Adopting Malay Syllable Structure for Syllable Based Speech Synthesizer for Iban and Bidayuh languages

Sarah F. S. Juan[1], Vyonne Edwin[1], Cheong Chai Yeen[1], Lee Jun Choi[1], Alvin W. Yeo[2]

[1]Faculty of Computer Science and Information Technology
[2]Institute of Social Informatics and Technological Innovations
Universiti Malaysia Sarawak (UNIMAS)
Kota Samarahan, Sarawak, Malaysia.
sjsflora@fit.unimas.my, vyonneedwin@gmail.com, chaiyeen@gmail.com, jclee@fit.unimas.my, alvin@isiti.unimas.my

Abstract— **Sarawak, Malaysia, has many under-resourced languages, which stands to become extinct if measures are not taken to preserve and maintain them. These languages are mostly spoken by the indigenous groups and not all of the languages are documented or studied. As an initiative to preserve, a Text to Speech (TTS) system has been built for Iban and Bidayuh languages, two out of 44 living languages in Sarawak. To expedite the development, we employed knowledge of closely-related language, i.e. Malay, which is the first language in Malaysia. In this paper, we employed a syllabification algorithm based on Malay syllable structure to build the Iban and Bidayuh syllable list and speech corpus. An accuracy test for the algorithm was conducted to determine the quality of the output from the TTS system using Categorical Estimation (CE). Test showed high percentage in accuracy and quality has a mean score of 3.07 out of 5, suggesting the approach works.**

Keywords-text to speech system; syllabification algorithm;

I. INTRODUCTION

A Text-to-Speech (TTS) system is an application to convert written text in a specific language, into artificial, machine-generated speech. User will input text into the system, then, the system will process or analyze the text in order generate speech. The ability of a machine or a computer to read any text aloud would enable the device to be used in many situations and purposes.

A text-to-speech system for Sarawak languages is proposed. This paper focuses two major languages in Sarawak, namely Iban and Bidayuh languages and the paper proceeds as follows. Section II details on the purpose of this study while section III covers current Malay TTS systems. In section IV, we present Iban and Bidayuh syllable structure and Section V and VI explains the syllabification and development of the algorithm, respectively. Section VII discusses briefly on the framework of our TTS system, followed by results of the experiments to evaluate the algorithm and the system in Section VIII and a conclusion in Section IX.

II. MOTIVATION

According to Lewis in [8], there are 46 languages in Sarawak, of which, 44 living languages and two languages have no known speakers. These languages are mostly spoken by the indigenous groups of Sarawak especially those living in the rural areas. Not many of these languages are written and some of these languages have dialects that are not mutually intelligible, as found in the Bidayuh language [14]. Inter-racial marriages results some families use common languages such as Malay or English to communicate. Situations such as these could possibly lead to extinction or purity of indigenous languages due to code switching in conversations. As such, indigenous languages need to be preserved to ensure they do not disappear.

In addition to language preservation, the disabled communities will also benefit. The visually impaired or the blind are able to use text to speech systems as an assistive tool to read text and education especially in helping children who suffers dyslexia in pronouncing words correctly [18]. However, current systems accommodate only dominant languages such as English, French and Japanese.

Thus, it is timely to develop a speech synthesizer that is able to read any of the Sarawak languages text aloud. Besides that, the system could also be one of the potential tools in bridging the digital divide in the communities of Sarawak.

III. EXISTING TEXT TO SPEECH SYSTEMS USING SYLLABIFICATION APPROACH FOR STANDARD MALAY

In preparing for our TTS, Malay speech syntheses using syllabification approach is studied. Say It! system uses segmentation technique to select the longest phoneme sequence and compare the selected sequence to the existing syllable database [21]. If a match is found, the sequence will be considered as a syllable unit and extracted from the database. Otherwise, the comparison process will be carried out with the reduction of the last phoneme list until a match is found. It is a simple implementation, which produces a quick result. However, authors found that parsing could be wrongly segmented.

Another attempt for Malay language was carried out by [3], by adapting a previously developed Arabic synthesizer. In this implementation, they considered linguistic analysis and phonetic analysis of Standard Malay sound system together with loan words. Moreover, they compared syllable structure and showed three syllable structures which mainly based on consonants (C) and vowels (V) sequence which are CV, CVC and CV and performed segmentation of syllables based on these rules. The authors claimed that this technique proved its ability to produce an acceptable result.

Aside from attempts mentioned, Universiti Sains Malaysia developed a syllable based speech synthesizer for Malay by adapting modules to existing synthesizer engines, which catered for other languages (American English and Indonesian) [16]. Words are arranged according to the syllable sequence based on Malay phonological rules. They presented four syllable structures which are CV, VC, CVC and V and found out that this

IEEE
computer
society