CHAPTER 2

WORD CLASSES AND INFLECTIONS

2.0. INTRODUCTION

The primary objective of this chapter is to provide the linguistic background required for the task of building a POS tagger for Bangla. The present chapter focuses on the linguistic analysis of the various word classes in Bangla. It also examines and discusses the inflectional endings of verbs, adverbs, nouns, classifiers, adjectives, and pronouns to provide a linguistic background of the suffixes that will be employed to formulate the disambiguation rules in the rule-based tagging system.

The chapter begins with an overall sketch of the word classes of Bangla parts-of-speech. The second section gives a comprehensive analysis and description of the inflectional endings of the Bangla verbs, adverbs, nouns, pronouns, adjectives and quantifiers. Finally, the discussions are summarized in the last section.

2.1. OVERVIEW OF THE WORD CLASSES IN BANGLA

The word classes or the lexical tags of Bangla have been divided into two broad categories: open class and closed class. The open class category forms the major section of any languages’ lexicon because the set is quite large and an unlimited number of new words can be created and added to the vocabulary of this class (Jurafsky & Martin, 2000).
The four major parts of speech in open class category are nouns, verbs, adjectives and adverbs.

The words of closed class type are a relatively small set and to this set no new words can be added\(^1\) (Jurafsky & Martin, 2000). The closed class includes words belonging to post-positions, determiners, conjunctions, auxiliary verbs, particles and pronouns.

2.1.1. NOUNS

Bangla nouns can be grouped into the following morphological and grammatical categories:

<table>
<thead>
<tr>
<th>CATEGORY</th>
<th>TYPES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>Singular/Plural</td>
</tr>
<tr>
<td>Countability</td>
<td>Count/Uncount</td>
</tr>
<tr>
<td>Animacy</td>
<td>Animate/Inanimate</td>
</tr>
<tr>
<td>Definiteness</td>
<td>Definite/Indefinite</td>
</tr>
<tr>
<td>Case Inflection</td>
<td>Nominative/ Accusative/ Locative/ Genitive/ Dative</td>
</tr>
</tbody>
</table>

Table 1: Types of Bangla Nouns (Ghosh, 1995), (Thompson, 2010).

Further, Thompson (2010) says that the genitive nouns can also be used as attributive adjectives to modify other nouns, as in, [rupor \(\ddot{t}h\text{ala}\)] (silver plate), and [\(\ddot{f}ur\text{er rofimt}\)] (sun rays) and the locative nouns help in forming post-positions and adverbs. Now, in order to distinguish these grammatical categories of number, definiteness and animacy, the nouns get inflected to classifiers, modifiers and case endings. For instance,
the classifiers like -ta, -gulo, -kʰani are used to determine whether the noun is a definite or indefinite, a singular or plural, animate or inanimate. Apart from the inflected nouns, there is another category of nouns, known as bare nouns. Thompson (2010) claims that the following types of nouns fall under the category of bare nouns which occur regularly without any classifiers or case markers.

a. Names and titles: [ʃær] (Sir), [mædəm] (Madam), [puja] (Puja).

b. Generic nouns: [manʊʃ] (human beings), [ʃɪbʌʊntu] (animals), [gacʰ] (tree).

c. Natural phenomena and objects: [ɔmaboʃʃɑ] (new moon), [nɔdɪ] (river), [bɾɪʃtɪ] (rain).

d. Abstract nouns: [ʃɔməu] (society), [bʰugol] (geography), [gʰrɪnə] (hatred).

2.1.2. VERBS

Bangla verbs are broadly classified into finite and non-finite verbs (Chattopadhyay, 2003), (Chatterji et al, 2012). The finite and non-finite verbs are further divided into main and auxiliary verbs. While the main verbs form the open category of word classes, the auxiliary verbs fall into the closed type of the word class. Chatterji et al (2012) states that the finite verbs indicate the features of person, tense, aspect and honorificity of the action of the corresponding clause, whereas the non-finite verbs are used to indicate the intermediate action of the clause. Bangla verbs do not show any number and gender agreement which means that the verbs are not inflected for number and gender features (Chatterji, 1926), (Chattopadhyay, 2003). Thus, the verbs are distinguished on the basis of person, tense, aspect and honorificity. Unlike English,
Bangla grammar does not have auxiliary verbs (Chatterji, 1926), (Chakrabarti, 2015). However, Chatterji et al (2012) points out that there are a few verbs in Bangla which function both as auxiliary and main verbs. In a verb chunk, these verbs occur in a sequence of a main and auxiliary verb respectively such as, [kʰetʰe hɔbe] (have to eat), [hoṭe pɔr̥to] (could have happened). (Chatterji et al, 2012).

2.1.3. ADJECTIVES

Adjectives in Bangla are indeclinable which behave nominally, when they get inflected to classifier like, -ta, -gulo etc, as in, [cʰoto-ta] (the small one), or [kalo-gulo] (the black ones) (Thompson, 2010). Bangla adjectives do not exhibit any regular morphological features. But, a few adjectives, which are derived from nouns and verbs, have suffixes which facilitate in identifying an adjective. For example, -nto is an adjectival ending which gets attached to verbs to form adjectives as in [pɔronto] (falling) from [pɔra] (fall) while, -t̥o gets inflected to nouns like [anond̥o] (joy) to form the adjective [anond̥t̥o] (delighted).

2.1.4. ADVERBS

Racova (1990) and Dash (2005) define adverbs as the auto-syntagmatic and auto-semantic inflectionless words which describe the circumstances of an action and the quality of verbs, adjectives, nouns and adverbs. According to Racova (1990), there are a few adverbs in Bangla that resemble adjectives on the ground of their basic grammatical properties like- the ability to express quality and inflectionless words. For instance, [bʰalo] indicates the meaning of both well (adverb) and good (adjective). So, in order to distinguish adverbs from the adjectives, the syntactic features of adverbs are taken into
consideration. However, Racova (1990) also made a point that adverbs refer to verbs more than adjectives because of their fundamental feature of characterizing an action. Bangla has a very few primary adverbs as most of the adverbs are derived from adjectives, nouns, and verbs (Racova, 1990). For instance, [kɔl] (tomorrow), [kʰub] (very), [ɔɹ] (today) are a few of the primary adverbs whose adverbial function is the sole and fundamental function. The secondary adverbs are formed from adjectives, as in, [ʃeʃe] (ultimately), from verbs, such as, [mɪle] (together), and from pronouns such as, [tæmon] (thus, so).

Based on the structural components, Dash (2015) categorizes three types of adverbs. The first type of adverbs is the simple and non-inflected adverbs such as, [mulɔtɔ] (mainly), [prəɾ] (often), [ækʰon] (now), [ɔɹ] (today), and [tɔkʰon] (then). The second type of adverbial forms are derived from nouns and adjectives by adding the suffix marker –e as in, [prot̚om-e] (firstly), [somən-e] (continuously), [dʰɪr-e] (gently) and more. The adverbial forms of the third type are obtained from the nouns by adding a specific set of adverb suffixes like -bʰabe as in, [bʰalo-bʰabe] (nicely), [soman-bʰabe] (equally), -mɔt̚o as in, [hisab-mɔt̚o] (accordingly), [ɪccʰe-mɔt̚o] (according to ones wish), -bəsət̚o as in [bʰul-bəsət̚o] (by mistake), [sɔbʰab-bəsət̚o] (by habit).

2.1.5. PRONOUNS

Thompson (2010) describes the pronouns as the sub-categories of nouns that substitute the nouns, especially the proper nouns. Chatterji (1926) and Chaki (1996) have classified the Bangla pronouns into eight categories according to their origin, meaning and the role they play in a sentence. The eight types of pronouns are as follows,
Table 2: List of Pronouns in Bangla

Dash (2000) describes two types of pronouns depending on their syntactic structure, the non-inflected pronouns and the inflected pronouns with case, number etc. The simple pronouns like, [e₁] (this), [ke] (who), [ʃe] (he/she) belong to the first category as they do not have any inflections and occur in bare forms in the corpus. The inflected pronouns like, [ɑmɑr] (my), [ɑmɑke] (to me), [t̰ɑd̯er] (their) include the inflectional endings such as case markers, number, honorific markers etc. The linguistic analysis of the inflected pronouns has been discussed in the following section (cf. 2.3).

<table>
<thead>
<tr>
<th>PRONOUN</th>
<th>EXAMPLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal</td>
<td><a href="I">ɑmɪ</a>, [tumɪ] (you)</td>
</tr>
<tr>
<td>Relative</td>
<td>[ʃɪnɪ] (who)</td>
</tr>
<tr>
<td>Interrogative</td>
<td>[ke] (who)</td>
</tr>
<tr>
<td>Reflexive</td>
<td>[nɪʃe] (ownself)</td>
</tr>
<tr>
<td>Demonstrative</td>
<td>[e₁] (this)</td>
</tr>
<tr>
<td>Indefinite</td>
<td>[keo] (someone)</td>
</tr>
<tr>
<td>Reciprocal</td>
<td>[pɔrɔʃpɔr] (each other)</td>
</tr>
<tr>
<td>Inclusive</td>
<td>[ʃɔkɔl] (everyone)</td>
</tr>
</tbody>
</table>

2.1.6. POST-POSITIONS

Post-positions belong to the word class of closed types. They form a small set and historically they are mostly derived from nouns and verbs, but now they form a separate word class of themselves (David, 2015). For, instance, the post-positions like, [tʰeke] (than, from), [nɪʃe] (regarding), [d̤eʃe] (with), [kæʃe] (near, by) and [pokkʰe] (on behalf
of) were frequently derived from nouns and verbs. Apart from these derived post-positions, there are underived post-positions present in the word class, such as, [upor] (on), [mod̄̄h̄e] (between), and [nagad] (around).

2.1.7. CONJUNCTIONS

Racova (1983) states that conjunctions in Bangla are classified according to the origin and the syntagmatic relations which they express. Based on the origin, conjunctions have been categorized into primary and secondary conjunctions. The primary conjunctions include [ba] (or), [ebon] (and), [ar] (and), [tobu] (but), [rotchoco] (yet), [rohi] (if), [pachre] (if) and more. The secondary conjunctions are derived from other parts-of-speech mostly, from verbs and nouns, but over time they have acquired the features of conjunctions and now in Modern Bangla they have become full-fledged conjunctions. [na] (or), [tai] (that’s why), [futoraŋ] (so), [tafe] (so that), [jeunno] (that’s why) belong to the class of secondary conjunctions.

In terms of function, the conjunctions are divided into coordinating, and subordinating conjunctions (Thompson, 2010). The coordinating conjunctions refer to the words that act as connectors that connect clauses and parts of complex sentences. Racova (1983) further states, that the coordinating conjunctions have been subdivided into adversative conjunctions such as [kintu] (still), [tobe] (then), [tobu] (still), copulative conjunctions as, [ar] (and), [ebon] (and), [o] (and), disjunctive conjunctions like, [ba] (or), [kimba] (or), [kna] (is it) , resultative conjunction as in, [futoraŋ] (so), [kure] (so), [tai] (that’s why), and finally, explicative conjunction, [noile] (otherwise).
The subordinating conjunctions refer to the words that connect the main and the subordinate clauses in a sentence. According to Racova (1983), the subordinating conjunctions in Bangla connect solely the clauses of content which are expressed through [k Sinatra], [[jæno]], and [je].

[je] (that) is the only finite complementizer found in the complement clauses of Bangla (Bayer, 1999), (Bhattacharya, 2000). David (2015) states that the complementizer [je] generally follows the verb to introduce the complement clause in a Bangla sentence, as in expressions like, [jani je] (know that). Bhattacharya (2000) claims that this complementizer can also occur in pre-verbal positions as in, [tumi je] (that you). Bayer (1999) points out that this complementizer also functions as a left-adjointed correlative in the Bangla sentences such as,

(1)  je  c'hleeta ok'hane  danie  chilo,  fe  amar b'ui

Who boy there stand was he my brother

‘The boy who was standing there, is my brother’.

2.1.8. PARTICLES

Bangla includes two types of particles in its corpus, which are used quite regularly with the conjugated forms of nouns, verbs, adverbs, pronouns and adjectives (Chatterji, 1926), (Dash, 2015). The Emphatic particles are used with all of the above mentioned parts-of-speech whereas the Negative particles are used only with verbs. There are three emphatic particles used in a Bangla corpus, -i, –o, and -fo which carry the meaning of ‘too’, ‘also’ and [+]emphatic. (Bayer & Lahiri, 1990), (Dash, 2015). Among these emphatic particles, -i and –o act as clitics which always get inflected to the words. Dash
(2015) states that -ɪ and -o are generally added immediately after the final conjugated form of words as in, [cʰeleke-ɪ] (to the boy only), [bole-o] (tells also), [bʰalo-bʰa-b-e-ɪ] (nicely [+emphatic]) and more. But, there are a few cases, especially in verbs, where the emphatic clitics get inserted in the middle of the inflected verb forms as in, [kore-o-cʰe] (have done indeed), and [peje-ɪ-cʰɪ] (have got indeed).

The emphatic particle -t̯o is used as a separate word which is tagged immediately after the conjugated word forms instead of being attached to the words. For instance, [mejetɑ t̯o] (the girl indeed), [bolecʰe t̯o] (have told indeed) proves that ‘t̯o’ occurs right after the conjugated noun and verb forms respectively. There are cases especially with post-positions and pronouns, where the emphatic clitics get directly added to the root forms of the word as in, [amr-o] (me too), [fe-ɪ] (he [+emphatic]), [pokkʰ-e-ɪ] (for [+emphatic]), [mɑte-o] (according to [+emphatic]).

The three types of negative particles -nɑ, -ni and -nai are used only with verbs and express the function of verbal negation (Racova, 1984), (Ramchand, 2004), (Dash, 2015). The negative marker -ni gets tagged to the inflected verbs and adds a sense of negation to the sense expressed by the inflected verb. For instance, in the verbs like [boli-ni] (did not tell), [jai-ni] (did not go), the verbs with the affixation of the negative marker express a negative meaning. The negative particle, -na can occur separately after the inflected verbs, such as [bolbo na] (won’t tell) and [kʰelʃ na] (do not play) besides getting attached to them such as, [parbo-na] (cannot do), [kʰa-nɑ] (do not eat) and so on. Furthermore, the negative particle, -nai also gets added separately after the conjugated verbs and adverbs to express negation as in, [kɔri nai] (did not do), and [ekʰane nai] (not here) respectively. Racova (1984) points out that the particles -na and -nai, apart from
expressing verbal negation also indicate adverbial and pronominal function of negation such as, [keo na] (no one), [ækɗomɪ na] (not at all), [ektuo na] (not a bit), [keo na] (no one), [okʰane nai] (not there) and more.

2.2. INFLECTIONS AND WORD CLASSES

Apart from the contextual occurrence, the category of a word can also be determined by looking at their inflectional endings or affixes. For instance, if a word like, ‘cat’ ends with the morpheme ‘-s’, then, it can be identified as a noun. Moreover, the morpheme -s also indicates a verb in 3rd person singular form. So, apart from its syntactic features, the analysis of the morphological features of the morpheme -s, which can either be a nominal inflection (plural marker) or a 3rd person singular verbal inflection, helps in identifying the tag of the word. Since, the morphemes attached to a word help in identifying the lexical tag in a tagging task; it becomes an essential requisite to linguistically analyze the morphemes and affixes to get a better grip on the performance of the tagger. Generally, a set of rules representing inflections and affixes is required to tag the corpus by a rule-based tagger. These hand-crafted rules can be used for POS disambiguation. In the present thesis, the inflectional endings are used to form rules to be fed to the tagger in order to identify POS tags of words. In this section, therefore, we shall discuss the verbal, adverbial, nominal, pronominal and adjectival inflections.

2.2.1. INFLECTIONS ON VERBS

A finite verb in Bangla constitutes root verb, followed by Causative, Aspect, Tense, Agreement, Honorificity and Emphatic suffixes. Let us briefly look at each of these here.
2.2.1.1. TENSE

There are three tenses in Bangla, namely, Present tense, Past tense and Future tense. However, only past tense and future tense have morphological realizations (Mondal, 2014). The present tense is marked null while the past and future tenses are marked with the affixes -l- and -b- respectively. For instance, in the verbs like, [kʰel-e] (plays) [kʰel-l-e] (played), and [kʰel-b-e] (will play), it can be observed that the root verb, [kʰel] (play) is followed by the agreement marker in the first example, but in the next two examples, the past tense and the future tense markers have been affixed to the verb root, followed by the agreement markers. So, it can be assumed that the -ɸ- marker is the present tense. In addition, the verbs like, [pɔr-cʰ-l-am] (was studying), [pɔr-cʰ-ɸ-i] (am studying) signify that a tense in Bangla occurs after the Aspectual marker and before the Agreement marker.

2.2.1.2. AGREEMENT INFLECTION

As pointed out in the previous paragraph, the agreement marker occurs after aspect and tense markers on a verb. In Bangla, subject-verb agreement is restricted only to Person agreement (Chatterji, 1926), (Chattopadhyay, 2003), (Mondal, 2011), (Mondal, 2014). That is, the verbs are inflected only for Person. Three forms of Person, namely, 1ˢᵗ Person, 2ⁿᵈ Person, and 3ʳᵈ Person are present and each form varies with Tense (Mondal, 2011).

The following table provides all the forms of agreement markers on a Bangla verb.
Table 3: Agreement markers

<table>
<thead>
<tr>
<th>PERSON</th>
<th>1 PERSON</th>
<th>2 PERSON</th>
<th>3 PERSON</th>
</tr>
</thead>
<tbody>
<tr>
<td>TENSE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRESENT</td>
<td>-i</td>
<td>-o</td>
<td>-e</td>
</tr>
<tr>
<td>PAST</td>
<td>-am</td>
<td>-e</td>
<td>-o</td>
</tr>
<tr>
<td>FUTURE</td>
<td>-o</td>
<td>-e</td>
<td>-e</td>
</tr>
</tbody>
</table>

2.2.1.3. HONORIFICITY

A finite verb in Bangla can also be marked for Honorificity. That is, there is honorific agreement between subject and verb (Chatterji, 1926). Note that, only 2nd and 3rd Person subjects show honorific agreement with the verb. The honorific markers representing 2nd person are –un in present tense as in, [kʰel-ϕ-un] (You(hon) play), –n in past and future tense as in, [kʰel-l-e-n] (He (hon) played), [kʰel-b-e-n] (He(hon) will play), and –en in past habitual like, [kʰel-t-en] (He (hon) used to play). These examples also imply that the honorific marker is always attached at the end of an inflected verb.

Chatterji (1926) and Chattopadhyay (2003) state that the non-honorific or the intimate markers, -ij and -i are found only in 2nd Person. These non-honorific markers have different forms in different tenses. -i gets inflected to the past and future tense verbs like, [kʰell-i] (You (nonhon) played), and [kʰelb-i] (you (nonhon) will play), whereas -ij occurs with the verbs of present and past habitual tense, such as, [kʰel-ij] (You (nonhon) play) and [kʰel.-ij] (You (nonhon) used to play) respectively. Since, the agreement
The following table shows the honorific and non-honorific markers of Bangla verbs.

<table>
<thead>
<tr>
<th>TENSE →</th>
<th>PRESENT</th>
<th>PAST</th>
<th>FUTURE</th>
<th>PAST HABITUAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>HONORIFICITY ↓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HONORIFIC</td>
<td>2P</td>
<td>3P</td>
<td>2P</td>
<td>3P</td>
</tr>
<tr>
<td>-un</td>
<td>-n</td>
<td>-n</td>
<td>-en</td>
<td>-n</td>
</tr>
<tr>
<td>NON- HONORIFIC</td>
<td>2P</td>
<td>2P</td>
<td>2P</td>
<td>2P</td>
</tr>
<tr>
<td>-if</td>
<td>-i</td>
<td>-i</td>
<td>-if</td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Honorificity

2.2.1.4. ASPECT

Chatterji (1926) and Chakrabarti (2008) have stated that the verbs in Bangla are inflected for four kinds of aspects, namely, Simple, Progressive, Perfect and Habitual. The aspect marker gets inflected to the root verb and is followed by tense, agreement and honorificity markers. According to Chattopadhyay (2003), the Progressive and Perfect aspects are found in both the present and the past tense but the Habitual aspect occurs only in past tense. In the verbs like, [kor-εʰ-ɪ] (doing), [kor-εʰ-l-am] (was doing), [kor-εcʰ-l-am] (had done) and [kor-ε-ɛm] (used to do/do), the highlighted affixes represent the present progressive, past progressive, past perfect and habitual aspectual
markers on a verb. Further, Mondal (2011) points out that the Simple aspect is marked null and - - indicates both past tense and habitual together. The following table displays all the aspectual markers in Bangla.

<table>
<thead>
<tr>
<th>TENSE</th>
<th>PRESENT</th>
<th>PAST</th>
</tr>
</thead>
<tbody>
<tr>
<td>PROGRESSIVE</td>
<td>-cʰ-/ccʰ-</td>
<td>-cʰr-/ccʰr-</td>
</tr>
<tr>
<td>PERFECT</td>
<td>-ccʰ-</td>
<td>-ccʰr-</td>
</tr>
<tr>
<td>HABITUAL</td>
<td>φ</td>
<td>-r-</td>
</tr>
<tr>
<td>SIMPLE</td>
<td>φ</td>
<td>φ</td>
</tr>
</tbody>
</table>

Table 5: Aspect markers

2.2.1.5. CAUSATIVES

Chakrabarti (2008) defines causative verbs as the ones that cause someone to do something. Mondal (2014) explains that a causative verb is formed by adding the affix -a- (if the verb root ends in a consonant) or -wa- (if the verb root ends in a vowel) to the verb root. For example, [pɔr] (to do) is a root word to which -a- affix can be added to make the root a causative root, [pɔrɑ] (to make someone read something). Similarly, [kʰa] (to give) is a verb root to which the affix -wa- is added to make it a causative verb, [kʰawɑ] (to feed). However, in the case of perfect aspect, the causative affix is -r-⁷⁸. The verbs, such as, [bʰor-t-ecʰ] (have made to fill), [lɪkʰ-t-ecʰlɑm] (had made to write) denote that the causative affix changes from -a- to -r- in perfect aspect. In addition, in the
progressive aspect, the causative morpheme -a- triggers in the change of the progressive aspect marker to -ccʰ-viii as in [kat-a-ceʰt] (making someone cut), [kat-a-ceʰt-lam] (was making someone cut) (Mondal, 2011), (Mondal, 2014).

2.2.1.6. EMPHATIC MARKERS

According to Chatterji (1926), Lahiri and Bayer (1990), and Dash (2015), Bangla has two emphatic particles -o and -i that exhibit the meaning ‘too’ and [+emphatic] respectively. The emphatic particles generally get attached at the end of an inflected verb after the agreement or the honorific markers. The boldfaced markers in [kʰa-be-i] (will eat[with emphasis]), [kor-be-o] (also do) indicate the emphatic markers. Lahiri and Bayer (1990) claims that in most of the cases, the emphatic markers cannot be added between the stem and the suffix endings. For example,

(2) am₁ oke *bol-cʰ-o-lam
   I him tell-PROG-EMP-PST-1P
   ‘I was telling him also’

(3) fe oke *mer-ecʰ-i-lo
   He him beat-PERF-EMP-PST-3P
   ‘He had beaten him only’

The examples (2) and (3) show that the sentences become ungrammatical when the emphatic markers get inserted between the stem and the endings. However, it has been observed that in case of a very few particular verbs and suffixes, the emphatic markers do come in between the root and suffix endingsix such as, [bole-i-cʰtlam] (already told), [kore-o-cʰtlam] (even did) (Dash, 2014).
2.2.2. INFLECTIONS ON ADVERBS

The adverbial suffixes can be divided into the morphological features of time, place, manner and degree. The suffix –e is one of the most common suffix marker which gets inflected to a noun or an adjective to form an adverb that includes the features of place, time and manner. For example [ekʰon-e] (here), [ţarpor-e] (after that), [dʰir-e] (slowly) represent the functional feature of place, time and manner respectively.

According to Racova (1990), quite a large number of adverbs are formed with the help of semi-suffixes like -bʰabe, -pokkʰe, -mɔto, such as, [bʰalo-bʰabe] (nicely), [nirde]-mɔto] (according to the instruction), [prokrito-pokkʰe] (in fact), which can have lexical meaning of their own. For instance, in the above mentioned examples, the adjective [bʰalo] (nice) gets inflected to the adverbial suffix -bʰabe, which also occur as a post-position. Likewise, in the adverbs [nirde]-mɔto], [prokrito-pokkʰe], both -mɔto and -pokkʰe are post-positions imparting the meaning ‘like’, and ‘on behalf of’ respectively. Since, the adverbial suffixes like these have a separate lexical entry, that’s why they are also known as semi-suffixes”. However, these inflectional endings denote the adverbs of manner.

The simple and underived bare adverbs refer to the morphological features of time and degree like, [ajkəl], (nowadays), [kəbe] (when), [ætɔ] (so much), [kʰub] (very). These adverbs do not get affixed to the adverbial suffixes. Hence, it can be assumed that the features of time and degree occur with the independent adverbial forms.

Now, Racova (1990) points out that, Bangla also uses adverbs borrowed from Sanskrit using the suffixes like, -aṭ, -fo, -fo, -cīṭ such as, [hɔtʰ-aj] (suddenly),
[ʃɑd̯ʰɑrɔn-ɪt̯o] (generally), [kromo-ʃo] (gradually), and [kɔd̯a-ɛt̯] (whenever). But, the adverbs formed with these suffixes are restricted to a very small list or they are just specific to few adverbs. It should be noted that, -ɑt, -t̯o are not considered to be suffixes anymore, because of their restricted and specific occurrence. They are treated as part of the adverbs itself, as in, [hɔtʰɑt̯], (suddenly), [ʃɑd̯ʰɑrɔnt̯o] (generally) and [kromoʃo] (gradually).

2.2.2.1. EMPHASIS AND ADVERBS

Both the derived adverbs and underived simple adverbs get inflected to the emphatic markers -ɪ and -o. In case of the derived adverbs, the emphatic markers get attached after the adverbial suffixes such as, –e, -mɔt̯o, -bɔʃot̯o, -mapʰɪk, and -bʰabe. However, the emphatic marker -ɪ gets attached to all the derived adverbs whereas the marker –o are attached only to those adverbs with the suffix –e and -bʱabe. Consider the following examples,

(4) ʃe bʱul-bɔʃot̯o-ɪ bolecʰe
He by mistake-EMP say
‘He said by mistake only’.

(5) ɭumi kajtə bʱalo-bʱabe-o kɔɾtɛ partɛ
You work nicely-EMP do-NF could have
‘You could have done the work nicely also’.

(6) ɭumi kajtə bʱalo-bʱabe-ɪ kɔɾtɛ partɛ
You work nicely-EMP do-NF could have
‘You could have done the work nicely also’.
In the sentence (4), the adverb \[bʰulbɔʃo]\textsuperscript{t} (by mistake) is inflected to \(-t\) marker to indicate the emphasis. On the other hand, in sentences (5), (6), (7) and (8) the adverbs \[bʰalobʰabe\] (nicely) and \[ʃekʰane\] (there) are inflected to both the emphatic markers \(-t\) and \(-o\) respectively.

Apart from the inflected adverbs, the emphatic markers get attached to the simple adverbs as well like- \([α-0]\) (even today), \[tɨmoq̪e-1]\textsuperscript{o} (in the mean time).

### 2.3. INFLECTIONS ON NOUNS

Bangla exhibits a strong and structural inflectional morphology for nouns. A nominal suffix refers to the inflected affix that gets attached to a noun expressing the grammatical function like definiteness, case, plurality and so on. Based on the function and features of nominal inflections, the Bangla nominal suffixes have been primarily classified into three categories, namely, Case, Classifiers and Emphasizers (Chatterji, 1926) and (Dasgupta, 1985). Apart from these three categories, the system of inflected nouns also includes the following substantive categories, that is, number, definiteness, and animacy (Chatterji, 1926), (Dasgupta, 1985), (David, 2015).
2.3.1. CASE

Bangla nouns are inflected for seven types of Case, which includes nominative, accusative, genitive, locative, dative, instrumental and ablative case (Chatterji, 1926), (Dasgupta, 1985), (Bhattacharya et al, 2005), (Deb, 2012).

2.3.1.1. NOMINATIVE CASE

The nominative case marker is used to mark the subject in a sentence (Chatterji, 1926), (Chakrabarti, 2008). In Bangla nouns, the nominative case is generally marked as null -ɸ- or with the marker –e-xi (Chatterji, 1926), (Chakrabarti, 2008).

2.3.1.2. ACCUSATIVE CASE

The accusative case markers are used to mark the direct object of a verb. The accusative case is realized as –ke with singular and animate nouns and pronouns as in, [mou-ke] (to Mou), [baba-ke] (to boy), [e-ke] (to him/her), [o-ke] (to him/her) (Bhattacharya &Simpson, 2007, 2011). It should be noted that this accusative marker occurs only with the animate human nouns. For instance, the animate non-human nouns like, [macʰ] (fish), [cʰəgol] (goat), [gacʰ] (tree) form ungrammatical expressions if they get affixed to -ke like, [*macʰ-ke], [*cʰəgol-ke], and [*gacʰ-ke] respectively.

2.3.1.3. DATIVE CASE

Chatterji (1926), and Deb (2012) state that the role of a dative case is to distinguish the recipient of something given, transferred and so forth. The dative case markers are used to express the relation of the indirect object of the verb. The accusative
marker –ke also acts as a dative case marker in Bangla nouns. Chatterji (1926) claims that, apart from the –ke marker, other markers such as, -re, -ere, -e also function as dative markers. However, these markers are generally found in different dialects of Bangla like, the –re marker gets inflected to the indirect nouns in Sylheti Bangla (Deb, 2012). The highlighted suffix in the sentence below represents the dative case marker in Bangla,

(9) ūṃ t̯umɪ bōlta manai-ke d̯ ɑo
You ball-ACC mania-DAT give
‘You give the ball to Manai’.

2.3.1.4. GENITIVE CASE

Chattopadhyay (2003), Chakrabarti (2008), and (Deb, 2012) have defined genitive case as the one that indicates the possessive relation of nouns and pronouns. It defines the possessor-possessed relationship between two objects. The genitive case in Bangla is realized through the two markers –r and –er such as, [rɔːɪ-r] (king’s), [rɔm-er] (Ram’s) (Chatterji, 1926), (Chattopadhyay, 2003). The –r marker occurs with the nouns ending with vowels while, the –er marker is inflected to the consonant final nouns. For example, the genitive markers, -r and -er occur with the vowel final head nouns, [lokɔ̃l-ɪr] (Lakshmi’s), [bɔba-ɪr] (father’s) and consonant final head nouns, [pukʊr-er] (pond’s), [tɛbɪl-er] (table’s) respectively.

2.3.1.5. LOCATIVE CASE

Deb (2012) states that the locative case indicates the place of action or event identified by the verb. According to Chatterji (1926), Chattopadhyay (2003), Bhattacharya and Simpson (2011), the locative case marker marks the location with both
the spatial and temporal references. In Bangla nouns, the locative case marker is formed by the affixes –e and -te. Similar to the genitive case markers, the locative case markers also get inflected according to the nature of the root noun. If the noun ends in a consonant, the marker is –e, like, [matʰ-e] (in the field) and if the root noun ends in a vowel the marker formed is -te, such as, [kolkata-te] (in Kolkata). Apart from these inflected case markers, the post-positions such as, [diye], [modʰe], [teke] and [maɭʰe] can also denote location.

2.3.1.6. INSTRUMENTAL CASE

According to Chatterji (1926), Chakrabarti (2008), and (Deb, 2012), instrumental Case refers to the instrument used in carrying out some action identified by the verb. In Bangla, the instrumental case markers are the same as the locative case markers, that is, the –e and -te markers like- [cokʰ-e] (through eyes), and [taka-te] (with money). However, apart from these inflected case markers, the post-positions like, [diye] and [daɾa] also indicate the instrumentality of a person through or by whom an action is accomplished as explained in the following set of data,

(10) amĩ cʰuɾi diye apel katlam
     I   knife的现象 by-PP apple cut
     ‘I cut the apple with the knife’.

(11) tɔmar daɾa karta hɔbe na
     You  by-PP work happen not
     ‘The work will not be possible by you’.

68
2.3.1.7. ABLATIVE CASE

Chatterji (1926) states that the ablative Case in Bangla specifies the person, thing or place from which the subject is removed, obtained or taken away. The main role of an ablative case is to indicate movement away from some location, person or thing (Deb, 2012). In Bangla, the ablative case is realized as -te, -t'eke, -er, -e, [dije] and so forth. Further, Chatterji (1926) also points out that since ablative case does not take any specific case markers, the genitive, locative and instrumental case markers as well as the postpositional words are used to mark the ablative case. The following examples validate the point,

(12) ɑmɪ ɔɪ ɡɑcʰ-er  pejərə _cmds̱  I that tree-ABL guava buy

‘I bought that tree’s guava’.

(13) ɗuɗb ɗije  cʰana  hɔɪ

Milk from-ABL cheese be

‘Cheese is made from milk’.

It is observed that in (12), the –er marker is the genitive marker inflected to the noun [ɡɑcʰ] (tree) to signify the ablative case. Similarly, the postpositional word [dije] (13) also indicates an ablative case marker.

Chatterji (1926) has observed that, Bangla does not have unique case markers. Each case has a number of case markers out of which few of the markers are used as common markers for more than one case. Therefore, due to the interchangeability among the four cases, markers like –e denotes the nominative case, locative case, and the dative case as well. Likewise, the ablative case is marked by post-positions like [dije], as well as
the locative, instrumental and genitive case markers like –er, and -e. Hence, we can claim that the case markers for Bangla are ambiguous in nature.

2.3.2. CLASSIFIERS

Racova (2007) states that classifiers are the affixes which categorize nouns into noun classes based on the shared, objectively perceived distinctive properties of objects they denote. Bangla has a rich classifier system which is restricted to the inflectional nominal suffixes of noun classes (Chatterji, 1926), (Chakrabarti, 2008). The Bangla classifier system uses a set of classifiers which includes the default classifier -ta, the human classifier -jon, the plural classifiers –gulo, -der and -ra, the inanimate count classifier -kʰa, -gacʰa, the inanimate mass classifier -kʰa, and the uncountable classifier –tu (Ghosh, 1995).

Besides nouns, Bangla classifiers can also be affixed to adjectives, numerals, quantifiers, and pronouns (Das, 2011). The following table summarizes the categories to which various classifiers can get affixed to.

<table>
<thead>
<tr>
<th>Classifier Category</th>
<th>-ta</th>
<th>-jon</th>
<th>–gulo</th>
<th>-der</th>
<th>-kʰa</th>
<th>-gacʰa</th>
<th>–tu</th>
<th>-ra</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noun</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Pronoun</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Adjective</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Quantifier</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Numeral</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 6: Word classes and their respective classifiers
2.3.2.1. -ta

The classifier -ta is said to be the default classifier of Bangla (Dasgupta, 1983), (Ghosh, 1995), (Bhattacharya, 2001). According to Bhattacharya (2001), the default form of the Bangla classifier -ta includes various allomorphs which are governed by phonological conditions. The allomorphs of -ta are, -ta, -te, -to and -tr (Dasgupta, 1983), (Ghosh, 1995). Now, Dasgupta (1983) in his paper has given a detailed account of the various morpho-phonologically conditioned environments in which the classifier -ta occurs. Firstly, it occurs with both the count nouns and the mass nouns such as, [lok-ta] (the man) as opposed to [dudh-ta] (the milk). [lok] (man) is a count noun whereas [dudh] (milk) is a mass noun, both of which get inflected to -ta. Secondly, the classifier -ta gets affixed to singular nouns only and not the plural form of nouns, as in [chele-ta] (the boy), [pakhi-ta] (the bird), [khabar-ta] (the food). The plural nouns like, [chele-ra] (the boys), [khabar-gulo] (the foods) become ungrammatical when they get affixed to the -ta marker as in [*chele-ra-ta] and [*khabar-gulo-ta]. It can be said that, -ta is singular in number and so, it cannot be inflected to a plural nounxiv.

Moreover, -ta is the only classifier which indicates the clock time when it gets attached to numerals like one, five, seven etc (Das, 2011). For instance,

(14) kokhon æk-ta babe?

When one-CL ring

‘When the clock will strike one?’.
The example above states that when the numeral [æk] (one) (14) gets affixed to -ta, it signifies the clock time of one. It can be concluded that the classifier -ta attaches to the numeral ‘one’ to indicate time.

Finally, -ta is used in informal speech indicating non-honorificity. Furthermore, -ti, the allomorph of -ta also exhibits the feature of both honorificity and non-honorificityxv.

Unlike -ta, -te and –to do not get inflected to nouns, pronouns and noun modifiersxvi (Dasgupta, 1983), (Ghosh, 1995), (Bhattacharya, 2001). For instance, [*lok-te] (the man), [t̯or-te] (yours’) are ungrammatical constructions as opposed to [lok-ta] and [t̯or-ta].

The allomorph -to occurs with only the numeral ‘two’ and –te gets affixed to the numerals, three, four, three and a half and so onxvii. Furthermore, -to and –te also denote time, such as,

(15) ʧumɪ ʤu-tɔ-r ʃomɔi ɛfo
    You two-CL around come
    ‘You come around at two’o clock’.

(16) ɑmɪ ɑrɑi-te-tʃabo
    I two-thirty-CL go
    ‘I will go at two-thirty’.

The allomorph -ti occurs in singular number and can be found with both inanimate and animate nouns as in [gacʰ-ti] (the tree), [boi-ti] (the book), [bacca-ti] (the
kid), [meje-ti] (the girl). It can also be realized with countable nouns such as, [cʰele-ti] (the boy) and non countable nouns like, [jɔl-ti] (the water).

It should be noted that in the present work, these allomorphs along with the classifier -ta have been grouped into an individual POS tag set Classifier (CL) while designing the tagset for CRF tagging. However, the tag CL refers to only those words where these classifiers get inflected to numerals, like, [tɪnte] (three), [d̯uto] (two), [æktɑ] (one) and so forth.

2.3.2.2. -kʰana

Bhattacharya (2001) refers to -kʰana as the inanimate count classifier. Both the classifiers, -ta and -kʰana share common morphosyntactic features. This classifier is restricted to inanimate count nouns, quantifiers and numerals. It also expresses definiteness when the noun is placed before the numeral-classifier combination (Ghosh, 1995). The allomorph of -kʰana is -kʰani. This allomorph is restricted to mass nouns. The -kʰana and -kʰani classifiers can denote both singularity and plurality.

Let us explain this classifier with examples. Nouns like, [boɪ] (book), [pen] (pen) are inanimate and count nouns to which -kʰana gets inflected to form [boɪ-kʰana] (the book), and [pen-kʰana] (the pen) respectively. These expressions also denote singularity and definiteness as the head noun gets directly affixed to the classifier. From the examples like, [tɪn-kʰana pen] (three pens), [d̯u-kʰana boɪ] (two books), it can be deduced that, when the numeral classifier combinations [tɪn-kʰana], [d̯u-kʰana] precede the inanimate, count nouns like, [pen] and [boɪ], they generate indefinite expression. On the contrary, if the head noun follows the numeral classifier combination as in, [boɪ
dq-[kʰən̥a] (two books), it expresses definiteness. In case of the allomorph -kʰun̥u, the nominal expressions, such as [one-kʰən̥i ʃo] (a lot of water), [ɪmʃ-kʰən̥i] (the thing) denote that this classifier can get affixed to quantifiers and mass nouns.

2.3.2.3. -gacʰa

Ghosh (1995) states that the classifier -gacʰa is used when the following noun denotes something of a thin, long, string like piece. -gacʰa can be realized in quantifiers and numerals only in the case of inanimate nouns as in, [kɔ-gacʰa ʃu] (few bundles of thread) [æk-gacʰa ʃu] (one bundle of thread), where [kɔ] (few) and [æk] (one) denote quantifier and numeral respectively. -gacʰa also refers to both singular and plural number like, [æk-gacʰa] (one bundle) and [dq-gacʰa] (two bundles). The allomorph of -gacʰa is -gacʰi.

2.3.2.4. -jɔn

The classifier -jɔn is known as the ‘human count classifier’ (Dasgupta & Bhattacharya, 1993). It gets affixed to human nouns, numerals and quantifiers. This classifier occurs with numerals in indefinite expressions and also expresses definiteness marking by altering the word order. It marks both the singular and plural number. It is also used to indicate both masculine and feminine gender. Similar to the default classifier -ta and its allomorphs, this human classifier marker is also included in the POS tag category of Classifier (CL). But, it specifically refers to the numerals which get attached to -jɔn, such as, [t̪u-n-jɔn] (three people), [æk-n-jɔn] (one person) and more.
2.3.2.5. -tu

The classifier –tu refers to a small amount or quantity of something. Its usage is restricted only in expressions with ‘ek’ (Ghosh, 1995). However, this ‘ek’ does not correspond to the numeral [æk] (one). It occurs only with uncountable nouns which suggest that the numeral-classifier combination is not possible with this marker (Ghosh, 1995). The expressions like, [ek-tu ḥol] (a little bit of water), [ek-tu tɔ ולהawa] (some fresh air) convey the meaning of ‘few’, ‘some’, where the ‘ek’ to which the classifier -tu gets affixed is not a numeral because numerals do not occur with uncountable nouns.

2.3.2.6. -tuku

The classifier –tuku is also used in uncountable nouns like –tu, and gets inflected to mass nouns (Ghosh, 1995), (Ghosh, 2011) like, [ɪcʰe-tuku] (small wish). This diminutive classifier also imparts the meaning of ‘some’, ‘little’. Besides noun, it also occurs with determiner such as, [er-tuku] (this much) which suggests the meaning of littleness.

2.3.2.7. PLURAL CLASSIFIERS

According to (Ghosh, 1995), Dayal (2014), Bangla does not mark the singular-plural distinction on the noun, but like the other classifier languages, it has markers that refer to pluralities. So, the Bangla classifier system marks three plural classifiers, -gulo, -ra and -der.

The classifier –gulo occur with both animate and inanimate nouns (Ghosh, 1995), (Dayal, 2014) as in, [cʰele-gulo] (the boys) and [cʰobt-gulo] (the pictures). It can be
inflected to count nouns, adjectives and pronouns. Furthermore, it does not occur with numerals and also does not express definite or indefinite expressions, unlike the other classifiers. 

–gulo signifies a collective classifier which leads to get affixed to quantifier as well.

The following examples explain the different occurrences of –gulo. The plural nouns like, \[kʰarap-gulo\] (bad ones), \[e-gulo\] (these), \[onek-gulo\] (many) signify that the -gulo classifier get inflected to adjectives, pronouns and quantifiers. Moreover, the expression \[*t̯ɪn-gulo\] (three) is ungrammatical because -gulo can never occur with quantifiers. -guli is realized as the allomorph of –gulo. It has similar properties as the –gulo classifier\textsuperscript{ix}.

-ra is the plural classifier which is restricted only to count animate nouns. According to Dayal (2014), –gulo can denote more than one entity, but -ra purely implicates plural reference. So, -ra is realized only with count animate nouns, adjectives and pronouns (Ghosh, 1995) such as, \[bacca-ra\] (children), \[lamba-ra\] (the tall ones), and \[tom-ra\] (you all) respectively. Ghosh (1995) has also claimed that, like –gulo, -ra also doesn’t occur with numerals. Furthermore, the -ra classifier does not occur with inanimate mass nouns either\textsuperscript{xx}.

Finally, the last plural classifier of Bangla is -der. Just like -ra, this classifier is also inflected to only count and animate nouns (Ghosh, 1995). For instance, -der can occur with nouns like [cʰele], which is a count animate noun, but, it cannot occur with inanimate nouns like [gɑrɪ] (car). So, -der can never generate nouns like, \[*gɑrɪ-der\] (the cars) or \[pʰul-der\] (flowers). But, the difference between -ra and -der lies in the case
marking. -der is marked for oblique case whereas -ra refers to nominative case. Similar
to -ra, -der also occurs with an adjective and a pronoun such as, [mota-der] (fat people),
[ta-der] (their) and so forth.

The table below summarizes the classifiers with their features.

![Table 7: Classifiers of Bangla](image)

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Feature</th>
<th>ta</th>
<th>te</th>
<th>to</th>
<th>t̯una</th>
<th>t̯un</th>
<th>gac̠u</th>
<th>gac̠</th>
<th>gulo</th>
<th>gur</th>
<th>j̯on</th>
<th>ra</th>
<th>der</th>
<th>tu</th>
<th>tuku</th>
</tr>
</thead>
<tbody>
<tr>
<td>Animate</td>
<td>+/-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Count</td>
<td>+/-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Singular</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

2.3.3. EMPHATIC MARKERS

Dash (2015) states that the two emphatic particles in Bangla, namely, -i and –o,
are tagged at the end of the nouns for emphasis such as, [bon-i bollo] (sister only told),
and [baba-o bollo] (father also told). However, Dasgupta (1984) claims that these
emphatic particles -i and –o are clitics and they do not partake in lexical morphology.

2.4. PRONOUNS

Pronouns are the sub-category of nouns which are anaphoric and referential in
nature. Generally, the pronouns are inflected for number, person and case. They also
encode the distinction of honorificity and non-honorificity as well as proximity^{xvi}. They
get inflected for singular and plural number. However, according to David (2015), unlike nouns, the non-human pronouns do not get attached to plural suffixes. The plural number is realized through the classifiers –gulo and -guli which get added to the pronoun stems. The pronouns are realized through the five cases, nominative, accusative, dative, genitive and locative.

The analysis of the pronouns is relatively simple. The stem carries the information on person, honorificity, and proximity, whereas the suffix marker denotes the case and number. David (2015) and Thompson (2010) have categorized the pronouns according to their features.

2.4.1. PERSONAL PRONOUNS

Thompson (2010) claims that the personal pronouns can function as subject or object of a sentence. The personal pronouns can be broadly classified into 1st person, 2nd Person and 3rd Person which includes the features of case, degree of formality and number. The 1st person personal pronouns are illustrated in the following table:

<table>
<thead>
<tr>
<th>Number Case</th>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominative</td>
<td>amā (I)</td>
<td>amra (we)</td>
</tr>
<tr>
<td>Accusative/Dative</td>
<td>amako (me)</td>
<td>amader (us)</td>
</tr>
<tr>
<td>Genitive</td>
<td>amar (mine)</td>
<td>amader (our)</td>
</tr>
<tr>
<td>Locative</td>
<td>amate (in myself)</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 8: 1st Person pronouns
The 2\textsuperscript{nd} person pronouns have the three-way distinction of honorificity- honorific, familiar and intimate (Chattopadhyay, 2003)\textsuperscript{xxii}. The table below summarizes the 2\textsuperscript{nd} person pronouns.

<table>
<thead>
<tr>
<th>Case</th>
<th>Number Honorificity</th>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominative</td>
<td>Intimate</td>
<td>tfoot (you)</td>
<td>tfoota (you all)</td>
</tr>
<tr>
<td></td>
<td>Familiar</td>
<td>tfoot (you)</td>
<td>tfoota (you all)</td>
</tr>
<tr>
<td></td>
<td>Honorable</td>
<td>tfoot (you)</td>
<td>tfoota (you all)</td>
</tr>
<tr>
<td>Accusative/Dative</td>
<td>Intimate</td>
<td>tfoot (to you)</td>
<td>tfoota (to you all)</td>
</tr>
<tr>
<td></td>
<td>Familiar</td>
<td>tfoot (to you)</td>
<td>tfoota (to you all)</td>
</tr>
<tr>
<td></td>
<td>Honorable</td>
<td>tfoot (to you)</td>
<td>tfoota (to you all)</td>
</tr>
<tr>
<td>Genitive</td>
<td>Intimate</td>
<td>tfoot (your)</td>
<td>tfoota (yours)</td>
</tr>
<tr>
<td></td>
<td>Familiar</td>
<td>tfoot (your)</td>
<td>tfoota (yours)</td>
</tr>
<tr>
<td></td>
<td>Honorable</td>
<td>tfoot (your)</td>
<td>tfoota (yours)</td>
</tr>
<tr>
<td>Locative</td>
<td>Intimate</td>
<td>tfoot (you)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Familiar</td>
<td>tfoot (you)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Honorable</td>
<td>tfoot (you)</td>
<td></td>
</tr>
</tbody>
</table>

Table 9: 2\textsuperscript{nd} Person Pronouns

The 3\textsuperscript{rd} person pronouns encode two levels of politeness- honorific and non-honorific. They are also marked for the three levels of spatial proximity- proximal, distal and unmarked. The last form of proximity is considered to be the form which denotes elsewhere or out of sight. The 3\textsuperscript{rd} person pronouns also act as demonstrative pronouns and demonstrative adjectives David (2015). This group of pronouns also includes the feature [+human]. Furthermore, the 3\textsuperscript{rd} person pronoun without the [+human] feature gets inflected with the classifier -ta and –gulo in their singular and plural forms respectively.
These are also known as inanimate pronouns which refer to facts, non-human objects and circumstances. In most of the cases, the pronoun forms in nominative and accusative cases are the same. The tables below present the 3rd Person pronouns with [+human] and [-human] feature.

<table>
<thead>
<tr>
<th>Number Case</th>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominative</td>
<td>Proximal</td>
<td>Distal</td>
</tr>
<tr>
<td>Honorific</td>
<td>mi (He/she)</td>
<td>mi (He/she)</td>
</tr>
<tr>
<td></td>
<td>enmi (He/she)</td>
<td>enmi (He/she)</td>
</tr>
<tr>
<td>Non-</td>
<td>e (He/she)</td>
<td>o (He/she)</td>
</tr>
<tr>
<td>honorific</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accusative</td>
<td>Proximal</td>
<td>Distal</td>
</tr>
<tr>
<td>Honorific</td>
<td>ěka (him/her)</td>
<td>ěka (him/her)</td>
</tr>
<tr>
<td></td>
<td>eneka (him/her)</td>
<td>eneka (him/her)</td>
</tr>
<tr>
<td>Non-</td>
<td>eke (him/her)</td>
<td>oke (him/her)</td>
</tr>
<tr>
<td>honorific</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genitive</td>
<td>Proximal</td>
<td>Distal</td>
</tr>
<tr>
<td>Honorific</td>
<td>ėr (his/her)</td>
<td>ėr (his/her)</td>
</tr>
<tr>
<td></td>
<td>enăr (his/her)</td>
<td>enăr (his/her)</td>
</tr>
<tr>
<td>Non-</td>
<td>ěr (his/her)</td>
<td>ěr (his/her)</td>
</tr>
<tr>
<td>honorific</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 10: 3rd Person pronouns with [+human] feature

80
Table 11: 3rd Person pronouns with [-human] feature

### 2.4.2. RELATIVE PRONOUNS

Thompson (2010) says that the relative pronouns are used in correlative constructions and the morphological pattern is similar to the 3rd person pronouns. So, the relative pronouns distinguish between both honorific/non-honorific as well as the human/non-human features (David, 2015). Moreover, they use the same plural classifier endings as the 3rd person pronouns. The following table provides the relative/correlative pair of pronouns.
2.4.3. REFLEXIVE PRONOUNS

David (2015) describes the reflexive pronouns in Bangla as the pronouns which are used when two core arguments of a transitive clause refer to the same referent. In Bangla, these are also used to indicate emphasis. The Bangla reflexive pronouns do not encode for person. It has the same form for all the three persons, that is, 1st person, 2nd person and 3rd person. However, the reflexive pronouns get inflected to the case markers of Nominative, Accusative and Genitive case. The table below lists down the reflexive pronouns of Bangla.

<table>
<thead>
<tr>
<th>Number Case</th>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominative</td>
<td>نুজে (ownself)</td>
<td>نুজের (ourselves)</td>
</tr>
<tr>
<td>Accusative</td>
<td>نুজেকে (myself)</td>
<td>نুজেগের (ourselves)</td>
</tr>
<tr>
<td>Genitive</td>
<td>نুজের (mine)</td>
<td>نুজেগের (ourselves)</td>
</tr>
</tbody>
</table>

Table 12: Relative pronouns

Table 13: Reflexive pronouns
2.4.4. INTERROGATIVE PRONOUNS

Thompson (2010) states that Bangla differentiates between singular and plural interrogatives for animate nouns. The interrogative pronouns include the declinable [kt] (what) and [ke] (who). This category of pronouns also includes the [+human]/[-human] feature. The derivational pronouns like [kɔtɔ] (how much) have been derived from adjectives and they combine with the classifiers -ta and –gulo to yield the singular and plural forms respectively. The following two tables list the interrogative pronouns with human and non-human features.

<table>
<thead>
<tr>
<th>Number Case</th>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominative</td>
<td>ke (who)</td>
<td>kara (who all)</td>
</tr>
<tr>
<td>Accusative</td>
<td>ka-ke (whom)</td>
<td>ka-der (whose)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ka-der-ke (whose)</td>
</tr>
<tr>
<td>Genitive</td>
<td>ka-r (whose)</td>
<td>ka-der (whose)</td>
</tr>
</tbody>
</table>

Table 14: [+human] interrogative pronouns

<table>
<thead>
<tr>
<th>Number Case</th>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominative/Accusative</td>
<td>kt (what)</td>
<td>kt (what)</td>
</tr>
<tr>
<td></td>
<td>kon-ta (which one)</td>
<td>kon-gulo (which ones)</td>
</tr>
<tr>
<td>Genitive</td>
<td>kif-c-r (of what)</td>
<td>kif-cr (of what)</td>
</tr>
<tr>
<td></td>
<td>kon-ta-r (of which one)</td>
<td>kon-gulo-r (of which ones)</td>
</tr>
<tr>
<td>Locative</td>
<td>kif-c (in what)</td>
<td>kif-c (in what)</td>
</tr>
</tbody>
</table>

Table 15: [-human] interrogative pronouns
2.5. ADJECTIVES

As it has been stated before (cf. 2.1.3) Bangla adjectives do not take any inflections. However, there are a few distinctive suffix endings in the case of adjectives which have been derived from verbs and nouns. So, if we are not sure of any word as an adjective, these inflectional endings help to identify them as one. Let us illustrate this further with few of the adjectival suffixes. For instance, -ɪk and -e are the adjective suffix endings which have been derived from nouns. The nouns like, /d̯ɪm/ (day) and /ʃekal/ (that time) changes to adjectives, [d̯oɪn-ɪk] (daily) and [ʃekel-e] (old-fashioned) with the addition of these adjectival suffixes. Similarly, the suffixes -ɪ and -ɪt̯o get inflected to nouns to become adjectives as in, [bʰər-ɪ ~bʰər] (heavy~weight) and [ʃɪkkʰ-ɪt̯o ~ʃɪkkʰɑ] (educated~education) respectively.

2.6. QUANTIFIERS

The quantifiers are classified under the category of nominal modifiers, that is, adjectives. In other words, quantifiers form sub-classes of adjectives which also imply that the quantifiers and the adjectives share similar morphological features. However, Thompson (2010) argues that, quantifiers differ from adjectives by their capability of retaining their word class, even after getting inflected to a classifier.

The quantifiers like, [ɔt̯o] (so much), [kɔt̯o] (how much), [ektu] (a bit) can function both as adverbs and adjectives depending upon the context. These quantifiers, except [ektu] can be used with both count and non-count nouns and can get inflected to singular as well as plural classifiers. The quantifier, [ektu] gets added only to the classifier -kʰɑn. Besides, the function of the quantifier [kɔt̯o] (how much) is restricted to
interrogative and exclamatory sentences. Apart from the classifiers the quantifiers also get attached to the emphatic markers as in expressions like, /æt̯otɑ-ɪ ɑcʰe/ (this much only is there).

2.7. SUMMARY

This chapter analyzes and discusses the word classes of Bangla and their inflectional suffix endings in relation to the word categories and inflectional endings. The word classes of Bangla have been categorized into two major and distinct groups, open and closed classes. The open class word category comprises nouns, verbs, adjectives and adverbs, while, the closed class of words are listed as, post-positions, conjunction, emphasis, negative particles and pronouns. The chapter also discusses inflectional morphology of the open class category of parts of speech, that is, verbs, adverbs, nouns, pronouns, adjectives and quantifiers. The inflected verbal suffixes in Bangla are distinguished on the basis of the morphological features of Tense, Aspect, Agreement, Honorificity, Causative and Emphasis. The causative markers are the first ones to get attached to the root verb and the emphatic markers come at the end of an inflected suffix.

The adverbial suffixes are analyzed based on the morphological functions of place, time, manner, and degree. The suffix –e is one of the most common suffix marker which gets inflected to a noun or an adjective to form an adverb and includes the features of place, time and manner. Most of the underived adverbs exhibit the features of time and degree. Moreover, the two emphatic markers, namely, -ɪ- and –o- get inflected to both derived adverbs as well as the bare adverbs.
In the nominal system, the nouns are inflected for Case, Classifiers, and Emphasizers. Seven distinct cases are present in Bangla, namely, Nominative, Accusative, Dative, Genitive, Locative, Instrumental and Ablative. The classifier system includes a closed class of affixes which categorize the noun classes based on their morpho-phonologically conditioned occurrences. The classifiers in Bangla consists of the default classifier -ta, the plural classifiers -gulo, -ra, -der, the human classifier -jon, and the inanimate classifiers -kʰana, -gacʰa, -tu, and -tuku that are categorized based on the features of animacy, definiteness, count and plurality. The emphatic markers, -ɪ and -o get affixed to the word final nouns to express emphasis.

Pronouns constitute the sub-class of nouns and are classified into eight types, namely, Personal pronouns, Relative pronouns, Demonstrative pronouns, Inclusive pronouns, Interrogative pronouns, Indefinite pronouns, Reflexive pronouns and Reciprocal pronouns. They exhibit the morphological features of person, case, number, honorificity and proximity. Even though the adjectives do not exhibit uniform morphological features, there are few derived inflectional suffixes that help in identifying the inflected adjectives.

Therefore, based on these analyses, the list of Bangla inflectional suffixes has been developed which act as the main input of rules to feed the rule-based analyzer. The following chapter expounds on the development of the Hybrid Tagger.
NOTES

i These words are termed as structural or function words.

ii A *bare noun* is considered as the root noun without any classifiers, modifiers, and case endings.

iii Syntactically, adjectives can function attributively and predicatively (David, 2015). An attributive adjective occurs before a noun as in, [baɾe kʰɔboɾ] (bad news), [miftʰ baccʰa] (sweet kid). On the other hand, a predicative adjective occurs as complements like, [ʃol noŋra] (the water is dirty), and in copular sentences, such as, [kɔtʰin cʰɪlo] (was difficult), [bʰalo hobe] (will be good).

iv Thompson (2010) classifies the adjectives into different types based on their semantic features, viz, qualifying adjectives, quantifying adjectives, and distributive adjectives. The qualifying or the distributive adjectives refer to the

a. Physical attributes such as, age, shape, size and color, as in, [mɪʃtʰ] (sweet), [lɔmba] (tall), [kalo] (black), [bʰoṭa] (blunt), [pracʰ:n] (ancient) and more.

b. Non-physical attributes like, [ɔmanobik] (inhumane), [ʃotil] (complex), [boka] (fool), [dɪʃɛhara] (confused), [alaḍa] (different) and so on.

c. Comments and opinions as in [projoŋno] (necessary), [uporukʔo] (suitable), [ʃunḍor] (pretty), [bɪpɔɔɲoŋk] (dangerous), [dʌɾun] (excellent) and so forth.
The quantifying adjectives refer to the numerals, and ordinals such as [protʰom] (first), [dɔṭ̚jo] (second), [τrɪt̚jo] (third), [sɔb] (all), and [bɛʃi] (more). Lastly, the distributive adjectives indicate the adjectives which refer to each person or thing separately, as in, [proṭṭek] (every), [onno] (other), [proṭi] (each) and so forth. Moreover, the comparison in adjectives is expressed through the use of post-positions like [tʰeke], [ceje] (than), [hoṭe] (from) and more. For instance,

(i) nɪl  ləl-er tʰeke  bʰalo.
   blue  Red  than  better
  ‘Blue is better than red’.

(ii) ɑmɪ  or  ceje  lɔmba
   I  him  than  tall
  ‘I am taller than him’.

Further, Dash (2015) indicates that according to the structural components, four types of adjectives exist in Bangla. First, the simple adjectives or the bare adjective forms which do not have any suffix or prefix inflected to them such as, [tʰanda] (cold), [kʰrap] (bad), [lɔmba] (tall), [bete] (short), [protʰom] (first), [coṭurtʰo] (fourth) and so forth. The adjectives derived from nouns and verbs along with the addition of the primary adjectival suffixes form the second group. [cɔloman] (moving), [bʰabuk] (thoughtful), [raṭrɪkaːlːn] (nocturnal) are few of the derived adjectives from nouns and verbs after getting inflected to primary adjectival suffixes like, -man, -uk and -ːn respectively. The third type constitutes the adjectives which are formed by adding the secondary adjectival suffixes to the nouns and verbs, as in, [brihoṭṭomo] (largest), [pranobɔntɔ] (lively), [kajahːn] (without body) etc. Finally, the fourth type of adjectives are those which are derived by adding a prefix to a base
adjective form like, [ʃucot̯ur] (shrewd), [ʃuprɑcɪn] (very ancient), [ɔkɑt̯or] (fearless) where
the prefix /u- and œ- are added to the adjective base [coṭur], [prɑcɪn], and [kɑt̯or] respectively.
Dash (2015) further points out that identifying adjectives is a tricky task since, a large
number of adjectives are found in bare forms, that, is without any suffix or prefix markers.

So, words like, [bɔro] (big), [cʰoto] (small), [gɔrom] (hot), [tʰɑndɑ] (cold) can be either a
noun or an adjective depending on the context.

It has been observed by Mondal (2011), that both the past habitual and the 3rd person past
tense honorificity is indicated by the marker –o and it is different from the rest of the
honorary markers. If we go by the order of suffixes attaching to a verb, then the 3rd person
past tense verb is supposed to be [*kʰelţoen], which is ungrammatical because two vowels
can never coexist. So, Vowel Deletion takes place deleting the agreement marker –o and
thus, changing the honorary marker to –en, thereby changing the verb to [kʰelţen] (used to
play) (Mondal, 2014). Likewise, the 2nd person present tense verb [kʰelun] has –un as the
honorary marker because the 2nd person present tense agreement marker is –o which gets
deleted by Vowel Deletion and gets replaced by –un (Mondal, 2014).

Mondal (2011) and Mondal (2014) have explained the discrepancy in the aspectual
morphemes in her work. Interestingly, it has been observed that the present progressive verb
like [tulcʰɪ] (picking), should ideally be [*tul-cʰ-ɪɪ] where -cʰɪ- is the progressive aspect, -ɪ
is the agreement marker and the null present tense marker. But, this doesn’t happen since both
the Progressive aspect affix and the agreement marker for 1st person present tense ends in a
vowel, i.e. -ɪ. Therefore, one of the vowels is deleted by the Vowel Deletion rule, thereby
changing the progressive aspect marker from -cʰɪ to -cʰ-. Likewise, for Perfect aspect, the
aspect marker in present tense is -ecʰ- but, in past tense it is -ecʰɪ- as in, [kor-ecʰɪlam] (have done) ~ ‘kor-ecʰɪlam] (had done). Furthermore, in case of verbs with vowel ending verb roots, the progressive affix changes from -cʰ(t)- to -ccʰ(t)- as in [kʰaceʰ] (eating), [gʰumaceʰɪle] (sleeping), due to the presence of the vowel in the vowel ending verb roots that necessitate Gemination.

vi Mondal (2014) has given a possible analysis for such a change. The causative morpheme is -a- if it is followed by Progressive affix, but it changes to -ɪ- and [j] is also added if it is followed by the perfect affix. The [j] is inserted between the two vowels of causative morpheme and the perfect aspect marker to avoid the succession of the vowels. It can be assumed that the presence of -ɪ- triggers Glide Insertion, thus, inserting [j] in between the two successive vowels.

vii When the causative morpheme -a- is followed by the progressive aspect markers in present and past tense respectively, then, the -cʰ(t)- gets geminated and changes to -ccʰ(t). So, the presence of the vowel morpheme -a- triggers the Gemination prompting the change (Mondal, 2011) and (Mondal, 2014).

ix Such instances take place due to the disparity in the orthographic representation of the verbs. In general, when a non-finite and finite verb occurs together, the non-finite verb takes the emphatic ending. But, sometimes both the non-finite and finite verbs are clubbed together resulting in a single finite verb. In such situations, the emphatic marker gets inserted between the root and the suffix ending, resulting in a clitic.
A distinctive feature of these *semi-suffixes* is that, a majority of these end in the –e marker, which also indicates the process of adverbialization from adjectives or post-positions. The examples below substantiate the argument,

(i) ʃoubʰaggo-*krom-e* ʾamɪ okʰɑn-e ʾchɪlɑm
Fortunately I there be
‘Fortunately, I was there’.
(ii) 忸ɪ sized-*rup-e* ʾkɔtʰɑ ʾbolo
You clearly word say
‘You say clearly’.

Consider the following examples,

(i) ʾram-ϕ bʰɑt kʰɑt
Ram rice eat
‘Ram eats rice’.
(ii) ʾlok-e mɑtʰe kʰæle
People playground play
‘People play on the playground’.

In sentence (i), [ɾam] (Ram) is the subject of the sentence which is marked as null. On the other hand, in (ii), [lok] (people) indicates the subject of the sentence which gets inflected with the nominative case marker –e. Hence, from these examples, it can be derived that –e and null markers represent the nominative case in the nominal system of Bangla.

(i) ʾhe ʾtɑr bʰɑu-*re* ækkʰɑn ʾbɔl ʾdɪʃe
He his brother-DAT a ball give
‘He gave a ball to his brother’.
We can see that the sentence is in Sylheti dialect of Bangla where the dative marker is –re which gets inflected to the indirect object [bʰɑ] (brother).

The examples below exemplify that Bangla classifier markers getting inflected to adjectives, pronouns, quantifiers and numerals.

(i) eɪ meje-ta bʰalo
   This girl-CL good
   ‘This girl is good’.
(ii) amɑɪ æk-ta pen dɑo
    Me one-CL pen give
    ‘Give me a pen’.
(iii) fe ɔnek-ta jɔl kʰelo
     He lot-CL water drink
     ‘He drank a lot of water’
(iv) oɪ nl-ta oke ɗe
    That blue-CL him give
    ‘Give him the blue one’.
(v) e-ta amɑr gɑrɪ
    This-CL my car
    ‘This is my car’.

The sentences above show that the classifier -ta gets affixed to nouns, adjectives and so on. [meje] (girl) (i) is a noun, whereas [æk] (one) (ii) denotes a numeral. In (iii), [ɔnek] (lot) is a quantifier, [nl] (blue) (iv) specifies an adjective and finally, [e] (this) (v) indicates a pronoun. Therefore, it has been proved that, apart from nouns, a classifier can also get inflected to adjectives, pronouns, quantifiers and numerals.

Biswa (2012) states that the pre-nominal occurrence of -ta leads to indefiniteness, whereas the definiteness is realized through NP fronting as in the following examples,
In the above sentences, the -ta is affixed to the numeral [ʃɑt̯] (seven). In (i), the numeral and the classifier precede the noun, thereby resulting in indefiniteness. However, in (ii) the numeral and the classifier occur after the noun, leading to definiteness. From the above analysis, it can be assumed that the sequence of the occurrence of the classifier -ta with respect to the head noun determines the definiteness or indefiniteness of the classifier.

Moreover, -ta can get attached to both animate and inanimate nouns (Biswas, 2012). The sentences below show that -ta can get affixed to both animate and inanimate nouns,

(iii) amí meje-ta-ke dɑklam

I girl-CL-ACC call

‘I called the girl’.

(iv) rɑʃt̯ɑ-ta fab̩hɑne par kɔro

Road-CL carefully cross do

‘Cross the road carefully’.

We can see, while the noun [meje] (girl) in (iii), is an animate noun, the noun [rɑʃt̯ɑ] (road) (iv) is an inanimate noun and both of them get affixed to -ta.

The following examples elucidate the presence of honorificity with the usage of -ta,
(i) **lok-**tɑ
daklo
Man-CL call
‘The man called’.

(ii) **lok-**tɪ
daklo
Man-CL call
‘The man called’.

(iii) **lok-**tɑ
(daklen
Man-CL call
‘The man called (hon)’.

The boldfaced suffixes specify the classifier -tɑ and its allomorph -tɪ occurring in non-honorific and honorific speech. (i) shows that, the noun [lok] (man) gets affixed to -tɑ in the informal speech, thereby exhibiting non-honorificity. On the other hand, (ii) and (iii) signify that -tɪ gets attached to the head noun in both non-honorific and honorific speech respectively.

 However, –to can occur with a noun and a pronoun, only in a few dialects in Bangla.

(i) **a**mɪ ɖu-**to** boī kɪnλɔm
I two-CL book buy
‘I bought two books’.

(ii) *okʰane ʈɪn-**to** cʰele ʔɛchê
There three-CL boy be
‘There are three boys’.

(iii) okʰane ʈɪn-**te** cʰele ʔɛchê
There three-CL boy be
‘There are three boys’.

(iv) ʃe ʃe ɾ**tj**e boī enecʰe
He four-CL book bring
‘He has bought four books’
The boldfaced suffixes in the above sentences are the –to and –te markers respectively.

From the examples, (i) and (ii), it can be seen that, the –to marker gets attached to only the numeral [ḍu] (two). So, the sentence becomes ungrammatical when it gets added to any other numeral apart from two, as in (ii). In sentences (iii) and (iv), the –te marker is seen to get affixed to numerals like [ṭn] (three) and [cər] (four) respectively.

Look at the following sentences,

(i) t̯ɪn- Jon lok boʃlo
   Three-CL man sit
   ‘Three men sat’.

(ii) meje t̯ɪn- Jon skul gælo
    Girl three-CL school go
    ‘Three girls went to school’.

(iii) ɔnek- Jon gəumoccʰe
    Many-CL sleep
    ‘Many people are sleeping’.

(iv) æk- Jon boʃlo
    One-CL sit
    ‘One person sat’.

(v) lok- Jon kɔ预料 bolcʰe
    People-CL word talk
    ‘People are talking’.

The first two sentences, i.e., (i) and (ii) suggest the indefinite and definite expression of the numeral-classifier combination of the classifier -Jon, preceded and followed by the noun respectively. In (i), the noun [lok] (man) occurs after the numeral-classifier combination resulting in an indefinite expression, whereas in (ii), the noun [meje] (girl) is followed by the human classifier -Jon and the numeral [ṭn] (three), to create a definite
expression. Example (iii) specifies that -jon can be used with quantifiers like, [ɔnɛk] (many). However, both the sentences (iii) and (iv) also indicate that this classifier can refer to both singular and plural form. While the quantifier [ɔnɛk] (many) is plural in number, the numeral [æk] (one) denotes singularity. Finally, (v) shows that -jon gets affixed to the noun [lok] (people) and denote plurality.

xix The set of data below compare the classifier –gulo with its allomorph -gulï.

(i) ʃe cʰobɪ-gulumi bʰalo ekečʰe
He portrait-CL nice draw
‘He has drawn nice portraits’.

(ii) e-gulumi bʰalo cʰobɪ
This-CL good portrait
‘These are good portraits’.

(iii) kʰarap-gulumi keo nebe na
Bad-CL nobody take not
‘Nobody will take the bad ones’.

(iv) meje-gulumi kʰub bʰalo
Girl-CL very good,
‘The girls are very good’.

(v) *țin-gulumi cʰobɪ bʰalo
Three-CL pictures good.
‘Three pictures are good’.

(vi) amɪ ɔnek-gulumi bʰalo cʰobɪ ɖekʰlam
I many-CL good movie see
‘I saw many good movies’.

The above mentioned examples compare with the sentences with –gulo marker to show that both the classifiers, –gulo and -gulï share similar features.
(i) *₂d̯u-ra baccã kʰečʰe
   Two-CL child play
   ‘Two kids are playing’.

(ii) *moner d̯ukkʰo-ra cole ġælo
    Heart sadness-CL go be
    ‘Heart’s sadness went away’.

Both the sentences mentioned above are ungrammatical. In (i), the -ra is inflected to the numeral [₂d̯u] (two) and in (ii), it is attached to an abstract and inanimate noun [d̯ukkʰo] (sadness). This goes on to prove that -ra does not get affixed to either a numeral or an abstract and inanimate noun.

The 2\textsuperscript{nd} person pronoun encodes a three-way distinction of honorificity, that is, honorific, familiar and intimate, while the 3\textsuperscript{rd} person pronoun encodes only the two-way distinction of honorificity and non-honorificity. In case of proximity, only the 3\textsuperscript{rd} person personal pronouns differentiate three-way distinction of proximity, near, far, and unmarked.

According to David (2015), the honorific pronouns are used when talking to people with superior status, like parents, teachers, elders and so on. The familiar pronouns are used in conversations between people of equal status like, spouses, friends etc. finally, the intimate form of pronouns are used to talk to children, to younger people, siblings, and people of lower status like servants.