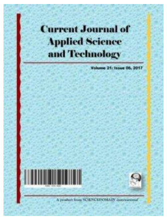

Current Journal of Applied Science and Technology

29(4): 1-10, 2018; Article no.CJAST.40568
ISSN: 2457-1024
(Past name: British Journal of Applied Science & Technology, Past ISSN: 2231-0843, NLM ID: 101664541)

Comparison of Syllable and Phoneme Modelling of Agglutinative Tamil Isolated Words in Speech Recognition

Ibralebbe Mohamed Kalith[1*], David Asirvatham[2], Ali Khatibi[3] and Samantha Thelijjagoda[4]

[1]Department of Mathematical Science, Faculty of Applied Sciences, South Eastern University of Sri Lanka, Sri Lanka.
[2]School of Computing and IT, Taylor's University, Subang Jaya, Selangor, Malaysia.
[3]Graduate School of Management, Management and Science University, Malaysia.
[4]Department of Information Systems Engineering, Sri Lanka Institute of Information Technology, Sri Lanka.

Authors' contributions

This work was carried out in collaboration between all authors. Author IMK designed the study, data collection, performed the statistical analysis and wrote the first draft of the manuscript. Authors DA and AK managed the analyses of the study. Author ST managed the literature searches. All authors read and approved the final manuscript

Article Information

DOI: 10.9734/CJAST/2018/40568
Editor(s):
(1) Dr. Orlando Manuel da Costa Gomes, Professor of Economics, Lisbon Accounting and Business School (ISCAL), Lisbon Polytechnic Institute, Portugal.
(2) Dr. Harry E. Ruda, Professor, Stan Meek Chair Professor in Nanotechnology, University of Toronto, Director, Centre for Advanced Nanotechnology, University of Toronto, Canada.
Reviewers:
(1) Mete Yaganoglu, Ataturk University, Turkey.
(2) A. Akila, Vels Institute of Science, Technology & Advanced Studies, Deemed University, India.
(3) Iroju Olaronke, Adeyemi College of Education, Nigeria.
Complete Peer review History: http://www.sciencedomain.org/review-history/26597

Original Research Article

Received 06 March 2018
Accepted 20 September 2018
Published 10 October 2018

ABSTRACT

Aim: In this paper, the emphasis was on improving the automatic speech recognition of Tamil speech by applying syllable and phoneme as a sub-word unit. Agglutinative complex words in Tamil are described by showing their element in the building of the sub-word such as the syllable and the phonemes. This present study used the Hidden Markov Model (HMM) based speech recognition system that was created using CMU Sphinx speech recognition toolkit. An effective consonant-vowel six-segment (CVS-6) algorithm was designed to syllabification of the Tamil

**Corresponding author: E-mail: imkalith@seu.ac.lk;*

isolated words and experimentally investigated its speech recognition accuracy. The database used in this study was designed using a maximum of 160 isolated words, representing 430 syllables and 216 unique syllables.

Results: Through the experiment, the syllable-based model achieved a mean recognition rate of 93.84 (standard deviation, 5.02) compared to 91.37 (standard deviation, 6.26) achieved by a phoneme-based model.

Conclusion: It was concluded from this research that the syllable-based model using the CVS-6 algorithm is a good choice and can be used in the development of sub-word modelling of isolated words, which is an effective sub-word modelling of medium and large vocabulary ASR Tamil language.

Keywords: Hidden Markov model; phoneme-based; syllable-based; ASR; Tamil language; Consonants Vowel Segmentation-6 Algorithm.

1. INTRODUCTION

Automatic Speech Recognition ASR has developed years of progress, including the successful introduction of commercial speech-based applications. Speech recognition requires that the speech waveform to be segmented into basic acoustic units. The word is the ideal and the usual unit of speech because ultimately it is the word that one tries to recognise. The dissemination of words in Tamil differs from other languages. It shows an agglutinative characteristic that makes up a large number of words as well as the size of the words. However, there are still unresolved problems in the speech recognition research community, one of which is how to design the internal structure of the words (sub-word). In most languages, the word is represented by the phonemes and the syllable sub-word unit. For ASR research work, each of these sub-word units is designed, parametrised and trained. This research highlights the modelling of the phoneme and syllable as a sub-word unit for Isolated Word Speech Recognition (IWSR) for the Tamil language and compares the effect of sub-words modelling to the rate of recognition.

1.1 Tamil Language

Tamil is one of the Dravidian language family, spoken by about 77 million people around the world, making it the 15th most widely spoken language [1]. It has been known as a common language in southern India, Sri Lanka, Malaysia and Singapore and has been widely spoken in the northern and eastern parts of Sri Lanka, where it has been recognised as an official language. There is some dialectic variation in Sri Lanka, with different speech patterns in rural areas of Nuwara Eliya, the northern regions of Jaffna and in the eastern parts.

1.2 Tamil Phonology

The phoneme is a small and fundamental unit of a speech. Tamil is a phonemic script. Tamil character sets include 12 vowels (V), 18 consonants (C), 216 compounds (VC) and a secondary character called Aythm. There are 247 characters in the standard Tamil alphabet. An allophone is an alternative sound to the standard phoneme. Including all allophones, the Tamil language has nearly 100 sounds [1]. The Tamil language phonology is characterised by the special feature existing in retroflex consonants and multiple rhotics. Special features in Tamil phonology include retroflex consonants and multiple rhotics [2].

1.2.1 Vowels and consonants

Tamil uses 12 vowel sounds, as shown in Table 1 [3]. Depending on the relative position, the expressions of the vowels are different. There are 18 basic consonant characters in the Tamil language. Considering these allophones [3], Tamil has approximately 25 distinct consonant phonemes (Table 2). The nasal consonants ங ஞ ண ந ம and ன are pronounced differently depending on the situation in which they occur, and this is illustrated in Table 2.

1.2.2 Problem of isolated words in Tamil ASR

Tamil word pattern explained by one of its parameters as a size of the word. Literacy in Tamil, the word is the basic unit of construction of long sentences. Tamil is a language intertwined with agglutinative inflectional morphology and morphophonology, where words are formed by combining a series of phonemes [4]. This means that the word may consist of many phonemes. This makes the challenges of defining complex word boundaries in ASR to the

Tamil language. As a result, a large data set is needed to train the audio model. One way to address this is to implement effective sub-word modelling. In this paper, the success of sub-word modelling, using phonemes and syllables was investigated in recognition of complex Tamil words. These complex words are frequent in the huge Tamil corpora. Agglutinative inflectional morphology, means that words combined the same root word with different prefixes and suffixes [4]. An example the verbs in Tamil form inflectional variations (Table 3). Two meaningful words are combined to form new words called morpho-phonology, where words are modified by deleting, inserting, or substitution of letters at the word boundary to form a new word, widely used in Tamil [4] (see Table 4).

1.3 Related Work

In the late 1970's, Itakura [5] performed speech recognition by calculating the prediction residual for 200 words. Myers et al. [6] proposed a level building Dynamic Time Wrapping (DTW) algorithm for connecting word recognition. Wilpon et al. [7] developed an automatic English word recognition using HMM. Many ASR research works in Indian languages have also focused on a syllable as a unit of speech. Thangarajan et al. [8] proposed a syllable-based continuous speech recognition in Tamil, attempting to solve the problem of "out of vocabulary" in Tamil by using the prosodic syllable as a sub-word unit for acoustic training. The proposed syllable-based model outperforms by 10% compared to the baseline triphone model. Radha et al. [9] proposed an isolated word speech recognition system for Tamil spoken language using backpropagation neural network based on the LPCC features that achieved a Mean Square Error (MSE) of 0.00515 / 0.01. Ganapathiraju et al. [10] presented the first successful, robust LVCSR system that used a syllable level acoustic unit in telephone- based spontaneous speech. The paper is concluded with a syllable-based system that will perform better than the existing tri-phone system. Sundarapandian et al. [11] suggested syllable-based continuous speech recognition of the Tamil language. The Modified Group Delay Function with Gammatone Cepstral Coefficients (MGDF-GWCC) are used to extract the feature from the acoustic signal, and these features are efficiently classified by Deep Neural Network with HMM (DNN-HMM). The IIIT-H Indic database was used to measure the system performance. Sakthi Vel [12] investigated the supra-segmental feature in the continuous speech of the Tamil language. The audio segmentation based on the fundamental frequency measurement and energy level functions provided promising results. In this research, manually segmenting and annotating the isolated word by analysing short-term energy spectrum analysis of voice signals were used as shown in Fig. 1. By analysing the formant frequency of the spectrum syllable boundary can be identified and labelled manually.

Table 1. Position and symbol of Tamil phonology

S. no	Vowels	Example
1	அ	அம்மா
2	ஆ	ஆங்கிலம்
3	இ	இன்பம்
4	ஈ	ஈட்டி
5	உ	உரிமை
6	ஊ	ஊக்கம்
7	எ	எண்பது
8	ஏ	ஏழ்மை
9	ஐ	ஐம்பது
10	ஒ	ஒற்றுமை
11	ஓ	ஓலை
12	ஔ	ஔடதம்

Table 2. Tamil consonants

S. no	Consonants	IPA	Example
1	க	k	கலை
2	க	g	அங்கே
3	க	h	பகல்
4	ங	ŋ	மங்கை
5	ச	tʃ	அச்சம்
6	ச	s	சாலை
7	ச	j	கஞ்சி
8	ஞ	ɲ	ஞானம்
9	ட	ʈ	தொடர்பு
10	ட	ɖ	பட்டம்
11	ண	n	வணக்கம்
12	த	t̪	பாதம்
13	த	d̪	புத்தகம்
14	ந	ɳ	நட்பு
15	ப	p	பாதம்
16	ப	b	அப்பம்
17	ம	m	மருத்துவம்
18	ய	j	யாசகம்
19	ர	R	அரசன்
20	ல	l	இலங்கை
21	வ	v	வடை
22	ழ	L	அழகு
23	ள	ɭ	குளவி
24	ற	r	குறள்
25	ன	N	கானகம்

Table 3. Example of agglutinative (inflectional variations) word formation for tense markings

No	English		Tamil			
	Noun	Verb	Noun	Verb with tenses		
				Past	Present	Future
1	I	Went	நான்	சென்றேன்	செல்கிறேன்	செல்வேன்
2	You	Went	நீ	சென்றாய்	செல்கிறாய்	செல்வாய்
3	They	Went	அவர்கள்	சென்றார்கள்	செல்கிறார்கள்	செல்வார்கள்
4	He	Went	அவன்	சென்றார்	செல்கிறான்	செல்வான்
5	She	Went	அவள்	சென்றாள்	செல்கிறாள்	செல்வாள்

Table 4. Examples of morpho-phonological word formation

Example	Two words	Combine to form one new word
Deletion	கருமை + விழி	கருவிழி
Insertion	மாலை + பொழுது	மாலைப்பொழுது
Substitution	பொன் + கொல்லன்	பொற்கொல்லன்

2. MATERIALS AND METHODS

The ASR system for recognising the Tamil language speech that has been used in all of the experiments on the ASR architecture has been described in Fig. 2. The system is based on HMMs using CMU Sphinx tools. The Sphinx is a portable toolkit for building and manipulating HMMs [13]. Sphinx tools are primarily used in speech recognition research developed by the speech group at Carnegie Mellon University in the United States of America. The data was sampled at 16 kHz. Frame features were extracted to reduce the amount of information in the input signal. 39 Mel-frequency cepstral coefficients MFCCs were extracted at a frame rate of 10 ms using a 25 ms Hamming window. Satori et al. [14] and Abushariah et al. [15] investigated ASR system for the Arabic language using CMU Sphinx tools to build a speech recogniser and a language model for the Arabic language.

2.1 Text Segmentation into Phonemes and Syllable

In the Tamil language, continuous speech is formed by combining the words into sentences. Most of the speech recognition system uses phonemes as a sub-word unit. This research used 37 phonemes to represent the entire database in isolated words (Tables 1 and 2). These phonemes were used in the tri-phone based acoustic training model. The HMM applied Gaussian mixture observation densities to solve the acoustic variation between the different existences of the same grapheme. In the different context, the same grapheme may be pronounced differently in words, as articulation is influenced by adjacent phones (Table 2). This explains the acoustic variability in the basic phonetic units is large and not well thought out for many languages.

Bahl et al. [16] described that, in such contexts, a word-based model performed significantly better than a phoneme-based model in dynamic time warping (DTW). Therefore, a phoneme-based model may not always be optimal when addressing highly context-sensitive language [17]. A syllable is an alternative sub-word unit that may be smaller than a word or larger than a single-phoneme. A syllable usually extends to several phonemes, making them easier to recognise in a speech signal [18]. However, defining a syllable is challenging in the Tamil language, as different languages have different rules of segmentation. The key is to define the syllable boundaries to form a syllable from the speech signal. In the short energy spectrum analysis, the formant value can be used to identify the boundary approximately [4]. This study introduced a simple solution for the segment of the Tamil text into syllable a consonant-vowel six-segment (CVS-6) algorithm for syllabification. In the short energy spectrum analysis (Fig. 1), formant value of the energy spectrum is shown to be almost co-extensive with a syllable segment represented by the CVS-6 algorithm (Fig. 3). Whole isolated words were used in this research segmented by this algorithm described in detail (Fig. 4)

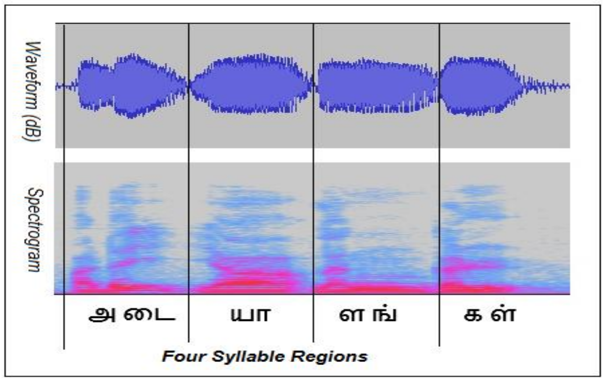

Fig. 1. Waveform and spectrogram analysis of complex word அடையாளங்கள் showing four syllable regions

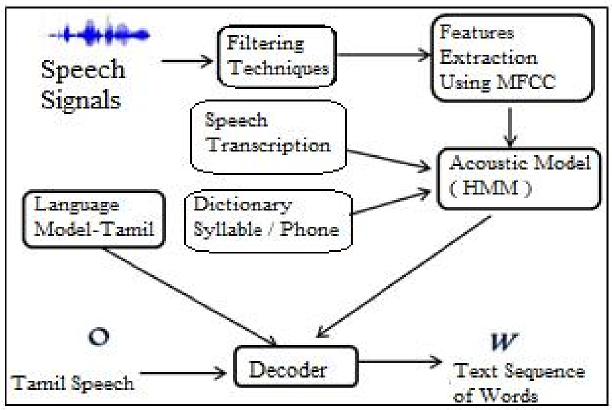

Fig. 2. ASR system for Tamil isolated word [4]

Count word length = n, Position of letter = P
For i = 1 to n do

- If P_i = V AND P_{i+1} = CV, Then //Condition1
  Check P_{i+2}, //Condition2
  If P_{i+2} = C, Then //Condition3
  P_{i+1}, P_{i+2} is a Syllable *[CV.C]* *//Statement1*
  Else
  P_i, P_{i+1} is a Syllable *[V.CV]* *//Statement2*
  Else If P_i = V AND P_{i+1} = C, Then //Condition4
  P_i, P_{i+1} is a Syllable *[V.C]* *//Statement3*
  Else
  P_i is a Syllable *[V]* *//Statement4*
- If P_i = CV AND P_{i+1} = C //Condition5
  Check P_{i+2}, //Condition6
  If P_{i+2} = C, Then //Condition7
  P_i, P_{i+1}, P_{i+2} is a Syllable *[CV.C.C]* *//Statement5*
  Else
  P_i, P_{i+1} is a Syllable *[CV.C]* *//Statement6*
  Else
  P_i is a Syllable *[CV]* *//Statement7*

Fig. 3. Proposed CVS-6 syllabification algorithm

Example	அ	டை	யா	ள	ங்	க	ள்
	V	CV	CV	CV	C	CV	C
	P_i	P_{i+1}	P_{i+2}	……………….			

Pass1 — Condition1 – True, Go to Condition2 – False, Execute Statement1, **[அ டை]**

Pass2 — Condition1 and 5 – False, Execute Statement7, **[யா]**

Pass3 — Condition1- False, Condition 5 and 6 -True Execute Statement6, **[ளங்]**

Pass4 — Condition1- False, Condition 5 and 6 -True Execute Statement6, **[கள்]**

Fig. 4. Explains the text segmentation of Tamil isolated complex word (அடையாளங்கள்) using CVS-6 segmentation algorithm

Table 5. Examples of Tamil words and its syllable constituent segmented by CVS-6 algorithm

Words	Syllable (V, CV, V.CV, CV.C, V.C, CV.C.C)	Syllable Count
அண்ணன்	அண், ணன்	2
சகோதரன்	ச, கோ, த, ரன்	4
பெரியம்மா	பெ, ரி, யம், மா	4
வெங்காயம்	வெங், கா, யம்	3
மருத்துவர்	ம, ருத், து, வர்	4
தென்கிழக்கு	தென், கி, ழக், கு,	4
நிறுத்தப்பட்டது	நி, றுத், தப், பட், ட, து	6
குடும்பத்தினர்	கு, டும், பத், தி, னர்,	5
ஆரோக்கியம்	ஆ, ரோக், கி, யம்	4
உத்தரவிட்டார்	உத், த, ர, விட், டார்	5
அடையாளங்கள்	அடை, யா, ளங், கள்	4

2.2 Performance Evaluation

Recognition rate given by Word Error Rate (WER) of the trained HMM is defined as follows:

$$WER = \frac{S + D + I}{N}$$

Where,

S is the number of substitutions,
D is the number of deletions,
I is the number of the insertions,
N is the number of words in the reference.

OR word accuracy (*WAcc*) is calculated as

$$WAcc = \frac{N-S-D-I}{N} = 1- WER$$

2.3 Database

The isolated word form was used to train the acoustic model to recognise isolated words. To test the objective of the sub-word effect in recognising isolated word speech in Tamil, a special database was built. The data were collected and analysed using the following steps. About 160 isolated words from the Tamil language were selected to train the acoustic speech model. These were collected from primary school textbooks, newspapers, articles, novels, and magazines.

The collected words were then organised into a batch system from Test1 to Test 8 (Table 7) for training and testing the acoustic model. These words were uttered by a single native speaker (age 32 years) and recorded. Each word was clearly articulated. The recording was performed under the conditions described in Table 6. Audacity software tools were used for recording and editing, as it is a powerful and sophisticated tool that allows speech segmentation and noise removal while maintaining speech parameters. The speaker recorded 160 isolated words. Each word was pronounced three times, and the most obvious version was chosen. The isolated word speech data were used in Experiments 1 and 2. The two sub-word methods were compared to assess their suitability for Tamil automatic speech recognition system development.

Table 6. Speech files parameters

Parameter	Value
File type	mswav
File Extension	wav
Sampling Rate	16 KHz
Depth	16 bits
Mono / Stereo	Mono
Feature file extension	mfc

3. RESULTS AND DISCUSSION

To evaluate the recognition accuracy of the syllable-based and phoneme-based sub-word model, the experiments used open source CMU Sphinx speech recognition toolkit. In this paper, two experiments were designed to evaluate the performance of the sub-word modelling of speech recognition. The first was the phoneme-based modelling of IWSR and the second was the syllable-based modelling of IWSR.

3.1 Experiment 1 (Phoneme-based Modelling of IWSR)

In this experiment, the selected words were segmented into the phoneme-based sub-word model (see Table 5). The acoustic training was done using a context- dependent or triphone-based isolated word modelling. The training was conducted one by one, and each batch of training data was tested with the recorded test data, as shown in Table 7. The experimental results are listed in Table 7. The scatter diagram in Fig. 5 shows the total presentation of the phoneme-based sub-word model. The mean speech recognition rate was 91.33%, with a standard deviation of 6.25%. A strong correlation can be observed between an increase in the size of the training database and a decrease in the recognition rate. The study assumes that this reflects the distribution of complex words in the test data set.

3.2 Experiment 2 (Syllable-based Modeling IWSR)

The second experiment dealt with the recognition of agglutinative complex words in the Tamil language. The acoustic model was trained and tested, and the results were briefed in Table 7. This experiment used isolated words and spoken speech data as experiment 1 but was segmented by the CVS-6 syllable-based sub-word model. Training was conducted in a context- dependent model based on isolated word modelling. The training was conducted one by one, and each batch of training data was tested with the recorded test data, as shown in Table 7. The experimental results are listed in Table 7. The scatter diagram in Fig. 5 shows the total presentation of the syllable-based sub-word model. The mean recognition rate was 93.85%, with a standard deviation of 5.02%. As can be seen from Fig. 5, a strong correlation was again observed between an increase in the size of the training database and a decrease in the recognition rate.

At present, some work has focused on developing a syllable segmentation algorithm in splitting text and speech. The syllable segmented from the text was determined using the syllabification algorithm, and the characteristic portion of the formant spectrum value in the speech signals. Unlike the algorithm that placed the boundaries of the syllable in the speech spectrum, these syllables were similar to a syllable in the transcripts. The first method of syllable segmentation followed this study. Jasleen et al. [19] developed a medium vocabulary commonly used in English word recognition in Indian English language speech recognition system, using HMM attained 85% accuracy in words for speaker independent. Venkataramana et al. [20] developed a speech recognition system using HMM and achieved 81% accuracy in an independent speaker. In this study, the Telugu morpheme was used as a basic sub-unit of the word for acoustic modelling. A morpheme is a unit smaller than a syllable, but larger than phoneme and the phonetic rules of the Telugu language is different from the Tamil language. Akila et al. [21] used time normalisation and speech rate to improve the performance of syllable-based Tamil language speech recognition. The rate of the speech depends on the syllable-based Tamil speech recognition system. The formant value of the character used to segment the syllable. The database used in this study was 200 syllables,

Table 7. Tamil speech recognition rate of syllable-based and phoneme-based model

Decode attempt	No of training word	No of decoding word	Recognition rate % phoneme (Exp. 1)	Recognition rate % syllable (Exp. 2)
Test 1	60	15	80.00	100.00
Test 2	120	30	100.00	100.00
Test 3	180	45	86.67	93.33
Test 4	240	60	95.00	95.00
Test 5	300	75	96.00	96.00
Test 6	360	90	93.33	92.22
Test 7	420	105	91.40	87.62
Test 8	480	120	88.24	86.55
Mean			91.33	93.85
Std.Div			6.25	5.02

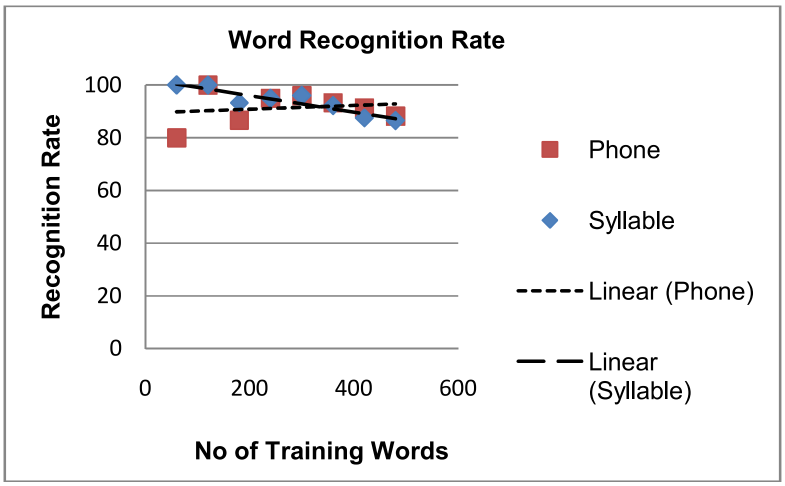

Fig. 5. Scatter diagram explains observation of phoneme and syllable-based speech recognition rates again the size of training data

and four speakers for training. The baseline showed the accuracy of 70% and the proposed model accuracy of 74%. The study was maintained an average speech rate, and the slower rate of speech determines the syllable boundaries.

4. CONCLUSION

This research compared two approaches to sub-word modelling for spoken isolated Tamil word recognition: phoneme and syllable modelling. The conventional phoneme-based model has faced challenges of recognising agglutinative complex words in isolated word speech. The study proposed a novel CVS-6 algorithm for syllable segmentation of the Tamil text corpus. From the experimental results, the syllable-based sub-word model showed a mean recognition rate of 93.48% and standard deviation of 5.02%. The phoneme-based showed a mean recognition rate of 91.33% and standard deviation of 6.25%. Overall, the proposed syllable-based model produced an improvement of 2.05% in recognition. The recognition speed is slower, because of the large size of syllables were used to trained, there were 160 unique isolated words with 430 syllables represents 216 unique syllables, but in a phoneme-based model there were 37 phones to represent whole words. From the scatter diagram, it is clear that as the increase the training data size, the phoneme-based sub-word modelling performed better for isolated word speech recognition in the Tamil language. The result clearly indicates that these two models can be recommended for small vocabulary isolated word speech in the Tamil language. Further work is needed to improve automatic speech recognition of Tamil language. There are four key area: the development of Tamil speech corpus, improved language modelling, the development of acoustic models for speech recognisers and the applications of those speech recognisers.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Harold F. Schiffman. A reference grammar of spoken Tamil. Cambridge University Press; 2006.
ISBN-10: 0521027527
2. Saraswathi S, Geetha TV. Comparison of performance of enhanced morpheme-based language model with different word-based language models for improving the performance of Tamil speech recognition system. ACM Transaction on Asian Language Information Processing. 2007; 6(3):Article 9.
3. Keane Elinor. Tamil. Journal of the International Phonetic Association. 2004; 34(1):111–116.
4. Mohamed Kalith I, David Asirvatham, Raisal I. Context-dependent syllable modeling of sentence-based semi-continuous speech recognition for the Tamil language. Information Technology Journal. 2017;16:125-133.
DOI: 10.3923/itj.2017.125.133
5. Itakura F. Minimum prediction residual applied to speech recognition. IEEE Transaction on Acoustics and Speech Signal Proc.1975;23(1):67-72.
6. Myers CS, Rabiner LR. A level building dynamic time warping algorithm for connected word recognition. IEEE Transaction on Acoustics and Speech, Signal Proc. 1981;ASSP-29(2): 284-297.
7. Wilpon G, Lawrence, Rabiner R, Chin-Hui Lee, Golden ER. Automatic recognition of keywords in unconstrained speech using hidden Markov models. IEEE Transaction on Acoustic Speech and Signal Processing. 1990;38(11).
8. Thangarajan, R. Speech Recognition for Agglutinative Languages. In: Modern Speech Recognition Approaches with Case Studies, Ramakrishnan, S. (Ed.), InTech, Croatia; 2012.
9. Radha V, Vimala C, Krishnaveni M. Isolated word recognition system for Tamil spoken language using back propagation neural network based on LPCC features. Computer Science & Engineering: An International Journal (CSEIJ). 2011;1(4).
10. Ganapathiraju A, Hamaker J, Picone J, Ordowski M, Doddington GR. Syllable based large vocabulary continuous speech recognition. IEEE Transactions on Speech and Audio Processing. 2001;9(4):358–366.
11. Sundarapandiyan S, Shanthi N, Mohamed Yoonus M. Syllable based Tamil language continuous robust speech recognition using MGDF-GWCC with DNN-HMM. International Science Press. 2016;9:3391-3400.
12. Sakthi Vel S. Rule-based Model for Automatic Segmentation of Tamil using Supra-Segmental Features. International Journal of Advanced Research in Computer and Communication Engineering. 2017;6(9).
13. Evandro Gouvêa. Robust group's Open Source Tutorial Learning to use the CMU SPHINX Automatic Speech Recognition system; 2012.
(Accessed 10 May 2017)
Available:http://www.speech.cs.cmu.edu/sphinx/tutorial.html
14. Satori H, Hiyassat H, Harti M, Chenfour N. Investigation Arabic speech recognition using CMU sphinx system. International Arab Journal of Information Technology. 2009;6(2):186-190.
15. Mohammad Abushariah, Raja Ainon, Roziati Zainuddin, Moustafa Elshafei, Othman Khalifa. Arabic speaker-independent continuous automatic speech recognition based on a phonetically rich and balanced speech corpus. The International Arab Journal of Information Technology. 2012;9(1).
16. Bahl LR, Brown PF, De Souza PV, Mercer RL. Acoustic Markov models used in the Tangora speech recognition system. Presented at the IEEE International Conference on Acoustics, Speech, and Signal Processing. 1988;497-500.
17. Karpagavalli S, Chandra E. A review on sub-word unit modeling in automatic speech recognition. IOSR Journal of VLSI and signal processing. 2016;6(6):77-84.
18. Natarajan VN, Jothilaksmi S. Segmentation of continuous Tamil speech into syllable like units. Indian Journal of Science and Technology. 2015;8(17).
DOI: 10.17485/ijst/2015/o8i17/61362
19. Jasleen K. Puneet M. On Developing an Automatic Speech recognition System for Commonly used English Words in Indian

English. International Journal on Recent and Innovation Trends in Computing and Communication. 2017;5(7):87-92.
20. Venkataramana K, Jayaprakash P, Nagendra Babu PS. Developing Telugu speech recognition system using Sphinx-4. International Journal of Innovative Research in Science, Engineering and Technology. 2015;4(11):41-48.
21. Akila A, Chandra E. Performance enhancement of syllable-based Tamil speech recognition system using Time normalization and rate of speech. CSI Transactions on ICT. 2014;2(2):77-84.

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sciencedomain.org/review-history/26597