Text-to-Speech in Learning Thai Language for Foreigners

*Faculty of Information Technology, King Mongkut’s Institute of Technology Ladkrabang, Bangkok, 10520, Thailand
**Faculty of Information Technology, Dhurakijpundit University, Bangkok, 10210, Thailand
***Faculty of Engineering, Khon Kaen University, Khon Kaen, Thailand
****School of Science and Engineering, Waseda University, Tokyo, Japan

Abstract: Text-to-Speech synthesis can be coupled with computer aided learning system and provide a helpful tool to learn a new language. The conversation dialogs are one of the important materials for practical used of each language. In the past, the learner could learn the correct pronunciation from tapes that come with the books. However, the learners could learn to pronounce only in the limited words in the book. When they find new words that do not have in the book, they must guess what is the correct one. In case of Thai language, the pronunciation of each word becomes more difficult because the Thai language is a tone language. This means that pitches are meaningful. A word pronounced with different pitches carry different meanings. This paper presents the work in progress on Thai text-to-speech for helping foreigners in learning Thai language. The aim is to build a system that could generate sound from the input Thai sentences.

Introduction

It is hard to learn a new language without adequate support. Now a day, computer becomes cheaper and more powerful. Therefore, it is used in many applications including language teaching. Text-to-Speech synthesis can be coupled with computer aided learning system and provide a helpful tool to learn a new language. The conversation dialogs are one of the important materials for practical used of each language. In the past, the learner could learn the correct pronunciation from tapes that come with the books. However, the learners could learn to pronounce only in the limited words in the book. When they find new words that do not have in the book, they must guess what is the correct one.

In case of Thai language, the pronunciation of each word becomes more difficult because the Thai language is a tone language. This means that pitches are meaningful. A word pronounced with different pitches carry different meanings. There is no spacing between words and no special mark to identify the end of a sentence. The Thai vowel forms do not follow initial consonants; some are placed before the initial consonants, some after the consonants, some above the consonants, and some underneath the consonants. The vowels that are “complex” forms (i.e. composed of more than one part) can be placed around the consonants.

There are 5 distinctive tones (pitches) in Standard Thai. They are

1. mid level tone (here represented with the number 1) for example: khaa1 (to be lodged in)
2. low level tone (represented with the number 2) for example: khaa2 (Galanga, an aromatic root)
3. falling tone (represented with the number 3) for example: khaa3 (I, slave, servant)
4. high level tone (represented with the number 4) for example: khaa4 (to sell)
5. rising tone (represented with the number 5) for example: khaa5 (leg) .
From the example mentioned above, it can be seen that it is very hard to the foreigners to pronounce a Thai word correctly.

**The proposed system**

This paper presents the work in progress on Thai text-to-speech for helping foreigners in learning Thai language. The aim is to build a system that could generate sound from the input Thai sentences. The synthesis system receives input Thai text and separates it into token according to the grammar of Thai language (Pantumetha, 1998, Dutoit, 1997). The sound of each word is generated.

**Experimental Results**

At this point of time, the system can generate sound from the variety of user input and shows how to produce sound for each word as shown in Fig. 1. However, some words are still not correctly produced and need to improve in the future.

Figure 1. The proposed Thai text to speech system.

**References**
