Introduction

1 Preview

This chapter is intended as a brief overview of the larger context in which the present work is embedded. Thus, Sections 2-4 take a brief look at the language and its variants as well as earlier studies on the language. Certain procedural issues are considered next in Sections 5-7 where the aims and theoretical orientation, sources of data and problems encountered during the analysis discussed. Finally, Section 8 attempts to place the grammatical units of Assamese in the context of the five-term grammatical hierarchy of sentence, clause, phrase, word and morpheme.

2. The Assamese language

Assamese is a language of eastern India specifically spoken in the present state of Assam and partially also in the nearby states of West Bengal, Arunachal pradesh, Meghalaya and Nagaland. According to the 1991 census report, the number of speakers of the language is 1,29,50,088. However, it is spoken as a second language by a considerable number of speakers of Tibeto Burman languages like Bodo, Mising and Karbi. Moreover, it has traditionally served as the lingua franca or pidgin in the neighbouring states of Nagaland and Arunachal pradesh.

Scholars are not unanimous in tracing the origin of Assamese language, though there is general consensus that it is an Aryan language with a high degree of influence of Tibeto Burman and Austric languages. George Grierson, Nathan Brown, Suniti Kumar Chatterjee and Banikanta Kakati have attempted to trace its origin from the Magadhi Prakrit of the MIA stage. The assumption of these scholars is based on the similarities of Assamese with Magadhi as well as with Bengali and Oriya, two NIA languages originating from Magadhi. On the other hand, Debananda Bharali, Dimbeswar Neog. and others have held the view that it emerged from a variety of MIA language of eastern Assam that is called Kamrupi Prakrit which the MIA grammarians did not mention.
third view, that of Kaliram Medhi also rules out the MIA origin of Assamese, but proposes influences from the eastern side of India. Serious diachronic research on this area is awaited.

The literary history of the language is available from last part of thirteenth century A.D, the first instance being *The Prāhrad Sorit* by Ram Saraswati. However the maturity of language and the poetic merit of the work suggests that there was a long literary tradition before it. Banikanta Kakati has extended the beginnings of the literary tradition of the language up to the time of the Charya literature of the Buddhist monks, which is supposed to have flourished between the 8th and 12th centuries A.D. The Charyas carry linguistic features that are predominantly associated with Assamese, though there are some features bearing similarity with Bengali and Oriya as well (Kakati 1972: 11).

3 **Dialectal variation**

Linguist usually divided Assamese into three dialect groups. The variety spoken from its eastern boundary to the undivided Nagaon district is called 'Eastern Assamese', the standard variety of the present time. The variety spoken in the undivided Kamrup district i.e. from Guwahati to Barpeta is called 'Kamrupi', and from there to the western boundary, the variety is called 'Goalporiya' or 'Rajbansi dialect'. 'Kamrupi' and 'Goalporiya' have further sub-dialectal variants. The variety chosen for this analysis is the standard variety.

4 **Earlier studies**

The tradition of linguistic studies of Assamese had relatively recent beginnings in comparison to its literary tradition. While the language attained its literary peak during the period between the fourteenth to sixteenth century under Madhab Kandali and Sankaradeva, little evidence is found of grammatical studies on it. It was only with the
work of William Robinson, namely, his *Grammar of the Assamese language*, in 1839, that the tradition of linguistic scholarship on Assamese began. Other works by foreign and indigenous scholars followed, the most prominent being Banikanta Kakati's doctoral dissertation, *Assamese: its formation and development* (1942). Kakati's work continues to be important for the various synchronic and diachronic concerns addressed in it.

The second half of the twentieth century saw several scholars notably Gollock Chandra Goswami, Upendranath Goswami and Rohini Kumar Mahanta making notable contributions to our understanding of the Assamese language. No attempt has been made here to assess the merit of the various linguistic works on Assamese: much of these address different issues of the language in rather global terms. It is only since recently that areas of specific concern have begun to be taken up for research.

**5. Aims of the study and theoretical framework**

The goal of the present study is to set out the salient features of nouns and nominalisations in Assamese in microlinguistic terms. Though it is one of the major Indo-Aryan languages, not much is known about the behaviour of Assamese nouns and the structures and processes in which they participate. The present effort is intended to fill that need. It is a language-oriented approach that seeks to contribute to our understanding of the Assamese language with regard to the chosen area. With this aim in view, abstract theorising, a valid intrinsic goal in itself, has been avoided and the analyses and conclusions presented proceed directly from the data.

The present work is also intended to fulfil part of the need for a typological comparison involving Assamese and one or more of the numerous non Indo-Aryan languages of the area. While the influence of these languages on Assamese as a consequence of contiguous existence has been investigated to some extent at the levels of phonology and
morphology, there a need for similar efforts at the level of syntax. The present work holds potential for a fruitful comparison with the neighbouring Tibeto-Burman languages insofar as the behaviour of determiners/classifiers is concerned.

Finally, the present work is intended to contribute to pedagogic efforts involving the teaching of Assamese as a mother tongue subject at various levels of the educational establishment. Due to the paucity of adequate materials, the curricula of Assamese at the undergraduate and postgraduate levels continue to have a strong historical bias to the near exclusion of synchronic concerns as far as Assamese syntax is concerned. It is anticipated that the present work will be of use as resource material on Assamese syntax.

It has been stated earlier that the emphasis of this work is exclusively language-centric. No attempt has been made to consciously vindicate the assumptions and postulations of one of the contemporary paradigms in the exploration of a specific aspect of Assamese. A major problem encountered in choosing a theoretical paradigm of the purpose of a language-centric analysis such as the present one is the following: there is a tendency to concentrate more on taking forward the hypotheses specific to the paradigms and on formalising devices, than to allow for the flexibility of expressing the genius and resources of individual languages.

For this reason, the researcher decided to opt for the theoretical framework envisaged in the contemporary scholarly British Grammatical Tradition initiated by Randolph Quirk, Sidney Greenbaum, Geoffrey Leech and Jan Svartvik in their Comprehensive Grammar of the English Language (1985) preceded by earlier and smaller versions. Needless to say, departures and modifications have been made as found appropriate and as necessitated by the area of the language being investigated. One area where such departure has been made is that of compounding, where certain insights from the
transformation lists have been drawn upon. Insights from the Indian grammatical tradition have also been incorporated in the body of the work.

6 Sources of the data

Since the present investigator is not a native speaker of the standard variety of Assamese, he has collected data from the following persons:

Mr. and Mrs. Arif Ahmed, Mr and Mrs. Ghana Gogoi, Simanta Kalita and Mr. and Mrs. M. Barmudoi,

While eliciting data, every attempt has been made to ensure homogeneity, both in terms of the linguistic background of the speakers as well as their educational and economic background.

7. Issues faced during the analysis

While carrying out the present investigation of nouns and processes of nominalisation in Assamese, certain issues of academic interest and sociolinguistic implication have come up.

One such issue is that of formal and informal varieties of the language. The formal-informal distinction correlates significantly with the phonology, morphology and syntax of the language. The formal varieties are heavily influenced by Sanskrit at all these levels. The formal-informal distinction is evident in usages such as the following:

1. I one-det film do-fut.1 where music director be-fut.3 Bhupen Hazorika instrument play-fut.3 Prabhat Sarma-gen party-nom
dialogue and direction Bhabendra Nath Saikia-gen
and acting do-fut.3 Nipan Goswami Biju Phukan etc-nom.
I shall make a film, where the music director will be Bhupen Hazarika, the instruments will be played there by Prabhat Sarma and his party, the dialogue writing and direction of the film will be of Bhabendra Nath Saikia, and Nipon Goswami, Bijuphukan etc will act.'

2. mō ekham kāt hasobi korim, zot ṣhongit porisaluk hobo bhupen hazorika, bādpēcēntc bōzābō prōbbhāt ṣārmar dēle ṣānnap aru nirdehōna bhābhendrē nāttb hōikijar aru obbēinc korīb nōpōn gōswami, bīzù pōukōn adije.

In (1), there are eight English words, all of which are replaced in (2) by Sanskrit words. In addition to such vocabulary-related differences, there are grammatical differences as well.

However, it is not possible to draw a clear line of difference between the formal and non-formal styles of the language consequently, no choice has been made between either of them. The structural complexities of the original language brought in by such items have been sought to be avoided in this way. For example the English words batting and volley ball as in (3) and (4) and the Sanskrit words pratiyogi and pītāmbara as in (5) and (6) are structurally complex units.

3. azi ḍōsinēr beting saisane?
   today sachin.gen batting see-perf.2fam?
   'Have you seen Sachin’s batting today?'

4. tumi bōlibol kāclane?
   you volleyball play-pr.2fam?
   'Do you play volley ball ?'

5. protizogi-ĥikol-ēk ah-ibōlōi kōa
   competitors-det-acc come-inf. tell-imp.2.fam
Tell the competitors to come.'

6. pitambara srikrisn-r an eta nam.

pitambara srikrisn-gen an eta nam.

'Pitambara is the another name of srikrishna.'

The structures of the words are as follows:

batting< bat+ing (affixation)

voleyball< voley ball(compounding)

pratiyogi< prati+yogi(prefixation)

pitambara< pita+ambara(compounding)

Such words are regarded as compositionally simple in Assamese: their use is regarded as natural in the language since they have not natural ready and native Assamese substitutes or counterparts. The strategy adopted during the present investigation has been to avoid data reflecting such usage.

A second issue faced during the analysis has been that of glossing the Assamese sentences. In Assamese morphology, many inflectional affixes are polysemous in nature. For example, as it will be evident in (1.2.3), the so called nominative marker –e/i is used to indicate both the subject as well as the temporal adverbial as in (7):

7. cha robibar-e ram-e gan ga-b-3

coming Sunday-loc Ram-nom song sing-fut-3

'Ram will sing on next Sunday.'

In case of verbal inflection, the problem is more acute as they have no traditional names.
8 Grammatical Units in Assamese

For the grammatical analysis of a language, the sentence and the word have been regarded as the more basic units in traditional grammar compared to the two other intermediate units – the clause and the sentence. The fifth and smallest unit of grammar, namely the morpheme was added to the hierarchy of units in the western tradition of linguistic studies. Since the orientation of this study is essentially structural, these five units of grammatical analysis generally recognized within that paradigm will be briefly considered here in the context of Assamese.

8-A The sentence

We can begin with John Lyons’ characterization of the sentence as “a grammatical unit between the constituent parts of which distributional limitations and dependencies can be established, but which can itself be put into no distributional class” (1979: 172) and consider it in the context of the following examples:

10. tumi ḍazarloī zaba ḍhōi?
   ‘Are you going to market?’

11. ḍānman roba
   ‘Please wait for a while.’

12. mōjco tcamar bgot zaco
   ‘I am also going with you.’

(10-12) are three different units within an Assamese utterance, which are distributionally independent from each other and therefore different sentences.

The limitations of this definition in the case of pronouns (anaphoric reference) etc is illustrated by (13), which is not absolutely independent:

13. hi māk kitap ekhōn dile
   ‘He gave me a book.’
The noun phrase replaced by pronoun /h/ in the sentence is traceable only from the preceding sentence(s). Keeping such possibilities in mind, Robins (1980:146) observes that “a sentence is by definition grammatically complete; it may, therefore, be preceded and followed by indefinite pause or silence.”

A sentence has phonological characteristics too. A grammatically complete sentence must have the characteristic intonation. Thus, four different sentences (15a-d) can be abstracted from (14), using four characteristic sentence-final intonations in place of pause pitches (orthographically represented by the comma).

14. moi ga dhulco, bhat khalco, basôn dhułko aru toponio marilko
   ‘I have had a bath, eaten my meal, washed the utensils, and also slept a little.’
15. a) moi ga dhułko
   ‘I have had a bath.’
   b) moi bhat khalko
   ‘I have eaten my meal.’
   c) moi basôn dhułko
   ‘I have washed the utensils.’
   d) moi dhułto toponio marilko
   ‘I also slept a little.’

There are four major sentence-final intonations in Assamese, and they are occur in the four following sentence types - a) declarative, b) interrogative, c) imperative and d) exclamative, as in (6-9), respectively-

16. moi bhat khaω
   ‘I eat rice.’
17. tumi bhat khaω ne?
   ‘Do you eat rice?’
18. toi bhat kha
   ‘(You) have your meal.’
19. toi bhat bọto mod khaω
'He drinks even with his meal!'

Sentence final intonations of interrogative sentences are also different according to their different structures. For instance, the intonations of interrogative sentence [20a—d] are different from each other:

20.a)  
\textit{tumi bhat khcoa?}  
'Do you eat rice?

b)  
\textit{tumi bhat khcoane?}  
'Do you eat rice?

c)  
\textit{tumi bhat khcoa ne nokhwa?}  
'Do you eat rice or not?

d)  
\textit{tumi ki bhat khcoa?}  
'What kind of rice do you eat?'

Exclamative sentences are of two types – exclamative-declarative and exclamative-interrogative, as in [21a-b]:

21 a)  
\textit{Iji bhatDr bgDt mDd khaj!}  
'He drinks with his meals!'

b)  
\textit{tumi mod khcoa?}  
'What, you drink too?'

Sentences are simple, complex or compound, depending on their structural complexity. Simple sentences are one-clause entities. Compound sentences are made up of more than one clause or simple sentence, which is functionally independent from other(s) i.e. they are grammatically coordinating. Complex sentences are made up of more than one sentence the grammatical relation between which are like that of a phrase to another phrase. Sentences (12-14) exemplify the three types respectively:

22. \textit{mDi ei kitapkhon porhisco}  
'I have read this book.'

23. \textit{mDi g\textsuperscript{b}lozlozlo zam aru deuta k\textsubscript{a}lozlozlo zab}
'I shall go home and father will go to college.'

24. tumi b̄at k̄ai ut̄i deutalo ei kap sah bōnaba

'After finishing your meal make a cup of tea for father.'

The relation between the two clauses in (24) is like the relation of an adverb phrase to the predicator of the sentence.

8-B The clause

As an intermediate unit, the clause differs from sentence in two features. Firstly, a sentence must have its characteristic sentence-final intonation; this is not compulsory for a clause. Secondly, in order to qualify as a sentence, the unit must have a verb fully inflected for the categories of tense, aspect, person and status. However, the verbs of clauses may or may not possess such categorial properties. Therefore, all sentences may be called clauses, but all clauses need not be sentences.

In terms of the nature of their verbs, clauses are either finite or non-finite. A finite clause is a clause containing a verb, which is inflected for required categories. On the other hand, clauses comprising verbs without categorial affixations are non-finite clauses. (15a-b) and (16a-b) respectively exemplify the finite and non-finite clauses in Assamese:

25. a) mōi adhunik kobitabor b̄alkoi buzī nepaw

'I don't understand modern poetry well.'

b) ram zāakali ijaloi ahisil

'Ram came here yesterday'

26. a) mōi ei kitap b̄on porhi

'After reading this book, I-----'

b) deutal b̄at k̄oa - (kahlīk b̄on)

'(The dish) used by father to have his meal'

In Assamese there are five different forms of non-finite verbs i.e. five suffixes to form non-finite verbs of non-finite clauses. They are -i, -a, -b, -h and -or. The non-finite forms
of the verbs $k^{h}$a ‘eat’ and mar ‘kill’, for example are $k^{h}$ai ‘eating’, $k^{h}$wa ‘eaten’, $k^{h}$ab ‘to
eat’, $k^{h}$al ‘if eaten’, $k^{h}$aw ‘on eating’ and mari ‘killing’, mari ‘killed’, marib ‘to kill’,
maril ‘if killed’, maraw ‘on killing’.

Modern grammars accord more importance to the clause unit than the sentence. Thus,
modern grammarians often use the term ‘clause’ instead of ‘simple sentence’ (Quirk et al
1985: 719). Moreover, even a complex sentence has been termed ‘superordinate clause’.
This usage reflects the insight that the subordinate clauses within a superordinate clause
are equivalent only to phrases or words realizing some clause element.

Clause elements and clause types:
A clause can consist of five different elements, namely - subject, object, verb, adverbial
and complement, abbreviated as S, O, V, A, C. Each of these elements is phrasally
realized by noun phrases, verb phrases, adjective phrases, adverb phrases and
postpositional phrases. Subject and object elements are realized by noun phrases, verb
elements by verb phrases, adverbial elements by adverb, noun, and postpositional phrases
and complement elements are realized by noun and adjective phrases. Sentences (27-41)
exemplify the element realization and basic clause types in Assamese:

27. [SV] ramcr deuktak (S) dhtukal (V)
   ‘Ram’s father has died.’
28. [SOV] ritue (S) nituk (O) bhal pai (V)
   ‘Ritu loves Nitu.’
29. [SOOV] tumi (S) mok (O) tomar kitapk (O) dibane (V)
   ‘Will you give me your book?’
30. [SC] mcr deuta (S) ezdn daktr (C)
    ‘My father is a doctor.’
31. [SCV] ram (S) czoddar rzza (C) asil (V)
    ‘Ram was the king of Ayodhya.’
From the above examples it is clear that in Assamese, the formation of a clause that is equivalent to a simple sentence is possible by excluding each of the elements (even verb) is possible. All these clauses types can be embedded within superordinate clauses (or complex sentences), which are structurally similar to these basic types. Chapter Four presents further discussion on clauses.

8-C The phrase

A phrase is a group of words the relation between which is not that of predicator-complement (Matthews 1981: 101); but of modifier-head or the like. It has been called a word like unit in comparison to the clause, the sentence like unit. A phrase is named
according to the word class of the headword, i.e. a noun phrase is headed by a noun, verb phrases are headed by verbs adjective phrases by adjectives and so on, e.g.

42. iman ıkınmani ıswaliznie ınor ıkget ıman ıkunjakoi ıkıh a pati ıkse

'This little girl is talking with me so nicely.'

iman ıkınmani ıswalizni- is a phrase headed by the noun ıswalizni, and therefore it is a noun phrase under which iman ıkınmani. Moreover, iman ıkınmani is another phrase headed by the adjective ıkınmani and therefore an adjective phrase. ınor ıkget is a postpositional phrase headed by the postposition ıkget and iman ıkıunjakoi is an adverb phrase headed by the adverb ıkunjakoi ıkıh a pati ıkse is a verb phrase headed by verb ıkıh a pat.

It was stated earlier that clause elements are realized by phrases. In addition to realization, the other two functions of phrases are modification and determination, These latter two are different from clause element realization in terms of the level of operation. Realizing a clause element is a clause level function of phrases, whereas modification and determination are phrase level functions. Adjective and adverb phrases function as modifiers in noun phrases and adjective phrases respectively. Noun phrases with the genitive -r inflection may function as determiners in larger noun phrases, e.g.

43. deutal məık ekıkın bər ıkunija kitap dibo

'Father will give me a very nice book.'

In sentence (43) the noun phrase ekıkın bər ıkunija kitap realizes the direct object element in the clause. bər ıkunija is another phrase within it headed by the adjective ıkunija and therefore an adjective phrase, which functions as a modifier.

44. bər kiskisija kola suli ıkıb al nẹpəw

'I don’t like jet black hair.'

Here the noun phrase bər kiskisija kola suli constitutes the modifier adjective phrase bər
The trees of that mountain are leafless.

In Assamese the last words of noun phrases may be inflected according to the clause elements they realize or the syntactic functions they perform within larger phrases. The last words of noun phrases may be inflected in the realization of subject, object and adverbial clause elements and while functioning as determiners. Noun phrases in complement (verbal) functions have no inflection for the same. Verb phrases, though inflected for the categories of tense, mood, person etc. have no inflection for the verb element realization. Adjective phrases, adverb phrases and postpositional phrases too have no inflection for their own functions in clauses. However, due to the ellipsis of the head noun, the modifying adjective (word or phrase) or the determiner may be inflected. In such instances of ellipsis, the genitive determiners have to carry double inflection, one for its own function and the other for its head’s function as a head. This is illustrated in (35) and (36):

35. ritur koinazonije sakori kore
   ‘Ritu’s bride is a serviceman.’

36. ziturzonie obhinoj kore
   ‘Jitu’s bride is an actress.’

The head koina has been ellipsed in the noun phrase in subject slot in (36) and therefore its genitive determiner zitur has to be inflected on behalf of its head.

An important point to be noted here is that, though higher in rank than a phrase, and a unit of one or more than one interrelated phrases, a clause may be itself a part or whole of a phrase. Moreover, it was stated above that a phrase may be a constituent of another,
larger phrase. A phrase, therefore, may comprise three units under it - the word, the phrase and the clause. All these three units under a phrase are, however, of equal grammatical status i.e. of a status of word, because, such clauses and phrases are substitutable for words. For instance,

\[ \text{ mdi zoakali dekha manuhzDn} \]

' \text{the man I saw yesterday}'

is a noun phrase in which the non-finite clause \text{mdi zoakali dekha} is a modifier and therefore substitutable by an adjective word. By using clauses as constituents of phrases, a phrase -- and consequently a clause or sentence -- may run to indefinite length.

8-D The word

The word is the most important unit of grammatical analysis both in traditional and modern grammar. As was stated earlier, the word is basic and smallest unit in traditional grammar. Even though the morpheme has emerged as the smallest grammatical unit in modern linguistics, the word still continues to be a basic unit of language analysis.

Bloomfield defines the word, as 'a minimal free form' i.e. a word has the potential to be a sentence on its own. If this definition is followed for Assamese, two kinds of words will have to be denied the status of word in the language. First, the postpositions - \text{dvara 'by'}, \text{pDra 'from'}, \text{karDne 'for'}, \text{babe 'for'}, \text{hoke 'on behalf of'}, \text{obihDne 'without'} etc. which cannot be used without a nominal word in the genitive form preceding it. For example, in (48) \text{karDne} and \text{dvara} cannot be used in the clause without the preceding nouns in genitive forms:

\[ 48. \text{ gDnDDntrD hoi manutor karDne, manutor dwara, manutor hajDn} \]

'Democracy is the rule of the people, by the people and for the people.'

Secondly, the components of conjunct verbs will not get the status of words. A conjunct verb in Assamese comprises two or more words the first member of which is a
noun, adjective or verb and the second member a verb. In the process of combination, both or all the words lose their own individual meanings and together refer to some action or so, as in (49) and (50):

49. \textit{uljai oniruddh} \textit{k b\textsuperscript{a} al paj}

‘Usha loves Aniruddha.’

50. \textit{rame horir bg\textsuperscript{a} t k\textsuperscript{a} t\textsuperscript{a} a patise}

‘Ram is speaking to Hari.’

\textit{b\textsuperscript{a} al pa} in (49) and \textit{k\textsuperscript{a} t\textsuperscript{a} a pat} in (50) are conjunct verbs.

While recognizing a word, the following four criteria are used by linguists:

8-D-i Criteria of word

The four criteria for word recognition are- (a) potential pause, (b) Internal stability, (c) Positional mobility and (d) Phonological co-relation (Lyons 1979: 199)

(a) Potential pause:

Under this criterion, the constituents of Assamese conjunct verbs (e.g. \textit{b\textsuperscript{a} al pa} ‘love’, \textit{bat sa} ‘wait’ etc.) are different words, they are not even compounds. Similarly, \textit{p\textsuperscript{\textcircled{c}ra} from’, \textit{dwara} ‘by’ \textit{babe} ‘for’ etc. in e.g. \textit{rubir p\textsuperscript{\textcircled{c}ra} from Rubi’, \textit{horir dwara} ‘by Hari’ \textit{mcor babe} ‘for me’, are different words.

(b) Internal stability:

A word is an uninterruptible unit i.e. no element, word or phrase can be inserted within it, nor any element within it interchanged. For example (51) cannot be rearranged as (51a):

51. \textit{manuh\textcircled{z}n e kamt\textcircled{w} m korib\textcircled{b}}

man the work the do will

‘The man will perform the work.’

51. a) \textit{\textcircled{z}nemanuh tokam ib\textcircled{k}r}

Under this criterion, \textit{manuh\textcircled{z}n} ‘the man’ is one word in Assamese, but \textit{manuh keiz\textcircled{z}n} ‘the men’ are two. We cannot insert any unit greater than the morpheme within the
word *manuh* but it is possible to do so between *manu* and *keizn*, e.g. *manuh b'ale*

*keizn* ‘a good number of people’ *manuh bohut keizn* ‘quite a few people’ etc.

(c) Positional mobility:

A word is though internally stable, positionally it is mobile. This criterion may be illustrated with reference to the following Assamese sentence in (52):

52. manuh-zDn-e saul-bosta bozar-3loi ni-b-ɔ

man-Det-nom rice-Det market-dat take-fut-3

1 2 3 4 5 6 7 8 9 10

'The man will take the bag of rice to the market.'

The sentence above comprises ten morphemes occurring in a particular order. The following permutations are possible for this sentence which will yield further acceptable variants

1 2 3 4 5 6 7 8 9 10

1 2 3 4 5 8 9 10 6 7

4 5 1 2 3 6 7 8 9 10

6 7 1 2 3 4 5 8 9 10

The following sequence, however, is not permissible:

*3 2 1 4 3 7 6 9 8 10 etc.*

This criterion is useful to differentiate the words from the morphemes. *manuh*, *zDn* and *-e* are not different words but different morphemes, and together they are one word. Even *saul* ‘rice’ and *bosta* ‘bag’, though free forms in other environments, are here one single compound word.

(d) Phonological co-relation:

There are some phonological features in some languages which, can be used as criteria for word status. (Robins-1980: 152) Kakati has observed that word stress falls on
the penultimate syllable in words comprising more than two syllables (1972: 116).

Goswami (1982: 131) maintains that every word in Assamese has specific primary stress, which is different from secondary stress of non-word elements. Sentences (53-4) illustrate:

53. ḥi ḡbɔ-r-ɔloi  go-1
   he home-dat go-past
   ‘He has gone home.’

54. ḥi-gbɔ-r-ɔloi  go-1
   next house-dat go-past
   ‘(He/Shie/They) has/have gone to the next house.’

In (53), ﳍ ‘he’ and ḡbɔ- ‘to home’ are two words because both of them contain primary stress. However, in (54) the same elements form one word due to the single primary stress for both the elements. It should be noted here that though grammatical and phonological criteria have co-relations, in the cases of some words they can be in conflict. In such cases, grammatical criteria are regarded to be decisive.

8-D-ii Ambiguity of the term ‘word’

The use of the term word in phonological/orthographic, grammatical and lexical senses is well known. Of these, phonological words represent grammatical words. However, the correspondence between phonological and grammatical words may not be one to one in all cases. In Assamese, for example, one phonological word may represent more than one grammatical word, as in (55a-b) and (56a-b)

55. a) tumi  bɔ at  kɔ-wa
    you rice eat-2.fam
    ‘Have your meal (please).’

   b) bɔ at  kɔ-wa mez-kɔn sapha  kɔ-r-a
    rice take-part table-det clean do-2.fam
    ‘(please), clean the dining table.’
56. a) tumi bagh dekh-is-a-ne?
   you tiger see-impv2.fam-interg?
   'Have you seen a tiger?'

b) bagh zöngöl-öt thak-e
   tiger forest-loc live-3
   'Tigers live in the forest.'

Phonologically *khōwa* in (55a) and (55b) are the same word. But in grammatical

...
verbs. Others are either invariable lexical words or non-lexical or function words. The nature of variation of words (or lexemes) is mostly agglutinating, though their inflecting nature is also occasionally apparent. Each class of words has its own characteristics of variation or inflection: some (e.g. nouns) are inflected for their own characteristic functions, while others (e.g. adjectives) on behalf of others. Even words of same class may be inflected differently and for different categories. In short, with so many differences, the variation of Assamese words is a most complex process. Some of this variation will be discussed in Chapter One (1.1.4) and others in Chapter Four, while yet others will remain untouched in this study.

Word classes or parts of speech:

Robins observes that “in the grammatical analysis of languages words are assigned to word classes on the formal basis of syntactic behavior, supplemented and reinforced by differences of morphological paradigms, so that every word in a language is a member of word class”(1980: 173). The words (or units that constitute the structures realizing sentence elements) in Assamese can be classified as follows:


vi. Determiner – ei ‘this’, ḫāl ‘all’, sariz ‘four’, ziz ‘which’


viii. Interjection – bah ‘hurrah’, sīh (how disgusting!), dhetteri (oh no!)

x. Correlators – ziman...jiman ‘as much...so much’, zenekoi...tenekoi ‘which way...that way’.

The examples are presented in their ‘dictionary form’ and not as they often appear in sentences when they function as constituents of phrases i.e. they are lexemes not grammatical words. The first four classes are open class items. Their numbers in the language are almost innumerable. The rest are closed class items. One can easily list out any of these classes: the creation of new items in these classes is very rare.

**8-E The morpheme**

The morpheme is the smallest unit of grammatical analysis in synthetic languages. The word, a unit of the next higher rank, is composed of morphemes. Alternatively, we can say that words are further analyzable into smaller units, which are called morphemes. The morpheme is not further analyzable into fragments having expression and content structure; i.e. morphemes are minimal meaningful units. Meaningful, of course, does not mean only lexical meaning but functional and grammatical meaning also. On the other hand, minimum meaningfulness does not imply having only a single meaning but the minimum forms that can express meaning. Such meaningful fragments within words are recurring elements in a language, and like word classes, phrase classes etc. these minimal units are also classifiable in terms of their distribution within words, e.g.

57. ram-e hori-k mar-il-e
   Ram-nom Hari-acc beat-past-3
   ‘Ram beat Hari.’

The sentence in (57) comprises three words but seven morphemes. Out of these, *ram*, *hori* and *mar* are lexically meaningful and the rest are functionally meaningful. These meanings may be called subjective case, objective case, past tense and third person
respectively. Each of these morphemes is a recurrent element in Assamese. Morphemes with functional meanings are more frequent in use than morphemes with lexical meanings.

Some morphemes in Assamese exhibit coalescence. For example, the morpheme /-zn/ in the word manuzn is a fusion of five different grammatical or functional meanings such as –definite reference, singular number, human class, superior status and masculine gender. The following four words are the same in their lexical meaning and definite reference, but differ from each other in number, class, status and gender:

- manuhzoni : man – definite.singular.human.inferior.feminine

Two aspects of Assamese morphemes merit discussion: their variations in phonemic form or morphophonemics and their classification. The same morpheme may exhibit different phonemic shapes in different environments. For instance, /hC/ ‘be’ is a morpheme which occurs in three different phonological shapes-/hC/, /hC/ and /hC/, as in

58. a) moi ramr deutak hC-ω
      ‘I am Ram’s father.’

b) tumi horir ki hω-wa
      ‘What is your relation with Hari?’

b) moi azir hC̄b̄ har hC̄b̄ apoti hω-m
      ‘I shall preside over today’s meeting.’

Each of such different phonemic variations of a morpheme is called an allomorph.

There are very few morphemes in the language that does not exhibit this kind of phonemic variation. There are two kinds of allomorphy in a language- phonologically
conditioned and morphologically conditioned allomorphy. The first type is numerous in Assamese and will be discussed in Chapter Five. The second type is relatively fewer. Some grammarians regard determinative affixes as morphologically conditioned allomorphs of same morpheme. If we adhere to the definition that a morpheme is a minimum meaningful unit, then each of these forms is a different morpheme. As shown above, they have one or more meanings in common, but different in at least one. However, the degree of difference between these classes of affixes is different from the degree of difference between other classes of affixes. Feminine suffixes are, on the other hand, purely allomorphs of same morpheme and are subject to both conditions. Morphologically they are of three different shapes—-\( i \), -\( ni \) and -\( ri \). -\( ni \) has four further variants that are phonologically conditioned. These are /-ni/, /-oni/, /-ini/ and /-uni/ (e.g. deka-\( ni \), kumar-\( oni \), bagh-\( ini \), sur-\( uni \)). -\( ri \) also has four such allomorphs—/-ri/, /-eri/, /-uri/ and /-ori/ (e.g. deka-\( ri \), dut-\( eri \), beja-\( uri \), kola-\( ori \))

Classification of morphemes:

Morphemes are classifiable in various dimensions. In terms of distribution, they are of two types—free and bound. A free morpheme has the potential to be a word in itself. Bound morphemes are bound to appear with at least one other morpheme. In (59) and (60), all the words comprise only free morphemes i.e. each of them is amorpheme as well as a word:

59. rakes bɔr d́ustɔ lora
   Rakesh very naughty boy
   ‘Rakesh is a very naughty boy.’

60. tɔi bʰat kʰa
   you rice eat
   ‘Have your meal.’

But in (61), all the words are bimorphemic and in (62) they are polymorphemic:
61. Ram-nom book-detl hari-acc give-3
'Ram gives the books to Hari'

62. mother-3(n)-nom boy-det-acc. stick-det-ins beat-impv-3
'The mother is beating the boy with the stick.'

The first morphemes in each of the words in (61) and (62) is a free morphemes and the rest are bound morphemes.

Typically, free morphemes are used as roots and bound morphemes as affixes; but a few bound morphemes are found to function as roots also. One of the finest examples of such a bound morpheme is -kei, the meaning of which is nearer to English 'some'. In (63) it is used as affix and in (64) as root:

63. this boy-det-det very intelligent.
'Vese boys are very intelligent.'

64. You-det how many there-dat. go-fut-2.fam?
'How many of you will go there?'

Root morphemes are potentially unlimited in number in Assamese.

Affixes are bound morphemes. They are limited in number but very frequent in use. One can easily exhaustively list all the affixes of the language.

It should be noted here that the construction of word need not be just root + affix; but root + root or root + root + affix as well. Words in such combinations are called compound words. These are again of two types- in one, two different roots are combined together, e.g. d' an-k' eti 'paddy crops', kahi-bati 'utensils', sau-l-bosta 'the sack of rice' etc. In the other type, the same root is reduplicated within the word, e.g. o k - o k 'knocking'. (1.6).

In terms of their positions, affixes in Assamese are of two types- prefixes and...
suffixes. Excluding prefixes of Sanskrit origin (upasargas), the number of prefixes is very few in Assamese. These are mostly used in the formation of negative and opposite words. All other affixes of Assamese are suffixes.

According to their function, affixes in Assamese are either derivational or lexical and inflectional. Lexical affixes derive new lexical items (lexeme) in the language. For instance, -oni, -orija and -ija constitute the new lexical items dh'anon' ‘paddy field’, navorija ‘boat man’ and lunija ‘salty’ respectively from the roots d'an ‘paddy’, na ‘boat’ and lon ‘salt’. Lexical affixes may be class changing or class maintaining. The first two of the above affixes are class maintaining. They derive noun lexemes from noun roots. The third one derives adjective lexemes from noun roots and therefore is a class-maintaining affix. The directions of class change are as follows:

Noun → Adjective, Verb [e.g. k'hɔŋ ‘anger’ + al = k'hɔnal ‘angry’, bol ‘colour’ + a = bol ‘to paint’]

Adjective → Noun, Verb, Adverb [e.g. murkʰɔ ‘foolish’ + ami = murkʰa ‘foolishness’, patɔl ‘light’ + a = patɔla ‘to lighten’, patɔl + koi = patɔlkoi ‘lightly’]

Verb → Noun, Adjective [e.g. nas ‘to dance’ + oni = nasoni ‘dancer’, zira ‘to take rest’ + onija = zironija ‘resting’]

Inflectional affixes form different grammatical words from lexemes. Case affixes, determinative affixes, gender affixes and a set of person affixes (only with a small class) are used with nouns (and also with pronouns and adjectives). Tense affixes, mood affixes, person affixes and non-finite formatives are used with verb words.

The derivational and inflectional division of affixes presents practical problems in Assamese. There are some borderline cases, which cannot be resolved in favour of one class or the other. Assamese feminine suffixes are the finest examples of such cases. bagh ‘tiger’ and baghini ‘tigress’ will be included in a dictionary as different word or lexical
items but 'kolita' and 'kolitan' will not. The genitive marker -r is another such affix by which a determiner word is formed out of a noun, pronoun or adjective but it cannot be regarded as derivational affix, or nobody has done so far.

Inflectional affixes in Assamese can be classified in terms of their functions into three classes—feature indicating, concord maintaining and element realizing. Singular and plural determiners (to, zon, jbn/on, bilak, hzt) with nominal words, and tense and aspect affixes (-is, -il, -im) with verb words indicate grammatical features of the respective words. Personal affixes of a small class of nouns (-r, -ra, -k), personal affixes with verb words (-co, -c, -a, and -i/e) and feminine suffixes are concord-maintaining affixes, and case affixes (-e/i, -k), are element realizing affixes. Element realizers are again subdivided into clause element realizers and phrase element realizers. The case affixes—nominative -e, accusative -k, instrumental -re, dative -loi and locative -t normally realize clause elements. But genitive -r normally realizes phrase elements.