Internal Use Only（非公開）

TR-SLT-0010

Thai Speech Recognition by Acoustic Models Mapped from Japanese

Sawit Kasuriya　Takatoshi Jitsuhiro　Genichiro Kikui

2002. 3.11

This report describes Thai speech recognition by using Japanese acoustic models. There was no Thai speech database for use in this study, and it takes a lot of time to correct a large number of speech database entries and label them. We also have only a small amount of Thai speech data, which was not alignment data. Therefore, we made initial models by mapping another language's acoustic models, and trained the initial models by using a small amount of data. We used phoneme mapping by the IPA table and made the initial Thai models from Japanese models. In our experiments, the initial models were trained by using isolated-word utterances by eleven Thai speakers. We evaluated the performance of isolated-word recognition with context-independent and context-dependent models.

（株）国際電気通信基礎技術研究所
音声言語コミュニケーション研究所
〒619－0288 京都府相楽郡精華町光台二丁目 2 番地 2　TEL：0774－95－1301

Advanced Telecommunication Research Institute International
Spoken Language Translation Research Laboratories
2-2-2 Hikaridai Seika-cho Soraku-gun Kyoto 619-0288,Japan
Telephone:+81-774-95-1301
Fax　　:+81-774-95-1308

Contents

1 Introduction

In a survey on Thai speech recognition, many researchers were found to have focused on word-base recognition [Ahkuputra, 1997] and phoneme-base recognition [Ekkarit, 2000]. Furthermore, research has never been done on continuous speech. The main problem was the need for a large Thai speech database. Therefore, National Electronics and Computer Technology Center, Thailand (NECTEC) were collecting and aligning a Thai speech database for speech recognition.

In this research, the target language is Thai. We also had only a small amount of Thai speech data and none of the utterances were aligned. Recently, many researchers have focused on the question of how to build a large-vocabulary continuous speech recognition (LVCSR) system for a new target language using speech data from varied source languages. They have tried to build a multilingual speech recognition system [Tanja, 2001] [Uebler, 2001]. Therefore, we used such a method and investigated Thai speech recognition. We made the initial Thai acoustic models by mapping from Japanese acoustic models. The Japanese acoustic models already exist. We used International Phonetic Alphabet (IPA) tables to map between Japanese and Thai phonemes. Our research was evaluated with Thai isolated-word recognition using Thai isolated-word utterances for training. They were spoken by eleven Thai speakers.

2 Making Thai Acoustic Models

We made the Thai acoustic models by using Japanese acoustic models because we did not have a large enough Thai database to train Thai acoustic models directly and the Thai database did not have segment information. The Japanese acoustic models were trained from a very large database. This should make good initial models for the Thai acoustic models. Therefore, we converted the Japanese acoustic models to from the initial Thai acoustic models. After we got the initial models for the Thai acoustic models, they were trained with some Thai utterances using the Baum-Welch algorithm.

2.1 Thai phonemes

We will briefly describe Thai phonemes. The Thai language is a tonal language. Thai syllables are C_iV, C_iVC_f. There are about 30,000 syllables [Sudaporn, 1993]. Each syllable has a tone and all tones have five several tones- a high tone, a middle tone, and a low tone in the static group, and a rise tone and a fall tone in the dynamic group.
The initial consonants (C_i) of the Thai language are 21 for single, twelve for double, and more than five initial double consonants for foreign languages. There are eight final consonants (C_f) and more than four for foreign languages. In Thai, there are nine short vowels, nine long vowels, and three short and three long double vowels. Thai vowels, single Thai consonants, and double Thai consonants are shown in Table 1, 2 and 3 as a sequence in appendix B. All of these Thai phonemes refer to studies by Thai linguists [Sudaporn, 1993].

Table 1. Single Thai consonants

21 Thai consonants

Place and Manner		Labial	Alveolar	Palatal	Velar	Glottal
Stop	Voiceless Unaspirated	p (ป)	t (ต ฏ)	c (จ)	k (ก)	ʔ (อ)
	Voiceless Aspirated	p^h (พ ภ ผ)	t^h (ท ธ ฒ ฑ ฐ ฏ)	c^h (ช ฌ ฉ)	k^h (ค ฆ ข)	
	Voiced	b (บ)	d (ด ฎ)			
Non-Stop	Nasal	m (ม)	n (น ณ)		ŋ (ง)	
	Fricative	f (ฟ ฝ)	s (ซ ศ ษ ส)			h (ฮ ห)
	Trill		r (ร ฤ)			
	Lateral		l (ล ฬ)			
	Approximant	w (ว)		j (ย ญ)		

Table 2. Double Thai consonants

12 double Thai consonants

Unaspirated Stop Set	pr (ปร)	tr (ตร)	kr (กร)
	pl (ปล)		kl (กล)
			kw (กว)
Aspirated Stop Set	p^hr (พร)	t^hr (ทร)	k^hr (คร)
	p^hl (พล)		k^hl (คล)
			k^hw (ขว)

Table 3. Thai vowels

24 Thai vowels

Tongue Height \ Tongue Advancement	Front	Central	Back
Close	i, iː (อิ, อี)	ɨ, ɨː (อึ, อือ)	u, uː (อุ, อู)
Mid	e, eː (เอะ, เอ)	ɜ, ɜː (เออะ, เออ)	o, oː (โอะ, โอ)
Open	æ, æː (แอะ, แอ)	a, aː (อะ, อา)	ɔ, ɔː (เอาะ, ออ)
Diphthongs	ia, iːa (เอียะ เอีย)	ɨa, ɨːa (เอือะ, อัวะ)	ua, uːa (อัวะ, อัว)

2.2 Mapping from Japanese phonemes to Thai phonemes

Japanese phonemes mainly have five vowels and 21 consonants for speech recognition, as defined by the Advanced Telecommunications Research Institute International (ATR). We have mapped the Japanese phonemes and Thai phonemes on the International Phonetic Alphabet (IPA) table, as shown in Figure 1 and Table 4.

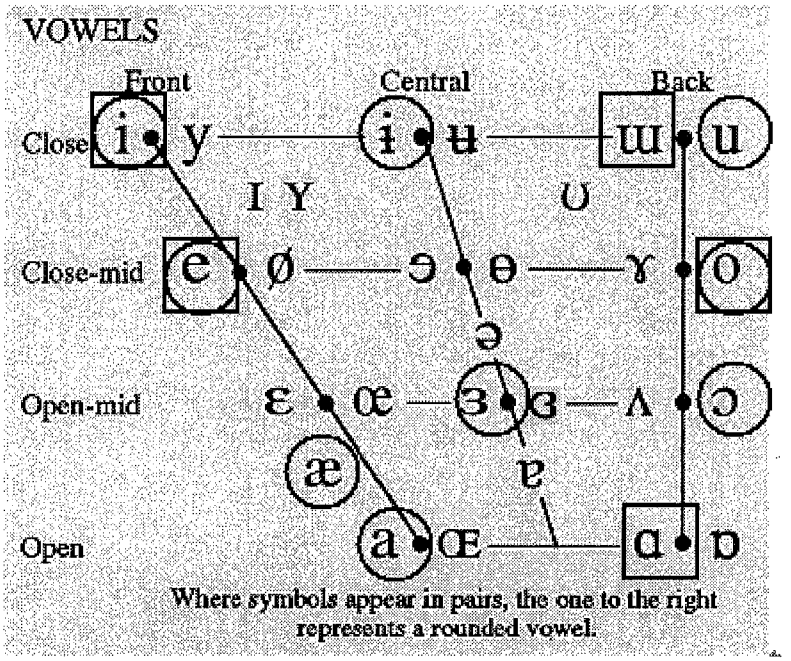

Figure 1. Japanese and Thai vowels on IPA

From Figure 1 and Table 4, the circles show the Thai phoneme positions, and the squares show the Japanese phoneme positions on the IPA table. Some Thai and Japanese phonemes have the same positions. That means a few phonemes should be similar to each other in occurrence and manner. It was easy to map from the Japanese phonemes to the Thai phonemes as shown in the highlight color of Table 5. On the other hand, we tried to match Japanese phonemes with Thai phonemes by occurence, manner, and the acoustic of the phonemes as much as possible.

This mapping matched only single Thai and Japanese phonemes as shown in Table 5. This conversion table did not consider the information from the Thai language tones. In the case of long vowels, double vowels, and double consonants, they were replaced with two single phonemes. For example, the long vowel of the Thai phoneme (/a:/) was defined by two Japanese vowels (/a/ + /a/). And the Thai double consonant (/kl/) was defined by the Japanese consonants (/k/ +/l/).

Table 4. Japanese and Thai consonants on IPA

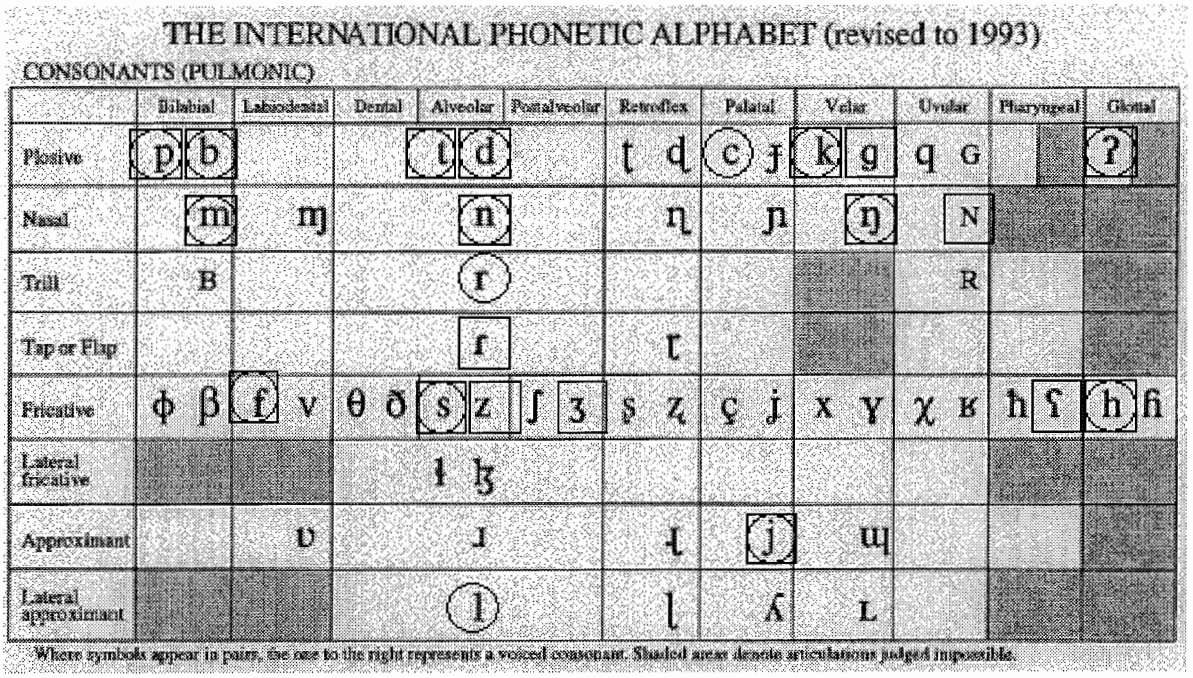

Table 5. Conversion Table

Phone	TH	JA	Phone	TH	JA
Vowels	i	i	Consonants	t^h	t
	ɨ	i		c^h	tʃ
	u	ɯ		k^h	k
	e	e		b	b
	ɜ	e		d	d
	o	o		m	m
	æ	e		n	n
	a	ɑ		ŋ	ŋ
	ɔ	o		l	ɾ
Consonants	p	p		ɾ	ɾ
	t	t		f	f
	c	tʃ		s	s
	k	k		h	h
	ʔ	ʔ		w	w
	p^h	p		j	j

3 Experiments and Results

In this research, all experiments were performed using ATRSPREC and a Japanese database constructed by ATR. The Thai database was collected by NECTEC. Context-independent and context-dependent acoustic models were used in these experiments.

3.1 Experimental Conditions

All utterances used in these experiments were spoken by native Thai speakers using the Thai middle dialect.

3.1.1 Training

In these experiments, 12-order cepstral coefficients, 12-order delta-cepstral coefficients, and delta-powers were extracted for features.

The Japanese acoustic models were trained from 503 Japanese phonetic balanced sentences, spoken by 168 males and 232 females, with 19,948 utterances. Each phoneme was modeled with context-independent HMMs with 32 Guassian mixtures and 81 states, and context-dependent acoustic models (HMnet) [M. Ostendorf, 1997] with 5 Guassian mixtures and 1,403 states. These acoustic models were gender models. Only a silence model was trained to Thai speech data because the recording environment between the Japanese speech database and the Thai speech database was very different. After we converted acoustic models, we got Thai acoustic models with 93 states for context-independent models and with 1,935 states for context-dependent models. This included three states for silence models.

We converted 26 Japanense phonemes to 30 Thai phonemes using the conversion table shown in Table 2. Thai acoustic models were trained by using 11,000 utterances from 5,000 Thai isolated-words of 11 Thai native speakers (6 females and 5 males). The training data required 3 hours and 53 minutes.

3.1.2 Evaluation

Two lexicons were used in our experiments: 50 isolated words and 250 isolated words. A male (M23) and a female (F03) were evaluated for the 50 isolated words. Two males (M05, M23) and two females (F03, F11) were evaluated for the 250 isolated words.

3.2 Initial Model Experiment

We evaluated the initial models that were converted from the Japanese acoustic models. Context-independent (CI) and context- dependent (CD) acoustic models were used in the experiments. The results indicated the performance of these initial models. As shown in Table 6, the context-dependent models exhibited a better performance than the context-independent models.

Table 6. Results of initial models

Lexicons	Speaker	Word Accuracy (%)	
		CI AM	CD AM
50 words	F03	58.00	58.00
	M23	50.00	72.00
	Average	54.00	65.00
250 words	F03	29.41	31.37
	F11	32.00	36.00
	M05	27.91	25.58
	M23	25.49	35.29
	Average	28.70	32.06

For some phonemes, these initial models can precisely segment utterances into phonemes. Therefore, these models can be trained to a certain extent.

3.3 Trained Model Experiment

The initial models were trained with 3 hours and 53 minutes of Thai utterances. Table 7 shows the results of this experiment.

These results indicate an improvement over the performance of the trained models. The improved performances of the context-independent models, compared with the previous experiments, were 17% and 30% for 50 and 250 isolated words. For example, the average word accuracy for the context-independent models with 50 isolated words was 54% when using the initial models and 71% when using the trained models, for a 17% increase. In the case of the context-dependent models with 50 and 250 isolated words, there was a 24% and 37% improvement. For example, the average word accuracy for the context-dependent models with 50 isolated words was 65% when using the initial models and 89% when using the trained models. That is a 24% increase.

Table 7. Results of trained models with 11 speakers

Lexicons	Speaker	Word Accuracy (%)	
		CI AM	CD AM
50 words	F03	76.00	88.00
	M23	66.00	90.00
	Average	71.00	89.00
250 words	F03	62.75	68.63
	F11	64.00	84.00
	M05	58.14	65.12
	M23	49.02	58.82
	Average	58.48	69.14

We add the training data to 23 speakers (8 females and 15 males). Therefore, the initial models were trained again. The results of this training data show in Table 8. The results of trained models with 23 speakers are small different from the results of 11 speakers. The results of the male speaker in the last models are lower than the trained models with 11 speakers such as, 54% was decreasing from 66% of M23 with context-independent models and 50 words lexicon, and 55.18% was decreasing from 58.14% of M05 with context-independent models and 250 words lexicon.

Table 8. Results of trained models with 23 speakers

Lexicons	Speaker	Word Accuracy (%)	
		CI AM	CD AM
50 words	F03	76.00	94.00
	M23	54.00	86.00
	Average	65.00	90.00
250 words	F03	64.71	76.47
	F11	62.00	82.00
	M05	55.81	74.41
	M23	45.10	52.94
	Average	56.90	71.46

When mapping the Japanese phonemes to Thai phonemes, it was difficult to map some of the Thai phonemes, such as "ɨ", "ɜ", "l", and "ɾ". They are very different from Japanese phonemes. However, we tried to map some Japanese phonemes to them. Therefore, the results of the initial models were much lower than the results of models

that were trained by a small amount of Thai speech data.

3.4 Model by Adapted Label Training Experiment

This experiment, we have adapted the initial models with 23 speakers with MAP estimation. After we got the speaker-dependent models of each speaker, we used those models to label the utterances of each speaker (23 speakers). Therefore, the label data has used for training in this experiment. The results show in Table 9.

Table 9. Results of trained models with label data

Lexicons	Speaker	Word Accuracy (%)	
		CI AM	CD AM
50 words	F03	74.00	86.00
	M23	40.00	80.00
	Average	57.00	83.00
250 words	F03	56.86	70.59
	F11	58.00	84.00
	M05	58.14	69.77
	M23	25.49	52.94
	Average	49.62	69.32

The results of this experiment are a lower than the results of trained models such as, the average of 50 words lexicon in context-independent models are 65% for trained models and 57% for adapted label training.

4 Conclusion

It takes a lot of time to correct a large amount of speech data and label them. We had only a small amount of Thai speech data, which was not alignment data. Therefore, we made initial models by mapping Japanese acoustic models, and trained the initial models by using a small amount of data. IPA tables were used for this mapping.

From our experiments, the context-dependent models of Japanese acoustic models that were converted to Thai acoustic models performed better than context-independent models in both experiments. Using Japanese acoustic models to convert Thai acoustic models was a good assumption as an indication in the first experiment.

The trained results of 11 speakers exhibited improvements in the word accuracy

rate under all conditions of the experiments, from about 17% to 40% when compared with the initial model experiments. The word accuracy rates of the 50 isolated words were better than those of the 250 isolated words, rising from about 13% to 33%.

There are small difference between the trained results with 23 speakers and the trained results with 11 speakers, from about 0% to 6% for female speakers and from about 3% to 12% for male speakers. The results of adapt label training in this research are lower than the results of trained models because we did not change some parameters of MAP estimation.

Acknowledgment

This research was supported in part by Asia Pacific Telecommunity (HRD Programme for Exchange of ICT Researchers and Engineers).

References

[Ahkuputra, 1997] Ahkuputra, V., Jitapunkul, S., Pornsukchandra, W., Luksaneeyanawin, S. 1997. "A Speaker-Independent Thai Polysyllabic Word Recognition Using Hidden Markov Model", *Proceedings of the 1997 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing*, 593-599

[Ekkarit, 2000] Ekkarit Maneenoi, Somchai Jitapunkul, Visarut Ahkuputra, Umavasee Thathong, and Boonchai Thampanitchawong. 2000. "Thai monophthong recognition using continuous density hidden Markov model and LPC Cepstral coefficients", The proceeding of the *6th International Conference on Spoken Language Processing*, IV: 620-623

[Tanja, 2001] Tanja Schultz, Alex Waibel. 2001. "Language-independent and language-adaptive acoustic modeling for speech recognition", *Speech Communication*, 35(2001): 31-51

[Uebler, 2001] Ulla Uebler, "Multilingual speech recognition in seven languages", *Speech Communication*, 35(2001): 53-69

[Sudaporn, 1993] Sudaporn Luksaneeyanawin, 1993. "Speech Computing and Speech Technology in Thailand", *Proceeding of the Symposim on Natural Language Proceeding in Thailand*, 276-321

[M. Ostendorf, 1997] M. Ostendorf, H. Singer. 1997. "HMM topology design using maximum likelihood successive state splitting", Computer Speech and Language, 11: 17-41

Appendix

A. Thai Speech Database

The Thai speech database that is used at the Advanced Telecommunications Research International Institute (ATR) consists of three principle sets: the isolated-words set, phonetic balanced sentences, and hotel reservation transcription (HRT). This database was collected by Information Research and Development Division, National Electronics and Computer Technology Center, Thailand (NECTEC). The details of these sets are as follows:

(1) The isolated-word set contains 5,000 daily words, 640 phonetic balanced words, and 131 extra words for hotel reservation transcription.
(2) A 390 phonetic balanced sentence set.
(3) Each database has utterances for 50 dialogues of hotel reservation transcriptions.

The reading speech data was spoken by 40 Thai native speakers (20 males and 20 females). They spoke in the middle dialect of the Thai language. All utterances were recorded in a quasi-quiet room.

The explanation of Thai Speech database is showing below.

- DB1 is isolated-word set.
- DB2 is Phonetic Balanced sentence (PB) set.
- DB3 is 50 dialogues of Hotel Reservation Transcriptions (HRT) set.

A.1 Isolated-word set (DB1)

This set has three subsets : 5,000 isolated words, Phonetic Balanced words (PB-word), and extra-word. Therefore, this set has three subsets : D0-D4 for 5,000 isolated-word set, D5 for PB word set and extra word set. Each D0 to D4 contains 1,000 Thai isolated words. Each speaker has spoken only one set (1,000 words) of 5,000 isolated words, PB-word set, and extra-word set.

Subdirectory : /DB_local/THAI/DB1

Example of file name and extension :

Subdirectory : /DB_local/THAI/DB1/F03/D2/WAV

File : F03_1_0001.16k

'F03' is speaker's name.

'1' is identification number of subset such as '1' is isolated words, 'b' is PB words, and 'e' is extra words, and 'n' is digits.

'0001' is identification number of utterance.

'16k' stand for the sampling rate of utterance (16,000 bits per second).

A.2 Phonetic balanced sentence set (DB2)

The PB set contains 390 Thai sentences that were collected from newspaper and journal. Each speaker was spoken all of these sentences.

Subdirectory : /DB_local/THAI/DB2

Example of file name and extension :

Subdirectory : /DB_local/THAI/DB2/F03/SD/WAV

File : F03_SD_001.16k

'F03' is speaker's name.

'SD' is identification number of subset.

'001' is identification number of sentence (001-390).

'16k' stand for the sampling rate of utterance (16,000 bits per second).

A.3 Hotel reservation transcription set (DB3)

The 50 dialogues of Hotel Reservation Transcription (HRT) are spoken in this database set. Each speaker was spoken five dialogues of HRT.

Subdirectory : /DB_local/THAI/DB3

Example of file name and extension :

Enviroment files

Subdirectory : /DB_local/THAI/DB3/F03/ENV/TAS22031

File : F03TAS22031.env

This subdirectory contains the enviroment of recording of each dialogues.

'F03' is speaker's name.

'TAS' is dialogue's name.

'22031' is dialogue's identification number.

'env' stand for enviroment.

Wave files

Subdirectory : /DB_local/THAI/DB3/F03/WAV/THAI/TAS22031

File : F03_TAS220310010b.16k

'F03' is speaker's name.

'TAS' is dialogue's name.

'22031' is dialogue's identification number.

'0010' is identification number of sentence in dialogue

'b' is clerk and 'a' is customer

'16k' stand for the sampling rate of utterance (16,000 bits per second).

B. Software

We developed some program to converse acoustic models (AM) from Japanese acoustic models to Thai acoustic models, to make the transcription file and Thai isolated-word lexicon.

B.1 Converse Acoustic Models

This program converses from Japanese acoustic models to initial acoustic models for Thai acoustic models. Using conversion table of Japanese phonemes and Thai phonemes as show in Table B-1. The usage of this program shows in the below.

USAGE: ./conv_model [conver_table] [source_HMnet] [output_HMnet]

Where: 'conv_model' is the command name of this program
'conver_table' is conversion file that using mapping from Japanese phonemes to Thai phonemes. This is text file same as Table B-1.
'source_HMnet' is the source acoustic models (Japanese acoustic models).
'output_HMnet' is the output acoustic models (Thai acoustic models).

For example, /home/ksawit/ThaiTool/conv_model/conv_model convert.tbl AM.F.5.1400 AM.THAI.F.5.1400

Table B-1. Conversion table

This file is conversion table file for using converse JA AM to TH AM

TH -> JA

conv_table

-	-
i	i
v	i
u	u
e	e
q	e
x	e
a	a
@	o
o	o
p	p
t	t
c	ch
k	k
?	q
ph	p
th	t
ch	ch
kh	k
b	b
d	d
m	m
n	n
ng	ng
l	r
r	r
f	f
s	s
h	h
w	w
j	j

B.2 Making Transcription File

This program makes the transcription file of each Thai utterance. The format of transcription file (TRS file), the input text file, and usage of this program are showed in the following.

USAGE: ./make_trs [Transcription File] [Raw File]

Where 'make_trs' is the command name of this program.
'Transcription File' is the input text file.
'Raw File' is the utterance files (*.16k).

For example, /home/ksawit/ThaiTools/make_trs/make_trs word5k_2_trn *.16k

Notice: Be careful to check the number of utterances (*.16k) and line of input text file (word5k_2_trn) should be equal amount.

Sample of transcription file (word5k_2_trn)
kot1ma:j4
k@:2ta:m0
krong0
k@:0ra3ni:0
kr@:p1
kra1sip3
kra1thop3kra1thv:an0
kra1nan3
kra:p3fik1
krung0rat3ta1na3ko:0sin4
ka1ru3na:0
klom0
klom0klv:n0
kon0la3wi3thi:0
klan1kr@:ng0
kla:ng0
kla:ng0khon0
...

Sample of output file (TRS file)

TRS file name	Transcription
F03_1_0001.TRS	0.00 -,k,o,t,m,a,a,j,- 1260.19
F03_1_0002.TRS	0.00 -,k,@,@,t,a,a,m,- 1126.19
F03_1_0003.TRS	0.00 -,k,r,o,ng,- 974.31
F03_1_0004.TRS	0.00 -,k,@,@,r,a,n,i,i,- 1300.56
F03_1_0005.TRS	0.00 -,k,r,@,@,p,- 832.00
F03_1_0006.TRS	0.00 -,k,r,a,s,i,p,- 876.25
F03_1_0007.TRS	0.00 -,k,r,a,th,o,p,k,r,a,th,v,v,a,n,- 1383.25
F03_1_0008.TRS	0.00 -,k,r,a,n,a,n,- 1053.25
F03_1_0009.TRS	0.00 -,k,@,@,r,a,n,i,i,- 1300.56
F03_1_0010.TRS	0.00 -,k,r,a,a,p,f,i,k,- 990.12

Definition of transcription

For example, 0.00 -,k,o,t,m,a,a,j,- 1260.19

'0.00' is the starting time

'-,k,o,t,m,a,a,j,-' are the sequence of phonemes. When '-' is silence.

'1260.19' is the ending time.

B.3 Making Thai Lexicon

In this research, we develop Thai speech recognition for isolated words. Therefore, we make the lexicon for Thai isolated words. We use the isolated-word database (DB1). From our experiment, we have two lexicons such as 50 isolated words and 250 isolated words. They were chosen from 5,000 isolated words. We have to prepare the input text file before making the lexicon.

USAGE: ./make_lex [type] [output_filename] [input file]

Where 'make_lex' is the command name of this program.

'type' is the lexicon type such as, '-nt' is no tone information, '-t' is the lexicon with tone information, '-n' is the Thai baseline lexicon, and '-s' is used for making syllable table only. In this research, '-n' is used to make the Thai lexicon.

'output filename' is the lexicon name.

'input filename' is the text file for making lexicon. It can use many files to make the lexicon (*.txt).

For example, /home/ksawit/ThaiTools/make_lex/make_lex -n Thai_lexicon word.txt

Sample of the input text file (word.txt)

เก็ง keng0
เก่ง keng1
เก็งกำไร keng0kam0raj0
เกณฑ์ ke:n0
เกตุ ke:t1
เก็บ kep1
เกม ke:m0
เกย์ ke:0
เกรง kre:ng0
เกรด kre:t1
เกราะ kr@1
เกรียงไกร kri:ang0kraj0
เกล็ดเลือด klet1lv:at2
เกล้า kla:w2
...

Definition of input text file

Example: เก็งกำไร keng0kam0raj0

'เก็งกำไร' is Thai word with Thai characters.

'keng0kam0raj0' is Thai syllable sequence. The digits represent Thai tones. In this case, they have three syllable, 'keng0', 'kam0', and 'raj0'

Sample of Thai lexicon for isolated words (Thai Lexicon)

```
6 [UTT_END] -    # UTT_END
5 [UTT_START] -    # UTT_START
10000 [เก็ง] k e ng 0 {|-}    # เก็ง|เก็ง|||
10001 [เก่ง] k e ng 1 {|-}    # เก่ง|เก่ง|||
10002 [เก็งกำไร] k e ng 0 k a m 0 r a j 0 {|-}    # เก็งกำไร|เก็งกำไร|||
10003 [เกณฑ์] k e: n 0 {|-}    # เกณฑ์|เกณฑ์|||
10004 [เกตุ] k e: t 1 {|-}    # เกตุ|เกตุ|||
10005 [เก็บ] k e p 1 {|-}    # เก็บ|เก็บ|||
10006 [เกม] k e: m 0 {|-}    # เกม|เกม|||
10007 [เกย์] k e: 0 {|-}    # เกย์|เกย์|||
10008 [เกรง] kr e: ng 0 {|-}    # เกรง|เกรง|||
10009 [เกรด] kr e: t 1 {|-}    # เกรด|เกรด|||
10010 [เกราะ] kr @ 1 {|-}    # เกราะ|เกราะ|||
10011 [เกรียงไกร] kr i:a ng 0 kr a j 0 {|-}    # เกรียงไกร|เกรียงไกร|||
10012 [เกล็ดเลือด] kl e t 1 l v:a t 2 {|-}    # เกล็ตเลือต|เกล็ดเลือด|||
10013 [เกล้า] kl a: w 2 {|-}    # เกล้า|เกล้า|||
```

Definition of Thai lexicon

It contains with word ID, Thai character and phoneme sequence of isolated-word. As the following:

[ID number] [Thai character] [Phonemes {|-}] [# Comment]

The comment of Thai isolated-word lexicon is same as Thai character.