INTERNATIONAL JOURNAL
FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Sentence Formation Using NLP on the Basis of American Sign Language

Kevin Gajjar[1], Aman Agrawal[2], Arran Gonsalves[3], Gargi Singh[4]

[1,2,3,4]*B.Tech, Student, Department of Computer Science and Engineering, Medi-Caps University, Indore, India.*

Abstract: Natural language processing at its core is a method to understand, process and utilize human language that helps in the development of different tools. One such field where this tool can be used for is sign language which is the primary method of communication for the impaired which usually requires a translator to interpret the meaning for those who do not have the knowledge. This paper aims to propose a method that can translate recognized signed words from ASL into proper grammatically correct English sentences with the use of different NLP techniques and parsing it into a Lexicalized Tree Adjoining Grammar (LTAG) using LALR parser. This approach uses a LTAG which is a lexicon that is organized on grammar and vocabulary of the English language and connects in a group of trees. The output matrix of words from the sign recognition is used as an input for the Parts of Speech (POS) tagger that will be parsed into the grammar tree giving a proper English sentence which will be verified by using Language Tool to check the grammar of the final sentence.

Keywords: Sentence Formation, American Sign Language, Natural Language Processing, Sign Language recognition, Grammar mapping

I. INTRODUCTION

Sign language is a visual means of communicating that is most natural and expressive for impaired people. It is an independent language and is used all over the world with native differences, similar to spoken languages [12]. Since there is not any international standard for sign language it makes it difficult for everyone to learn thus creating a gap in communication between communities [1]. To bridge this gap there is plenty of work being carried out in the field of Computer Science such as creating databases, developing recognition models and making it easier to access these resources to reduce the use of a human interpreter. This research work will focus on sentence formation from the words extracted after sign recognition for the impaired community.

Natural language Processing (NLP) is concerned with giving computers the ability to understand human language in the same way we do. It combines linguistics and rule-based human language to achieve a system that can understand and generate language used by humans [2]. The first step, Parts of speech (POS) tagging is a method of NLP that retrieves grammatical information about the words depending on the definition and their context which will be used to tag the matrix of words extracted from sign recognition. The second step is parsing which will be used in determining the syntactic structure of the sentence.

Since both spoken language and ASL have different grammar rules the words drawn from the recognition would need to be translated using rules to accurately parse the words based on the syntactic structure. The prominent difference between the two is English relies on Subject-Verb-Object (SVO) sentence structure, while ASL more frequently uses Topic-Comment structure. ASL also places temporal markers at the beginning of sentences, while English places them at the end. Additionally, verbs in ASL can be modified based on location or target [3]. To translate the grammar will need a well-defined grammar map. Grammar itself is a complex language phenomenon and usually language specific but world languages share similarities and a vast amount of differences [17]. Traditional Grammar is more concerned about the structure and formation of sentences while modern grammar leans towards the connections that exist in the structure of a language. The translation between the two languages is achieved through Lexicalized tree Adjoining grammar (LTAG) which consists of lexicons that make up parts of speech along with the tenses and syntactic structure for sentence planning [4]. Parsing will take place on a map with LALR (Lookahead LR) parser which can handle large grammar classes while being reliable in giving accurate results. This will be discussed further section of the paper.

The origination of the paper is as follows, section II contains the literature survey. Section III describes the proposed methodology. Section IV discusses the result. Finally, section V will conclude the work.

II. LITERATURE SURVEY

[1] Ankit Ojha, Ayush Pandey, Shubham Maurya, Abhishek Thakur, proposed a desktop application that uses a computers webcam to capture a person signing gestures for American sign language (ASL), and translate it into corresponding text and speech in real time using Convolutional neural network (CNN) A finger spelling sign language translator is obtained which has an accuracy of 95%.

[4] Anne Abeillé, Kathleen Bishop, Sharon Cote, and Yves Schabes, discusses a sizable grammar for English written in the Tree Adjoining grammar (TAG) formalism. The grammar uses a TAG that is both lexicalized and feature-based. It describes how to deal with semantic non compositionality in verb-particle combinations, light verb constructions and idioms, without losing the internal syntactic composition of these structures.

[6] R.Agarwal, Sumeet & Agrawal, Sagarkumar & Latif, Akhtar, presents the development of an application which will enable a hassle free communication between physically disabled person and normal person. The application is responsible for the conversion of hand gestures into a meaningful sentence which can be read out by a normal person. In this approach, gestures based on Indian sign language are used. The major stapeps associated with the application are gesture recognition and natural language processing as most important module.

[12] Sugandhi, Kumar, P., & Kaur, presented a system features a rich corpus of English words and commonly used sentences. It consists of components such as an ISL parser, the Hamburg Notation System, the Signing Gesture Mark-up Language, and 3D avatar animation for generating SL according to ISL grammar. The results proved that the proposed system is highly efficient and achieves an average score of accuracy. The performance of proposed system has also been evaluated using the BiLingual Evaluation Understudy score, which results in 0.95 accuracy.

[16] Stone, Matthew and Christine Doran, present a Lexicalized Tree Adjoining Grammar algorithm for simultaneously building both the syntax and semantics of a sentence (LTAG). This method captures the connection between pragmatic and syntactic constraints on descriptions in a sentence, as well as the inferential interactions between several descriptions in a sentence, in a natural and beautiful manner. At the same time, it makes contextually appropriate syntactic choices based on linguistically motivated declarative definitions of syntactic constructions' discourse functions..

[17] Ishola, Ọlájídé, discusses what a language at its core, how grammar is different from past till present and NLP intersection for language with computers and with the help of concise grammar rules, machines can process human languages.

III. METHODOLOGY

The impaired and normal users are the two end users of this application.

A. Sign Recognition

The given fig.1 shows the workflow of the model First performs hand gestures will be captured through a camera, and the images obtained are compared against the database American Sign Language Lexicon Video Dataset (ASLLVD) which consists of >3,300 sign videos for almost 9,800 tokens [5]. The signs are recognized using a CNN model which has an accuracy of 89.2%. Since there exists a difference in grammar in sign language and spoken English we need to add proper syntax and fill the stop words that do not exist in sign language.
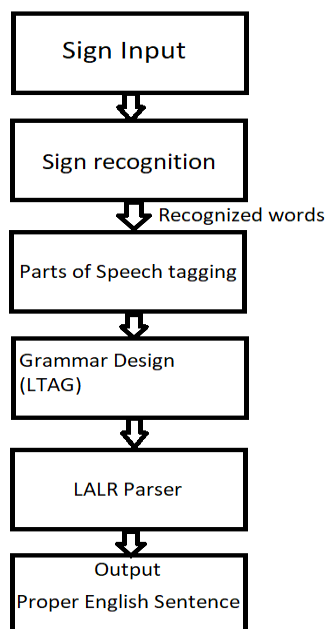
Sign Input
↓
Sign recognition
↓ Recognized words
Parts of Speech tagging
↓
Grammar Design (LTAG)
↓
LALR Parser
↓
Output
Proper English Sentence

Fig. 1: Flow of work

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue IV Apr 2022- Available at www.ijraset.com

The procedure can be better understood by using an example.

Consider a sentence in English "My teacher gave me an apple and I will be sharing it with my friends."

In sign language it would be written as "Teacher give me apple and I share with friends."

Here the difference in grammar is clearly stated like self-referring pronouns like "me, my, I" are expressed similarly expressed in sign, the tense in the first part is in past tense and in the second part is in continuous future but in sign language, it is not clear without specifically stating tenses.

B. Parts of Speech Tagging

The words that are recognized are then passed to the NLP engine where each word is given some description and this process is called tagging [6]. POS tagging is divided into two stages tokenization and tagging. The raw text output gets tokenized and then each word is tagged associated part of speech. The fig.2 shows the Penn Treebank tags that are used to tag the words.

CC	Coordinating conj.	TO	infinitival *to*
CD	Cardinal number	UH	Interjection
DT	Determiner	VB	Verb, base form
EX	Existential there	VBD	Verb, past tense
FW	Foreign word	VBG	Verb, gerund/present pple
IN	Preposition	VBN	Verb, past participle
JJ	Adjective	VBP	Verb, non-3rd ps. sg. present
JJR	Adjective, comparative	VBZ	Verb, 3rd ps. sg. present
JJS	Adjective, superlative	WDT	Wh-determiner
LS	List item marker	WP	Wh-pronoun
MD	Modal	WP$	Possessive *wh*-pronoun
NN	Noun, singular or mass	WRB	Wh-adverb
NNS	Noun, plural	#	Pound sign
NNP	Proper noun, singular	$	Dollar sign
NNPS	Proper noun, plural	.	Sentence-final punctuation
PDT	Predeterminer	,	Comma
POS	Possessive ending	:	Colon, semi-colon
PRP	Personal pronoun	(	Left bracket character
PP$	Possessive pronoun	)	Right bracket character
RB	Adverb	"	Straight double quote
RBR	Adverb, comparative	'	Left open single quote
RBS	Adverb, superlative	"	Left open double quote
RP	Particle	'	Right close single quote
SYM	Symbol	"	Right close double quote

Fig. 2: Types of POS tags [7]

C. Grammar Design

To construct a valid sentence we require a set of rules of the particular language these rules can be defined under Design grammar. It enables the formal representation of a vocabulary and rules that describe how designs can be synthesized [6]. The simplest and easiest form of grammar is context free grammar (CFG) [9]. It is easy to deal with and write CFG for natural languages as it is comprised of written rules which can be written in the form of notation:

G = {V, T, P, S}

Where,

V = final set of non-terminal symbols

T = final set of terminal symbols

P = set of production rules

S = the start symbol

The Context free grammar can be represented in the form of trees which is used to Lexicalized tree adjoining grammar. The formation of LTAG from CFG is possible when provided we integrate constraints from syntax, semantics and pragmatics. The syntax is giving a sentence structure and rules of grammar. Semantics is providing the meaning of the sentence and pragmatics is taking the meaning of the sentence within a certain context.

A TAG grammar consists of a finite set of elementary trees, which can be combined by these substitutions and adjoining operations to produce derived trees recognized by the grammar. In substitution, the root of the first tree is identified with a leaf of the second tree, called the substitution site [4]. Adjoining is a more complicated splicing operation, where the first tree replaces the subtree of the second tree rooted at a node called the adjunction site; that subtree is then substituted back into the first tree at a distinguished leaf called the FOOT node. Our grammar structure for translation incorporates two additional principles. First, the grammar is lexicalized each elementary structure in the grammar contains at least one lexical item. Second, our trees include features that add articles of speech and characteristics which do not exist in ASL but are present in English. For example, American Sign Language does not consist of gestures for verbs and articles like is, am, are, was, were, for. Therefore to construct a meaningful sentence from sign gestures there is the need for a language processor which will insert these features in appropriate places.

D. Parsing

In the fig.3, it explains LALR bottom up parsing. The inputs to the parser are the parts of speech of the English language namely verb, noun, adjective, pronoun, conjunction, etc. Since stop words and articles are not included in the sign language we will be inserting them into appropriate places using the lexical grammar tree for constructing a meaningful sentence. Consider the impaired user inputs. Here "Teacher give me apple and me share-with friends" The inputs to the parser are the parts of speech of English language viz. Pronoun, noun, adjective, verb, conjunction, etc. This word matrix S will be then tokenized into words S1, S2, S3… each Start symbol consists of its own separate grammar and the words are parsed into their respective position. This is done by the LALR parser which uses to separate and analyze the text according to the production rules specified in the formal grammar. Since the Signed inputs lack some verbs and articles of speech so to construct a meaningful sentence LTAG will be used to insert words into the appropriate places. After parsing the output comes to be of varying results show in table 1.
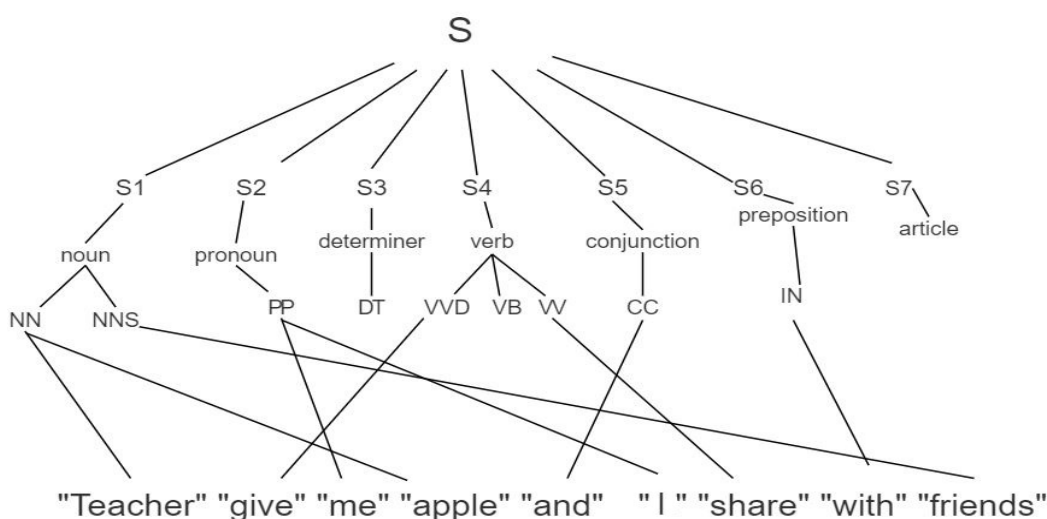

Fig. 3: Derivation of mapping for Teacher give me apple and I share with friends

E. Grammar Check

After the sentence is formed we will verify the integrity and check where it contains any error. To accomplish this by using LanguageTool [10] in python which is an open-source grammar tool library that allows us to make, detect grammar errors and spelling mistakes through a python script.

TABLE I

RESULT DERIVED FOR THE INPUT "TEACHER GIVE ME APPLE AND ME SHARE-WITH FRIENDS"

Sign Input	English Text Output
Teacher give me apple and me share-with friends	My teacher gives me an apple and I will share it with my friends
	My teacher gave me an apple and I have shared it with my friends
	My teacher gives me an apple and I am going to share with my friends

IV. RESULT AND DISCUSSION

The method discussed above is used for translating American Sign Language into English with the core focus on sentence formation using Natural language processing. Fig 3 shows how a sentence is parsed and a proper English sentence is formed.

As it is difficult to handle different types of sentences and tenses the method was not able to accurately construct a sentence with much accuracy. This issue can be looked upon by asking the users to provide an input specifically for tense in further versions of the method. A grammar-developed sentence generates only for those inputs which provide some structure as per the grammar rules. Hence in some cases meaningful sentences may not be constructed.

V. CONCLUSION

The major objective of this work is to give the significance of sign language as a language and focus on methods available for Sentence formation using Natural Language processing. As the method focus on grammar not only it can be used for sentence formation in English but in other languages as well, as long as we can provide a proper grammatical structure and tagging for other sign languages. The method can be implemented to work as an interpreter for the communication between a deaf and dumb person and a normal person.

REFERENCES

[1] Ankit Ojha, Ayush Pandey, Shubham Maurya, Abhishek Thakur, Dr. Dayananda P, 2020, Sign Language to Text and Speech Translation in Real Time Using Convolutional Neural Network, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) NCAIT – 2020 (Volume 8 – Issue 15),

[2] Liddy, E.D. 2001. Natural Language Processing. In Encyclopedia of Library and Information Science, 2nd Ed. NY. Marcel Decker, Inc.

[3] https://www.nidcd.nih.gov/health/american-sign-language

[4] Anne Abeillé, Kathleen Bishop, Sharon Cote, and Yves Schabes, "A Lexicalized Tree Adjoining Grammar for English", . March 1990.

[5] http://www.bu.edu/asllrp/av/dai-asllvd.html

[6] R.Agarwal, Sumeet & Agrawal, Sagarkumar & Latif, Akhtar. (2015). Sentence Formation in NLP Engine on the Basis of Indian Sign Language using Hand Gestures. International Journal of Computer Applications. 116. 18-22. 10.5120/20428-2757.

[7] Taylor, Ann, Mitchell P. Marcus and Beatrice Santorini. "The Penn Treebank: An Overview." (2003).

[8] https://www.ling.upenn.edu/courses/Fall_2003/ling001/penn_treebank_pos.html

[9] Shrawankar, Urmila & Dixit, Sayli. (2016). Framing Sentences from Sign Language symbols using NLP.

[10] https://pypi.org/project/language-check/

[11] Sampada S. Wazalwar & Urmila Shrawankar (2017) Interpretation of signlanguage into English using NLP techniques, Journal of Information and Optimization Sciences,38:6, 895-910, DOI: 10.1080/02522667.2017.1372136

[12] Sugandhi, Kumar, P., & Kaur, S. (2020). Sign Language Generation System Based on Indian Sign Language Grammar. ACM Transactions on Asian and Low-Resource Language Information Processing, 19(4), 1–26. doi:10.1145/3384202

[13] Bragg, Danielle & Koller, Oscar & Bellard, Mary & Berke, Larwan & Boudrealt, Patrick & Braffort, Annelies & Caselli, Naomi & Huenerfauth, Matt & Kacorri, Hernisa & Verhoef, Tessa & Vogler, Christian & Morris, Meredith. (2019). Sign Language Recognition, Generation, and Translation: An Interdisciplinary Perspective.

[14] https://www.cse.scu.edu/~m1wang/projects/NLP_English2IndianSignLanuguage_18w.pdf

[15] Kulkarni, Aishwarya. "Dynamic sign language translating system using deep learning and natural language processing." (2021).

[16] Stone, Matthew and Christine Doran. "Sentence Planning as Description Using Tree Adjoining Grammar." ACL (1997). DOI:https://doi.org/10.3115/976909.979643

[17] Ishola, Ọlájídé. (2019). Universal Dependencies for Yorùbá. 10.13140/RG.2.2.22165.14564.

[18] Kanakaraddi, Suvarna G; Ramaswamy, V (2014). [IEEE 2014 IEEE International Advance Computing Conference (IACC) - Gurgaon, India (2014.02.21-2014.02.22)] 2014 IEEE International Advance Computing Conference (IACC) - Natural language parsing using Fuzzy Simple LR (FSLR) parser. , (), 1337–1341. doi:10.1109/IAdCC.2014.6779521

[19] DeRemer, Frank, and Thomas Pennello. "Efficient computation of LALR (1) look-ahead sets." ACM Transactions on Programming Languages and Systems (TOPLAS) 4, no. 4 (1982): 615-649.

[20] Deborah I. Fels, Jan Richards, Jim Hardman, Sima Soudian, and Charles Silverman. 2004. American sign language of the web. In <i>CHI '04 Extended Abstracts on Human Factors in Computing Systems</i> (<i>CHI EA '04</i>). Association for Computing Machinery, New York, NY, USA, 1111–1114. DOI:https://doi.org/10.1145/985921.986001

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089   (24*7 Support on Whatsapp)