



23rd November 2015



# VoiceTra Speech to Speech Translation App for Myanmar Language

*Press Conference, UCSY, Yangon, Myanmar*

Ye Kyaw Thu,

Researcher,

Multilingual Translation Lab.,

Advanced Speech Translation Research and Development Promotion Centre,

NICT, Kyoto, Japan

Win Pa Pa,

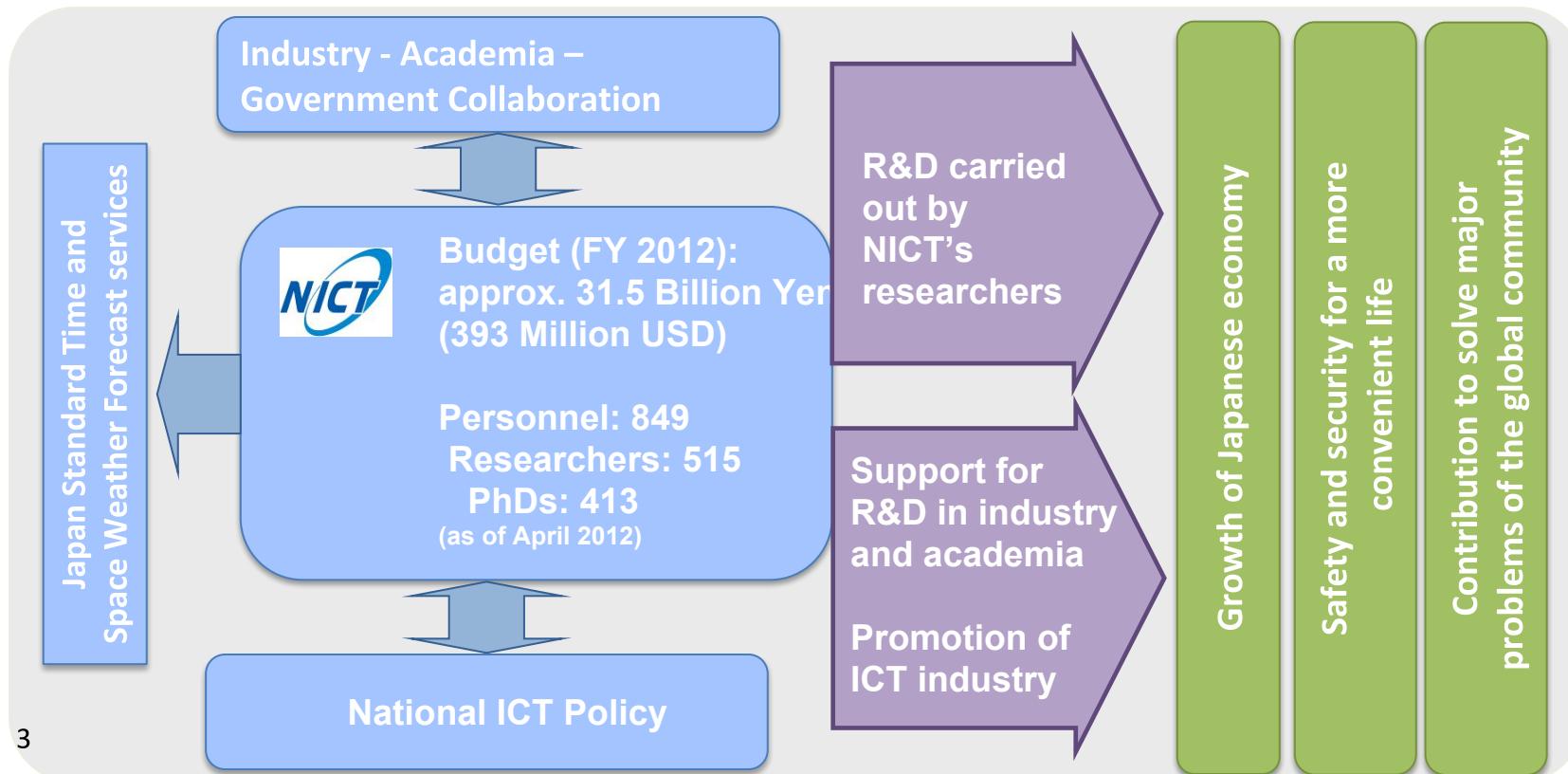
Associate Professor,

# Table of Contents

- Introduction to NICT
- Introduction to UCSY
- ASR Development for Myanmar
- Statistical Machine Translation for Myanmar Language
- TTS Development for Myanmar
- User Interface of VoiceTra App
- Demo of VoiceTra App
- Q&A

# NICT's Role

As the sole national research institute in the field of information and communications technologies (ICT), NICT strives to advance technologies as well as contributes to national ICT policies, by promoting its own research and development and by cooperating with and supporting industry and academia.



# Location of NICT Facilities

## Koganei City, TOKYO

Headquarters  
 Photonic Network Res. Institute  
 Network Security Res. Institute  
 Applied Electromag. Res. Institute  
 Advanced ICT Res. Institute  
 (Quantum, Terahertz)  
 Terahertz Technology Res. Center  
 Cybersecurity Research Center

## Noumi City, ISHIKAWA

Hokuriku StarBED Techn. Center

## Keihanna Region, KYOTO

Universal Comm. Res. Institute

## Kobe City, HYOGO

Advanced ICT Res. Institute  
 (Bio, Nano)

## Saga City, SAGA / Itoshima-City, FUKUOKA

Hagane-yama LF Standard Time  
 and Frequency Transmission  
 Station

## Onna Village, OKINAWA

Okinawa Electromag. Techn. Center

## Sendai City, MIYAGI

Resilient ICT Research Center

## Tamura City,

Ohtakadoya-yama LF  
 Standard Time and Frequency  
 Transmission Station

## Kashima City, IBARAKI

Kashima Space Techn. Center

## Chiyoda Ward, TOKYO

Network Testbed R&D Promotion  
 Center

## Yokosuka City, KANAGAWA

Wireless Network Res. Institute

## Suita City, OSAKA

Center for Information and Neural  
 Networks

- Headquarters
- Research Institute
- Promotion Center / Research Center
- Technology Center
- Standard Time and Frequency Transmission Station

## Overseas Centers

North-America Center  
 Europe Center  
 Asia Center

# For this VoiceTra Project

- Spoken Language Communication Lab.,

<http://www2.nict.go.jp/univ-com/slc/en/index.html>

- Multilingual Translation Lab.,

[http://www2.nict.go.jp/univ-com/multi\\_trans/index-e.html](http://www2.nict.go.jp/univ-com/multi_trans/index-e.html)



# University of Computer Studies, Yangon (UCSY)



- Founded as Universities Computer Center (1971)
- Changed to UCSY in 1998
- Home page: [www.ucsy.edu.mm](http://www.ucsy.edu.mm)

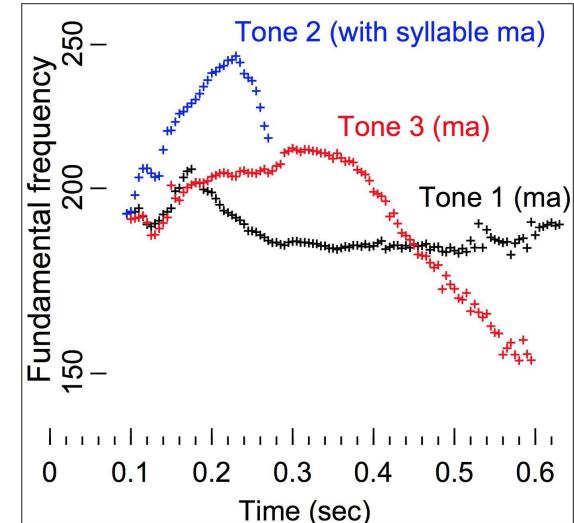
# Natural Language Processing Lab

- 2 Researchers and 1 Student visited NICT for 9 months and a year (July 2014 – June 2015) for developing Myanmar Automatic Speech Recognition (ASR) and Text to Speech (TTS)
- Home page: [www.nlpresearch-ucsy.edu.mm](http://www.nlpresearch-ucsy.edu.mm)



# Myanmar Tone

- Myanmar is a tonal language
- Tone is carried by syllable and is featured by both fundamental frequency and duration of syllable



Tone Name	Vowel	Symbol	Example
Low (Tone 1)	- ၁	a	မှ /m a/ 
High (Tone 3)	- ၃	a:	မှုံ /m a:/ 
Creaky (Tone 2)	-	a.	ဥ /m a./ 

# Text Preprocessing

- Syllable segmentation
  - Rule based syllable segmentation  
(Ye et al., ICCA2013)
  - Segmented based on pronunciation  
မြန်စကား  မ\_ြ\_န\_စ\_ကား
- Word Segmentation
  - Used Conditional Random Fields (CRF)  
(Win et al., ICGEC2015)  
မ\_ြ\_န\_စ\_ကား  မြန်\_စကား

# Text Normalization

ဖေဖော်ဝါရီ ၁၃ ဟာ ပိုလ်ချုပ်မွှေးနှေပါ  
ဖေဖော်ဝါရီ ဆယ့်သုံး ဟာ ပိုလ်ချုပ်မွှေးနှေပါ

ဖုန်းနံပါတ်က ၁၂၃၄၅၆ ဖြစ်ပါတယ်  
ဖုန်းနံပါတ်က ထစ်နှစ်သုံးလေးငါး ဖြစ်ပါတယ်

၈:၃၀ မှာ လေယာဉ်ထွက်ပါမယ်  
ရှစ်နာရီ သုံးဆယ်မြို့နှစ်မှာ လေယာဉ်ထွက်ပါမယ်

# Grapheme to Phoneme (G2P) conversion

- Used phrase based statistical machine translation (PBSMT) (Ye et al.,  
PAC ၁၀၀။ (sa. ka:) => za- ga:  
ပျဉား (pa. nja) => pjin nja  
စာရေး (sa jei:) => sa- jei:
- Phoneme correctness accuracy = 94.29 %

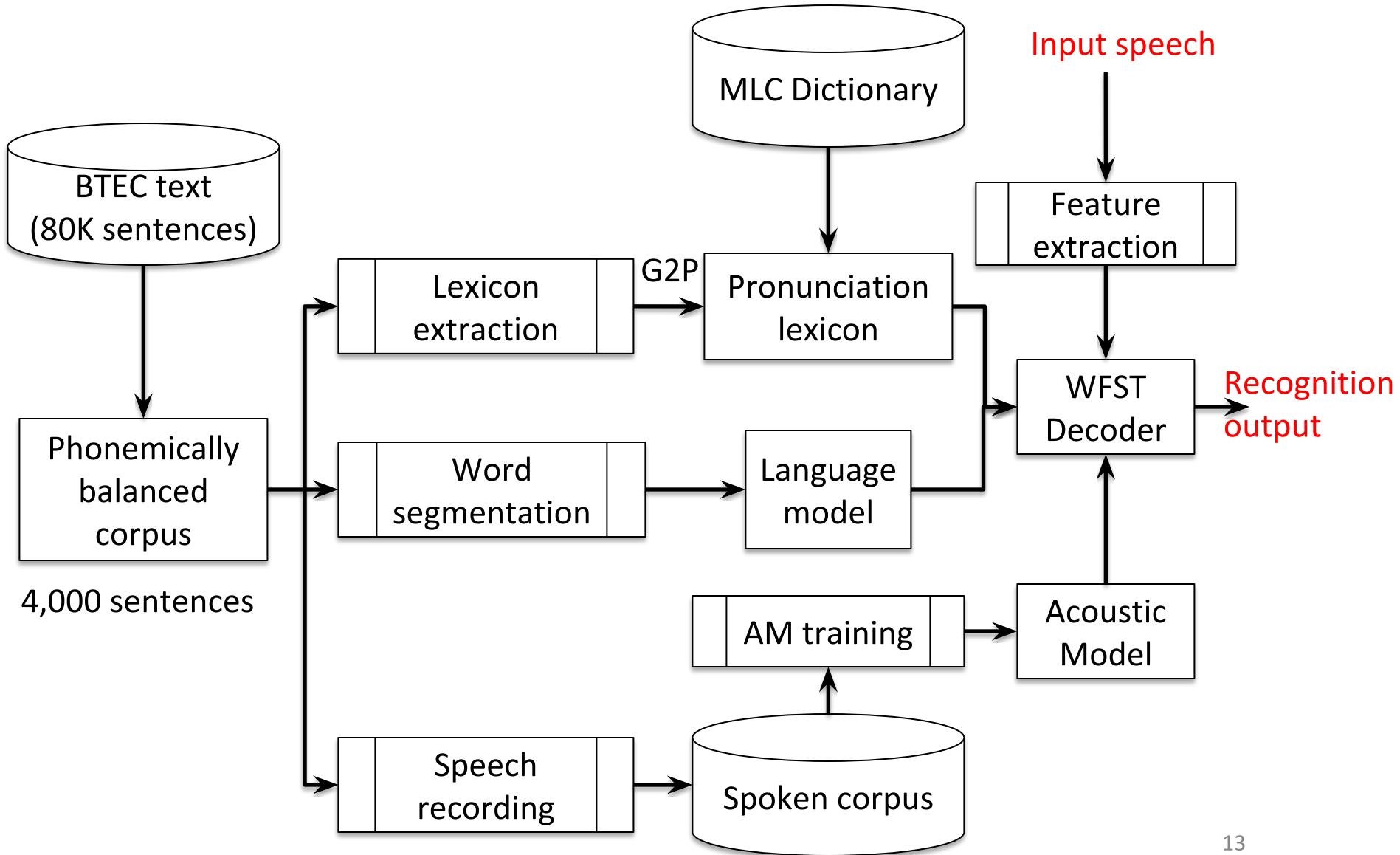
# Speech Corpus Building

- Speech corpus building  
(Read speech)
  - 3.59 hrs of female
  - 3.35 hrs of male
- Used Marantz recording device
- Recording with maximum decibel (dB) is -12
- Training with HMM based Ossian Speech synthesizing tool

<https://www.inf.ed.ac.uk/research/isdd/admin/package?view=1&id=184>



# Developing Myanmar ASR



# Corpus Preparation

こんにちは

မင်္ဂလာပါ

どうもありがとうございます

ကျေးဇူးတင်ပါတယ်။

バスで行けます

ဘက်စ်နဲ့သွားနိုင်တယ်

30分くらいです

မိန်စ်သုံးဆယ်လောက်ပါ

- Parallel text for Japanese and Myanmar
- For Japanese to Myanmar SMT
- For Myanmar to Japanese SMT

# Word Segmentation

prazer em conhecê-lo .

es freut mich , sie zu treffen .

pleased to meet you .

encantado de conocerle .

enchanté .

आपसे मिलकर खुशि हूँइ .

saya senang bertemu dengan anda .

piacere di conoscerla .

初めてまして。

만나 서 반갑 습니다 .

уулзахад таатай байна .

saya senang bertemu dengan anda .

ထွေ ရတာ ဝမ်းသာ ပါတယ် ။

तपाईंलाई भेटेर खुसी लाग्यो ।

очень рад с вами познакомиться .

କର୍ତ୍ତାଙ୍କର ଅଳ ହମ୍ମିଲେନ୍ଦନାର ।

- How to break  
for Myanmar?!

# Alignment

- Example alignment for ja-my, my-ja combination of GIZA++

	1	2	3	4	5	6	7
	私	は	文子	と	いい	ます	。
1	私						
2	は						
3	文	子					
4	と						
5	い	い	い				
6	す	す	す				
7	。				ま	す	
8	。				す	す	

# Phrase Table

```
" congratulation " from ||| ରାଖ୍ଯପ୍ରି ||| 0.0073484 2.69058e-11 0.073484 1 ||| 1-0 ||| 30 3 1 |||
" congratulation " from ||| ରାଖ୍ଯପ୍ରି ପିତାଯ ||| 0.00165753 2.69058e-11 0.073484 0.0080614 ||| 1-0 ||| 133 3 1 |||
" congratulation " from ||| ରାଖ୍ଯପ୍ରି ପିତାଯ ଲ୍ଲି ||| 0.036742 2.69058e-11 0.073484 9.19274e-05 ||| 1-0 ||| 6 3 1 |||
" congratulation " ||| ରାଖ୍ଯପ୍ରି ||| 0.0073484 1.66363e-08 0.073484 1 ||| 1-0 ||| 30 3 1 |||
" congratulation " ||| ରାଖ୍ଯପ୍ରି ପିତାଯ ||| 0.00165753 1.66363e-08 0.073484 0.0080614 ||| 1-0 ||| 133 3 1 |||
" congratulation " ||| ରାଖ୍ଯପ୍ରି ପିତାଯ ଲ୍ଲି ||| 0.036742 1.66363e-08 0.073484 9.19274e-05 ||| 1-0 ||| 6 3 1 |||
" congratulation ||| ରାଖ୍ଯପ୍ରି ||| 0.0073484 7.53522e-06 0.073484 1 ||| 1-0 ||| 30 3 1 |||
" congratulation ||| ରାଖ୍ଯପ୍ରି ପିତାଯ ||| 0.00165753 7.53522e-06 0.073484 0.0080614 ||| 1-0 ||| 133 3 1 |||
" congratulation ||| ରାଖ୍ଯପ୍ରି ପିତାଯ ଲ୍ଲି ||| 0.036742 7.53522e-06 0.073484 9.19274e-05 ||| 1-0 ||| 6 3 1 |||
```

Fig. Inside phrase-table.1.gz

- A big dictionary for Machine Translation
- You can also combine more than one phrase tables (several techniques)

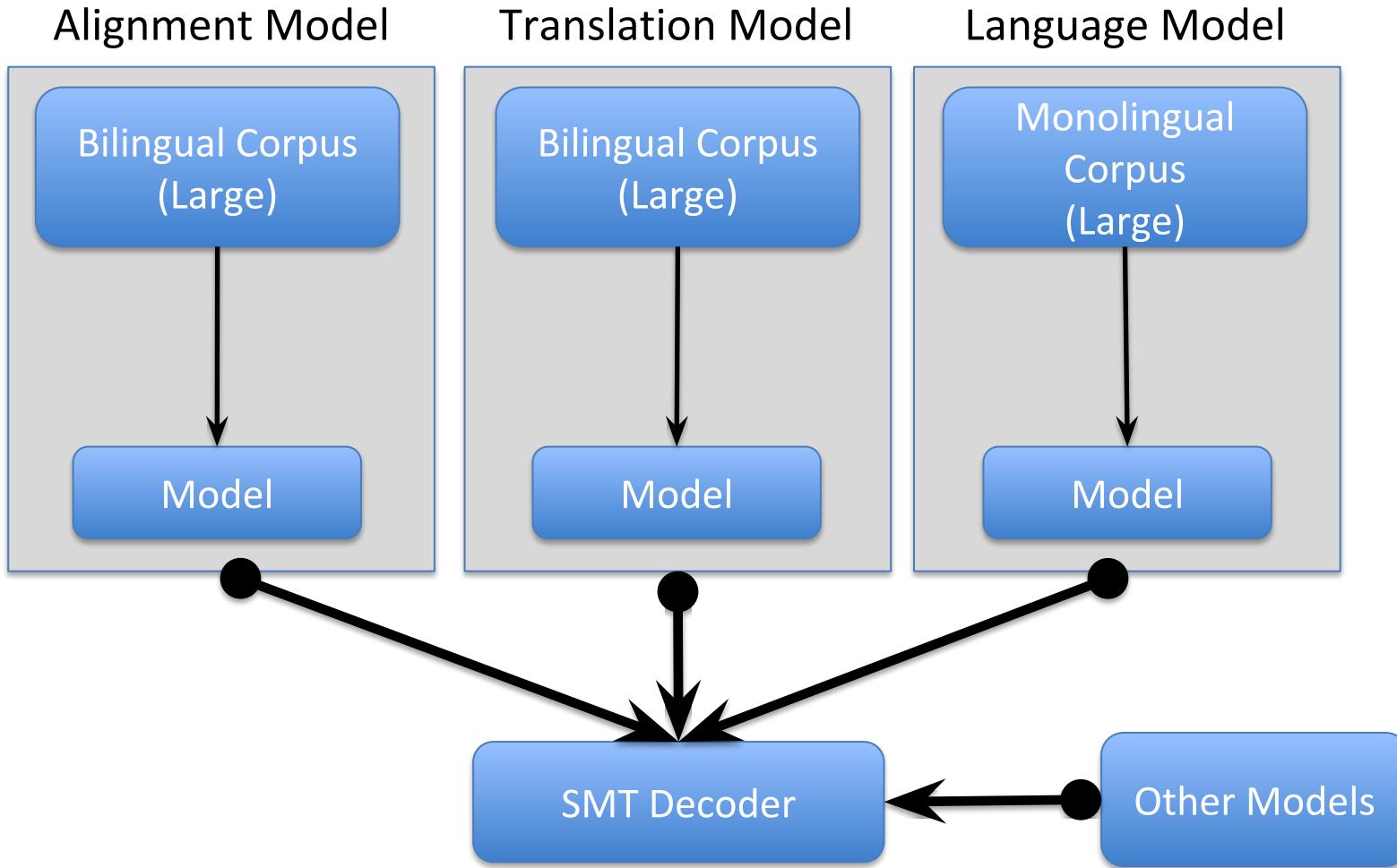
# Language Model

$$P(w_1 w_2 \dots w_n) = \prod_i P(w_i | w_1 w_2 \dots w_{i-1})$$

- Joint probability of words in a sentence

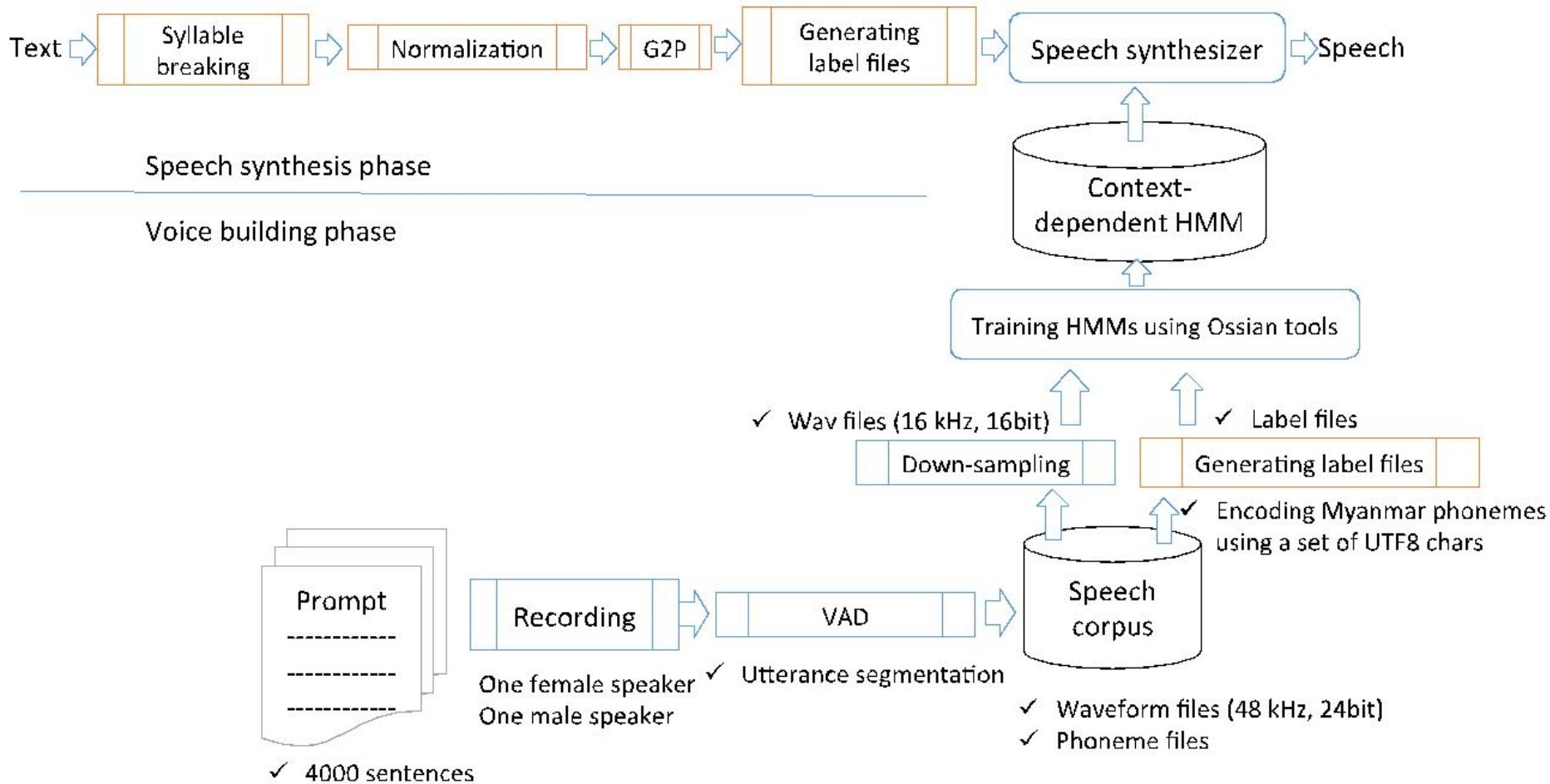
$$\begin{aligned} P(\text{မင်းကဘယ်သူလဲ။}) &= P(\text{မင်း}) \times P(\text{က} | \text{မင်း}) \times P(\text{ဘယ်သူ} | \text{မင်း က}) \\ &\quad \times P(\text{လဲ} | \text{မင်း က ဘယ်သူ}) \times P(\text{။} | \text{မင်း က ဘယ်သူ လဲ}) \end{aligned}$$

# Overall SMT



For VoiceTra App, we used inhouse SMT System

# TTS Development for Myanmar



# Some Achievements

- Best Paper Award for G2P conversion paper at ICCA 2015 Conference, Myanmar



# Some Achievements



- Tokyo Broadcasting System (TBS) TV Program:  
Yume no Tobi+ (in English: Next Door)

# Some Achievements

- Hay Mar Soe Naing, Ye Kyaw Thu, Win Pa Pa, Hiroaki Kato, Andrew Finch, Eiichiro Sumita, Ciori Hori. “Rule Based Katakana to Myanmar Transliteration for Post-editing Machine Translation” 言語処理学会第21回年次大会 発表論文集, March 16~21, 2015, pp. 257-260
- Ye Kyaw Thu, Win Pa Pa, Andrew Finch, Aye Mya Hlaing, Hay Mar Soe Naing, Eiichiro Sumita and Chiori Hori, Syllable Pronunciation Features for Myanmar Grapheme to Phoneme Conversion. In Proceedings of the 13th International Conference on Computer Applications (ICCA 2015), pp. 161-167, Yangon, Myanmar, Feb  
[Best Paper Award]
- Ye Kyaw Thu, Win Pa Pa, Andrew Finch, Jinfu Ni, Eiichiro Sumita and Chiori Hori, 2015, The Application of Phrase Based Statistical Machine Translation Techniques to Myanmar Grapheme to Phoneme Conversion, In Proceedings of the Pacific Association for Computational Linguistics Conference (PACLING 2015), pp. 170-176, Legian, Bali, Indonesia, May

# Some Achievements

- Win Pa Pa, Ye Kyaw Thu, Andrew Finch, Eiichiro Sumita. Word Boundary Identification for Myanmar Text Using Conditional Random Fields. In Proceedings of the Ninth International Conference on Genetic and Evolutionary Computing, pages 447-456, Yangon, Myanmar, August 26-28, 2015
- Ye Kyaw Thu, Win Pa Pa, Jinfu Ni, Yoshinori Shiga, Andrew Finch, Chiori Hori, Hisashi Kawai, Eiichiro Sumita. HMM Based Myanmar Text to Speech System. In Proceedings of the 16th Annual Conference of the International Speech Communication Association (INTERSPEECH 2015), pages 2237-2241, Dresden, Germany, September 6-10, 2015

# User Interface of VoiceTra App

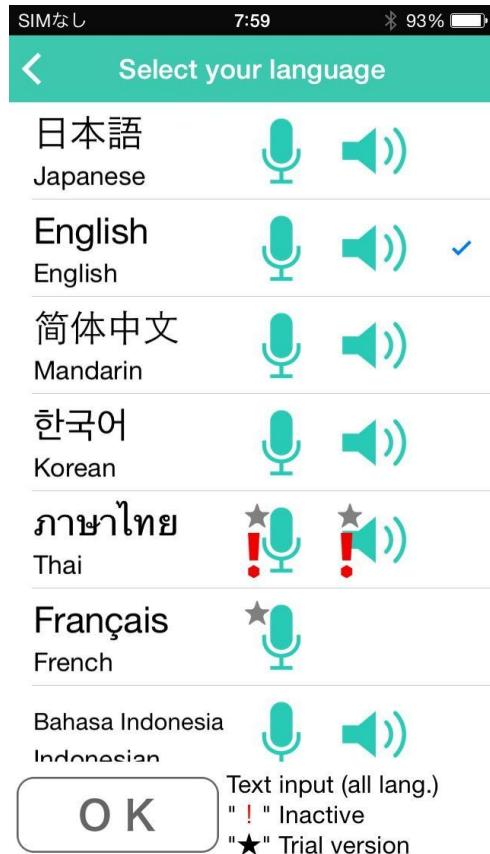


Fig. UI of Language Selection

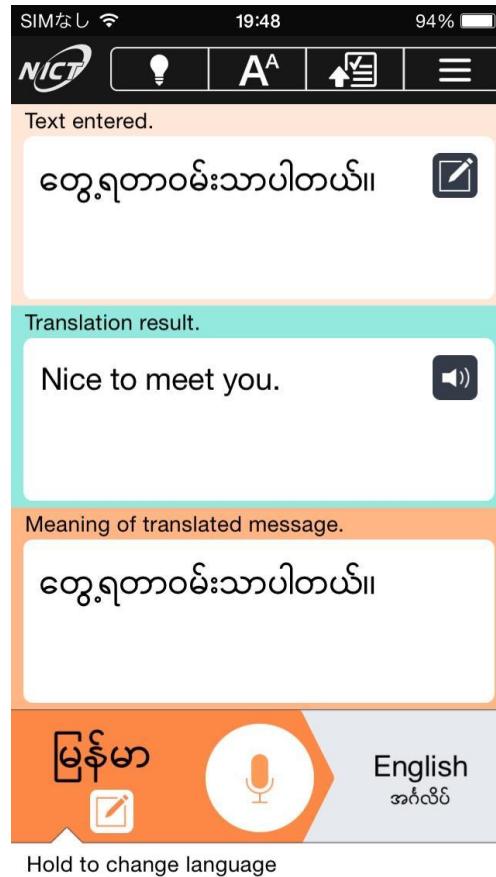


Fig. Main UI of VoiceTra App

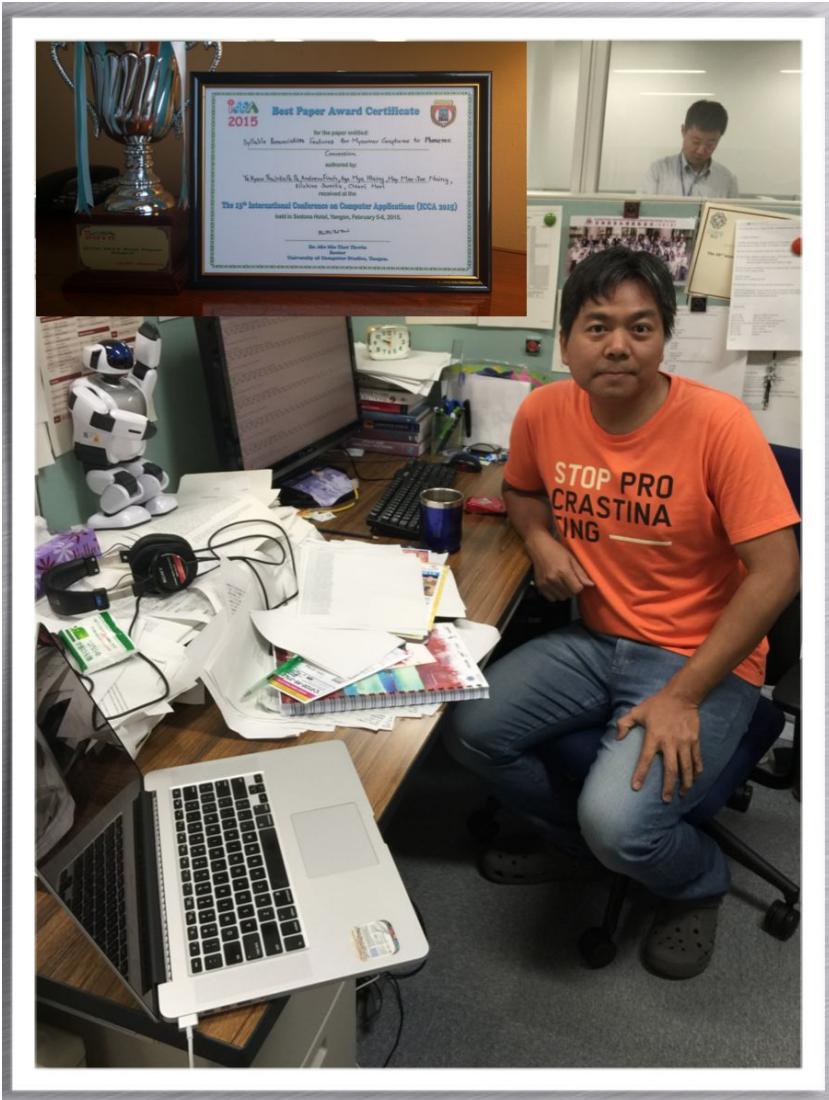
# Demonstration of VoiceTra App

# Thank You!

# Introduction to UCSY

- Introduction to NICT
- Introduction to UCSY
- ASR Development for Myanmar
- Statistical Machine Translation for Myanmar Language
- TTS Development for Myanmar
- UI of VoiceTra App
- Demo of VoiceTra App
- Q&A

# >whoami



- A native of Myanmar
- Doctor's of Global Information and Telecommunication Studies from Waseda University
- Researcher at NICT
- Invited researcher at Waseda University

# Word Segmentation

ရာ	သီ	ပု	တူ	တော်	တော်	ကောင်း	တယ်
<	-	>		+			
ရာ	သီ	ပု	တူ	တော်	တော်	ကောင်း	တယ်
<	-	>		<			
ရာ	သီ	ပု	တူ	တော်	တော်	ကောင်း	တယ်
<	-	-		<			
ရာ	သီ	ပု	တူ	တော်	တော်	ကောင်း	တယ်
-	-	-		-			

Fig. Syllable tagging with different tag sets