

Paper Title	Publication Year and Author	Scope	Methods used	Evaluation Criteria	Evaluation Results	Difficulties Faced / Disadvantages	Future Avenues	Datasets Used
A Bangla Semantic Parser using Context-Free-Grammar [1]	M. F. Mridha et al. [2017]	<b>Target Population:</b> Bengali language  <b>Scope of Studies:</b> Determining the meaningfulness and grammatical structure of the language, specifically for the purpose of natural language processing tasks.	Using a three-stage procedure [preprocessing, a linguistic database, and the development of context-free grammar rules], semantic parsing is performed.	Authors have employed a measure of accuracy to assess the efficacy of proposed approaches. The accuracy of the system was determined by evaluating how well the input sentences aligned with the predetermined rules, and by addressing any potential ambiguities within the sentences.	91.57% accuracy in matching input sentences with defined rules, resolving ambiguities.	Inadequate language pre-processing mechanisms, determining meaningful sentences and grammatical structural inconsistencies of Bengali language.	Development of a semantic parser using semantic rules and Bangla WordNet.	Over 10,000 Bengali sentences. [No information on dataset availabilities]
A Rule-Based Bengali Grammar Checker [2]	A.N.M Fahim Faisal et al. [2021]	<b>Target Population:</b> Bengali corpus for grammar  <b>Scope of Studies:</b> development of rule-based Bengali grammar checker tool.	Used Tokenization and POS tagging. Then Rule matching for the identification of Adjective Phrases and Noun Phrases.	The system's performance was measured by comparing its output with a set of manually annotated sentences.	The average accuracy of the grammar checking system was 86%.	<b>Not Mentioned</b>	An improved POS tagger could be introduced in the system which will make the system much more efficient. It was suggested that a system could be created using the statistical method.	To test the grammar checking system, two sets of 50 sentences was used. No information about training data was mentioned. [No information on dataset availabilities]
An Empirical Framework for Parsing Bangla Assertive, Interrogative and Imperative Sentences [3]	Mohammed Safayet Arefin et al. [2015]	<b>Target Population:</b> Different Types of Bengali Sentences.  <b>Scope of Studies:</b> Parsing Bengali sentences, specifically assertive, interrogative, and imperative sentences.	The researchers used Context-sensitive grammar [CSG] approach using four stages: Input sentence, Tokenizer, Lexicon, and Rule Generator.	The success rate of parsing different types of sentences [assertive, interrogative, and imperative] against test dataset hence accuracy.	The researchers state that 83.09% of the 420 tested sentences generated correct outputs. It also provides success rates for different types of sentences [assertive: 81.2%, interrogative: 85%, imperative: 80%].	The difficult grammar of Bengali language requires a specific domain expert and lack of structured data was a major obstacle for the authors.	The authors mentioned that the clear lexicon representation and accurate parsing of sentences can be used for preprocessing of other NLP tasks to further improve the accuracy.	420 Bengali sentences of different types. [No information on dataset availabilities]
An exploratory research on grammar checking of Bangla sentences using statistical language models. [4]	Rahman, Md Riazur et al. [2020]	<b>Target Population:</b> Different Types of Bengali Sentences.  <b>Scope of Studies:</b> Bengali grammar checking and performed a comparative performance analysis on previous research	Authors proposed an improved method using Kneser-Ney smoothing based n-gram language model for grammar checking and performed a comparative performance analysis between Kneser-Ney and Witten-Bell smoothing techniques for the same purpose	The system's performance was measured by precision, recall and f-score.	The proposed method performed f-scores of 72.92% and 68.51% for grammatical and ungrammatical data respectively	<b>Not Mentioned</b>	The authors suggested more features such as parts of speech tags and other linguistic features can be added to improve the performance of the system	Authors used a large corpus of 20 million words containing 181820 grammatically correct sentences. Around 80% of the corpus is used for training purposes. The validation set consists of 20000 correct sentences and 20000 incorrect sentences. [No information on dataset availabilities]
An Investigative Design Based Statistical Approach for Determining Bangla Sentence Validity [5]	Md. Riazur Rahman et al. [2016]	<b>Target Population:</b> Bengali language sentence  <b>Scope of the Studies:</b> Improvement in performance for grammar checking in the Bengali language	Authors have used statistical method using an n-gram based language model with Witten-Bell smoothing and Backoff language modeling strategy.	Investigators have used precision and recall as their validation techniques.	Authors found that 4-gram model performing best among other methods and proposed system performing with 82% precision and 81% recall.	Lack of resources for grammar checking in the Bengali language was one of the difficulties authors faced. Moreover, data scarcity and difficulty in annotation with linguistic complexity was also a big barrier.	Authors suggest some combination of linguistic information into their statistical system may produce better performance.	Corpus consisting of 10 million-word tokens of Bengali language [No information on dataset availabilities]
Bangla Natural Language Processing: A Comprehensive Analysis of Classical, Machine Learning, and Deep Learning Based Methods [6]	Ovishake Sen et al. [2022]	<b>Target Population:</b> Bengali-speaking people  <b>Scope of Studies:</b> Analysis of Bengali NLP tools and methods	The researchers classified different techniques into classical, machine learning and deep learning based approaches based on the type of NLP tasks	Techniques used for different NLP tasks in Bengali language	The researchers showed the usage pattern of different algorithms in this field and how those perform for different tasks and provided with a discussion of the possible future improvements	The authors indicated that there is a very limited support for Bengali language in regards to corpus and overall research work using different techniques.	<b>Not Mentioned</b>	Published research work within a timeline of 1999 to 2021, among which 50% were published after 2015. [No information on dataset availabilities]
Bangla Parts-of-Speech Tagging using Bangla Stemmer and	Md. Nesarul Hoque et al. [2015]	<b>Target Population:</b> Bengali language.	The methodology includes stemmer development, rule-based analysis, and	Accuracy of correctly identifying the parts of speeches of different word tokens	The accuracy of the proposed system was 93.7% where it identified	Authors pointed out that the method they have used is computationally	<b>Not Mentioned</b>	8155 word tokens for POS tagging. [No information on dataset availabilities]

Rule based Analyzer [7]		<b>Scope of Studies:</b> To develop an automated POS tagger for the Bengali language.	the use of Bangla vocabulary and a Bangla verb dataset.		7637 POS correctly out of 8155.	expensive for implementation. However they have also faced difficulties because of lack of quality training data		
Bangla Syntax Analysis: A Comprehensive Approach [8]	Lenin Mehedy et al. [2003]	<b>Target Population:</b> Bengali language Syntax  <b>Scope of Studies:</b> Bengali Syntax Analysis.	Using Context Free Grammar to generate rules for parsing Using three stages for generation: lexical analysis, syntax analysis and semantic analysis	Capacity of the proposed model to produce proper parsing of Bengali sentences	The researchers did not provide any numerical metrics for the result analysis.	According to the author it has its limitations such as the constraint in space that prevented the inclusion of grammar rules to handle various Bangla punctuation symbols, the focus on syntax analysis only and the lack of analysis on concepts such as voice, narration, and other special concepts of Bangla grammar. Additionally, other types of clauses such as adverbial or adjective clauses were not parsed further.	Author suggested future researchers work with what they have left or didn't consider during their studies.	<b>Not Mentioned.</b>
Building a Foundation of HPSG-based Treebank on Bangla Language [9]	Altaf Mahmud et al. [2007]	<b>Target Population:</b> Bengali corpus for grammar  <b>Scope of Studies:</b> Development of large and syntactically annotated corpus for Bengali.	The methodology for implementing a Bangla language grammar involves customizing the Lingo Grammar Matrix as the starting point. The grammar is developed on the LKB system by modifying the matrix. The word order in Bangla is free due to the nonconfigurational nature of the language and uses the "Monosydeton" coordination strategy.	Capacity of the grammar representation to produce syntactical construction of correct sentences.	The authors claimed that the HPSG-based treebank could generate syntactical construction of a large number of individual Bengali sentences and that the annotator objects of word ordering could navigate quickly through the parse forest to identify the correct or preferred analysis.	The authors identified that there were a number of ambiguous parsings and those parsing's were unpredictable.	The authors mentioned that they were still reviewing and improving the proposed system to reduce these unpredictable behaviors and improve the accuracy.	<b>Not Mentioned.</b>
Complexity Guided Active Learning for Bangla Grammar Correction [10]	Bibekananda Kundu et al. [2013]	<b>Target Population:</b> Bengali corpus for grammar  <b>Scope of Studies:</b> complexity guided active learning for Bengali grammar correction.	An example-base of sentences and their complexity scores is used where users are asked to rate sentence complexity in a multiple-choice format. Complexity estimation is based on the k-NN algorithm and user input, and an active learning procedure is used to improve accuracy. The system provides explanations, receives feedback, and updates its example-base	During training, sentences of different complexity are provided to the user in a multiple-choice format, and the system estimates the complexity score of the same sentence based on the extracted feature values and k-NN algorithm.	The active learning-based complexity estimation prototype which is available online has been found to have a relative error of 0.16.	<b>Not Mentioned</b>	<b>Not Mentioned</b>	Bangla Unicode sentences were collected from multiple sources including Bengali online newspapers ["Ananda Bazar Patrika"], Bengali literature written by Rabindranath Tagore, Sarat Chandra Chattopadhyay and Bankim Chandra Chattopadhyay and Bangla blogs. [No information on dataset availabilities]
Detection of Semantic Errors from Simple Bangla Sentences [11]	K. M. Azharul Hasan et al. [2014]	<b>Target Population:</b> Bengali corpus for grammar.  <b>Scope of Studies:</b> detection of semantic errors in the Bengali language.	Categorization of words based on semantic relationships and creation of categorization tables for identifying classes. Validation and checking of relationships between subjects, verbs, and objects. Construction of a semantic knowledge base. Lastly, an automatic error-detection algorithm is developed	The testing was done using sample contents built by experts. The performance of the technique relies on the validation and object-verb relation tables for all the verbs and nouns.	Testing using sample contents built by experts showed that out of 6 sentences in the SOV format, one had a semantic error and the rest were correct.	<b>Not Mentioned</b>	<b>Not Mentioned</b>	<b>Not Mentioned</b>
Feature Unification for Morphological Parsing in Bangla [12]	Sajib Dasgupta et al. [2004]	<b>Target Population:</b> Derivationally Bengali complex words.  <b>Scope of Studies:</b> Feature Unification based Grammar model.	Feature-based approach that uses morphotactic and orthographic rules and lexicon to tokenize an input word into morphemes. A chart parser to generate the parse tree, and feature unification and feature collection are performed.	Improvements of results of the existing techniques by using morphological parser	The paper discussed how the process of grammar correction can be improved by using the feature unification techniques for morphological parsing.	The authors could not demonstrate any numeric evaluation of their proposed methodology.	The authors proposed that it is important to further research this approach to improve the process and gather quantitative evidence.	<b>Not Mentioned</b>
Graph theoretic interpretation of Bangla traditional grammar [13]	Samir Karmakar et al. [2016]	<b>Target Population:</b> Bengali grammar representation.  <b>Scope of Studies:</b> Graph	The study proposes a methodology for capturing syntactico-semantic peculiarities of the individual constituents of a sentence using graph theory. The	Prospect of improvement in case of graph theoretical representation of Bengali Grammar	The paper discussed the possibilities that instead of using distinct vertices/nodes to represent syntactic and semantic	The authors identified that the identification of a vertex as either maximal or elaborator has nothing to do with their relative hierarchies in a tree	The authors proposed a new possible grammar representation which can improve the performance of	<b>Not Mentioned</b>

		theoretic interpretation of Bengali Grammar.	methodology includes substituting identical vertices with a single alphabet and capturing dissimilarities in terms of their respective syntactic and semantic properties through certain conceptual measures.		properties, one can distinguish non-identical vertices in terms of their respective maximality. This can be achieved by substituting identical vertices with a single alphabet and capturing dissimilarities in terms of conceptual measures.		the grammar correction methods and proposed that this approach can be improved further.	
Identifying the Writing Style of Bangla Language Using Natural Language Processing [14]	Syeda Farjana Shetu et al. [2020]	<b>Target Population:</b> Bengali written text. <b>Scope of Studies:</b> Bengali writing style identification using NLP techniques.	Using a rule based model for data processing which involves tokenization and creating a Sadhu extensions library	Accuracy of the model in identifying writing style of paragraphs [sadhu or cholito]	The research reported a classification accuracy of 90% for Sadhu style and 82% for Cholito style.	Data Scarcity was one of the difficulties the authors mentioned. Other than that annotation difficulty and linguistic complexities was also a barrier to authors.	<b>Not Mentioned</b>	The research collected 60 paragraphs from Tagore, 40 paragraphs from Nazrul, 50 paragraphs from a newspaper, and 50 paragraphs from Humayun Ahmed and Mohammad Zafar Iqbal. The researchers also collected various newspaper articles in Bengali language.
In search of a suitable method for disambiguation of word senses in Bengali [15]	Alok Ranjan Pal et al. [2019]	<b>Target Population:</b> Polysemous words in Bengali language texts. <b>Scope of Studies:</b> Studies related to Word Sense Disambiguation [WSD] in Bengali.	Lesk algorithm with a context expansion feature in a lemmatized environment with sentence clustering and similarity measuring.	Accuracy of different techniques for meaning identification	Researchers compared the performance of different approaches and mentioned that combining the supervised and knowledge-based methodologies resulted in the highest accuracy of 92%.	Linguistic complexity was a major barrier for authors.	Combining supervised and knowledge-based approach to have a better understanding of the complex nature of the language	
Incorporating Supervised Learning Algorithms with NLP Techniques to Classify Bengali Language Forms [16]	Abdul Bari Parves et al. [2020]	<b>Target Population:</b> Bengali literature <b>Scope of Studies:</b> Classification of Sadhu-bhasha and Cholito-bhasha in Bengali literature.	Bag of Words, Term Frequency-Inverse Document Frequency [TF-IDF] was used with supervised learning algorithms.	Accuracy, precision, recall, AUC score and F1 score.	Multinomial Naive Bayes [MNB] performed the best with an accuracy of 99.95% and the highest scores for precision, recall, AUC score, and F1 score.	Lack of Training Data	<b>Not Mentioned</b>	<b>Not Mentioned</b>
Performance Analysis of Different Word Embedding Models on Bangla Language [17]	Zakia Sultana Ritu et al. [2018]	<b>Target Population:</b> Bengali word clustering techniques. <b>Scope of Studies:</b> Focus on comparing the performance of three dynamic word clustering models for producing accurate Bengali word clusters.	Dynamic word clustering models were employed by the authors.	Accuracy and efficiency was used as an evaluation criteria.	FastText-Skip Gram was found to be the most accurate and efficient model for building Bengali word clusters. limitations	Lack of work in determining an efficient and accurate method for producing word clusters in the Bangla language using dynamic models.	FastText-Skip Gram model suggested as the most accurate and efficient model for building Bangla word clusters if the dataset is prepared properly	<b>Not Mentioned</b>
Recognizing Bangla grammar using predictive parser [18]	K. M. Azharul Hasan et al. [2011]	<b>Target Population:</b> Bengali language. <b>Scope of Studies:</b> Context-Free-Grammar and Predictive Parser for grammar checking.	Using Left-Factoring Techniques for developing CFG and developing Predictive parser using CFG.	Total number of sentences/paragraphs, the number of sentences/paragraphs correctly detected, and the acceptance rate	The success rate for traditional sentences was 83.33%, with 100 out of 120 sentences being correctly detected. For nontraditional sentences, the success rate was 72.5%, with 58 out of 80 sentences being correctly detected. For paragraphs, the success rate was 72%, with 18 out of 25 paragraphs being correctly detected	Low Accuracy for non-traditional sentences and paragraphs Lack of quality training corpus	Proposal for introducing advanced NLP techniques to improve accuracy and efficiency	<b>Not Mentioned</b>
Semantic Role Labeling for Bengali Noun using 5Ws [19]	Amitava Das et al. [2010]	<b>Target Population:</b> Bengali corpus development research. <b>Scope of Studies:</b> Building annotated corpus and acquiring tools	Using a hybrid approach for annotating Bengali nouns [machine learning techniques added with rule based limitations] and a rule based post-processor to reduce false hits	Precision, Recall and F-measure of the systems for successful annotation of the nouns	The results indicate that the rule-based post-processor has improved the performance of the system, as evidenced by the increase in Precision, Recall, and F-Measure values.	Resource shortage of Bengali language in the form of a digitized corpus and lack of research in different NLP tasks in Bengali language.	Rule based post processing for reducing wrong identification and mapping the 5W rules to gain better insight into annotation techniques.	The researchers used an annotated corpus for training the model in annotating the nouns

		for feature extraction of Bengali language.			The F-Measure values are provided for each semantic role tag, indicating the performance of the system for each tag.			
Structural Analysis of Bangla Sentences for Machine Translation [20]	Md. Musfique Anwar et al. [2010]	<b>Target Population:</b> Bengali language corpus  <b>Scope of Studies:</b> Bengali sentence analysis with machine translation using fuzzy logic.	Generating structural representation, language modeling using a bilingual corpus, calculating fuzzy membership values.	The success rate of the designed machine translation system	The success rate for translating simple sentences was around 98%, for both complex sentences and compound sentences, it was around 95%.	Lack of a corpus-based automatic Bangla machine translator, syntactic transfer and generation of simple, complex, and compound sentences, extracting additional information from sentences	Fuzzy logic-based approach, parsing the finite verb for extracting information about tense, person, and mode of the verb	Bengali Natural sentences and Bilingual corpus [Bengali and English]
A Generalized Lexical Functional Grammar-Based Processing of An Indian Language - Bangla [21]	P. Sengupta et al. [1998]	<b>Target Population:</b> Bengali language.  <b>Scope of Studies:</b> Effective computer implementation of four Lexical Functional Grammar [LFG]-based parsers targeting Indian languages.	Identification of forward reference points by introduction of a new metavariable and broadening the scope of the find operator. Forward reference precipitation using special schemata termed "m^structure" schemata projected by the verb.	The approach was put to the test using a random sample of different sentences and the successful interpretation rate was calculated.	All simple sentences in active voice [144, or approximately 57% of the samples] were successfully interpreted. Most of the neuter voice [total 76 sentences [i.e. about 30% of the samples]] sentences were also correctly parsed. From the 28 [app. 11%] active-voice sentences with embedded clauses, barring those which used verb forms not already stored in the lexicon, the remaining were correctly parsed. The parser in its present form could not parse neuter voice sentences with embedded clauses that constituted the remaining 2% of the samples.	The parser could not interpret neuter voice phrases with embedded clauses.	The author stated the best possible way for preventing wrong acceptance in parsing is to make the explicit rules as strict as possible in the future.	A random sample of 252 simple and difficult sentences drawn at random from newspaper clippings.
Probabilistic Approach of Parsing Bengali Sentences [22]	Ayesha Khatun et al. [2018]	<b>Target Population:</b> Bengali language.  <b>Scope of Studies:</b> probabilistic context-free grammars [PCFGs] based on the intonation of sentences and structures used for Bayesian inference.	This paper proposed five methods with input and output representation. They are Syntax analyzer, Rule generator, Statistical parser, Error Handler, and Lexicon.	Accuracy of predicting the probability of the parser for all kinds of Bengali sentences	The paper achieved an accuracy of 85% for all kinds of Bengali sentences. Highest Probability: The highest probability of the parse tree is 1.65E-8%.	The problem faced by the paper, which it aims to avoid through statistical parsing		3500 sentences [10648 words] for testing purposes.
Text Normalization for Bangla, Khmer, Nepali, Javanese, Sinhala, and Sundanese TTS Systems [23]	Sodimana et al. [2018]	<b>Target Population:</b> Low Resource languages.  <b>Scope of Studies:</b> Text normalization tasks for low-resource languages	Text normalization system using the "Thrax" grammar compiler to generate finite-state transducers from grammars.	Precision, recall, and F-score measures on test datasets for each language.	Not explicitly mentioned in the review.	Data scarcity for low-resource languages was a prime difficulty faced by the authors.	The paper shows that the system performs well for most semiotic classes and languages, but there is still room for improvement in some cases using more advanced concepts and quality dataset.	<b>Not Mentioned</b>
Transformational Generative Grammar [TGG]: An Efficient Way of Parsing Bangla Sentences [24]	Mohammad Kamrul Huq Maroof et al. [2016]	<b>Target Population:</b> Bengali corpus for grammar.  <b>Scope of Studies:</b> Development of an effective parser for Bengali language.	Tokenization and implementation of lexicon with rule generation and parsing algorithm and parsing with TGG	The proposed natural language parsing system was tested and testing criteria include success rate for assertive, interrogative, imperative, optative sentences, and exclamatory sentences.	The success rate for assertive, interrogative, imperative, optative sentences, and exclamatory sentences are 84%, 90.5%, 80%, 87.5%, and 77% respectively. The overall success rate of the proposed system is 84.4%.	The use of TGG in BLP is still at its rudimentary stage and the implemented system mainly deals with parsing structurally simple sentences. The system is not able to handle more complex sentences and more research needs to be done to improve this.	Future study is possible for parsing of complex and compound sentences that can improve the system.	770 different kinds of sentences collected from Bangla grammar books.
Unsupervised morphological	Sajib Dasgupta et al. [2007]	<b>Target Population:</b>	Unsupervised morphological parsing	Accuracy and F-score was used as an	The research states that the proposed	The research mentions several linguistic	<b>Not Mentioned</b>	A large, unannotated corpus for inducing

parsing of Bengali [26]		Bengali Language. <b>Scope of Studies:</b> Unsupervised morphological parsing in Bengali.	method, including the steps involved in inducing prefixes, suffixes, and roots, as well as the scoring metrics used for ranking candidate affixes.	evaluation matrix.	unsupervised morphological parser outperformed the baseline system in terms of exact accuracy and F-score.	characteristics of Bengali that complicate its automatic processing, such as highly productive morphology, a large number of compound words, and a more or less free word order.		prefixes, suffixes, and roots
A Proficient autonomous Bangla Semantic Parser for Natural Language Processing [27]	M. F. Mridha et al. [2015]	<b>Target Population:</b> Bengali corpus for grammar. <b>Scope of Studies:</b> Developing autonomous Bengali Semantic Parser for Natural Language Processing.	Design of the Semantic Parser was completed with prior discussion of the known approaches. Creation of semantic parser rules to employ rules using the rule-based technique	During training and testing data, the words were classified according to the part of speech and the accuracy of each type of sentences were calculated.	The accuracy of 86.25% was achieved from the Semantic Parser.	The accuracy obtained initially was declining during the experiment when ambiguous words and sentences were being added.	Further testing with the model can be planned for future work with a larger corpus and added features and variables.	A total of over 2000 short sentences were chosen for testing, and they contain over 9250 tagged terms.
An Empirical Framework of Idioms Translator From Bengali to English: Rule Based Approach [28]	Ayesha Khatun et al. [2007]	<b>Target Population:</b> Bengali corpus for grammar. <b>Scope of Studies:</b> translating Bengali to English by offering a translation method for idioms found in sentences.	Development of an idioms checker and an idioms translator, a tokenizer, a rule generator, and a parser. Characterizing rules for the source and destination languages Development of a machine translator and generator	With roughly 15000 sentences, the suggested system was put into practice and tested.	The results of the trial demonstrated that the suggested method provided an accuracy of 85.33%.	The authors identified that for some sentences, the system could not receive the exact parse tree. However, Too simple and short Bangla sentences were chosen to test the parsing model.	It was feasible to create a more robust parser for Bangla sentences to update CSG rules. This can be accomplished by using semantic characteristics in further study.	With roughly 15000 sentences, the suggested system was put into practice and tested.
An investigative design of optimum stochastic language model for bangla autocomplete [29]	Md.Iftakher Alam Eyamin et al. [2019]	<b>Target Population:</b> Bengali grammar representation. <b>Scope of Studies:</b> design of optimum stochastic language model for Bengali autocomplete.	Implementation of a novel method for determining the best language model based on performance indicators using a big corpus of distinct word kinds in Bangla. Stochastic word prediction on Bangla sentences using N-gram-based language models. Automatic completion of a phrase by forecasting a group of words instead of a single word.	A thorough research was carried out to assess the effectiveness of the classifier for each model by adjusting the length of the test sentences, including unigram, bigram, trigram, backoff, and linear interpolation.	Here, the average accuracy of five languages i.e. unigram, bigram, trigram, backoff and linear interpolation were 38.62%, 79.12%, 83.38%, 85.96% and 91.58% respectively with the optimum prediction length being 7 [n=7] .	Over the course of the several experimental stages, the linear interpolation surpasses the other models in terms of accuracy and failure rate.	This research may be utilized to create adaptive software for Bangla automatic word completion, which may also be applied to Bangla language correction techniques in the future. Further testing with other well-known machine learning models can be planned for future work with a larger corpus.	The dataset was constructed from the Bangla newspaper the "Daily Prothom Alo". The corpus contains more than 12 million words and about 1 million [937,349] sentences, where the total number of unique words is 294,371 and average word length is 7.
Rule Based Replacement of Pronoun by Corresponding Noun for Bangla News Documents [30]	Md. Majharul Haque et al. [2017]	<b>Target Population:</b> Bengali language. <b>Scope of Studies:</b> Rule based replacement of pronoun by corresponding noun.	Usage of general and special tagging. Dependency parsing and full name identification was used with rule-based replacement of pronoun by corresponding noun.	Utilizing existing dataset, the system had been trained and tested. The evaluation criteria were success rate of word identification and percentage of words tagged.	Using the lexicon database and "SentiWordNet", 65.13% of words were generally tagged. With addition of a list of suffixes, 66.73% of the words were tagged. After special tagging was implemented, 72.53% of words were recognized. Later, with the implementation of Dependency parsing, 79.50% of words were recognized. The proposed method displayed 71.80% accuracy in the performance evaluation.	Only singular pronouns could only be modified in this study project and some pronouns were replaced improperly.	It is proposed that future work must focus on expanding this approach to make it more reliable and versatile.	A data file has been used here for first name, middle name and last name with 7500 entries. A table has been maintained with 700 entries for the list of division, district, upazila and municipality as name of places of Bangladesh [Bangladesh Post Office, 2016]. With 3,000 news documents and 200 news articles from Bangladesh, the system had been trained and tested using help from the lexicon database and SentiWordNet.
Analysis of Bangla Transformation of Sentences Using Machine Learning [31]	Rajesh Kumar Das et al [2023]	<b>Target Population:</b> Bengali Sentences. <b>Scope of Studies:</b> Transformation of Sentences Using Machine Learning.	The model accepts a Bangla sentence as input, determines the sentence construction type, and outputs the sentence type. Total six supervised machine learning algorithms were used to classify three types of sentence formation: Sorol Bakko [simple sentence], Jotil Bakko [complex sentence] and Jougik Bakko[compound sentence]	The dataset is analyzed and accuracy, precision, recall, f1-score and confusion matrix was calculated..	The highest accuracy with the decision tree classifier, which is 93.72%.	<b>Not Mentioned.</b>	<b>Not Mentioned.</b>	Total 2727 numbers of data from various sources
Statistical parsing of Bangla sentences by	Ayesha Khatun et al. [2017]	<b>Target Population:</b> Bengali	Statistical parser using probabilistic version of Cocke-Younger-	The ratio of the total number of sentences is applied in the system	The average correctness of a basic sentence is	The proposed system performed better on simple sentences than	Proposal to include statistical parsing with	Sentences were collected from popular newspapers, textbooks, blogs, and

CYK Algorithm [32]		language. <b>Scope of Studies:</b> Statistical parsing of Bengali sentences by CYK Algorithm for grammar checking.	Kasami [CYK] algorithm and left binarization technique to grammar is used to improve parsing efficiency. Rule probability and word probability is used to generate different probabilities	and the number sentences which were successfully parsed.	92.75%, while the average accuracy of a difficult statement is 83.75%. The average accuracy of a compound sentence was 76.66%.	the complex and compound sentences. It could not get the same probability for the same structure as the model considered both lexicon probability and rule probability.	lexicalized PCFGs. Inclusion of more Bangla idioms and phrases.	literature of Bangla. Total 2025 of different kinds of sentences.
Development of Bangla Spell and Grammar Checkers: Resource Creation and Evaluation [33]	Nahid Hossain et al [2021]	<b>Target Population:</b> Bengali Grammar. <b>Scope of Studies:</b> Development of Bengali Spell and Grammar Checkers.	The spell checker uses the Double Metaphone algorithm and Edit distance based on the distributed lexicons and numerical suffix dataset to detect all types of Bangla spelling mistakes	Accuracy was used as evaluation metrics.	The spell checker is 97.21% and the grammar checker is 94.29% .	The lack of adequate corpus and lexicon is one of the major hindrances that slowed down the progress of NLP researchers in Bangla language	<b>Not Mentioned.</b>	Over 100 million words from reputed and diversified online sources
Relation between Subject and Verb in Bangla Language: A Semantic Analysis [34]	Mozammel Haque et al [2016]	<b>Target Population:</b> Bengali language. <b>Scope of Studies:</b> Find out the main verb in a sentence that is semantically valid for a certain subject during machine translation.	Authors used a subject-verb algorithm.	Accuracy is used as an evaluation criteria.	The algorithm gives appropriate results for 598 sentences between 600 sentences. The accuracy of the algorithm is 99.67% during testing.	Bangla language has two types of format for writing; [Shadhu] and [Colito]. Bangla follows Subject + Object + Verb [SOV] structure and English follows Subject + Verb + Object [SVO] structure.	In future, the proposed algorithm will be useful in various fields of computational linguistics for Bangla language.	<b>Not Mentioned.</b>
NLG Approach for Bangla Grammatical Error Correction [35]	Bibekananda Kundu et al. [2014]	<b>Target Population:</b> Bengali Grammar. <b>Scope of Studies:</b> The procedure for the automatic detection and correction.	Authors used lemmatization with over generation optimization and best path selection	Authors used precision, recall and f-score	This paper tested 4 lakh 80 thousand collected well-formed Bangla sentences.	Authors have faced sparseness problem and constraint satisfaction problem.	To avoid the sparseness problem of data Jelinek Mercer Smoothing can be applied. Using linguistic fitness functions to optimize the search space can also be applied.	For this analysis, Bangla well-formed Unicode sentences were collected from the web of various domains including literature, science, sports, music and news wire [2005-2010]
Banglalm: Data mining based bangla corpus for language model research [36]	Md Kowsher et al [2021]	<b>Target Population:</b> Bengali Language. <b>Scope of Studies:</b> Data mining based Bengali corpus for language model research.	The authors released two preprocessed versions of this dataset that is especially suited for training both core machine learning-based and statistical-based models.	Not related.	Not related.	Not related.	The proposed dataset will significantly contribute to the Bengali machine learning and NLP community	The dataset is available online at the link below. <a href="https://shorturl.at/aKV04">https://shorturl.at/aKV04</a>
Vacaspati: A Diverse Corpus of Bangla Literature. [37]	Bhattacharyya, Prमित et al. [2023]	<b>Target Population:</b> Bengali Language. <b>Scope of Studies:</b> Datasets for Bengali corpus	The authors offer Vacaspati: A Diverse Corpus of Bangla Literature. VACASPATI ˆ is varied from multiple aspects, including type of composition, topic, author, time, space, etc. It contains more than 11 million sentences and 115 million words. We also built a word embedding model, VAC ˆ - FT, using FastText from VACASPATI ˆ as well as trained an Electra model, VAC ˆ -BERT, using the corpus. VAC ˆ -BERT has far fewer parameters and requires only a fraction of resources compared to other state-of-the-art transformer models and yet performs either better or similar on various downstream tasks.	Authors used macro-F1 score as a performance measure.	The overall average performance was 78.27 according to the authors.	Investigators believe the dataset is unlikely to contain highly objectionable content as they used only published research for review.	Since VACASPATI is a corpus made from only literary data, VAC -BERT did not use newspaper or social media articles during pre-training. Most of the other BERT models had most likely used that data during pre-training. Hence, the performance of VAC -BERT on non-literary data is significant to establish the quality of VACASPATI as a corpus	Not related.
Preprocessing the Bengali Text: An Analysis of Appropriate Verbs [38]	Mozammel Haque et al. [2018]	<b>Target Population:</b> Bengali corpus for grammar. <b>Scope of Studies:</b> Preprocessing	This paper proposed an algorithm that finds the unique verb root based on the user input. They used Modified Verb Root for creating the	Accuracy is used as an evaluation criteria.	The testing accuracy of this paper is 98.66%.	There are some exceptional cases [in verb] that the proposed algorithm can't handle.	<b>Not Mentioned.</b>	<b>Not Mentioned.</b>

		Bengali verbs and finding a unique and meaningful word.	appropriate verbal form.					
Automatic Extraction of Bengali Root Verbs Using Paninian Grammar. [39]	Arijit Das et al et al. [2017]	<b>Target Population:</b> Bengali corpus for grammar. <b>Scope of Studies:</b> Extract the root forms of the Bengali verbs.	This proposed algorithm is based on supervised learning. First, the surface-level forms or inflected forms of the verbs have been classified into a certain number of groups of similar tense and person. Secondly, a set of rules have been applied to extract the root form from the surface-level forms of a verb.	Accuracy is used as an evaluation criteria.	In Bangla Sadhu bhasa accuracy is 88%. Furthermore, in the Colloquial Bengali Language of Bangladesh the accuracy is 85%.	The morphological varieties of the words depend on geographical locations.	<b>Not Mentioned.</b>	<b>Not Mentioned.</b>
Representation of Bengali into UNL: An Analysis of Appropriate Bengali Verbs [40]	Mozammel Haque et al. [2018]	<b>Target Population:</b> Bengali corpus for grammar. <b>Scope of Studies:</b> Representing the Bengali verb into UNL[Universal Networking Language].	The proposed algorithm converts the Bengali verb into Universal Networking Language [UNL]. The proposed algorithm also detects equivalent UWs based on the user input.	Accuracy is used as an evaluation criteria.	The accuracy of the proposed algorithm was 95.33%.	The proposed algorithm cannot be directly applicable in exceptional cases.	The proposed algorithm cannot be directly applicable in exceptional cases.	The algorithm's proficiency is tested using 150 [30 * 5] verbs and get 143 accurate results among them.
Automatic Part-of-Speech Tagging for Bengali: An Approach for Morphologically Rich Languages in a Poor Resource Scenario. [41]	Sandipan Dandapat et al. [2007]	<b>Target Population:</b> Bengali corpus for grammar. <b>Scope of Studies:</b> To build a Part-of-Speech [POS] tagger for the Bengali language	First tagger HMM-S makes use of the supervised HMM model parameters. Second tagger HMM-SS uses the semi-supervised model parameters. Third tagger ME based model to find the most probable tag sequence for a given sequence of words.	Accuracy is used as an evaluation criteria.	The accuracy of this paper is 76%..	The problem arises due to the over-fitting of the supervised models in the case of small training data.	Supervised learning model along with suffix information. Morphological restriction on the possible grammatical categories of a word. The problem is reducing to increase annotated data	The training data includes manually annotated 3625 sentences. All the models have been tested on a set of randomly drawn 400 sentences [5000 words] disjoint from the training corpus
Towards POS Tagging Methods for the Bengali Language: A Comparative Analysis [42]	Fatima Jahara et al [2021]	<b>Target Population:</b> POS Tagging for the Bengali Language. <b>Scope of Studies:</b> To investigate the performance of different types of POS tagging techniques under a supervised approach.	Authors have used N-gram with HMM and TrT. They have also used CRF tagger.	Accuracy is used as an evaluation criteria.	Highest accuracy of 91.83% [for 11 tagset] and 84.5% [for 30 tagset] among all the tagging techniques	The proposed method used a very small data set containing 2000 sentences, and it faced difficulties due to the limited rules.	Syntactic parser can be a good future direction according to the authors.	The training corpus consists of 98K tagged tokens and the test corpus contains 17K tokens.
Parsing Bengali Text - an Intelligent Approach [43]	Goutam Kumar Saha [2006]	<b>Target Population:</b> Bengali text. <b>Scope of Studies:</b> Identify a parser	Authors have proposed an intelligent Bengali parsing algorithm	<b>Not Mentioned.</b>	<b>Not Mentioned.</b>	Find difficulties in understanding whether a word is a proper noun or not.	Semantic issues are also important with syntactic issues.	<b>Not Mentioned.</b>
A Computational Linguistic Analysis of Bangla[44]	Zeeshan Rahman Khan et al. [1995]	<b>Target Population:</b> Represent an implementation of a parser for the Bengali language. <b>Scope of Studies:</b> Parser includes a mechanism to handle the interaction between different syntactic phenomena.	Authors proposed Earley parser algorithm and lexicon is generated using the feature-based representation	<b>Not Mentioned.</b>	<b>Not Mentioned.</b>	limitations in analysis complex sentences.	Performance of the parser implementation isn't very fast on complex sentences	<b>Not Mentioned.</b>
Punctuation Prediction in Bangla Text[45]	Rahman, Habibur et al. [2023]	<b>Target Population:</b> Bengali text corpus. <b>Scope of Studies:</b> Punctuation	Authors proposed a bidirectional recurrent neural network [BRNN] along with Attention model with a plausibly large Bangla dataset.	F-1 score is used as an evaluation criterion.	Model's performance using a balanced Bangla dataset achieve performance scores 62.2 for question.	<b>Not Mentioned.</b>	In the future, incorporating more punctuation symbols may result in a more accurate model.	<b>Not Mentioned.</b>

		Prediction in Bengali Text						
A dependency annotation scheme for Bangla treebank[46]	Sanjay Chatterji et al. [2014]	<b>Target Population:</b> Bengali language corpus <b>Scope of Studies:</b> To develop a syntactic-semantic relation set for the Bengali language.	Authors have created a treebank called "KGPBenTreebank".	The scheme was tested through the electronic annotation of 4167 Bangla sentences by three annotators. Accuracy is used as an evaluation criteria.	The accuracy was 92.88%.	The accuracy of the annotated corpus might not be perfect and needs further improvement.	<b>Not Mentioned.</b>	<b>Not Mentioned.</b>
A Finite Automaton to Represent Structure of Conjugate and Single Bangla Letters [47]	Sanjib Kumar Saha et al. [2010]	<b>Target Population:</b> To design and implement a finite-state automaton for Bengali script. <b>Scope of Studies:</b> Determine the validity of a given sequence of characters according to the rules of the Bengali language.	Authors proposed an automaton-based approach.	<b>Not Mentioned.</b>	<b>Not Mentioned.</b>	It may not be able to handle all possible combinations of letters and words	<b>Not Mentioned.</b>	<b>Not Mentioned.</b>
A Hybrid Dependency Parser for Bangla[48]	Arnab Dhar et al. [2021]	<b>Target Population:</b> to present a two-stage hybrid framework for dependency parsing of Bengali sentences. <b>Scope of Studies:</b> Data-driven parser trained on a Bengali Treebank	Authors suggested a data-driven parser trained on a Bangla treebank. A demand frame-based parser that rectifies mistakes made by the first parser in identifying relations	Test data set of 150 sentence	Test data set of 150 sentence	Achieving higher accuracy in identifying certain types of relations, such as Subject, Object, Locative, and Relation	Both a data-driven parser and a grammar-driven parser, which may increase the computational complexity and resource requirements	<b>Not Mentioned.</b>
DPCSpell: A Transformer-based Detector-Purificator-Corrector Framework for Spelling Error Correction of Bangla and Resource Scarce Indic Languages [49]	Bijoy et al [2022]	<b>Target Population:</b> Bengali and Resource Scarce Indic Languages. <b>Scope of Studies:</b> Spelling Error Correction of Bengali and Resource Scarce Indic Languages	A Transformer based Detector-Purificator-Corrector Framework for Spelling Error Correction of Bangla and Resource Scarce Indic Languages	Authors evaluated the performance of our method using Precision, Recall, F-scores, Exact Match, and Modified Accuracy.	The exact match is 94.78% and modified accuracy is 95.16%. Moreover, the precision is 0.9487, recall is 0.9478 and F-scores is 0.948.	Authors stated that their new model may need more testing.	Authors suggest that future investigators can introduce newer concepts.	<b>Not Mentioned.</b>
A Proposed Automated Extraction Procedure of Bangla Text for Corpus Creation in Unicode [50]	D. S. H. Pavel [2006]	<b>Target Population:</b> Bengali Language. <b>Scope of Studies:</b> The paper focuses on proposing an automated extraction procedure for Bangla text in order to create a corpus in Unicode.	Authors configured a web crawler to locate Bangla text on the WWW. Authors applied language detection to ascertain Bangla content. Authors converted non-Unicode Bangla text to Unicode. Authors utilized search engines like Google to enhance text retrieval efficiency. Authors extended their method to offline Bangla documents.	Authors employed N-gram statistics for language detection, compared with an existing Bangla dictionary.	Results were not provided	Authors faced challenges in converting web documents to Unicode. Identifying encodings of Bengali e-documents was difficult. Dealing with Bangla documents in image format was out of scope. Risk of discarding genuine Bangla text due to errors or non-Bangla elements.	Authors plan to use the corpus for creating a Bangla lexicon, grammar extraction, machine translation, and word frequency analysis.	Authors referred to collecting online and offline Bangla documents from various sources.
A Review of Bangla Natural Language Processing Tasks and the Utility of Transformer Models[51]	Firoj Alam et al. [2021]	<b>Target Population:</b> Current state of Bengali natural language processing. <b>Scope of Studies:</b> using current state-of-the-art algorithms to encourage future research in the field.	Authors have used deep learning models like CNN and LSTM	<b>Not Mentioned.</b>	<b>Not Mentioned.</b>	Data scarcity for Bengali languages was a prime difficulty faced by the authors.	Encourage the use of transformer models for BNLN tasks and enrich the research in this field	<b>Not Mentioned.</b>
An approach of handling verbal inflections of bengali text: Conversion of	Rahman et al. [2019]	<b>Target Population:</b> Bengali Text. <b>Scope of Studies:</b>	Combination of rule-based and statistical approaches, including pattern matching and machine learning techniques, to identify	Accuracy was used to evaluate the effectiveness of the algorithm in identifying verb roots with their	In identification of verb roots the algorithm achieved 99.5% accuracy. The fluency of the Cholito text was	The accuracy of the system may decrease when dealing with complex sentence structures or rare verb forms not present in	<b>Not Mentioned</b>	Corpus of Bengali text used to build the algorithm and test its performance. Manually annotated dataset was used to

sadhu to cholito form of language [52]		Handling verbal inflections of Bengali text and converting Sadhu to Cholito form of language	verb roots and inflections in Bengali language.	inflections. Other than that fluency was used to evaluate the converted "Cholito" text produced from the Sadhu text	evaluated manually on a scale of 0 to 5, with an average score of 96.805%.	the training data.		evaluate the accuracy of the algorithm in identifying verb roots with their inflections.
Clause identification and classification in bengali [53]	Ghosh et al. [2010]	<b>Target Population:</b> Bengali language. <b>Scope of Studies:</b> Recognizing and categorizing clause techniques in Bengali	Establishing clause boundaries through a syntax-based rule model. Determining clause types using a statistical model based on Conditional Random Field [CRF].	The effectiveness of the techniques is evaluated based on the accuracy of clause recognition and classification.	There is no explicit comparison of the performance of different techniques in the review.	Data scarcity was mentioned indirectly as the accuracy of the system depends on the quantity and quality of annotated training data. Annotation difficulties not explicitly mentioned in the review but implied as dependency parsing and clause-level annotation require effort. Other than that complexity of handling natural language sentences and clauses was also a difficult hurdle.	<b>Not Mentioned</b>	The dataset used for training consists of 980 sentences.
Designing a Bangla parser using non-deterministic push down automata [54]	Rahman et al. [2017]	<b>Target Population:</b> Bengali language. <b>Scope of Studies:</b> Developing a Bengali parser using Non Deterministic Push-Down Automata [NPDA] for detecting grammatical and syntactical errors in Bengali language sentences	Converting Bangla language Context Free Grammar [CFG] into Non Deterministic Push-Down Automata [NPDA] by removing left recursion, eliminating null and unit productions, and converting CFG into Chomsky Normal Form [CNF].	Running time for parsing, accuracy in detecting grammatical and syntactical errors all are stated as an evaluation criteria.	NPDA parser requires less time for parsing compared to the predictive parser and has a lower overall running time	<b>Not Mentioned</b>	<b>Not Mentioned</b>	POS tagger with 1,20,000 Bangla words.
Parsing Bangla Using LFG: An Introduction. [55]	Haque et al. [2005]	<b>Target Population:</b> Bengali language. <b>Scope of Studies:</b> Parsing simple sentences in Bengali using Lexical Functional Grammar [LFG]s	Lexical Functional Grammar [LFG] is used to parse Bangla sentences by adding another level to Context-Free Grammar [CFG] and using functional schemata expressions associated with grammar rules.	<b>Not Mentioned</b>	The paper describes the success of LFG in parsing simple sentences in Bangla and its ability to incorporate semantic information into the syntactic structure. However, detailed quantitative evaluation results are not provided in the review.	Data scarcity for Bangla language, linguistic complexity, or difficulties in developing an effective LFG model for free-word-order.	The review mentions the future work of extending the basic LFG structure to incorporate the free-word-order of the Bangla language, which can be considered as a proposed solution.	<b>Not Mentioned</b>
Bangla Spelling Error Detection and Correction Using N-Gram Model [56]	Promita Bagchi et al [2023]	<b>Target Population:</b> Bengali Text. <b>Scope of Studies:</b> Bengali Spelling Error Detection and Correction	Usage of a combination of edit distance and n-gram language model to detect an erroneous target word and provides improvement suggestions	Accuracy was used as evaluation metrics.	They have compiled 11,500 Bangla sentences to apprise our model and the performance rate of our proposed model is 96%	Data sparsity and not enough previous research with the complex nature of the language was barriers faced by authors.	<b>Not Mentioned</b>	1,00,000 articles from various online sources
Bangla real-word error detection and correction using bidirectional lstm and bigram hybrid model.[57]	Noshin Jahan et al.[2020]	<b>Target Population:</b> Bengali language. <b>Scope of Studies:</b> Bengali real-word error detection and correction	This paper proposes a hybrid method which involves a combination of two different approaches like N-gram language model such as bigram and bidirectional long short-term memory [LSTM], a special type of recurrent neural network [RNN].	Accuracy was used as evaluation metrics.	The system shows significantly promising performance conveying 82.86% accuracy.	<b>Not Mentioned</b>	<b>Not Mentioned</b>	<b>Not Mentioned</b>
An Empirical Framework for Semantic Analysis of Bangla Sentences [58]	Parijat Prashun Purohit et al. [2014]	<b>Target Population:</b> Bengali language. <b>Scope of Studies:</b> a structure for a semantic analyzer that can-do semantic parsing of Bengali sentences.	Tokenization of sentences with POS analysis and the tagging. CFG rules with semantic analyzer was also used.	Accuracy was used as evaluation metrics.	The authors divided their analysis of the results of the presented framework into three sections. There is a near 99% success rate when sentences are simple, with word lengths ranging from 3 to 6. The average	Limited development in the field of Bangla language grammar checking was a difficulty the authors faced.	<b>Not Mentioned</b>	<b>Not Mentioned</b>

					accuracy for both complex and compound sentences with a range of 5-9 words is 90.2% and 90.7%, respectively.			
Grammar Checking System Using Rule Based Morphological Process for an Indian Language [59]	Lata Bopche et al. [2012]	<b>Target Population:</b> Hindi Language.  <b>Scope of Studies:</b> Grammar Checking System Using Rule Based Morphological Process.	Rule-based morphological method POS analysis and the tagging pattern analysis	Not Mentioned	Not Mentioned	Data sparsity and not enough previous research with the complex nature of the language was barriers faced by authors.	Not Mentioned	Not Mentioned
N-gram based Statistical Grammar Checker for Bangla and English [60]	Md. Jahangir Alam et al. [2007]	<b>Target Population:</b> Bengali language.  <b>Scope of Studies:</b> N-gram-based statistical approach to automating Bengali grammatical checking.	N-gram-based statistical approach with POS analysis and the tagging.	Accuracy was used as evaluation metrics.	In terms of accuracy, the approach presented by the authors gives an accuracy of 63% for English and roughly 58% for Bangla. Additionally, they have achieved 38% accuracy in automatically tagging POS for Bengali.	The tag set they used for Bangla do not have enough agreement features. Moreover limited development in the field of Bangla language grammar checking was a difficulty authors faced.	Hybrid system combining both statistical and rule based approaches was suggested by the authors.	Not Mentioned
A Corpus Based Unsupervised Bangla Word Stemming Using N-Gram Language Model [61]	Tapashee Tabassum Urmi et al. [2016]	<b>Target Population:</b> Bengali corpus for grammar.  <b>Scope of Studies:</b> Developing corpus based unsupervised Bengali word stemmer using n-gram language model.	Creation of a large corpus of Bangla stems and their accompanying inflectional forms. Development of an effective unsupervised method for creating a 6-gram model-based stemmer for the Bangla language. Root identification is accomplished using a straightforward threshold methodology.	Training of a larger corpus of data for unsupervised learning model. Then, the system was tested based on the accuracy level of the stem identification procedure.	For the stem identification process, the 6-gram model's accuracy was 40.18%.	The method is dynamic but requires a significant amount of data to get outcomes that are more appropriate.	A large volume of Bangla text should be employed and good threshold values should be produced for this method to be able to create a better root word dictionary for the language.	1,593,398 sentences from a variety of themes, including news, sports, blogs, websites, business periodicals, journals, etc.
Bangla Grammatical Error Detection Using T5 Transformer Model [62]	H.A.Z. Sameen Shahgir et al. [2023]	<b>Target Population:</b> Bengali Grammar.  <b>Scope of Studies:</b> Bengali Grammatical Error Detection.	This paper presents a method for detecting grammatical errors in Bangla using a Text-to-Text Transfer Transformer [T5] Language Model, using the small variant of BanglaT5, fine-tuned on a corpus of 9385 sentences where errors were bracketed by the dedicated demarcation symbol. The T5 model was primarily designed for translation and is not specifically designed for this task, so extensive post-processing was necessary to adapt it to the task of error detection.	Corrected output versus total tested data was used as evaluation metrics.	The final average Levenshtein Distance after post-processing the output of the fine-tuned model was 1.0394 on a test set of 5000 sentences	The challenges of adapting a translation model for grammar.	The approach can be extended to other languages, demonstrating the potential of T5 models for detecting grammatical errors in a wide range of languages.	A corpus of 9385 sentence
Effectiveness Analysis of Different POS Tagging Techniques for Bangla Language [63]	Jueal Mia, Md et al. [2022]	<b>Target Population:</b> Bengali Grammar.  <b>Scope of Studies:</b> Bengali parts-of-speech [POS] tagging.	The authors have compared some well-known supervised POS tagging approaches [Brill, HMM, unigram, bigram, trigram, and recurrent neural network] for Bangla languages.	Authors have used confusion matrix as an evaluation metric.	Accuracy for Brill, HMM, Unigram, Bigram, Trigram and RNN are respectively 95.07%, 94.27%, 95.82%, 94.62%, 93.07%, and 95.60%.	Not Mentioned	Not Mentioned	Not Mentioned