Amharic-English Speech Translation in Tourism Domain

Michael Melese Woldeyohannis,
Addis Ababa University,
Addis Ababa, Ethiopia

Laurent BESACIER,
LIG Laboratory, UJF,
Grenoble, France

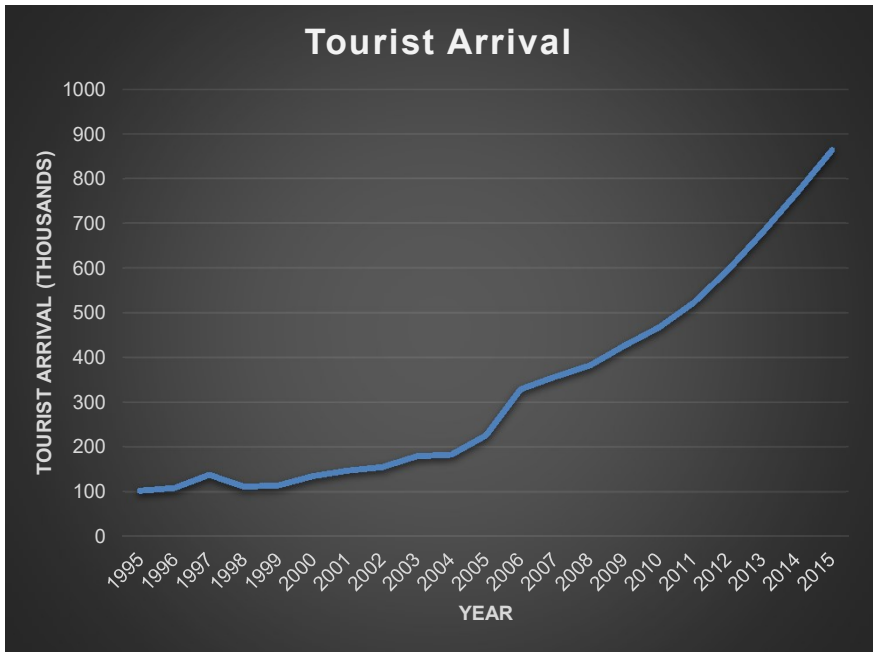
Million Meshesha,
Addis Ababa University,
Addis Ababa, Ethiopia

Overview of speech translation

- Speech translation research for major and technological supported languages has been conducted since the 1983s by NEC corporation when they demonstrate as an approach
  - English, European languages (like French and Spanish) and Asian languages (like Japanese and Chinese)
- Computer with the ability to understand natural language promoted the development of man-machine interface people to communicate effectively in public.
  - This can be extended through different digital platforms such as radio, mobile, TV, CD and others.

Ethiopia and Tourist attraction

- Ethiopia has much to offer for international tourists. These include;
  - peaks of the rugged Semien mountains to the lowest points on earth called Danakil Depression which is more than 400 feet below sea level
  - Tourist attraction including world heritages, which are registered by UNESCO
- Since the year 2010 until 2015, the average number of tourist flow increase by 13.05% per year to visit different location in Ethiopia.
- Amharic is the official language of the government of Ethiopia and means of communication by the society among the 89 language in the country.



Amharic language

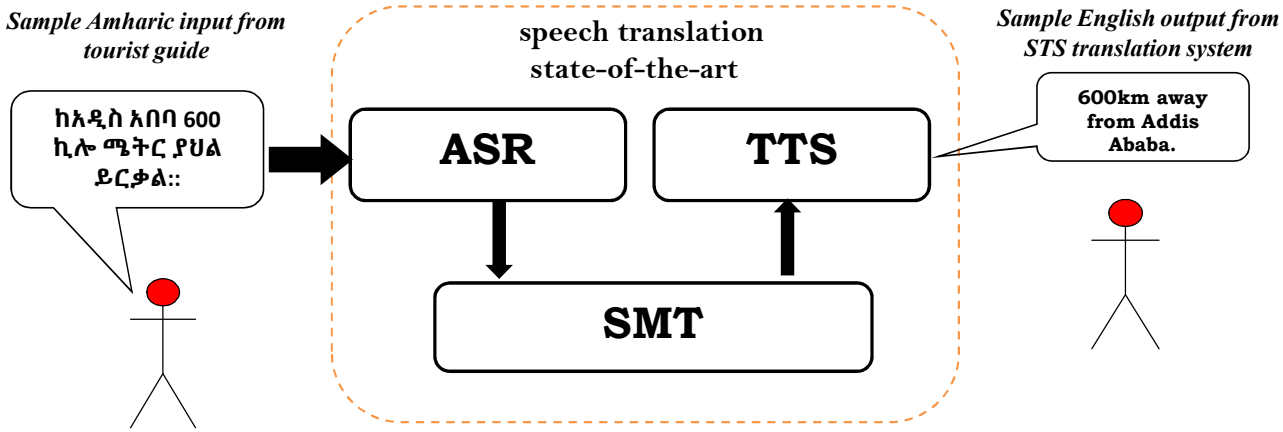
- Amharic is the 2nd largest spoken Semitic languages among 89 registered languages in the country with up to 200 different spoken dialects.
  - Unlike other Semitic languages, such as Arabic and Hebrew, Amharic (አማርኛ) script uses a grapheme called fidel (ፊደል).
- Amharic language is under-resourced

	ə	u	i	a	ē	ɨ	o	wə	wi	ua	wē	wɨ
1	ሀ	ሁ	ሂ	ሃ	ሄ	ህ	ሆ					
2	ለ	ሉ	ሊ	ላ	ሌ	ል	ሎ			ሏ		
3	ሐ	ሑ	ሒ	ሓ	ሔ	ሕ	ሖ			ሗ		
4	መ	ሙ	ሚ	ማ	ሜ	ም	ሞ			ሟ		
5	ሠ	ሡ	ሢ	ሣ	ሤ	ሥ	ሦ			ሧ		
6	ረ	ሩ	ሪ	ራ	ሬ	ር	ሮ			ሯ		
7	ሰ	ሱ	ሲ	ሳ	ሴ	ስ	ሶ			ሷ		
8	ሸ	ሹ	ሺ	ሻ	ሼ	ሽ	ሾ			ሿ		
9	ቀ	ቁ	ቂ	ቃ	ቄ	ቅ	ቆ	ቈ	ቊ	ቋ	ቌ	ቍ
10	በ	ቡ	ቢ	ባ	ቤ	ብ	ቦ			ቧ		
11	ቨ	ቩ	ቪ	ቫ	ቬ	ቭ	ቮ			ቯ		
12	ተ	ቱ	ቲ	ታ	ቴ	ት	ቶ			ቷ		
13	ቸ	ቹ	ቺ	ቻ	ቼ	ች	ቾ			ቿ		
14	ኀ	ኁ	ኂ	ኃ	ኄ	ኅ	ኆ	ኈ	ኊ	ኋ	ኌ	ኍ
15	ነ	ኑ	ኒ	ና	ኔ	ን	ኖ			ኗ		
16	ኘ	ኙ	ኚ	ኛ	ኜ	ኝ	ኞ			ኟ		
17	አ	ኡ	ኢ	ኣ	ኤ	እ	ኦ			ኧ		

18	ከ	ኩ	ኪ	ካ	ኬ	ክ	ኮ	ኰ	ኲ	ኳ	ኴ	ኵ
19	ኸ	ኹ	ኺ	ኻ	ኼ	ኽ	ኾ	ዀ	ዂ	ዃ	ዄ	ዅ
20	ወ	ዉ	ዊ	ዋ	ዌ	ው	ዎ					
21	ዐ	ዑ	ዒ	ዓ	ዔ	ዕ	ዖ					
22	ዘ	ዙ	ዚ	ዛ	ዜ	ዝ	ዞ			ዟ		
23	ዠ	ዡ	ዢ	ዣ	ዤ	ዥ	ዦ			ዧ		
24	የ	ዩ	ዪ	ያ	ዬ	ይ	ዮ					
25	ደ	ዱ	ዲ	ዳ	ዴ	ድ	ዶ			ዷ		
26	ጀ	ጁ	ጂ	ጃ	ጄ	ጅ	ጆ			ጇ		
27	ገ	ጉ	ጊ	ጋ	ጌ	ግ	ጎ	ጐ	ጒ	ጓ	ጔ	ጕ
28	ጠ	ጡ	ጢ	ጣ	ጤ	ጥ	ጦ			ጧ		
29	ጨ	ጩ	ጪ	ጫ	ጬ	ጭ	ጮ			ጯ		
30	ጰ	ጱ	ጲ	ጳ	ጴ	ጵ	ጶ			ጷ		
31	ጸ	ጹ	ጺ	ጻ	ጼ	ጽ	ጾ			ጿ		
32	ፀ	ፁ	ፂ	ፃ	ፄ	ፅ	ፆ					
33	ፈ	ፉ	ፊ	ፋ	ፌ	ፍ	ፎ			ፏ		
34	ፐ	ፑ	ፒ	ፓ	ፔ	ፕ	ፖ			ፗ		

Problems

- Non-resident tourist speak foreign languages hindering them to communicate with the local guide.
  - As a result, they look for bilingual guide or bilingual system.



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.

Related Works

	Author	Problem Solved	Performance	Research Direction
ASR	**Solomon Birhanu (2001)**	Investigate the Consonant-Vowel syllable recognition for the Amharic language	Recognition accuracy of 87.68 for Speaker Dependent and 72.75 Speaker independent	towards speaker independent recognition of speech and tuning the model to diverse environment including.
	Solomon Teferra (2005)	Develop a large vocabulary, speaker independent continuous Amharic speech recognition using syllable and triphone.	Recognition accuracy of 90.43 % for Syllable based and 91.31% for Tri-phone.	Improving the performance of syllable and triphone ASR for Large Vocabulary.
	Tachbelie, et. al, (2014)	Selecting acoustic, lexical and language modeling units for Amharic ASR	3% absolute WER reduction as a result of using syllable acoustic units in morpheme-based LM.	syllable AM in morpheme-based speech recognition to be tested for other morphologically rich language
SMT	**Sisay Adugna (2009)**	English-Afaan Oromo machine translation system to assist professional translators.	BLEU Score of 17.74%	possibility of exploring for other local language to make the information available in all local language.
	Mulu Gebreegziabher, et. al, (2012)	Preliminary experiments on English-Amharic statistical machine translation	BLEU score result is 35.32	The experiment have been extended to get a better result out of translation.
	Mulu Gebreegziabher, et. al, (2015)	Phoneme-based English-Amharic SMT	BLEU score of 37.53 for the phoneme-based EASMT system	Further improvement of English-Amharic SMT though different technique
TTS	**Henock Leulseged (2003)**	Concatenative Amharic TTS synthesis for Amharic Language	88% using Diphone and 75% for syllable based recognition	Overcome the problems of germinated sounds for syllable and diphone based synthesis.
	Sebsibe et. al, (2004)	Unit Selection Voice For Amharic Using Festvox	Perceptual evaluation of the synthesizer showed that the quality of the voice is good	Improving by proper selection of unit and optimal corpus which covers all basic units and variations.

Speech translation corpus

- A 20hr Amharic read speech prepared by Solomon T. et al, (2005) is used for training which is available at https://github.com/besacier/ALFFA_PUBLIC/tree/master/ASR
- Testing data BTEC 2009 available through IWSLT (Kessler, 2010).
  - English corpus is translated to Amharic to prepare parallel Amharic-English BTEC using a bilingual speaker.
  - Amharic speech data is recorded using Lig-Aikuma under normal office environment from eight native Amharic speakers (4 male and 4 female) with different age range.

Speech translation corpus

For Amharic ASR, a total of 10,875 taken from (Solomon T. et al, 2005) for training and 8112 sentences has been recorded under a normal working environment for testing.

- A total of 7.43hr read speech corpus collected with an average speech time of 3297 ms. Out of these utterance 98.54% of the speech data fall below 7sec.

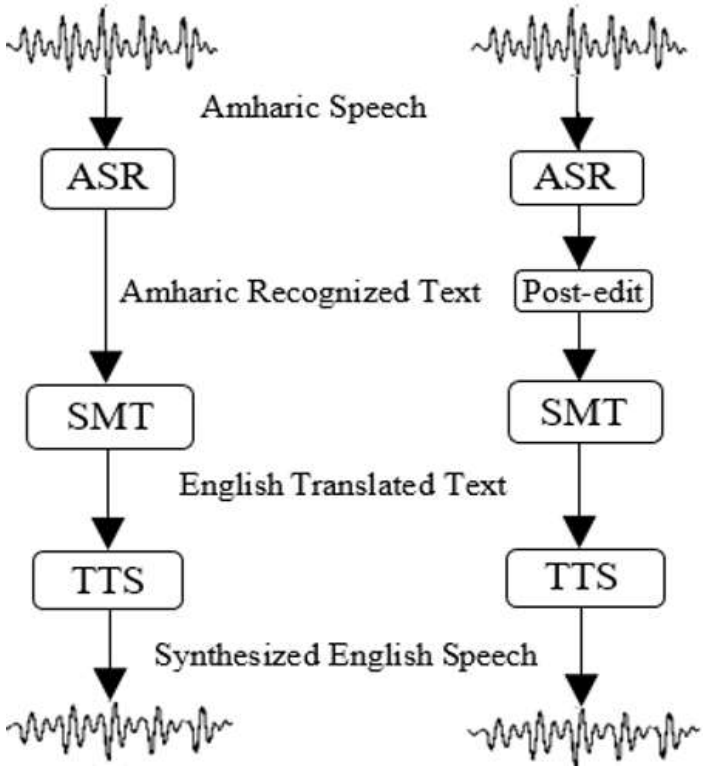
	Train	Test	LM	
			Word	Morpheme
Sentence	10,875	8,112	261,620	261,620
Token	145,404	50,906	4,223,835	5,773,282
Type	24,653	4,035	328,615	141,851

- For Amharic-English SMT, A total of 19472, 500 and 8112 sentence have been used for training, development and testing respectively.

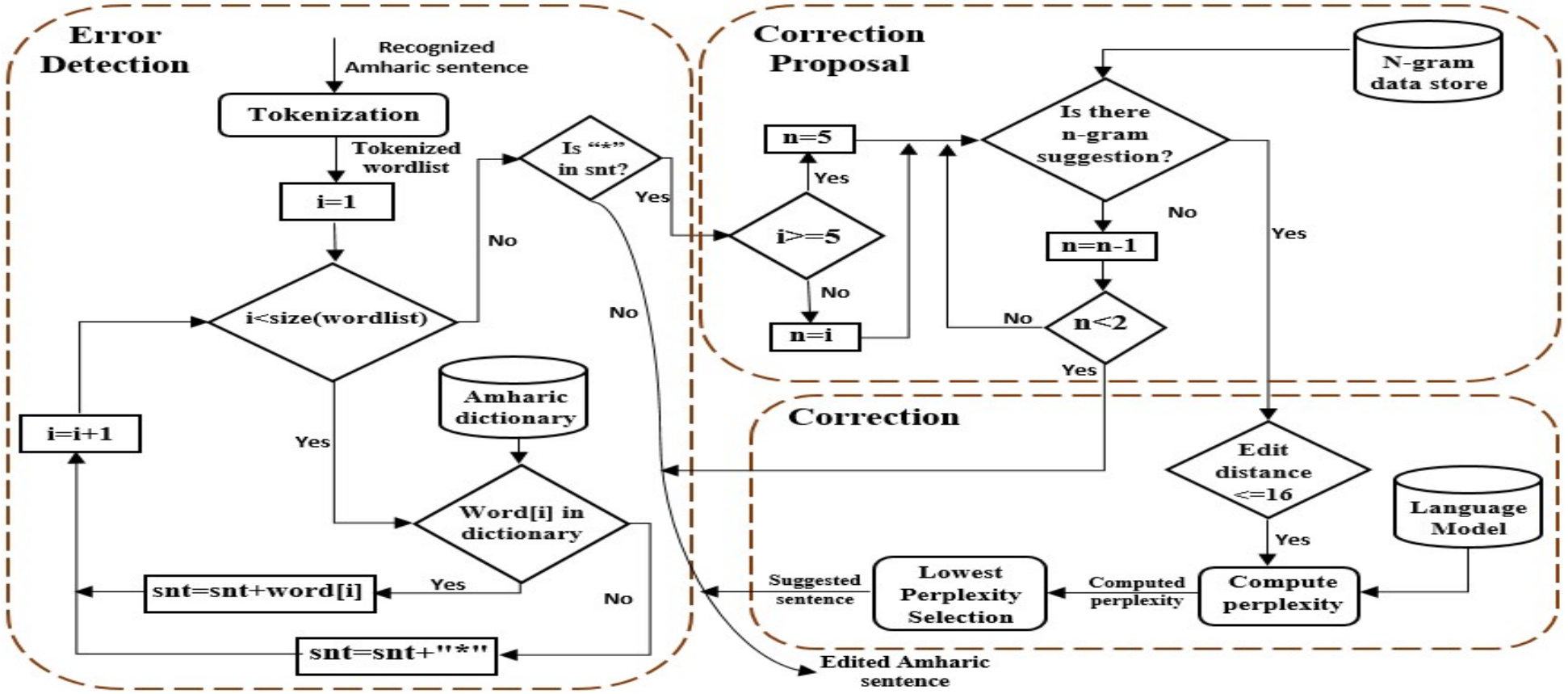
Language	Units		Train	Dev	Test
Amharic	Word	Sentence	19,472	500	8,112
		Token	107,049	2,795	37,288
		Type	18,650	1,470	4,168
	Morpheme	Sentence	19,472	500	8,112
		Token	145,419	3,828	50,906
		Type	15,679	1,621	4,035
English	Word	Sentence	19,472	500	8,112
		Token	157,550	4,024	55,,062
		Type	10,544	1,227	3,775

Speech Translation Components

- State-of-the-art of speech translation suggest to apply through the integration of cascading components; ASR, SMT and TTS
  - The output of a speech recognizer contains more and presents a variety of errors. These errors further propagates to the succeeding component which results in low performance.
  - Hence, in this study we propose an Amharic ASR post-editing module that can detect an error, identify possible suggestion and finally correct.
- Post-edit is conducted using a corpus based n-gram approach containing 681,910 sentences (11,514,557 tokens) of 582,150 type data crawled from web including news and magazine.
  - The n-gram has 5,057,112 bigram, 8,341,966 trigram, 9,276,600 quadrigram and 9,242,670 pentagram word sequences.



Post-edit


Error Detection
Recognized Amharic sentence
Tokenization
Tokenized wordlist
i=1
i<size(wordlist)
Is "*" in snt?
Yes
No
Amharic dictionary
Word[i] in dictionary
Yes
No
snt=snt+word[i]
snt=snt+"*"
i=i+1
Correction Proposal
N-gram data store
Is there n-gram suggestion?
Yes
No
n=n-1
n<2
No
Yes
n=5
Yes
i>=5
No
n=i
Correction
Edit distance <=16
Yes
Language Model
Compute perplexity
Computed perplexity
Lowest Perplexity Selection
Suggested sentence
Edited Amharic sentence

Possible suggestion list	Distance
የስጦታ እቃ ተስፋ አደርጋለሁ ብለዋል	5
የስጦታ እቃ ብየ ተስፋ አደርጋለሁ	5
የስጦታ እቃ አንደማይ ተስፋ አደርጋለሁ	5
የስጦታ እቃ ተስፋ አደርጋለሁ ይላሉ	5
የስጦታ እቃ ተስፋ አደርጋለሁ ለፊገግታ	5
የስጦታ እቃ ብዬ ተስፋ አደርጋለሁ	5
የስጦታ እቃ ተስፋ አደርጋለሁ ብሏል	5
የስጦታ እቃ ተስፋ አደርጋለሁ ይላሉም	5
ማንኛውም የስጦታ እቃ ተስፋ አደርጋለሁ	5
የስጦታ እቃ የሚሸጥበት ተስፋ አደርጋለሁ	6
የስጦታ እቃ አንደሚሆን ተስፋ አደርጋለሁ	6
የስጦታ እቃ አንደሚተካ ተስፋ አደርጋለሁ	6
የስጦታ እቃ አንደደረሰ ተስፋ አደርጋለሁ	6
የስጦታ እቃ አንደሚወጣ ተስፋ አደርጋለሁ	6
የስጦታ እቃ አንደሚመጣ ተስፋ አደርጋለሁ	6
የስጦታ እቃ አንደሚጠጣ ተስፋ አደርጋለሁ	6

Sample suggestion for "የስጦታ እቃ +ተዘነጉ ተስፋ አደርጋለሁ"
For equivalent English "Am hoping to buy some souvenirs"

No	Type	Sentence recognized and corrected
1	Raw	አራስን ጭምሮ ጉበኝት እገዛ+ ዋጋ ይኖረል
	Edited	አራስን ጭምሮ ጉበኝት በዛ ዋጋ ይኖረል
2	Raw	አባክዎን ተሸካሚ+ ጥሩት
	Edited	አባክዎን ተሸካሚውን ጥሩት
3	Raw	አባኮን+ ሌላ ፎጣ ማግኘት እችላለሁ
	Edited	አባክዎን ሌላ ፎጣ ማግኘት እችላለሁ
4	Raw	አባክዎን ሻንጣዎን ይክፈቱ+
	Edited	አባክዎን ሻንጣዎን ይክፈቱት
5	Raw	የስጦታ እቃ +ግዛት እያሰብኩ ነው
	Edited	የስጦታ እቃ ለመግዛት እያሰብኩ ነው
6	Raw	ይህ የጉዞ ላይ ህመም -ያጋጥም ይችላል
	Edited	ይህ የጉዞ ላይ ህመም ሊያጋጥም ይችላል
7	Raw	-ህ ባቡር ዶቨር በስንት ሰላት ይደርሳል
	Edited	ይህ ባቡር ዶቨር በስንት ሰላት ይደርሳል
8	Raw	ሻንጣዎን ይክፈቱ -ም
	Edited	ሻንጣዎን ይክፈቱ
9	Raw	አባክዎን ተሸካሚውን +ጥሩት
	Edited	አባክዎን ተሸካሚውን ይጥሩት

Sample raw and post-edited sentence

Experimental Result

	Amharic-English SMT	
	Word-Word	Morpheme-Word
BLEU	14.72	11.24

Amharic-English Statistical Machine Translation

		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

Preliminary experiment for Unit Selection for Amharic Speech Recognition (Melese et. al 2016)

Cont'd

	Before post edit		After post edit
	Word-Word	Morpheme-Word	Word-Word
Recognition Accuracy (%)	77.4	76.4	78.5
Translation in BLEU	12.83	6.29	13.08

Amharic Speech to English Text Translation

Concluding remarks

- Our experiments show that after post-editing the performance of translation improved by 1.95% (from 12.83 to 13.08) as a result of advancing ASR out put by 1.42%.
  - This implies that, minimizing broadcast error improves the accuracy of cascading components.
- The result we found from the experiments is promising to design well performing Amharic-English speech translation.
- Further works need to be done to apply post-editing at the translation stages of speech translation to reduce error broadcasting to the next stage.

አመሰግናለሁ
Thank you!