Literature Review

SNo	Authors	Title	Major findings	Limitations
1	Fredrik Jørgensen, 2007	Clause Boundary Detection in Transcribed Spoken Language	Partially Solving the clause detection by exploiting the special role coordinating conjunctions play and reducing clause boundary detection to a classification problem.	Results reported in this paper onlystates how many of the conjunctions are classified correctly. It does not state anything about the proportion of clause boundaries.
2	Sanjeev Kumar Sharma,2016	Clause Boundary Identification for Different Languages: A Survey	Each language uses its own methods and techniques for identification and making sentences simpler for use. Various techniques such as Conditional Random Field (CRF), Support Vector Model (SVM), HiddenMarkov Model (HMM), Rule based approaches have been used for identification of clauses.	A survey
3	Heiki-Jaan Kaalep, Kadri Muischnek	Robust clause boundary identification for corpus annotation	Detecting an "ordinary" clause boundary (recall 88%, precision 86%) is an easier task than detecting an embedded clause (recall 69%, precision 90%)	The system needs information about parts of speech and grammatical categories coded in the word-forms, i.e. it takes morphologically annotated text as input.

SNo	Authors	Title	Major findings	Limitations
4	Rahul Sharma, Soma Paul	A rule based approach for automatic clause boundary detection and classification in **Hindi**	The System performance for finite clauses in terms of F1 scores of 91.53% for clause boundary identification and 80.63% for clause Classification, while giving inadequate results for nonfinite clauses with 60.57% accuracy.	System doesn't handle classification of different instances of 'to' (else, then, or etc.) and of coordination where a punctuation serves as a coordinator.
5	Daraksha Parveen, Ratna Sanyal, and Afreen Ansari	Clause Boundary Identification Using Classifier and Clause Markersin **Urdu** Language	Conditional Random Fields are used for classification of clause boundary beginning and ending andalso detecting the type of subordinate clause. Here, linguistic rules are given higher priority, hence misclassification is corrected via these rules.	It is highly dependent on linguistic rules. Missing of these rules may lead to wrongly classified data. An improvement can be achieved in the proposed clause boundary identifier by including more sophisticated linguistic rules, clause markers for different subordinate clauses and also for those clauses which are embedded in the main clause.
6	Poornima C, Dhanalakshmi V,Anand Kumar M, Soman K P	Rule based Sentence Simplification for English to **Tamil** Machine Translation System	Rule based technique is proposed to simplify the sentences based on connectives like relative pronouns,coordinating and subordinating conjunction. Sentence Simplification is expressed as the list of sub-sentences that are portions of the original sentence.	200 sentences were tested with simplification. 115 sentences are translated correctly after simplification.

SNo	Authors	Title	Major findings	Limitations
7	Richard Evans, Constantin Orasan	Relative clause extraction for syntactic simplification	Investigates non–destructive simplification, a type of syntactic text simplification which focuses on extracting embedded clauses from structurally complex sentences and rephrasing them without affecting their original meaning.	It is highly dependent on linguistic rules. Missing of these rules may lead to wrongly classified data.
8	Sandeep Saini ,Vineet Sahula,Umang Sehgal	Relative clause based text simplification for improved English to **Hindi** translation	Uses syntactic parse tree to analyze English sentences and extract relative clauses which are then used for translation purposes.	Topicalization and Inconsistencies in the treebank were some of issues.Accuracy for Start of clause 97.91% and End of Clause 94.44%
9	Anand Kumar M	Clause Boundary Identification for **Tamil** Language	Uses the dependency parser to identify the boundary for the clause. The dependency tag set contains 11 tags, and is useful for identifying the boundary of the clause along with the identification of the subject and object information of the sentence.	The developed model showed better accuracy for the sentences of smaller length.
10	Rahul Sharma, Soma Paul, Riyaz Ahmad Bhat and Sambhav Jain	Automatic Clause Boundary Annotation in the **Hindi** Treebank	Showed how implicit clausal information is captured in a dependency tree. Used HDTB and automatically added the clausal information using the dependencies between constituents in the treebank.	Obtaining a correct parse tree is a difficult task so all the parse trees have to be manually checked for errors before operating on them.