Amharic Speech Recognition for Speech Translation

Michael Melese, Laurent Besacier, Million Meshesha

Amharic Speech Recognition for Speech Translation

Michael Melese Woldeyohannis [1], Laurent Besacier [2], Million Meshesha [1]
(1) Addis Ababa University, Addis Ababa, Ethiopia
(2) LIG Laboratory, UJF, BP53, 38041 Grenoble Cedex 9, France
michael.melese@aau.edu.et, laurent.besacier@imag.fr,
million.meshesha@aau.edu.et

ABSTRACT

The state-of-the-art speech translation can be seen as a cascade of Automatic Speech Recognition, Statistical Machine Translation and Text-To-Speech synthesis. In this study an attempt is made to experiment on Amharic speech recognition for Amharic-English speech translation in tourism domain. Since there is no Amharic speech corpus, we developed a read-speech corpus of 7.43hr in tourism domain. The Amharic speech corpus has been recorded after translating standard Basic Traveler Expression Corpus (BTEC) under a normal working environment. In our ASR experiments phoneme and syllable units are used for acoustic models, while morpheme and word are used for language models. Encouraging ASR results are achieved using morpheme-based language models and phoneme-based acoustic models with a recognition accuracy result of 89.1%, 80.9%, 80.6%, and 49.3% at character, morph, word and sentence level respectively. We are now working towards designing Amharic-English speech translation through cascading components under different error correction algorithms.

KEYWORDS: Amharic Speech Recognition, Speech Translation, Under Resourced Languages, Amharic Speech Corpus

1. Introduction

According to the official site of the Ethiopian Embassy in the USA, Ethiopia has much to offer for international tourist [1]. It is a land of natural contrasts, ranging from the peaks of the rugged Semien Mountains to the depths of the Danakil Depression, which is one of the lowest points on earth more than 400 feet below sea level.

According to report of United Nation 2013 world tourism (UNWTO, 2013) and World Bank[2], a total of 770,000 non-resident tourists come to Ethiopia to visit different locations; out of more than 1 billion international tourists for the year 2015 to visit several tourist attraction including world heritages, which are registered as Ethiopian tourist attractions by UNESCO. In fact, most of non-resident visitors speak foreign languages hindering them to communicate with the local tourist guide as language barrier is a major problem for today's global communication. As a result, they look for

---

[1] Investing in Ethiopia available at http://www.ethiopianembassy.org/PDF/investingtourism.pdf
[2] http://data.worldbank.org/indicator/ST.INT.ARVL

bilingual guide or bilingual system which serve as intermediary between the tourists/visitors and that of the guide.

In an increasingly globalized economy and humanitarian service, speech to speech translation attracts communication between people who speak different languages by making computers to understand speech (Honda, 2003; Gao et. al, 2007). As most natural form of communication, speech allows human beings including tourist, travel agencies, tour operators, hotels, transport user and other people to communicate effectively in public at large.

The state-of-the-art speech translation can be seen as a cascade of three major components (Gao et. al, 2006); Automatic Speech Recognition (ASR), Statistical Machine Translation (SMT) and Text-To-Speech (TTS) synthesis. ASR is the process of converting speech input into its equivalent textual representation. Whereas, TTS is the process of producing a synthesized speech for a text given to synthesize. Between speech recognition and synthesis, SMT takes the result of speech recognizer as an input and convert the text into target language based on which the speech synthesizer generates a synthetic sound.

Therefore, there is a need to develop a speech translation system so that tourists can effectively communicate with the tourist guide regardless of the language that they speak. As a matter of fact the success of such a system greatly depends on speech recognition. Hence in this study an attempt is made to select the best unit to use for acoustic and language model units that helps to design an optimized Amharic speech recognizer in tourism domain.

2.   Related works

Research in speech translation started in 1983 by NEC Corporation in the ITU Telecom World (Karematsu et. al, 1996), when they demonstrate speech translation as an approach for selected languages. Currently speech translation aimed at translating a speech signal of a source language to another speech signal in a target language using cascading speech translation components (Xiaodong et al,. 2011); Subsequently, a number of speech translation research have been attempted for resourced and technological supported language as discussed in (Vemula et. al, 2010, Chiori 2012, Nakamura 2014, He and Deng 2011; Gao et. al 2006).

A number of speech translation research have been conducted for technological supported languages like English and French. On the other hand, attempts in this field for under resourced languages like Amharic, in particular is not yet started so far. For Amharic, most of the research is conducted on ASR rather than on SMT and TTS due to the comparative availability of resources as compared to other Ethiopic languages.

Amharic speech recognition started in 2001 when Berhanu, (2001) developed an isolated Consonant-Vowel syllable recognition system. Subsequently, several attempts have been made in the academic research for speech recognition as discussed in (Tachbelie, et. al, 2014). These researches were conducted using different methods and techniques to solve a number of problems in the process of recognition without taking speech translation into account. Besides ASR as a cascading component, preliminary experiments were conducted for English-Amharic SMT (Teshome, et. al, 2012) and encouraging result were found. Later on, the result obtained improved by using phonemic transcription on the Amharic corpus (Teshome, et. al, 2015). As a last component of speech translation, a number of TTS research have been attempted using a number of

techniques and methods (Anberbir 2009, Leulseged 2003 and Sebsibe et. al 2004) to solve different problems. Among these, concatenative, cepstral, formant and a syllable based speech synthesizers were used. However, all the above research cannot be directly used for this research work due to a number of reasons. These reasons include research domain, resource unavailability, different methods and techniques used to solve the problem, size of data used and the continuity of research attempted for speech translation.

Therefore, the aim of this research is to select the best unit for acoustic models and language models for designing a speech recognizer with minimal error that can help in developing an optimized speech translation system.

3. Amharic language

Amharic is the official working language of government of Ethiopia, among 89 languages which are registered in the country. Amharic is second largest spoken Semitic language in the world next to Arabic (CSA, 2007). The majority of the speakers of Amharic can be found in Ethiopia, but there are also speakers in a number of other countries, particularly Israel, Eritrea, Canada, the USA and Sweden. It has five dialectical variation across different parts of the country (Paul, 2009). These includes dialects like Addis Ababa, Gojam, Gonder, Wollo and Menz.

3.1 Amharic writing system

Unlike other Semitic languages, such as Arabic and Hebrew, Amharic /ˈämarǝña/ script uses a grapheme based writing system called fidel /fidälǝ/ which is written and read from left to right (Grover, 2009). Modern Amharic has inherited its writing system from Ge'ez /gǝˈǝzǝ/, which is still the classical and ecclesiastical language of Ethiopia (Abyssinica, 2015).

An Amharic character represents a consonant vowel (CV) sequence and the basic shape of each character is determined by the consonant, which is modified for the vowel. There are speech sounds of Amharic that are specific and not found in any other foreign language (Leslau, 2000). These include sounds such as ፕ/p'/, ፕ/t'/, ፅ/s'/, ጭ/ʧ'/, and ቅ/k'/ which have a sharp click–like characters beside glottalized voice articulating at different places. Amharic symbols are categorized into four different categories consisting 276 distinct symbols; these are core character, labiovelar, labialized and labiodental. The detail category is presented in Table 1.

Category	Character set	Order	Total
Core characters	33	7	231
labiovelar	4	5	20
labialized	18	1	18
labiodental	1	7	7
		Total	**276**

TABLE 1: Distribution of Amharic character set

Amharic has a total of 231 (33*7) distinct core characters, 20 (4*5) labiovelar symbols, 18 labialized consonants and 7 labiodental. The first category possess 33 primary characters each representing a consonant having 7 order in form to indicate the vowel which follows the consonant to represent

CV syllables. Table 2 shows sample core characters used in Amharic writing system with their seven orders.

	Order						
	1st	2nd	3rd	4th	5th	6th	7th
	ə	u	i	a	e	ɨ	o
h	ሀ	ሁ	ሂ	ሃ	ሄ	ህ	ሆ
l	ለ	ሉ	ሊ	ላ	ሌ	ል	ሎ
h	ሐ	ሑ	ሒ	ሓ	ሔ	ሕ	ሖ
m	መ	ሙ	ሚ	ማ	ሜ	ም	ሞ
s	ሠ	ሡ	ሢ	ሣ	ሤ	ሥ	ሦ
r	ረ	ሩ	ሪ	ራ	ሬ	ር	ሮ
s	ሰ	ሱ	ሲ	ሳ	ሴ	ስ	ሶ

TABLE 2: Sample Amharic core characters

In the same way, labiodental category contains a character ቨ/v/ with 7 order (ə, u, i, a, e, ɨ, o) borrowed from foreign languages and appears only in modern loan words like ቫይታሚን /vajitamini/. Similarly, the labiovelar category contains 4 (ቅ/k'/, ኅ/h/, ክ/k/ and ግ/g/) characters with 5 orders (ʷə, ʷi, ʷa, ʷe, and ʷɨ) that generates 20 distinct symbols. Furthermore, there are labialized 18 characters for instance ሏ/l ʷa /, ሟ/m ʷa /, ሯ/r ʷa / and ሷ/s ʷa /.

In Amharic writing, all the 276 distinct symbols are indispensable due to their distinct orthographic representation. Whereas for speech recognition, we mainly deal with distinct sound rather than with orthographic representation; thus, among the given character set, different graphemes generate the same sound and this greatly minimizes the number of sounds to be modelled in speech recognition. Table 3 presents graphemes that have been normalized into common graphemes.

Number of Graphemes	Graphemes to be normalized	Equivalent sound	Normalized Graphemes
4	ሀ, ሐ, ኀ, and ኸ	/h/	ሀ /h/
2	አ and ዐ	/ʔ/	አ /ʔ/
2	ሠ and ሰ	/s/	ሰ /s/
2	ፀ and ጸ	/ ts'/	ጸ/ts'/

TABLE 3: List of normalized Amharic speech sounds

Among the given 33 core character set, graphemes with multiple variants have to be normalized into their sixth order to generate equivalent sound as shown in Table 3. The selection of graphemes is made based on the usage of character in Amharic document. Thus, as a result of normalizing the sound (ሀ, ሐ, ኀ, ኸ) to ሀ /h/, (አ, ዐ) to አ /ʔ/, (ሠ , ሰ) to ሰ /s/ and (ፀ, ጸ) to ጸ/ts'/ the total number of distinct sounds reduced from 33 to 27 from the models which is 18.18%.

4. Corpus Preparation

One of the most fundamental resources for any ASR system and development is speech and text corpora. Collecting standardized and annotated corpora is one of the most challenging and expensive task when working with under resourced languages (Lewis et al., 2012). Unlike English and European languages such as French, Spanish, Amharic can be considered as an under-resourced and technologically less supported language that suffers from devising text and speech corpora in digital format.

The speech corpus used for the development and training of speech recognition system is a 20hr Amharic read speech[3] prepared by Tachbelie et al, (2014) from EthioZena website. Whereas for testing, inaccessibility of standardized digital corpora in tourism domain were the challenge for the researcher. However, to overcome the problem that arise from unavailability of the resource in tourism domain, a parallel English-Arabic text corpus was acquired from BTEC 2009 available through International Workshop on Spoken Language Translation (Kessler, 2010). BTEC corpus contains basic traveler expression corpus. The initial English corpus is translated to Amharic to prepare parallel Amharic-English BTEC using a bilingual speaker; and this data is used for the development of speech corpus for ASR and Amharic-English parallel corpus for statistical machine translation (Gao et al., 2006).

A large amount of speech data can be collected using mobile phones which speeds up data collection as compared to traditional methods of data collection (Davel et. al, 2014). Mobile and handheld device is becoming increasingly available and sharply decreasing cost even for the developing country to collect speech data. As a result, Amharic speech data is recorded using smartphone based application for speech data collection tool Lig-Aikuma (Blachon, et. al, 2016) under normal office environment. The speech data is collected from eight native Amharic speakers (4 male and 4 female) with different age range. The speakers read each aligned sentence with the possibility to record again the sentence anytime they mispronounced the sentence. Table 4 shows the age and gender distribution of the speakers.

	Age and Gender			
	Male		Female	
	18-30	31-50	18-30	31-50
Number of Utterances	1000	1112	1000	1000
	1000	1000	1000	1000
Total	2000	2112	2000	2000

TABLE 4: Distribution of utterance per age and gender

A total of 8112 sentences with a length ranging from 1 to 28 word length (average of 4 to 5 word per sentences) have been recorded. A total of 7.43hr read speech corpus ranging from 1020 ms to 14633 ms with an average speech time of 3297 ms was collected. The distribution of speech length across sentence is presented as shown in Figure 1.

---

[3] Amharic Speech data available at https://github.com/besacier/ALFFA_PUBLIC/tree/master/ASR
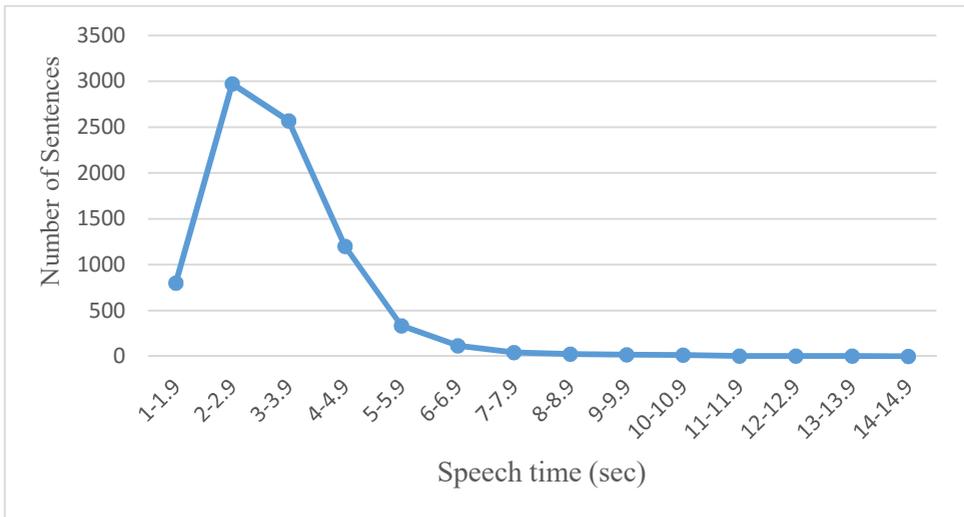
FIGURE 1: Speech length vs sentence distribution

As we can see from Figure 1, 98.54% of the speech data fall below 7sec. For this paper, to avoid bias in test set selection proportional amount of data selected from each speaker. Thus, a total of 507 utterances have been selected from a total of 8112 sentences that consist of 16 different groups.

The language model data has been collected from different sources. A text corpus collected for Google project (Tachbelie et al., 2015) used as an out-domain data and separately translated BTEC has been used for the purpose of domain adaptation. The out-domain data consist of 219,631 sentences (4,003,956 tokens) of 319,858 types. On the other hand, in-domain data contains 22,616 sentences having 113,903 tokens of 17,694 types. A total of 242,247 sentences (4,117,859 tokens of 326, 630 type) have been used to train 3-gram language models. Compared to other standard corpora, this corpus is very small in terms of size and accordingly the models will suffer from lack of sufficient training data.

5.   Experimentation and discussion

The ASR experiments have been conducted at different acoustic, lexical and language model units. In all experiments, we used Kaldi[4] tool for speech recognition and SRILM[5] for language model. Furthermore, for morpheme-based recognition, a word segmenter is required to split word forms into sub-word units for speech recognition. A number of research attempted for developing Amharic morphological analyzer as discussed in (Tachbelie et al., 2009), but none of them can be used directly for this work due to the size of lexicon used for segmentation, unavailability of data, etc. As a result, a corpus-based, language independent and unsupervised segmentation tool morfessor

---

2.06 (Mathias, 2002) is used to segment word forms. The segmentation is then applied on the training, testing and language model data.

5.1 Acoustic modeling units

Speech recognition requires segmentation of speech data into fundamental acoustic units (Abate, 2005). The ideal and natural acoustic unit is the word (Tachbelie et al., 2014). On the other hand, the use of words as acoustic units large vocabulary in speech recognition systems is impractical because of the need for a very large data to train models sufficiently. Accordingly, syllable and phonemes units were used at the acoustic model (AM) level. Besides acoustic modeling units, the speech corpus used to develop the systems is an Amharic read speech corpus consisting of 10,875 sentences (28,666 tokens) for training and 507 (2470 word or 3464 morpheme token) for testing after translating the English text to Amharic from a parallel English-Arabic BTEC corpus.

5.2 Lexical and Language modeling units

Automatic speech recognition systems works with a pre-defined lexicon, i.e., the number of distinct words it contains, which is an important parameter for an ASR system. If these words are not in the lexicon; then the word is considered as out-of-vocabulary (OOV) which is one of the main source of error in automatic speech recognition. Thus, as a result of OOVs, the word might be recognized to other similar units, which will lead the adjacent words to be misrecognized to different word.

In the lexicon, the pronunciation of each word is defined. For this paper, separate phoneme and syllable pronunciation dictionary have been prepared for both word and morpheme based recognition. A 28,861 words pronunciation dictionary prepared from separate BTEC training data is used for word based recognition. In the same way, after segmenting the words using morfessor, 14,132 words dictionary is used for morpheme based recognition. The OOV rate of 28.18% and 6.28% achieved for word based and morpheme based recognition, respectively. The lower rate of OOV in morpheme-based recognition is obtained as a result of morfessor based segmentation.

Besides the lexical model, we used word and morpheme based language models. The language model used for our ASR experiment is an interpolation between in-domain and out-domain data. A weight of 0.9 given to small in-domain LM. The LMs has a perplexity of 49.3 and 24.7 for words and morphemes on testing dataset respectively. Despite the difference in lexicon and rate of OOV, the Amharic speech recognition system has been tested using phone and syllable based acoustic unit with words and morphemes based language models.

5.3 Experimental result

As Amharic is a morphologically rich and complex language, presenting the evaluation result in different units is important. Accordingly, the experimental result is presented in terms of word recognition accuracy (WRA), morph recognition accuracy (MRA), Character Recognition Accuracy (CRA) and Sentence Recognition Accuracy (SRA).

---

[6] The unit obtained with Morfessor segmentation is referred here as morpheme without linguistic definition of morpheme.

The WRA accuracy for morpheme-based recognition has been computed after words have been obtained by concatenating the recognized morph sequence. Whereas, the MRA of word-based recognition result are obtained by segmenting the result of recognition and reference of test set. Then, the result for each system is computed using NIST Scoring Toolkit (SCTK[7]). Table 5 below shows the comparison of the ASR experiments conducted with respect to phoneme and syllable as acoustic unit against morpheme and word based language model units with respect to our interpolated LMs.

		Phoneme	Syllable
Morpheme based LM	CRA	89.1	85.5
	MRA	80.9	75.8
	WRA	80.6	75.8
	SRA	49.3	43.4
Word based LM	CRA	70.1	69.7
	MRA	52.3	50.9
	WRA	56.0	54.7
	SRA	13.2	13.2

TABLE 5: Recognition accuracy of phonemes and syllable based recognition

The performance of systems has been computed with respect to each unit as shown in Figure 2. Under morphemes-based LM as a unit; 80.6% for WRA, 80.9% MRA, 89.1% for CRA and 49.3% for SRA results were found using phonemes as AM unit. In the same way, for using syllable as AM with morphemes-based LM, a 75.8% for WRA, 75.8% MRA, 85.5% for CRA and 43.4% SRA recognition result found.

In addition to this, using words as LM unit, 56.0% for WRA, 52.3% MRA, 70.1% for CRA and 13.2% SRA recognition result were obtained using phonemes as AM unit. Similarly, using syllable as AM unit, 50.9% for WRA, 54.7% MRA, 69.7% for CRA and 13.2% SRA results were achieved under the same rate of OOV.

Moreover to this, recognition accuracy of word and morph are almost indistinguishable in morph-based LM and rather different in word-based LM. This is due to the segmentation of a word into sub-word units which removes the morphological variations encountered during recognition. Figure 2 presents recognition accuracy result achieved for morpheme-based and word-based LMs against phoneme and syllable based AMs.
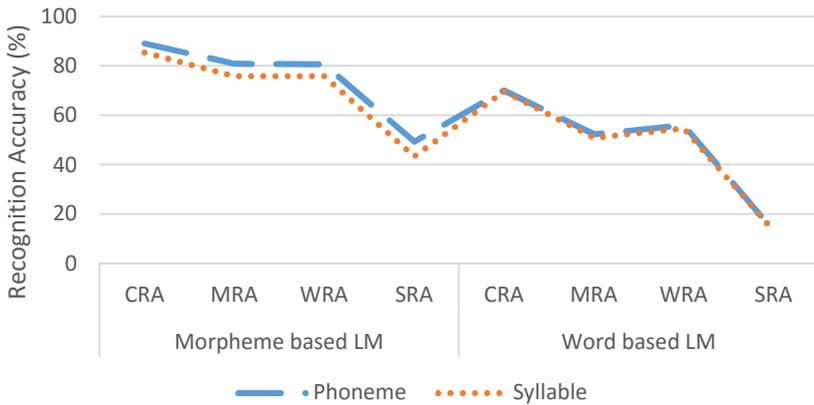
---

[7] http://www.openslr.org/4/

FIGURE 2: Experiment result for phoneme and syllable based recognition.

The performance of morpheme-based LM with phoneme AM outperforms others with a performance improvement by at least 3.6% at character level, 5.1% at morph level, 4.8% at word and level 5.9% at sentence than other unit of representation.

The performance difference between phoneme and syllable based recognition under the same LM unit appeared as a result of context-dependency of triphones under the same rate of OOV. Correspondingly, performance improvement at morpheme based recognition resulted because of segmenting words into their constituent's level which provides less rate of OOV.

The result obtained from the experiment shows that using morpheme as LM achieved better recognition due to the low rate of OOV which is true for morphological rich language like Amharic. In addition to this, phonemes based recognition provides a better performance than syllable due to the context-dependency.

6. Conclusion and further work

Speech recognition and translation is a field which has been and being researched for more than a decade for most of the resourced languages like English and most European languages. On the other hand, attempts in this area for under resourced languages like Amharic speech translation, in particular, not yet started. To facilitate Amharic speech recognition for speech translation, an attempt is made to construct a read-speech corpus of 7.43hr in tourism domain. The Amharic speech corpus has been recorded after linguist translate a standard Basic Traveler Expression Corpus (BTEC) under a normal office working environment. In addition to this, in-domain language model data is prepared to adapt domain. Our experiments show that the best recognition results achieved at morpheme based LM with phoneme based AM recognition which is acceptable for morphological rich language like Amharic. The result we found from the experiments is promising to design Amharic-English speech translation by means of cascading components through different error correction algorithms at different stage, which is our next research direction.

References

[Abate, 2006] Solomon Teferra Abate. (2006). Automatic Speech Recognition for Amharic. Ph.D. Thesis. Available at: http://www.sub.uni-hamburg.de/opus/volltexte/2006/2981/pdf/thesis.pdf.

[Abyssinica, 2015] Abyssinica dictionary (2015). Amharic, the Official Language of Ethiopia available at http://dictionary.abyssinica.com/amharic.aspx.

[Anberbir, 2009] Tadesse Anberbir, Tomio Takara (2009), Workshop on Language Technologies for African Languages – afLat 2009, pages 46–52, Athens, Greece.

[Berhanu, 2001] Solomon Berhanu, (2001) "Isolated Amharic Consonant-Vowel syllable recognition: An Experiment using the Hidden Markov Model," MSc Thesis, School of Information Studies for Africa, Addis Ababa University, Ethiopia.

[Blachon et. al, 2016] Blachon, D., Gauthier, E., Besacier, L., Kouarata, G.-N., Adda-Decker, M., and Rialland, A. (2016). Parallel speech collection for under-resourced language studies using the lig-aikuma mobile device app. SLTU.

[Chiori, 2012] HORI Chiori, (2012) "Network based Speech-to-Speech Translation," Journal of the National Institute of Information and Communications Technology, vol. 59, pp. 231–237.

[Creutz et al., 2002] Mathias Creutz, and Krista Lagus (2002). Unsupervised Discovery of Morphemes. Proceedings of the Workshop on Morphological and Phonological Learning of ACL-02 pp. 21–30. Philadelphia, Pennsylvania.

[CSA, 2007] CSA (Central Statistics Agency) (2007). Addis Ababa, Ethiopia: Central Statistics Agency. http://www.csa.gov.et

[Davel et. al, 2014] Marelie H. Davel, Jaco Badenhorst, Willem D. Basson, Febe De Wet, Etienne Barnard, Alta De Waal, Nic J. De Vries (2014). A Smartphone-Based ASR Data Collection Tool for under-Resourced Languages 56: 119–131.

[Gao et al., 2006] Yuqing Gao, Liang Gu, Bowen Zhou, Ruhi Sarikaya, Mohamed Afify, Hong-kwang Kuo, and Wei-zhong Zhu, Yonggang Deng, Charles Prosser, Wei Zhang and Laurent Besacier (2006). IBM Mastor: Multilingual Automatic Speech-To-Speech Translator. 2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings vol. 5, p. 5.

[Gao et. al, 2007] Yuqing Gao, Liang Gu and Bowen Zhou (2007). Speech-to-Speech Translation. World Scientific.

[Grover, 2009] Hudson, Grover. (2009). The World's Major Languages : Amharic. In The World's Major Languages. Second Edition. Pp. 594–617. Oxon and New York: Routledge.

[He and Deng 2011] Xiaodong He and Li Deng, (2011) "Speech recognition, machine translation and speech translation a unified discriminative learning paradigm," IEEE Signal Process. Magazine, vol. 28, no. 5, pp. 126–133.

[Honda 2003] Masaaki Honda (2003). Human Speech Production Mechanisms. NTT Communication Science Laboratories 1(2): 24–29.

[Karematsu et. al, 1996] Akira Karematsu and Tsuyoshi Morimoto, (1996) Automatic Speech Translation : Fundamental Technology for future Cross-language Communication, vol. 28. Gordon and Breach Publishers.

[Kessler, 2010] Fondazione Bruno Kessler (2010), International Workshop on Spoken Language Translation; Paris, France 2-3.

[Leslau, 2000] Wolf Leslau (2000). Introductory Grammar of Amharic. Wiesbaden: Harrassowitz.

[Leulseged 2003] Henok Leulseged. (2003): Concatinative Text-to-Speech synthesis for Amharic Language, msc. Thesis, Addis Ababa University, Addis Ababa,Ethiopia,.

[Lewis et al., 2012] William D. Lewis and Phong Yang (2012). Building MT for a Severely Under-Resourced Language: White Hmong. In. USA: Association for Machine Translation in the America.

[Nakamura 2014] Satoshi Nakamura, "Towards real-time multilingual multimodal speech-to-speech translation," in SLTU-2014 Workshop Proceedings, St. Petersburg, Russia, 2014, p. 268.

[Paul, 2009] Lewis Paul (2009). Language of the World. Sixteenth edition. Dallas, Texas: SIL International Publications. Https://www.ethnologue.com/world, accessed January 26, 2014.

[Sebsibe et al 2004] Sebsibe H., S P Kishore, Alan W Black, Rohit Kumar, and Rajeev Sangal, (2004) "Unit Selection Voice For Amharic Using Festvox," in 5th ISCA Speech Synthesis Workshop, Pittsburgh,.

[Tachbelie et. al, 2009] Martha Yifiru Tachbelie, Solomon Teferra Abate, and Wolfgang Menzel (2009), Morpheme-Based Language Modeling for Amharic Speech Recognition, In Proceedings of the 4th Language and Technology Conference - LTC-09, pp. 114– 118.

[Tachbelie et. al, 2014] Martha Yifiru Tachbelie and Solomon Teferra Abate and Laurent Besacier (2014). Using different acoustic, lexical and language modeling units for ASR of an under-resourced language - Amharic, Speech Communication, vol. 56.

[Tachbelie et. al, 2015] Martha Yifru Tachbelie and Solomon Teferra Abate (2015). Effect of Language Resources on Automatic Speech Recognition for Amharic. In pp. 1–5. Addis Ababa, Ethiopia: IEEE.

[Teshome et. al, 2012] Mulu Gebreegziabher Teshome, and Laurent Besacier (2012). Preliminary Experiments on English-Amharic Statistical Machine Translation. In pp. 36–41. Cape Town, South Africa.

[Teshome et. al, 2015] M. G. Teshome, L. Besacier, G. Taye and D. Teferi: Phoneme-based English-Amharic Statistical Machine Translation, AFRICON, 2015, Addis Ababa, Ethiopia 2015, pp. 1-5.

[UNWTO, 2013] United Nation World Tourism Organization (UNWTO) annual report 2012: (2013) World Tourism Organization (UNWTO), Madrid, Spain.

[Vemula et. al, 2010] Venkata Vinay Babu Vemula, Pradeep Kumar Narne, and Mrudula Kudaravalli, Poornima Tharimela, Kishore Prahallad, ANUVAADHAK (2010): A Two-way, Indian language Speech-to- Speech Translation System for Local Travel Information Assistance, Int. J. Eng. Sci. Technol., vol. 2, no. 8, pp. 3865–3873.