VOIS for Hearing-Impaired Children in Myanmar

Aye Thida
Professor
Head of Artificial Intelligent Research Lab
University of Computer Studies, Mandalay
ayethida@ucsm.edu.mm

Project Leader

- Dr. Aye Thida

Project Members

- Nway Nway Han
- Sheinn Thawtar Oo

Background

- Children's language acquisition starts from the early years.
- The first few years, usually before age 5, are called the critical period.
- After that period language acquisition is more difficult and less successful.
- The language acquired early in a child's life is the first language.
- Exposure to a natural language in the critical period is important to the development of cognitive activities, such as:
  - ✓Literacy
  - ✓Memory organization
  - ✓Number manipulation

Eight main ethnic groups in Myanmar

- 135 ethnic groups in Myanmar
  - Kachin
  - Kayah
  - Karen
  - Chin
  - Burma
  - Mon
  - Rakhine
  - Shan

How to Pronounce “Chair” and “Good morning”

Main ethnic group	Good morning	Chair
Kachin	Kaj ai i/ka gya ai i/	La hkum/ la kwan/
Kayah	တယ်ရဘယ်ဆီ/te ra be si/	နွိုး/now/
Karen	Gaw mu mu/gaw mu mu/	Law say naw/law se naw/
Chin	Zing Pha/zin pa/	Tut Phah/tut pa/
Burma	မင်္ဂလာပါ/mingalar par/	ခုံ /khoun/
Mon	မငေးရအောင်/ma nge ra aung/	ခိန်းဟဂျော့/kein ha gyaw/
Rakhine	သာလီစွပါ / tha li swa pa/	ခုံ /khoun/
Shan	မိုဲ့သုင်ခၢး /mai son ka/	တင့်ဆင်း /tan nan/

Problems and Challenges

- Hearing aids and Cochlear Implants (CI) may help the child hear but they do not restore normal hearing.
- Oral educational support is needed to learn a language.
- Speech perception always precedes speech production.
- Oral instructions and speech therapy is needed to develop a speech understandable to common people.
- Speech therapy is somewhat dull and repetitive for a child.

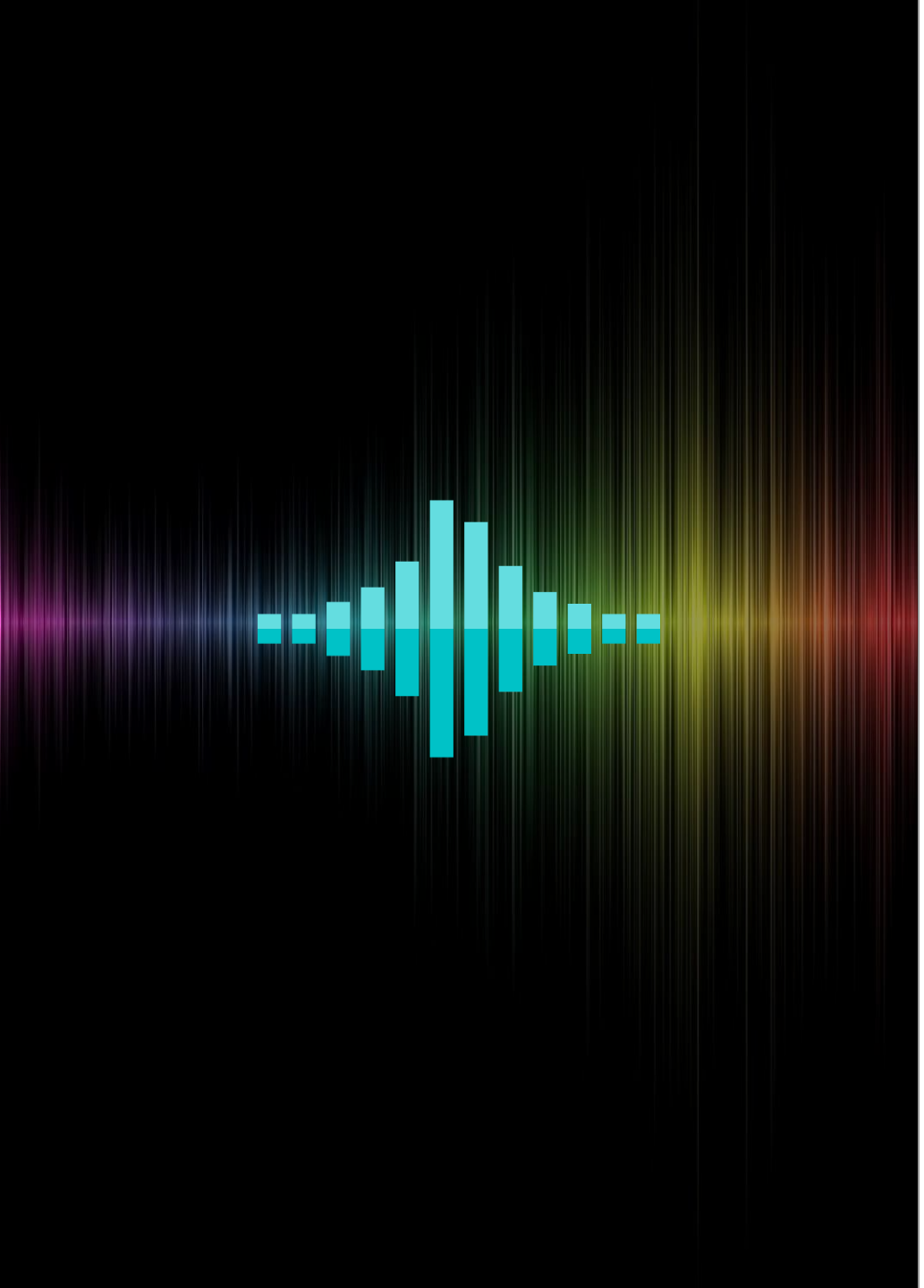
via [mandalaydeafchild.org]

Targets:

- In Myanmar, 80% of mobile phone users use smartphones.
- We develop an application for hearing-impaired children with hearing aids to practice their speech.
- If the app can recognize their words in the respective language, it can help the children to pronounce words correctly.
- Instead of traditional flash card method, we can use interactive animations to engage the children's attention for long.
- Rewarding with unique animations when a word is pronounced correctly will encourage the children to learn the correct pronunciation.

What is VOIS?

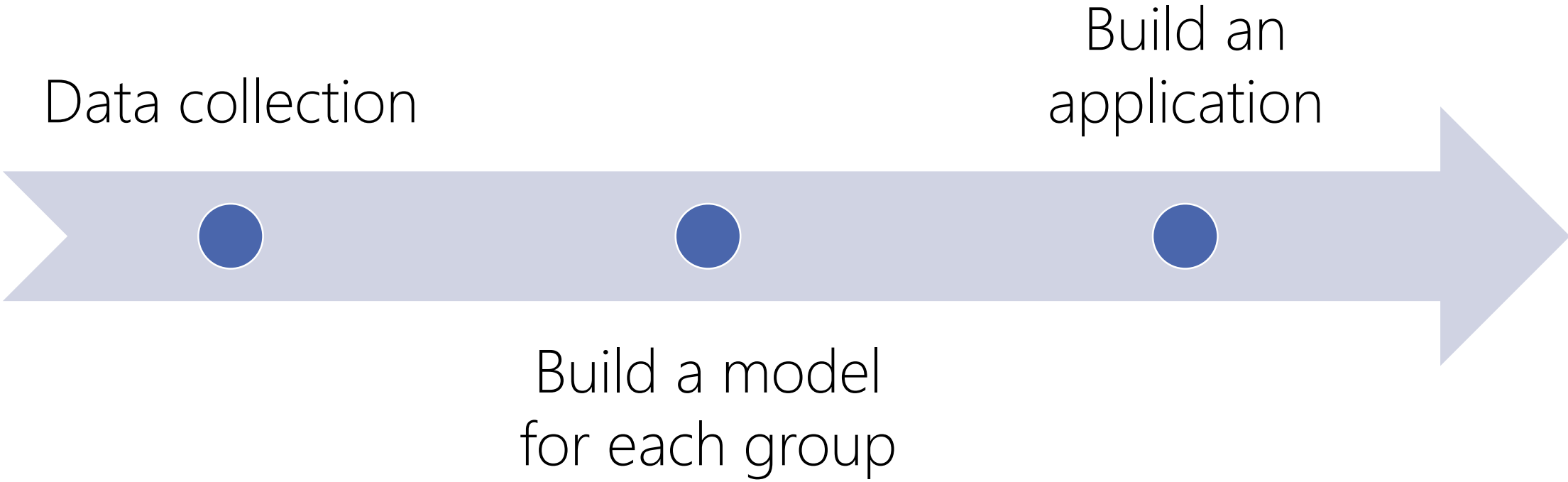
- Voice data are collected from children with normal and acceptable speech.
- The data are used to train a deep learning model to recognize each word in its correct pronunciation.
- The trained models are built into a mobile application.
- We decided to build the speech recognition engine into the app to be able to recognize the speech offline.
- The application is designed to be fun and to provide interactions for the children.



OBJECTIVES OF VOIS

- To use smartphones as an **additional tool** for speech therapy.
- To make speech practice **interactive** and **fun**.
- To **maximize the exposure of a natural language** to a hearing impaired child and constantly **correct their pronunciation**.
- To **encourage** hearing impaired children with hearing aids to speak to **develop** their **speech production** skills.

Proposed System Overview



Data collection

- The words are divided into 10 categories.
  - food, toy, animals, environment, body, places, objects, vehicles, flowers and house
- The voice data are collected with one syllable /ကား /ka:/car/, two-syllable /ရေခွက် /jei gwe'/glass/ and three syllables /သွားတိုက်တံ/ thwar: tike tan/ toothbrush/ from the children.
- The voice data are collected from children with typical and acceptable speech.
- The age of the children are between 4 and 10 years.

Build a model for each group

- The hearing-impaired children's education is challenging because they are unlikely to develop normal speech and language ability.
- The speech data are trained by Convolutional Neural Network (CNN) to recognize each word in its correct pronunciation.
- The trained models are built into a mobile application.
- The speech recognition is implemented into the application to be able to recognize speech offline for eight ethnic groups.
- When connected to the Internet, Neural Network models can be updated to recognize new words.

Cont'd:

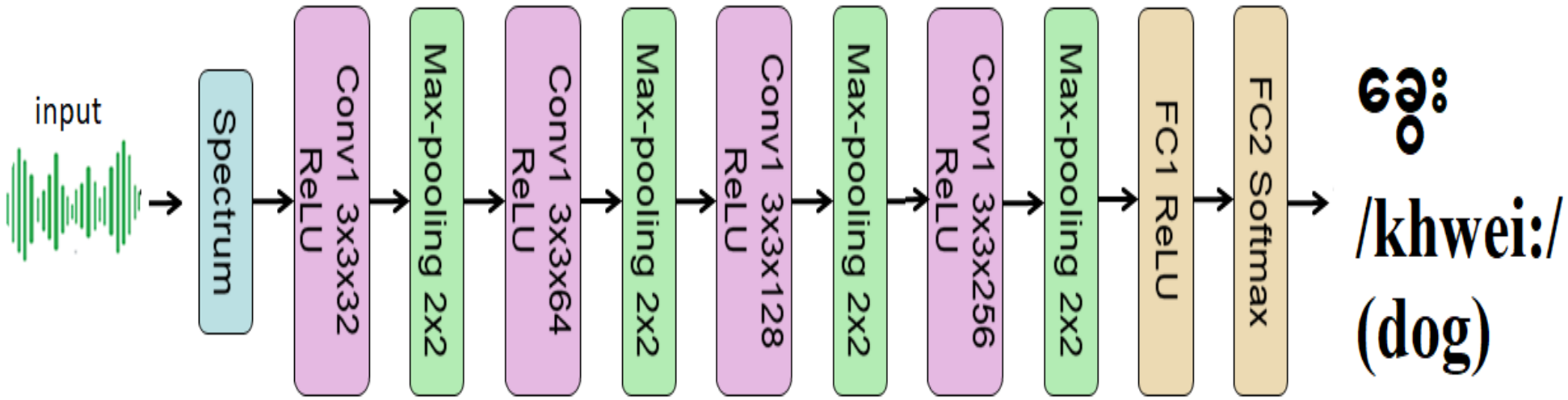


Fig. 1: The speech recognition engine of VOIS.

VOIS Application

Fig 2. The user interface of VOIS

- The children can press the mic button to record their voice, for example, ကား /ka:/ (car), pronouncing the word for the picture.
- They can also press the picture to hear the correct sound. They can practice the word until they can speak the correct pronunciation.
- If their pronunciation is correct, the ကား /ka:/ (car) will animate and play some car sound.
- The children can use a sticker board in this app. The sticker board keeps therapy fun, while the activity report shows the child's progress.
- If they get the pronunciation wrong, the app will show the children their pronunciation is wrong and do not play fun animations.

Impact

- The proposed project develops a **speech therapy app** for children with hearing defect, which can also be used as Speech Language Pathology (SLP) app in the general context.

- It can help **hearing-impaired children** to train with the language pre-requisites at their own space.
- The proposed project is an important step for **language acquisition early in childhood and** it can be extended for other languages.

Cont'd:

- Our proposed Project is in need of support from the ASEAN-IVO funding.
- We would like to bring together researchers in the field of Natural Language processing and Speech Synthesis.
- We would like to develop and encourage more research works with developed country, Japan and other developed and developing countries of ASEAN.

Output/Outcome

- The hearing-impaired children can learn and operate the language freely through a simple practice using this application.

- The expectation is that this application can bring both opportunities and life-quality improvements for children with hearing loss in Myanmar.

Conclusion

- VOIS is the prototype of speech therapy of hearing-impaired children.
- The proposed system also encourages further researches in Natural Language Processing (NLP) and speech synthesis.
- We will develop an app and need to do with high accuracy and the performance of CNN in terms of validation.
- We will implement this proposed system accessible to all hearing-impaired children in Myanmar.