**ASSOSA UNIVERSITY**



**COLLAGE OF COMPUTING AND INFORMATICS**

**DEPARTMENT OF COMPUTER SCIENCE**

**RESEARCH**

**On**

**KBS Amharic** **Speech to Text Recogication**

**Prepared by:**

**Name**  **ID**

Bulcha Abdi………..ETR/816/06

Agerie wudu…………ETR /766/06

Besufekade Assefa……ETR /806/06

Dawit Assefa…………ETR/837/06

Haile Tiruneh ……..ETR/413/05

**SUBMITTEDTO:ins Woldue**

**INTRODUCTION**

In the society every one (either human or animals) wishes to interact with each other and tries to convey own message to others. The opponents receive the one’s messages and may get the exact and full idea of the senders. He may get the partial idea or sometimes cannot understand anything out of it. The last case may happen only when there is some lacking in communication (i.e. when a child convey message, the mother can understand easily while others cannot).If we put our eyes in the speech history, we find that initially the human was not able to speak, so for the interaction they were using expressions and they tried to convey their feelings. After some time they had converted these expression into different sounds i.e. for enjoyment they used some type of sound, for afraid they used another type of sound and so on. The time has converted the things and gave the birth to different characters and collectively all the characters had created the words and finally the sentence were formed.

Speech-to-text-translation (audiovisual translation) of spoken language into written text is an upcoming field since movies on DVDs are usually sold with subtitles in various languages.While the original language is given auditorily, subtitles provide a translated version in another language at the same time visually. The audiovisual transfer from the spoken original language into other languages which are presented in the subtitles can be called an interlingual audiovisual translation. Interlingual translation aims at transferring messages from one language into another language. This translation process combines classical interpreting with a transfer from spoken language patterns into written text patterns. Auditory events which are realized as noises or speech melodies would often not be transferred because normally hearing people can interpret them by themselves.

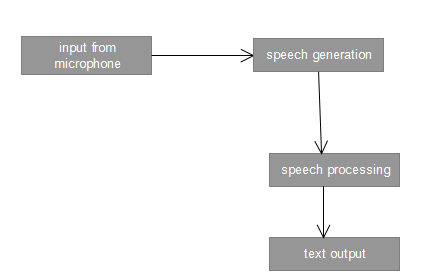


Fig 1.1 Speech-to-text translation

Language is a very fast and effective way of communicating. To use language means to express an unlimited amount of ideas, thoughts and practical information by combining limited amount of words with the help of a limited amount of grammatical rules. The result of language production processes are series of words and structure. Series of words are produced – i.e. spoken or signed – in a very rapid and effective way. Any person can follow such language production processes and understand what the person wants to express if two preconditions are fulfilledtherecipientsmust.  
1. know the words and grammatical rules the speaker uses and  
2. be able to receive and process the physical signal.

**Statement of the problem**

As far as the researcher’s knowledge on NLP is concerned, researches made in the area of  
Amharic natural language processing are very limited in number.

1.To be fast enough in producing written language that.  
2. It becomes possible to meet the expectations of the audience with respect to the characteristics of a written text. Word-by-word transfer enhanced by a description of auditory events from the surroundings as well as adaptations of the original wording into easier forms of language must be possible. Moreover,  
3. A successful real-time presentation must match the reading abilities of the audience, i.e. the written words must be presented in a way that is optimally recognizable and understandable for the readers.

**Objectives of the study**

## General Objective

* The general objective of this study is to develop speech recognition for speech to text conversion for Amharic language.

## Specific Objectives

* To identify and gather information about research on speech to text
* To analyze and specify research on speech to text
* To design a way to solve the problem

## Scope and limitation of the research

**Scope of the research**Speech recognition the word leads the readers mind in a lot many different paths, but here the study is concerned only with speech pattern recognition, for conversion of the inputted speech into equivalent text.

**Limitation of the study**

* The software cannot catch the entire word as well as the continuous speech.
* It does not contain any word dictionary so spell check is not possible.
* It does not contain the language grammar rule, hence grammar check is also not possible.
* The software compares the spoken character with the attribute of the store database character and generates the ranking and highest ranked characters will be returned. So sometimes it may be possible that two different characters get the same rank, in such situation, the software may get confused for right selection.

**Methodology**

In order to develop the research we will use different methodology.

**Literature review**

So far many Speech-to-text *r*ecognition researches have been done and used Knowledge based

Approach this approach focuses on to mechanize the speech recognition process according to

the way a person applies intelligence in visualising, analyzing, and characterizing speech based on a set of measured acoustic features. The Artificial Intelligence approach is a hybrid of the acoustic phonetic approach and pattern recognition approach. Both acoustic phonetic and template based approach failed at their own to explore considerable insight into human speech processing. As a result, error analysis and knowledge based system enhancement couldn‟t get strength. In traditional Knowledge based approach, the production rules are created heuristically from empirical linguistic knowledge or from the observations from the speech spectrogram. Knowledge helps the algorithm to perform better and also in the selection of a suitable input representation, the definition of units of text.

**Data Collection Methodology**

To develop the Speech-to-text *r*ecognition System, data collection is one of the important tasks which must be carried out to gather information in the research. The data for developing the Speech-to-text *r*ecognition system will be obtained from different sources. In gathering and collecting necessary data and information needed from the research and used secondary data collecting metdology. These are:-

**Secondary Source**: The secondary data were obtained from document analysis of the research.

**Document Analysis:** manually collected files will be analyzed for collecting and specifying research. Since there are forms used to fill information related to text *r*ecognition document analysis will be more appropriate to gather information.

**Significant of the research**

More recently student with learning or physical disabilities have been able to use speech to text conversation. Those with learning disabilities that affect their ability to write can now complete exam via voice recognition technology and those with physical disabilities such as upper body paralysis can use speech to text conversation to communicate effectively with others and Even the most experienced typists will occasionally have a spelling blunder, the a verage person is likely to make several mistakes in his or her composition.SRS always provides the ever the most experienced typists will occasionally have a spelling blunder,the average person is likely to make several mistakes in his or her composition.SRS always provides the correct splling of a word (assuming it translated it accurately in the first place),thus eliminating the need to spend time running spell checkers.