Deep Learning Based Text Translation and Summarization tool for hearing impaired using Indian Sign Language

Anurag Kumar Jha[1], Kabita Choudhary[1] and Dr. Sujala Shetty[1]
[1]*Birla Institute of Technology and Science, Pilani, Dubai Campus, DIAC Dubai 34055, UAE*
{f20200128, h20200905, sujala}@dubai.bits-pilani.ac.in

Keywords: Natural La-finalnguage Processing (NLP), Natural Language Generation (NLG), Bidirectional Auto-Regressive Transformers (BART), Multilingual Bidirectional Auto-Regressive Transformers (mBART), Signing Gesture Markup Language (SiGML), HamNoSys, Indian Sign Language, Transformer-based Highlights Extractor (THExt), Automatic text summarization (ATS), Machine Translation(MT)

Abstract: There have been multiple text conversions emerging with time but there has hardly been any work in the field of sign language. Even in the field of sign language multiple methods have been proposed to convert it into text via image detection, but due to the rarity of sign language corpus not much work has been put into text or speech to sign language. The proposed project intends to create a translation model to convert text or audio into sign language with its designated grammar. The process includes translation of any language to English followed by summarization of a big article or text, removal of stopwords, reordering the grammar form and stemming words into their root form. The translation is performed by mBART model, summarization is performed using BART model, conversion into animation is done via mapping words into a dictionary and replacing words by letters for unknown words. The paper uses HamNoSys (Regina et al., 1989), SiGML, BART, mBART and NLP to form the translation system. The paper aims to establish better means of communication with the deaf, dumb and people with hearing issues.

1. INTRODUCTION

1.1. Sign Language generation

A large chunk of information on the web consists of visual information which makes it less accessible for deaf people. The problem is less persistent amongst those who lost the hearing capacity in later stages of life but not for those born with the condition. There have been multiple studies (Conrad 1979, Holt et el. 1996 Allen 1986) which have depicted a poor reading power amongst deaf children when compared to their peers. Around only 25% of these children are capable of reading at a level above that of a 9-year-old hearing child.

Unlike conventional languages sign language utilises hand gestures, body movements and facial expressions for conveying any information. Translation systems exist between almost all existing languages using machine learning, but sign language stands as an exception. Even in sign language conversion of text into sign language has seen very little development over the years. The objective of the paper is to create a translation system which converts provided text into animated sign language (Indian Sign Language) using animated human figures.

There has not been much work in the field of sign language computerization and those done are mostly in American (Matthew et al., 2003) or British sign language (E´va Safar, 2003). The underlying architecture for these systems is mostly based on (R. San et al., 2004):

- **Direct translation of input into target words:** The biggest drawback of this system is that output is not grammatically correct and difficult to understand.
- **Statistical machine translation** which is ruled out in our case because of the lack of a large parallel corpus
- **Transfer based:** These include proper grammatical rules in place from proper translation from one language to another

As discussed before the existing methods have not been developed in terms of Indian Sign Language and our work makes an effort to fix this issue.
The proposed method is to create/collect video animation for the entire pool of ISL words which are around 1500 in total. The input text is manipulated to abide by the grammatical syntax of ISL and then mapped to the dictionary of the video animations. Words not present are broken into letters and shown one by one. These can be for things such as a name or a place.

The method includes displaying sign language using an avatar after translating SiGML to motion data.

A major challenge in the system is the conversion of one language to another with a completely different set of grammatical syntax in place.

Also, the feature of text summarization has been added to deliver large volume of data in smaller amount of time.

1.2. Text Summarization

There is enormous volume of textual content that is generated on the Internet and in the numerous archives of headlines, academic papers, government documentation, etc., automatic text summarization (ATS) has become more crucial. With the enormous amount of textual content, manual text summarizing takes a lot of time, effort, money, and sometimes impracticable (El-Kassas et al., 2021). A verity of task can be done using ATS like generation of summary for a scientific paper, news articles, creating summary of audio podcast etc.

A strategy to extracting highlights based on a recognized contextualized embedding architecture (Moreno et al., 2022)., especially the transformers, is known as a Transformer-based Highlights Extractor (THExt, in short). BART is a sequence-to-sequence model trained as a denoising autoencoder, It is applicable to many types of tasks like sequence classification (categorizing input text sentences or tokens), summarizing text, machine translation like translation between multiple language, question answering. Its pretraining has mainly two phases. Assign corrupted text with an arbitrary noise and sequence-to-sequence model is learned to rebuild the actual text. It is evaluated with a different noise approach as shown in Figure 1, like randomly shuffling the order of the original text and using a novel in-filling scheme (in this scheme length of span of text are replaced with mask token). It is an unsupervised language model which can be fine-tuned to a specific application like medical chatbots, generating summary of meeting, natural language to programming language, language translation etc. As it is already pretrained with very large amount of data, a small data set can be used to fine-tune it.

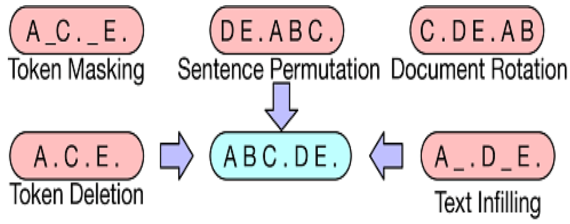

Figure 1. Transformations for noising the input for BART fine-tune

By using the BART transformer model, we can automate the text summarization task. Text summarization can be done in 2 ways.

- Extractive summarization: It provides the important text present in the given input.
- Abstractive summarization: It provides the actual summary of the given input. So, it is more challenging as it has to read complete text and understand the meaning of the text and provide us with the summary.

1.3. Text Translation

The ability to translate moderate languages has significantly improved due to training a universal translation system between different languages (Firat et al., 2016). Recent research (Arivazhagan et al., 2019; Fan et al., 2020) has also shown that multilingual translation models in a single model have a great potential for performance improvement. Using the pretrain and fine-tune approach common to NLP, recent pretrained multilingual sequence-to-sequence (seq2seq) models have made it simple to build neural machine translation (MT) systems (Liu et al., 2020). Pretrained models are excellent candidates for MT domain adaption tasks, where domain-specific bitext is typically less accessible than general bitext, because fine-tuning these models typically requires less data than is required for from-scratch translation models.

For the translation of text from any language to English, we are using mBART (Yuqing et al., 2020). It is a multilingual neural machine translation model. mBART supports up to 50 languages. Initially, the mBART model trains in 25 different languages and is fine-tuned for different tasks. For translation, it was fine-tuned on bitext (bilingual finetuning). Later in mBART50, the multilingual nature of the pretraining is used for fine-tuning the model.

2. BACKGROUND AND LITERATURE REVIEW

About 63 million people, which accounts for about 6.3% of the Indian population, suffer from hearing issues (Indian Census, 2011). A vast majority ranging from 76% to 89% of this population have no knowledge of sign language. The low literacy rate can be attributed to lack of work put into this field and absence of proper translation systems. Compared to this massive population who face this issue, the number of certified ISL translators in India is very less. This huge gap calls for being bridged.

One of the major works in the field of generating animations for english words has been done by J.R. Kennaway et al. However the work is centred around American Sign Language (ASL) and British sign language (BSL). Moreover, the aspect of grammar in sign language has not been taken much into account. We try to address both these issues in our proposed method by forming an algorithm to convert text into ISL grammar syntax and then to corresponding Indian sign language.

Another work more related to ISL was by Khushdeep et al. conversion of HamNoSys to SiGML for sign language animation. However, this method too failed to take into consideration the grammar aspect of sign language.

There has been some research (Pamela et al., 1999) into machine translation used in other sign languages which are:

- **Direct Translation:** The architecture works on word-to-word translation. The biggest drawback is the lack of context and meaning. There is no syntactical analysis and grammatical syntax is ignored. There is direct translation without any reordering which has a massive issue in the sense that ordering of sentences is completely different in sign language as compared to English. The format used in the English language is Subject-Verb-Object compared to Subject-Object-Verb in ISL.
- **Transfer based (Rule based):** In this architecture input is passed through syntactic and semantic transformation to convert it into intermediate text which is then converted into target language using linguistic rules.
- **Interlingua based:** It is an alternative to the above architectures and is based on Interlingua which is a language independent semantic structure formed by the semantic analysis of the input. This is then used to generate the target language.

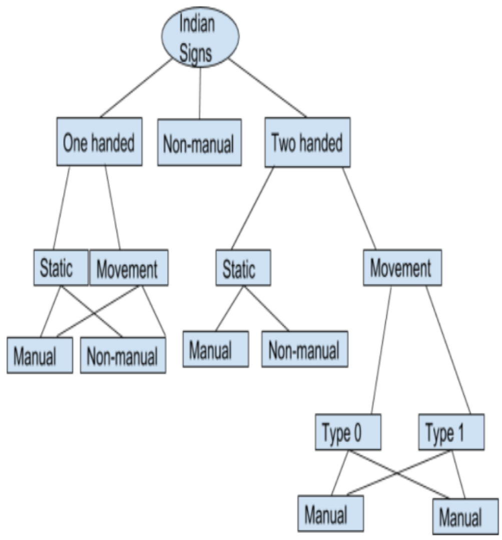

Figure 2: ISL Type Hierarchy(Type 0 refers to use of both hands and in type 1, use of one hand is dominant)(Ulrike, 2003)

In order to formulate an algorithm to translate English text to sign language, the following table of sign language details must be kept in mind:

Table 1: Important details of sign languages.

1	NOT the same all over the world
2	NOT just gestures and pantomime but do have their own grammar
3	Dictionary is smaller compared to other languages
4	Fingerspelling for unknown words
5	Adjectives are placed after the noun for most of the sign language
6	Never use suffixes
7	Always sign in present tense
8	Do not use articles
9	Do not uses I but uses me
10	Have no gerunds
11	Use of eyebrows and no-manual expression
12	Not been invented by hearing people

As mentioned before, sign language grammar is not similar to conventional languages and has certain distinct features which are:

Table 2: Features of sign languages

1	Number presentations are done with hand gestures for each hand.
2	Signs for family relationships are preceded by male or female.
3	In interrogative sentence, all the WH questions are places in the back of the sentence.
4	It also consists of pf many non-manual gestures such as mouth pattern, mouth gestures etc.
5	The past, present and future tense is presented by signs for before, then and after.

3. BLOCK DIAGRAM OF THE MODEL

The complete sequence for the conversion of any language to Indian sign language is shown in Figure 3.

The input text is first summarized to reduce the volume of content in case the amount of information is huge. This text is then transformed to match the syntax of Indian sign language. Once we have the sequence of words generated in ISL grammar these can be used to generate HamNoSys. HamNoSys is then converted into XML form known as SiGML. This is then processed further to produce the animation. The sequence of steps for animating have been inspired by Khushdeep et al., with the drawback of lacking ISL grammar resolved and feature of text summarization resolved.

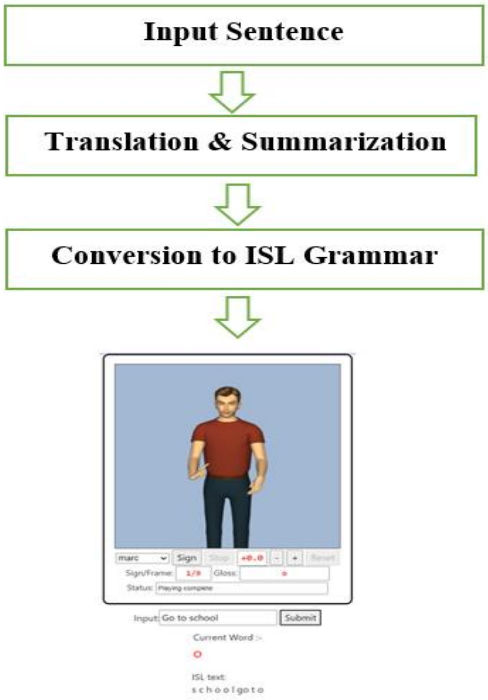

Figure 3: Block diagram for text Translation and Summarization in ISL

4. MEDIA USED

For transfer of text to sign language we will be using a dictionary having English words and their equivalent sign. The sign can be in the format of video, images or code signs. All these are compared in table 3.

Table 3: Comparison between different media representing sign languages.

Kind of media	Pros	Cons
Video Signs	• Realistic • Easy to create	• Time consuming to create • High memory consumption • Not supported by translation system
Pictures	• Very less memory consumption	• Time consuming to create • Not realistic as compared to videos • Not supported by translation system
Code Sign Language Text	• Minimal Memory Consumption • Supported by translation system as it is the written form and can be processed very easily	• Very difficult to read and understand • Required to be learnt

An analysis of the table gives the estimate that although videos are more time consuming to create and require a higher amount of memory, they are best suitable for easy understanding.
Hence the final output in the method proposed will be in the format of animated videos. To reduce the time to deliver these in case of huge amount of information we reduce it by text summarization.

5. METHODOLOGY

We have used http://www.indiansignlanguage.org/ to download video clips to map to our English word dictionary. These videos are then manually labelled.

Input is taken in the form of English text. Text parsing is done with the help of Stanford parser (Xu. H et al., 2011) which creates its grammatical phase structure. This is reordered in accordance with ISL. English Language grammar follows Subject-Verb-Object structure which is Subject-Object-Verb in the case of ISL. The irrelevant stop words are removed.

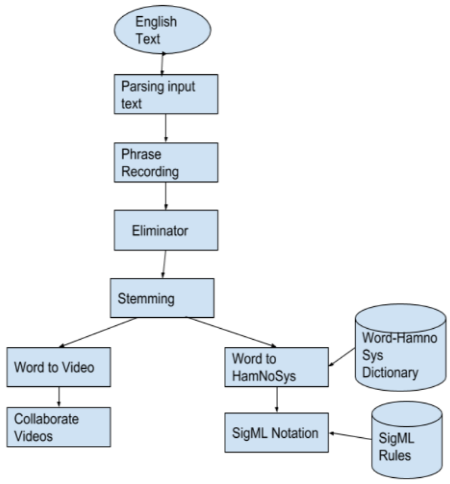

Figure 4: Algorithm for text to ISL

5.1. Solution Structure

- **Parsing the input:** To carry the translation of one language to another, both their grammatical structure must be known. Parsing is used to obtain this grammatical structure. For parsing the input Stanford parser (Xu.H et al., 2011) is used which breaks input into part of speech tagged text, CFG and type dependency representation.
- **Grammar rules for conversion from English to ISL:** The grammatical pattern of ISL varies from that of English language. ISL requires the verb patterns to be shifted after nouns as shown in the Table 4.
- **Eliminating stop words:** The English language includes words which don't have any meaning in ISl which include modals, foreign words, possessive ending, coordinating conjunction, determiners, adjectives, comparative and superlative, nouns plural, proper plural, particles, symbols, interjections and non-root verbs.
- **Stemming and Lemmatization:** Stemming is used to convert words into their root form using Port stemmer rules. Each of the words is checked in the dictionary, if it doesn't exist, it is tagged to its synonym.

Table 4: Grammatical reordering from English to ISL

Verb Pattern	Rule	Input Sentence	Parsed Sentence	Output Sentence
verb+object	VP TO NP	Go to school	(VP (VB Go)(TO to) (NP (NN school)))	School to go
Subject + verb	NP V	Birds fly	(NP (NNS birds)) (VP (VBP fly))	Birds fly
subject + verb + subject complement	NP V NP	His brother became a soldier	(NP (PRP$ his) (NN brother)) (VP (VBD became) (NP (DT a) (NN soldier)))	His brother a soldier became
subject + verb + direct object + preposition +prepositional object	NP V NP PP	She made coffee for all of us	(NP (PRP She)) (VP (VBD made) (NP (NN coffee)) (PP (IN for) (NP (NP (DT all)) (PP (IN of) (NP (PRP us))))))	she coffee for all of us made

- **Output generation:** Upon the execution of the above steps we receive the ISL equivalent of the input. It is then checked to corresponding keys in our text-animation dictionary. If a word is found it is displayed as video by passing it through a HamNoSys (Regina et al., 1989), generator, otherwise the word is broken and finger-spelling used to express the word.

5.2. Output Analysis

In order to judge the accuracy of the grammar and syntax a total of 100 English sentences were taken, 50 simple and 50 complex. They were passed through our proposed system and output validated by language expert.

The simple sentences fared an accuracy of 100% whereas the complex sentences were 96% accurate.

One drawback of the system however is to handle exclamation words like Oh! Alas! Since they don't have direct translation in ISL.

Also words having more than one parts of speech, example book has both verb and noun form cannot be handled well by the conversion system, since the original input sentence after being parsed and processed is converted into ISL grammar format which does not have a parsing structure, which indeed becomes difficult to identify the nouns and verbs from a given sentence, eventually causing the wrong form of the sentence being selected.

However, since less research has been carried out in this field, there are many pitfalls that needs to be worked on and some scenarios which needs to be thought of.

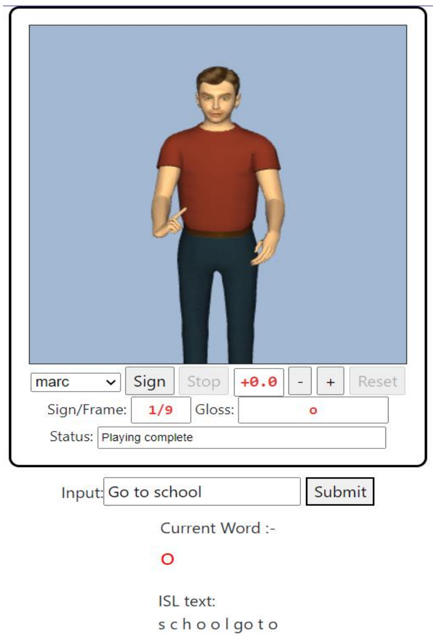

Figure 5: Conversion of English to ISL Grammar

6. BART ARCHITECTURE

The purpose of a NLP model is to not only understand the whole text given to read but also to understand the sequence of the text, like what comes before and after a token. The sequence of input tokens plays a very important role. For example, let’s say the statement is, “We are going to the theatre to watch a movie.” So if we mask “theatre ” by adding some noise and pass it to the model like:” We are going to the [mask]to watch a movie” The BART model should read thoroughly the provided text and also understands the sequence of words to Predict the masked word.

6.1. Architecture

As with GPT, BART modifies ReLU activation functions to GeLUs and initializes parameters from N, utilizing the conventional sequence-to-sequence Transformer architecture from. There are 6 layers in the encoder and decoder in basic model and 12 layers large model.
The architecture is very similar to that of BERT, with the following exceptions: (1) Each layer of the decoder performs cross-attention over the final

6.2. Pre-Training BART

In Bart, training is done by masking or currupting the data in different ways , and then optimizing the loss for recontraction .Cross entropy is calculated between the decoders output and original data.

Bart used different noising schemes for masking, such as:

- Token masking: Some random tokens are replaced with masks in a sentence and the model is trained to predict the single token based on the rest of the sequence.
- Token deletion: Some random tokens are deleted in a sentence and the model learns to find the deleted token and from where it was deleted.
- Text Infilling: Some contiguous tokens are deleted and replaced with a single mask and the model learns the missing token and the content.
- Sentence permutation: Sentences are permuted, and the model learns the logical implication of the sentence.
- Document Rotation: Here the documents are rearranged randomly. This helps the model to learn how the documents are arranged.

hidden layer of the encoder in addition (as in the transformer sequence-to-sequence model); and (2) Unlike BART, BERT employs an additional feed-forward network prior to wordprediction. BART

has approximately 10% more parameters overall than BERT mode of the same size.N, utilizing the conventional sequence-to-sequence Transformer architecture from. There are 6 layers in the encoder and decoder in basic model and 12 layers large model.
The architecture is very similar to that of BERT, with the following exceptions: (1) Each layer of the decoder performs cross-attention over the final hidden layer of the encoder in addition (as in the transformer sequence-to-sequence model); and (2) Unlike BART, BERT employs an additional feed-forward network prior to wordprediction. BART has approximately 10% more parameters overall than BERT mode of the same size.

For a NLP model it is imperative to completely read the sentence and understand each and every token in context of their sequence. Such a case, the input sequence can be interpreted by using a bi-directional approach.

Bart uses the bi-directional approach as shown in, Figure 5 to find the masked token. Hence the first part of the BART model is to use bi-directional encoder of BERT to find the best representation of its input sequence. In the second part It uses an autoregressive model which uses only past input sequences to predict the next word.

.

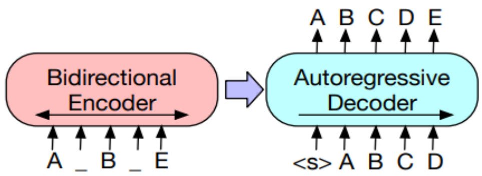

Figure 6: Semantic representation of Bart

Table 5: Example summaries from the BART model on News article

Text	BART Summary
Acid reflux troubles many people. It can lead to difficulties or problems like heartburn. How much stomach acid is produced affects you. To manage the acid reflux, it is necessary to choose the right food to eat. In an Instagram post, the nutritionist, Lavnit Bhatra, has shared 5 foods that can help control and avoid acidity. She writes, "Unhygienic eating can still make you experience acidity, while atrazide and other OTC products can exacerbate stomach acidity. Dietary changes will help control symptoms of acidity and relieve acid reflux. "	Nutritionist Lavnit Bhatra has shared 5 foods that can help control and avoid acidity. Unhygienic eating can still make you experience acidity, she writes. Dietary changes will help control symptoms of acidity and relieve acid reflux.
A message was sent to Prime Minister Narendra Modi through the first major political rally in Srinagar after the redrawal of statehood and special status from Jammu and Kashmir to restore statehood and hold elections without delay. Thousands of people attended the rally organized by Jammu and Kashmir's own party (JKAP) on Saturday morning at Srinagar' s Sher - e - Kashmir International Cricket Stadium. The party's chief, Altaf Bukhari, said he wanted a large crowd to fulfil the promises made to the prime minister.	Thousands of people attended the rally organized by Jammu and Kashmir's own party (JKAP) The party's chief, Altaf Bukhari, said he wanted a large crowd to fulfil the promises made to the prime minister.
The participants, as per an official statement, will discuss and deliberate on ways to increase the number of women in higher echelons of teaching, research and industry, along with trying to find ways to provide women with equal access to STEM (Science, Technology, Engineering, Mathematics) education, research opportunities and economic participation.A special programme to showcase the contribution of women in science and technology will also be held, which will also witness lectures by renowned women scientists, it added.	The participants will discuss and deliberate on ways to increase the number of women in higher echelons of teaching, research and industry. They will also try to find ways to provide women with equal access to STEM education, research opportunities and economic participation.

HuggingFace provides us the platform to use Bart model for both pretrained and fine-tuned version. We can also use the API for models to summarize text and translate in some other language as well. For our model we are using the "facebook-bart-large-cnn" model for text summarization and the "facebook/mbart-large-50-many-to-many-mmt" model for text translation.

7. DATASET SUMMARY

7.1. Bart

Bart is fine-tuned with CNN, Daily mail dataset which has 300k news articles and all are unique sets, those are written by journalists at CNN and Daily Mail. It supports both type of summarization that are abstractive and extractive summarization.

7.2. mBart

Initially the mBART.cc25 checkpoint (Yinhan et al., 2020) available in the fairseq library is (Myle et al., 2019) to continue the pretraining process. The monolingual data from XLMR (Alexis et al., 2019) is used to extend the pretraining to a set of 25 languages in addition to the 25 languages mBART model. To be consistent mBART, 250K sentence piece model which was trained using monolingual data for 100 languages from XLMR is used, and thus already supports languages beyond the original 25 mBART was trained on. For pre-training, mBART50 was trained for an additional 300K updates with a batch size of 1700 tokens.

- By using huggingface api, we are translating text into English as shown in Figure 7.

```
import requests
r = requests.post\
(url="https://kabita-choudhary-translationmodel.\
hf.space/api/predict",
 json={"data": ['''संयुक्त राष्ट्र के प्रमुख का \
  कहना है कि सीरिया \
 में कोई सैन्य समाधान नहीं है''',"Hindi"
,"English" ]})
r.json()
```

```
{'data': ['The head of the United Nations says there is no military solution in
Syria'],
 'is_generating': False,
 'duration': 27.792309284210205,
 'average_duration': 38.51862472663691}
```

Figure 7: Sample of text-translation from Hindi to English

- By using huggingface api, we are summarizing text .In Table 5 there are some examples for text summarization that is done by using BART.

8. CONCLUSION

We demonstrate that we can translate and summarize any text to English and transform the generated text into Indian sign language. This system can be used to benefit the hard hearing people having communication difficulties. Additionally due to the feature of text summarization a large volume of information can be delivered in lesser time which helps them to keep pace with everyone else. For translation and summarization, the language generation model is used from the huggingface platform. The future scope of this project is gathering information from an audio and video file and converting it into sign language.

9. REFERENCES

Regina Leven, Heiko Zienert, Thomas Hamke, and Jan Henning Siegmund Prillwitz.(1989). *HamNoSys Version 2.0: Hamburg Notation System for Sign Languages: An Introductory Guide in Proceedings of International Studies on Sign Language and Communication of the Deaf.* vol. 5, Hamburg, Germany.

El-Kassas, W.S.Salama, C.R.Rafea, A.A Mohamed ,H.K.(2022) *Automatic text summarization: A comprehensive survey.* Expert Syst. Appl. 2021, 165, 113679

Moreno La Quatra, Luca Cagliero. (2022). *Transformer-based highlights extraction from scientific papers.* Knowledge-Based Syst. 2022, 252, 109382.

Orhan Firat, Kyunghyun Cho, and Yoshua Bengio. (2016). *Multi-way, multilingual neural machine translation with a shared attention mechanism.* In NAACL.

Naveen Arivazhagan, Ankur Bapna, Orhan Firat, Dmitry Lepikhin, Melvin Johnson, Maxim Krikun, Mia Xu Chen, Yuan Cao, George Foster, Colin Cherry, *(2019). Massively multilingual neural machine translation in the wild: Findings and challenges.* arXiv preprint arXiv:1907.05019.

Ulrike Zeshan. (2003) . *Indo-Pakistani Sign Language Grammar:A Typographical Outline*. Sign Language Studies, vol. 3, no. 2, pp. 157-212,.

E´va Safar, Ian Marshall. (2003). *A Prototype Text to British Sign Language (BSL) Translation. System*Proceedings of the 41st Annual Meeting on Association for Computational Linguistics - Volume 2 (ACL '03), vol. 2, pp. 113-116

Matthew Huenerfauth. (2003). *A Survey and Critique of American Sign Language Natural Language Generation and Machine Translation Systems*. Computer and Information Sciences, University of Pennsylvania, Philadelphia, September.

Pamela W.Jordan, John W.Benoit, Bonnie J.Dorr. (1999). *A Survey of Current Paradigms in Machine Translation, Advances in Computers*. vol. 49, pp. 1-68

R. San-segundo , J. M. Montero , J. Macías-guarasa , R. Córdoba , J. Ferreiros and J. M. Pardo. (2004). *Generating Gestures from Speech*

Xu, H, Abdel Rahman, S Jiang, M Fan, J.W. and Huang Y, (2011). November. *An initial study of full parsing of clinical text using the Stanford Parser*. In 2011 IEEE International Conference on Bioinformatics and Biomedicine Workshops (BIBMW) (pp. 607-614). IEEE.

Yinhan Liu, Jiatao Gu, Naman Goyal, Xian Li, Sergey Edunov, Marjan Ghazvininejad, Mike Lewis, and Luke Zettlemoyer, (2020). *Multilingual denoising pre-training for neural machine translation*.

Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdel Rahman Mohamed, Omer Levy, Ves Stoyanov, Luke Zettlemoyer. (2019). *BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension*,

Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova. (2019). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*, Google AI Language.

Colin Raffel,Noam, Shazeer,Adam, Roberts, Katherine Lee, Sharan Narang ,Michael Matena , Yanqi Zhou , Wei Li , Peter J. Liu . (2020). *Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer*

Myle Ott, Sergey Edunov, Alexei Baevski, Angela Fan, Sam Gross, Nathan Ng, David Grangier, and Michael Auli. (2019) *FAIRSEQ: A fast, extensible toolkit for sequence modeling.* In North American Association for Computational Linguistics (NAACL): System Demonstrations.

Alexis Conneau, Kartikay Khandelwal, Naman Goyal, Vishrav Chaudhary, Guillaume Wenzek, Francisco Guzman, Edouard Grave, Myle Ott, Luke Zettle- ´ moyer, and Veselin Stoyanov. (2019). *Unsupervised cross-lingual representation learning at scale*.

Yuqing Tang, Chau Tran, Xian Li, Peng-Jen Chen, Naman Goyal, Vishrav Chaudhary, Jiatao Gu, Angela Fan. (2020). *Multilingual Translation with Extensible Multilingual Pretraining and Finetuning*.

Lalitha Natraj, Sujala D. Shetty (2019) A Translation System That Converts English Text to American Sign Language Enhanced with Deep Learning Modules, International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8 Issue-12,

J.R. Kennaway et al., "Providing signed content on the Internet by synthesized animation," ACM Trans. Comput.-Hum. Interact. 14, 3 (2007), 15–es, https://doi.org/10.1145/1279700.1279705.