Contents

[Introduction 4](#_Toc461114694)

[Background 4](#_Toc461114695)

[Motivation 4](#_Toc461114696)

[Problem statement 5](#_Toc461114697)

[Language Analysis 9](#_Toc461114698)

[Natural Language 9](#_Toc461114699)

[Grammatical Structure for English Language: 9](#_Toc461114700)

[Sign Language 14](#_Toc461114701)

[Translation 14](#_Toc461114702)

[Machine Translation Architecture 14](#_Toc461114703)

[Rules and it's need 17](#_Toc461114704)

[Grammatical Structure for Pakistan Sign Language: 17](#_Toc461114705)

[Related Work 19](#_Toc461114706)

[Systems under consideration 19](#_Toc461114707)

[Proposed Methodology: Part1 24](#_Toc461114708)

[Overview of the Proposed System 25](#_Toc461114709)

[Data Gathering 27](#_Toc461114710)

[Deduced Observation: 30](#_Toc461114711)

[Our Proposed Grammar for PSL 31](#_Toc461114712)

[Proposed Methodology: Part2 33](#_Toc461114713)

[Grammar Processing of a Natural Language 33](#_Toc461114714)

[Analysis of English Language Grammar 33](#_Toc461114715)

[Getting Output from Natural Language Parser 34](#_Toc461114716)

[Saving Grammar Information in a Tree Data-Structure 35](#_Toc461114717)

[Getting Word Dependencies from Typed Dependencies Tool 36](#_Toc461114718)

[Annotated Tree 37](#_Toc461114719)

[Tense Identification 39](#_Toc461114720)

[Sentence Structure Identification 39](#_Toc461114721)

[Algorithm for Structure Type Identification 41](#_Toc461114722)

[Translation of Compound, Complex and Complex Compound Sentences 41](#_Toc461114723)

[Renaming of English Tree according to PSL Grammar 43](#_Toc461114724)

[Construction of PSL Tree 44](#_Toc461114725)

[Translated Sentence 50](#_Toc461114726)

[Evaluation 52](#_Toc461114727)

**List of Figures**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

[Figure 1: Over View of Proposed System 6](#_Toc461009320)

[Figure 2: Stanford Parser Output 7](#_Toc461009321)

[Figure 3: Updating in Stanford Tree 7](#_Toc461009322)

[Figure 4: PSL Equivalent Sentence 8](#_Toc461009323)

[Figure 5: Machine Translation Architecture 10](#_Toc461009324)

[Figure 6: English and Chinese Equivalent Sentence 11](#_Toc461009325)

[Figure 7: English and PSL Equivalent Sentence 12](#_Toc461009326)

[Figure 8: Types of English Sentences in Terms of Structure 13](#_Toc461009327)

[Figure 9: Types of English Sentences in Terms of Function 15](#_Toc461009328)

[Figure 10: Types of English Sentences in Terms of Tense 17](#_Toc461009329)

[Figure 11: TESSA Proposed System 19](file:///C:\Users\12003065137\Downloads\FYP_doc.docx#_Toc461009330)

[Figure 12: Major Components of English to ASL by Grieve-Smith 20](file:///C:\Users\12003065137\Downloads\FYP_doc.docx#_Toc461009331)

[Figure 13: Proposed System of English to ISL 21](#_Toc461009332)

[Figure 14: Pros and Cons of deferent translation systems 23](#_Toc461009333)

[Figure 15: Proposed System 24](#_Toc461009334)

[Figure 16: PSL structure 27](#_Toc461009335)

[Figure 17: English Sentence Categories We Catered 28](#_Toc461009336)

[Figure 18: bar chart for followed word-order in Pakistan 29](#_Toc461009337)

[Figure 19: English and its equivalent PSL sentence 29](#_Toc461009338)

[Figure 20: English Tree constructed from NLP Output 36](#_Toc461009339)

[Figure 21: Dependency Graph 37](#_Toc461009340)

[Figure 22: Input Tree 38](#_Toc461009341)

[Figure 23: Output Tree (Annotated) 38](#_Toc461009342)

[Figure 24: Parsing of a Compound Sentence 42](#_Toc461009343)

[Figure 25: Parsing of a Compound Sentence 42](#_Toc461009344)

[Figure 26: Input English Tree 43](#_Toc461009345)

[Figure 27: Output renamed English Tree 44](#_Toc461009346)

[Figure 28: Evaluation 52](#_Toc461009347)

Chapter1

Introduction

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

## Background

A language is mod of communication among human beings. There exists numerous languages for human being to interact with each other. These languages can be written or spoken but among these languages, there is a special kind of language which is neither written nor spoken. This is known as Sign Language. It is a visual language which means it is gestured language, it uses signs and gestures to communicate.

Pakistan, with the population of 130.58 million, suffers from about 2.49% of disability out of which 7.40% are impaired of hearing. According to World Health Organization the commonness of disability is 10% which makes Pakistan seem as if they have very fewer people suffering from any kind of disability but in fact most of the parents deny the presence of disability among their children or conceal the fact that their children are disabled.

Sign Language is a general term used for gestured language but it can be made more specific if we categories it geographically like American Sign Language, British Sign Language, Indian Sign Language etc. Like other sign languages Pakistan Sign Language has its own lexeme, syntax and semantics. We could use English structure for sign language but unfortunately deaf people are usually unaware of English grammar and sentence structure.

Although the deaf community communicates with each other using sign language without any difficulty but the problem arises when they have to communicate with their non-disabled peers. Like in universities when a lecture is delivered in spoken language, deaf students are unable to understand, for this a sign language translation medium is highly needed. Commonly human interpreters are used for this purpose but due to the limited and costly resources such solution is discouraged. Another approach to bridge this gap is to automate such system which will do the interpreter work for us.

## Motivation

Los Angeles times, "A Child Literacy Initiative for the Greater Los Angeles Area" states that "No skill is more crucial to the future of the child, or to a democratic and prosperous society, than literacy."

Pakistan is ranked seventh among the world's most populated countries with a 55 percentile literacy rate around the world. Keeping in view the low literacy rate it is advised to work for the prosperity of our country while consuming each and every way opportunity, one of which is educating the disabled persons.

The history of Special Education is not very old in Pakistan. At the time of independence three such schools were working with the special need. Now the Department of special education runs 48 special schools under the supervision of Chief Minister.

Deaf children are educated separately in special schools from the beginning in Pakistan. Such segregated systems keeps them away from their non-disabled peers, shielding them through the social emotional experience of hearing students but there is a big flaw in this approach and that is it limit their contact with other human being and making them feel isolated from the outside world. So that is why we wanted to minimize this gap as much as possible.

## Problem statement

To bridge the above mentioned gap between hard of hearing and a normal person, we have to come up with a system which can translate on language to another. Till now interpreters were doing this work for us but now we are going to automate it because of limited resources. Unfortunately very little work is done on natural language to sign language conversion and that is there is only one website (www.PSL.org.pk) available for such purpose. It contain word by word translation stored in the form of videos.

**Existing Issues**

* Such approach is good for learning signs but it is not build for translation propose as translation is mostly done on sentence level so this approach fails when it comes to translating some texts into sign language.
* Storing 5,000 videos isn't memory efficient.
* These videos require a lot of bandwidth for downloading.
* Difficult to maintain huge amount of data.
* Time consuming if we want to use it for translating.

For example:   
Input: I go to school.

We have to look for the signs for "I", "go", "to", "school".

And also there will be no consistency in the data as the videos stored in the PSL database it not signed by only one person.

* Moreover, Pakistan Sign Language sentence structure does not resembles with English language structure.

For example:   
Input: "I go to school."  
In PSL, above sentence is sign as "I school go."

Keeping in view the sentence structure issue mentioned about, there is a need to study Pakistan Sign Language's sentence structure so that we could translate according to the PSL grammar. Without grammar, sentences can't convey their proper meanings.

After document the grammar we now need to automate the function of converting English sentence into its Pakistan Sign Language equivalent sentence.

For this we have come up with a system which take an English sentence as an input and will generate its equivalent Pakistan Sign Language's sentence.

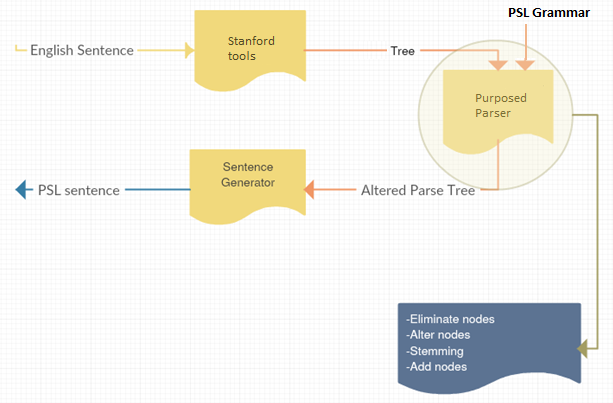


Figure 1: Over View of Proposed System

Our proposed system uses some third party Stanford tools which parse the given sentence to give English language lexicons along with their parts of speech and dependency tags, in the form of a tree. The outputted tree, along with our documented grammar is passed to our proposed parser which then alter the given tree nodes based on those grammar rules. Last but not the least a PSL sentence is generated from the altered tree.

When a sentence for example "I have not seen this movie before” is passed through it, it first passes through Stanford English tagger it will output the following tree.

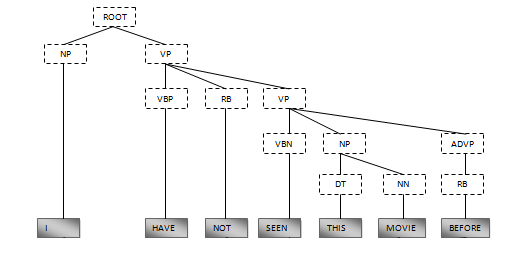


Figure 2: Stanford Parser Output

This tree is then passed onwards to our proposed parser which will either eliminate some nodes, add some nodes, stem them or alter them based on the PSL grammar. Its output will also be a tree.

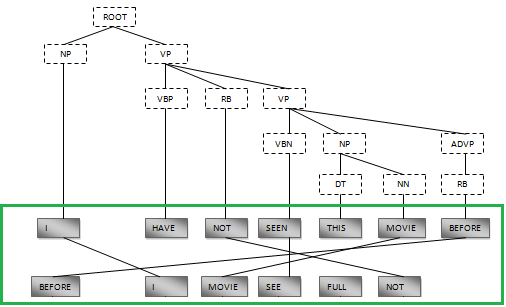


Figure 3: Updating in Stanford Tree

This above drawn tree is then make an input to the sentence generator which will form the PSL equivalent sentence.

Below is the **Transformation of** "I have not seen this movie before" **to** "Before I movie see full not.

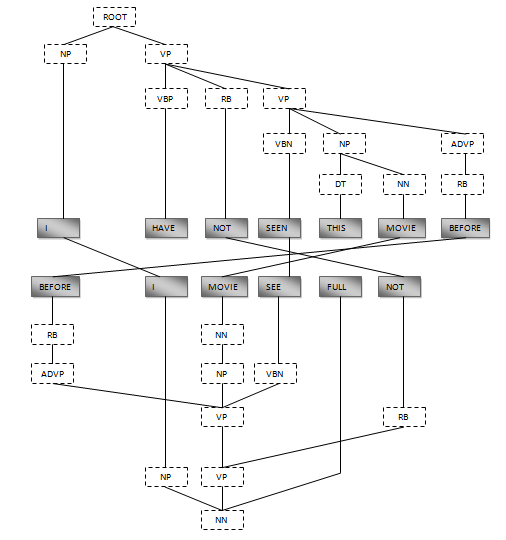


Figure 4: PSL Equivalent Sentence

# 

Chapter2

Language Analysis

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

## Natural Language

The word natural language corresponds to language that has developed naturally in use i.e. we do not invent them but acquire them. Some of the examples of Natural languages are English, German, French etc. natural languages are spoken and written. Natural languages are boundless in a sense that there is no restriction on the length of the sentence or the phrase that means you can never run out of new sentences. They are also riddled with ambiguity, this is because a single word, phrase or a sentence has multiple meaning depending on the context.

Before focusing our attention to anything else, we first have to study the structure for the source language i.e. English language. because when we make a transfer based MT, we have to completely explore both the languages and make some transfer rules depending upon the structure of both these languages.

English language structure states that adjective is always placed before a noun and in PSL we could have a structure where adjective is placed after the noun. So in this way our transfer rule will be if there is an adjective before a noun then swap their position.

### Grammatical Structure for English Language:

Textual Analysis of English language is in fact our first phase of the proposed methodology which is carried out in the pattern given below.

**Types of English Sentences in Terms of Structure:**

1. Simple
2. Complex
3. Compound
4. Compound complex

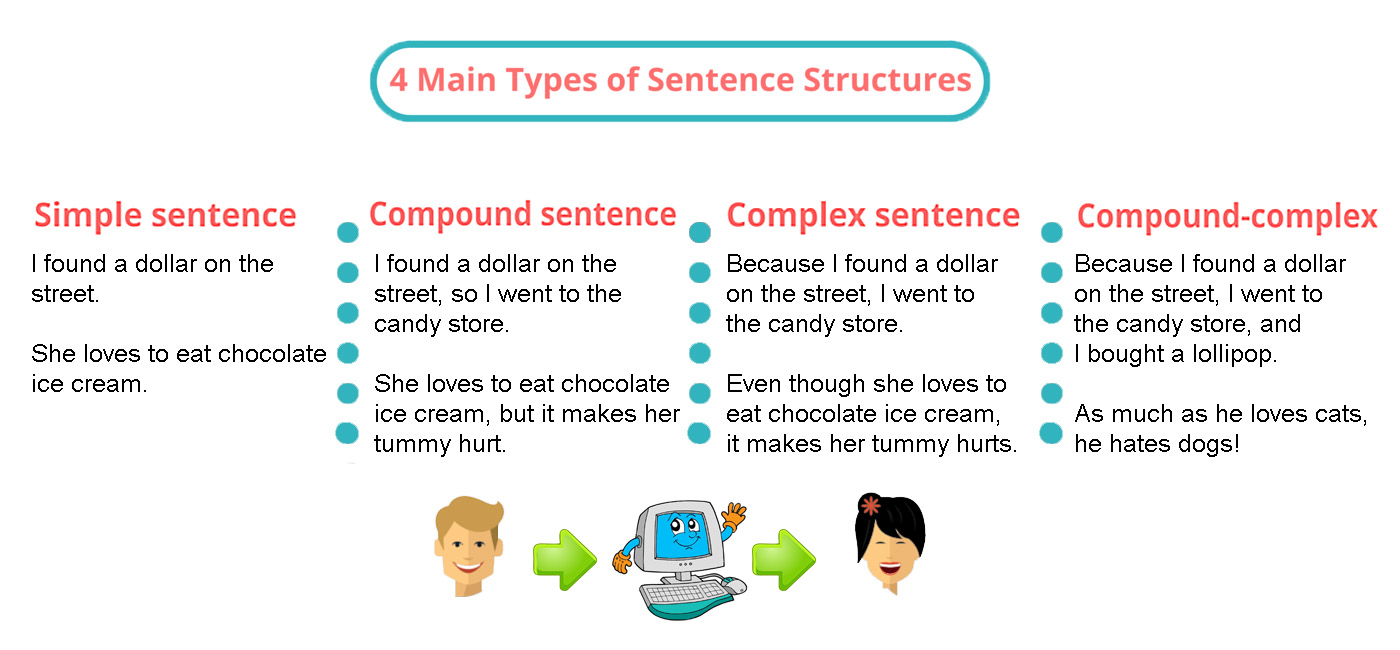


Figure 8: Types of English Sentences in Terms of Structure

**Simple Sentence**

A simple sentence consists of one independent clause.

Now there is a need to answer “what is an independent clause”?

An independent clause is one which is contain a subject followed by a verb or a verb phase and also it expresses a single idea.

For example:

* My new computer has an Intel core processor.

**Compound Sentence**

Compound sentence refers to a sentence containing two independent clauses joined by a coordinating conjunction which could be "and", "but", "or", "so", "yet", "however" etc.

For Example:

* I wish to buy an Apple laptop but I don't have enough money.

As compound sentence consists of two independent sentences that is why it has two subjects and two verb phrases.

**Complex Sentence:**  
A complex sentence is a type of sentence which has one independent clause and one or more dependent clauses.  
Whereas a dependent clause is opposite to an independent clause. Dependent clause cannot be a standalone sentence by itself.

For Example:

* That boy in red shirt, who is standing outside dean's office, is my student.

**Compound-Complex Sentence**

The compound sentence means it has two independent clauses and complex sentence means it has one independent clause and one or more dependent clauses. so after summing up both of these definitions

A compound-complex sentence contains 3 or more clauses: 2 independent and at least 1 dependent clause.

For example:

* That boy in red shirt, who is standing outside dean's office, is my student and he is one of the genius people I know.

Independent clause: That boy in red shirt is my student.

He is one of the genius people I know.

Dependent clause: who is standing outside dean's office?

**Types of English Sentences in Terms of Meaning /Function:**

1. Declarative (Statement)
2. Interrogative(Question)
3. Imperative (Command)
4. Exclamatory(Emotions)

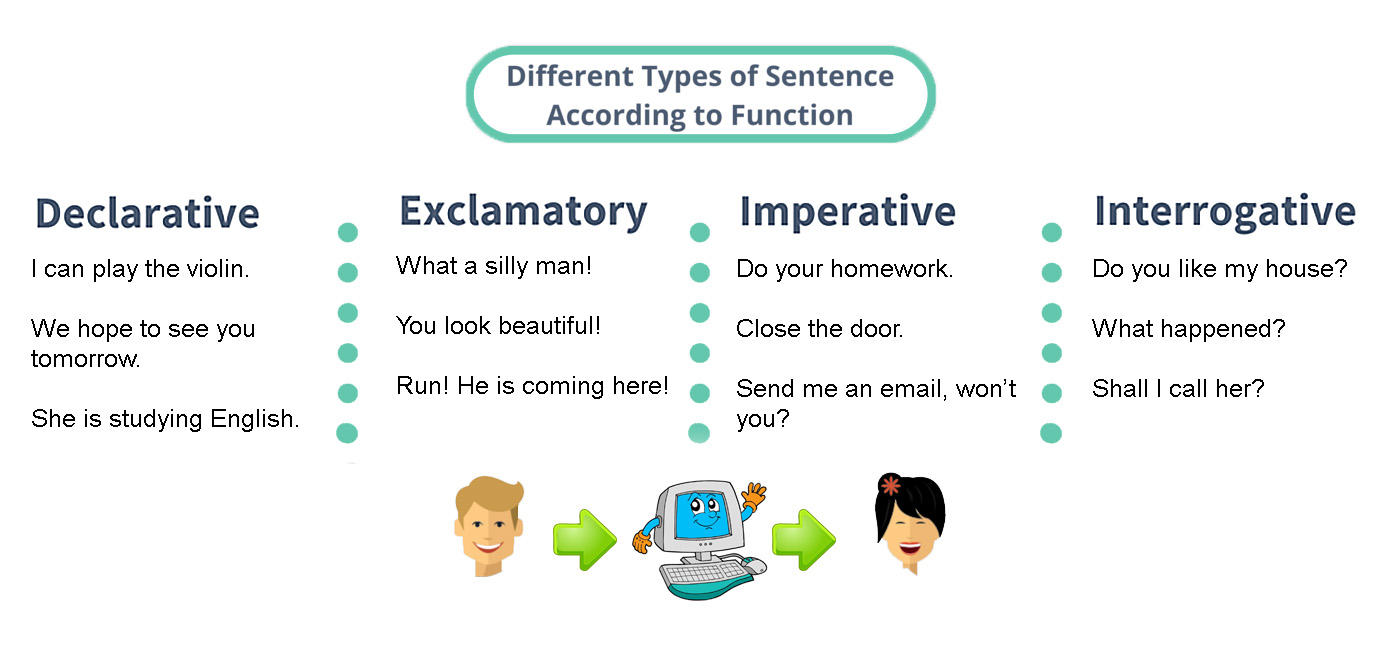


Figure 9: Types of English Sentences in Terms of Function

**Types of Sentences by Function:**

1. Declarative sentence: A declarative sentence is a statement which is followed by a period.

For example: I have studied Compiler Construction course in my BS. Software Engineering degree.

2. Interrogative sentence: An interrogative sentence is a questioning sentence and it is followed by a question mark.

For example: Where do you live?

3. Imperative sentence: An Imperative sentence is a commanding or requesting sentence. Like declarative it is also followed by a period.  The subject is always "you" understood.

For example: Shut the door.

4. Exclamatory sentence: An exclamatory sentence is a sentence that expresses a feeling.  It is followed by an exclamation mark or a period.

For example: Alas! We missed your performance.

**Types of English Sentences in Terms of Tense:**

1. Present
   1. Simple
   2. Indefinite
   3. Continuous
   4. Perfect
   5. Perfect Continuous
2. Past
   1. Simple
   2. Indefinite
   3. Continuous
   4. Perfect
   5. Perfect Continuous
3. Future
   1. Simple
   2. Continuous
   3. Perfect
   4. Perfect Continuous

**Type of sentence by tense:**

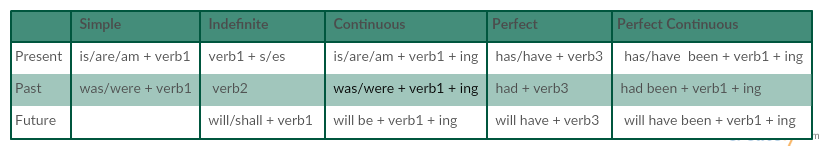


Figure 10: Types of English Sentences in Terms of Tense

## Sign Language

Sign Language mainly uses manual communication to express meaning. This can involve simultaneously combining hand shapes, orientation and movement of the hands, arms or body, and facial expressions to express a speaker's thoughts.

Sign Language is a general term used for gestured language but it can be made more specific if we categories it geographically like American Sign Language, British Sign Language, Indian Sign Language etc. Like other sign languages Pakistan Sign Language has its own lexeme, syntax and semantics. But Pakistan Sign Language’s syntax has no form.

## Translation

Plenty of work is done on sign language regarding corpus, grammar, machine translation etc. but unfortunately very little work is done on Pakistan Sign Language so far. American and British sign languages are the most developed ones. The work done on Machine Translation from English to American or British sign language mostly follows the described below architecture.

### Machine Translation Architecture

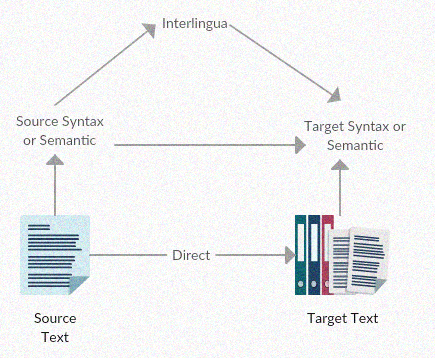
Dorr el at [1] has roughly organized MT architecture into the following three categories. 

Figure 5: Machine Translation Architecture

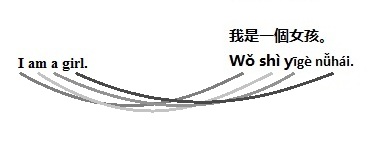
The tree levels of the above pyramid diagram correspond to different layer of transfer. At the bottom of the pyramid is the direct approach which consist of the most basic form of transfer i.e. word-to-word replacement. Note: word order of the target language is maintained the same as the source language.

At top is the Interlingua approach which consist of the most degenerate for of transfer i.e. Source-language text should be transformed into an Interlingua (an abstract language-independent representation). The target language is then generated from that Interlingua. It is a costly process as we have to maintain the semantics of the source language.

The third level which lies between these two extreme is the transfer architecture, most translation systems fall into this category. The initial intent of transfer approach is to provide syntactically correct target language text by transforming source-language representations into suitable target-language syntactic representations. Although the transfer rules that perform this conversion depends on both the source and target languages, some of the rules may need only slight modification when a MT system is developed for a new target language linguistically related to an existing one. Note: for applying transfer approach we need some "linking rules" that map between the source language text and target language text.

When one Natural Language is translated into another Natural Language, this is mostly done on the basis of one to one mapping that is the direct approach which is because the word-order does not differ from each other. The result of a direct translation architecture is a string of target-language words directly replaced from the words of the source language. Generally, the word order of the target-language text is the same as the source language.

For example:



|  |  |
| --- | --- |
| English | Chinese |
| I | Wǒ. |
| am | shì |
| a | yīgè |
| girl | nǚhái |

Figure : English and Chinese Equivalent Sentence

Unfortunately the target language with which we are dealing is not a Natural language but Sign language. Sign Language differs from Natural language as natural language is spoken and written but sign language is visual. There is a lot of information which can be expressed more effortlessly when using visual language instead of written or spoken language.   
For example:

He angrily said "Close the door".

When signing this sentence we do not need to sign the word "angrily" but have to express the by signer's face expressions.

So in this way when signing a sentence a lot of information is skipped and only those are sign which can convey our meaning of the sentence.

When we are deal with American Sign Language mostly the word-order of ASL and English is the same so direct approach is applicable in such translators but in Pakistan Sign Language we can't continue SVO word-order as in Pakistan different word-order is being used.

That is why direct approach fails while translating natural language to sign language because there isn't one to one mapping of the words and also the word-order can't be maintained.

For example:

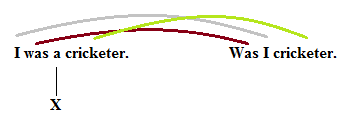


Figure 7: English and PSL Equivalent Sentence

When we translate one language to another our main aim is to maintain its semantics at any cost even if we have to alter its syntax. Interlingua machine translation gives the most accurate results as it deals will the semantics and not the syntax. Unfortunately Interlingua machine translation is difficult and may be even impossible for a wider domain. There may be a very few systems which can maintain the semantics while translation. So Transfer based Machine translation seems to be another choice it is not efficient as Interlingua MT but is better than direct translation. This approach make use of some linking rules. The effect of these rules could be swapping the object and the subject, eliminate articles etc. Unlike direct approach it accommodates complex mapping. However, a common criticism of this approach is that a large set of transfer rules must be constructed for each source-language to target-language pair. If we compare the cost of writing these transfer rules with the cost of implementing Interlingua MT, we would go for Transfer based MT approach.

Now the next challenge is that what are these rules which will help us to translate to spoken language to sign language?

## Rules and it's need

These rules are the grammatical structure. Every existing language has a grammatical structure on the basis of which it is constructed. Grammar is defined as the set of rules that explain how words are used in a language, how words and their component parts to form sentences. Without grammar, clear communication is nearly impossible. Grammar keeps you from being misunderstood while expressing your thoughts and ideas.

Like for English, we have subject-verb-object, adjective-subject-verb or subject-verb-adverb etc but Pakistan Sign Language has no formalized grammar structure.

Our one of the contribution to this project that we have finally documented the Pakistan Sign Language grammar.

## Grammatical Structure for Pakistan Sign Language:

Sentence structure of Pakistan Sign language differs from English Language as it has its own rules and glossary. English language follows a *Subject + Verb + Object* structure whereas Pakistan Sign Language follows various structures*,* one of which is known as Topic-Comment structure where the topic submit to subject or the object and the comment refers to the information regarding that topic described earlier. Unlike other Sign Languages very little work is done on Pakistan Sign Language that is why PSL is not that much stable as American Sign Language or other developed Sign languages.

Grammatical Structure of Pakistan Sign Language along with its documented form is further discussed in the next Chapter.

**Challenges in forming a grammatical structure for PSL:**

A system face many problems while translating an English sentence into equivalent PSL sentence. Some of them are listed below.

* As mentioned above Sign Language is a visual language it has its own grammar and structure which makes it distinct from other spoken languages like English.
* Those deaf persons who do not attain sign language from their parents, develop their signing skills later in life unlike those who inherit it from their parents. In this way the earlier category get to learn some spoken language for writing but the others don't.
* There is great diversity in terms of hearing perception, their style, their language background, their inheritance of sign language as mentioned above, and so on.
* Till now no free standard written grammar exists for PSL nor has much progress been made in this field. A simple English subject-verb-object sentence is signed differently by different people in the same demographic area such as:  
   -Subject-Object-Verb  
   -Subject-Verb-Object

Chapter3

Related Work

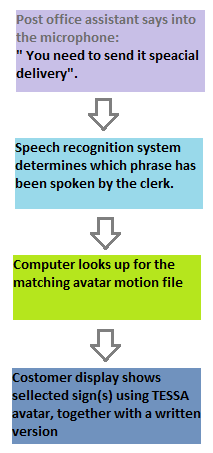
\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

## Systems under consideration

Some of the MT systems for other sign language are briefly described below. The underlying MT architecture can be classified into

1. Direct translation system
2. Transfer based architecture and
3. Statistical MT.

Some of the direct translation systems include:

* The TExt and Sign Support Assistant commonly known as TESSA have been developed to assists hearing impaired or hard of hearing people at the post office during transactions.

TESSA system combines with the speech recognition system and the visual avatar to communicate with a deaf person. This system has already stored some pre-defined phrases into its database along with its equivalent British Sign Language translation. When an assistant speaks into the microphone which is then recognized by the computer speech recognition system. After that the system looks up for the phrase into system's database and if the phrase is found, it then uses its equivalent British translation to display the avatar signing to the customer's display screen along with its written version.

TESSA is a very domain specific system. It store a small set of phrases which are already translated into is British Sign Language version due to which very restricted conversation could be carried out between the clerk and the costumer.

Figure : TESSA Proposed System

* Angus B. Grieve-Smith from the University of New Mexico, 1999 have worked on SignSynth project. In this project Grieve-Smith proposed an English to American Sign Language Machine translation for Weather Reports. He stated that his proposed application consists of four components: lexical analyzer, parser, a transfer module and a generation module.

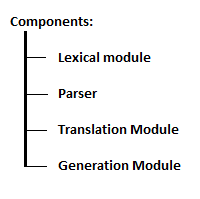
In this project lexical module tag the text with the lexical category for each word; for example: partly cloudy was tagged as <sky>. The tagged report is then parsed through RecDescent parser, developed by Damian Conway which will then produce parse tree based on an input text in a particular language but this step is often skipped it if the source text is not syntactically complicated. The transfer is done by a simple look up table. This table consists of the lexical analyzer items, each with its translation in ASL, written in Newkirk notation. The output from this phase is then entered into the generation module which will produce a fluent ASL based on the formulas, all of which uses a topic-comment structure and non-manual grammatical morphemes.

Figure : Major Components of English to ASL by Grieve-Smith

The transfer architecture systems include:

* Prototype Machine Translation System From Text-To-Indian Sign Language was proposed by Tirthankar Dasgupta,Sandipan Dandpat and Anupam Basu in 2008. According to their paper they have followed the transfer based Machine Translation as their frame architecture. They proposed a system with the following components.

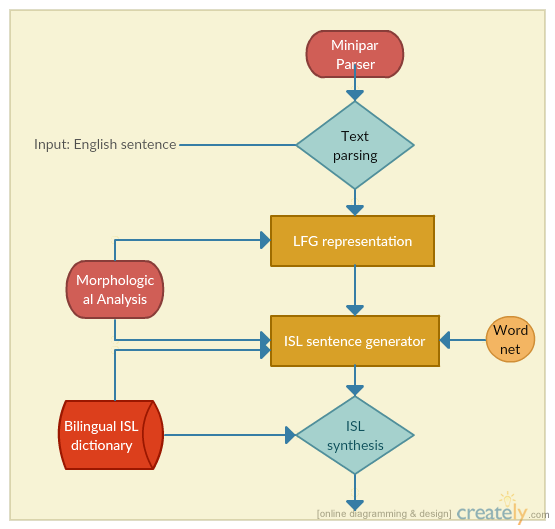


Figure 13: Proposed System of English to ISL

Input English text is parsed from a third party tool "Minipar" and from the outputted parse tree a dependency tree is constructed. This dependency structure is then passed to the next phase i.e. LFG representation phase. In this phase the given sentence is represented as the set of attribute-value pair where the grammatical symbol or a syntactic function is known as the attribute and the value is the related feature possessed by the concerning element. In the generation phase the English structure is converted into the ISL equivalent sentence by applying two operations on the sentence

1. Lexical selection
2. Word order correspondence

Lexical selection phase is done by using English-ISL bilingual lexicon. For example in ISL word "EVERYONE" is replaced by "ALL".

ISL follows the Subject-Object-Verb order. For example: English sentence "I eat meat." is equivalent to "I meat eat." in ISL.

* Speers in 2001 proposed and implemented an American Sign Language machine translation system called the ASL Workbench [Speers 2001]. He uses the LFG (lexical functional grammar) for analysis of the English text into f-structure. Then based on the transfer rules, constructed, English functional structure into an ASL. And again LGF rules are used to produce and American Sign Language output.
* Another English to American Sign Language Translation system was TEAM which was built by the University of Pennsylvania that used Synchronous Tree Adjoining Grammar rules to construct an American Sign Language syntactic structure while a dependency tree was built during analysis of English sentence. This system does not replace the bilingual lexicons to identify the applicable word-sign pair. Lingual module outputs a written American Sign Language gloss notation. This ASL gloss notation contains the manual and non-manual information along with the morphological variation. This gloss notation is then forwarded to the Synthesis module which then output an animated human model.

Following below is the table briefing about the pros and cons of the system mentioned above.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| System | Domain Restricted | Systematic Approch | Limited Data Bank | Insufficeint Grammer Research | Achieved Objective |
| TESSA | ✔ |  | ✔ | ✔ |  |
| SignSynth | ✔ | ✔ |  | ✔ | ✔ |
| Prototype MT System From Text-To-ISL |  | ✔ |  | ✔ | ✔ |
| ASL Workbench |  | ✔ |  |  | ✔ |
| TEAM |  | ✔ |  |  | ✔ |

Figure 14: Pros and Cons of deferent translation systems

Our system resembles a lot with the “Prototype Machine Translation System From Text-To-Indian Sign Language by Tirthankar Dasgupta,Sandipan Dandpat and Anupam Basu" as it used Stanford parser to parse the sentence, Stanford dependency tool to get the dependency graph merge the tree and the graph to get an annotated tree. Inputting that annotated tree to the translation module where rules are applied to the tree and these rules are in the form of CFG. These rules results the following operations:

* Add Node
* Delete Node
* Alter Node

The output of the translation module is a PSL sentence.

Chapter4

Proposed Methodology: Part1

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

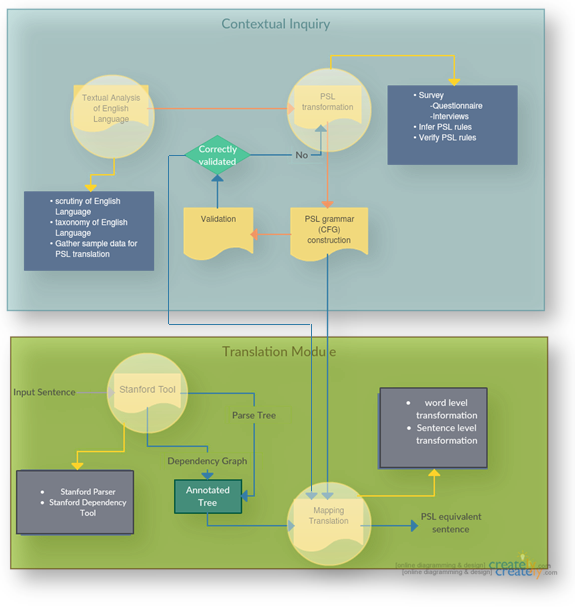
The preceding chapters have discussed the need for a translation component (English to Pakistan Sign Language) which will enable a normal hearing person to communicate with a hard of hearing person. In this chapter, we will discuss the proposed system: two major components   
1. Contextual Inquiry   
2. Translation Module   
along with their sub-components. 

Figure 15: Proposed System

## Overview of the Proposed System

**Contextual Inquiry:**

As experts has defined contextual inquiry as a semi-structured interview approach to gather data about the use of context, where the interviewer speaks to the interviewee, aiming to gather as much data as he/she can for later analysis.

We are making use of such method for our data gathering part. We seek to collect as much data as we can in a neutral environment from our interviewee (deaf interpreters).

Textual Analysis of English Language:

* Scrutiny of English Language: In this phase of the Context Inquiry section we have had a critical examination of English language. Observed its structure, its formation.
* Taxonomy of English Language: After having keenly observed the English language structure we came to know the basic classification on English language structure. i.e.
  + Classification on the Basis of structure:
    - Simple
    - Complex
    - Compound
    - Compound Complex
  + Classification on the Basis of Meaning:
    - Declarative
    - Interrogative
    - Exclamatory
    - Imperative
  + Classification on the Basis of Tenses
    - Present
      * Simple
      * Indefinite
      * Continuous
      * Perfect
      * Perfect Continuous
    - Past
      * Simple
      * Indefinite
      * Continuous
      * Perfect
      * Perfect Continuous
    - Future
      * Simple
      * Continuous
      * Perfect
      * Perfect Continuous

As discussed in Chapter 2.

* Gather Sample data for PSL Structure Survey: Carrying on this structured approach, we have gathered some sample data, categorized into above mentioned classifications. Furthermore, it is discussed later in this chapter.

PSL transformation:

* Survey: In this phase we have developed some questionnaires on the foundation of above gathered sample data and get them filled from the interpreters of the deaf community residing in University Of Management and Technology, Lahore. This survey also includes some unstructured interviews for a more flexible data collection as our interview depends on the respondent's answers, allowing the respondent to talk in some depth about the translation of English Language to Pakistan Sign Language and also this approach increases the validity of our data as we give the interpreters opportunity to probe for deeper understanding**.**
* Infer PSL rules: After the surveys we then have some raw data for English to Pakistan Sign Language conversion. Exploring and analysis of this raw data then lead us towards inferring some rules for this conversion.
* Verify PSL rules: "Verification is better than firing arrows into the sky." So it's a better practice to get those rules verified from the interpreters to check if are heading towards the right direction.

PSL Grammar Construction: Once our inferred rules are verified we then form a Context Free Grammar for Pakistan Sign Language.

Validation: Dry run the constructed CFG for few sentences to check the accuracy of our CFG. If we had been facing some errors then we have head back towards the PSL transformation Phase to again consider the deduced rules.

**Translation Module:**

It is the implementation Section where we are provided with an English sentence as an input and we have to perform some operations to convert it in PSL equivalent sentence. This phase will be discussed in more detail in the next chapter.

Stanford Tool: Provided English sentence is passed through the Stanford tool where it get parsed from the Stanford Parser, resulting a tree and a dependency graph is created when processed through the dependency tool.

Annotated tree: Both the outputs from the Stanford tool are merged to give an annotated tree.

Mapping Translation:

* Word Level Transformation: When an annotated tree is provided to this phase as an input, a pre-order traversal is taken and on the basis of each reduction some rules are fired which are built on the CFG constructed in the contextual Inquiry section. Note that this transformation is done on node level.
* Sentence Level Transformation: After traversing through the whole tree we then apply some sentence level transformation rules which are applicable only on the identification of the tense and its category (Meaning vise classification).

Output of this phase is our PSL equivalent sentence.

Data Gathering

A well-defined grammar for English language is available but there is no formalized grammar or sentence structure for Pakistan Sign Language. For Example a sentence from indefinite category **she likes badminton** is signed in differently in various parts of Pakistan as shown in table below.

|  |  |
| --- | --- |
| Like badminton she. | V+O+S |
| She badminton like. | **S+O+V** |
| She like badminton. | **S+V+O** |
| Badminton like she. | **O+V+S** |
| Badminton she like. | **O+S+V** |
| Like she badminton. | **V+S+O** |

Figure : PSL structure

There was a great need of formalizing the grammatical structure of Pakistan Sign Language as without grammar, clear communication is nearly impossible. Grammar keeps you from being misunderstood while expressing your thoughts and ideas.

**Contribution:**

We took 10 sentences of each below mentioned categories and ask students and interpreters of our universities hard of hearing community to write the PSL sentence against our sample sentences. Other than the Questionnaire, we have conducted some unstructured interview and held some discussion sessions with the interviewee of ours in neutral environment, to get a clear insight of the sign language grammar.

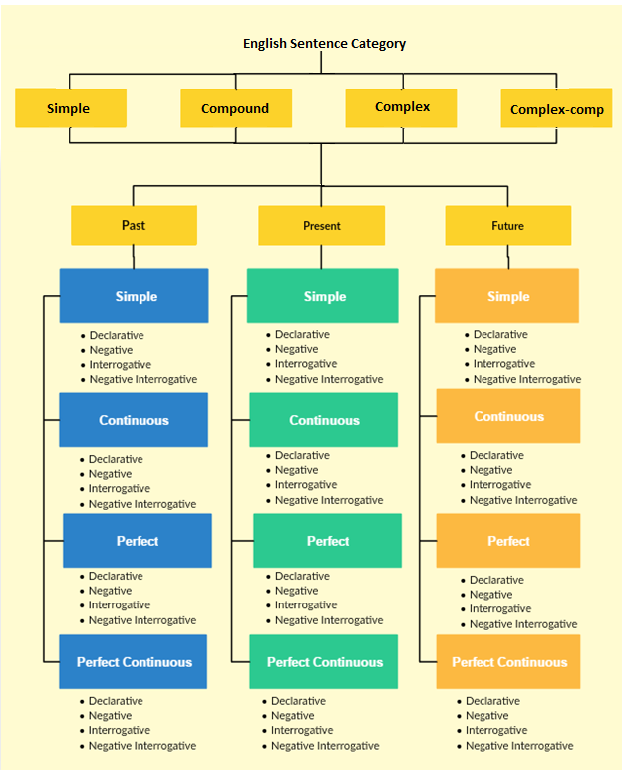


Figure 17: English Sentence Categories We Catered

After the survey it was noticed that different sentence structures were followed by different people. A bar chart for the basic word order is shown in the following figure.

Figure : bar chart for followed word-order in Pakistan

From the above chart we observed that for simple category sentences most frequently followed sentence structure is **S+O+V** .In the same way we have catered almost all categories mentioned aboveand followed the highly used sentence structures for each category.

Some of the categories are mentioned below with their equivalent English and PSL sentence.  
Note: They are all declarative sentence type.

|  |  |  |
| --- | --- | --- |
| Category | English Sentence | PSL Sentence |
| Present simple | She is poor. | She poor. |
| Past simple | She was poor. | Was she poor. |
| Future simple | She will be poor. | She poor after. |
|  |  |  |
| Present Cont. | I am enjoying the rain. | I rain enjoy now. |
| Past Cont. | I was enjoying the rain. | Was I rain enjoy. |
| Future Cont. | I will be enjoying the rain. | I rain enjoy after. |
|  |  |  |
| Present perfect | I have played football. | I football play full. |
| Past perfect | I had played football. | Was I football play full. |
| Future perfect | I will have played football. | I football play full after. |
|  |  |  |
| Present perfect cont | I have been eating. | I eat full. |
| Past perfect cont. | I had been eating. | Was I eat full |
| Future perfect cont. | I will have been eating. | I eat full after. |
|  |  |  |
|  |  |  |

Figure : English and its equivalent PSL sentence

### Deduced Observation:

After collecting and examining all the raw data on English to PSL conversation, we have concluded our observation into mentioned below rules.

* **Complex and Complex-Compound Sentences:**Complex and complex compound sentences are broken down into simple or compound sentences.
* **Sentence structure:**PSL follows many sentence structures but the most widely used sentence structure is Subject-Object-Verb.
* **Time Words:**words like tomorrow, yesterday, or soon are used in English to indicate the time of an event. If any of time words exists in the English sentence, it must be placed at the beginning of the sentence.
* **Articles:**   
  Eliminate articles from the sentence.
* **Negation:**Negation is sign in PSL by shaking head i.e. side to side. In PSL negation is placed at the end of the sentence.
* **Adjective:**   
  Adjectives are always placed before the noun however some sign them after the noun.
* **Adverb:**Position of adverb is not altered.
* **Indication of Tense:**for indication that a sentence is in Past tense, PSL place "was" at the beginning of the sentence. For Future tense" After" is placed at the end of the sentence and for present no such indicators are use i.e. if there are no tense indicators that means sentence is in present form.
* **Interrogative:**
  + **Wh question:** If the source sentence is an interrogative then its wh-questioning part is moved to the end of the sentence. Except if it starts with "when" then word "time" is placed at the end of the sentence.
  + **Auxiliary Question:** If the source sentence is an interrogative then its Auxiliary questioning part is replaced by "yes no" and is moved to the end of the sentence.
* **Indication of sub-tense:**For indication continuous tense "now” word is used in the equivalent PSL sentence and for perfect tense "full" is used.

Our Proposed Grammar for PSL**:**

Observing all these sentence structures we proposed a Context Free Grammar for PSL which cover all categories mentioned in Table above**.**

**Note:**  following grammar is ambiguous as left factoring has to be done on it.

**Sentence** 🡪 Sentence conj Sentence   
 | Sentence

**Sentence**🡪 Present  
 |Future  
 |Past

**Present** 🡪 TimeClause Subject Object Verb TenseIndicator neg Question?

**Past** 🡪 *was* TimeClause Subject Object Verb TenseIndicator neg question?

**Future** 🡪 TimeClause Subject Object Verb TenseIndicator *full* Question?

**Subject** 🡪 Adj Noun  
 | Adj Proper-Noun  
 |Adj Pronoun  
 |Subject and Subject

**Object**🡪 Adj Noun  
 |Adj Proper-Noun  
 |Object and Object  
 | €

**Sentence** 🡪 Sentence conj Sentence   
 | Sentence

**Sentence**🡪 Present  
 |Future  
 |Past

**Present** 🡪 TimeClause Subject Object Verb TenseIndicator neg question?

**Past** 🡪 *was* TimeClause Subject Object Verb TenseIndicator neg question?

**Future** 🡪 TimeClause Subject Object Verb TenseIndicator *full* question?

**Subject** 🡪 Adj Noun  
 | Adj Proper-Noun  
 |Adj Pronoun  
 |Subject and Subject

**Object**🡪 Adj Noun  
 |Adj Proper-Noun  
 |Object and Object  
 | €

**Sentence** 🡪 Sentence conj Sentence   
 | Sentence

**Sentence**🡪 Present  
 |Future  
 |Past

**Present** 🡪 TimeClause Subject Object Verb TenseIndicator neg question?

**Past** 🡪 *was* TimeClause Subject Object Verb TenseIndicator neg question?

**Future** 🡪 TimeClause Subject Object Verb TenseIndicator *full* question?

**Subject** 🡪 Adj Noun  
 | Adj Proper-Noun  
 |Adj Pronoun  
 |Subject and Subject

**Object**🡪 Adj Noun  
 |Adj Proper-Noun  
 |Object and Object  
 | €

**Verb🡪** Adv Verb  
 | allEnglishVerbs | €

**TenseIndicator🡪** complete  
 |full  
 |now  
 |€

**neg🡪** Not| €

**TimeClause🡪** Tomorrow  
 |yesterday  
 |today  
 | €  
**Adj🡪** allEnglishAdjectives | €

**Adv🡪** allEnglishAdverbs | €

**Noun🡪** allEnglishNouns | €

**ProperNoun🡪** allEnglishProperNouns | €

**Pronoun🡪** allEnglishPronouns | €

**Question🡪** what   
 |why  
 | time   
 |where   
 |which  
 | AuxInterr  
 | €

**AuxInterr🡪** yes no

**Conjunction 🡪** and | or | but

**Sentence** 🡪 Sentence conj Sentence   
 | Sentence

**Sentence**🡪 Present  
 |Future  
 |Past

**Present** 🡪 TimeClause Subject Object Verb TenseIndicator neg question?

**Past** 🡪 *was* TimeClause Subject Object Verb TenseIndicator neg question?

**Future** 🡪 TimeClause Subject Object Verb TenseIndicator *full* question?

**Subject** 🡪 Adj Noun  
 | Adj Proper-Noun  
 |Adj Pronoun  
 |Subject and Subject

**Object**🡪 Adj Noun  
 |Adj Proper-Noun  
 |Object and Object  
 | €

Chapter 5

Proposed Methodology: Part2

For implementation of a system for translation of English language grammar to Pakistan Sign language grammar, structure/composition of both grammars was analyzed with the help of collected data.

## Grammar Processing of a Natural Language

Every language has a grammar to define the composition/structure of its sentences/statements.

In computer science, we classify grammars as

1. Ambiguous
2. Non-Ambiguous

Ambiguous grammars are those grammars in which the meaning or constitution of an element of grammar can vary with respect to the given context. Humans can identify which meaning is conveyed, from their previous experience with the grammar of that language. But computers cannot identify the correct meaning unless they are provided with some rules to distinguish between different meanings of the same combination of terms/words.

Programming languages are usually made to be non-ambiguous, to make their processing easier for compiling and interpreting software.

Natural languages commonly have ambiguous grammars.

## Analysis of English Language Grammar

An analysis of Natural language grammar tells us that:

* Natural Languages are composed of sentences.
* A sentence can be broken down to phrases.
* A phrase can be further broken down to other phrases or words.
* Words can be thought of as lexical units of natural language.

This view of grammatical structure gives us a nested composition of sentences, which can have several composition levels. Level 1 is the sentence itself, and the words are placed in the last level, as they are the atomic components of a natural language grammar which cannot be divided further. Following figure shows the grammatical composition of an English sentence.

Naughty children make noise in the classroom.

Sentence

Naughty children

Noun Phrase

make noise in the classroom.

Verb Phrase

Naughty

Adjective

children

Noun

make

Verb

noise

Noun

in the classroom

Preposition Phrase

in

Preposition

the

Determiner

classroom

Noun

Software programs that parse sentences of natural language and identify their structure and composition, are known as Natural Language Parsers (NLP).

## Getting Output from Natural Language Parser

Our project uses Stanford's NLP for processing sentences to get the composition of sentence and identification of its constituent elements.

The parser gives its output in a bracketed notation. An example sentence and its corresponding parser output is given below:

**Sentence:**   
My dog is chasing a black cat.

**Natural Language Parser Output:**   
(ROOT (S (NP (PRP$ My) (NN dog)) (VP (VBZ is) (VP (VBG chasing) (NP (DT a) (JJ black) (NN cat)))) (. .)))

This output can be viewed in a nested hierarchy as:

(ROOT

(S

(NP (PRP$ My) (NN dog))

(VP (VBZ is)

(VP (VBG chasing)

(NP (DT a) (JJ black) (NN cat))))

(. .)))

Where the indentation (space from the start of line) denotes the level of sentence components. Each starting brace denotes the start of a new level.

Stanford's Natural Language Parser uses the grammar tags from Penn Treebank Tag-set for tagging linguistic structure of natural languages. These tags are given in Appendix 1a.

## Saving Grammar Information in a Tree Data-Structure

To process the grammar information of English Sentences, and translating them into Pakistan Sign Language sentences. The output from Natural Language Parser is stored in a tree data-structure.

For the sentence example given in previous section, the NLP output is

**Natural Language Parser Output:**   
(ROOT (S (NP (PRP$ My) (NN dog)) (VP (VBZ is) (VP (VBG chasing) (NP (DT a) (JJ black) (NN cat)))) (. .)))

The tree data-structure constructed from this output can be viewed as given in following diagram:

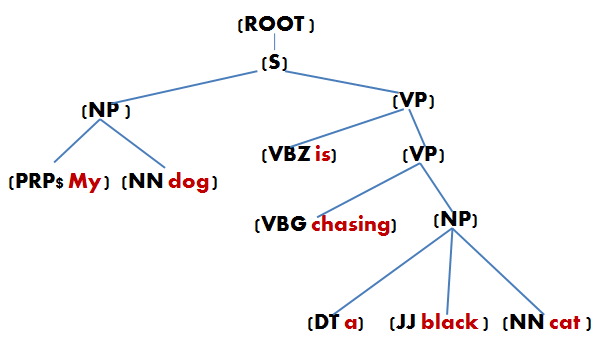


Figure : English Tree constructed from NLP Output

## Getting Word Dependencies from Typed Dependencies Tool

Another tool called Stanford's typed dependences Tool, gives information about the dependencies between the words of input sentence, in the following format:

**Input Sentence:**   
My dog is chasing a black cat.

**Output: Typed Dependencies**   
nmod:poss(dog-2, My-1)   
nsubj(chasing-4, dog-2)   
aux(chasing-4, is-3)   
root(ROOT-0, chasing-4)   
det(cat-7, a-5)   
amod(cat-7, black-6)   
dobj(chasing-4, cat-7)

This dependency output actually represents a directed graph. Graphical representation of the graph is given in the following diagram:

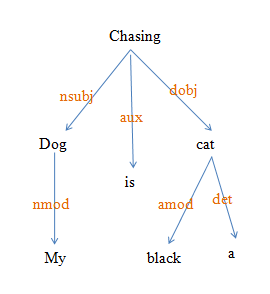


Figure : Dependency Graph

In this graph, nodes represent the words, and the edges represent the dependencies between the words.

## Annotated Tree

The tree data structure constructed from NLP output, is annotated by adding the word dependency information in it.

The following diagrams show the input tree and the output tree after the addition of typed dependencies information.

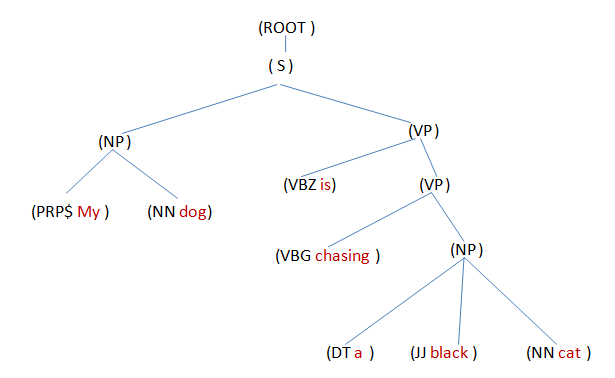


Figure : Input Tree

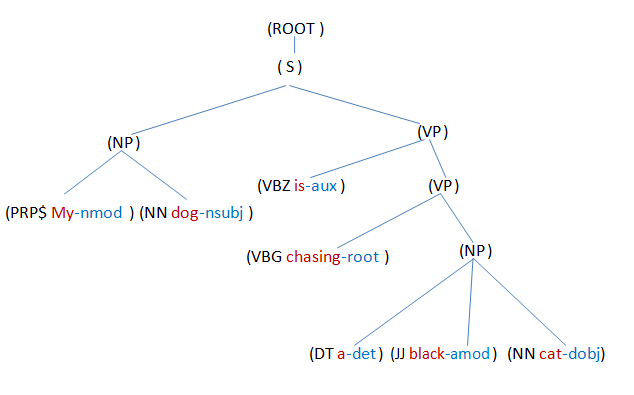


Figure : Output Tree (Annotated)

As you can see that our tree data structure now has additional information stored in word-level nodes.

## Tense Identification

The English language sentences are categorized into 14 categories with respect to tense.

|  |  |
| --- | --- |
| **Tense No** | **Tense** |
| 1 | Present Indefinite Tense |
| 2 | Present Continuous Tense |
| 3 | Present Perfect Tense |
| 4 | Present Perfect Continuous Tense |
| 5 | Past Indefinite Tense |
| 6 | Past Continuous Tense |
| 7 | Past Perfect Tense |
| 8 | Past Perfect Continuous Tense |
| 9 | Future Indefinite Tense |
| 10 | Future Continuous Tense |
| 11 | Future Perfect Tense |
| 12 | Future Perfect Continuous Tense |
| 13 | Simple Present Tense |
| 14 | Simple Past Tense |

For an input sentence, its tense is identified by the identification of types of helping verbs and verbs present in the composition of that sentence. The constructed annotated tree of the sentence is traversed for the search of helping verbs and main verbs.

## Sentence Structure Identification

The English sentences have four categories with respect to sentence structure.

1. Simple Sentence
2. Compound Sentence
3. Complex Sentence
4. Complex-Compound Sentence

For definitions of these categories, we first define two types of clauses, i.e. Independent clause and dependent clause:

**Independent Clause**

An independent clause contains a subject followed by verb or verb phrase. It can convey its full meaning by itself.

**Dependent Clause**

A dependent clause cannot convey its full meaning by itself.

Now we define the before mentioned four sentence categories, in terms of dependent and independent clauses:

**1. Simple Sentence**

A simple sentence contains only one independent clause.

**2. Compound Sentence**

A compound sentence contains two or more independent clauses.

**3. Complex Sentence**

A complex sentence contains one independent clause, and one or more dependent clauses.

**Complex-Compound Sentence**

A complex-compound sentence is a sentence which is a compound sentence as well as a complex sentence.

Therefore, to identify the sentence structure type, we will have to identify the number of independent and dependent clauses in it.

The independent and dependent clauses can be differentiated on the basis of type of conjunction used in the sentence.

A **conjunction** is a part of speech that connects words, sentences, phrases, or clauses.

**Coordinating conjunctions** are conjunctions that join two or more items of equal syntactic importance. The acronym FANBOYS can be used to remember these conjunctions, which are

* for
* and
* nor
* but
* or
* yet
* so

Conjunctions other than coordinating conjunctions are **subordinating conjunctions**.

If a clause contains coordinating conjunction at its start, it is an independent clause.

Clauses containing subordinating conjunctions are dependent clauses.

### Algorithm for Structure Type Identification

1. Traverse the tree data-structure of a sentence. Identify the number of clauses and their starting positions in the sentence.

if (number of clauses == 1)

return: sentence is a simple sentence

end if

2. Identify the positions of all coordinating conjunction used in the sentence.

3. for each starting position of clause 'S'

for each position of coordinating conjunction 'CC'

if (position of 'S' - position of 'CC' == 1)

Mark clause as an independent clause.

end if

end for

end for

if (number of independent clauses == number of clauses)

return: sentence is a compound sentence

else if (number of independent clauses > 2)

return: sentence is a compound-complex sentence

else

return: sentence is a complex sentence

end if

### Translation of Compound, Complex and Complex Compound Sentences

For translation of compound, complex and complex compound sentences, the sentence is first broken down in simple sentences. The each simple sentence is sent to the translation function.

The following diagrams show the break-down of a compound sentence into simple sentences:



Figure : Parsing of a Compound Sentence

Similarly, Complex sentences can be broken down to make simple sentences too:



Figure : Parsing of a Compound Sentence

## Renaming of English Tree according to PSL Grammar

In our constructed English language tree, the words are tagged using Treebank Tag-set. These tags are different than the tags (names) in the grammar that we have developed for Pakistan Sign Language, which contains simplified and easy to understand tag-set.

Therefore, after the identification of sentence type and tense, the Penn tree-bank tags are renamed to the PSL grammar tags using a mapping function. The following diagrams show the input English tree and the renamed English tree.

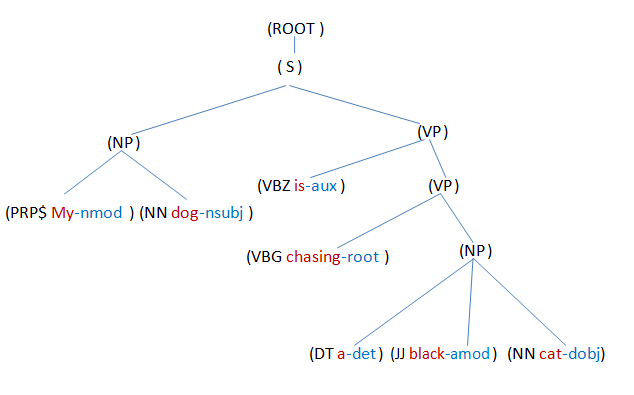


Figure : Input English Tree

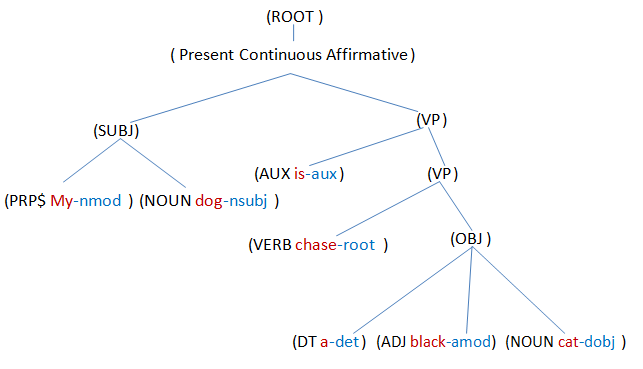


Figure : Output renamed English Tree

The tags renamed in this example are given in the following table:

|  |  |
| --- | --- |
| English Tag | PSL Tag |
| S | Present Continuous Affirmative |
| NP | SUBJ |
| PRP$ | PRPS |
| NN | NOUN |
| VBZ | AUX |
| VBG | VERB |
| NP | OBJ |
| JJ | ADJ |

## Construction of PSL Tree

After the renaming of English sentence tree, the tree is parsed from top to bottom (pre-order traversal).

PSL grammar rules are input to the program in form of a text file. This file contains PSL grammar productions written in Backus-Naur form. The format for writing grammar rules is explained in Appendix 1b.

At each non-terminal (non-leaf) node, the children of that node are compared by the PSL grammar productions for that non-terminal and a best matched production is chosen.

The tree is modified according to the matched PSL production. The modification operations on a non-terminal node are of three types:

1. Addition of a new child node
2. Removal of a child node
3. Re-ordering of children nodes

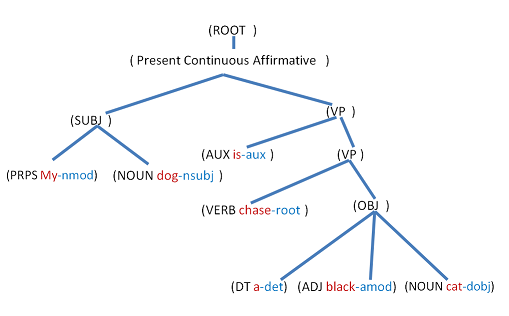
The new tree thus constructed have all productions according to PSL grammar.

Let us take the example of the sentence used in previous sections:

**Input Sentence:**   
My dog is chasing a black cat.

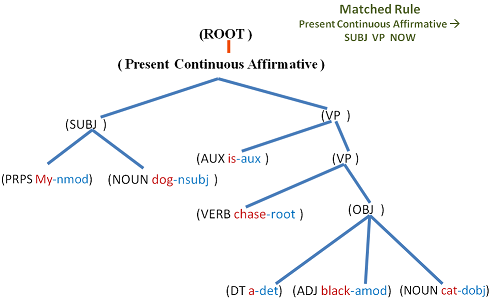
The series of operations performed on this sentence's English tree to convert it to an equivalent PSL tree, are explained below with diagrams:

The renamed English tree of the sentence is



The tree will be traversed from top to bottom in a pre-order traversal order (The node itself is visited first, then its children from left to right)

The (ROOT) node is visited first. Our PSL grammar has no rule for this node, therefore no changes will be made to its children. The next node visited is (Present Continuous Affirmative) node.

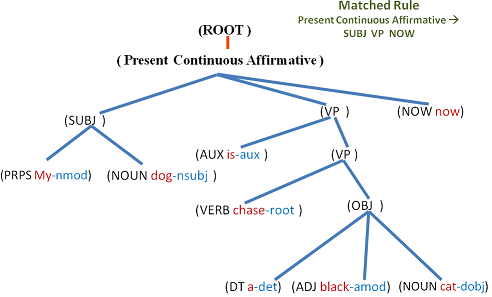


As we can see that in this tree, the grammar production for this node is

Present Continuous Affirmative --> SUBJ VP

But in our PSL grammar, the matched production for (Present Continuous Affirmative) non-terminal is: Present Continuous Affirmative --> SUBJ VP "now"

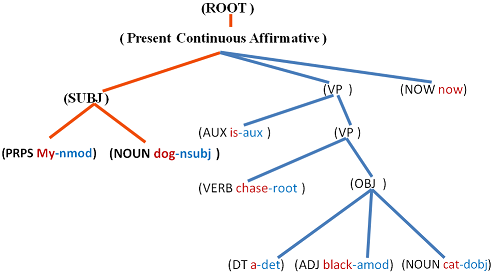
Which contains an additional token "now" at the end. (As for Present Continuous Affirmative sentence in PSL, we use the word "now" after the Subject and Verb Phrase). Therefore, the program will add an additional child with the value "now", to the node being visited. The new tree is given in the following diagram:



The next node visited is (SUBJ) and the matched PSL grammar production is:

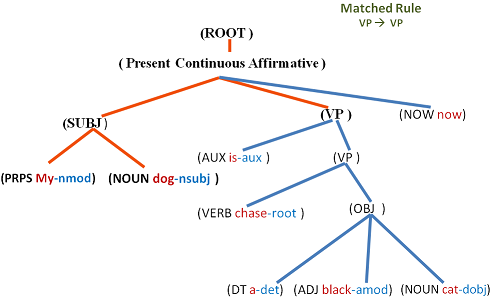
SUBJ --> PRPS NOUN

Which is same as in the current tree, therefore no changes will be made to the node (SUBJ)



The next visited node is (VP) and the matched PSL grammar production in this case is

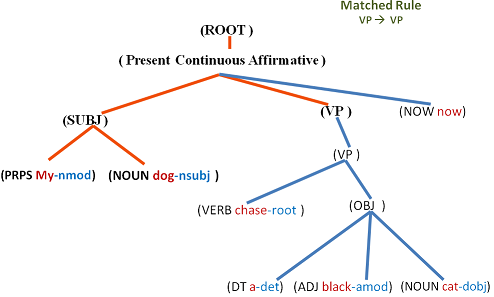
VP --> VP



But as we can see, the current tree contains the order at node (VP) as

VP --> AUX VP

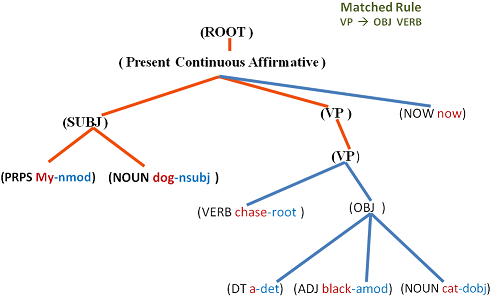
Therefore, the (AUX is-aux) node will be removed from the node (VP), as shown in diagram below:



The next node to be visited is the node (VP), child of current node (VP)

The matched PSL grammar production in this case is

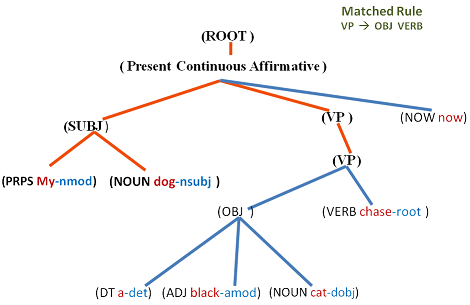
VP --> OBJ VERB



but in the current tree, the order of (VP) node's children is

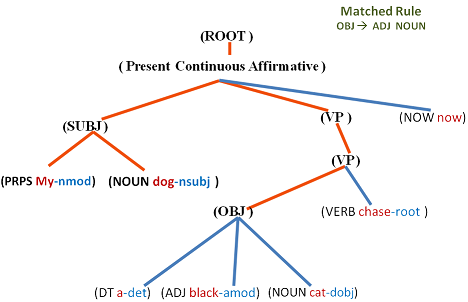
VP --> VERB OBJ

Therefore, the children nodes of (VP) will be reordered to match the PSL grammar.



The next node to be visited is (OBJ) node, which has the production order as

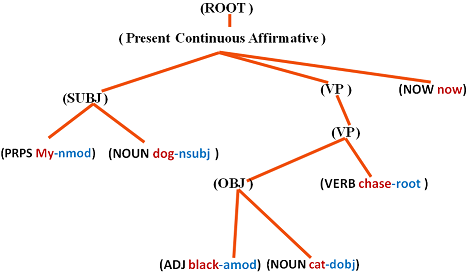
OBJ --> DT ADJ NOUN



But the matched PSL grammar production in this case is

OBJ --> ADJ NOUN

(As PSL does not contain determiners), Therefore the (DT) node will be removed from tree. The remaining nodes are leaf nodes which do not have any grammar rules. Therefore the final PSL tree will be:

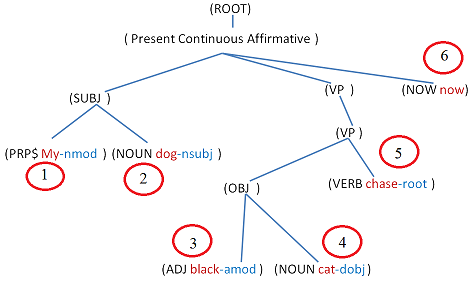


## Translated Sentence

The PSL tree is traversed again for a leaf-only traversal. And the value (word) encountered on each leaf node is concatenated in a resultant string variable.

After the traversal, we get PSL sentence in the resultant string variable.

The following diagram shows the order in which the leaf nodes will be visited.



The resultant PSL sentence will be:

My dog black cat chase now.

Chapter5

# Evaluation

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

As PSL lacks the written orthography, this makes it difficult to evaluate a Text-to-PSL MT system by conventional ways because of which the standard techniques for evaluating Text- Text MT systems are not applicable for Text-to- PSL systems. We have evaluated our system by getting the feedback of PSL experts and users.

Evaluation of our system is done on two levels

* Sentence level
* Paragraph level

As we have discussed the literacy rate in chapter1, most of the deaf in Pakistan are not familiar with English Language because of which we can’t evaluate our system’s output directly from them, as the outputted text is also in English. For the evaluation we have used a third party tool known as “siGML server player”. PSL sentence, we get from our parser is then passed to this tool which will then sign the PSL sentence using an avatar.

We have taken approx. 500 sentences from the categories that we have catered in our grammar document and pass it to our proposed system. Meanwhile we have presented the same set of data in the form of questionnaire to our PSL users and experts. The generated outputs from our system and deaf community are then compared and accuracy of our system is calculated. The system was evaluated on a set of 480 sentences. The overall system performance is around 93%. (See appendix 2)

After the evaluation on sentences we took some stories from the internet and from a grammar book of 5th class and did the same procedure. (See appendix 3)

Following are the results.

|  |  |  |  |
| --- | --- | --- | --- |
| Stories | Total sentences | No. of Correct sentences | Accuracy  (%) |
| Thirsty Crow taken from internet | 16 | 14 | 87 |
| Thirsty Crow taken from 5th class grammar book | 11 | 11 | 100 |
| Greedy Dog | 11 | 11 |
| Clever Fox | 15 | 15 |

Figure : Evaluation

**Conclusion and Future Work**

The research analysis shows various different dimensions in which people have been working to resolve these issues. The unfortunate part is that no significant work exists for Pakistan sign language. We have proposed the solution to one of the task (English text to Pakistan Sign Language translation) mentioned in the previous paper by documenting the PSL grammar and by proposing a translator.

In future, we intend to address all these challenges by implementing all remaining components in the proposed architectural framework especially the Pakistan Sign Language to English text translation to develop a two way communication tool. As we have also proposed a general architectural framework, in the next stage of our work it can help translating Urdu text/voice to animations of Pakistan sign language, and vice versa.

# Appendix A: Penn Treebank II Tags

Note: This information comes from "Bracketing Guidelines for Treebank II Style Penn Treebank Project" - part of the documentation that comes with the Penn Treebank.

**Contents:**

Bracket Labels  
Clause Level  
Phrase Level  
Word Level  
Function Tags  
Form/function discrepancies  
Grammatical role  
Adverbials  
Miscellaneous

**Bracket Labels**

**Clause Level**

S - simple declarative clause, i.e. one that is not introduced by a (possible empty) subordinating conjunction or a wh-word and that does not exhibit subject-verb inversion.  
SBAR - Clause introduced by a (possibly empty) subordinating conjunction.  
SBARQ - Direct question introduced by a wh-word or a wh-phrase. Indirect questions and relative clauses should be bracketed as SBAR, not SBARQ.  
SINV - Inverted declarative sentence, i.e. one in which the subject follows the tensed verb or modal.  
SQ - Inverted yes/no question, or main clause of a wh-question, following the wh-phrase in SBARQ.

**Phrase Level**

ADJP - Adjective Phrase.  
ADVP - Adverb Phrase.  
CONJP - Conjunction Phrase.  
FRAG - Fragment.  
INTJ - Interjection. Corresponds approximately to the part-of-speech tag UH.  
LST - List marker. Includes surrounding punctuation.  
NAC - Not a Constituent; used to show the scope of certain prenominal modifiers within an NP.  
NP - Noun Phrase.  
NX - Used within certain complex NPs to mark the head of the NP. Corresponds very roughly to N-bar level but used quite differently.  
PP - Prepositional Phrase.  
PRN - Parenthetical.  
PRT - Particle. Category for words that should be tagged RP.  
QP - Quantifier Phrase (i.e. complex measure/amount phrase); used within NP.  
RRC - Reduced Relative Clause.  
UCP - Unlike Coordinated Phrase.  
VP - Verb Phrase.  
WHADJP - Wh-adjective Phrase. Adjectival phrase containing a wh-adverb, as in how hot.  
WHAVP - Wh-adverb Phrase. Introduces a clause with an NP gap. May be null (containing the 0 complementizer) or lexical, containing a wh-adverb such as how or why.  
WHNP - Wh-noun Phrase. Introduces a clause with an NP gap. May be null (containing the 0 complementizer) or lexical, containing some wh-word, e.g. who, which book, whose daughter, none of which, or how many leopards.  
WHPP - Wh-prepositional Phrase. Prepositional phrase containing a wh-noun phrase (such as of which or by whose authority) that either introduces a PP gap or is contained by a WHNP.  
X - Unknown, uncertain, or unbracketable. X is often used for bracketing typos and in bracketing the...the-constructions.

**Word level**

CC - Coordinating conjunction  
CD - Cardinal number  
DT - Determiner  
EX - Existential there  
FW - Foreign word  
IN - Preposition or subordinating conjunction  
JJ - Adjective  
JJR - Adjective, comparative  
JJS - Adjective, superlative  
LS - List item marker  
MD - Modal  
NN - Noun, singular or mass  
NNS - Noun, plural  
NNP - Proper noun, singular  
NNPS - Proper noun, plural  
PDT - Predeterminer  
POS - Possessive ending  
PRP - Personal pronoun  
PRP$ - Possessive pronoun (prolog version PRP-S)  
RB - Adverb  
RBR - Adverb, comparative  
RBS - Adverb, superlative  
RP - Particle  
SYM - Symbol  
TO - to  
UH - Interjection  
VB - Verb, base form  
VBD - Verb, past tense  
VBG - Verb, gerund or present participle  
VBN - Verb, past participle  
VBP - Verb, non-3rd person singular present  
VBZ - Verb, 3rd person singular present  
WDT - Wh-determiner  
WP - Wh-pronoun  
WP$ - Possessive wh-pronoun (prolog version WP-S)  
WRB - Wh-adverb

**Function tags**

**Form/function discrepancies**

* ADV (adverbial) - marks a constituent other than ADVP or PP when it is used adverbially (e.g. NPs or free ("headless" relatives). However, constituents that themselves are modifying an ADVP generally do not get -ADV. If a more specific tag is available (for example, -TMP) then it is used alone and -ADV is implied. See the Adverbials section.
* NOM (nominal) - marks free ("headless") relatives and gerunds when they act nominally.

**Grammatical role**

* DTV (dative) - marks the dative object in the unshifted form of the double object construction. If the preposition introducing the "dative" object is for, it is considered benefactive (-BNF). -DTV (and -BNF) is only used after verbs that can undergo dative shift.
* LGS (logical subject) - is used to mark the logical subject in passives. It attaches to the NP object of by and not to the PP node itself.
* PRD (predicate) - marks any predicate that is not VP. In the do so construction, the so is annotated as a predicate.
* PUT - marks the locative complement of put.
* SBJ (surface subject) - marks the structural surface subject of both matrix and embedded clauses, including those with null subjects.
* TPC ("topicalized") - marks elements that appear before the subject in a declarative sentence, but in two cases only: if the front element is associated with a *T* in the position of the gap.  
  if the fronted element is left-dislocated (i.e. it is associated with a resumptive pronoun in the position of the gap).
* VOC (vocative) - marks nouns of address, regardless of their position in the sentence. It is not coindexed to the subject and not get -TPC when it is sentence-initial.

**Adverbials**

Adverbials are generally VP adjuncts.

* BNF (benefactive) - marks the beneficiary of an action (attaches to NP or PP).  
  This tag is used only when (1) the verb can undergo dative shift and (2) the prepositional variant (with the same meaning) uses for. The prepositional objects of dative-shifting verbs with other prepositions than for (such as to or of) are annotated -DTV.
* DIR (direction) - marks adverbials that answer the questions "from where?" and "to where?" It implies motion, which can be metaphorical as in "...rose 5 pts. to 57-1/2" or "increased 70% to 5.8 billion yen" -DIR is most often used with verbs of motion/transit and financial verbs.
* EXT (extent) - marks adverbial phrases that describe the spatial extent of an activity. -EXT was incorporated primarily for cases of movement in financial space, but is also used in analogous situations elsewhere. Obligatory complements do not receive -EXT. Words such as fully and completely are absolutes and do not receive -EXT.
* LOC (locative) - marks adverbials that indicate place/setting of the event. -LOC may also indicate metaphorical location. There is likely to be some varation in the use of -LOC due to differing annotator interpretations. In cases where the annotator is faced with a choice between -LOC or -TMP, the default is -LOC. In cases involving SBAR, SBAR should not receive -LOC. -LOC has some uses that are not adverbial, such as with place names that are adjoined to other NPs and NAC-LOC premodifiers of NPs. The special tag -PUT is used for the locative argument of put.
* MNR (manner) - marks adverbials that indicate manner, including instrument phrases.
* PRP (purpose or reason) - marks purpose or reason clauses and PPs.
* TMP (temporal) - marks temporal or aspectual adverbials that answer the questions when, how often, or how long. It has some uses that are not strictly adverbial, auch as with dates that modify other NPs at S- or VP-level. In cases of apposition involving SBAR, the SBAR should not be labeled -TMP. Only in "financialspeak," and only when the dominating PP is a PP-DIR, may temporal modifiers be put at PP object level. Note that -TMP is not used in possessive phrases.

**Miscellaneous**

* CLR (closely related) - marks constituents that occupy some middle ground between arguments and adjunct of the verb phrase. These roughly correspond to "predication adjuncts", prepositional ditransitives, and some "phrasel verbs". Although constituents marked with -CLR are not strictly speaking complements, they are treated as complements whenever it makes a bracketing difference. The precise meaning of -CLR depends somewhat on the category of the phrase.
* on S or SBAR - These categories are usually arguments, so the -CLR tag indicates that the clause is more adverbial than normal clausal arguments. The most common case is the infinitival semi-complement of use, but there are a variety of other cases.
* on PP, ADVP, SBAR-PRP, etc - On categories that are ordinarily interpreted as (adjunct) adverbials, -CLR indicates a somewhat closer relationship to the verb. For example:
  + Prepositional Ditransitives  
    In order to ensure consistency, the Treebank recognizes only a limited class of verbs that take more than one complement (-DTV and -PUT and Small Clauses) Verbs that fall outside these classes (including most of the prepositional ditransitive verbs in class [D2]) are often associated with -CLR.
  + Phrasal verbs  
    Phrasal verbs are also annotated with -CLR or a combination of -PRT and PP-CLR. Words that are considered borderline between particle and adverb are often bracketed with ADVP-CLR.
  + Predication Adjuncts  
    Many of Quirk's predication adjuncts are annotated with -CLR.
* on NP - To the extent that -CLR is used on NPs, it indicates that the NP is part of some kind of "fixed phrase" or expression, such as take care of. Variation is more likely for NPs than for other uses of -CLR.
* CLF (cleft) - marks it-clefts ("true clefts") and may be added to the labels S, SINV, or SQ.
* HLN (headline) - marks headlines and datelines. Note that headlines and datelines always constitute a unit of text that is structurally independent from the following sentence.
* TTL (title) - is attached to the top node of a title when this title appears inside running text. -TTL implies -NOM. The internal structure of the title is bracketed as usual.

# Appendix B: How to write rules in rules file

## Grammar Rules Format

Canonical forms of grammar should be written in rules.txt file in the format:

**NonTerminal ::= Production1 | Production2**

Every non-terminal should be enclosed in angular brackets(<>) without space. e.g.

**<VP>**

Terminals (string literals) should be additionally enclosed in quotes ("") without space. e.g.

**<"full">**

All productions of a non-terminal, should be declared in single line, separated by OR sign (|).

Comments are allowed but they should not contain these symbols:

::, <, > , |

Spaces between tokens are allowed.

A sample rule is given below:

**<FUTURE-PERFECT-AFFIRMATIVE> ::= <NP> <VP> <"full"> <"after"> | <NP><ADJP><"full"><"after">**

If a production contains two or more non-terminals with the same name, they should be numbered. e.g.

<VP> ::= <PP><PP><PP><VERB>

should be written as

<VP> ::= <PP1><PP2><PP3><VERB>

(change the order if required)

# Appendix C: Sentence level Test Results

|  |  |  |  |
| --- | --- | --- | --- |
| Category No | Input | Output | Result |
| s-1 | He is clever. | He clever. | Correct |
| s-2 | They drink. | They drink. | Correct |
| s-3 | They drink water. | They water drink. | Correct |
| s-4 | He is not clever. | He clever not. | Correct |
| s-5 | White people drink water. | White people water drink. | Correct |
| s-6 | You ask strange questions. | You strange questions ask. | Correct |
| s-7 | He walks slowly. | He slowly walk. | Correct |
| s-8 | White people do not drink water. | White people water drink not. | Correct |
| s-9 | You do not ask strange questions. | You strange questions ask not. | Correct |
| s-10 | He does not walk slowly. | He slowly walk not. | Correct |
| s-11 | Is he clever? | he clever yes-no? | Correct |
| s-12 | Does he study? | he study yes-no? | Correct |
| s-13 | Do they drink water? | they water drink yes-no? | Correct |
| s-14 | Is he not clever? | he clever not yes-no? | Correct |
| s-15 | Does he not study? | he study not yes-no? | Correct |
| s-16 | Do they not drink water? | they water drink not yes-no? | Correct |
| s-17 | Do white people drink water? | white people water drink yes-no? | Correct |
| s-18 | Do white people not drink water? | white people water drink not yes-no? | Correct |
| s-19 | Do you ask strange questions? | you strange questions ask yes-no? | Correct |
| s-20 | Do you not ask strange questions? | you strange questions ask not yes-no? | Correct |
| s-21 | He is studying. | He study now. | Correct |
| s-22 | He is drinking water. | He water drink now. | Correct |
| s-23 | He is not studying. | He study now not. | Correct |
| s-24 | He is not asking questions. | He questions ask now not. | Correct |
| s-25 | White people are drinking water. | White people water drink now. | Correct |
| s-26 | You are asking strange questions. | You strange questions ask now. | Correct |
| s-27 | You are asking questions slowly. | You questions slowly ask now. | Correct |
| s-28 | White people are not drinking water. | White people water drink now not. | Correct |
| s-29 | You are not asking strange questions. | You strange questions ask now not. | Correct |
| s-30 | You are not eating food slowly. | You food slowly eat now not. | Correct |
| s-31 | Are you listening? | you listen now yes-no? | Correct |
| s-32 | Is he eating food? | he food eat now yes-no? | Correct |
| s-33 | Is he not listening? | he listen now not yes-no? | Correct |
| s-34 | Is he not eating food? | he food eat now not yes-no? | Correct |
| s-35 | Are white people drinking water? | white people water drink now yes-no? | Correct |
| s-36 | Are you asking strange questions? | you strange questions ask now yes-no? | Correct |
| s-37 | Is he eating food slowly? | he food slowly eat now yes-no? | Correct |
| s-38 | Are white people not drinking water? | white people water drink now not yes-no? | Correct |
| s-39 | Is he not eating good food? | he good food eat now not yes-no? | Correct |
| s-40 | Is he not eating food slowly? | he not yes-no? | Incorrect |
| s-41 | He has eaten. | He eat full. | Correct |
| s-42 | He has taken tea. | He tea take full. | Correct |
| s-43 | He has not eaten. | He eat full not. | Correct |
| s-44 | He has not taken tea. | He tea take full not. | Correct |
| s-45 | White people have taken tea. | White people tea take full. | Correct |
| s-46 | He has consumed healthy food. | He healthy food consumed full. | Correct |
| s-47 | He has eaten food slowly. | He food slowly eat full. | Correct |
| s-48 | White people have not taken tea. | White people tea take full not. | Correct |
| s-49 | They have not shot a red cow. | They red cow shoot full not. | Correct |
| s-50 | They have not eaten food slowly. | They food slowly eat full not. | Correct |
| s-51 | Has he killed? | he kill full yes-no? | Correct |
| s-52 | Has he killed people? | he people kill full yes-no? | Correct |
| s-53 | Has he not killed? | he kill full not yes-no? | Correct |
| s-54 | Has he not killed people? | he people kill full not yes-no? | Correct |
| s-55 | Have white people taken tea? | white people tea take full yes-no? | Correct |
| s-56 | Have they shot a big elephant? | they big elephant shoot full yes-no? | Correct |
| s-57 | Have they taken tea slowly? | they tea slowly take full yes-no? | Correct |
| s-58 | Have white people not taken tea? | white people tea take full not yes-no? | Correct |
| s-59 | Have they not killed a brown camel? | they brown camel kill full not yes-no? | Correct |
| s-60 | Have they not eaten food slowly? | they food eat full not yes-no? | Incorrect |
| s-61 | They have been crying. | They cry full. | Correct |
| s-62 | They have been killing animals. | They animals kill full. | Correct |
| s-63 | They have not been shouting. | They shout full not. | Correct |
| s-64 | They have not been killing animals. | They animals kill full not. | Correct |
| s-65 | White people have been killing animals. | White people animals kill full. | Correct |
| s-66 | He has been asking strange questions. | He strange questions ask full. | Correct |
| s-67 | He has been eating food slowly. | He food slowly eat full. | Correct |
| s-68 | White people have not been killing animals. | White people animals kill full not. | Correct |
| s-69 | He has not been asking strange questions. | He strange questions ask full not. | Correct |
| s-70 | He has not been driving a car slowly. | He car slowly drive full not. | Correct |
| s-71 | Have they been walking? | they walk full yes-no? | Correct |
| s-72 | Has Aslam been giving money? | Aslam money give full yes-no? | Correct |
| s-73 | Have they not been walking? | they walk full not yes-no? | Correct |
| s-74 | Has Aslam not been giving money? | Aslam money give full not yes-no? | Correct |
| s-75 | Have white people been killing animals? | white people animals kill full yes-no? | Correct |
| s-76 | Have you been asking strange questions? | you strange questions ask full yes-no? | Correct |
| s-77 | Have you been eating food slowly? | you food slowly eat full yes-no? | Correct |
| s-78 | Have white people not been killing animals? | white people animals kill full not yes-no? | Correct |
| s-79 | Have you not been asking strange questions? | you strange questions ask full not yes-no? | Correct |
| s-80 | Have you not been eating food slowly? | you eating food full not yes-no? | Incorrect |
| s-81 | He was clever. | was He clever. | Correct |
| s-82 | He slept. | was He sleep. | Correct |
| s-83 | He took tea. | was He tea take. | Correct |
| s-84 | He was not clever. | was He clever not. | Correct |
| s-85 | Old men made weapons. | was Old men weapons make. | Correct |
| s-86 | You asked an interesting question. | was You interesting question ask. | Correct |
| s-87 | He drove fast. | was He fast drive. | Correct |
| s-88 | Old men did not make weapons. | was Old men weapons make not. | Correct |
| s-89 | You did not ask an interesting question. | was You interesting question ask not. | Correct |
| s-90 | He did not drive fast. | was He fast drive not. | Correct |
| s-91 | Was he clever? | was he clever yes-no? | Correct |
| s-92 | Did he drive? | was he drive yes-no? | Correct |
| s-93 | Did he drive a car? | was he car drive yes-no? | Correct |
| s-94 | Was he not clever? | was he clever not yes-no? | Correct |
| s-95 | Did he not drive? | was he drive not yes-no? | Correct |
| s-96 | Did he not drive a car? | was he car drive not yes-no? | Correct |
| s-97 | Did old men make weapons? | was old men weapons make yes-no? | Correct |
| s-98 | Did old men not make weapons? | was old men weapons make not yes-no? | Correct |
| s-99 | Did he drive a red car? | was he red car drive yes-no? | Correct |
| s-100 | Did he not drive a red car? | was he red car drive not yes-no? | Correct |
| s-101 | You were dancing. | was You dance. | Correct |
| s-102 | You were making excuses. | was You excuses make. | Correct |
| s-103 | You were not dancing. | was You dance not. | Correct |
| s-104 | You were not making excuses. | was You excuses make not. | Correct |
| s-105 | Young boy was driving a car. | was Young boy car drive. | Correct |
| s-106 | He was driving a red car. | was He red car drive. | Correct |
| s-107 | He was driving a car slowly. | was He car slowly drive. | Correct |
| s-108 | Old men were not making weapons. | was Old men weapons make not. | Correct |
| s-109 | He was not driving a red car. | was He red car drive not. | Correct |
| s-110 | He was not driving a car slowly. | was He car slowly drive not. | Correct |
| s-111 | Were you dancing? | was you dancing yes-no? | Correct |
| s-112 | Were you making excuses? | was you making excuses yes-no? | Incorrect |
| s-113 | Were you not dancing? | was you dancing not yes-no? | Correct |
| s-114 | Were you not making excuses? | was you excuses make not yes-no? | Correct |
| s-115 | Were old men making weapons? | was old men making weapons yes-no? | Incorrect |
| s-116 | Was he driving a red car? | was he red car drive yes-no? | Correct |
| s-117 | Was he driving a car slowly? | was he car slowly drive yes-no? | Correct |
| s-118 | Were good people not giving money? | was good people money give not yes-no? | Correct |
| s-119 | Was he not driving a red car? | was he red car drive not yes-no? | Correct |
| s-120 | Was he not driving a car slowly? | was he car slowly drive not yes-no? | Correct |
| s-121 | They had won. | was They win full. | Correct |
| s-122 | They had won the match. | was They match win full. | Correct |
| s-123 | They had not won. | was They win full not. | Correct |
| s-124 | They had not lost the match. | was They match lose full not. | Correct |
| s-125 | Good players had won the match. | was Good players match win full. | Correct |
| s-126 | They had won a difficult challenge. | was They difficult challenge win full. | Correct |
| s-127 | They had won the match successfully. | was They match successfully win full. | Correct |
| s-128 | Rich people had not given money. | was Rich people money give full not. | Correct |
| s-129 | They had not won a difficult challenge. | was They difficult challenge win full not. | Correct |
| s-130 | They had not won the match successfully. | was They match successfully win full not. | Correct |
| s-131 | Had they won? | was they win full yes-no? | Correct |
| s-132 | Had they won the match? | Had they win? | Incorrect |
| s-133 | Had they not won? | was they win full not yes-no? | Correct |
| s-134 | Had he not brought money? | was he full not yes-no? | Incorrect |
| s-135 | Had black people won the match? | was black people match win full yes-no? | Correct |
| s-136 | Had he killed a big lion? | was he big lion kill full yes-no? | Correct |
| s-137 | Had he eaten the food slowly? | was he food slowly eat full yes-no? | Correct |
| s-138 | Had lazy girls not brought money? | was lazy girls money bring full not yes-no? | Correct |
| s-139 | Had he not killed a big lion? | Had he not kill? | Incorrect |
| s-140 | Had he not eaten the food slowly? | Had he not food slowly eat? | Correct |
| s-141 | You had been dancing. | was You dance full. | Correct |
| s-142 | You had been watching television. | was You television watch full. | Correct |
| s-143 | You had not been dancing. | was You dance full not. | Correct |
| s-144 | You had not been watching television. | was You television watch full not. | Correct |
| s-145 | Good men had been helping the poor. | was Good men poor help full. | Correct |
| s-146 | They had been helping old men. | was They old men help full. | Correct |
| s-147 | They had been eating cakes slowly. | was They cakes slowly eat full. | Correct |
| s-148 | Rich people had not been giving money. | was Rich people money give full not. | Correct |
| s-149 | They had not been helping old men. | was They old men help full not. | Correct |
| s-150 | They had not been eating cakes slowly. | was They cakes slowly eat full not. | Correct |
| s-151 | Had you been dancing? | was you dance full yes-no? | Correct |
| s-152 | Had you been making money? | was you money make full yes-no? | Correct |
| s-153 | Had you not been watching? | Had you not ? | Incorrect |
| s-154 | Had you not been watching him? | was you full not yes-no? | Incorrect |
| s-155 | Had rich people been giving money? | was rich people money give full yes-no? | Correct |
| s-156 | Had they been helping poor men? | was they poor men help full yes-no? | Correct |
| s-157 | Had he been driving a car slowly? | was he car slowly drive full yes-no? | Correct |
| s-158 | Had little boy not been watching cricket? | Had little boy not cricket watch? | Incorrect |
| s-159 | Had they not been helping poor men? | was they full not yes-no? | Incorrect |
| s-160 | Had he not been driving a car slowly? | Had he not car slowly drive? | Correct |
| s-161 | He will be clever. | He clever be after. | Correct |
| s-162 | He will serve. | He serve after. | Correct |
| s-163 | He will serve the queen. | He queen serve after. | Correct |
| s-164 | He will not serve. | He serve after not. | Correct |
| s-165 | Little boy will make money. | Little boy money make after. | Correct |
| s-166 | They will read this boring article. | They boring article read after. | Correct |
| s-167 | They will read slowly. | They slowly read after. | Correct |
| s-168 | Bad students will not read books. | Bad students books read after not. | Correct |
| s-169 | They will not read this boring book. | They boring book read after not. | Correct |
| s-170 | I shall not walk slowly. | I slowly walk after not. | Correct |
| s-171 | Will he be clever? | he clever be after yes-no? | Correct |
| s-172 | Will he serve? | he serve after yes-no? | Correct |
| s-173 | Will he serve the emperor? | he emperor serve after yes-no? | Correct |
| s-174 | Will he not be clever? | he clever be after not yes-no? | Correct |
| s-175 | Will he not serve? | he serve after not yes-no? | Correct |
| s-176 | Will he not serve the king? | he king serve after not yes-no? | Correct |
| s-177 | Will bad students make a noise? | bad students noise make after yes-no? | Correct |
| s-178 | Will bad students not make a noise? | bad students noise make after not yes-no? | Correct |
| s-179 | Will they take the red flower? | they red flower take after yes-no? | Correct |
| s-180 | Will you not smell this beautiful flower? | you beautiful flower smell after not yes-no? | Correct |
| s-181 | He will be speaking. | He speak after. | Correct |
| s-182 | He will be teaching English. | He English teach after. | Correct |
| s-183 | He will not be speaking. | He speak after not. | Correct |
| s-184 | He will not be teaching English. | He English teach after not. | Correct |
| s-185 | Bad students will be making a noise. | Bad students noise make after. | Correct |
| s-186 | They will not be driving a red car. | They red car drive after not. | Correct |
| s-187 | I will be walking slowly. | I slowly walk after. | Correct |
| s-188 | Bad students will not be making a noise. | Bad students noise make after not. | Correct |
| s-189 | You will not be touching this beautiful flower. | You beautiful flower touch after not. | Correct |
| s-190 | I will not be walking slowly. | I slowly walk after not. | Correct |
| s-191 | Will they be cooking? | they cook after yes-no? | Correct |
| s-192 | Will they be cleaning the window? | they window clean after yes-no? | Correct |
| s-193 | Will he not be cooking? | he cook after not yes-no? | Correct |
| s-194 | Will he not be cleaning the window? | he window clean after not yes-no? | Correct |
| s-195 | Will black students be eating this food? | black students food eat after yes-no? | Correct |
| s-196 | Will you be touching this beautiful flower? | you beautiful flower touch after yes-no? | Correct |
| s-197 | Will I be walking slowly? | I slowly walk after yes-no? | Correct |
| s-198 | Will bad students not be making a noise? | bad students noise make after not yes-no? | Correct |
| s-199 | Will teachers not be giving good grades? | teachers good grades give after not yes-no? | Correct |
| s-200 | Will I not be walking slowly? | I slowly walk after not yes-no? | Correct |
| s-201 | I will have spoken. | I speak full after. | Correct |
| s-202 | He will have opened the door. | He door open full after. | Correct |
| s-203 | I will not have stopped. | I stop full after not. | Correct |
| s-204 | He will not have opened the door. | He door open full after not. | Correct |
| s-205 | Little boy will have opened the door. | Little boy door open full after. | Correct |
| s-206 | He will have closed the brown window. | He brown window close full after. | Correct |
| s-207 | I will have spoken loudly. | I speak full after. | Incorrect |
| s-208 | Little boy will not have opened the door. | Little boy door open full after not. | Correct |
| s-209 | He will not have shut the brown window. | He brown window shut full after not. | Correct |
| s-210 | I will not have spoken loudly. | I speak full after not. | Incorrect |
| s-211 | Will you have slept? | you sleep full after yes-no? | Correct |
| s-212 | Will you have spent a dollar? | you dollar spend full after yes-no? | Correct |
| s-213 | Will you not have slept? | you sleep full after not yes-no? | Correct |
| s-214 | Will you not have changed clothes? | you clothes change full after not yes-no? | Correct |
| s-215 | Will little boy have opened the door? | little boy door open full after yes-no? | Correct |
| s-216 | Will he have closed the brown window? | he brown window close full after yes-no? | Correct |
| s-217 | Will I have spoken loudly? | I speak full after yes-no? | Incorrect |
| s-218 | Will little boy not have opened the door? | little boy door open full after not yes-no? | Correct |
| s-219 | Will he not have closed the brown window? | he brown window close full after not yes-no? | Correct |
| s-220 | Will I not have spoken loudly? | I speak full after not yes-no? | Incorrect |
| s-221 | You will have been sleeping. | You sleep full after. | Correct |
| s-222 | You will have been spending money. | You money spend full after. | Correct |
| s-223 | You will not have been sleeping. | You sleep full after not. | Correct |
| s-224 | You will not have been spending money. | You money spend full after not. | Correct |
| s-225 | Little boy will have been opening the door. | Little boy door open full after. | Correct |
| s-226 | They will have been taking beautiful pictures. | They beautiful pictures take full after. | Correct |
| s-227 | I will have been speaking louldy. | I louldy speak full after. | Correct |
| s-228 | Little boy will not have been following orders. | Little boy orders follow full after not. | Correct |
| s-229 | He will not have been driving a red car. | He red car drive full after not. | Correct |
| s-230 | I will not have been speaking loudly. | I loudly speak full after not. | Correct |
| s-231 | Will you have been sleeping? | you sleep full after yes-no? | Correct |
| s-232 | Will you have been creating trouble? | you trouble create full after yes-no? | Correct |
| s-233 | Will you not have been sleeping? | you sleep full after not yes-no? | Correct |
| s-234 | Will you not have been taking notes? | you notes take full after not yes-no? | Correct |
| s-235 | Will little boy have been watching cars? | little boy cars watch full after yes-no? | Correct |
| s-236 | Will they have been growing big vegetables? | they big vegetables grow full after yes-no? | Correct |
| s-237 | Will I have been speaking loudly? | I loudly speak full after yes-no? | Correct |
| s-238 | Will little boy not have been watching cars? | little boy cars watch full after not yes-no? | Correct |
| s-239 | Will they not have been growing big fruits? | they big fruits grow full after not yes-no? | Correct |
| s-240 | Will I not have been speaking loudly? | I loudly speak full after not yes-no? | Correct |

# Appendix D: Paragraph level Test Results

**Thirsty Crow version1**One hot day, a thirsty crow flew all over the fields looking for water. He did not find any water for a long time. He felt very weak, almost lost all hope. Suddenly, he saw a water jug below the tree. He flew straight down to see if there was any water inside. Yes, he could see some water inside the jug. The crow tried to push his head into the jug. Sadly, he found that the neck of the jug was too narrow. He tried to push the jug so that water flows out, but the jug was too heavy. The crow thought hard for a while. When he looked around, he saw some pebbles. Suddenly, He had a good idea. He picked up the pebbles, one by one, and dropped them into the jug. When pebbles filled the jug, the water level rose up. The crow could drink the water now. His plan had worked.

was One hot day , thirsty crow over fields for water look fly. was He for long time water find not .was He very weak feel , almost hope lose. was Suddenly , he below tree water jug see. was He fly straight down see if was there water inside. he inside jug water see. was crow into jug his head push try. was he find that was of jug neck too narrow. was He jug push try but was jug too heavy. was crow for while hard think. When was he look around, was he pebbles see. was Suddenly , He good idea had. was He the pebbles one by one picked up , into jug them drop and. When was pebbles jug fill , was water level rise up. crow now water drink. was His plan work full.

Result :

14/16 = 87 %

**Thirsty Crow version 2**There was a crow. He was very thirsty. He flew here and there in the search of water. At last, he reached a garden. He saw a pitcher there. The water was very low in it. His beak could not reach the water. He put some pebbles in the pitcher. The water rose up. Now his beak could reach the water. The crow drank water and flew away.

**Output:**was There crow. was He very thirsty. was He in of water search here and there fly. was At last , he garden reach. was He there pitcher see. was water in it very low. His beak water reach not. was He in pitcher pebbles put. was water rise up. his beak water reach. was crow water drink and fly away.

Result:

100% accuracy.

**Greedy Dog**There was a dog. He was very hungry. He got a piece of meat from a butcher's shop. He took that piece in his mouth and ran to the jungle. He was crossing the bridge of a stream. He looked into the water. He saw his reflection in it. He thought that a second dog was holding a piece of meat in the water. He opened his mouth and barked at the dog. His own piece of meat fell down into the water. He lost his own piece of meat too.

**Output:**was There dog. was He very hungry. was He from butcher shop of meat piece get. was He in his mouth piece take and to jungle run. was He of stream bridge cross. was He into water look. was He in it his reflection see. was He think that was second dog of in water meat piece hold. was He his mouth open and at dog barked. was of meat His own piece into water fall. was He of meat his own piece too lose.

Result:

100% accuracy.

**Clever Fox**One day, a hungry fox was wandering in the forest. She was searching for food. Suddenly, she saw a crow sitting in a tree. It was holding a piece of meat in its beak. The big piece of meat attracted the hungry fox. She sat down under the tree and thought about some clever trick to get the meat. Suddenly, a clever idea came to her mind. She began to praise the crow. She said that crows are very lovely birds and they have beautiful voice. The crow got trapped in her trick. He looked down to the fox. Fox requested him to sing her a song. When the crow opened his beak to sing, the piece of meat fell down. The fox took the meat and ran away. She had made a fool of the crow.

**Output:**was One day , hungry fox in forest wander. was She for food search. was Suddenly , she crow in tree sit see. was It of in its beak meat piece hold. was of meat big piece hungry fox attract. was She under tree sit and about clever trick think meat get. was Suddenly , clever idea to her mind come. was She crow praise begin. was She say that crows very lovely birds and they beautiful voice have. was crow in her trick get. was He to fox look. was Fox him her a song sing request. When was crow his beak open , was of meat piece fall down. was fox meat take and run away. was She of crow fool make full.

Result:

100% accuracy.