

Contents

Contents	vii
List of Figures	ix
1 Introduction	1
1.1 Phrases and Idioms	1
1.2 Linguistic Characteristics of Assamese	2
1.3 Context Free Grammar	3
1.4 Purpose	4
2 Related Work	6
2.1 English Phrase Parsing	6
2.2 Parsing Idioms in Lexicalized TAGs	7
2.3 Identification of Non compositional Phrases	8
2.4 Parsing Noun Phrase Structure with CCG	9
3 Assamese Phrases and Idioms	10
3.1 Phrases in Assamese	10
3.1.1 Noun Phrase (NP)	10
3.1.2 Verb Phrase (VP)	12
3.1.3 Prepositional Phrase /Postpositional Phrase (PP)	14
3.1.4 Adjective Phrase (AP)	15
3.1.5 Adverb Phrase (AdvP)	16
3.2 Idioms in Assamese	17
3.3 Difficulties in Identification	18
4 Our Proposed Method	23
4.1 Introduction	23
4.2 Production Rules	24
4.2.1 Noun Phrase	25

4.2.2	Verb Phrase	25
4.2.3	Prepositional Phrase /Postpositional Phrase (PP)	26
4.2.4	Adjective Phrase (AP)	26
4.2.5	Adverb Phrase (AdvP)	26
5	Implementation	27
5.1	Flex	27
5.2	Bison	28
5.3	Assamese Phrase Parser	29
5.4	Result	30
6	Conclusion and Future Work	32
6.1	Conclusion	32
6.2	Future Work	32
	Bibliography	34

List of Figures

1.1	A parse tree showing the derivation bc from A	4
2.1	A parse tree for “A flight”	7
3.1	A parse tree for an NP	12
3.2	A parse tree for a VP	14
3.3	A parse tree for a PP	15
3.4	A parse tree for an AP	16
3.5	A parse tree for an AdvP	17
3.6	An incorrect parse tree	20
3.7	A parse tree	21
4.1	A parse tree for মানুহ দুজন	24
4.2	A parse tree for দুজন মানুহ	24
5.1	Block diagram of the Parser	30