[1]Hem Chandra Das

[2]Kshirod Sarmah

[3]Raju Narzary

[4]Utpal Bhattacharjee

Development of an Automated Assamese Dialect Translation System Based on Rule-Based Techniques: Conceptualization and Preliminary Implementation

JES

Journal of Electrical Systems

Abstract: - To translate a particular word or collection of words from one dialect to another, a system called the dialect translation module is used in this work. The suggested system's results can help with understanding and analysing the dialectal differences between the dialects. This will also benefit individuals keen on acquiring communication skills in one or more of the dialects. When translating, consider both mainstream Assamese (standard Assamese) and the Kamrupia (Palasbari) dialect of Assamese. This study delves into the analysis of text, word construction, morphology, grammar, and ambiguity. Towards the conclusion, a method for developing a deliberate translation system is suggested.

Keywords: Assamese, dialect, Rule-based dialect translation, morphology, translation system

I. INTRODUCTION

An application known as a Rule-based Dialect Translation (RBDT) system is designed to translate words, phrases, and/or sentences from one dialect of a language to another, employing a set of predefined rules. These rules might pertain to grammar, morphology, or other aspects inherent to natural languages. Inter-dialect translations prove beneficial for individuals familiar with either dialect or those desiring proficiency in both.

Assamese, belonging to the Indo-European language family, is spoken by most of the Assamese population in the North-Eastern region of India. However, there are differences in the language spoken in different parts of Assam, particularly in terms of pronunciation, intonation, and vocabulary. This has resulted in the emergence of several regional dialects. The Central Assamese dialect is spoken in the Nagaon district of Assam and its neighbouring regions. The Eastern Assamese dialect is spoken in the vicinity of the Sibsagar district. The Kamrupia dialect is spoken in the districts of Kamrup, Nalbari, Barpeta, Darrang, Kokrajhar, and Bongaigaon. People in the relevant regions also speak Assamese varieties such as Nalbaria, Palashbaria, Barpetia, Kamrupia, Goalparia, Jorhatia, Sibsagaria, and so forth.

The dialect Eastern group spoken in and around former undivided Sivasagar district areas of currently Golaghat, Jorhat and Sivasagar, and another dialect Kamrupi spoken within the Kamrup region i.e., Barpetia, Nalbaria, Palasbaria. Both these two dialects are derived from Standard Assamese. But the Eastern group dialects are very similar with Standard Assamese language. Both dialects share words that originated from the Kamrupia script, which, in turn, has its roots in the Gupta script. This script is commonly recognized as the Assamese script, representing a modern Eastern script that evolved from the Brahmic script. However, the progressive shift to dialect has resulted in a major reduction in the use of its script and grammar, as well as a large vocabulary overlap with Assamese. Consequently, the script is not widely known, and the scarcity of a substantial number of manuscripts or

[1]Department of Computer Science and Technology, Bodoland University, Kokrajhar, 783370, Assam, India, hemchandradas78@gmail.com

[2]Department of Computer Science, Pandit Deendayal Upadhyaya Adarsha Mahavidyalaya (A Govt. Model College), Goalpara, 783124, Assam, India,

kshirodsarmah@gmail.com

[3]Department of Computer Science and Technology, Bodoland University, Kokrajhar, 783370, Assam, India, rajnarz@gmail.com

[4]Department of Computer Science and Engineering, Rajiv Gandhi University, Papum Pare, 791112, Arunachal Pradesh, India. utpal.bhattacharjee@rgu.ac.in

Corresponding Author: Kshirod Sarmah

*Department of Computer Science, Pandit Deendayal Upadhyaya Adarsha Mahavidyalaya (A Govt. Model College), Goalpara, 783124, Assam, India, kshirodsarmah@gmail.com

texts further compounds the issue. As a result, rather than using the script of two dialects, the suggested system employs the Assamese script. This paper outlines the text-based processing of both dialects.

The Assamese phonetic inventory comprises eight oral vowel phonemes, three nasalized vowel phonemes, and twenty-two consonant phonemes. Eight Assamese vowel's organizational structure is shown in Table 1.

Table 1: Assamese Language Vowels.

	FRONT	CENTRAL	BACK
High	i		u
			ʊ
High-Mid	e		o
Low-Mid	ɛ		ɔ
Low		a	

II. LITERATURE SURVEY

In the realm of dialect translation and rule-based machine translation (RBMT), there has not been a substantial and organized investigation into the Assamese language dialect. A system is proposed, titled "Malay Dialect Translation and Synthesis System," consisting of three modules: the Malay Dialect Translation module, the Malay Grapheme to Phoneme module, and the Malay Dialect Synthesis module [1]. "An Improved Grapheme to Phoneme rules for Assamese Language" [1] has also been developed to set grapheme to Phoneme fundamental rules by using different 7000 phonetically rich words of Assamese language. The rules have undergone testing in a speech synthesis process based on Hidden Markov Models, specifically in the context of HMM-based speech synthesis known as HTS [2]. The approach employed is a direct machine-to-machine method derived from word-to-word translation, as discussed in "A Machine Translation System for Standard Punjabi to Malay Dialect" [3]. This approach encompasses pre-processing of the source text (standard Punjabi), Grammatical Translation Rules, Lexicon Lookup, and the translation of standard Punjabi text to text in the Malwai dialect. The paper titled "A Bengali-Sylheti Rule-Based Dialect Translation System: Proposal and Preliminary System" [4][5] suggests a solitary module—the dialect translation module—that performs the translation of a specified word or group of words between Bengali-Sylheti and Chalita Bangla, both ways. The paper titled "Entity Recognition in Assamese Text" [6][7] presents the implementation of entity recognition in Assamese text through Conditional Random Fields (CRF), comprising two modules: Parts of Speech (POS) and Entity Recognition. Regarding the Assamese language, limited efforts have been made in machine translation between Assamese and English, employing various approaches. [8] presents an approach for "Assamese-English Bilingual Machine Translation" by using statistical phrase-based translation toolkit Moses such as BLEU, IRSTLM, GIZA. [8][9] focus on "Assamese Word Sense Disambiguation through Supervised Learning." Their work involves determining the correct sense of a word within a context where ambiguity arises due to multiple possible meanings for a given word [10][11].

Some partially related work, like "Vaasaanubaada Automatic Machine Translation of Bilingual Bengali-Assamese News Texts" [13][14], deals with bilingual text at the sentence level. The paper titled "A Common Parts-of-Speech Target Framework for Indian Language" [14] addresses tagging through the adoption of a hierarchical and decomposable tag set schema.

The suggested Rule-Based Dialect Translation (RBDT) System

Suggested concept for translation

In the envisioned system considering text as input from both dialects, prepare it as per rules and ambiguity, and at last construct he output from one dialect to another. Different steps have been followed by using the following manner Fig.1
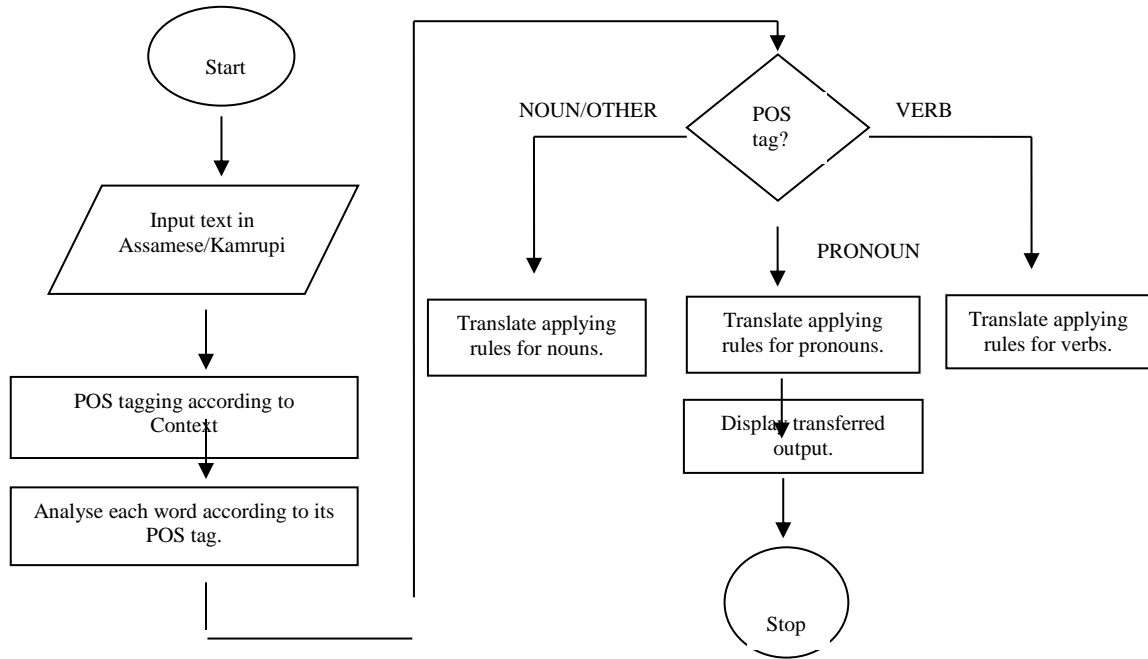
Figure 1: The structure of the suggested Rule-based dialect Translation System (RBDT)

III.   METHODOLOGY

Text-Based Analysis

Initial Experiment

Standard Assamese alphabet comprises 12 vowel graphemes and 52 consonants graphemes. Additionally, there are numerous conjunct consonants. The primary distinction between the two lies in the completely different pronunciation of words. The drawback of text-based processing is the difficulty in incorporating translation based on the speaker's tone, as the tone may change the tense and other characteristics of a sentence during speech.

In contrast to English, Assamese, like all Indian languages, is a free-word-order language, meaning that altering the positions of words in a sentence does not alter its meaning. As an illustration, cattle eat grass can be written correctly as গৰুৱে ঘাঁহ খায়/gôrʊ -e gʰãh kʰa -e/ (Cows eat grass) or ঘাঁহ গৰুৱে খায় / gʰãh gôrʊ-e kʰa-e /(Cows eat grass)or ঘাঁহ খায় গৰুৱে/ gʰãh kʰa-e gôrʊ-e /(Cows eat grass).
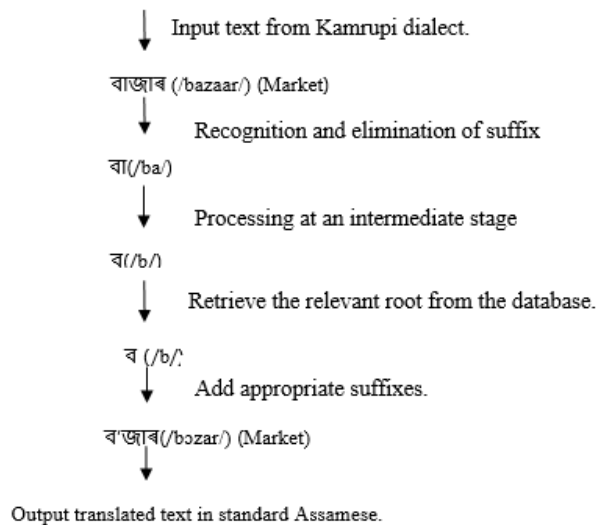

Figure . 2: Process Flowchart

Grammar and the classification of words into parts of speech (POS)

The standard Assamese dialect is grammatically like the Kamrupi dialect. In Assamese, the penultimate stress system is followed, while Kamrupia aligns with Bengalis in sharing the initial stress, establishing the initial stress system as the predominant feature. The initial emphasis observed in the Kamrupia dialect could indicate a linguistic shift resulting from contact with another language group. For instance, while the sound /x/ is not consistently present in Kamrupia, it does appear occasionally in standard Assamese. As an example, আখা/a -kʰa/ (Hope) (Kamrupia) and আশা/a –xa/(Hope)(standard Assamese). Manually, in Parts-of-Speech, a list of terms was created based on knowledge, memory, and input from domestic people. Assamese has a large morphological vocabulary. The key dialectal difference between Standard Assamese and Kamrupia lies in the morphology of words, encompassing tense forms and ambiguities based on parts of speech. To create a set of morphological rules that would be applicable to all dialects and hence facilitate translation was therefore the main objective.

The following is an example of POS:

সিহতে কলমে ৰে   যুদ্ধ   কৰিব

/xihote/ /kolom-ere /   /yuddha /   /koribo/

(They will fight with pen and ink)

Pronoun   Noun Post-position   Noun   Verb

Morphology

Assamese is a language with rich morphology. The primary distinctions between Standard Assamese and Kamrupia dialects are found in the morphology of words, which encompasses variations in tense forms and ambiguities related to parts of speech. Hence, the primary objectives were to establish a set of morphological rules applicable to both dialects, with the aim of aiding in the translation process.

Assamese has a few unique morphological features [3]. The following are a few of those that have been made public:

i.   In the Assamese language, numbers are not properly marked. Singular and Plural numbers are the two types of numbers in Assamese.
ii. In Assamese, grammatical expression of gender is also absent. Gender in this language is classified into two categories: Masculine and Feminine, with additional utilization of genders like common and neuter.
iii. For personal pronominal possession, kinship nouns are inflected.

iv. Assamese features two kinds of affixes: prefixes and suffixes, both of which are prevalent in the language. Within the Assamese language, each prefix and suffix is widely used. Suffixes encompass derivational, inflectional, and conjugational forms.

iv. Assamese grammar includes six case types: Nominative, Accusative, Instrumental, Dative, Ablative, and Locative. Additionally, Assamese recognizes six parts of speech (POS) [10]. These are:

    • Noun (Common, Proper, Collective, Material, Abstract.)

• Pronoun (Personal, Demonstrative, Inclusive, Relative, Indefinite, Interrogative, Reflexive)

    • Verb (Transitive, Intransitive)

    • Adverb (Manner, Place, Time)

    • Adjective (Nominal, Qualifying)

    • Indeclinable (Conjunction, Interjection etc.)

    For instance, the statement "He has gone home" can be expressed in the Standard Assamese dialect as সি ঘৰলৈ/ঘৰত গৈছে and in the Kamrupia dialect as সি ঘৰত/ঘৰক গেছি. This illustrates clear distinctions in terms of morphology.

Creation of rules for translation

It has been observed that each dialect of Assamese language has a characteristic pattern of speaking. From the initial study, it has been observed that the changes in the pronunciation of the same word in different dialect are coming due to anyone or more of the following reasons:

a)   Shortening or lengthening of the final vowel.

b)   Adding one special vowel after the final vowel.

c)   Silencing some consonants if they appear before or after some vowels.

From the above study, it has been observed that some rule-based transition mechanism may be developed for converting words pronounce in one dialect to some other dialect. A thorough investigation will be carried out in this work to establish such a transition mechanism.

 Case markers are linked to the basic rules for nouns and pronouns. Single letters or groups of letters attach at the end of a noun/pronoun to identify their functional relationship in the sentence are known as case markers. In Assamese, cases are termed as "কাৰক," and their markers are referred to as "বিভক্তি," always affixed at the end of a noun or pronoun. Another characteristic that sets it apart is the utilization of markers, which is unique to Assamese and a few other dialects and languages and signals a shorter sound of O. The frequent usage of velar nasal /ŋ/ is another remarkable phonetic trait of the Assamese language. This /ŋ/ is consistently linked to a homorganic sound such as /g/ in other New Indo-Aryan languages. In Assamese, on the other hand, it is always used alone [4]. The absence of the voiceless velar fricative /x/ in any other language in the country makes it a distinctive feature of the Assamese language. It sounds comparable to the European velar sound. It could be an Indo-European feature that the Assamese have kept [4]. In addition, both dialects tense forms of verbs account for a large portion of morphological distinctions. All the rules employed in the suggested system are detailed in the tables (Tables 1 and 2) provided below.

Designing the database.

In this implementation, the dictionary (database) named Lexicon can store diverse root words, specific suffixes from one dialect, and their corresponding counterparts from the other dialect. Additionally, it includes parts-of-speech tags and brief descriptions of the root words. The schema for Lexicon is outlined in Table 3.

TABLE 1: LIST OF RULES FOR NOUNS AND PRONOUNS WITH EXAMPLES

Kamrupi Marker		*Assamese Marker*	*Usage example*
Direct translation from DB, both noun & pronoun			আম→আম(Mango)
ঈ/ী(/i/)		উ/ঈ/ী(/u//i/)	মেকুৰী→মেকুৰী(Cat)
আ/া(/a/)		ও/আ/া(/ʊ//a/)	কুম্ৰা→কোমোৰা
এ/য়/ে(/ɛ//j/)		*Null/*ে(/e/)	কুত্তায়→কুকুৰে(Dogs)
আ/া(/a/)		উ/া(/u/)	দুআৰ→দুৱাৰ(The door)
ে া *vowels+*(/ʊ/)	তো	জন	আপাতো→লৰাজন(Boy)
Other vowels+		ী/জনী	আপীতো→ছোৱালীজনী(Girl)
Noun/Pronoun word ending in a pronoun-suffix must have both root and suffix in DB			নুন→নিমখ(Salt)

TABLE 2: COMPILATION OF RULES FOR VERB TENSE FORMS ALONG WITH EXAMPLES.

KamrupiSuffix	*IPA*	*StandardAssameseSuffix*	*IPA*	*Example*
উ/ু	/u/	ও/ে া	/ʊ/	কৰু/কৰো মই কামতো কৰু/কৰো (I will do this work)
অ	/ɔ/	আ/া	/a/	কৰ/কৰা তুমি এইটো

				কৰা/কৰ (You do it)
আ/া	/a/	অ	/ɔ/	কৰা/কৰক এইতো কৰা/কৰক (Do this)
উ/ু	/u/	ও/ো	/ʊ/	কল্লু/কৰিলো মই কামতো কল্লু/কৰিলো (I did the work)
ঈ/ী+ল্লা	/i/+(lô + lô)+ /a/	(ই)+লি	/i/+li	কল্লি/কৰিলি তুমি তোমাৰ কামতো কৰিলি/কল্লি (Did you do your work)
অ+ল্লা	/ɔ/+(lô + lô)+/a/	(ই)+লা	/i/+/l/+/a/	কল্লা/কৰিলা তুমি তোমাৰ কামতো কল্লা/কৰিলা (Did you do your work)
(ই)+ম	/i/+/m/	(ই)+ম	/i/+/m/	কৰিম/কৰিম মই কৰিম/কৰিম (I will)
অ+বি	/ɔ/+/b/+/i/	(ই)+বি	/i/+/b/+/i/	কৰবি/কৰিবি তুমি কৰবি/কৰিবি (You'll do it)
বো	/b/+/ʊ/	(ই)+ব	/i/+/b/	কৰবো/কৰিব সিহতে কৰবো/কৰিব (They will do)
বা	/b/+/a/	(ই)+বা	/i/+/b/+/a/	কৰবা/কৰিবা তুমি কৰবা/কৰিবা (You will do)
(উ/ু)+চ+ছ	/u/+(sô + shô)	(ই)+ছো	/i/+/s/+/ʊ/	কচ্ছু/কৰিছো মই কচ্ছু/কৰিছো (I've done)
ই/ি	/i/	ই/ি	/i/	কৰি/কৰি মই কৰি আছো (I'm doing it)
(অ)+চ্ছি	/ɔ/+(sô + shô)+/i/	(ই)+ছে	/i/+/s/+/e/	কচ্ছি/কৰিছে সি কচ্ছি/কৰিছে (He has done)
(অ)+চ্ছা	/ɔ/+ (sô + shô)+/a/	(ই)+ছা	/i/+/s/+/a/	কচ্ছা/কৰিছা তুমি কচ্ছা/কৰিছা (You've done it)
(অ)+ছিলু	/ɔ/+/s/+/i/+/l/+/u/	(ই)+ছিলো	/i/+/s/+/i/+/l/+/ʊ/	কচ্ছিলু/কৰিছিলো মই

				কচ্ছিলু/কৰিছিলো (I did)
(অ)+ছিল	/ɔ/+/s/+/i/+/l/	(ই)+ছিল	/i/+/s/+/i/+/l/	কচ্ছিল/কৰিছিল সি কচ্ছিল/কৰিছিল (He did)
(অ)+ছিলি	/ɔ/+/s/+/i/+/l/+/i/	(ই)+ছিলি	/i/+/s/+/i/+/l/+/i/	কচ্ছিলি/কৰিছিলি তুমি কচ্ছিলি/কৰিছিলি নেকি? (Did you?)
(অ)+ছিলা	/ɔ/+/s/+/i/+/l/+/a/	(ই)+ছিলা	/i/+/s/+/i/+/l/+/a/	কচ্ছিলা/কৰিছিলা তুমি কচ্ছিলা/কৰিছিলা (You did)
(অ)+বা	/ɔ/+/b/+/a/	আ/◌া+বা	/a/+/b/+/a/	কৰবা/কৰাবা তুমি কৰবা/কৰাবা (You'll do it)
উ/◌ু+ওউ	/u/+/ʊ/+/u/	ও/◌ো+ৰাওক	/ʊ/+/w~β/+/a/+/ʊ/+/k/	কৰুওউ/কৰোৰাওক তাৰ দ্বাৰা কৰুওউ/কৰোৰাওক (Do it by him)
উ/◌ু+ৰা	/u/+/w~β/+/a/	ও/◌ো+ৰা	/ʊ/+/w~β/+/a/	কৰুৰা/কৰোৰা তাৰ দ্বাৰা কৰুৰা/কৰোৰা (to be made by him)
চুন	/s/+/u/+/n/	চোন	/s/+/ʊ/+/n/	কৰিমচুন/কৰিমচোন মই কৰিমচুন/কৰিমচোন (I'm going to do it)
অ+বিচুন	/o/+/b/+/i/+/s/+/u/+/n/	ই/◌ি+বাচুন	/i/+/b/+/a/+/s/+/u/+/n/	কৰবিচুন/কৰিবাচুন তুমি কৰবিচুন/কৰিবাচুন (You're going to do it)
অ+বিচুন	/o/+/b/+/i/+/s/+/u/+/n/	ই/◌ি+বিচোন	/i/+/b/+/i/+/s/+/u/+/n/	কৰবিচুন/কৰিবিচোন তুমি কৰবিচুন/কৰিবিচোন (You're going to do it)
অ+বিচুন	/o/+/b/+/i/+/s/+/u/+/n/	ই/◌ি+বচোন	/i/+/b/+/s/+/u/+/n/	কৰবিচুন/কৰিবচোন তুমি

|				কৰিবিচুন/কৰিবচোন (You're going to do it)

TABLE 3: DATABASE TABLE LEXICON STRUCTURE

Name	Description
Assamese_base(root)	The base-word and/or suffix in standard Assamese dialect
Kamrupia_base(root)	The Kamrupia dialect's base word and/or suffix
English_description	English definitions and brief summaries of words
POS	Tag for parts of speech

System Setup

Matching of Suffixes

The base (root) word and its prefix and/or suffix are the greatest tools for managing the morphological entity. Each of the two dialects' appended suffixes can be seen in the example found in the specified rules. They all adhere to the following rule: distinct root word + similar suffix= many variations of the same tense, case markers, etc.
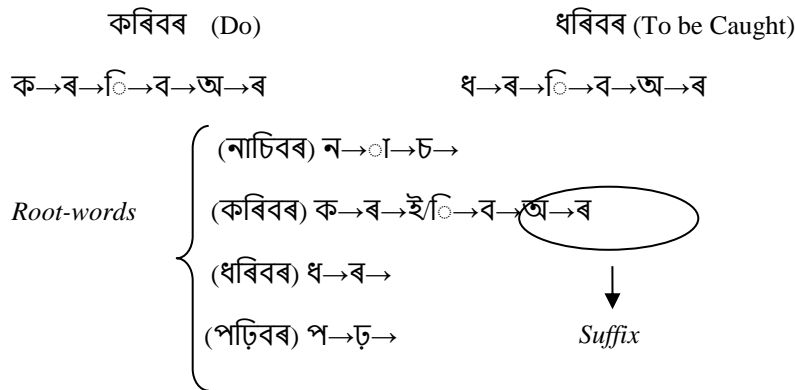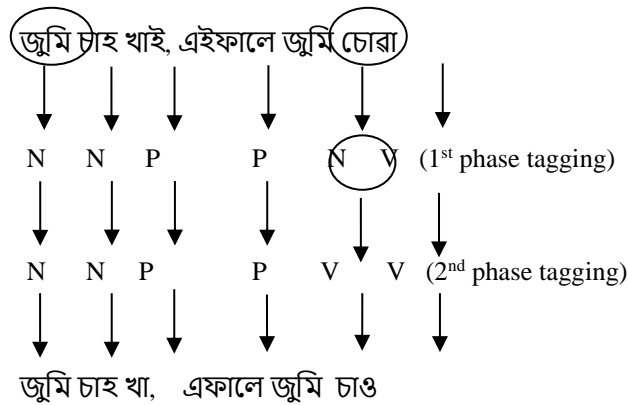

In consequence, the suffix-matching approach is established as valuable and well-organized technique. To determine which rule is most appropriate, this system obtains the suffix and compares it to the list of rules. The target suffix is substituted for the extracted root word of the source dialect after its matching POS tag is also taken from the lexicon and exchanged with the root word of the opposite accent in the database. The POS tagging can be used to get the principles of verbs (tense rules) and/or nouns/pronouns (case marker rules).

Resolving Ambiguity

When a word has two or more meanings, it is said to be ambiguous in the context of word sense. There are bound of words that are spelled an analogous however meaning is completely different in several circumstances. Though human have the power to grasp the ambiguity of words reckoning on the tone of the speaker present in numerous sentences. Thus, applying the guidelines directly to text-based processing led to incorrect outcomes. As an example, the ambiguity of Assamese dialect word জুমি shows within the following-

তুমি এইফালে জুমি চোৱা →*Used as verb here, meaning to peep*

জুমি চাহ খোৱা→ *Name of girl used as noun in this instance.*

The computer program uses a separate list or pre-identifies words that generate ambiguity manually before using them. The suggested approach handles ambiguity as part of a two-step POS tagging process. The first stage involves searching the database for the terms and categorizing them using their fundamental tagging scheme. The term is then tagged once more if it is still unclear after utilizing the other tags in the phrase and/or according to grammatical norms.

জুমি চাহ খাই, এইফালে জুমি চোৱা

N   N   P   P   N   V   (1ˢᵗ phase tagging)

N   N   P   P   V   V   (2ⁿᵈ phase tagging)

জুমি চাহ খা,   এফালে জুমি চাও

Framework for Assessment

The database satisfies our needs despite its limited scope because the data has been carefully selected to verify all potential regulations and several ambiguities. To test the words, single words from different speakers of both dialects were inserted together as a group. As a result, the following outcomes were attained.

In this case, the word(s) in the input text that are stored in the database are referred to as the only database. Refers to any text in general, regardless of whether it is included in the database.

Table 4: Findings from the assessment

Database	Accuracy (%)
Words (only the database)	96
Words (generally speaking)	54
Phrases (including words from a database)	85
Phrases (generally speaking)	41.5

Input Standard Assamese text-আমি সৰু থাকোতে আম খাইছিলো(We used to eat mangoes when we were little)

Input Kamrupia text- আমি সৰু থাকাংতে আম খাছু (We used to eat mangoes when we were little)

Correct Kamrupia text-আমি সৰু থাকাংতে আম খাছিলু (We used to eat mangoes when we were little)

The tiny size of the database, or the lack of vocabulary, is one factor contributing to the low accuracy in the usual cases. Since there was little time for study, planning, and execution in this simulation project, larger databases typically yield better results. Because Kamrupia is a dialect, users may insert text according to their speech pattern rather than the specific syllable structure, which could lead to an incorrect outcome. Additionally, because the emphasis in Kamrupia is mostly on the speaker's tone, users may do so. The above example will be written as or based on the sentence's tense, which in turn depends on the speaker's tone.

IV.   CONCLUSION

This work aimed to provide a method of text-to-text translation between the Kamrupia dialect and Standard Assamese by presenting the translation rules. Clear translations that fall within the parameters of the database suggest an even larger database that contains every term that can be used in the language. Since Assamese has unique rules for speech, tone, and prosody that set it apart from other Assamese dialects like Kamrupia, Barpetia, Goalporia, etc., not all the rules are examined in detail. For instance, in Kamrupia, হাতি (Elephant) is typically spoken as হাথি, but it is not written that way. Therefore, the next big accomplishment should be rules for understanding and producing speech, as well as a study of basic speech aspects and dialect prosody, in addition to expanding the database to include additional words.

REFERENCES

[1]   T.-P. Tan, S.-S. Goh and Y.-M. Khaw, "A Malay Dialect Translation and Synthesis System: Proposal and Preliminary System," in 2012 International Conference on Asian Language Processing, Hanoi, Vietnam, 2012.

[2]   B. Sharma, P. B. Acharjee and P. H. Talukdar, "AN IMPROVED GRAPHEME TO PHONEME RULES FOR ASSAMESE LANGUAGE," International Journal of Advanced Research (IJAR), vol. 3, no. 12, pp. 1460-1464, 31 Dec 2015.

[3]   H. Singh, R. Khanna and V. Goyal, "Research Cell: An International Journal of Engineering Sciences," 2013.

[4]   S. Chakraborty, A. Sinha and S. Nath, "A Bengali-Sylheti Rule-Based Dialect Translation System: Proposal and Preliminary System," in Proceedings of the International Conference on Computing and Communication Systems, NEHU, Shillong, India, 2016.

[5]   N. Mahanta, S. Dhar and S. Roy, "Entity Recognition in Assamese Text," in 2016 IEEE International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2016.

[6]   K. K. Baruah, P. Das, A. Hannan and S. K. Sarma, "Assamese-English Bilingual Machine Translation," International Journal on Natural Language Computing (IJNLC), vol. 3, no. 3, pp. 227-231, 2014.

[7]   P. P. Borah, G. Talukdar and A. Baruah, "Assamese Word Sense Disambiguation using Supervised Learning," in IEEE 2014 International Conference on Contemporary Computing and Informatics (IC3I), Mysore, India, 2014.

[8]   K. Vijayanand, S. I. Choudhury and P. Ratna, "VAASAANUBAADA: automatic machine translation of bilingual Bengali-Assamese news texts," in IEEE Computer Society Language Engineering Conference, 2002. Proceedings, Hyderabad, India, 2002.

[9]   B. Sankaran, M. Choudhury, T. Bhattacharya, and K. Bali, "A Common Parts-of-Speech Tagset Framework for Indian Languages," in Proceedings of the International Conference on Language Resources and Evaluation, LREC 2008, Marrakech, Morocco, 2008.

[10] G. Goswami and M. B. Emeneau, "Structure of Assamese," Gauhati: Dept. of Publication, Gauhati University, 1982.

[11] U. Bhattacharjee, "Recognition of Assamese Phonemes Using RNN Based Recognizer," in International Symposium on Computer Music Modeling and Retrieval on Speech, Sound and Music Processing: Embracing Research in India, 2011.

[12] M. K. Sinha, R. M., "Elderly Assistant Based on Facial Emotion and Posture Analysis," IJRDO - Journal of Electrical and Electronics Engineering, vol. 6, no. 6, pp. 13-20, 2020.

[13] A. Kale, "Optimisation of Pulse Forming Networks," IJRDO - Journal of Electrical and Electronics Engineering, vol. 7, no. 7, pp. 1-13, 2021.

[14] M. S. Javed, "The Impact of Language on Culture and To Understand Second Language for Communication, Human Life, And Society," EPH - International Journal of Humanities and Social Science, vol. 9, no. 1, pp. 1-7, 2024. DOI: https://doi.org/10.53555/eijhss.v9i1.175.