



***Languages and Peoples of the Eastern Himalayan Region (LPEHR) comprises an annual special issue of *Himalayan Linguistics* together with occasional special publications. Peer-reviewed and open-access, it focuses on linguistic and ethnographic documentation and description in the Eastern Himalaya.***

## **Languages and People of the Eastern Himalayan Region (LPEHR)**

*A study of the Morphological Patterns of Collocation in Assamese: A Thematic Overview*

---

**Mouchumi Handique**

Guwahati, Assam

### **ABSTRACT**

Collocation in general, refers to the habitual and more predictable co-occurrences of lexical items in a syntactic construction. These lexical items co-occur by choice. Certain lexical items pull each other towards themselves and so their occurrence is more frequent than others. There is no explanation to why we say *black tea* but *dark chocolate*, or why *blond* goes only with *hair* and not with any other similar object. Collocation adorns an utterance with a more specific and unique sense. J.R Firth (1957) observed collocation as a part of the meaning of a word. This paper explores the basic morphological patterns of collocation in Assamese, a language of the Indo-Aryan family, spoken in the North-East Indian State of Assam. As collocations give interesting insights about how a language community perceives life and the world, a few themes have been chosen to see how creatively Assamese uses collocations and to observe their mappings with the morphological patterns. This paper also throws light on how ‘gender’ and the ‘formal-informal distinction’ affect collocations in Assamese.

### **KEYWORDS**

Collocation; Assamese; Morphological patterns; Thematic arrangements

This is a contribution from *Himalayan Linguistics* Vol 24(2): 25–40.

ISSN 1544-7502

© 2025. All rights reserved.

Tables of contents, abstracts, and submission guidelines are available at [escholarship.org/uc/himalayanlinguistics](https://escholarship.org/uc/himalayanlinguistics)

# *A study of the Morphological Patterns of Collocation in Assamese: A Thematic Overview*<sup>†</sup>

**Mouchumi Handique**

Guwahati, Assam

## 1 Introduction

Collocation is a sub-type of multiword expressions, formed by words having a relative semantic transparency where the collective meaning reflects the whole or part of the meanings of the individual members. It is a syntagmatic attraction among lexical items which co-occur by choice. They may serve as each other's qualifier, quantifier, intensifier, and so on. Although they do not restrict other lexical items in a similar environment, they always tend to attract a particular set of words towards themselves. The word *rancid* occurring only with *butter* but not with *milk* is an example of this attraction. Collocation doesn't form just by the associating related ideas together. For example, the earth is round, but we don't often say *round earth* in English. However, we can't deny that in some other language, *round earth* might be a collocation.

Collocation, as a term was first introduced by J.R Firth (1957) who rightly argued that 'You shall know a word by the company it keeps' (Palmer, 1981, p. 76). Firth gave examples of collocation examining the occurrence of *ass* in constructions like '*You silly ass*' and with a select set of adjectives such as *silly*, *stupid*, *obstinate*, etc. Precisely, the co-occurring elements of a lexical unit give insight about the semantic, syntactic, and pragmatic assets of a word. Hence, Firth identified collocation as a part of the meaning of a word. He brought about this idea at a time when the Saussurian notion of 'langue' and 'parole' was a central topic of discussion in the contemporary linguistic literature which later inspired the Chomskian notions of 'competence' and 'performance'. Firth observed that an utterance is inseparable from the context of its occurrence and so, it derives meaning from the context as much it does from the string of sounds that it is formed of. However, Firthian notion of collocation was practically adopted in linguistics only after his death by a group of linguists, often referred to as the 'Neo-Firthian School'.

The phraseological approach to 'collocation', developed rigorously by linguists such as Hausmann slightly differs from the general Firthian notion in the sense that it focuses on the lexical co-occurrences in the phrase level. Examples are- *a flock of bird*, *a school of fish*, *a heavy smoker*, etc. This approach undoubtedly marks a significant advancement in the area, but it deviates from the general notion of collocation in some ways. From definitional point of view, this approach apparently denies the collocational relation among lexical items in a clause- '*when butter becomes rancid...*'. To deal with this problem, recently in linguistic literature, various new terms have been adopted, such as, *multi-word strings* (Mauranen, 2000), *multi-word expressions* (Evert, 2009), *lexical collocations* (Krenn (2000) in (Evert, 2009)), etc. Bartsch (2004) offered a more compact and reasonable definition of collocation that it refers to the 'lexically and/or pragmatically constrained recurrent co-occurrences of at least two lexical

---

<sup>†</sup>acknowledgements go here

items which are in a direct syntactic relation with each other’ (Bartsch, 2004, p. 76).

At present, contextual semantics and lexicography are areas where study of collocation plays a vital role. Firth himself observed that collocation is an excellent method to be adopted in descriptive lexicography. Of late, collocation has been adopted in another branch of linguistics called pedagogical linguistics. The large corpora based dictionaries namely, *The Collins Cobuild English Dictionary* (Sinclair, 1995), *Dictionary of Collocation* (Kjellmer, 1994), etc. used collocation as an extensive tool for analysing and organising senses. Sinclair (1991) used a set of terminologies to study collocations. The word under discussion, i.e., the word whose collocations are discussed is called the **node** (N), the words on either side of the node, bearing the semantic expense of the collocation form the **span** and each word within the span is a **collocate**. The collocates are marked as N+1, N-1, etc., depending on their position and distance from the node. As per this terminology, the collocate to the immediate left of the node is N+1 and to the immediate right is N-1.

For this paper, a written Assamese corpus from different genres as well as the EMILLE-CIIL Assamese Corpus on CQP Web, maintained by the University of Lancaster were used. The paper focuses mostly on phrasal collocations where the words follow a strict word order. The various morphological patterns and morpho-syntactic and semantic variations of collocation will also be examined in this paper. However, keeping in mind that collocations give insights about how a language community perceives the life and the universe, this paper tries to explore the creative use of collocations in a few themes (areas of life) and their mappings with the morphological patterns.

## 2 Functions of collocation

In most collocative constructions, a collocate serves as a modifier, quantifier, qualifier, and intensifier. It also puts forward an analogy of the item with something else, by functioning as a simile, metaphor and sometimes as a metonymy. Let’s take some examples (1-4).

1. *nisei*            *kɔm*  
very.much    little/less  
‘very little/less’

Here, *nisei* ‘very much’ functions as a quantifier that modifies *kɔm* ‘little/less’. It is to be noted that *nisei* occurs only with adjectives that stand for the concepts of ‘lesser’, ‘lower’, ‘slower’, ‘shorter’, ‘smaller’, etc. Although the core meaning of *nisei* is ‘very much’, it never modifies adjectives bearing the opposite senses of ‘greater’, ‘larger’, ‘bigger’, ‘higher’, ‘taller’, etc.

2. *ziɔni-ja*    *mou*  
rest-ADJVL    honey  
‘Resting honeybees (figurative- unmarried girls in their parents’ house)’

Example 2 is a collocation functioning as a metaphor where an unmarried girl living at her parent’s house is compared to a honeybee taking rest for a while. The ground of comparison is that a girl takes a brief rest at her parents’ place as a guest before she gets married.

3. *tilp<sup>h</sup>ul-ija*                            *nak*  
sesame.flower-ADJVL    nose  
‘A beautiful nose like a sesame flower’

The collocation in example 3 functions as a simile that compares a beautiful ‘nose’ with a sesame flower.

4. *haliki-zūtija tāt*  
martin-tangled weaving.mill  
‘A weaving mill tangled by a martin’

Here in example 4, *halikizūtija* ‘tangled by a martin’ apparently modifies the whole weaving mill, but it refers to the threads that are part of the mill. It is an example of a metonymic collocation. Another point to note is that the threads of the weaving mill are not necessarily tangled by a martin; they got tangled somehow.

### 3 The morphological patterns of collocation in Assamese

Following Sinclair, in this paper, the word whose collocations are discussed is tagged as N (Node) and the collocates as N+1 and N-1, as per their position from the node. The word to the immediate left of the node is N-1 and to the immediate right is N+1. Usually, either of them presupposes the other.

The chief morphological patterns of Assamese collocations are broadly of three types with several sub-types. Those patterns are as follows.

#### A. N-1\_N

- i. NP (Adjective\_Noun)
- ii. Adj.P (Adverb\_Adjective)
- iii. VP (Adverb/NF\_Verb(F))

#### B. N\_N+1

- i. NP (Noun-Locative\_Adjective)
- ii. NP (Noun-Instrumental\_Adjective/Noun)
- iii. NP (Noun-Genitive\_Noun)

#### C. N\_N

- i. (Noun/Nominal-Noun/Nominal)
- ii. (Adjective-Adjective)

#### 1. Pattern A: N-1\_N

1. Pattern A (i) NP (Adjective\_Noun):

### 3.1. Pattern A: N-1\_N

#### 3.1.1. Pattern A (i) NP (Adjective\_Noun)

The most common morphological pattern of collocation in Assamese is a noun phrase where an adjective pre-modifies the head noun. The adjectival modifiers in the attributive positions can be of different types. They are, core or non-derived adjective, derived adjective, negative adjective, and echo words. Here the examples are arranged as per the type of the adjective in the N-1 position.

### 3.1.1.1. Core or non-derived adjective\_noun

Core, non-derived adjectives collocating with nouns in the attributive position essentially presuppose the occurrence of the node.

5. *soup<sup>h</sup>lija*      *ɔd*  
hot.and.bright    sun  
‘Hot sunshine’

6. *ɔzanti*      *muluk*  
unknown    region  
‘A place not known’

The adjectives in the N-1 positions have no occurrence or meanings in isolation or with other words. Moreover, they never occur in the predicative positions. Hence, we can never find constructions like \**ɔd-zak soup<sup>h</sup>lija*, \**muluk-k<sup>h</sup>ɔn ɔzanti*, etc. It can be noted here that *-zak* and *-k<sup>h</sup>ɔn* are classifiers.

### 3.1.1.2. Derived adjective\_noun

Assamese has a set of denominal adjectives that collocate with a particular set of nouns. Usually, nouns take the *-ija* suffix to derive adjectival forms. Examples are as follows.

7. *hendur-ija*      *ali*  
vermilion-ADJVL    path  
‘A path as beautiful as if coloured by vermilion’

8. *tekeli-ja*      *kɔdɔ*  
round.earthen.pot-ADJVL    wasp  
‘A kind of wasp that builds a round hive that resembles a kind of earthen pot’

In example 7, *hendurija* is a derivative of *hendur* ‘vermilion’ that is derived by the adjectival suffix *-ija*. *tekeli-ja* ‘like a round earthen pot’ of example 8 is derived from *tekeli* ‘a round earthen pot with a neck’ by the adjectival suffix *-ija*.

### 3.1.1.3. Negative adjectival forms\_noun

A set of prefixes *ɔ-*, *ɔnɔ-*, *ni-* and their allomorphic variations function as negativisers in Assamese. There is a bunch of nouns that pick up a particular set of words taking any of the prefixes mentioned above to collocate with the nominal nodes as pre-modifiers. Here are a few examples.

9. *ɔ-kat<sup>h</sup>ɔ*      *zukti*  
NEG-able.to.be.cut    argument  
‘An irrefutable point/logic/argument’

10. *ɔ-kɔt<sup>h</sup>ɔ*      *ni.ɔzətɔn*  
NEG-explainable    torture

‘An unspeakable torture’

As shown in examples 9 and 10, the words *ɔkatiɔ* ‘irrefutable, undeniable’ occurs only with *zukti* ‘point, reason, logic, argument’ and *ɔkɔtʰjɔ* ‘unspeakable’ occurs only with *niɔzətɔn* ‘torture’. Moreover, they never occur predicatively.

**3.1.1.4. Echo word\_Noun**

Assamese has a bunch of echo words that go with particular concepts. Such words constitute a type of disyllabic reduplication. In writing, they are usually written without a space, but some words are written with a space, for no specific reason. Some of these words echo a real-world sound and some do not. Let’s take the following examples.

11. *.ib.ib*            *bɔtah*  
 light/gentle    wind  
 ‘Light/gentle wind; breeze’

12. *silmil*    *tɔpɔni*  
 light    sleep  
 ‘Light sleep’

The words *.ib.ib* and *silmil* in examples 11 and 12 stand for the same sense ‘light’. In example 11 it is ‘light and pleasant’ as opposed to ‘strong and cold’, and in 12, it is ‘light’, as opposed to ‘deep’. They never modify any other word and also, they always occur in the attributive positions. So, we won’t find any construction like \**bɔtah-zak .ib.ib* or \**tɔpɔni-tɔ silmil*. Here, *-zak* and *-tɔ* are classifiers.

**3.1.2. Pattern A (ii) Adj.P (Adverb\_Adjective):**

In Assamese, there are select set of adjectives and adverbial modifiers that form collocations. The adverbs *nisei* ‘very’ and *tenei* ‘very’ in the N-1 positions in examples 13 and 14 mostly occur with a select set of adjectives bearing the concepts of lesser, smaller, shorter, slower, lower, lighter, softer, younger, nearer, etc. Though the core senses of *nisei* and *tenei* are ‘very’, ‘utterly’, etc., they never co-occur with an adjective carrying the senses of bigger, larger, faster, harder, taller, heavier, better, far, etc. Let’s look at examples 13 and 14.

13. *nisei*    *ħɔ.ɯ*  
 very    small  
 ‘Very small’

14. *tenei*    *takɔ.ɯ*  
 very    less  
 ‘A very small amount’

There is no predicative use of *nisei* ‘very’ and *tenei* ‘very’.

### 3.1.3. Pattern A (iii) VP (Adverb-Verb)

A certain number of verbs take a select set of adverbs or non-finite forms of verbs functioning as adverbs. These constructions are different from conjunct and compound verbs in the sense that unlike phrasal verbs, they don't function as single lexical units. Examples (15-17) present collocations whose collocates occur only with these particular nodes.

15. *kʰɔkmɔk-koi*    *xaɪ pa*  
sudden-ADVL    wake up  
'To wake up suddenly'

16. *ɸɪɪai*                    *lɔɪ de*  
fast.and.quick    run  
'To run fast'

17. *dɔkɔsi*                    *kʰa*  
extremely/too.much    eat  
'To eat a lot'

There is no predicative use of *kʰɔkmɔk-koi* 'suddenly' (example 15) *ɸɪɪai* 'fast and quick' (example 16) and *dɔkɔsi* (example 17). It is to be noted that, *dɔkɔsi* (Example 17) has a negative connotation.

### 3.2. Pattern B: N\_N+1 (Noun-case\_N+1)

Collocations of this patterns are noun phrases with a case marked nominal head and a predicative compliment that can be either an adjective or a noun. The following discussion highlights the sub-types of this pattern.

#### 3.2.1. Pattern B (i) NP (Noun-Locative\_Adjective)

In this pattern, the node NP is a locative case marked form that takes the *-t*, *-ɔt* or *-e* locative markers. It takes a predicative adjectival compliment to form collocation. The *-e* locative marker is an old Assamese form that is currently used in restricted contexts, only. It is mostly used in reduplications. Examples are as follows.

18. *pijah-ɔt*    *atuɪ*  
thirst-LOC    anxious  
'Very thirsty'

19. *buddʰhi-t*                    *b.ɪhɔspɔti*  
intelligence-LOC    lord.Brihaspati  
'As intelligent as lord Brihaspati'

20. *bi.ɪɔh-e*    *akul*  
agony-INST    distressed  
'(Somebody) down with sorrow'



27. *mota maiki*  
 male female  
 ‘Male and female’

28. *pit.i mat.i*  
 father mother  
 ‘Parents’

Examples 25-28 are more like compound nouns. This clarifies that in certain contexts, compound nouns also form collocations. The collective meanings of the collocates match the individual meanings of the nodes. However, the collective meaning of the nodes in example 26 is different from the individual meanings of the participants. Here, ‘attention’ is a collective metaphorical sense of the individual physical activities of the organs, ‘eye’ and ‘ear’. These words always occur in this order. So, we never find a collocation like *nizɔɔa noi* ‘stream and river’ (example 25). Some collocations follow a gender hierarchy, allowing the male-referent precede the female referent. Therefore, in Assamese, *mota* ‘male’ comes before *maiki* ‘female’ (example 27) and *pit.i* ‘father’ comes before *mat.i* ‘mother’ (example 28). Members of these collocates never alter their positions. However, another collocation *ma-deuta* ‘mother and father, parents’ is an exception, as *ma* ‘mother’ always occurs before *deuta* ‘father’.

### 3.3.2. Pattern C (ii) (Adjective\_Adjective)

In this sub-type, both the lexical items are adjectives, and both are nodes. Let us take examples 29 and 30.

29. *b<sup>h</sup>al beja*  
 good bad  
 ‘Good and bad’

30. *tita mit<sup>h</sup>a*  
 bitter sweet  
 ‘The bitter and sweet tastes (of life)’

## 4 Compound and Conjunct verbs: Are they collocations?

Researchers of collocation often encounter the question of whether compound and conjunct verbs can be classified as collocations. From one point of view, the phrasal verbs are also a type of collocation, but the most remarkable difference between collocative constructions and phrasal verbs is marked by the fact that in a phrasal verb, the co-occurring lexical items serve as a single lexical unit, unlike other collocational phrases. In most collocations, both units retain their individual meanings, but when combined, they create an expression with a more specific and distinctive meaning. Assamese has a variety of compound and conjunct verbs which form a large part of the lexicon of the language. They are a separate topic of discussion in syntax for their vast morpho-syntactic and semantic variations. For these reasons, phrasal verbs are not taken for discussion as a type of collocation in this context.

## 5 Words collocating with multiple nodes

Some words collocate only with a particular node, but some others collocate with more than one node. Table 1 presents a few words collocating with more than one node.

N-1		Node		Construction	
<i>kɔna</i>	‘blind’	<i>bid<sup>hi</sup></i>	‘luck’	<i>kɔna bid<sup>hi</sup></i>	‘mischievous/ill luck’
		<i>bid<sup>h</sup>ata</i>	‘destiny, the almighty’	<i>kɔna bid<sup>h</sup>ata</i>	‘the ill destiny’
<i>gela</i>	‘rotten’	<i>gɔ.ɔm</i>	‘heat, hot’	<i>gela gɔ.ɔm</i>	‘unbearable heat’
		<i>gɔp</i>	‘pride’	<i>gela gɔp</i>	‘hollow pride’
<i>p<sup>h</sup>uta</i>	‘with a hole’	<i>kɔpal</i>	‘forehead’	<i>p<sup>h</sup>uta kɔpal</i>	‘ill fate’
		<i>kɔ.ɔi</i>	‘an old coin’	<i>p<sup>h</sup>uta kɔ.ɔi</i>	‘a useless coin with holes in it’
<i>niseitenei</i>	‘very much, utterly’	<i>ħɔ.ɔu</i>	‘small’	<i>nisei ħɔ.ɔu</i>	‘very small’
		<i>sapɔ.ɔ</i>	‘short’	<i>tenei patɔl</i>	‘very light’
		<i>kɔm</i>	‘lesser’	<i>nisei kɔm</i>	‘very less’
		<i>takɔ.ɔ</i>	‘lesser’	<i>tenei sapɔ.ɔ</i>	‘very short’

Table 1 – Words that collocate with multiple nodes

## 6 Suffixal patterns of collocations

Some noun roots derive adjectives by taking a select set of adjectival suffixes. Table 2 presents a set of nouns taking a particular set of adjectival suffixes.

Suffix		Noun		Derived adjective	
-zɔnɔk	‘that gives birth	<i>apɔtti</i>	‘complaint’	<i>apɔttizɔnɔk</i>	‘worthy of criticism’
	to something’	<i>ɔhɔsɔjɔ</i>	‘mystery’	<i>ɔhɔsɔjɔzɔnɔk</i>	‘mysterious’
-mulɔk	‘with the purpose	<i>ħɔmad<sup>h</sup>an</i>	‘solution’	<i>ħɔmad<sup>h</sup>anmulɔk</i>	‘meant for a solution’
	of, as a means of’	<i>ħasti</i>	‘punishment’	<i>ħastimulɔk</i>	‘as a means of punishment’
-pɔɔħu	‘giving birth’	<i>p<sup>h</sup>ɔl</i>	‘fruit, result’	<i>p<sup>h</sup>ɔlɔpɔɔħu</i>	‘fruitful’
-ħɔm.ɔidd <sup>h</sup> ɔ	‘prosperous, wealthy’	<i>tɔit<sup>h</sup>ɔ</i>	‘information’	<i>tɔit<sup>h</sup>ɔħɔm.ɔidd<sup>h</sup>ɔ</i>	‘very informative’

Table 2 – Nouns taking select set of adjectival suffixes

## 7 Thematic distribution of collocations and their mapping with the morphological patterns

Every language has a unique way of looking at life and the universe. This is captured in the way meaning is conceived in the language. Collocations add a special, more colourful, and specific sense to the regular meanings of the individual items. The Assamese collocations are tried to be organised in a few broad thematic classes. Here, only a few themes have been selected to see their alignment with the morphological patterns that we explored so far. It also aims to see the lexical creativity of collocation that adds to the semantic variations. The themes include the following concepts:

- i. Concrete concepts such as human body parts, material objects, utensil, etc.

- ii. The earth and the cosmic phenomena such as rain, sunshine, darkness, heat, etc.
- iii. Temporal concepts such as day, night, morning, noon, etc.
- iv. Sensory concepts such as taste, feeling, state of mind, behaviour, etc.

Tables (3-7) present a few examples of collocations from all these themes along with their corresponding morphological patterns. It is to be noted that some examples don't show any morphological pattern, for example, there is a difference between the patterns for *tilp<sup>h</sup>ulija nak* 'nose like a sesame flower' and *g<sup>h</sup>ola soku* 'pale, sad eyes'. In the first one, the adjective, *tilp<sup>h</sup>ulija* (*tilp<sup>h</sup>uln-ija<sub>ADJVL</sub>*) reflects the process of suffixation, but in the latter, *g<sup>h</sup>ola* 'pale' is an example of common adjective.

## 7.1. Material or physical concepts

### 7.1.1. Human body parts

Pattern	Node(s)	Construction
A (i) NP (adjective_noun)	<i>suli</i> 'hair'	<i>meg<sup>h</sup>ali suli</i> 'hair like dark cloud'
	<i>gal</i> 'cheek'	<i>oup<sup>h</sup>ulija gal</i> 'beautiful cheek like the flower of elephant apple'
	<i>soku</i> 'eye'	<i>g<sup>h</sup>ola soku</i> 'pale, sad eyes'
	<i>dāt</i> 'teeth'	<i>dalimgutija dāt</i> 'teeth like the seeds of a pomegranate'
	<i>aṅuli</i> 'finger'	<i>lihiri aṅuli</i> 'beautiful slim finger'
	<i>kōkal</i> 'waist'	<i>k<sup>h</sup>amusija kōkal</i> 'beautiful slim waist'
	<i>nak</i> 'nose'	<i>tilp<sup>h</sup>ulija nak</i> 'nose like a sesame flower'
B (iii) NP (Noun-Genitive_Noun)	<i>k<sup>h</sup>opa</i> 'hair bun'	<i>kōldilija k<sup>h</sup>opa</i> 'a hair bun that looks like a banana flower'
	<i>buku</i> 'chest, heart'	<i>buku<sub>1</sub> kutum</i> 'very close to heart'
	<i>hija</i> 'heart'	<i>hija<sub>1</sub> amōt<sup>h</sup>u</i> '(literal- the softest part of the heart) the most loved thing or person'
Pattern C: N_N (Noun-Noun)	<i>soku</i> 'eye'	<i>soku<sub>1</sub> mōni</i> '(literal- the pupil of the eye) the beloved'
	<i>kan</i> 'ear'	<i>soku kan</i> 'attention'
	<i>ha<sub>1</sub></i> 'bone',	
	<i>mu<sub>1</sub></i> 'head'	<i>ha<sub>1</sub> mu<sub>1</sub></i> 'all the bones of the body'

Table 3 – Human body part collocations

### 7.1.2. Miscellaneous objects, items, etc

Items in Table 4 are under the token utensils, objects, etc.

Pattern	Node(s)	Construction
A (i) NP (adjective_noun)	<i>bati</i> 'bowl'	<i>hatik<sup>h</sup>uzija bati</i> 'a bowl as large as an elephant's foot'
	<i>si<sub>1</sub>a</i> 'pressed rice'	<i>tēteliṇōtija si<sub>1</sub>a</i> 'pressed rice'

	<i>tāt</i> ‘weaving mill’	that looks like tamarind leaves’ <i>halikizutija tāt</i> ‘the badly entangled threads of a weaving mill’
	<i>nəlija</i> ‘with a handle’	<i>nəlija kətəli</i> ‘a knife with a handle’
C (i) (noun_noun)	<i>gʰəli</i> ‘house’, <i>ba.li</i> ‘land, farming field’	<i>gʰəli ba.li</i> ‘a house and land’
	<i>ko.li</i> ‘spade’, <i>kutʰəli</i> ‘axe’	<i>ko.li kutʰəli</i> ‘spade and axe, etc., working tools as a whole’
	<i>pʰəli</i> ‘slate’, <i>putʰi</i> ‘book’	<i>pʰəli putʰi</i> ‘the book and the slate’

**Table 4 – Collocation of miscellaneous objects, items, etc.**

## 7.2. The Earth; the universe

Pattern	Node(s)	Construction
A (i) (Adjective-Noun)	<i>bəməhun</i> ‘rain’	<i>kinkinija bəməhun</i> ‘light rain’ <i>sipsip bəməhun</i> ‘light rain’ <i>dʰəməhəmə bəməhun</i> ‘heavy rain’
	<i>ɔod</i> ‘sunshine’	<i>souppʰəlija ɔod</i> ‘hot and bright sunshine’ <i>pɔəkʰəli ɔod</i> ‘hot and bright sunshine’
	<i>bətah</i> ‘wind’	<i>ɔibɔib bətah</i> ‘light, pleasant wind’ <i>pʰi.pʰi.ɔija bətah</i> ‘light, pleasant wind’ <i>pʰe.pʰe.ɔija bətah</i> ‘strong, cold wind’
	<i>pəhəli</i> ‘light’	<i>dʰimik dʰamak pəhəli</i> ‘faint light’
	<i>endʰəli</i> ‘darkness’	<i>gʰitmitija endʰəli</i> ‘very intense darkness’
C (i) (Noun <sub>Noun</sub> )	<i>paha.li</i> ‘hill’	<i>paha.li bʰoijam</i> ‘hills and plains’
	<i>ɔo</i> ‘sunshine’	<i>ɔod bəməhun</i> ‘sunshine and rain’
	<i>bəməhun</i> ‘rain’	
	<i>habi</i> ‘wood’	<i>habi bənəni</i> ‘wood, forest’
	<i>bənəni</i> ‘grass’	
	<i>gəs</i> ‘tree’	<i>gəs bən</i> ‘the plants’
	<i>bən</i> ‘grass’	

**Table 5 – Collocation of the concepts from the earth and the universe**

## 7.3. Time/Season

Token	Pattern	Node	Construction
Time	A (i) NP (Adjective_Noun)	<i>puwa</i> ‘morning’	<i>bəhmə puwa</i> ‘very early in the morning’ <i>upəzi puwa</i> ‘early in the morning’ <i>kahili pəwa</i> ‘early in the morning’
		<i>dupəlija</i> ‘noon’	<i>bʰəli dupəlija</i> ‘mid noon’
		<i>gədhuli</i> ‘evening’	<i>pʰut gədhuli</i> ‘early in the evening’
		<i>həndʰiə</i> ‘evening, dusk’	<i>kal həndʰiə</i> ‘an inauspicious moment in the evening’

	<i>niha/ati</i>	<i>mazniha/mazati</i> ‘midnight’ <i>dupɔi niha/dupɔi ati</i> ‘midnight’ <i>dɔb<sup>h</sup>ag ati</i> ‘midnight’ <i>b<sup>h</sup>ɔi ba.ihā</i> ‘when the summer is at its peak’
Season	<i>ba.ihā</i> ‘summer’	

Table 6 – Collocation of temporal concepts

## 7.4. Sensory concepts

Token	Pattern	Node(s)	Construction
Physical senses	A (i) (Adjective_Noun)	<i>gɔm</i> ‘heat’	<i>utkɔt gɔm</i> ‘unbearable heat’ <i>gela gɔm</i> ‘scorching heat’
		<i>sēsa</i> ‘cold’	<i>him sēsa</i> ‘cold like snow’
		<i>tita</i> ‘bitter’	<i>ɔkɔta tita</i> ‘very much bitter’
State of mind, feelings, etc.	A (i) (Adjective_Noun)	<i>anɔndɔ</i> ‘happiness, joy’	<i>ɔnabil anɔndɔ</i> ‘an incredible joy’
		<i>mɔm</i> ‘love’	<i>ɔp<sup>h</sup>uɔntɔ mɔm</i> ‘lots of love’
		<i>k<sup>h</sup>ɔŋ</i> ‘anger’	<i>pɔsɔndɔ k<sup>h</sup>ɔŋ</i> ‘very strong anger’
		<i>ħuk<sup>h</sup></i> ‘happiness’	<i>pɔm ħuk<sup>h</sup></i> ‘extreme happiness’
		<i>sneh</i> ‘love’	<i>ɔpɔt<sup>h</sup> sneh</i> ‘very intense love’
	B (i) NP (Noun-Locative_Adjective)	<i>ħɔk</i> ‘sadness, pain’	<i>ħɔkɔt mɔijɔman</i> ‘broken in sadness’
		<i>b<sup>h</sup>agɔi</i> ‘exhaustion’	<i>b<sup>h</sup>agɔi kahil</i> ‘exhausted’
		<i>k<sup>h</sup>ɔŋ</i> ‘anger’	<i>k<sup>h</sup>ɔŋɔt ɔgniħɔma</i> ‘in fire in anger’
		<i>pijah</i> ‘thirst’	<i>pijahɔt atu</i> ‘very thirsty’
		<i>bedɔna</i> ‘sorrow’	<i>bedɔnat zɔzɔ.iti</i> ‘down with sadness’
		<i>anɔndɔ</i> ‘happiness, joy’	<i>anɔndɔt atmɔha.ia</i> ‘forgetting the self in joy’
		<i>b<sup>h</sup>ɔj</i> ‘fear’	<i>b<sup>h</sup>ɔjɔt bitɔt</i> ‘in extreme fear of something’
	C (i) (Noun_Noun)	<i>b<sup>h</sup>ɔkti</i> ‘worship’	<i>b<sup>h</sup>ɔktit gɔdgɔd</i> ‘devoted in worship’
		<i>ħuk<sup>h</sup></i> ‘happiness’, <i>duk<sup>h</sup></i> ‘sadness’	<i>ħuk<sup>h</sup> duk<sup>h</sup></i> ‘happiness and sadness’

**Table 7 – Collocation of sensory concepts**

One purpose of arranging collocations thematically in tables is to see which morphological pattern they mostly follow. The tables show that the most frequently occurring morphological pattern of collocation in Assamese is Pattern A (i) NP (Adjective\_Noun). So, we can hypothesise at this stage that Pattern A (i) NP (Adjective\_Noun) is the dominant morphological structure for Assamese collocation. Another purpose of this classification is to thematically categorise nodes and collocates which is a very helpful method of organising data for lexicographic purposes. Such classification helps lexicographers in sense identification, organisation and defining concepts in a dictionary.

## 8 Gender in Collocation

Gender in Assamese is not grammatical, but lexical and there are separate sets of adjectival collocates that go with different gender words. They are gender-specific adjectives and they always occur attributively in a construction. Table 8 Collocation and gender presents a list of gender specific collocates.

Gender	N-1	Node(s)
Female	<i>muk<sup>h</sup>soka</i> ‘sharp mouthed’ <i>ɔsɔki</i> ‘intelligent, sharp mouthed’ <i>t<sup>h</sup>upɔ.i</i> ‘not knowing how to do household things’ <i>hadɔ.i</i> ‘loving, welcoming’ <i>kazi</i> ‘very skilful’ <i>g<sup>h</sup>ɔɔd<sup>h</sup>ɔɔɔ</i> ‘one who keeps the in-law’s family together’	<i>sowali</i> ‘girl’, <i>tiɔta</i> ‘woman’, <i>ti.i</i> ‘woman’, <i>g<sup>h</sup>oini</i> (derogative) ‘wife’, etc.
Male	<i>mɔtgɔz</i> ‘brave’ <i>dɔŋuwa</i> ‘good for nothing’ <i>sɔp<sup>h</sup>ɔl</i> ‘young’ <i>hɔlɔŋgɔz</i> ‘careless’ <i>k<sup>h</sup>amidat<sup>h</sup></i> ‘courageous’	<i>loɔa</i> ‘boy’, <i>manuh</i> ‘man’, <i>deka</i> ‘a young man’ <i>pu.uh<sup>h</sup></i> ‘male’
Children/babies	<i>p<sup>h</sup>ulkumɔlija</i> ‘as soft as a flower’	<i>hihu</i> ‘children’

**Table 8 – Collocation and gender**

## 9 Formal-informal polarity and collocation

Assamese lexicon maintains a stylistic polarity of formality that affects the collocational patterns, also. Here there are a few points to note, first, most words in the lexicon are formality neutral words that are used in both formal and the written variety of the language. Second, there are words in Assamese that are used only in very formal contexts. They are not used in a home-environment. Third, there are very informal words that are never used in a formal context. Even in a home-environment, their use is restricted. Most words in the language are formality neutral. The formal and formality-neutral words take separate sets of collocates.

Table 9 Formal informal distinction in *Collocation* shows that the formal and formality neutral words are collocating with separate sets of words. They never overlap in any context.

So, we never find constructions like \**bibah gʰɔɪ* or \**bija bʰɔbɔn* for ‘marriage hall’ and \**mɔɪa deh* or \**m.ɪtɔ hɔ* for ‘dead body’.

## 10 Findings

This brief study on the basic morphological patterns and thematic distribution of Assamese collocations has the following findings:

There are three broad morphological patterns of collocations in Assamese. They are:

- (A) N-1\_N (i) NP (Adjective\_Noun) (ii) Adj.P (Adverb\_Adjective) (iii) VP (Adverb/NF\_Verb)
- (B) N\_N+1 (i) NP (Noun-Locative\_Adjective) (ii) NP (Noun-Instrumental\_Noun), (iii) NP (Noun-Genitive\_Noun)
- (C) N\_N (i) (Noun\_Noun) (iii) (Adjective\_Adjective)

Out of them, A (i) NP (Adjective\_Noun) is found to be the dominant pattern. All the themes explored here have the maximum number of collocations from this pattern.

The members of a collocational construction presuppose each other. Sometimes the node presupposes the collocate and sometimes the collocate presupposes the node and sometimes both presuppose each other. The collocate functions as a modifier, intensifier, qualifier, and quantifier and sometimes puts forward an analogy of the item with another. Collocations function as a simile as well as a metaphor, also. For example, *dalimgutija dāt* ‘beautiful teeth like the pomegranate seeds’ is a simile where someone’s teeth are said to be alike with pomegranate seeds. Another collocation, *sɔkuɪ mɔni* ‘(literal- the pupil of the eye) the beloved’ is a metaphor where the beloved is compared to ‘the pupil of the eye’ on the ground of ‘preciousness’. Similarly, some collocations have metonymic functions.

Collocations in Assamese maintain a male-female hierarchy as well as a formal-informal polarity. We have seen that except for one collocation *mɔ-deuta* ‘mother and father’, in all the other male-female constructions, the male referent comes first, for example, *loɪa sɔwali* ‘boy and girl’. There are lexical items that presuppose words from a specific gender. For example, *sɔpʰɔl* ‘young, energetic’ occurs only with *deka* ‘a young boy’. In terms of formality and informality, collocates are selective about the words they co-occur with. The formal nodes take formal collocates and formality neutral or informal words take formality neutral or informal collocates.

Words are not only selective about co-occurring lexical items, but also about suffixes. Nouns take a select set of adjectival suffixes to derive adjectives. We have come across examples of noun-adjectival compounds such as *lɔzza-zɔnɔk* [*lɔzza* ‘shame’, *-zɔnɔk* ‘that which causes to’] ‘shameful’. We won’t find a similar construction \**hukʰ-zɔnɔk* [*hukʰ* ‘happiness’, *-zɔnɔk* ‘that which causes to’] ‘causing happiness’. Collocation is a creative lexical process. The thematic analysis of collocation in different domains of life and the universe is meant for the study of this creativity in collocational constructions.

## 11 Concluding remarks

This paper is a preliminary study of collocations in Assamese, especially of their basic phraseological aspects. Collocation is an essential part of study of modern lexicography, contextual semantics and recently, of pedagogical linguistics.

This topic has drawn attention for two main reasons. Firstly, collocations exhibit a wide range of creativity in every language and offer valuable insights into how a language conceptualizes the world. They act as unique linguistic embellishments, reflecting the distinct character of a language—something that the study of individual lexical items alone cannot fully capture. Secondly, with the ongoing development of machine translation engines for various languages, including Assamese, a systematic exploration of collocations has become increasingly essential. That said, collocation is a broad field, and it is not feasible to address all its aspects within a single paper. This research will therefore be extended using data from larger corpora and through more detailed thematic analysis.

### Abbreviations

ADJ	Adjective
ADJVL	Adjectival
ADVL	Adverbial
GEN	Genitive
INST	Instrumental
LOC	Locative
N	Node
NEG	Negative
NF	Non-finite
V	Verb

### References

- Bartsch, Sabine. (2004). *Structural and Functional Properties of Collocations in English*. Tübingen : Gunter Narr Verlag.
- Evert, Stefan. (2009). Corpora and collocations. In A. L. Kytö, *Corpus Linguistics: An International Handbook* (pp. 1212–1248). Berlin, New York: Mouton de Gruyter.
- Firth, John Rupert. (1957). *Papers in Linguistics 1934–1951*. London: Oxford University Press.
- Kjellmer, Göran. (1994). *A Dictionary of English Collocations, Vol. 1–3*. . Oxford: Clarendon Press.
- Mauranen, Anna. (2000). Strange strings in translated language: A study on corpora. In Maeve. Olohan, *Intercultural Faultlines Research Models in Translation Studies* (pp. 119-141). Manchester: St. Jerome Publication.
- Palmer, Frank R. (1981). *Semantics (Second Edition)*. Cambridge: Cambridge University Press.
- Sinclair, John. (1995). *The Collins COBUILD English Dictionary*. London: Harper Collins.
- Sinclair, John. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.