

Technologies and AI/ML Applications for Persons with Disabilities

Chatchawarn Hansakunbuntheung, PhD.

Assistive Technology and Medical Devices Research Center (A-MED)
National Science and Technology Development Agency (NSTDA)



My Background

1998	B.Eng. Electrical Engineering, Chulalongkorn University
2000	M.Eng. Electrical Engineering, Chulalongkorn University
2001	D.Sc. Global Information and Telecommunication Studies, GITI, Waseda University
2001 Japan	Visiting Researcher, Electrotechnical Laboratory (ETL), Tsukuba Science Center,
2000 – 2018	Researcher National Electronics and Computer Technology Center (NECTEC) National Science and Technology Development Agency (NSTDA)
2018 – Present	Researcher Assistive Technology and Medical Devices Research Center (A-MED) National Science and Technology Development Agency (NSTDA)

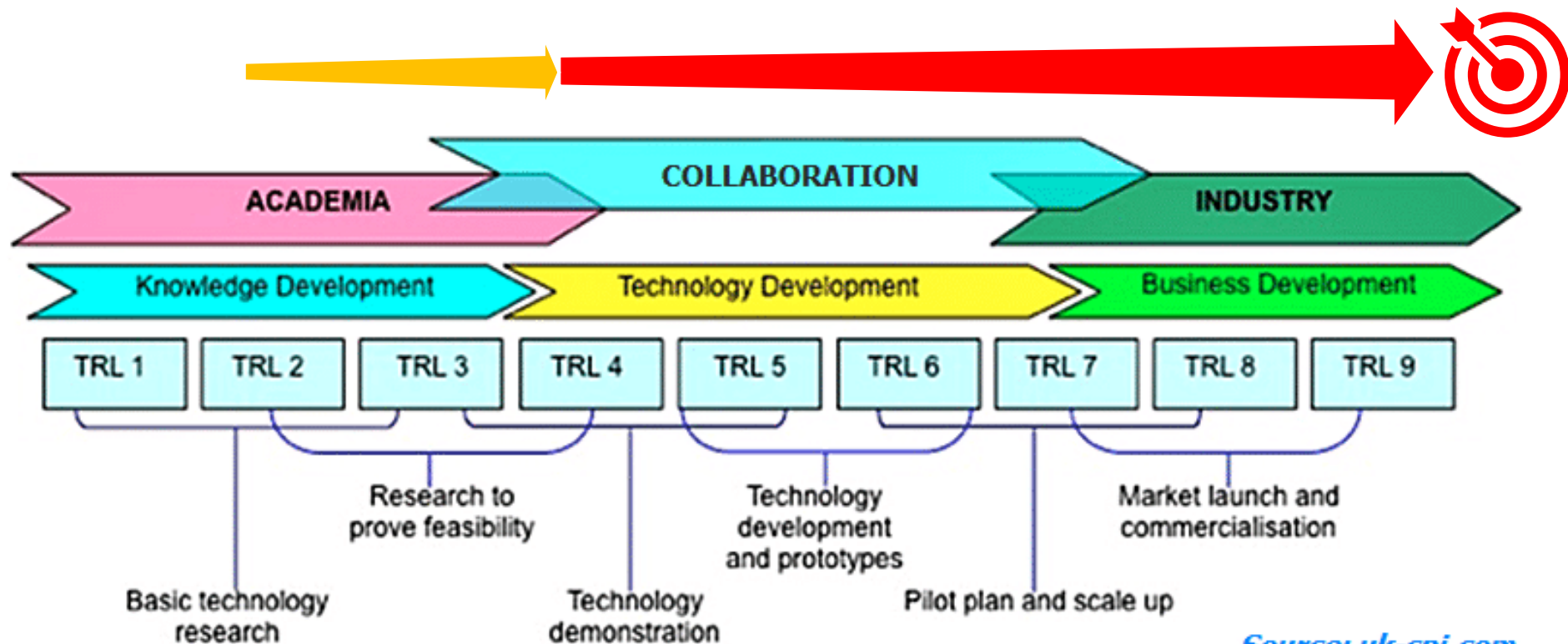
Works and Fields of Research

The background of the slide features a large, light blue silhouette of a human head in profile, facing right. The interior of the head is filled with a dense, overlapping pattern of numerous clock faces of various sizes and colors (shades of blue, green, and purple). The clock faces are arranged in a way that they appear to be floating or contained within the head, creating a complex, textured effect. The overall color palette is cool, with various shades of blue and green.

- Speech and Language processing
- ML/AI
- Speech synthesis
 - Thai, Myanmar, Lao, Malaysian
- Assistive Technology
- Accessible Media
- Sign Language Processing
 - Sign-to-Text Text-to-Sign
 - Motion Capture
- Language Resource
 - Multilingual Text and Speech corpus
 - Sign Language corpus

Focus to deliver Industry/Public Products and Services

- Medical devices, Health Care System
- Assistive Technology, Accessible Services



Today's Topics

- Let's know “Persons with Disabilities (PwD)”
- Assistive Technology : Technologies for PwDs
- Related Works@NSTDA
- From Lab To Market

Disabilities

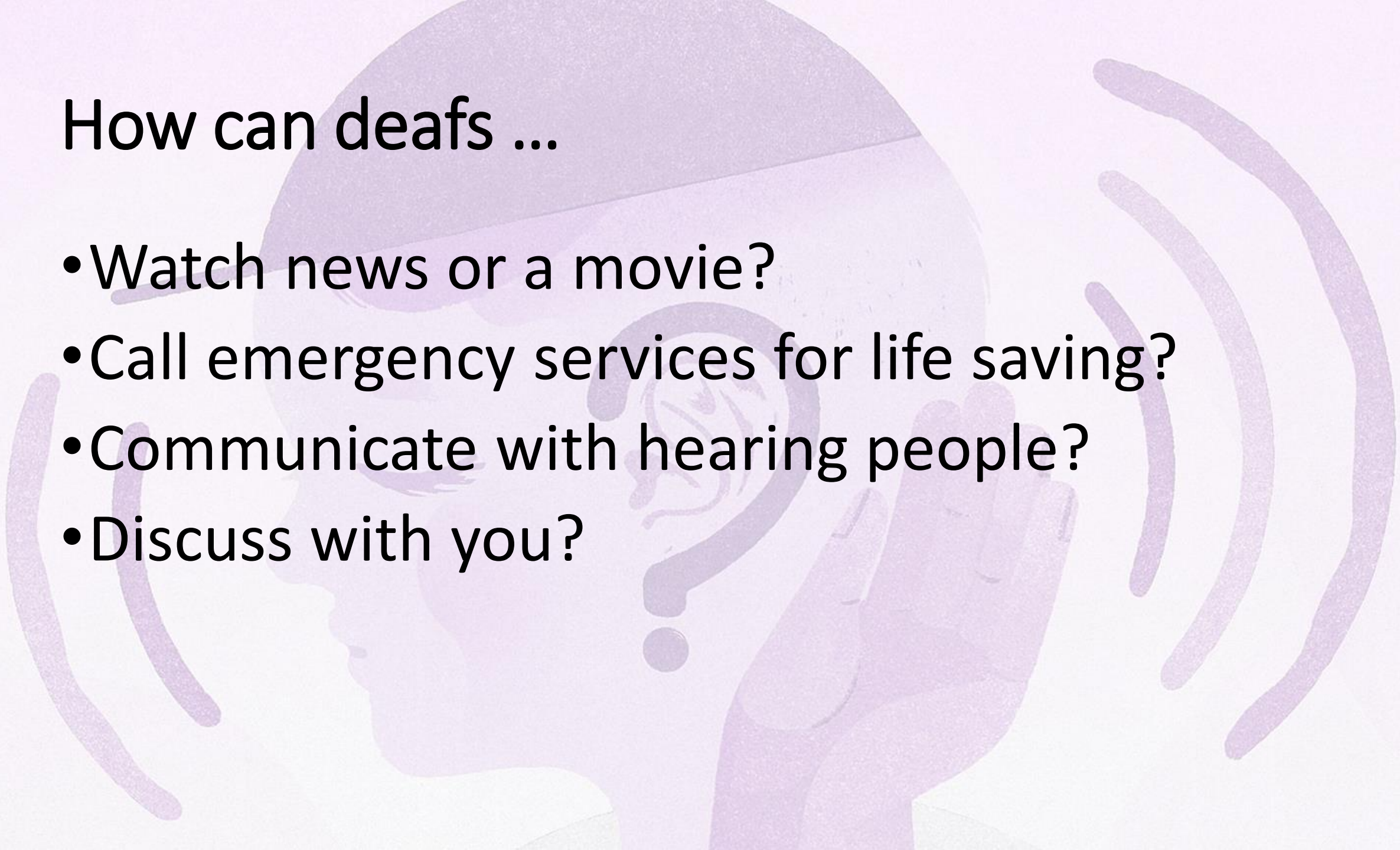


How can blinds ...

- Read a book?
- Preparing a report on a computer?
- See a movie?
- See this presentation?

How can deafs ...

- Watch news or a movie?
- Call emergency services for life saving?
- Communicate with hearing people?
- Discuss with you?



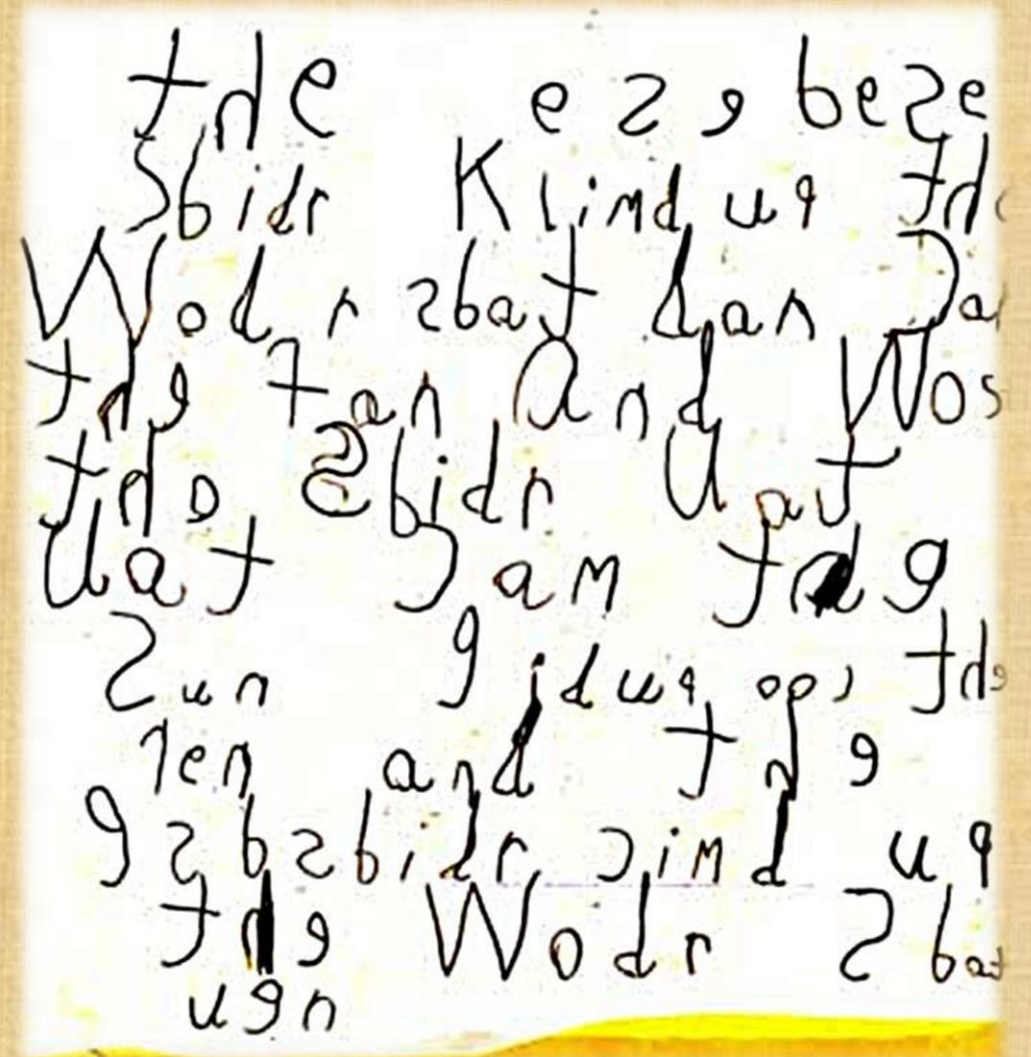
A young child with light brown hair and large, expressive eyes is the central focus. They are wearing a light pink short-sleeved shirt and have a clear tracheostomy tube secured with a grey strap around their neck. The child's hands are raised in front of them, palms facing forward. In the background, a woman with blonde hair, wearing a dark blue uniform and a lanyard, is crouching down. The setting appears to be a clinical or hospital room with light blue walls and medical equipment visible in the distance.

How can people with speech impairment ...

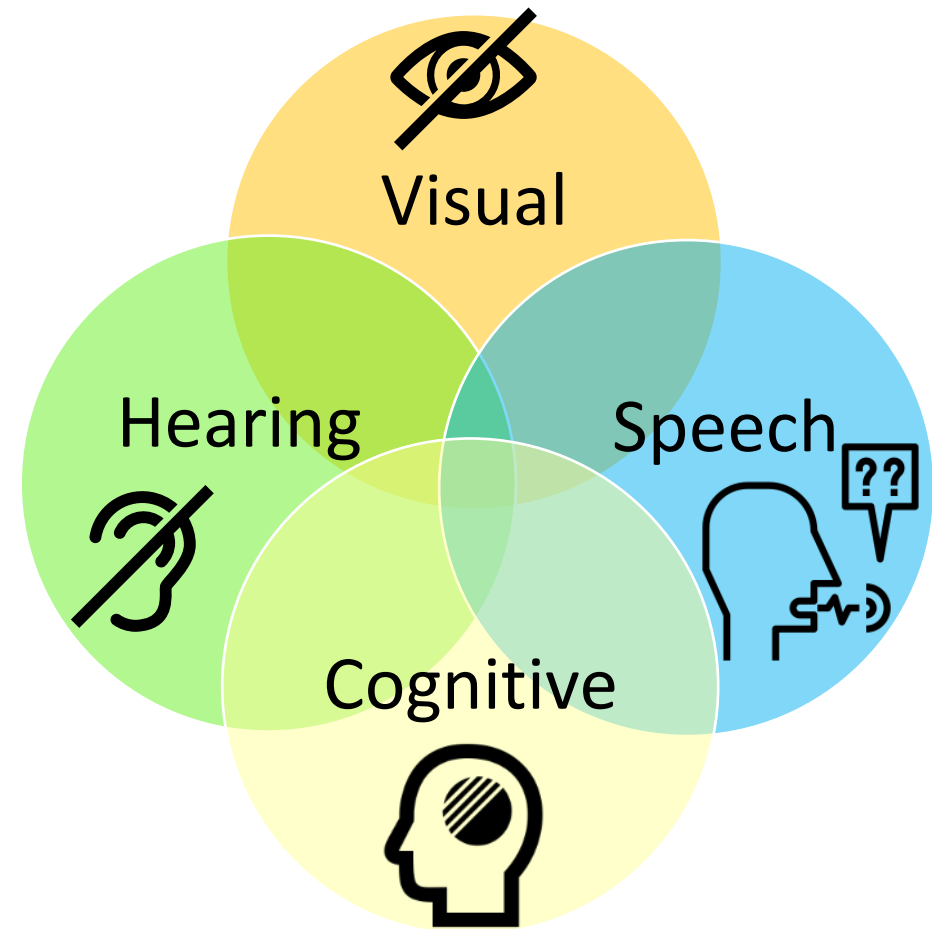
- Call 911?
- Ask for help?
- Talk with their parent?

How can people with learning disability...

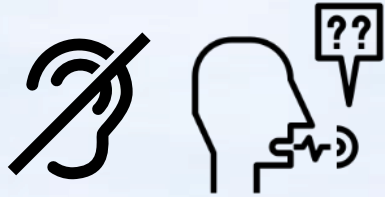
- Writing a report?
- Chat with you?
- Read a book?



Human as a social animal



Disabilities and Accessibility



Accessible communication



Accessible Media



Accessible Education



Assistive Technology

Accessible Media



Accessible Communication



Accessible Education



Core Technologies

Frontend

Automatic
Speech
Recognition
(ASR)

Text-to-Speech
Synthesis
TTS

Gesture
Recognition
(GR)

Pose
Generation
(PG)

Standard/Format

WCAG

EPUB

Backend

Natural
Language
Understanding
(NLU)

Machine
Translation
(MT)

Dialogue
Management
(DM)

Accessibility
Conformance

Text Suggestion

Assistive Technology

Accessible Media



Accessible Communication



Accessible Education



VAJA : Thai Text-to-Speech @NECTEC

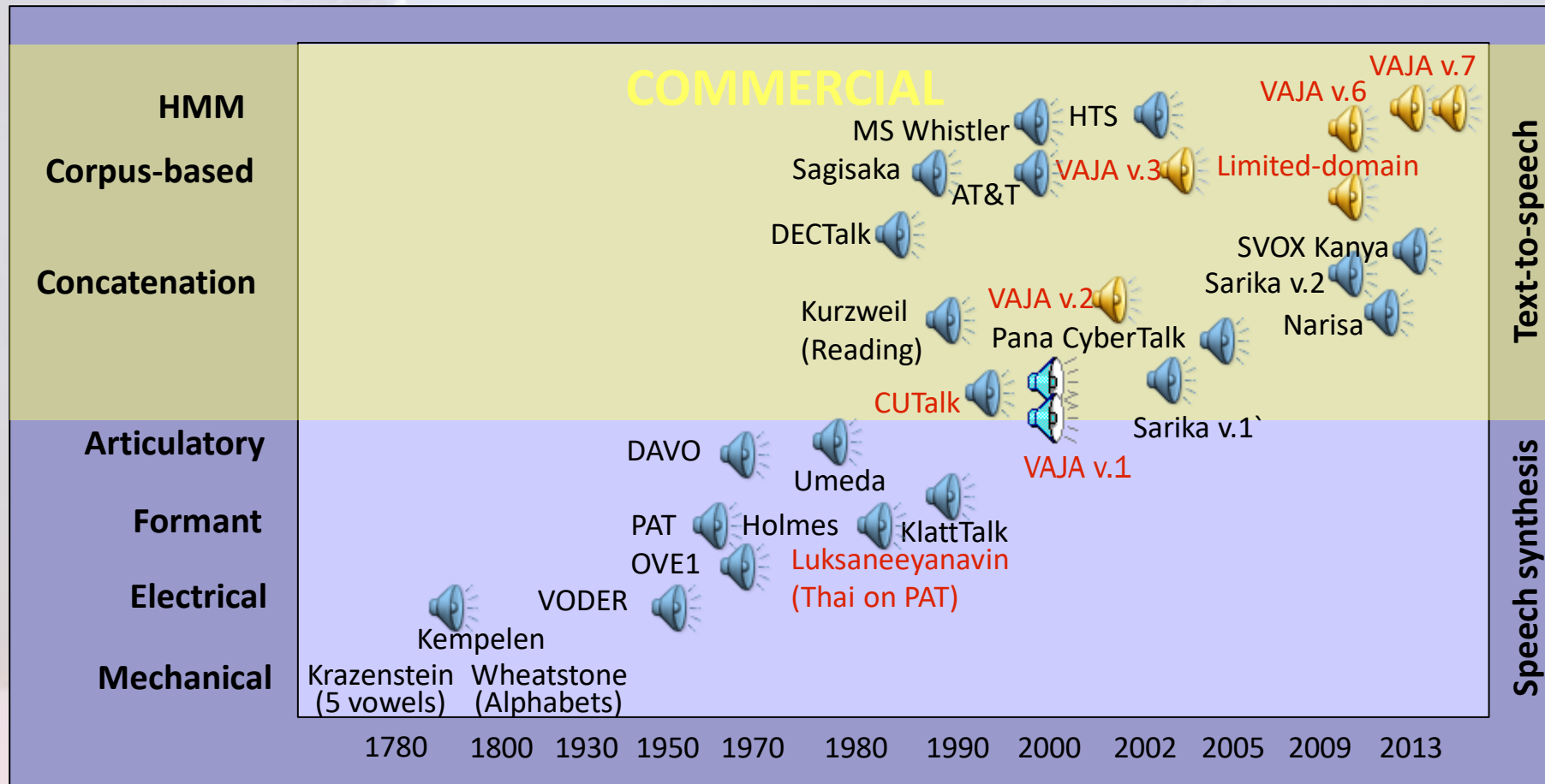
VAJA v.1
Simple Concatenation

VAJA v.2
Concatenation
+Signal Smoothing

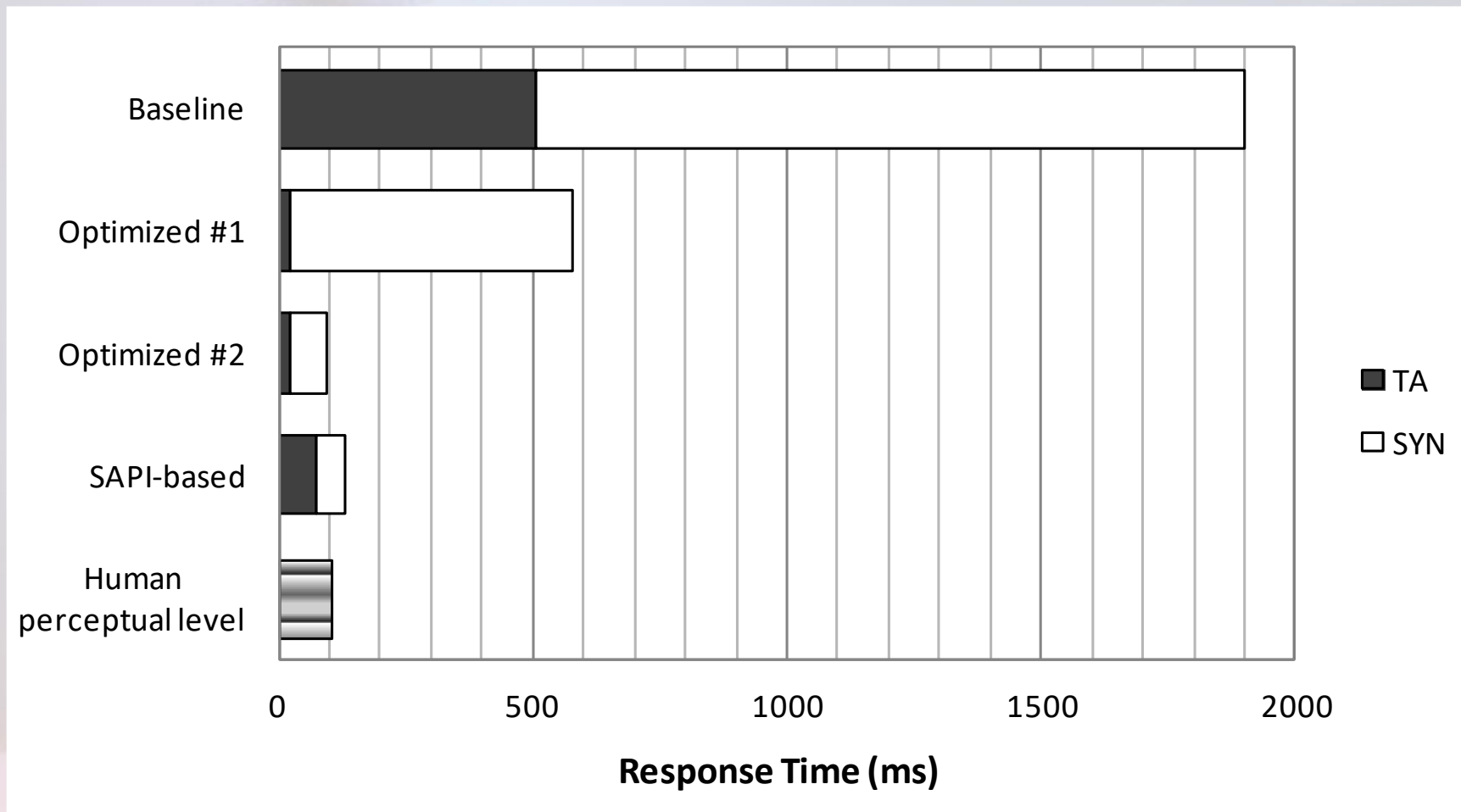
VAJA v.3
Unit Selection

VAJA v.7
HMM-based

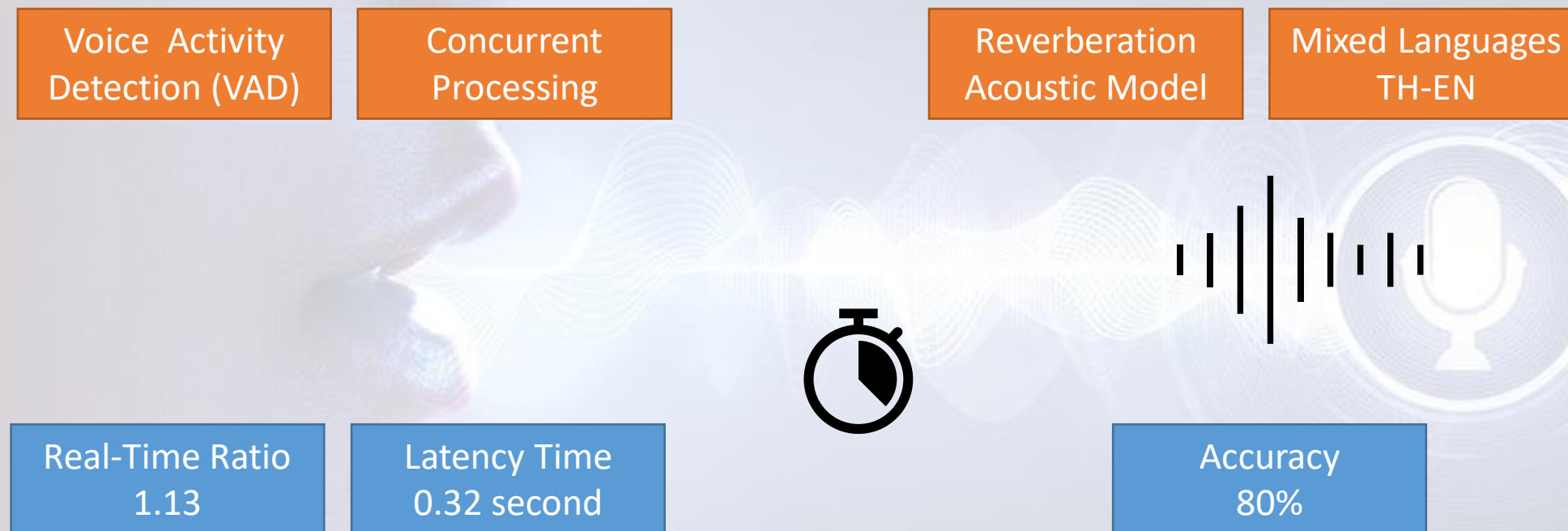
Latest VAJA
End-to-End TTS



VAJA 7 : Latency Time



Thai ASR : Partii @NECTEC



Under Real Office Environment from a TV Station

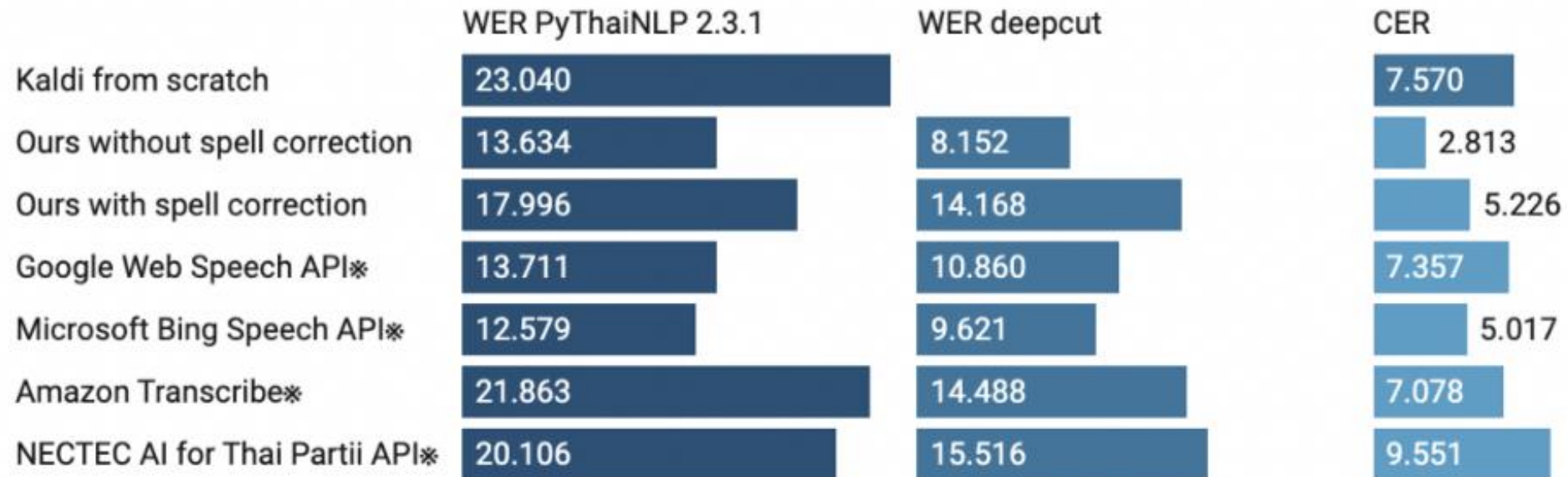
Word Error Rate : Training

ชุดทดสอบ	รายละเอียด
LOTUS-BN (BN)	<ul style="list-style-type: none">- Broadcast news speech- 3,140 utterances, 3 speakers
VoiceTra4u- Log (VT)	<ul style="list-style-type: none">- VoiceTra-4u application log- Sport and travel domains- 1,916 utterances, real environments
LOTUS-SOC (SOC)	<ul style="list-style-type: none">- Twitter text- 5,586 utterances, 7 noisy environments, Average 8.24 signal-to-noise ratio (SNR)
Lecture (LT)	<ul style="list-style-type: none">- Talking speech 1 speaker- 1 long speech (1hours 22 minutes), real environment s

Performance Evaluation by AIResearch.in.th and PyThaiNLP

Results on `test-unique` split of Thai Common Voice 7.0

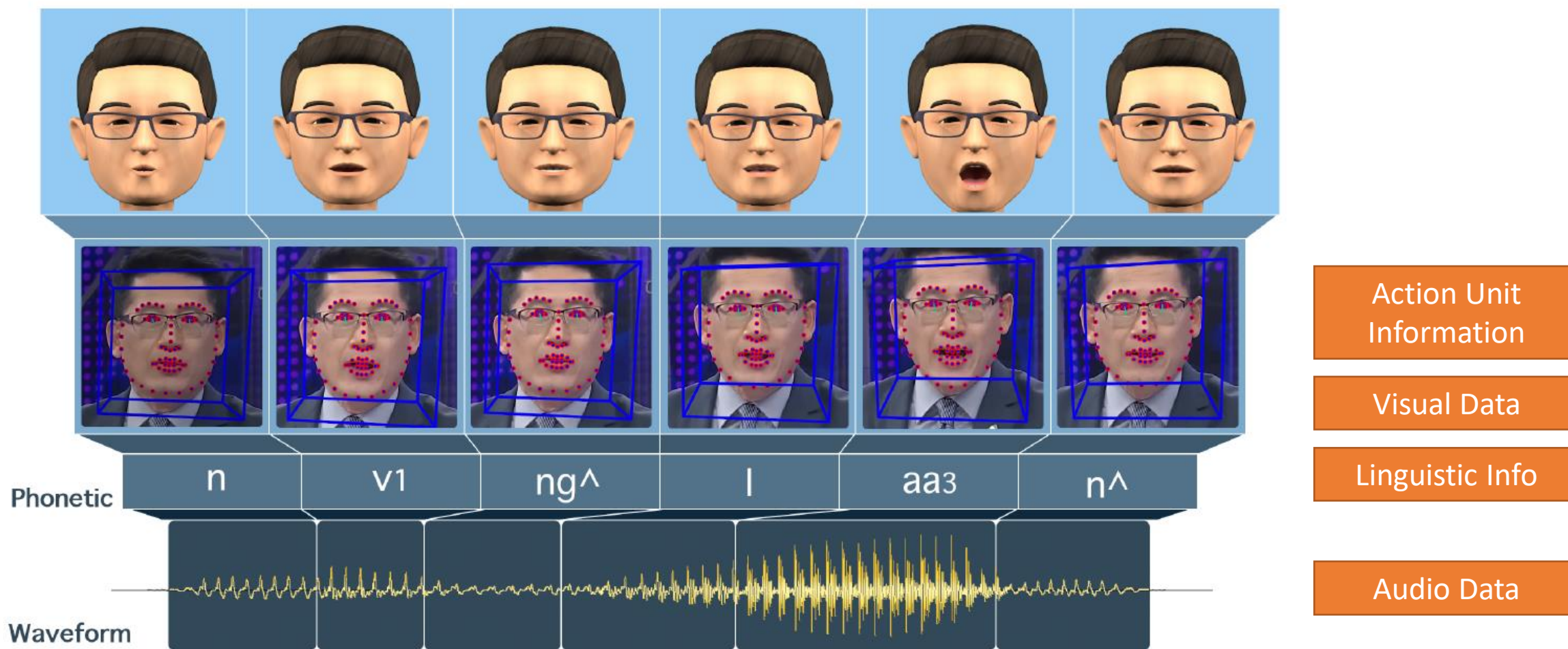
Word Error Rate (WER) and Character Error Rate (CER); the lower the better



* APIs are not fine-tuned with Common Voice 7.0 data

[Get the data](#) • Created with [Datawrapper](#)

Speech2Lip @NECTEC



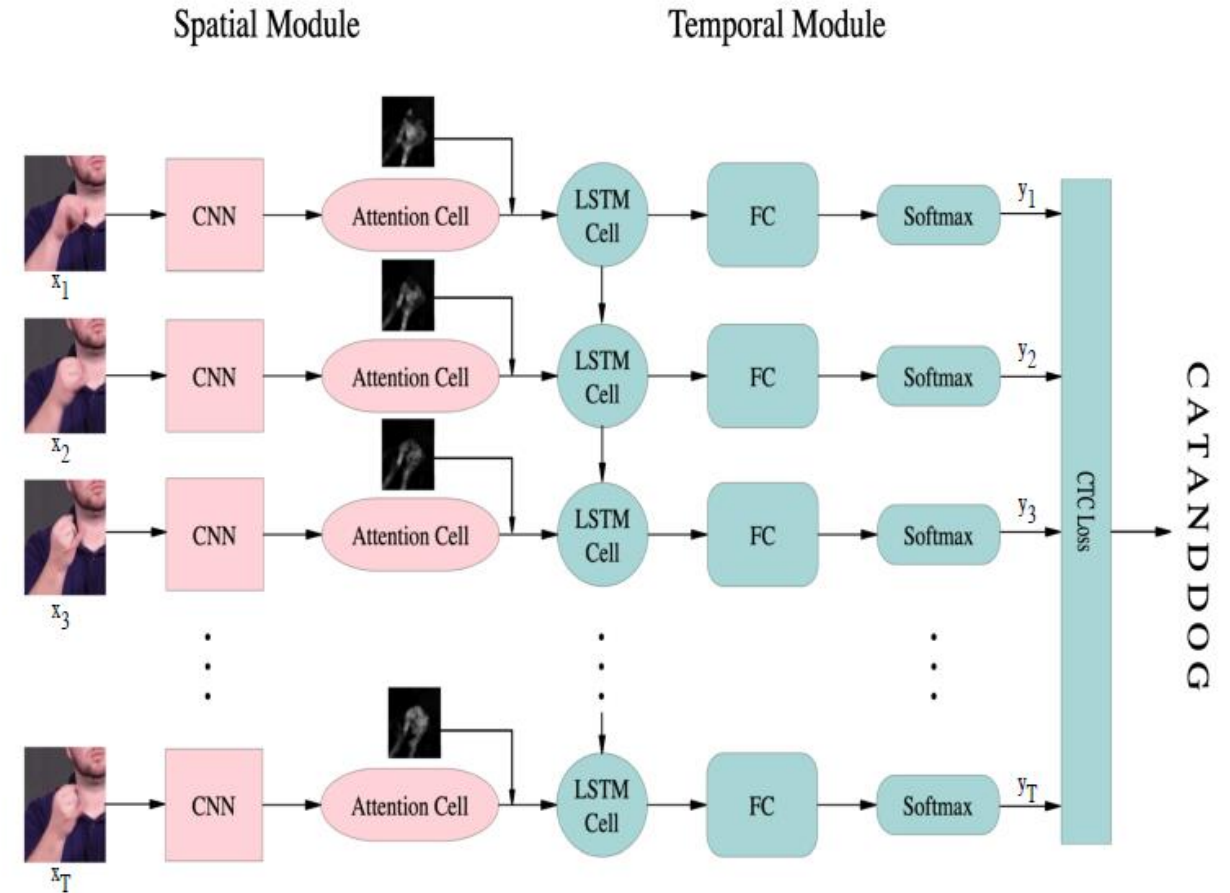
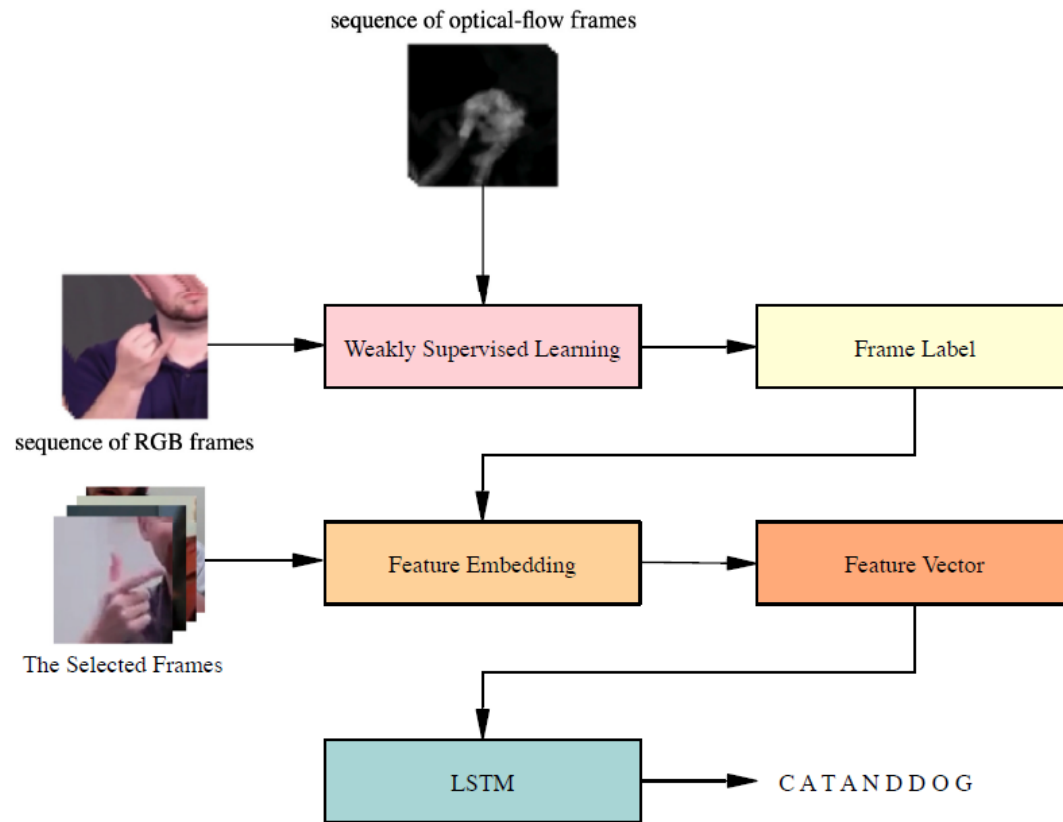
Sign Language Recognition Using Weakly Supervised Learning @A-MED



ChicagoFSWild The largest real-world Sign Language Corpus

P. Pannattee, W. Kumwilaisak, C. Hansakunbuntheung, and, N. Thatphithakkul, “Novel American Sign Language Fingerspelling Recognition in the Wild with Weakly Supervised Learning and Feature Embedding”, ECTI-CON 2021, p.291-294, 2021.

Sign Language Recognition Using Weakly Supervised Learning @A-MED



Hand Cropping using Iterative Attention and Cost Functions

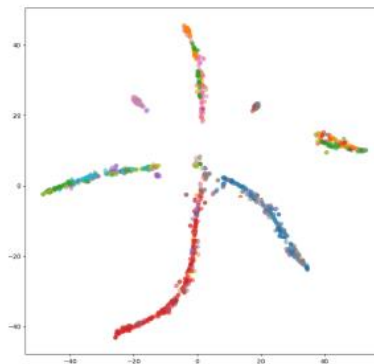
Optical Flow-based
Attention



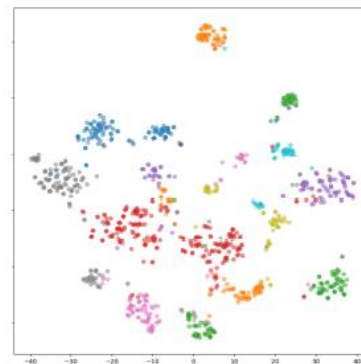
zoom factor = 0.72

zoom factor = 0.72

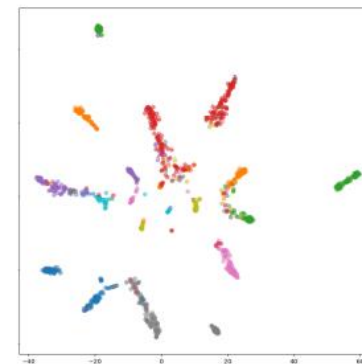
zoom factor = 0.81



(a)



(b)



(c)

(a) Cosine Embedding, (b) Cross-Entropy, and (c) Cosine Embedding conjunction with Cross-Entropy.

SL Recognition : Letter Error Rate (LER)

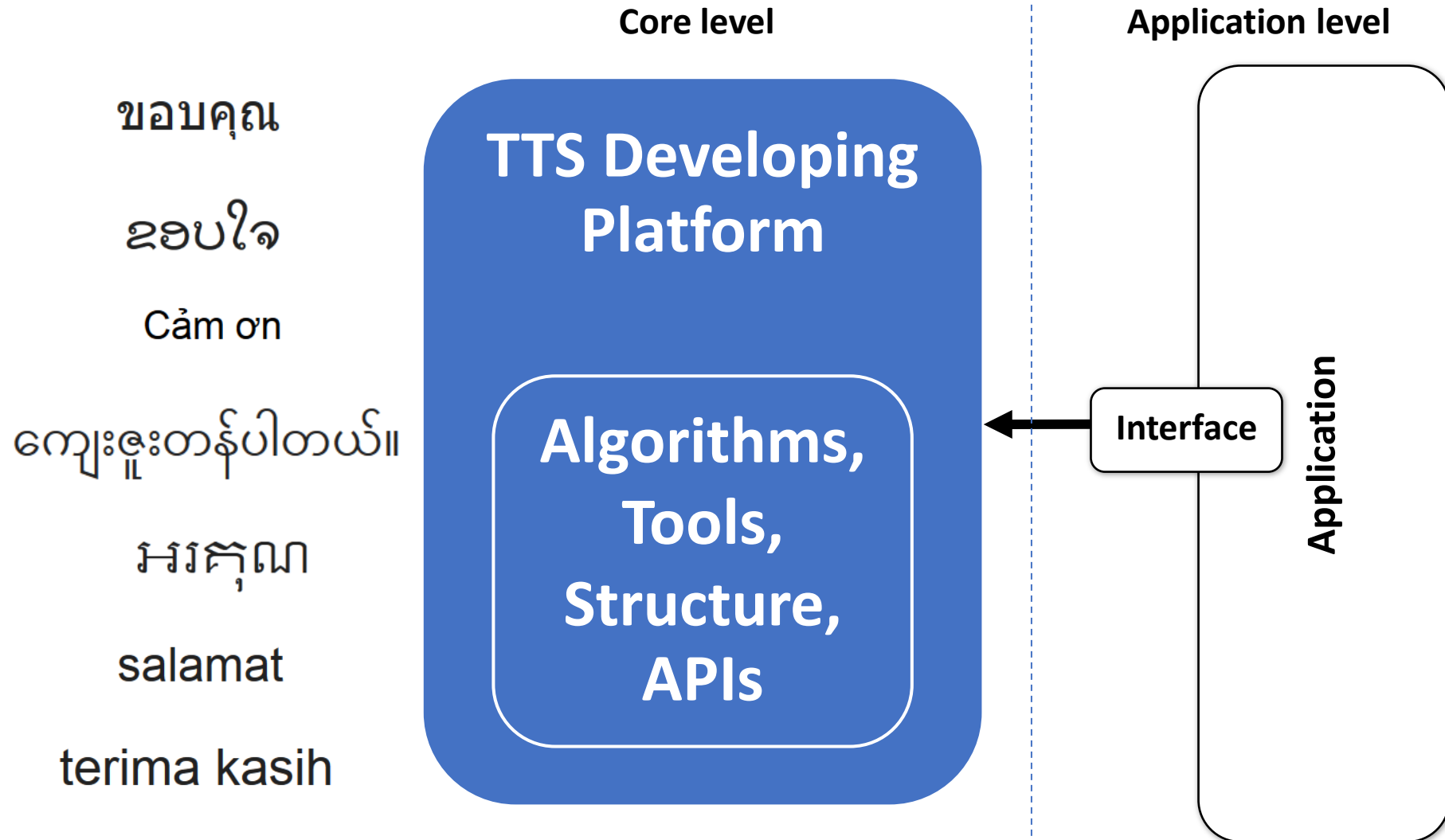
TABLE I
PERFORMANCE EVALUATION IN THE DEVELOPMENT SET.

Method	Cropping Method	$1 - LER$ (Percent)
AlexNet	-	11
Spatial Attention	-	23
Spatial Attention	Hand ROI	43.1
AlexNet [4]	Hand ROI	42.8
Spatial Attention [1]	Iterative Attention	46.8
The proposed method	Iterative Attention	50.5

TABLE II
PERFORMANCE EVALUATION IN THE TESTING SET.

Method	Cropping Method	$1 - LER$ (Percent)
AlexNet [4]	Hand ROI	41.9
Spatial Attention [1]	Iterative Attention	45.1
The proposed method	Iterative Attention	48.0

Multilingual TTS @NSTDA

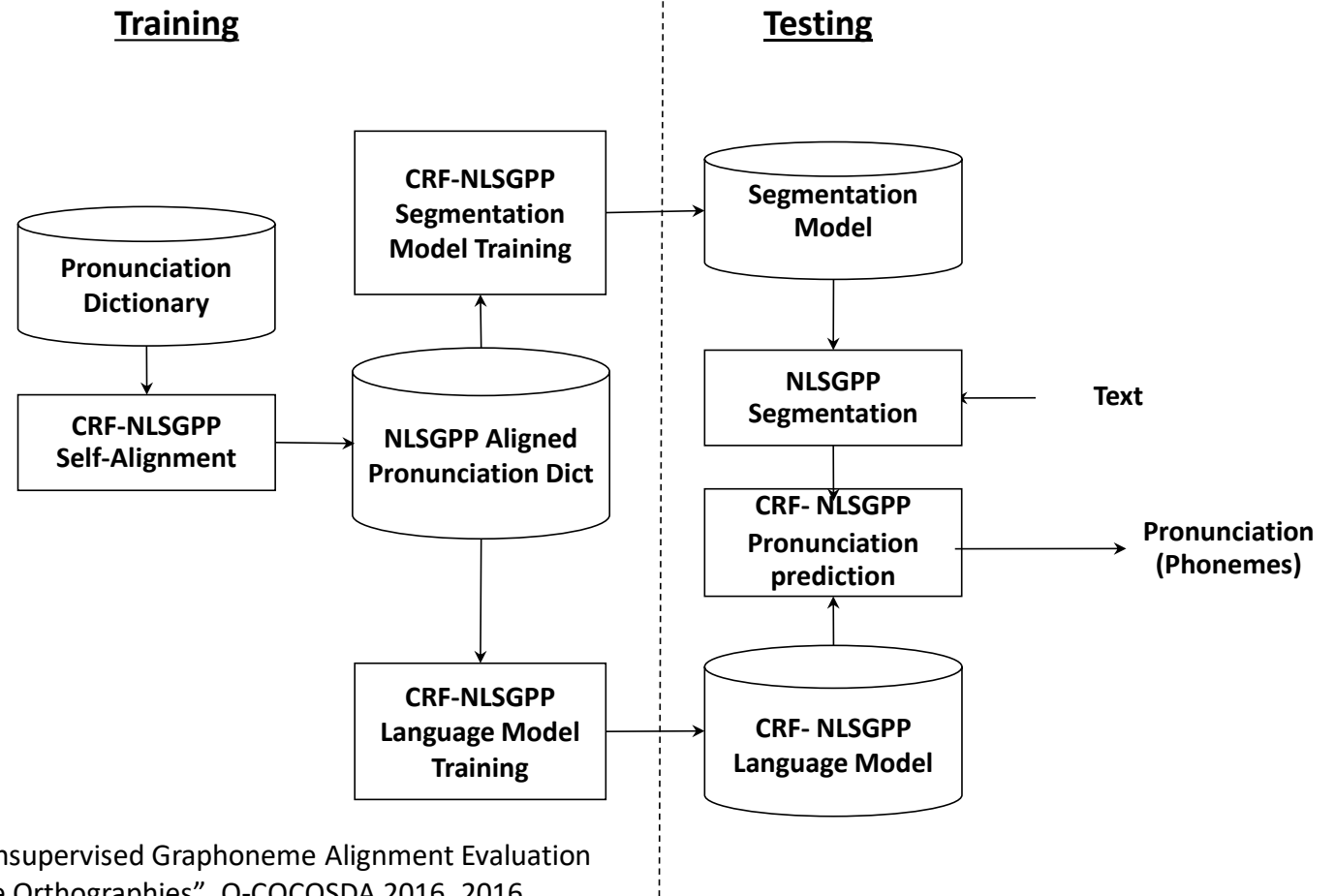
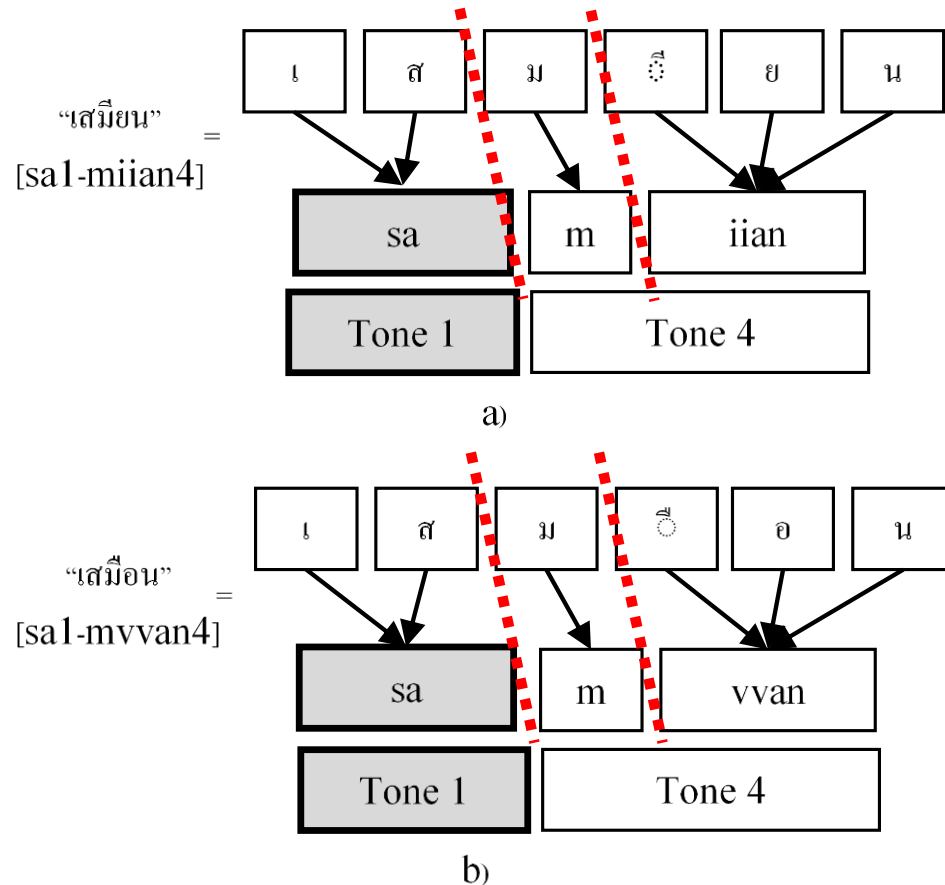


ກຮັມ	kr a0L m^
ກຮາຟ	kr aa3 f^
ກຮາຟິກ	kr aa0L f i1 k^
ກຮາມ	kr aa0L m^
ກຮິດາ	kr i0H d aa0L
ກຮິສຕ໌	kr i0H s^
ກຮີງ	kr i3 ng^
ກຮິກ	kr ii3 k^
ກຮິນທິ	kr ii0L n^ th ii0H
ກຊງສີ	kr u0L ng^ s ii4

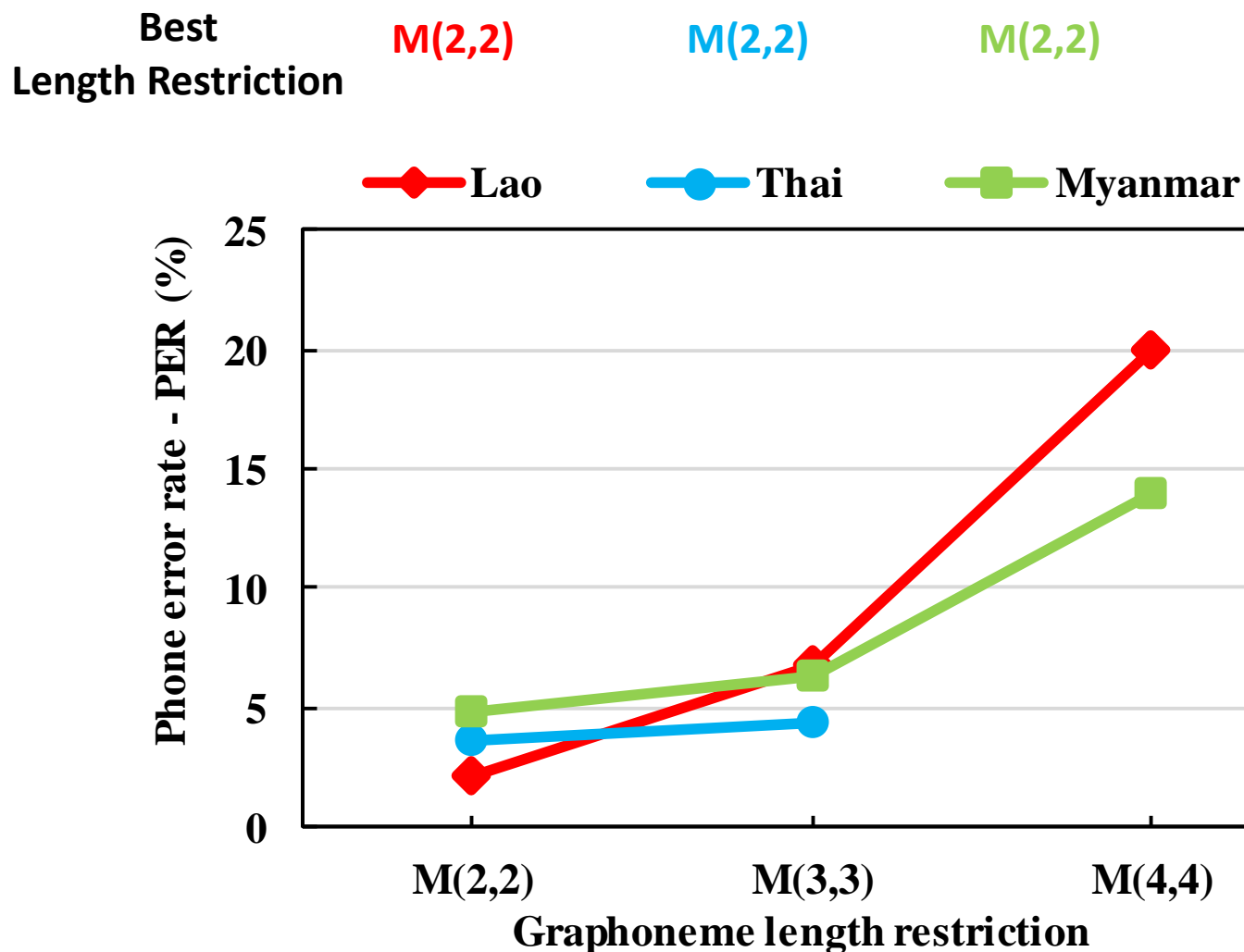
ກກຸດຊັ	kā ku θaː
ກກຸຮາ	keʔ kə̀rə
ກກຸຮຸ	keʔ kə̀ru
ກກຸກຸມຸນຸ	keʔ ku kə̀myiː
ກກຸກຸຊຸ	keʔ ku zú
ກກຸກກຸລຸ	keʔ keʔ laː
ກກຸກກຸສຸ	keʔ kiː daʔ
ກກຸກກຸຕກຸ	keʔ kə̀ dəgaː
ກກຸກກຸ	keʔ sʰeʔ
ກກຸກກຸກກຸ	keʔ təlaʊʔ
ກກຸກກຸກກຸ	keʔ pətèi
ກກຸກກຸກກຸ	keʔ bi neʔ
ກກຸກກຸກກຸ	keʔ θəliʔ

mikhael	m-i x-a e-l
mikhail	m-i x-ai-l
miko	m-i k-o
miky	m-i k-i
mil	m-i-l
mila	m-i a
milah	m-i a-h
milan	m-i a-n
miley	m-i e-j
milial	m-i -i j-a
milik	m-i e-hz
miliki	m-i -i k-i
milikku	m-i -i-hz k-u
milikmu	m-i -i-hz m-u

Non-uniform Length-restricted Syllable-unbounded Grapheme-Phoneme Pair (NLSGPP) using Conditional Random Field



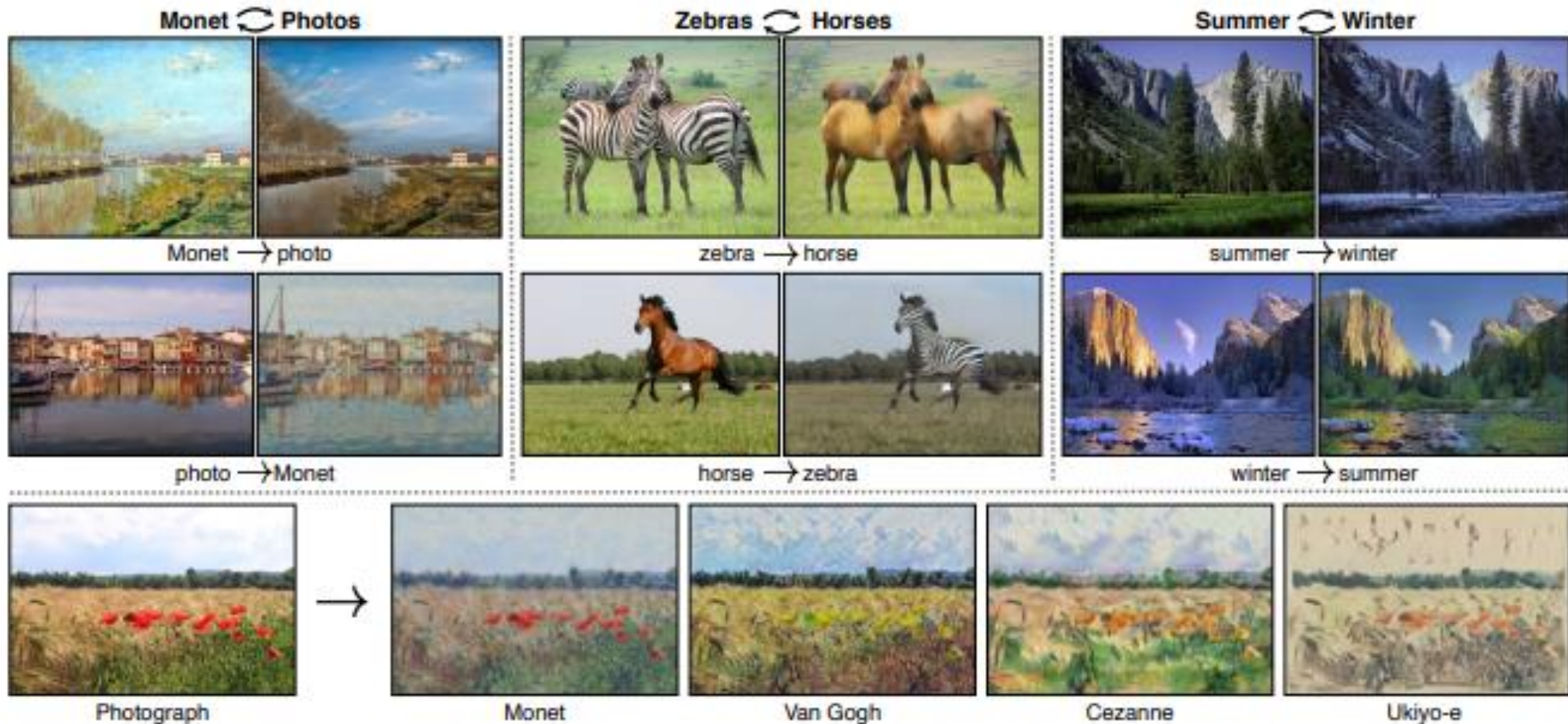
Length-restriction Effects in Phone Level



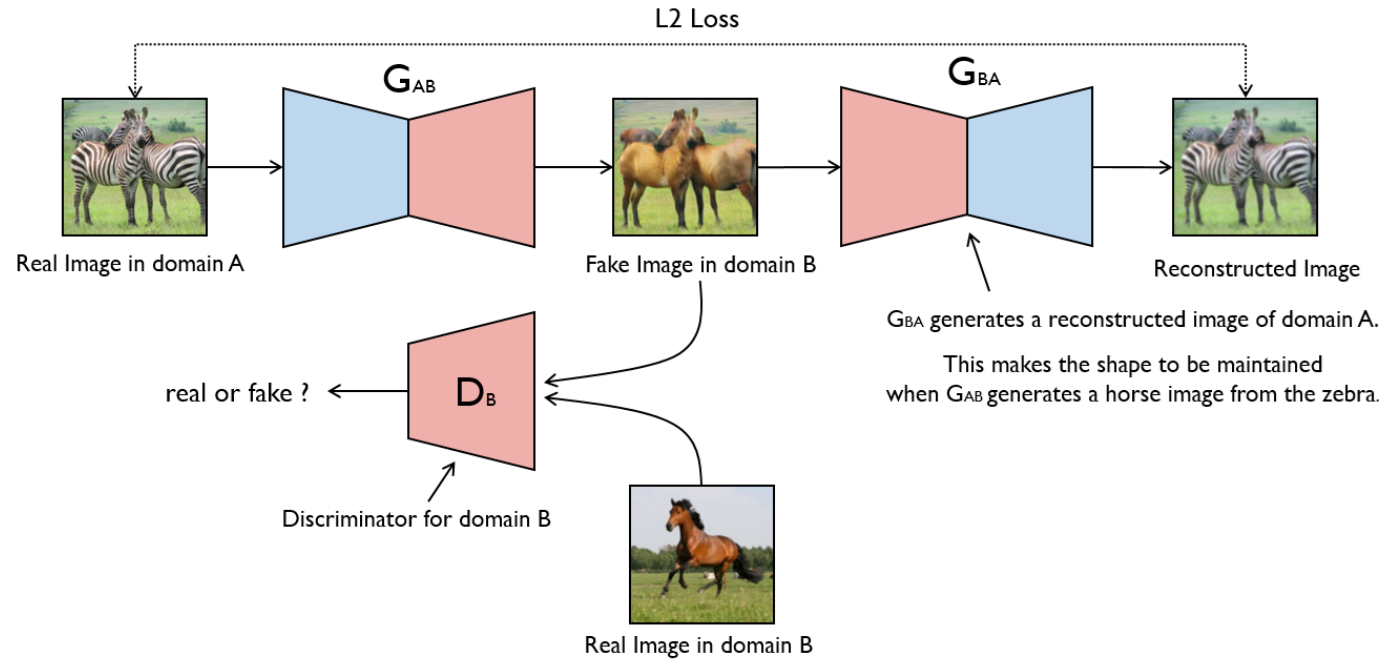
Speech Enhancement using CycleGAN

@NSTDA

CycleGAN



Speech Enhancement using CycleGAN



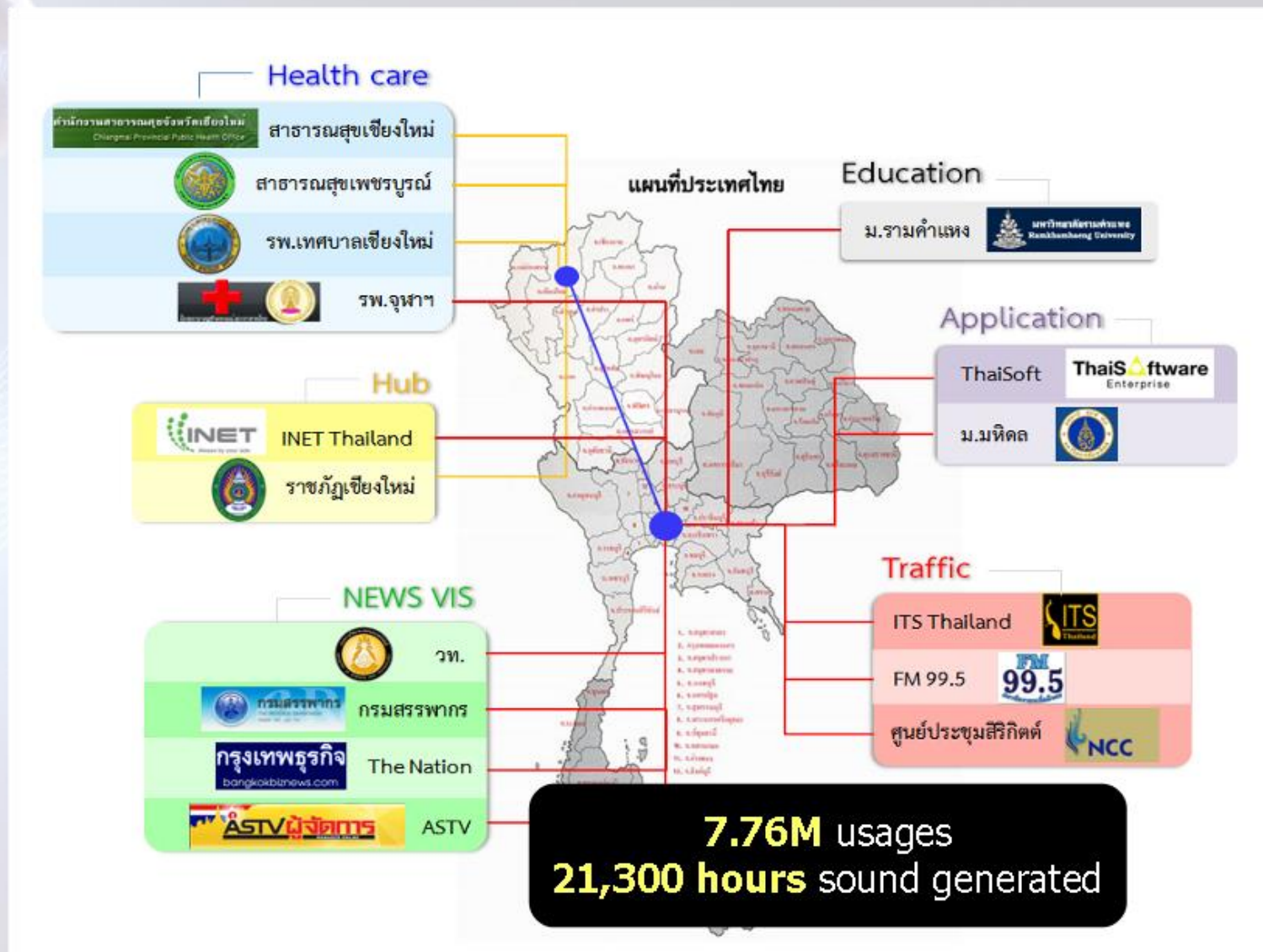
	Original	Enhanced
Example 1		
Example 2		

Alternative Keyboard for Person with Cerebral Palsy



From Lab To Market

VAJA : From Lab to Market



TTS in a self-management support system for diabetes patients over telephones @Chulalongkorn Hospital



เช็กเบาหวานผ่านมือถือ

เทคโนโลยีเพื่าระวังโรคภัย

เตือน และเกิดผลที่เห็น ค่ะ การนำเทคโนโลยีเข้ามาประยุกต์ใช้ในรูปแบบการดูแลรักษาโรคเรื้อรัง โดยมุ่งเน้นการให้ความรู้กับผู้ป่วย เพื่อให้มีความรู้ความเข้าใจ สามารถดูแลตนเองในเบื้องต้นได้ พร้อมทั้งนำที่กระตุ้นเตือนเมื่อถึงกำหนดการติดตาม จะช่วยให้การดูแลสุขภาพของผู้ป่วยทำได้ดียิ่งขึ้น

ทีมวิจัยเนคเทคยังได้พัฒนาระบบส่งเสริมการจัดการดูแลตนเองทางคอมพิวเตอร์และอินเทอร์เน็ตสำหรับผู้ป่วยเบาหวาน เพื่อให้ผู้ป่วยสามารถกำหนดวิธีการดูแลสุขภาพ และประมวลผลการรับประทานอาหาร และออก

ถ้ามีสายโทรเข้าโทรศัพท์มือถือ และเสียงจากปลายสายเป็นบันทึกเสียงอัตโนมัติที่คอยตอบคำถามเหมือนคนจริง คำสั่งชักประวัติผู้ป่วย อย่าเพิ่งแปลกใจ หรือเข้าใจผิดว่าเป็นสายจากมิจฉาชีพ หลอกให้โอนเงินเข้าบัญชี

แต่เป็นเทคโนโลยีที่โรงพยาบาลจุฬาลงกรณ์ และทีมวิจัยจากศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ หรือเนคเทค พัฒนาเพื่อติดตามอาการโรคแบบถึงเนื้อถึงตัว เพื่อเฝ้าระวังความเสี่ยงให้แก่ผู้ป่วยโรคเบาหวานที่รักษาตัวอยู่กับโรงพยาบาลจุฬาฯ

นพ.วิโรจน์ เจียมจรัสรังษี ผู้ช่วยอธิบดีฝ่ายส่งเสริมสุขภาพ คณะแพทยศาสตร์ โรงพยาบาลจุฬาลงกรณ์ ระบุว่า ทีมวิจัย นอกว่า เบาหวาน

พัฒนาเป็นระบบส่งเสริมการจัดการดูแลตนเองทางโทรศัพท์สำหรับผู้ป่วยเบาหวาน "ระบบนี้ช่วยให้ผู้ป่วยสามารถดูแลตนเองได้จากที่บ้าน โดยผลจากการตอบคำถามผ่านโทรศัพท์แบบตอบกลับ จะสามารถวิเคราะห์ข้อมูลสุขภาพได้ทันที ในกรณีที่ผู้ป่วยมีความเสี่ยงเกิดโรคแทรกซ้อน

กำลังกาย โดยระบบจะแจ้งเตือนให้รีบพบแพทย์ผ่านแอปพลิเคชัน "ระบบดังกล่าวคาดว่าจะเริ่มต้นใช้งานภายในเดือนกันยายน โดยนำร่องกับกลุ่มผู้ป่วยเบาหวานที่สมัครใจเข้าร่วมโครงการ เพื่อทดสอบประสิทธิภาพการใช้งาน นอกจากนี้ยังได้เห็นหน้าคือเพื่อผสม

A screenshot of a mobile application interface. It shows a flowchart with several steps, including '1. ตรวจสอบระดับน้ำตาลในเลือด' (Check blood sugar level), '2. บันทึกผลลงในสมุดบันทึก' (Record results in the notebook), and '3. แจ้งเตือนเมื่อถึงกำหนดการติดตาม' (Alert when it's time for follow-up). There are also icons of people and a calendar.

Lao TTS @ E-Government Center, Lao PDR



Thailand Transcription Service Center

Closed Caption (CC) and Sign Language (SL) Services

- National Broadcast Service of Thailand (NBT) station
- Thai PBS station
- etc.



Thai Telecommunication Relay Service (TTRS)



Web and Mobile Apps



Kiosk



Video Phone



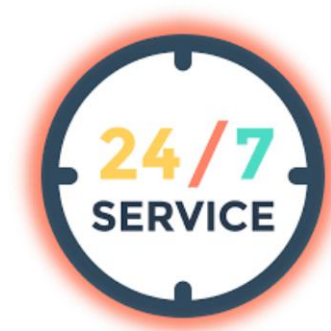
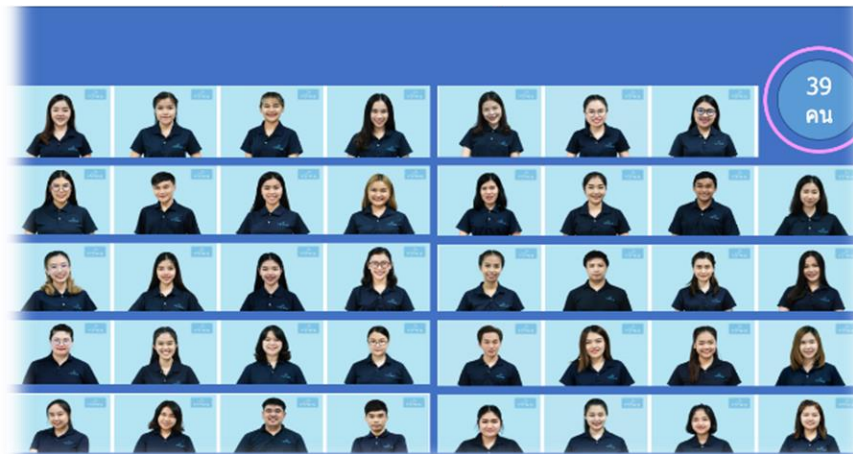
Video Relay Service



Text Relay Service



ล่ามภาษามือ ศูนย์ TTRS



Thai Telecommunication Relay Service (TTRS)

Web and Mobile Apps



Kiosk



180 sets

Video Phone



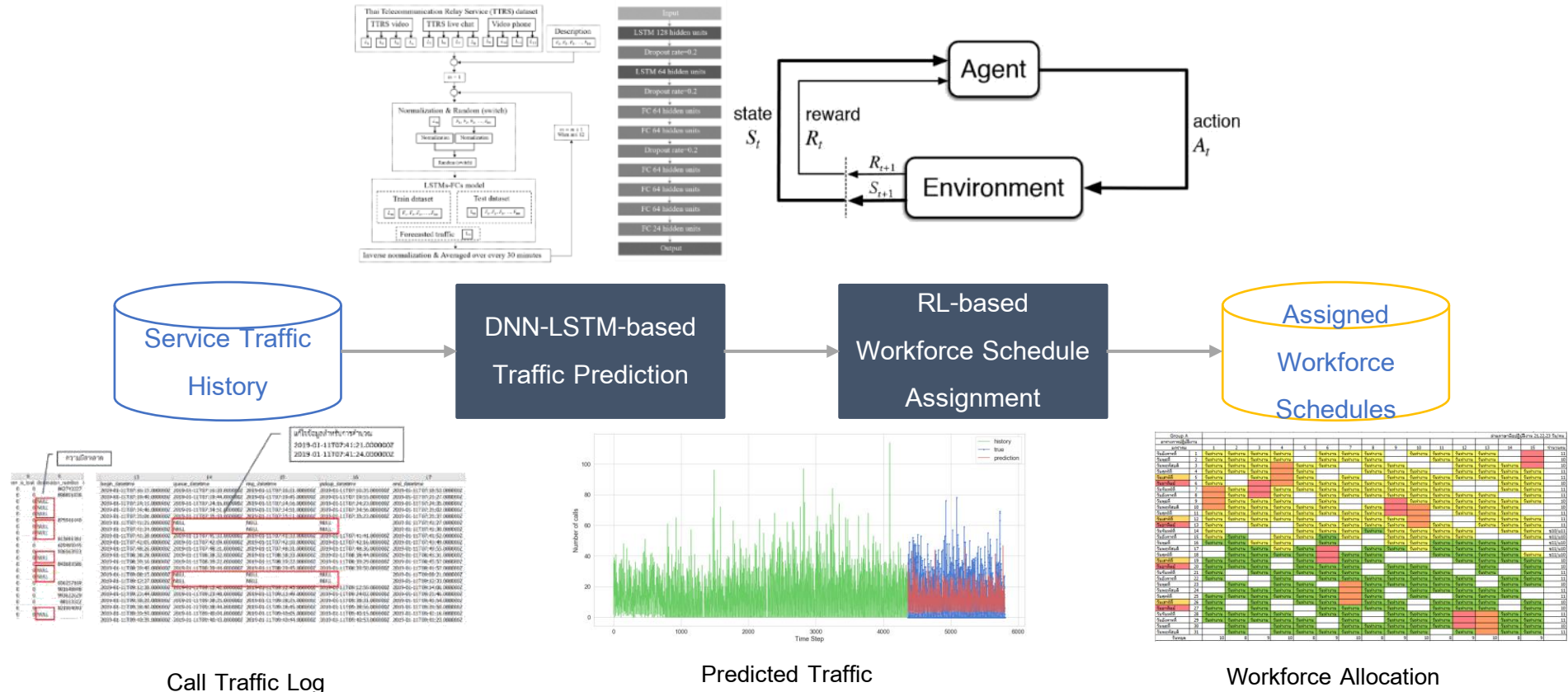
127 sets

Over

177,003 Calls



Workforce Management using DNN-LSTM and Reinforcement Learning



S. Phikulngoen W. Kumwilaisak, N. Thatphithakkul, C. Hansakunbuntheung, and W. Phantachai, "Call Center Workforce Management with Neural Networks and Reinforcement Learning Reinforcement Learning", 2nd ICA-SYMP 2021.

Workforce Management Evaluation

TABLE III: Forecasting error comparison.

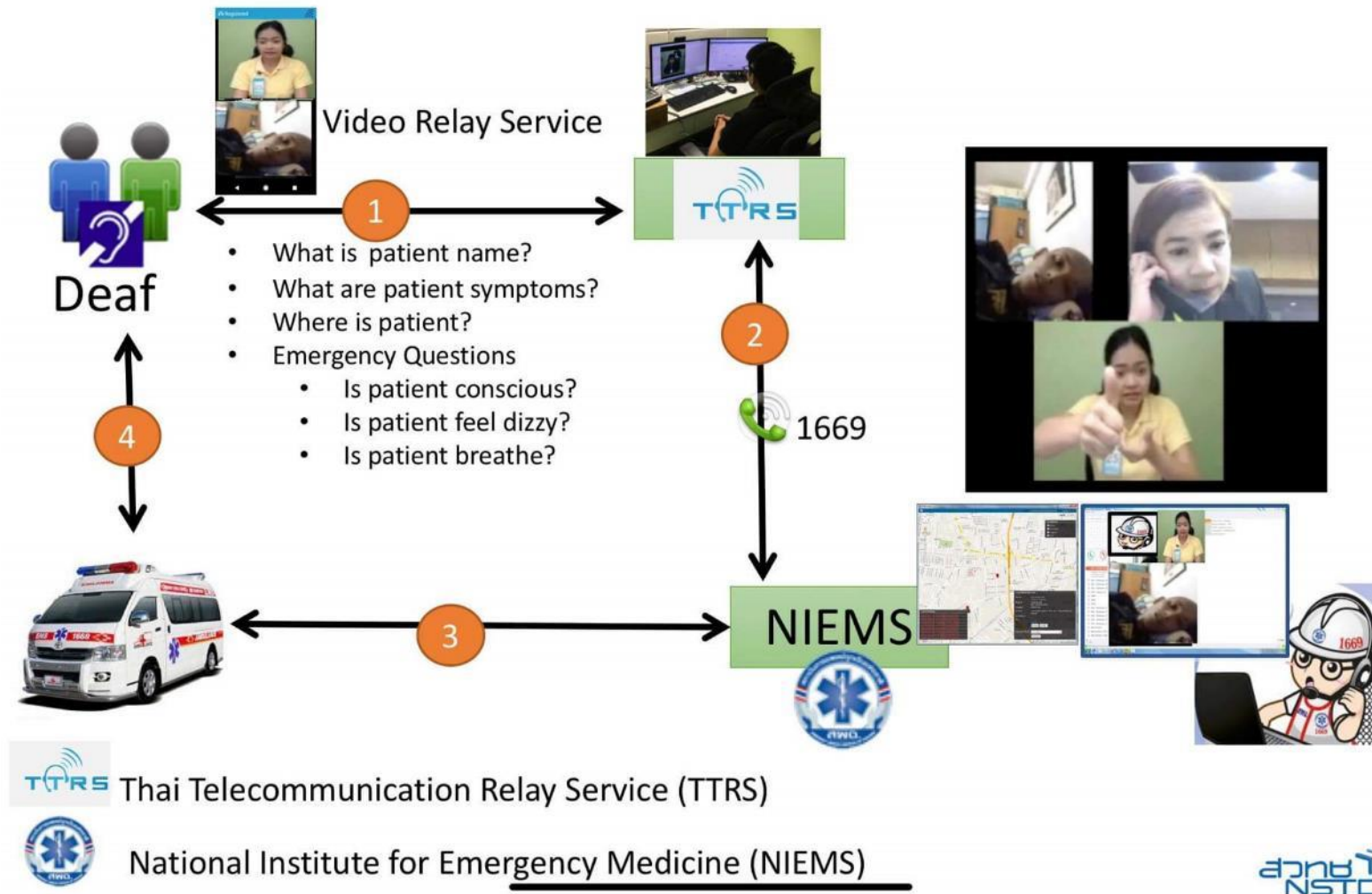
	LR	SES	HW	LSTM	Proposed
Incoming(%)	52.12	45.37	29.72	6.28	0.98
Received(%)	49.69	42.12	28.69	5.32	1.56
$T_s(\%)$	50.81	43.77	29.94	7.45	1.98
$T_w(\%)$	49.84	44.93	28.77	8.94	1.23

TABLE IV: Service performance comparison.

	LR	SES	HW	LSTM	Proposed
Abandon(%)	42.71	38.84	19.42	9.62	3.54
ASA(sec)	43	41	24	17	11
SLA(%)	36	32	61	77	93

S. Phikulngoen W. Kumwilaisak, N. Thatphithakkul, C. Hansakunbuntheung, and W. Phantachat, "Call Center Workforce Management with Neural Networks and Reinforcement Learning Reinforcement Learning", 2nd ICA-SYMP 2021.

TTRS for Emergency Medical Services



From Lab To Market



Thank you for your attention

Q&A

Chatchawarn Hansakunbuntheung
email: chatchwarn.han@nstda.or.th