Chapter 4

Link Grammar Formalism for Marathi Language

4.1 Introduction

This chapter focuses on developing Marathi Link Grammar. The designing of Marathi links has been carried out by two approaches: Phrasal approach and *Karaka* approach. Phrasal approach follows the English Link Grammar [Sleator and Temperley, 1991] which is already fully developed. We use it as guidelines for developing Marathi links. This approach has an advantage that it makes conceptualizing the linking scheme for Marathi sentences easier. As a first step towards our intended goal, we therefore analyzed various grammatical aspects such as word order, noun phrase, verb phrase, subject/object to verb agreement, and develop corresponding Marathi links accordingly. However, adaption of English links to Marathi is not so straightforward because of the syntactic differences of the two languages. Hence we have followed *Karaka* approach of Paninian grammar to develop the links. Computational Paninian Grammar [Akshar et al., 1990] has already been adopted for parsing natural languages especially Indian languages. We introduced meaningful new Marathi links using both these approaches. This chapter is organized as follows. In Section 4.2 and 4.3, we present some details of Word Order and phrases and their structure. In Section 4.4, we give details of Marathi sentence types which are taken as primary scope of our work, it helps in modeling the structures of particular sentence types. In Section 4.5, we give brief idea of proposed Marathi Link Grammar Parser. In Section 4.6 we present the phrasal approach links developed and in Section 4.7, links formulated through *Karaka* approach. In Section 4.8, we discuss links developed at sentential level i.e. studying sentence type structures of complex and compound followed by concluding remarks in Section 4.9. The structures modeled for parsing are discussed in next chapter. This Section also discusses about cross linking in Marathi sentences with possible solution to avoid it.
4.2 Word Order

Word order is a major typological property of any language. Any study on syntax will be incomplete if there will be no discussion on word order. Marathi is a verb final language with relatively free word order language. The unmarked order is SOV [Subject Object Verb]. Still in general the positions of various words in a Marathi sentence do not follow the strict ordering.

1. Lili Minila udya pustak deil लिली मिनीला उदया पुस्तक देईल.
2. Lili udya pustak minila deil लिली उदया पुस्तक मिनीला देईल.
3. Lili udya minila pustak deil लिली उदया मिनीला पुस्तक देईल.
4. Minila udya Lili pustak deil मिनीला उदया लिली पुस्तक देईल.
5. Udya Minila Lili pustak deil उदया मिनीला लिली पुस्तक देईल.
6. Udya Lili Minila pustak deil उदया लिली मिनीला पुस्तक देईल.

The sentences 1 to 6 are the variations of (Tomorrow Lili will give a Book to Mini). Thus the positions of some of the chunks of a Marathi sentence are interchangeable. On the other hand, in a Marathi sentence there are many units in which the words must occur in a fixed order, for example,

- Auxiliary verbs always follow the main verb (e.g. jayala hava जायला हवा but not hava jayala हवा जायला)
- Case suffixes and postpositions always follow the nouns (e.g. Lili Ne लिली ने but not Ne Lili ने लिली)
- Adjectives precede the noun they modify (e.g. pandhari gadi पांढरी गाडी but not gadi pandhari गाडी पांढरी)

Hence we observe that the position of various phrases in a sentence is, in general, not fixed, but the position of the constituent words is fixed in a phrase. Below we discuss the phrase structure of Marathi.

4.3 Phrases and their Structure

The syntactic study of the structure of any natural language begins with the details of the phrases. This section presents with the description of phrases and their structures of Marathi language.
4.3.1 Noun Phrase

The minimal unit of a Noun Phrase (NP) can be a noun or a pronoun. A noun phrase functions as an argument of the main verb in a sentence. The Unmarked relative order of the modifiers in a noun phrase is as follows,

| Demonstrative adjective; Possessive adjective; Quantifiers; Intensifiers; Adjective; Relative clause; Noun; Head noun; Emphatic particle |

There has not been any systematic study of the preferred order of the modifiers relative to one another. However following are some observations where some modifiers shall be in fixed order.

4.3.2 Verb Phrase

According to Damale [Damle, 1965] a compound verb is composed of a sequence of a sentence of two verbs V1+V2. The primary verb V1 is non finite while V2 is categorized as an auxiliary verb or helping verb. The primary meaning of this Verb-Verb compound is determined by the lexical meaning of the main verb. The auxiliary verb adds a special or additional meaning to the meaning of the main verb. According to Pandharipande [Pandharipande, 1997] there are three types of verb groups based on their formal difference which are

- **Simple**: Selects only one verb root in a sentence.
- **Conjunct Verb**: This type of verb comprises a noun/adjective/adverb and verb in a sentence. They are formed by adding verbs such as होणे (to happen), करणे (to do), वाटणे (to feel), देणे (to give), घेणे (to take) etc. This process is extremely productive in Marathi.
- **Compound Verb**: This involves a sequence of two verbs or sequence of verbs. As explained in [Damle, 1965].

Let us observe the verb or Verb Phrase (VP) in the following sentences.

7. *Ram gaana gaato* राम गाणं गातो
   (Ram sings a song)

8. *Ram gaana gaat aahe* राम गाणं गात आहे
   (Ram is singing a song)

9. *Ashokne fal khaala* अशोकने फल खाला
   (Ashok ate a fruit)
10. *Ashok fal khat hota* अशोक फल खाता होता
   (Ashok was eating a fruit)

11. *Aasha udya yeil* अशा उद्या येईल
    (Aasha will come tomorrow)

12. *Aasha udya yet rahil* अशा उद्या येंत राहिल
    (Aasha will be coming tomorrow)

13. *Usha chhan gaate* उषा खात गाते
    (Usha sings well)

14. *Usha hushar aahe* उषा हुशार आहे
    (Usha is clever)

15. *Usha sundar diste* उषा सुंदर दिसते
    (Usha looks beautiful)

From the above examples sentences 7,9,11 have only one verb root, sentences 8,10,12 have main verb as well as auxiliary verb i.e these sentences have compound verb and sentences 13,14,15 have adverb and adjectives linked with verb so they are conjunct verbs. From the above analysis the surface structure of the Marathi verb phrases can be shown by the following diagram,

![Verb Phrase Diagram](image)

The most difficult part of any language is verb. Learning a language means how to operate verbal forms of a particular language. The following categories that are related with the verb phrase have to be considered in detail while processing the language. They are
(a) GNP (Gender, Number, Person):

Gender: The Marathi verb forms reflect gender inflectionally. The concord is between the gender of the subject or object NP and the verb form. The only instances where gender is not changed are the forms of the verb *aahe* आहे*(to be)* and the future forms of the verbs.

Number: The sense of number is technical and it is different from the sense of number in relation to nouns etc. In verbs, it is dependent grammatical function, a feature of concord, since it depends on the number of the subject NP or the object NP and not on anything inherent in the verb.

Person: Person is a way of classifying referents in relation to the speaker; it can take various forms, but the one found in Marathi. Pronouns is traditionally interpreted as distinguishing between speaker (first person speaker), person(s) addressed (second person), and referent (personal or otherwise) spoken of (third person). Indeed it is a category rather of concord between subject NP and VP than one appropriate to a single constituent of a sentence. The following table of the verb *ga* गळ (sing) in relation with the Marathi personal pronoun will make the point clear.

<table>
<thead>
<tr>
<th>Type</th>
<th>First Person</th>
<th>Second Person</th>
<th>Third Person</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>M</td>
</tr>
<tr>
<td>Singular</td>
<td><em>mi gatoo</em></td>
<td><em>tu gatos</em></td>
<td>to gato</td>
</tr>
<tr>
<td>Plural</td>
<td><em>aamhi gatoo</em></td>
<td><em>tumhi gatat</em></td>
<td>te gatat</td>
</tr>
</tbody>
</table>

Table 4.1: Verb Relation with Pronoun

(b) TAM (Tense, Aspect, Mood):

Tense: All the Marathi verbs have forms in present, past and future tense. Tense is marked by a suffix which immediately follows the verb stem and precedes all other suffixes. The agreement markers for Person, Number and Gender are different in different tenses. Three tenses Present, Past and future are formally distinguished. These tenses are marked by the following suffixes. The tense aspect and mood system of Marathi is extremely complex due to the following facts (i) tense and aspect markers
often merge (ii) several explicator verbs carry the function of aspect / tense markers and (iii) time reference is not always determined by tense markers on a verb within the same clause.

<table>
<thead>
<tr>
<th>Tense</th>
<th>Marker</th>
<th>Example</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present</td>
<td>-t</td>
<td>Basato</td>
<td>He sits</td>
</tr>
<tr>
<td>Past</td>
<td>-l</td>
<td>Basala</td>
<td>He sat</td>
</tr>
<tr>
<td>Future</td>
<td>Vowel + -l-n</td>
<td>Basel</td>
<td>He will sit</td>
</tr>
</tbody>
</table>

Table 4.2: Tense Marker on bas

**Aspect**: The notion of tense is different from the notion of aspect. Conceptually tense expresses time relation between events while aspects express event interval time. In Marathi aspects resides in what are fundamentally tense expressions and at times boundaries between tense and aspect are blurred. There are altogether eight aspect markers in Marathi they are perfective, imperfective, second perfective or inceptive, infinitive and a completive.

**Mood**: Marathi has a large range of modal forms which are expressed by (i) the markings on the verb. (ii) the use of the modal auxiliaries and (iii) a combination of both (i) and (ii). Moods such as indicative, conditional, imperative, optative, intentional, debitative etc. Express conditions as well as goals of the action expressed by the verbs.

(c)**Finite and Non-finiteness**:

Marathi makes a clear distinction between finite and non finite forms. Finite forms can be used independently as the main verb in a simple or a complex sentence as well as in embedded clauses. Non finite forms are used only in subordinate clauses. The derived non-finite forms are infinitives, gerunds and participles. Infinitives and gerunds are not inflected for tense, voice, aspect and mood etc. However the participle (present, past and future) maintains their aspectual reference.
4.4 Sentence Types

4.4.1 Simple Sentences

A simple sentence as a rule consists of a predicate and one or more arguments [Dhongade and Wali, 2009]. Predicates may be copular, intransitive, transitive, causative and dative type. There is also a group of impersonal predicates that are devoid of any arguments. Verb less sentences are mostly confined to exclamations and elliptic questions.

Copular sentences: A copula is a linking verb with no independent meaning of its own. Marathi has two copulas आहे to be and होणे to become, happen. They vary for e.g.

16. *tu lekhika aahes* तु लेखिका आहेस [Copula आहेय]  
you are a writer

17. *tu lekhika zaalis* तु लेखिका झालीस [Copula होणे]  
you became a writer

Declarative sentences: A declarative sentences falls into five basic categories on the basis of the classification of verbs (i) intransitive (ii) transitive and causative (iii) impersonals (iv) datives and (v) passives Intransitive have only one argument. Intransitive verbs inflect for tense, aspect and mood and agree with the nominative subject in gender, number and person. An intransitive sentence may be extended with adverbs, adjectives or postpositional phrases.

18. *Lili ratri baara vajta nijte* लिली रात्री बारा वाजता निजते  
(lili sleeps at twelve at night)

19. *Liline zaad todale* लिलीने झाळ तोडले  
(Lili cut the tree)
20. *Minine babula hasawale* मिनीने बाबुला हसवले  
(Mini made/ caused Babu to laugh)

21. *Minine lili kadun babuche karvi kaam karawale* मिनीने लिली कडून बाबुचे कार्य काम करवले  
(Mini got Lili to do the work through Babu)

Simple transitive verbs take two arguments while double transitive verbs take three. Causative sentences are derived from both intransitive and transitive verbs by the suffix *aw*-व. Such causative are complex structures but interesting feature of these structures is mediators like कडून, कार्य, द्वारा come in particular fixed order.

**Impersonal Sentences:** Such sentences lack a referential subject. A large group of these sentence belong to weather verbs. A small group of health expression and impersonal passive sentences also belongs to this group. Weather expression like *dhaagaalale* दगाळले (become cloudy), *ujaadale* उजाले (become light/dawn), *sanjawale* साजाले (it is evening), and health expressions like *malmaalty* मालमालत्या (it feels queasy in the stomach), *gudmartay* गुडमरत्या (it feels suffocated) etc. are some of the verbs come under such sentences.

**Dative Sentences:** Sentences with dative subjects are found with a class of verbs such as inchoative, stative, psych-verbs and some others. The structure of dative sentences is extremely controversial. Dative sentences obligatorily take a dative subject.

22. *तिला थंडी बाजते*  
She feels cold

23. *तिला लाज बाजते*  
She feels shy

**Modal Sentences:** Modality is notional category associated with aspects of reality and its refraction in mind and language. Modality serves to express the attitudes and
opinion of the speaker to the prepositional content of a sentence. Modal sentences thus contrast with declarative sentences, which are truth oriented. Modals express wish, intention, possibility, necessity, commands and other subjective attitudes.

24. *aapan madivar jaau ya* आपन माइवर जाओ या [Imperative mood]
   (Let us go upstairs/ Should we go upstairs)

25. *lili baget javi* लिली बागेत जावी [Optative]
   (May Lili go in the garden)

26. *lili babula baget jaau dele* लिली बाबुला बागेत जाओ देने [Permissive Complex]
   (Lili allows Babu to go in the garden)

27. *lili ne baget jaau naye* लिली ने बागेत जाओ नये [Obligative modal]
   (Lili should not go in the garden)

28. *lili khelu pahate / ichhite* लिली खेलू पाहते / इच्छिते [Desiderative]
   (Lili wants to play)

29. *lilila nachtaa yete* लिलिला नाचता येते [Capability/Potential]
   (Lili knows how to dance /Lili can dance)

### 4.4.2 Complex Sentences

Complex sentences are either of the complement or the correlative type [Dhongade and Wali, 2009]. In both there is certain interdependence between the main and the dependent clause. A complement clause is embedded under a main clause and may be finite, non finite or a small clause. Non finite clauses are derived from infinitive and participles. The correlative structure consists of a pair of clauses containing relative and correlative elements in mutual relationship. The relative clause is considered subordinate to the correlative. It usually precedes the correlative though other orders are also found.

**Complement Structure:** There are three varieties of the structure let us see it in
following, Marathi complement system is rather complex. The principal Complementizer is  की. It marks all finite complement clauses.  की  precedes the complement clause. In addition the main clause contains words such as असं / हे / ही गोट / (असं) माणस (so/this/this story/saying so) clearly these words belong to the pronominal or the noun category. Their function is to relate the main clause to the complement clause. Complement schema is shown in following tree chart.

Figure 4.2: Complement Schema

Following are some sentences of a particular type

30. ही गोष्ट विचित्र आहे कि लिली लग्न केल<br>
   \( A1 \)

   (The story that lili got married is strange)

31. हे विचित्र आहे कि लिली लग्न केल [A2]

   (It is strange that lili got married)

32. असा दिसतय की आज पाउस पडेल [B1]

   (It seems that it will rain today)

33. असा लिलिला वाटत कि मिनी इधे गाठी [B2]

   (Lili believes/thinks that Mini is not here)
34. asaa maza andaz aahe ki udya paus padel असा माझा अंदाज आहे की उदया पाउस पडेल [B3]
   (In my opinion it will rain tomorrow)

35. liline minila asa sangitala ki babu bajarat gela लिलीने मिनिला असं सांगितला की बाबू बाजरात गेला [C]
   (Lili told mini that Babu went shopping)

**Small clause complement**: A small clause lies midway between a finite and a non finite clause. It consists of an attributive clause embedded under certain predicates such as maan, samaz, mhan, wat, dis मान, समज, महं, म्याण, वाट, दिस (Consider, think, say, believe, seem).

36. lili samzate / maanate ki mini sundar aahe लिली समजते / मानते की मिनी सुंदर आहे (Lili thinks that mini is beautiful)

**Non Finite Complement Clause**: Non finite complement clauses are derived by means of the infinitive in ne -ने, the inceptive participle u -उ, or the past participle in l/le -ल/ल सुफ़िज़्द with adverbial markers. These clauses are essentially nominal in character they function as subject or object of certain predicates. Subject clauses mark the predicate with ne/na -ने/ना.

37. tina tithe rahaana yogya navhate तिने तिथे रहाणे योग्य नव्हाते [subject complement clause]
   (Her staying there / for her to stay there was not proper)

38. liline minila nachayale shikvala लिलीने मिनिला नाचायला शिकवला [Object Complement Clause]
   (Lili taught Mini to dance)

39. liline lagna kelyachi varta khoti nighali लिलीने लग्न केल्याची वार्ता कोटी निघाली [Complex NP]
   (The rumour about Lili getting married turned out to be false)
Correlative structure: Correlative structure includes various adjectival and adverbial clauses. In a basic correlative structure the relative and the main clause are juxtaposed forming a relative correlative pair. Each clause carries its own relative and correlative marker. The correlative pattern is mostly confined to finite clauses, it is absent in non finite and certain adverbials. Non finite clauses employ participles.

Adjectival correlative: Adjectival correlative are either restrictive or appositive. Restrictive correlative clause exhibit four main pattern namely full correlative, gap and free relatives and multiple headed correlatives.

Full Correlatives: In the full correlative both the relative ‘J’ and the correlative ‘T’ vary with the case, gender and number of their respective head nouns. In its basic form the relative clause precedes the correlative with several variations such as a nouns follows both the relative and the correlative marker.

40. *ji mulgi ghari geli ti dha aahe* जी मुलगी घरी गेली ती मुलगी द आहे
   (The girl who went home is stupid)

   The correlative noun can be deleted like,

41. *ji mulgi ghari geli ti dha aahe* जी मुलगी घरी गेली ती द आहे
   (Variation of sentence 40)

   The correlative marker ‘T’ must not be deleted, following sentence hence becomes ungrammatical.

42. *ji mulgi ghari geli mulgi dha ahe* जी मुलगी घरी गेली मुलगी द आहे *

   A postposed relative marker is more acceptable if it is followed by an additional correlative marker as follows,

43. *ti mulgi dha aahe ji mulgi ghari geli* ती मुलगी द आहे जी मुलगी घरी गेली

   This construction also allows correlative noun to delete as follows,

44. *ti dha aahe ji mulgi ghari geli(ti)* ती द आहे जी मुलगी घरी गेली (ती)
But not the correlative marker ‘T’ as

45. dha aahe ji mulgi ghari geli # # द आहे जी मुलगी घरी गेली

In short in a correlative clause the noun may delete but not the correlative marker ‘T’. Similarly the relative noun may delete as in 43. But the marker ‘J’ must not delete when the relative clause is postpose as follows,

46. ji mulgi dha aahe ji mulgi ghari geli ti ती मुलगी द आहे जी मुलगी घरी गेली ती *

To summarize (i) When a relative clause precedes the correlative clause, the relative marker ‘J’ may delete along the noun giving rise to a gap relative. (ii)The correlative marker ‘T’ must not delete. (iii)When the correlative precedes the relative clause the relative ‘J’ and the correlative ‘T’ must not be deleted. (iv) When the correlative head precedes the relative clause the noun must delete in the relative clause. In short the noun may be repeated in both relative and correlative only when the relative clause precedes the correlative.

**Gap relatives**: Gap relatives are generated by deleting the relative marker ‘J’ along with the noun as follows,

47. ghari geli ti mulgi dha ahe # घरी गेली ती मुलगी द आहे
   (The girl who went home that girl is stupid)

48. ti mulgi dha aahe # ghari geli ti ती मुलगी द आहे # घरी गेली ती
   (The girl is stupid, the one that went home)

**Free Relatives**: Free relatives also called as headless or fused relative refer to an identified antecedent and optionally allow interrogative pronoun to follow the relative and correlative pronoun as in following e.g.

49. jo hasato to shevti radto जो हसतो तो शेवटी रडतो
   (He who laughs weeps at the end)
Multiple headed relatives: In a multiple headed relative clause several NPs are simultaneously relativized. The correlative clause contains a correlative NP corresponding to each relative NP. The relative clause may precede or follow the correlative clause.

50. jya mulaane je pustak vaachale tya mulala te pustak aavadate ज्या मुलाने जे पुस्तक वाचले त्या मुलाना ते पुस्तक आवडते

(Which boy read which book, he liked it)

Appositive correlatives: Appositive correlatives offer information that is incidental to main message for e.g.

51. to maza mitra jo hindi bolto, amerikan aahe तो माझा मित्र जो हिंदी बोलतो, अमेरिकन आहे

(That friend of mine, who speaks hindi, is an American)

The appositive correlative differs from the restrictive correlative in several respect such as

(a) In a basic appositive correlative the relative clause follows the head, while in a basic restrictive correlative clause is in preceding position.

(b) In the appositive correlative the relative ‘J’ is obligatory while the correlative ‘T’ is optional as follows,

52. # maza mitra jo hindi bolto, amerikan ahe # माझा मित्र जो हिंदी बोलतो, अमेरिकन आहे

(A friend of mine/my friend, who speaks Hindi is an American)

(c) In the appositive correlative, though not in the restrictive correlative, the correlative ‘T’ may be replaced by हा, ही, हे (This (m,f,n)); अशा जो, अशी जी (such that (m,f,n)) as follows

53. ha maza mitra jo hindi shikvato. ek prasidha lekhak aahe हा माझा मित्र जो हिंदी शिकवतो, एक प्रसिद्ध लेखक आहे

(This friend of mine, who teaches Hindi, is a famous writer)
An appositive correlative but not a restrictive correlative may be antecedced by
quantifiers, proper nouns and first and second person pronoun as follows,

54. joshi buwa je aamchya kade rauhta far majedaar maanuus ahahe जोशी बुवा जे आमच्या कडे राहताने फार मजेदार मानुस आहेघ

Non finite adjectival relative clauses: Non finite adjectival relative clauses
utilize present, past and future verbal participles. These constructions delete the cor-
relative marker, and the final verb or the auxiliary is put in its corresponding participle
form. The correlative noun follows the participle. The word order within the participle
clause is fixed. The participle shows gender number agreement with the one it modifies.

55. ji mulgi hasat aah he ti mulgi ithe rauhte जी मूलगी हसत आहे ती मूलगी इथे राहतेह (The girl who is laughing lives here)

56. hasat asaleli mulgi ithe rauhte हसत असलेली मूलगी इथे राहतेह (Laughing girl lives here)

Adverbial correlative clauses: Adverbial clauses may be finite or non finite.
Most finite adverbial clause are correlative type. The pair J...T connects the subor-
dinate and the main clause. Adverbial finite clauses may precede or follow the main
clause. They maintain the verb final order. The subordinate connective J may move
within the clause but must precede the verb. Word movement within the clause is
also possible with certain restrictions. Non finite adverbial clauses utilize participle or
infinitival verb forms as noted below,

57. jevha lili jail tevha mini jail जेव्हा लिली जाईल तेव्हा मिनी जाईल [Temporal clause]
(Mini will go when lili will go)

58. jase to wedewakade bolto tase ti bolate जसे तो बेडेवाकडे बोलतो तसे ती बोलतेह [Manner
clause]
(She speaks the same crazy way he does)
59. *jithe tu jatos tithe mini jaate* जिथे तू जातो तिथे मिनी जाले [Locative clause]

(Mini goes wherever you go)

60. *lili jitaki sundar aahe titki mini naahi* लिली जितकी सुंदर आहे तितकी मिनी नाही [Equative Clause]

(Mini is not as beautiful as lili)

61. *jar lili aali tar tila mi pustak dein* जर लिली आली तर तिला मी पुस्तक देईं [Conditional clause]

(If Lili comes, I will give her the book)

62. *jari tu mazi bahin aslis tari mi tula madat karnaar nahi* जरी तु माझी बहिण असलिय तरी मी तुला मदत करणाऱ नाही [Concessive clause]

(I will not help you even though you are my sister)

**Conjunctive adverbial clause:** Adverbial clauses expressing cause, result, intention and purpose do not use the correlative pattern. They employ conjunction particles such as *karan, mhanun, sabab, hetu, ya hetu ne* कारण, महनुन, सबब, हेतु, या हेतु ने (because, so, therefore, reason, intention, with this intention) in their finite version as follows,

63. *lili ushira aali karan (ka / ka ki ) ti kamat hoti* लिली उशिरा आली कारण ( का / का की ) ती कामात होती [Cause]

(Lili came late because she was working / busy)

64. *lili kamat hoti mhanun ushira aali* लिली कामात होती महनुन उशिरा आली [Reason]

(Because Lili was working so she came late)

65. *lilila bhetave mhanun mi aalo aahe* लिलिला भेटावे महनुन मी आलो आहे [Intention/reason]

(I have come in order to see Lili)
4.4.3 Compound Sentences

There are three major coordinators namely conjunctive, disjunctive and adversative [Dhongade and Wali, 2009] in compound sentences.

**Conjunctive coordinators:** The basic conjunctive coordinator is आणि, meaning and with alternative such as अं , अणि, अणि, अणि, अणि (alternatives of and).

**Sentence Coordination:** Any number of sentences can be coordinated with आणि which is placed before the last conjunct. In a sequence of more than two sentences all preceding sentences before the last are simply juxtaposed as in following example,

66. **babu gela aani lili ghari aali** बाबू गेला आणि लिली घरी आली

   (Babu left and Lili came home)

67. **babu gela, lili ghari aali aani lagech minila phone kela** बाबू गेला, लिली घरी आणि आणि लगेच मिनिला फोन केला

   (Babu left, Lili came home and immediately phoned Mini)

**Constituent Coordination:** In constituent coordination in most cases आणि is placed before the last conjunct.

**Noun Coordination:** Nouns of all categories may be coordinated. However the coordination is complicated by the fact that in Marathi both case and agreement systems are complex. Noun coordination requires coordinate nouns to bear the same case.

68. **lili, sudha aani mini gharat hotye** लिली, सुधा आणि मिनी घरात होत्या

   (Lili, Sudha and Mini were in the house)

**Pronoun coordination:** All pronouns in any person and in any form can be coordinated agreement takes the plural form of the highest person irrespective of the pronominal order and number are as follows coordination of adjectives, adverbs and active and passive verbs.
69. mi aani tu udya ghari jaau मी आणि तु उधा घरी जाऊ
(I and you will go home tomorrow)

70. lili jara baawali aani vedi aahe लिली जरा बावळी आणि वेडी आहे [Adjective]
(Lili is little bit disorderly and crazy)

71. lili halu halu aani mand swarat bolte लिली हाळ हाळ आणि मंद स्वरात बोलते [Adverb]
(Lili speaks slowly and in a low voice)

72. lili chendu fekte aani zelte लिली चेंडू फेकते आणि जेलते [verb]
(Lili throws the ball and catches it)

73. chor kholit shirla aani lagech pakadala gela चोर खोलत शिरला आणि लागेच पकडला गेला
(thief entered the room and was immediately caught)

**Disjunctive Structure:** There are three disjunctives किंवा, का/की, अथवा all expressing the sense of or. The first किंवा is prevelant. The second का/की is used in interrogative and in subordinate clauses expressing the sense of whether. In both sentence and constituent coordination किंवा is placed immediately before the last sentence or constituent as the case may be. It may also appear before each sentence or sentential constituent. It is never placed in the beginning of the first sentence or first sentence constituent. Although किंवा allows a juxtaposed sequence like आणि, unlike आणि it may however not be totally absent from the sequence. The last placement of किंवा is obligatory. Agreement in the disjunctive structure is always with the last element of the disjunction as is to be expected.

74. lili ghari geli asel kiwa baget basali asel लिली घरी गेली असेल किंवा बागेत बसली असेल [Sentence Coordination]
(Lili may have gone home or she may be sitting in the garden)

75. All sentence constituents employ kiwa,

*lili kiwa babu ghari jail* लिली किंवा बाबू घरी जाईल [Noun Coordination]
(Lili or Babu will go home)

76. *lilene halu halu kiwa jorane aani thasun bolayala have* लिली ने हल्ल हल्ल किवा जोराने फांसून बोलायला हवे [Adverb Coordination]

(Lili should speak either slowly or loudly and forcefully)

77. *lili balala uchalate kiwa khali thevate to radayacha thambat nahi* लिली बाळाला उचलते किवा खाली ठेवते तो रदायचा ठांबत नाही [Verb Coordination]

(Lili lifts baby or puts him down / whether Lili lifts the baby up or puts him down, he does not stop crying)

**Adverbial Structure:**

The three adversative coordinators *pan, parantu and tathapi* पण, परतू, तथापि meaning but are semantically identical except in their usage. The last one is used mostly in formal contexts. The first two are nearly exchangeable.

78. *lili hushaar aahe pan abhyas karat nahi* लिली हुशार आहे पण अभ्यास करत नाही

Lili is intelligent but does not study

79. *babu khup abhyas karto parantu paas hoil ka nahi te sangta yet nahi* बाबु खूप अभ्यास करतो परतू पास होईल का नाही ते सांगता येत नाही

Babu studies a lot it cannot be predicated whether he will pass or not

### 4.5 Proposed Marathi Link Grammar Parser Framework

This parsing process in our proposed framework involves three important phases; preprocessing, parsing and post-processing. The block diagram of proposed framework is as given in Figure 4.3.

**Pre-processing:** We have seen that Marathi is a morphologically rich language. All words are not in their root form. Mostly they are derived from the root word for example in English the root word go has several derivations for e.g. go, goes, going, gone
et. Hence storing all the derivatives of all words is necessary but leads to a lexicon to be of very huge size.

Similarly part of speech tagging of every word is necessary. Major categories in Marathi are noun, verb, adjective, adverb etc. These word classes are not just the idle invention but are useful categories for further processing. Hence accordingly to keep the lexicon size optimum, root word along with other information like saamanya roop inflected form / word, root word, POS are stored in lexicon itself. There is a lot complexity involved in morphological structure as well as in POS tagging of words which is attributed to irregularities of languages.

**Parsing:** Parsing is the process by which an input sentence is analyzed and assigned a suitable structure. Parsing process makes use of two components: a parser which is a procedural component (a computer program), and a grammar which is declarative. The function of the parser is to accept individual tokens and opt appropriate linking requirement of each word from the Link Dictionary and building a valid linkage in
the sequence of words given as a input. Link Dictionary so have special importance. Link dictionary contains the linking requirement of each word with another word. To identify linking requirement we have taken two approaches the first one is phrasal approach and the other one is Karaka approach. The basic assumption in Marathi Link Grammar parser is; each word is linked with one or the other way to the other word in a sentence. Identifying this appropriate linking requirement is the major task of the proposed algorithm. The linking requirement are so proposed or modified from English Link types that they are self explanatory and gives the relation between that linked words.

**Post-processing:** Post processing phase ensures the condition needs to be satisfied as per Link Grammar formalism i.e. exclusion and satisfaction [Sleator and Temperley, 1991]. This phase also produces linkage which is generated in the parsing phase. The detail description of experiment is explained in next chapter. Following section cover the Marathi Link Grammar formulation which is the foundation for the parsing through link grammar formalism.

### 4.6 Link Development using Phrasal Approach

At the earlier stage in research, we attempted to formulate Marathi Grammar by studying phrases of Marathi. The literature available for Marathi [Damle, 1965, Pandharipande, 1997, Dhongade and Wali, 2009] study the language in the form of western grammarians pattern. Phrasal link development is influenced by English Link Grammar. English and Marathi language similarities and differences were studied to propose the Marathi Phrasal links.

#### 4.6.1 Noun Phrase Links

The Noun phrase links developed [Patil and Pawar, 2012] for the Marathi Link Grammar formulation purpose are discussed in this section.
**D**: Marathi does not have a separate category of articles. The function of articles is carried out by numerals एक (one) and the indefinite pronoun कोणी (Some one). This link connects एक or कोणी which behaves as article with the noun. e.g. Relative order for Marathi noun phrase has been discussed in above Section 4.3.1. Considering the facts following link types are proposed

![Diagram](image)

**Figure 4.4: Link D**

![Diagram](image)

**Figure 4.5: Link D**

**A**: This link connects the attributive adjective with the noun. Adjectives precede the noun they modify. English adjectives are uninflected they do not change morphology with the variations in the noun they qualify. But in Marathi adjectives could be inflected or uninflected. The link A is proposed for the uninflected adjectives of Marathi for e.g. *hushaar हुशार* (intelligent) is an uninflected adjective (e.g. *hushar mulga हुशार मुलगा* (intelligent boy), *hushar mulgi हुशार मुलगी* (intelligent girl)). A link is followed by three suffixes, the first one can either be ‘m’ male or ‘f’ female or ‘n’ neuter denoting the gender of the head noun, the second suffix gives the number information i.e. ‘s’ for singular and ‘p’ for plural and the third suffix ‘j’ can be used as a case ending marker of the head noun. Case ending marker are associated with head noun like मुला ने मुला ने here ने is case ending marker.
MA: This link connects the pre modifying adjective with the noun it modifies. MA link is proposed for the inflected adjectives and they too depend on gender, number and case ending of the head noun. Following are some examples of inflected adjectives changala mulga चांगला मुलगा (nice boy), changali mulgi चांगली मुलगी (nice girl), changale zaad चांगले झाड (nice tree). The notion of suffix is described as in following link type of gender. Neuter is additional in Marathi for e.g. झाड is neuter and so adjective used for झाड too gets inflection चांगले.

Figure 4.8: Link MA with Suffix m and s

Figure 4.9: Link MA with Suffix f and s
In Marathi certain form of adjectives do exists like reduplicated adjectives, some substrings like *saa / shi / shya* is attached with adjectives which do not have an equivalent structure in English so for such kind of adjectives we propose various links and they are as given below,

**AS**: This link connect the adjective with *saa / shi / shya* this link follows by three suffixes as discussed in link A.

![Figure 4.11: Link AS](image)

**AR**: In Marathi reduplicated adjectives are used like *chhan chhan goshti* (nice nice stories). **AR** link is used to connect the first repetition of adjective and then second adjective links with noun followed by it.

![Figure 4.12: Link AR](image)

**J**: This link connects the nouns with case ending of possessive adjectives. Possessive adjectives precede the head noun as modifier case ending. Case ending *cha / chi / che / chya* can be used as possession of noun for e.g. *shyamcha ghoda* (shams horse), *sitechi pati* (sitas slate). Suffix ‘p’ is proposed
to show the possession of noun with case ending *cha/chi/che/chya*.

**DC**: This link is proposed to connect the case ending *cha / chi / che / chya* च / ची / चे / च्या with the head noun. **DC** link uses suffix as explained for link **A**. Let us see the example of using **J** and **DC**.

```
raam    cha           ghoda aath mail dhaavala
rama      caa             GaaoDa Aaz maOla Qaavalaa
Raam's horse ran eight miles
```

Figure 4.13: Link DC and J

Quantifiers are also the part of order in noun phrase morphology in Marathi quantifiers consists of the following co-occurring categories,

```
(Approximate)----(cardinal)------(collective)
(ordinal)            (Multiplicative)      (Measure)
                   (Fractional)
```

Figure 4.14: Categories of Quantifiers

We propose following few link types for the categories of quantifiers,

**ENi**: This link connects the indefinite words like *sumare / jawal jawal* सुमारे / जवल जवल (approximately) to expression of quantity. It uses suffix ‘i’ to show the indefiniteness of the word.

```
sumaare      chaalis mule ranget ubhe hote
saumaaro           caaLIsa maulao raMgaot ]Bao haoto
approximately fourty boys were standing in a queue
```

Figure 4.15: Link EN with Suffix i

**ENo**: This link connects the ordinal words like *pahile / shevatche / madhale* पहिले (first)/ शेवटचे (last) / मध्ये (in-between) for expression of quantity.
ENoc: This link connects the ordinal words with the fractional measure like *ardha* (half), *paav* (quarter) etc. here ‘c’ suffix used as count for measure.

**Figure 4.17: Link EN with Suffix o and c**

EM: This link connects the fractional with measure of noun.

**Figure 4.18: Link EM**

ET: This link connects the definite words *te* / *ti* / *tya* / *ティ* / *त्या* with the ordinal words.

**Figure 4.19: Link ET**

Emphatic particles such as *ch* (only) / *pan* (also) / *sudha* (as well) follow the noun phrase as discussed in relative order. We propose following links,

**M:** this link connects nouns/pronouns to quantifier *sagale* (all) or *thode* (little).
All people had came

Figure 4.20: Link M

we all came yesterday

Figure 4.21: Link M

**Me**: This link connects emphatic particles *ch च* (only) / *hi ही* (also) / *sudha सुधा* (also)/*tar तर* (indeed).

**Mc**: The object of comparison may precede or follow the head noun. The suffix *peksha पेखा* is used as object of comparison. This link connects the suffix *peksha पेखा* (compared to) with noun.

### 4.6.2 Verb Phrase Links

The links which are developed for the Marathi verb phrase [Patil and Pawar, 2011b] are as follows with the brief description, before explaining verb phrase links let us discuss about two links developed for present participle and past participle which modifies...
noun but are inflecting verbs,

**Present Participle:**

We propose link type MN for present participle which takes two suffixes to give more information. The first suffix is for gender and second one is for number information. For this we need to separate the words for ex. *naach naara mor* नाचणारा मोर (dancing peacock), *rad naari mulgi रुद्धारी मुलगी* (crying girl) here *naach नाच* (to dance) or *rad रा* (to cry) are verbs but modifying nouns.

```
naach  naara
naaca  naara
dancing
```

**Figure 4.24: Link MN with Suffix m and s**

```
gaa   naari
gaa  naarI
singing
```

**Figure 4.25: Link MN with Suffix f and s**

**Past Participle:** Link ML is proposed to connect past participle which take two suffixes like MN link. Example of this category is *baslela mulga बसलेला मुलगा* (sited boy), *umalaleli kali उमलालेली काळी* (blossomed bud) here again *basne* (to sit), *umalne (to blossom)* are the verbs but modifying nouns like *mulga* (boy) and *kali* (bud).

```
bas  lela
sat / seating
```

**Figure 4.26: Link ML with Suffix m and s**

```
umal  leli
blossomed
```

**Figure 4.27: Link ML with Suffix f and s**
P: Complements of आहे / आहेत / असतील / असेल (present form of is); होत / होती / होते / होत्या (past form of was) are linked to these morpho words through P links. Various suffixes may follow this link to differentiate between different types of complements suffix ‘a’ and ‘p’ are used for adjective and post position respectively.

That flower is beautiful

Figure 4.28: Link P with Suffix a

Tanuja is in the room

Figure 4.29: Link P with Suffix p

PT: This link connects ta/te/to form of verb with auxiliary verb आहे / आहेत / असतील / असेल ; होत / होती / होते / होत्या . Suffixes for this link depends upon the auxiliary verb being used as shown in following table. This verb construction is used in case of present or past indefinite. e.g.
chalat चालत (walking), uthat उठत (standing), basat बसत (seating), gaat गात (singing), yet येत , are main verb.

<table>
<thead>
<tr>
<th>Morpho Word</th>
<th>Suffix</th>
<th>Morpho Word</th>
<th>Suffix</th>
</tr>
</thead>
<tbody>
<tr>
<td>आहे</td>
<td>*s</td>
<td>होत</td>
<td>ms</td>
</tr>
<tr>
<td>आहेत</td>
<td>*p</td>
<td>होती</td>
<td>fs</td>
</tr>
<tr>
<td>असेल</td>
<td>*s</td>
<td>होते</td>
<td>mp</td>
</tr>
<tr>
<td>असतील</td>
<td>*p</td>
<td>होत्या</td>
<td>fp</td>
</tr>
</tbody>
</table>

Table 4.3: Suffixes for Morpho-words

In this entries * implies that the corresponding suffix may be either ‘m’ or ‘f’ depending on the gender of the subject.
**Figure 4.30**: Link PT with Suffix m and s

**Figure 4.31**: Link PT with Suffix f and s

**PE**: The aa/e/ii verb form is connected to the auxiliary verb आहे / आहेत / अस्तीत / असेल / होत / होती / होते / होत्या. This link may be followed by two suffixes according to above table

**Figure 4.32**: Link PE with Suffix m and p

**CJ**: This link connects noun/adjective/adverb/pronoun with the verbalizer in a conjunct verb. This link takes five suffixes to give additional information as follows, First suffix can be ‘n’ or ‘a’ or ‘p’ or ‘d’ which mean noun or adjective or pronoun or adverb respectively Second suffix gives information on verbalizer it can be ‘t’ or ‘e’ or ‘r’ ता / ते / तो form or आ / ए / ई form and root form of verbalizer respectively Next three suffixes give the Gender, Number and Person information

**Figure 4.33**: Link CJ with Suffix n, t, f, s and i
CD: This link is defined to join various components of a compound verb. It joins the lexical verb with the intensifiers. Since a verb phrase may involve more than one intensifier, the lexical verb can have more than one CD link. This link also have five suffixes as of CJ link. First suffix tells about the primary/main verb. It can be either ‘i’, ‘t’, ‘a’ or ‘r’ for infinitive, ता / ते / तो form, आ / इ / ई form and root form of the main verb respectively. Next four links are same as last four links of CJ link.

CJ1: This link connects noun/adjective/pronoun/adverb with infinitive form of verb. This type of construction is used in verbal. This link may followed by two suffixes. First suffix is either ‘n’, ‘a’, ‘p’, ‘d’, respectively. Second suffix is ‘j’, If the oblique form of the verb infinitive is used. It may be noted that oblique form is used if the verbal is followed by some case ending.
CDR: Reduplicating verbs are also used as reduplicating adjectives *chhan, chaan goshti*, such verbs are connected with CDR link

4.6.3 Subject/Object to Verb Links

The morphological dependence of verb phrases of Marathi language on subject/object is quite different than English Language. Marathi links need appropriate suffixes on the link to check the agreement between verb morphology and subject/object of the sentence. Verb Morphology dependence on Subject/Object

In case of nominative and dative subjects, the Marathi verb morphology dependent on object instead of subject as discussed below. In general, subjects must be encoded in the form of verb by means of gender, number and person suffixes.

**Nominative Subject**

Intransitive verb has agreement markers of gender and number of the subject where as the subject of a transitive verb in the perfective aspect is marked with the postposition *ne ने*. The first and second person pronouns do not take *ne ने*. These are called as ergative subjects i.e the verb does not agree with the subjects; rather it agrees with the direct object.

**Dative Subjects**: There is a class of verbs in Marathi that requires to be marked with the dative postposition *la ला*. For such subjects too the verb morphology is independent of subject. e.g. Subject is feminine singular, object is masculine plural,
verb is dependant on object Ex. *Reema la vees hazar rupaye dile.* रिमा ला वीस हजार रुपये दिले (Twenty thousand rupees has been given to Rima)

Also, if the direct object is followed by a postposition the verb does not agree with any element in the sentence. In this case, it takes the neuter, third person, singular marking i.e the default agreement.

Following table illustrates the verb dependence in various sentences.

<table>
<thead>
<tr>
<th>Sentence</th>
<th>Subject</th>
<th>object</th>
<th>Verb Dependence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ti Zopali</td>
<td>F,S</td>
<td>-</td>
<td>Subject</td>
</tr>
<tr>
<td>To Zopala</td>
<td>M,S</td>
<td>-</td>
<td>Subject</td>
</tr>
<tr>
<td>Tyane amba khalla</td>
<td>M,S</td>
<td>M,S</td>
<td>Object</td>
</tr>
<tr>
<td>Tine amba khalla</td>
<td>F,S</td>
<td>M,S</td>
<td>Object</td>
</tr>
<tr>
<td>Tyane ambe khalla</td>
<td>M,S</td>
<td>M,P</td>
<td>Object</td>
</tr>
<tr>
<td>Tine ambe khalla</td>
<td>F,S</td>
<td>M,S</td>
<td>Object</td>
</tr>
<tr>
<td>Rama ne Ravana la marle</td>
<td>M,S</td>
<td>M,S</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 4.4: Verb Dependence

**Subject/Object Verb Links:** As per above discussion, we propose following links to connect subject/object with the verb [Patil and Pawar, 2011a].

**Subject Links:** Marathi verb morphology may or may not depend on the subject of the sentence. We, therefore propose three types of subject links for subjects without any case ending, nominative subject and dative subject.

**S:** It connects subject (without any case ending) to the main verb. S link may have three suffixes representing the gender, it can either be ‘m’ for male or ‘f’ for female. The Second suffix gives the number information i.e. ‘s’ for singular and ‘p’ for plural. The third suffix gives the information about the person of the subject. This is used in case of the first person (मैं) or second person (तू). Suffixes ‘i’ for (मैं) (symbolic I) and ‘u’ for तू (symbolic you) are used.

**SN:** If the subject is followed by case-ending -ने then the verb morphology is independent of subject. In this case the subject is linked to the case ending -ने with Jn link and ने is linked to main verb via SN link. Ex Raam ne aamba khalla. रामने आँबा खाला
SL: This link is used in the dative subject i.e. subject followed by the case ending ना. In this case also the verb morphology is independent of the subject. Here the subject is linked to the case ending ना with Jl link and ना is linked to the main verb via SL link Ex.  

```
Aarun la chinch dili अरुणला चिंच दिली .
```

Figure 4.40: Link SL

Object Links: Depending upon whether the verb morphology depends or not on the object, we propose the following two object links.

O: It connects the transitive verbs to direct or indirect objects when the object is not followed by any case ending. If the main verb makes SN link with the subject, the gender and number information has to be attached with the O link using two suffixes. Person information is not important for object.

Ex.  

```
Kavita ne chinch khalli कविताने चिंच खाली.
```

Figure 4.41: Link O with Suffix f and s

OL: This link connects the-la with the main verb when the object is followed by the case ending-la. Object is connected to the case ending -la with Jl link.
4.7 Karaka Approach

The second approach for formulating the links in our framework is *Karaka* relations. *Karaka* relations are the syntactic-semantic relations which are defined in a Paninian Grammar. The goal of Paninian approach is to construct a theory of human natural language communications. Paninian grammar is particularly suited to free word order languages. It makes use of *vibhakati* information for mapping to semantic relations and uses positions information only secondarily. The Paninian grammar was originally designed more than two millennia ago for writing a grammar for Sanskrit. It has been adopted by [Akshar et al., 1990] to deal with modern Indian languages in Computational Paninian Grammar. The notion of *Karaka* relations is central to the Paninian model. It gives relations between verbal and other related constituents in a sentence. They capture certain level of semantics. The levels in the Paninian model are shown in the following figure,

- - - - - Semantic level (what the speaker has in mind)
- - - - - Karaka level
- - - - - Vibhakti level
- - - - - Surface level (uttered sentence)

Figure 4.43: Semantic Level

The surface level is the uttered or written sentence. The *vibhakati* level is the level at which there are words or local word groups based on case endings, prepositions or
postposition markers. A noun group is a unit containing a noun, its vibhakti and possibly some adjective. vibhakati for verbs can also be followed by auxiliary verbs. Such information is collectively called vibhakati for the verb. The vibhakati level abstracts away from many minor differences among languages. The topmost level relates to what the speaker has in his mind. This may be considered to be the ultimate meaning level. Between this level and vibhakati level is the Karaka level. It includes Karaka relations and a few additional relations such as tadarthya (purpose). One can imagine several levels between the Karaka and the ultimate level, each containing more semantic information. Thus, Karaka level one in the series of levels, but one has relationship to semantics on the one hand and syntax on the other.

There are around six different types of Karaka relations. It is clearly not possible for them to capture the innumerable types of semantic relations among all possible actions or states (written as verbs) and all possible objects (written as nominals) in the world. What the Karaka relation do specify with respect to a verb, however, are six or so relations of nominals (or verbals) that participate in the action specified by the particular verb. These are sufficient for providing the maximum necessary information relative to verb. Following section give details of each Karaka relation links proposed [Patil and Pawar, 2013] for the parser.

4.7.1 Karaka Links

**Ka_ Karta** - Out of all participants in an action there is one who is Swatantra or most independent. The action denoted by verb in the sentence is acted by something or someone and is called Karta Karaka. In our system we are labeling such relation i.e. Karta with verbs as Ka_ Karta here Ka stands for Karaka relation and the relation is Karta. Karaka relations are the relations of nominal that participate in the action specified by the particular verb mentioned in the sentence.

80. mazyaa shaletil rajune suvarnpadak jinkale माझ्या शाळेतील राजुने सुवर्णपदक जिकले  
(Raju who from is my school won goldmedl)
Here राजु ने and सुवर्ण पदक are the words which relate directly with the verb जिकले. So by asking question ‘who won the gold medal ?’ the Karta gets identified so we connected this two words by Ka_ Karta. Karaka relation shown as below,

\[
\text{माझ्या शालेत राजु ने सुवर्ण पदक जिकले}
\]

Raaju from my school won gold medal

Figure 4.44: Link Ka Karta

**Ka_Karma का कर्मः:** In this same example we could get one more Karaka relation which is Karma Karaka. Above we saw that Karta the swantantra or the most independent participate in an action, is the ashray (or the locus) of activity. In other words the activity resides in or moves forth from the Karta, the result of which is called Karma Karaka for finding Karma (if one exists) in sentence is to ask questions like in same example जिकले क्या करणावर झाले ? then the answer can be drawn as सुवर्ण पदक.

To show this relation we proposed Ka_Karma का कर्म link to show relation between Karma and Kriyapad i.e. Verb. So now we can draw Ka_Karma link in same example as shown below.

\[
\text{माझ्या शालेत राजु ने सुवर्ण पदक जिकले}
\]

Raaju from my school won gold medal

Figure 4.45: Ka Karma

**Ka_Karan का करण :** The next Karaka is Karan which is instrument. Karan is called to that word which shows the instrument of the activity in a sentence. e.g.

81. मी सुरीने सफारचंद कापले
(I cut an apple with knife)

Here in the sentence the instrument used is सुरी so the words सुरीने and कापले are connected with Ka_Karan का करण which gives following representation.
I cut an apple with knife

**Figure 4.46: Link Ka Karan**

**Ka_. Sampradan का. संप्रदान** - When activity in the sentence gives the meaning of donating something then the word on which this donating activity happened is called sampradan for e.g.

82. *shivaji maharajani dadoji konddevana gurudakshina dili* शिवाजी महाराजानी दादोजी कॉंडेवाना गुरुदाक्षिणा दिली

(Shivaji Maharaj offered Gurudakshina to Dadoji Konddev)

Here the donating गुरुदाक्षिणा happened to दादोजी कॉंडेव by शिवाजी महाराज दादोजी कॉंडेव and दिली verb gets connected with Ka_sampradan का. संप्रदान.

**Figure 4.47: Link Ka Sampradan**

**Ka_. Aapadan का. अपादन** - Any activity starts at some point and then from that point if the person or things departs or separates from it such separation action is called as Apadan. Apadan Karaka participates in an action involving separation takes place. This relation is shown by Ka_apadan का. अपादन like among those words.
83. zadavarun fule baget padli झाड़वरुन फूले बागेत पदली
(Flowers fall down from tree into the garden)

![Diagram of ka_aapadan]

Ka_ Adhikaran का. अधिकर्ण - Adhikaran Karaka is the locus of Karta or Karma i.e. the place or time of the activity mentioned in the sentence is called as Adhikaran. In the example 83 word बागेत is Adhikaran and hence बागेत and पदली are connected with Ka_ Adhikaran का. अधिकर्ण which is shown in following figure.

![Diagram of ka_adhikaran]

There are major six Karaka relations exists as per Panini framework and there are many Up-pad relations occur in a sentence. Up-pad relations are the relations of those words which are not directly participate with verb of a sentence. e.g. Consider sentence 80 here, माझ्या शाळेतील are the words which are related to noun i.e. Karta and which are not participate in relating to verb जिकलेले of the sentence. Hence they have concerned up-pad relations as shown below,

![Diagram of u_swamitwa]

Figure 4.48: Link Ka Aapadan

Figure 4.49: Link Ka Adhikaran

Figure 4.50: Link Up Pada
4.8  Sentential Link Development

The proposed Marathi Link Grammar parser framework has been set up in the scope of Simple sentences, Complex sentences and Compound sentences. In above section we have discussed about proposed link types for various phenomenon of the Marathi Grammar which includes linking mechanism for simple sentences but complex and compound sentences need to have additional linking requirements along with links developed for simple sentences hence in this section we explore proposed link types of complex and compound sentences.

4.8.1  Crossing of the Link

Before exploring complex and compound sentence link types, we would throw light on why we need such sentential links additionally. Marathi complex or compound sentences can usually be expressed in more than one way. The linking scheme for Marathi complex sentences is developed so that linking of all types of structure is consistent. The biggest challenge dealing with complex and compound sentences is crossing of the links. That is planarity rule. We observed that, in general planarity cannot be maintained for Marathi complex and compound sentences. Following complex sentence e.g. violet the planarity rule if system builds links in its usual manner. The crossing of the links occurs because of the correlative structure. In sentence number 41, since ‘mulgi’ is subject मुलगी of the verb phrase धा आहे द आहे ka.Karta link is also required in it and so crosses the correlative marker Ti ती.

![Crossing of the Links](image)

Figure 4.51: Crossing of the Links
To avoid such crossing of links complex as well as compound sentences can be parsed in two levels: the first level giving the clausal links and the second level giving the internal clause links. That is splitting the parse structure in two levels the upper level deals with relative-correlative marker and chunks of clauses and lower level deals with the words within the clause. New links are proposed to have valid and functional linkage between the words of complex sentences.

```
<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PMR</td>
<td>RCM</td>
<td></td>
<td>RCM</td>
<td>PMR</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mulgi ghari gehi</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ka_karta</td>
<td>ka_adhikaran</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mulgi</td>
<td>ghari</td>
<td>geli</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dha</td>
<td>aahe</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dha</td>
<td>aahe</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cmc</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dha</td>
<td>aahe</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

Figure 4.52: Two Level Linking Scheme

The links proposed as shown in above figure are RMR which connects relative marker to relative clause, link RCM connects relative clause to correlative marker and link CMC connects the correlative marker to correlative clause.

### 4.8.2 Complex Sentence Links

In Marathi language complex sentences are either of the complement or the correlative type. In both the types there is certain interdependence between the main and the dependent clause. A complement clause is embedded under a main clause and may be finite, non finite or small clause. Marathi complement system is complex. The Principal Complementizer is की कि. कि (precedes the complement clause and in main clause words such as असा/हे/हि goshta/ (asa) mhanun असा / हे / हि goshta / (संस्कृत द्वितीय मृदु) ध्वनि महत्त्व (so/this/this story/ saying so) are included. There exist many variations of complement structure. A correlative structure consists of a pair of clauses containing relative and correlative elements in mutual relationship. The relative clause is considered subordinate to the
correlative. It usually precedes the correlative though other orders are also found. Each clause carries its own relative and correlative marker. Relative and correlative markers handled in our system are $ji-ti$, $jar-tar$, $jehva-tevha$, $jyane-tyane$, $jo-to$, $jya-tya$, $jine-tine$, $jase-tase$ etc. The proposed links for complex sentences [Patil and Pawar, 2015] are discussed in this section.

### 4.8.2.1 Proposed Links for Complex Sentences

**HM**: This link connects Header $hi$ to the main clause of the sentence for e.g in sentence number 30 The header $hi$ precedes the main clause whenever such order is encountered header and main clause (here the main clause is *goshta vichitra aahe* गोष्ट विचित्र आहे) is connected through the link HM. The link is shown if following figure.

**HC**: This link connects header $hi$ to the Complementizer $ki$. Complementizer marks all finite complement clauses. The main clause contains header $hi$ clearly these words belong to the pronominal or the noun category. Their function is to relate the main clause to the complement clause. So to relate this HC link is proposed.

**MCO**: Here M stands for main clause and CO stands for Complementizer. This link is proposed to connect Main clause to Complementizer $ki$. 

**COC**: This link is proposed to connect the Complementizer $ki$ to the Complement clause. All the above links is the part of one sentence only. So when all the chunks i.e Header, main clause, Complementizer and complement clause is present all this link will appear at sentential level.
**CH**: This link is proposed to connect Complementizer to header. Consider the variation of sentence number 30

84. *llila lagna kela hi goshta vichitra aahe* लिली लग्न केला ही गोष्ट विचित्र आहे

here the complement clause *llila lagna kela* लिली लग्न केला precedes the header ही in such order CH link is used.

![Figure 4.54: Link CH](image)

**SH**: Link SH connects subject of the main clause to the header. Consider the sentence,

85. *llila mini ithe nahi asa vatata* लिली मिनी इथे नाही असा वाटत

(Lili thinks that Mini is not here)

the sentence analysis says that *asa* असा is a header which precedes the main clause *vatata वाटत*, *mini ithe nahi* मिनी इथे नाही is a complement clause and main subject i.e. *Karta* of a main clause is *llila* लिली in this case subject *lili* लिली has to be connected with the header by that means to main clause, hence link SH is used. This kind of sentence is defined as center embedded complement clause.

![Figure 4.55: Link SH](image)
CAM: Consider the sentence,

86. lili baralat hoti ki mala maru naka asa mhanun  लिली बरछल होती कि मला मारु नका असं महून

(Lili was whispering that do not beat her)

Certain complex sentence take quotative word mhanun महून or asa mhanun असं महून. These construction optionally utilize ki की just in case the main clause is in preceding position. To connect complement clause in this case with the quotative word asa mhanun असं महून CAM link is proposed.

OM: Link OM connects object of the main clause of the complex sentence. Similar to SH link object of main clause may precede complement clause which precedes main clause such construction is treated as complex sentence, which comes under non finite complement clause as discussed in Section 4.4.2.

SM: Link SM connects subject of the main clause with main clause. It is the same link as OM only the difference is subject connects except object. Consider the sentence,

87. lilila babune minishi lagna kelyach aavadala nahi  लिलीला बाबुने मिनीशी लगन केल्याच आवदला नाही

(Lili does not like that Babu got married to Mini)

here main clause is aavadala nahi आवदला नाही, babune minishi lagna kelyach बाबुने मिनीशी लगन केल्याच is a complement clause and lilila लिलीला is a subject of the main clause. Hence lilila लिलीला to aavadala nahi आवदला नाही will be connected through SM link. This is also a non finite complement clause.
**RMR:** The first RM stands for relative marker and the other R is for relative clause. The link **RMR** connects Relative marker to relative clause. There are various relative marker *ji, jevha, jar,jitaki, jari, jya, jo, joparyant* etc. *जी, जेब्हा, जर, जितकी, जरी, ज्या, जो, जोपर्यंत* etc, all such markers precede relative clause. Hence in such structure RMR link is used.

**RMCM:** Here ‘CM’ stands for correlative marker and as stated above ‘RM’ is for relative marker so ‘J’. ‘T’ pair of relative marker and correlative marker is connected with **RMCM**. Pairs such as *jar tar, jitaki titaki, jevha tevha* etc जर तर, जितकी जितकी, जेब्हा तेब्हा. are connected with this link.

**RCM:** Link **RCM** connects relative clause to correlative marker. At certain correlative structure deletion of relative markers are allowed and hence in such case relative clause connects with the correlative clause.

**CMC:** Correlative marker to the correlative clause is connected through the link **CMC**. Correlative markers are *ti, tevha, tar, titaki, tari, tya, to, toparyant* etc. ती, तेब्हा, तर, तितकी, तरी, त्या, तो, तोपर्यंत etc.

**CMRM:** The link **CMRM** connect the correlative marker to the relative marker. Consider the variation of sentence number 43,

88. *ti mulgi dha aah jīghari geli* ती मुलगी द आहे जी मुलगी घरी गेली

here correlative clause is preceding the relative clause with their markers preceding
to their own clauses. Hence in such case **CMRM** link is used.

**CRM**: Correlative clause to relative marker is connected with the **CRM** link. This link becomes the part of links of a sentence discussed in **CMRM** link.

![Diagram](image)

**CMS**: Link **CMS** connects the correlative clause to subject of the correlative clause. Consider the sentence number 51, which is a relative clause with the relative marker and *to maaza mitra माझा मित्र* is a correlative clause which precedes the relative clause. But *to maaza mitra amerikan aahe माझा मित्र अमेरिकन आहे* is the complete correlative clause which is broken up in two predicates. So the subject of a correlative clause *mitra मित्र* is connected to the correlative clause with **CMS** link.

**RC**: Link **RC** connects the relative clause to correlative clause. Consider the sentence number 51 again in which relative clause precedes the part of correlative clause in such case link **RC** is used. Such kind of construction of complex sentence is defined as appositive correlatives. This link is depicted in Figure 4.45.

**HS**: There is a variation of the appositive correlative is also possible i.e. the correlative marker can be replaced by header *ha/hi/he हा, ही, हे* (this (m,f,,n)). Link **HS** is used to connect such header to the subject of the correlative clause.

**SC**: Link **SC** connects subject of the subject of the correlative clause to the correlative clause. Consider the sentence number 54 here *je aamchya kade rahtaat जे आमच्या कडे राहतात* is a relative clause and *phar majedaar manus aahe फर मजेदार मानुस आहे* is a correlative clause and *joshi buwa जोशी बुवा* is subject of this clause. Hence to connect
subject of correlative clause SC link is used.

---

**ADM:** This link is used to connect adverbial clause to the main clause. Adverbial clauses expressing cause, result, intention and purpose do not use correlative pattern. They employ conjunctive particles such as karan, mhanun, sabab, hetu, ya jetu ne कारण, म्हणून, सबब, हेतु, या हेतू ने in their finite version. Consider variation of sentence number 64,

89. **kamat aslya-mule lili ushira aali कामात अस्लयांमुळे लिली उशीरा आली**

(As Lili was working / busy, she came late)

here **kamat aslya-mule कामात अस्लयांमुळे** is a adverbial clause and **lili ushira aali लिली उशीरा आली** is main clause. Such adverbial clause is connected with main clause by **ADM** link.

---

**MCP:** As discussed in above example of conjunctive adverbial clause main clause is connected with conjunctive particles explained above. Consider sentence number 63, here **lili ushira aali लिली उशीरा आली** is a main clause, **karan ki कारण कि** is a conjunctive particle and **ti kamat hoti ती कामात होती** is a adverbial clause so MCP will link **lili ushira aali लिली उशीरा आली to karan ki कारण कि**.

**CAD:** This link connects conjunctive particle to adverbial clause. As per above example link **CAD** connects **karan ki कारण कि to ti kamat hoti ती कामात होती**. Following figure shows **MCP** and **CAD** link

---
4.8.3 Compound Sentences

In Marathi language, coordination is of two types: sentence coordination and constituent coordination. There are three major coordinators namely Conjunctives, Disjunctive and Adversative. Any number of sentences can be coordinated with aani आणि which is always placed before the last conjunct. In a sequence of more than two sentences, all preceding sentences before the last are simply juxtaposed. Various parts of speech can also be coordinated at constituent level. Nouns of all categories may be coordinated. Pronouns, adjectives, adverbs and active and passive verbs can also be coordinated. While coordinating within a sentence part of speech follows certain agreement rules on the conjoining category. Thus we are treating coordination at sentential level and on constituent or word level. Following are the links proposed [Patil and Pawar, 2014] for both this categories of coordination.

4.8.3.1 Proposed Links for Compound Sentences

SeCC: Link SeCC connects the sentence which precedes the conjunctive coordinators. This is a sentential level link. Consider the compound sentence number 66, here babu gela बाबू गेला is a separate sentence and lili ghari aali लिली घरी आली is another sentence which is combined one with conjunct aani. So to link babu gela बाबू गेला to aani आणि link SeCC is used.

CCSe: Link CCSe is the further part of above link. This link connects conjunctive coordinator aani आणि to lili ghari aali लिली घरी आली for the sentence stated above. Here CC is initials taken of conjunctive coordinator and Se stands for sentence.
**SeDC**: This is similar to **SeCC**. Link **SeDC** connects Sentence to Disjunctive coordinators like *kiwa, ka / ki, and athava* किवा, का / की, अथवा.

**DCSe**: Link **DCSe** links disjunctive coordinator to the sentence which precedes it.

**SeAC**: Link **SeAC** connects sentence with adversative coordinator like *pan* पण, *parantu* परंतु or *tathapi* तथापि. Consider the sentence number 78, here *lili hushar aahe* लिली हुऱ्यार आहे is one sentence which is combined with *pan* पण to another sentence *abhyas karat nahi* अभ्यास करत नाही to link such sentence with adversative coordinator link **SeAC** is used.

**ACSe**: Link **ACSe** connects adversative coordinator with the sentence which is followed by it.

**SCC and CCS**: There are word level constituents too. To connect two subjects in one sentence link **SCC and CCS** is used. Consider sentence,

90. *lili ani sudha bajarat gelya* लिली आणि सुधा बाजऱत गेल्या  
    (Lili and Sudha went into the market)
is noun coordination. To connect lili लिली with aani आणि link SCC is used and aani आणि and sudha सुधा is connected with link CCS.

**OCC and CCO:** This two links are used to connect two objects coordinated with conjunctive coordinator for e.g. consider example,

91. *lili ne aambe aani keli khaale* लिली ने आंबे आणि केळी खालें (Lili ate mangoes and bananas)

here objects *aambe aani keli* आंबे आणि केळी is connected through links CCO and OCC link.

**AjCC and CCAj:** Adjectives which are coordinated with conjunctive coordinator are linked with AjCC and CCAj link respectively. Consider example sentence number 70, here *bawali aani wedi बावली आणि वेडी* are adjectives and are combined with conjunctive coordinator *aani आणि*.

**AvCC and CCAv:** Link AvCC links adverb which precedes the conjunctive coordinator. Link CCAv connects conjunctive coordinator which precedes the adverb.
Consider sentence number 71, here *halu halu aani mand* हलू हलू आणि मंद is adverb of verb *bolate* बोलते.

SDC and DCS: ‘S’ stands for subject and ‘DC’ stands for disjunctive coordinator. Consider sentence number 75, here *lili kiwa babu* लिली किवा बाबू is a subject coordination with disjunctive coordinator *kiwa* किवा so links SDC connects *lili* लिली with *kiwa* किवा and link DCS connects *kiwa* किवा to *babu* बाबू.

DCO and ODC: Link DCO and ODC are used to connect two objects with disjunctive coordinator. For example consider the sentence

92. *lili pen kiwa pensil aanele* लिली पेन किवा पेंसिल आनेल

(Lili will bring pen or pencil)

here *pen kiwa pensil* पेन किवा पेंसिल is object of the sentence and connect this object DCO and ODC links are used.

AjDC and DCAj: Consider sentence,

93. *te biskit god kiwa kharat asu shakate* ते बिस्किट गोद किवा खारत असू शकते

(That biscuit can be sweet or sour)
here \textit{god kiwa kharat} गोड किवा खारट is adjective of the sentence and is combined with the disjunctive coordinator. At such adjective conjunction link \textit{AjDC} and \textit{DCAj} is used.

![Figure 4.71: Link AjDC and DCAj](image)

\textbf{AvDC and DCAv}: Adverbs conjunction by using disjunctive coordinator is also possible i.e in sentence number 76, \textit{halu halu kiwa jorane} हलु हलु किवा जोराने is the adverb of verb \textit{bolayala pahije} बोलयाला पाहिजे. In such type of adverb conjunction link \textit{AvDC} and \textit{DCAv} is used.

![Figure 4.72: Link AvDc and DCAv](image)

We point out to certain observation as follows,

- All word level conjunctions are the inner links and are take part in whole sentence linking once they are chunked as one whole unit and then it is linked with the verb.

For example,

\textit{94. lili ani babu bajarat gele} लिली आणि बाबू बाजारत गेले

(Lili and Babu went into the market)

here \textit{lili ani babu} लिली आणि बाबू is the \textit{Karta} of the \textit{Kriyapada gele} गेले.

So subject to conjunctive coordinator links are opted at inner level and the whole unit \textit{lili ani babu} लिली आणि बाबू gets connected with \textit{gele} गेले.

- Adversative coordinators are used in sentential conjunction only. They do not take part in word level linking.

For example,
4.9 Concluding Remarks

In this chapter, we made an exhaustive study on word order, their phrase structure and sentence structure of Marathi grammar. Thereafter we discussed idea of a proposed Marathi Link Grammar Framework and Link types are proposed. Links proposed have been broadly classified into three categories.

- **Category 1 (Phrasal Links):** Phrase structure of Marathi grammar is studied in detail. Few English links are used as it is for Marathi Link Grammar and few others are modified as per detail study. Total 31 links are proposed in this category.

- **Category 2 (Karaka Links):** *Karaka* relation from Paninian Grammar were studied and new meaningful links based on it are proposed. Total six *Karakas* can directly be identified by using the theory and hence we propose six *Karaka* links.

- **Category 3 (Sentential Links):** At sentential level to preserve Planarity rule, we proposed sentential level links. Specifically crossing of the links occur in complex and compound type of sentences hence, we proposed total 44 links under this category.

The above observations shall be helpful in developing link grammar for any language in general. This paves the way for critical analysis of various linguistic aspects of the language which in turn helps in formalizing its linking requirements.

For quick reference a comprehensive list of all Marathi links discussed in this work and brief description of their usage is given in Appendix 1, 2 and 3 for phrasal, Karaka
and sentential links respectively. It must be observed that the list of Marathi links developed in this work is not complete, however we have developed the Marathi links required to parse more common sentences of Simple, Complex and Compound type. This chapter dealt with the first and important step towards achieving the goal of development of Marathi links. In the following chapter, we discuss algorithm proposed for Marathi Link Grammar parsing, the experimental setup and results.