in language preservation works as Kelabit is an under-resourced language.**

Index Terms—**Transcription Tool; Kelabit Language; PRISMA.**

I. INTRODUCTION

Software users have expectations when using software, and when their expectations are not met, users may find it difficult to use the software, and may abandon using the software out of frustration with the difficulties encountered. The digital divide between urban and rural population could also affect the way users handle software – for those who are used to using computer technology, they may find it easier to navigate around when using software. The users who do not use much computer technology may not be able to do so. Hence the researchers of this project believe that it is important to obtain user requirements, in order to find out the users' expectations and their level of computer technology proficiency, before designing and developing the software for the target group of users.

Transcribing is the process of making a full written copy of spoken or dictated material (transcriptions). In computer science, transcribing is the process of transferring information from one recording and storing system to another. This involves using a type of software called transcribing tool to do the transcription work: the spoken or dictated material that is in audio file is listened to by the person doing the transcribing (called the transcriber), who will then type the transcription text on the tool's text editor. This will then be saved, and the transcription can be stored in the form of text files for further usage in other endeavours. The transcriptions stored can be considered as language resources that can be used in human language technologies such as machine translation, speech recognition etc. These technologies require large amounts of resources (data) in order for it to be developed and move forward. With the availability of these resources, it is hoped the growth of language technologies can be spurred on. In time, the language can be preserved, documented and ultimately – considering its threatened status – prevent its extinction.

There are 63 known indigenous languages in the Malaysian state of Sarawak on Borneo island. One of these languages is the Kelabit language, spoken by the Kelabit community, which originates from the remote mountainous region of Bario in northern Sarawak. The Kelabit language is considered an under-resourced language, which means the language do not have sufficient resources to maintain itself [1]. Examples of language resources are speech and text data, linguistic resources (including computational linguistic resources such as corpora, pronunciation dictionaries, vocabulary lists, grammars, taggers, morphological analyzers, parsers, speech recognizers, annotation standards and tools etc.), and literature as well as language technologies (automatic speech recognition, machine translation, etc.) [2]. In addition, there are not many linguistic experts in this language, as well as the language lacks presence on the World Wide Web. All these contribute to the language being under-resourced, and may lead it to extinction if there are no efforts to preserve it in digital formats, and if no language technologies are implemented on time for the language. As of August 2016, Ethnologue [3] classifies the Kelabit language as "Level 6b – Threatened" on the Expanded Graded Intergenerational Disruption Scale (EGIDS). This denotes that while the language is spoken within all generations, it is facing declining usage among its speakers, especially among the youth. Some members of the older generation in Bario have stated in the interviews that the Kelabit language spoken by their younger generation is either "diluted", "mixed" or "not the pure Kelabit language anymore", which comes as a result of assimilation with other languages such as Malay and English.

This paper describes the efforts undertaken to design a transcription tool software for the Kelabit community in Bario, Sarawak, Malaysia. The rest of the paper is divided into the following sections: related works, method, discussion, acknowledgements and references.

II. RELATED WORKS

Transcribe is a simple Chrome web app that allows audio uploading and the transcribing process does not need the switching between the audio player and the text editing document. Transcribe is preferable in Journalism.co.uk such that they opt to record Skype and phone interviews to focus on the conversation and refer back in other time [2]. This tool is a time-saving tool as user can transcribe quotes or sections from an audio interview and toggle between the audio player and the text-editing document. This web app offers an audio player which is integrated with a text editor on the screen interface. User can upload an mp3, mp4, m4a, amr, wma, aac,

or wav file, listen to the audio file and transcribe within the box underneath the player. User can slow down, pause, or rewind audio and insert timestamp anytime by using a single-keypress keyboard shortcut. Apart from that, this app allows user to speak thru the microphone as the dictation engine supports fourteen languages mostly on the major European and Asian languages. Other than that, Transcribe works offline too as everything works locally on the browser.
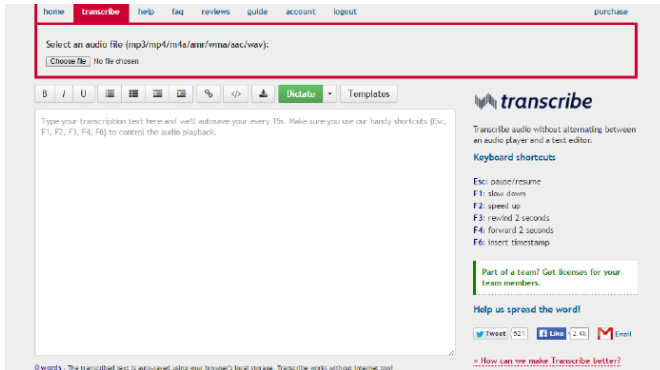

Figure 1: Transcribe

oTranscribe was created by Elliot Bentley, a journalist who aims to make transcription easier [4]. This application has an audio player and also a text editor as such the other transcribing tools. It does not need to toggle between the audio player and the text editor during the file transcription process. oTranscribe allows the audio uploading which supports mp3, ogg, webm, and wav file formats. In addition, this app has an in-built file converter too. This application also allows the usage of keyboard shortcuts to play, pause, add time stamps, slow down, and speed up the audio play. Moreover, the transcription is auto-saved. Hence, if the internet connection lost or the browser is closed, the transcript file is not lost.
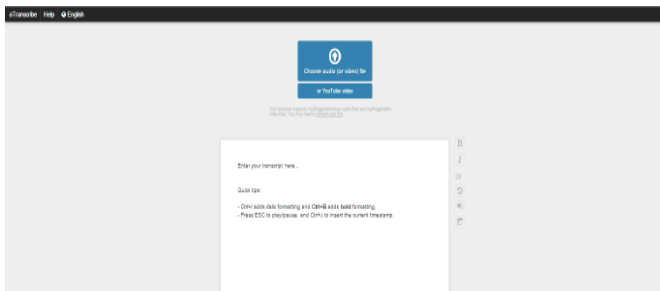

Figure 2: oTranscribe

Listen N Write a free tool designed specifically for transcription. It doesn't use speech recognition or automatic transcription. It's simply a media player bundled with a text editor, but a media player with subtle but clever differences that make it a great tool for playing back recorded lectures and speeches (for example) and writing down what you hear. Listen N Write's interface consists of three pieces: the media player and the optional text editor and bookmarks box. By default, the tool only opens the media player and text editor. The bookmarks and text editor can be hidden through the media player's view menu.
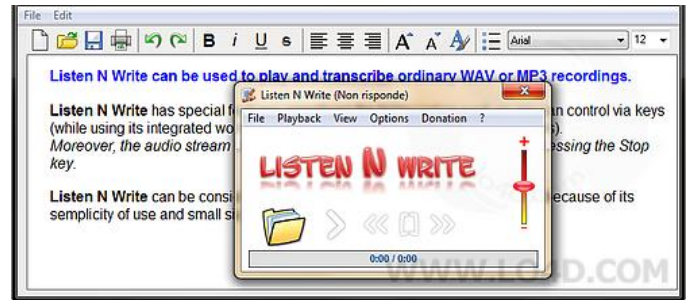

Figure 3: Listen N Write

III. METHOD

As we believe that community participation is important, we used the PRISMA model for this project. PRISMA is short for Participatory Action Research in Software Development Methodology Augmentation [5]. PRISMA combines two parts – technical and non-technical, namely the software development process (the technical process), and the social change process (the non-technical process). The software development process used here is the conventional Software Development Life Cycle (SDLC), while the social change process includes "the change the community wants, the reasons they want it, as well as the roles for people inside the indistinct world of political and social systems, multiple disciplines, environments and multiple stakeholders". These two processes are encompassed by eight steps: community selection, rapport building, problem exploration, hypothesis formation, design/methodology, implementation/testing, and evaluation/reflection.

The following describes how we carried out the steps in PRISMA in developing the transcribing tool.

Step 1: Community selection

The first step in PRISMA is selecting the community to work with. The Bario Radio, a community-run radio was launched in 2011 in Bario to serve the Kelabit community in broadcasting contents such as news, entertainment, information and talk shows with visitors [6]. This radio project is part of the successful e-Bario project. Volunteers perform broadcasting duties twice from Mondays to Saturdays. The contents broadcasted on air can be captured in audio files, and in turn these audio files can be transcribed. Hence, the Kelabit community in Bario is selected for this project.

Step 2: Rapport building

Next, after the selection of community is done, the rapport between researchers and community needs to be build. UNIMAS's e-Bario project has already established rapport between researchers and the community for the past few years. Both researchers and the community are comfortable with each other to discuss and exchange feedbacks and knowledge. The community understands that by having this transcribing tool, it can assist in helping them create digital versions of their oral contents, and in turn, assist in preserving their language and culture.

Step 3: Problem exploration

Then, the problem that is being faced needs to be identified, explored and confirmed. Kelabit language is an under-resourced language, and there is a need for software for building language resources. There are existing

transcribing tools/methods that are currently available, but these tools may not be suitable to be used by the community for various reasons (level of ICT skills, tool's features that may be complex to the users, etc.). Four existing transcribing tools/methods were selected to be tested by the respondents – manual transcribing, using a word processor (Microsoft Word) while listening to the audio on a sound player (Windows Media Player), Transcriber, and Transcription Aid. During the sessions, the respondents were interviewed before they tested out the tools/methods, observed when they were carrying out the testing, and then interviewed again after they have completed testing the tools/methods.

Step 4: Hypothesis formation

After the problem has been confirmed, the next step would be to form the hypothesis for this research. The research question used is: can PRISMA be used to build software for building language resources (namely this transcribing tool)? The expected outcome would be to assess existing transcribing tools/methods, produce the design of a transcribing tool that is customized for the selected community, and to perform transcription of the captured audio files.

Step 5: Design/methodology

In this step, the solution to the problem identified is designed, as well as the approach to the solution is identified. The solution is identified as building the transcribing tool (the software for building language resources) based on the needs of the community. The design of the transcribing tool is obtained through interview sessions with the participants, and then analyzing the results to come up with the design of the transcribing tool. Participants are asked to test out the four existing transcribing tools/methods and then interviewed on their experiences using the tools/methods. They are also asked for their feedbacks and recommendations to build proposed transcribing tool. These data collected are then analyzed to get the user requirements for developing the transcribing tool.

Step 6: Tools & techniques

Participants are asked to test out the existing tools/methods on paper (for manual transcribing) and a laptop computer. The interview data is recorded on a voice recorder, and as well as written down. Once the user requirements have been obtained, the tool will be designed and developed using JavaScript, HTML5 and CSS.

Step 7: Implementation/testing

The prototype of the tool will be developed and tested in this step. The same group of participants interviewed during data collection for user requirements will be asked to test the tool and give their feedbacks. The tool will be developed using the conventional Software Development Life Cycle (SDLC) methodology.

Step 8: Evaluation/reflection

Finally, as a way to determine if this project has achieved its goals, an evaluation and reflection will be carried out. Participants will be asked to evaluate the transcribing tool to check if the tool is usable and acceptable by the community. Researchers will reflect on the lessons learnt throughout this project.

IV. FINDINGS AND DISCUSSION

In order to collect the requirements data for this study, the respondents – Bario villagers and native Kelabit speakers – were asked to try out transcribing, and then they were interviewed after they have conducted the transcribing task. This data collection process was conducted over four days' period from 13 to 16 April 2016. A total of six people – 4 females and 2 males, aged in their 40s to 70s – participated in this data collection process. The participants are all native Kelabit speakers. Three of these respondents are volunteers with Bario Radio; one of these three respondents is also the station manager. These respondents all have basic computer skills. Three respondents have worked as teachers and principals/headmaster, two respondents worked in human resources and administration while another worked as a matron at a boarding school hostel. All respondents are currently living in Bario post-retirement from their previous jobs.

The transcribing carried out by the respondents are done via both manual (handwritten) and computer. For the manual way, respondents listened to a short audio file containing recording of a conversation between two Kelabit speakers, and then wrote down the transcription of the conversation on paper. The audio file is played on Windows Media Player. After that, the respondents tried out doing transcription on a computer using software – a word processor, and specific transcribing tools. The first software used is Microsoft Word. Respondents listened to the same audio file on Windows Media Player, then typed out the transcription on Microsoft Word. Then, respondents tried doing the transcription using two types of transcribing tools, namely Transcriber, and Transcription Aid. Both Transcriber and Transcription Aid are free transcribing tools that are available on the Web.

When the respondents have completed the transcription, they were interviewed on their experience using the four methods of transcription. They were asked to rank the methods from easy, moderate, hard to very hard. The interviews were also conducted to get their suggestions and feedbacks to help in building the transcription tool for their community, as a needs analysis to obtain user requirements for the design of the transcribing tool.

All 6 respondents ranked the Transcription Aid as the easiest transcription tool used. Reasons given include the interface is simple and easy to navigate, both the audio control buttons and space to transcribe are on the same interface, the black colour of the interface is pleasing to the eyes. However 2 respondents commented that the button sizes could be made bigger.

There are two tools ranked to be moderate: 4 respondents chose Transcriber while the remaining 2 respondents chose the manual method. Comments on Transcriber include that it is more complex than Transcription Aid and they need more time to understand how to use it. Respondents also mentioned that some parts of the interface are difficult to control (moving the audio, the cursor). As for the signal wave display, some respondents said they do not need the display. The respondents who chose manual method said that they prefer writing over typing, hence the choice of manual method.

Half the respondents ranked using Microsoft Word and Windows Media Player to be hard. The reason given is that it is difficult, inconvenient and can be frustrating to switch back and forth between two interfaces. The act of switching

between interfaces also slows down the transcribing process. Another two respondents said that Transcriber is hard for them (due to complex interface), while one respondent said that manual is hard (inconvenient to move between paper and computer screen to control the Windows Media Player).

Lastly, the tool/method ranked very hard is a tie between the manual method and the Microsoft Word and Windows Media Player. Each tool/method has three respondents ranking it as very hard. Reasons given for manual method being very hard are typing is preferred over writing, the audio can be very fast so need to rewind, this method is better when working with older people (listening to them and then transcribing what they said) and that manual method is good for those who can write fast. As for Microsoft Word and Windows Media Player ranked very hard, respondents said

that they have difficulty in typing, switching between two windows and they felt that having to switch windows slows them down.

As for the feedbacks and suggestions given by the respondents, they suggested having simple and clean interface with bigger buttons. They also requested to have both the audio controls and space for typing transcription in one page, along with the audio slider control. They stated that they do not need the audio's wave signals to be displayed. The respondents also requested for the screen to be in black colour, as they stated that black screen is easier on the eyes for them compared to white screen.

The following figure shows the design of the transcription tool, based on the requirements gathered from the interviews with the respondents.
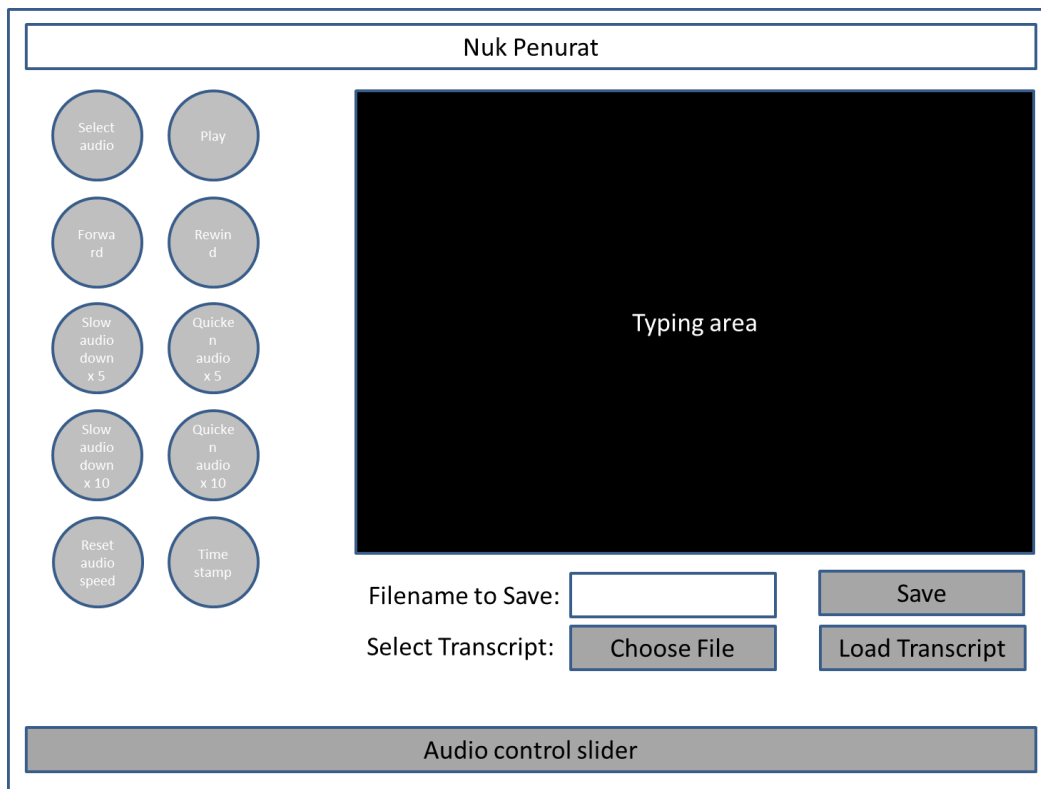

Figure 4: Design of the transcription tool

V. Conclusion and Future Works

This paper describes the efforts to apply PRISMA in collecting user requirements for the design of a transcription tool that can be used to build language resources for the Kelabit community.

References

[1] E-L. Ng, B. Chin, A. W. Yeo, and B. Ranaivo- Malançon, "Identification of Closely-Related Indigenous Languages: An Orthographic Approach", *International Journal on Asian Language Processing*, vol. 20, no 2, pp. 43-61, 2010.

[2] L. Besacier, E. Barnard, A. Karpov, and T. Schultz, "Automatic speech recognition for under-resourced languages: A survey", *Speech Communication*, vol. 56, pp. 85-100, 2013.

[3] Ethnologue, "Kelabit" [online] available: https://www.ethnologue.com/language/kzi, 2016

[4] H. Jaakkola and B. Thalheim, "Architecture-driven modelling technologies," In: *Proceedings of the 2011 conference on Information Modelling and Knowledge Bases XXII*, Anneli Heimburger et al. (eds). IOS Press, p. 98, 2011.

[5] S-T. Siew, A. W. Yeo, and T. Zaman, "Participatory Action Research in Software Development: Indigenous Knowledge Management Systems Case Study", in *Human-Computer Interaction, Part 1, HCII 2013*, LNCS 8004, pp. 470-479, 2013.

[6] M. C. Samani, P. Bala, J. Maliki, and R. Harris, "Bario Community Radio: Engaging People Making News," *Proceedings of the 8th International Conference on Information Technology in Asia (CITA)*, Kuching, Sarawak, pp. 1-6, 2013.