International Journal of Scientific Engineering and Applied Science (IJSEAS) – Volume-2, Issue-3, March 2016
ISSN: 2395-3470
www.ijseas.com

TAMIL LANGUAGE TRANSLATION USING PORTABLE SPEECH TRANSLATION SYSTEM

[1]J.Poornakala,[2]A.Maheswari
[1]PG Students, Jeppiaar Engineering College, Chennai,Iindia.
[2]Professor, Jeppiaar Engineering College, Chennai,India.

Abstract-Nowadays in most of the places people face problems while speaking with other people, who does not know their languages or other languages, in existing system and technique they record the interaction and translate into another language using manual transaction, in order to avoid the difficulties, In this paper we are going to automatically recognize the speech (in English)and translate to another language(Tamil).In this device it consists of three parts namely speech recognition device, English to Tamil translation and Tamil speech generation ,it first recognize the speech(in English) using speech reorganization device and it displayed it on the screen in English(text) then translate into Tamil language(text) and displayed it, after that it convert into Tamil speech and it should be heard at the other end of the device .

Keyword: Speech Recognition Device, English To Tamil Translation, Tamil Speech Generation.

I. INTRODUCTION

The foreign population in South Korea has steadily increased for the last five years. According to the Annual Foreign Residents report, statistic shows that by the end, South Korea's foreign population reached 1.5 million. As Korean is the official language spoken in Korea, almost all foreigners need language assistance. Thus, for this growing number of foreign visitors, advance translation services or technologies are in high demand. Among these technologies, translation machines are advantageous as they are available twenty-four hours every day for users. Also, for buyers, translation machines may cost less than hiring interpreters for long-term services[3].

In this system, we are going to implement the new system for speech to speech for one language to another language. This process we using voice reorganization board is used to collect voice information from external source and processing and converting into another language. Liquid crystal display is used to display the information. Voice board is used for play output voice[3] .

Speech recognition technology has been studied since the 1960's; however, the technology actually began being used in the 1990's. Since the 2000's, speech recognition technology became popularized as the collection of a corpus was made possible through the internet while computing.power made remarkable advancement. Lately, starting with automobile navigation system, speech recognition technology is applied to various devices including digital camera, smart robot, refrigerator and smart TV. Especially, as mobile terminal with built-in microphone and wireless data network, namely a smart phone, became rapidly popularized, speech recognition technology is being applied to mobile terminals in a wide variety includingvoice search services and personal secretary services. One of the most notable examples is with the speech-to-speech translation technology where speech recognition technology, machine translation technology and speech synthesis technology all came together. Speech-to-speech translation technology represents a technology which automatically translates one language to another language in order to enable communication between two parties with different native tongues. To translate a voice in one language to another voice in a different language, speech-to-speech translation technology is comprised of three core technologies that had been previously mentioned: speech recognition technology, which recognizes the utterance of a human and converts it into a text; machine translation technology, which translates the text in a certain language into a text in another language; and speech synthesis technology,which converts the translated text into a speech. Additionally,the technology to understand the natural language and the user interface-related technology integrated with the UI also play an important role in this speech-to-speech translation system[1].

Research on speech translation has shown that it is possible to build systems that translate spontaneously spoken utterances from one language to another. To handle the challenge of ambiguity introduced by spontaneous speech, most systems introduce semantic constraints by limiting the domain of discourse, thereby reducing the number of suitable interpretations. For many applications, constraining the domain (hotel reservation, scheduling, travel planning, etc.) is quite acceptable and can provide practical translation devices. Over the years, we have developed a large number of such systems for many languages, domains, and platforms. These efforts have shown that acceptable performance can be obtained for spontaneous speech input, but also that practical concerns such as portability and reconfigurability become increasingly important. The two main challenges are in maintaining the existing system and obtaining enough data for each new language and domain.

II. RELATED WORKS

Seung Yun et al [1] established a massive languageand speech database closest to the environment where speech-to-speech translation device actually is being used after mobilizing plenty of people based on the survey on users'demands. Through this study, it was made possible to secure excellent basic performance under the environment similar to speech-to-speech translation environment, rather than just under the experimental environment. Moreover, with the speech-to-speech translation UI, a user-friendly UI has been designed; and at the same time, errors were reduced during the process of translation as many measures to enhance user satisfaction were employed. After implementing the actualservices, the massive database collected through the service was additionally applied to the system following a filtering process
in order to procure the best-possible robustness toward both the details and the environment of the users' utterances. By applying these measures, this study is to unveil the procedures where multi-language speech-to-speech translation system has

been successfully developed for mobile devices.

Thomas Meyer et al[2] shows that the automatic labeling of discourse connectives with the relations they signal, prior to machine translation (MT), can be used by phrase-based statistical MT systems to improve their translations. This improvement is demonstrated here when translating from English to four target languages–French, German, Italian and Arabic–using several test sets from recent MT evaluation campaigns. Using automatically labeled data for training, tuning and testing MT systems is beneficial on condition that labels are sufficiently accurate, typicallyabove 70%. To reach such an accuracy, a large array of features for discourse connective labeling (morpho-syntactic, semantic and discursive) are extracted using state-of-the-art tools and exploited in factored MT models.

Sangmi Shin et al [3] proposed system is a one-way translation that is designed to help English speaking patients describe their symptoms to Korean doctors or nurses. A humanoid robot is useful because it can be extended to reach out to people in need first and may substitute the role of human workers, unlike laptops or tablets. The system consists of three main parts - speech recognition, English-Korean translation, and Korean speech generation.

M.D. Faizullah Ansariet al[4] proposed Voice Translator is speech to speech translation application for android mobile phone, which translates English speech to Hindi speech and vice versa. Voice Translator includes three modules, Voice Recognition, Machine Translation and Speech Synthesis.Voice Recognition module captures the voice or speech from the mobile user through speaker, identifies then converts the speech into text and then the text send to Machine Translation for further process.MachineTranslation module does the process of translation i.e. this module consists of library for both language and when text is received by this module, it converts the text of one language to another as per user choice and thus it sends the translated text to last module. Speech Synthesis module acts as the text to speech translator i.e. when itgets the translated text. This module processes on translated text which converts it into speech and then makes it as user output.

Shigeki Matsuda et al [5] prposedan overview of VoiceTra, which was developed by NICT and released as the world's first network-based multilingual speech-to-speech translation systemfor smartphones, and describes in detail its multilingual speech recognition, its multilingual translation, and its multilingual speech synthesis in regards to field experiments. We show the effects of system updates using the data collected from field experiments to improve our acoustic and language models.

Ping Xu et al [6]proposes using cross-lingual languagemodeling with syntactic information for low-resource speech recognition. We propose phrase-level transduction and syntactic reordering for transcribing a resource-poor language and translating it into a resource-rich language, if necessary. The phrase-level transduction is capable of performing - cross-lingual transduction. The syntactic reordering serves to model the syntactic discrepancies between the source and target languages.Our purpose is to leverage the statistics in a resource-rich language model to improve the language model of a resource-poor language and at the same time to improve low-resource speech recognition performance. We implement our cross-lingual language model using weighted finite-state transducers (WFSTs), and integrate it into a WFST-based speech recognition search space to output the transcriptions of both resource-poor and resource-rich languages. This creates an integrated speech transcription and translation framework.

OzlemKalinli et al [7] propose a noise adaptive training (NAT)algorithm that can be applied to all training data thatnormalizes the environmental distortion as part of the model training. In contrast to feature enhancement methods, NAT estimates the underlying "pseudo-clean" model parameters directly without relying on point estimates of the clean speech features as an intermediate step. The pseudo-clean model parameters learned with NAT are later used with vector Taylor series (VTS) model adaptation for decoding noisy utterances at test time.

IlknurDurgar El-Kahlout et al[8] present the results of our work onthe development of a phrase-based statistical machine translation prototype from English to Turkish—an agglutinative language with very productive inflectional and derivational morphology.We experiment with different morpheme-level representations for English–Turkish parallel texts. Additionally, to help with word alignment, we experiment with local word reordering on the English side, to bring the word order of specific English prepositional phrases and auxiliary verb complexes, in line with the morpheme order of the corresponding case-marked nouns and complex verbs, on the Turkish side. To alleviate the dearth of the parallel data available, we also augment the training data with sentences justwith content word roots obtained from the original training datato bias root word alignment, and with highly reliable phrase-pairs from an earlier corpus alignment. We use a morpheme-based language model in decoding and a word-based language model in re-ranking the -best lists generated by the decoder. Lastly, we present a scheme for *repairing* the decoder output by *correcting*words which have incorrect morphological structure or which are out-of-vocabulary with respect to the training data and language model, to further improve the translations**.**

AarthiReddyetal[9]presents a model for machine-aided human translation (MAHT) that integrates source language text and target language acoustic information to produce the text translation of source language document. It is evaluated on a scenario where a human translator dictates a first draft target language translation of a source language document. Informationobtained from the source language document, including translationprobabilities derived from statistical machine translation(SMT) and named entity tags derived from named entity recognition(NER), is incorporated with acoustic phonetic information obtained from an automatic speech recognition (ASR) system.

Enrique Vidal et al [10] presentscurrent machine translation systems are far from being perfect. However, such systems can be used in computerassistedtranslation to increase the productivity of the (human) translation process. The idea is to use a text-to-text translation system to produce portions of target language text that can be accepted or amended by a human translator using text or speech. These user-validated portions are then used by the text-to-text translation system to produce further, hopefully improved suggestions. There are different alternatives of using speech in a computer-assisted translation system: From pure dictated translation to simple determination of acceptable partial translations by reading parts of the suggestions made by the system. In all the cases, information from the text to be translated can be used to constrain the speech decoding search space. While pure dictation seems to be among the most attractive settings, unfortunatelyperfect speech decoding does not seem possible with the current speech processing technology and human error-correcting would still be required.

Tanja Schultz et al [11] proposed strategies to overcome the limits of today's speech translation systems. In the first part, we describe our layered system architecture that allows for easy component integration, resource sharing across components, comparison of alternative approaches, and the migration toward hybrid desktop/PDA or stand-alone PDA systems. In the second part, we show how flexibility and configurability is implemented by more radically relying on learning approaches and use our English–Thai two-way speech translation system.

III. SYSTEM DESCRIPTION

1. Hmm Human Computer Interface Model

HMM algorithm [1][5] is used to translate efficiently from one language to another language which is used in Android, iPhone OS, it enhance user satisfaction through additional features. Improving the performance of speech-to-speech translation engine by continuously reflecting text and acoustic logs collected from smart mobile devices of the users on the system. Speech to speech translation based on user demands.

2. Phrase-Based Statistical MT Systems

Phrase-Based Statistical MT Systems [2]measured withthe dedicated ACT metric and it is anautomatic labeling of discourse connectives(English to French,German, Italian and Arabic), hence the performance can be gained by syntactic and dependency structures, improve the coherence andreadability of SMT output.

3. CMU Sphnix-4

CMU Sphinix-4 [3] used rule based translation for automatic recognition of one way speech English to Korean in Doctor's office. It can recognize native English speaker up to 84% and not native English speaker up to 64%.it has disadvantage that it is veryComplex, high human interruption.

4. Speech Synthesis

Speech Synthesis [4] used in mobile application for android platform. It make easy to understand and interact with others, it removes the barrier of language between two or more different linguistic people. It can translate from only English to Hindi not to any other languages.

5. Voicetra

VoiceTra [5] used both speech and text as an input in Smartphone and android OS. World's first network based multilingual speech to speech translation system for smartphones, Understandable translation results up to 62.3%. It does not contain large amount of collected speech data.

6. Weighted Finite-State Transducers (WFSTS)

Weighted Finite-State Transducers (WFSTS)[6] is used to translate Cantonese to standard Chinese translation Cross-lingual language model can be easily applied to speech translation of other language pairs for efficient direct decoding from source speech to target text. It is not efficient direct decoding from source speech to target speech.

7. Vector Taylor Series (VTS)

Vector Taylor Series (VTS) [7] is used to automaticspeech recognition in order for checking word accuracy. Experimentsperformed on the Aurora 2 and Aurora 3 tasks demonstratethat the proposed NAT method obtain relative improvements of18.83% and 32.02%, respectively, over VTS model adaptation.

8. Statistical Machine Translation (SMT)

Statistical Machine Translation (SMT) [8] [10] used in speech recognition from English to Turkish. Local word reordering on the English side, to bring the word order of specific English prepositional phrases and auxiliary verb complexes, in line with the morpheme order of the corresponding case-marked nouns and complex verbs, on the Turkish side Translationinto Turkish seems to involve processes that are somewhatmore complex than standard statistical translation models.

9. Automatic Speech Recognition (ASR)

Automatic Speech Recognition (ASR) [4] [9] is used in source language textand target language acoustic information to produce the texttranslation of source language document. The system combination used here is that words that are not included in the ASR vocabulary can be correctly decoded by the combined system.

10. Multilingual Speech-To-Speech Translation

Multilingual Speech-To-Speech Translation [5][11] used in speech-to-speech translation system, speechrecognition, both speech and text as an input and it can be used in smartphone, Android OS. It does not contain large amount of collected speech data.

11. Machine-Aided Human Translation (MAHT)

Machine-Aided Human Translation (MAHT) [9], source language textand target language acoustic information to produce the texttranslation of source language document. The system combination used here is that words that are not included in the ASR vocabulary can be Correctly Decoded By The Combined System.

12. Computer-Assisted Translation (CAT)

Computer-Assisted Translation (CAT) [10] text-to-text translationsystem to produce portions of target language text that can beaccepted or amended by a human translator, private companies (user's manuals, newspapers, etc.). Official institutions (European Union parliament, the CanadianParliament, U.N. sessions, Catalan and Basque parliaments inSpain, etc.) and the decoding of a human translator utterance is part of a prefix of the final target sentence.After user-validation or amending, this prefix is used by the CAT system to suggest a new suffix that the human translator canaccept or modify using speech and/or typing in an iterative way until satisfactory, correct target sentence is finally produced the translation restrictions can be includedinto the speech decoding system.

13. Text To Speech (TTS)

Text To Speech (TTS) [4]is used in Mobile application for android platform. It make easy to understand and interact with others, it removes the barrier of language between two or more different linguistic people. It can translate from only English to Hindi not to any other languages.

IV. RESULTS AND DISCUSSION

In this survey paper we have listed many approaches based on speech recognition system which can able to recognize the speech and translate into another language using speech recognition device, language translation and speech generation, in this survey we can able to translate speech to text, text to speech and speech to speech can be done efficiently.

V. CONCLUSION AND FUTURE WORK

There are many approaches available for speech Recognition and speech translation each approach has its advantages and disadvantages. This survey paper has provided a comprehensive overview of various hang speech recognition approaches and speech translation approaches. The experiment for the proposed system demonstrates reliable performance, but it was done in a limitedenvironment. Thus, further improvements are required to resolve the restricted conditions of the current system. The next goal of the system is to improve the quality of the work so that the system can be tested with real world example. To deal with more expressions and to improve the accuracy of speech recognition.

REFERENCES

[1] Seung Yun, Young-Jik Lee, and Sang-Hun Kim "Multilingual Speech-to-Speech Translation System for Mobile Consumer Devices" *IEEE Transactions on Consumer Electronics, Vol. 60, No. 3, August 2014.*

[2] Thomas Meyer, NajehHajlaoui, and Andrei Popescu-Belis "Disambiguating Discourse Connectives for Statistical Machine Translation" *IEEE/Acm Transactions On Audio, Speech, And Language Processing, Vol. 23, No. 7, July 2015*.

[3] Sangmi Shin, Eric T. Matson "Speech-to-Speech Translation Humanoid Robot in Doctor's Office".

[4] M.D. Faizullah Ansari, R.S. Shaji, T.J.SivaKarthick, S.Vivek ,A.Aravind "Multilingual Speech to Speech Translation System in Bluetooth Environment",*2014 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT).*

[5] Shigeki Matsuda, Xinhui Hu, Yoshinori Shiga, Hideki Kashioka, Chiori Hori, Keiji Yasuda,
Hideo Okuma, Masao Uchiyama, EiichiroSumita, Hisashi Kawai, and Satoshi Nakamura "Multilingual speech-to-speech translation system: VoiceTra" *2013 IEEE 14th International Conference on Mobile Data Management.*

[6] Ping X , Pascale Fung "Cross-Lingual Language Modeling forLow-Resource Speech Recognition", *IEEE Transactions On Audio, Speech, And Language Processing, Vol. 21, No. 6, June 2013.*

[7]OzlemKalinli, *Member, IEEE*, Michael L. Seltzer, *Senior Member, IEEE*, JashaDroppo, *Senior Member, IEEE*, and Alex Acero, *Fellow, IEEE* "Noise Adaptive Training for Robust Automatic Speech Recognition"*IEEETransactions On Audio, Speech, And Language Processing, Vol. 18, No. 8, November 2010.*

[8] ˙IlknurDurgar El-Kahlout and Kemal Oflazer "Exploiting Morphology and Local Word Reordering in English-to-Turkish Phrase-Based Statistical Machine Translation" *IEEETransactions On Audio, Speech, And Language Processing, Vol. 18, No. 6, August 2010.*

[9] Aarthi Reddy and Richard C. Rose, Senior Member," Integration of Statistical Models for Dictation of Document Translations in a Machine-Aided Human Translation Task" *IEEETransactions On Audio, Speech, And Language Processing, Vol. 18, No. 8, November 2010.*

[10] Enrique Vidal, Francisco Casacuberta, Luis Rodríguez, Jorge Civera, and Carlos D. MartínezHinarejos" Computer-Assisted Translation Using Speech Recognition" *IEEE Transactions On Audio, Speech, And Language Processing, Vol. 14, No. 3, May 2006.*

[11] Tanja Schultz, Alan W. Black, Stephan Vogel, and Monika Woszczyna "Flexible Speech Translation Systems" *IEEE Transactions On Audio, Speech, And Language Processing, Vol. 14, No. 2, March 2006.*