

PRONOMINALIZATION IN BENGALI

HUMAYUN KABIR

PhD

UNIVERSITY OF EDINBURGH

1976



ACKNOWLEDGEMENTS

I should like to express my indebtedness to Dr Edward Keith Brown, my supervisor, who suggested the topic of this work. I am profoundly grateful to him for his instruction and encouragement, which have made this work possible.

I would like to thank my friend, Robert Calder, who read an early version of this work, for his suggestions.

This work was made possible by a scholarship from the Commonwealth Scholarship Commission in the United Kingdom, and the study leave granted by Jahangirnagar University, Bangladesh. I am indebted to both.

GLOSSES

Each sentence in this work contains a gloss in English. In certain cases, literal translations have been given to facilitate a clear understanding of the syntax and semantics of the Bengali sentence and this occasionally leads to an infelicitous English gloss. In a set of synonymous examples, only the first sentence has been glossed. In important examples, the words (except proper names) in the transliterated sentence and its English gloss have been numbered on a word for word correspondence.

SUMMARY

This work is the first investigation into the syntax of pronouns in Standard Colloquial Bengali (SCB) using a transformational generative model of syntax. The model used is an adaptation of that proposed in Chomsky (1965) and Fillmore (1968a), and resembles the model developed in Stockwell et al (1973). One claim made in this work is that not everything which can be used to refer deictically or anaphorically is a pronoun. This work makes a clear distinction between pronouns and other elements which can function pronominally. Pronominalization, in a broad sense, has been taken to mean the derivation of pronouns in syntactic structures. This can be either pronouns introduced in the underlying structure or pronouns introduced in a transformational derivation. The process of pronominalization has been broken mainly into pronominalization proper, reflexivization and relativization. This work proposes a transformational treatment of anaphoric pronouns, but attempts to show that not all anaphoric pronouns arise due to transformational reduction of coreferential noun phrases. The basic condition for the application of any sort of pronominalization rule is coreference, which has been treated in a Chomskyan (1965) fashion. This work deals also with so-called 'sentence pronominalization', and 'reciprocal pronominalization'. It shows that the concept of 'sentence pronominalization' is untenable, and reciprocal structures in Bengali do not involve pronominalization. Moreover, this work shows that head noun deletion under lexical identity is a sort of, nonanaphoric, pronominalization.

CONTENTS

ACKNOWLEDGEMENTS

GLOSSES

SUMMARY

CHAPTER 1: <u>GENERAL INTRODUCTION AND BASE RULES</u>	1
1.0 The Model	1
1.1 Types of Rules	5
1.1.1 Deep and Surface Structure Constraints	6
1.1.2 Morphophonological Component	6
1.2 THE BASE RULES	6
1.2.1 Notes to the Base Rules	8
1.3 VERB	12
1.3.1 VERB ROOT (VB)	13
1.3.2 Tense and Aspect	21
1.3.3 AUX	23
1.4 INTERROGATIVE	27
1.5 NEGATIVE	29
1.6 SUBJECT-AUX CONCORD	39
1.7 AUX-MOVEMENT	41
CHAPTER 2: <u>NOUN PHRASES</u>	43
2.0 Introduction	43
2.1 Subcategorization of Nouns	43
2.2 General Characteristics of Noun Phrases	46
2.3 Number	49
2.3.1 Plurality by Plural Markers	50
2.3.2 Plurality by Quantifiers	52
2.4 Underlying Structures of Noun Phrases	54

2.4.1	NP --> N	54
2.4.2	D(eterminer)	55
2.4.3	SPEC --> QUANT (CL)	55
2.4.4	SPEC --> PL	63
2.4.5	NP --> (DEF)(SPEC)N	66
2.4.6	NP --> (DEIC)(SPEC)N	71
2.4.7	PART	74
2.4.8	NP --> CONJCO NP NP*	75
2.4.9	NP --> S	78
2.4.10	NP --> D N S	80
2.5	Summary of the Major Rules Discussed	81
 CHAPTER 3: <u>CASES AND CASE MARKERS</u>		 84
3.0	Introduction	84
3.1	Case Grammars	85
3.2	CASES	90
3.2.1	AGT and INS	91
3.2.2	DAT and NEUT	95
3.2.3	LOC and ESS	97
3.3	An Outline of Constituent Order in Bengali	103
3.4	Subject Marking and Subject Placement	107
3.4.1	Active Subject Marking and Placement	111
3.5	Object Marking and Objectivalization	113
3.6	CASE MARKERS	121
3.6.1	CAM for AGT	122
3.6.2	CAM for DAT and NEUT	123
3.6.3	CAM for INS	125
3.6.4	CAM for LOC and ESS	127

CHAPTER 4: <u>PRONOUNS AND PRONOMINALIZATION</u>	133
4.0 Introduction	133
4.1 Types of Bengali Pronoun	134
4.1.1 Personal Pronouns	134
4.1.2 Abstract Pronoun	136
4.1.3 Interrogative Pronouns	136
4.1.4 Locative and Temporal Pronouns	137
4.1.5 Deictics and Pronouns	137
4.2.0 Forms and Feature Specifications of Pronouns	143
4.2.1 Personal Pronouns	143
4.2.2 Interrogative and Indefinite Pronouns	148
4.2.3 Abstract, Locative and Temporal Pronouns	151
4.3 Pronominalization: Some Approaches and Problems	152
4.3.1 Our Approach in Outline	160
4.3.2 What is Pronominalizable?	166
4.3.3 Pronominalization and Cases	167
4.3.4 Coreferentiality versus Definitization	173
CHAPTER 5: <u>PRONOMINALIZATION PROPER</u>	175
5.0 Introduction	175
5.1 Antecedent and Anaphora: Rules of Coreference	177
5.1.1 Coreference in Coordinate Conjoined Structure	178
5.1.2 Coreference in Subordinate Conjoined Structure	179
5.1.3 Coreference in Complex Structures with Direct Discourse Constituent Sentence	181
5.2 Antecedent and Anaphora: Hierarchy of Antecedents	186
5.3 Application of the Pronominalization Proper Rule	188
5.3.1 Application of the Pronominalization Proper Rule in [-HUM] and [-ANI] Noun Phrases	196

5.4	Forward and Backward Pronominalization	202
5.4.1	Constraints on Pronominalization	204
5.5	Pronouns, Definite Descriptions and Epithets	212
5.6	Pronominalization in Coordinate Conjoined Structure	214
5.6.1	Pronominalization in Coordinate Conjoined Structures with Coordinate Noun Phrases	219
5.6.2	Pronoun Conjunction	221
5.7	Pronominalization in Subordinate Conjoined Structure	228
5.8	Pronominalization in Complex Structures with Direct Discourse Complement Clause	231
5.8.1	Indirect Discourse Formation and Pronoun Symmetry	234
5.9	The Derivation of Anaphoric Locative Pronouns	244
5.10	The Derivation of the Abstract Pronoun <u>tā</u>	248
5.11	Summary of the Major Rules Discussed	249
 CHAPTER 6: <u>REFLEXIVIZATION</u>		 253
6.0	Introduction	253
6.1	Reflexivization and the Simplex Structure Constraint	254
6.2	Reflexive Pronouns in Bengali	262
6.3	Reflexivization is a Partly Optional Rule	264
6.4	Application of the Reflexive Rule	266
6.5	Reflexivization and Coordination	269
6.5.1	Reflexivization in Coordinate Noun Phrases and Derived Reflexives: Pronoun Conjunction	270
6.6	Reflexivization versus Emphasis	277
6.6.1	Possessive Emphatic <u>nij</u>	279
6.6.2	Reflexivization, Emphasis and the 'Picture Nouns'	280

CHAPTER 7: <u>RECIPROCAL STRUCTURES</u>	289
7.0 Introduction	289
7.1 The Semantics of Reciprocity	290
7.2 Constraints on Reciprocity	294
7.3 Reciprocal Items in Bengali: Are They Pronouns?	298
7.4 Approaches to Reciprocity and Problems	307
7.5 The Derivation of Reciprocal Structures	310
CHAPTER 8: <u>RELATIVIZATION</u>	313
8.0 Introduction	313
8.1 The Relative Deictic <u>ye</u> and the Relative Pronouns	317
8.2 Restrictive and Nonrestrictive Relative Clauses	325
8.2.1 Restrictive Relative Clauses in Bengali	325
8.2.2 Proper Nouns, Unique Reference and Restrictive Relativization	332
8.2.3 Nonrestrictive Relative Clauses in Bengali	334
8.3 Containing Noun Phrases and Restrictively Relativizable Noun Phrases (NP-Rel's)	336
8.3.1 Containing Noun Phrases	338
8.3.2 Relativizable Noun Phrases (NP-Rel's)	339
8.3.3 How Much 'Load'?	341
8.4 The Derivation of Restrictive Relative Clause Sentences	350
8.4.1 Elaborate Restrictives and 'Anglicized' Restrictives	362
8.4.2 Quantifiers, Generic Noun Phrases and Restrictive Relativization	366
8.5 The Derivation of Nonrestrictive Relative Clause Sentences	369

CHAPTER 9: <u>IDENTICAL HEAD NOUN DELETION</u>	375
9.0 Introduction	375
9.1 Direction of Deletion	376
9.2 Identical Head Noun Deletion	379
9.3 Identical Modifier and Head Noun Deletion	382
9.4 Deletion of Identical Head Noun with Nonidentical modifiers	385
 CHAPTER 10: <u>SENTENCE PRONOMINALIZATION</u>	 386
10.0 Introduction	386
10.1 The Abstract Pronoun <u>tā</u>	387
10.2 Constraints on Sentence Pronominalization	389
10.3 An Outline of Complementation in Bengali	391
10.4 Sentence Pronominalization	407
 CHAPTER 11: <u>THE LEXICON</u>	 421
11.0 Introduction	421
11.1 The First Lexicon	422
11.1.1 Lexical Insertion Rule	423
11.1.2 Forms of Lexical Entries	425
11.1.3 Feature Specification	426
11.1.4 Redundancy Rules	428
11.1.5 Sample First Lexicon	429
11.2 The Second Lexicon	435
11.2.1 Sample Second Lexicon	437
11.3 The Morphophonological Component	442

APPENDIX: A: <u>ON THE BENGALI LANGUAGE</u>	444
A.0 Introduction	444
A.1 The Phonemes of Bengali, Orthography and Transliteration	445
A.2 <u>Sādhu</u> and <u>Calit</u> Bengali	447
A.3 Works on Bengali	452
APPENDIX: B: <u>ABBREVIATIONS</u>	457
<u>BIBLIOGRAPHY</u>	459
A. Theoretical	459
B. On Bengali	471

GENERAL INTRODUCTION AND BASE RULES1.0 The Model.

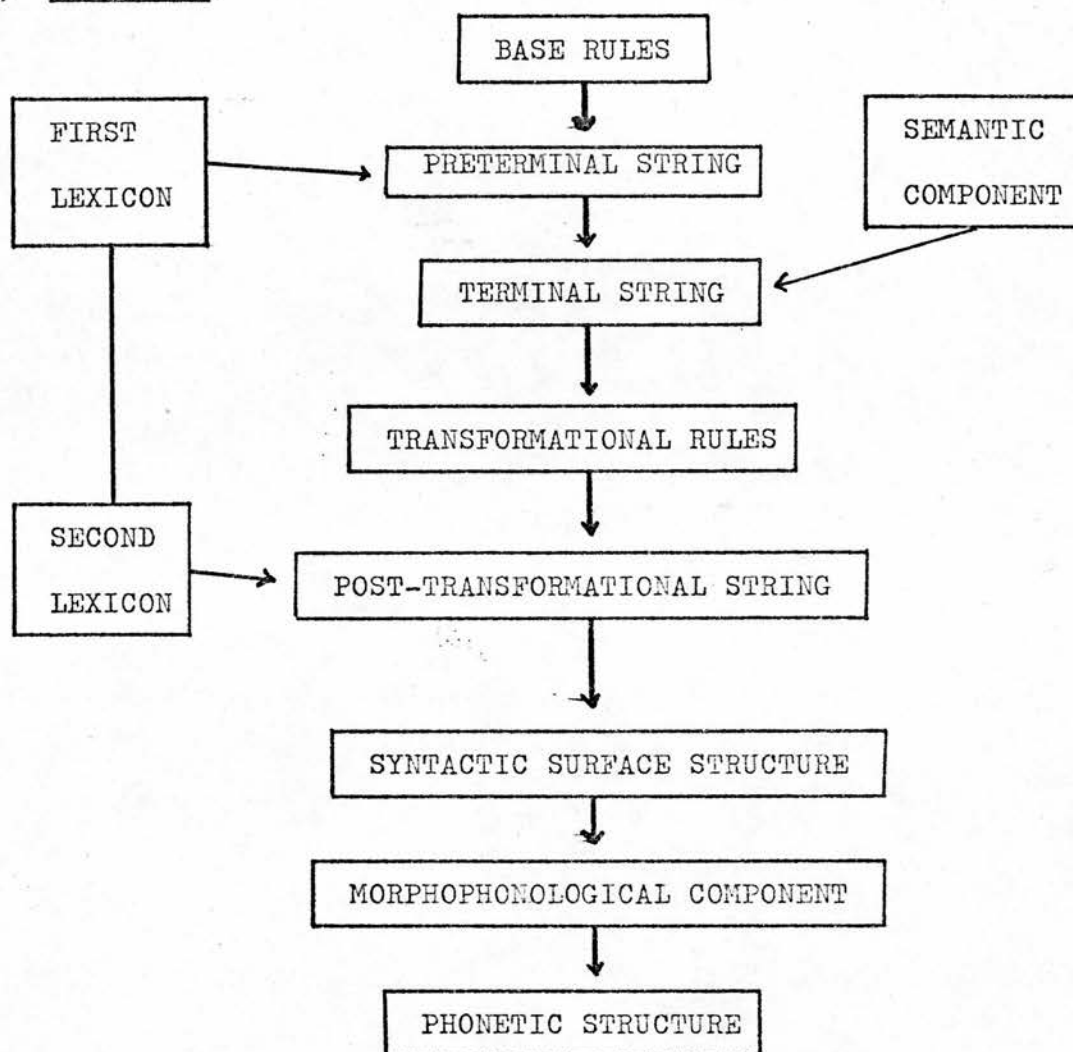
The model used in this work can be called a transformational case grammar, transformational in the sense of Chomsky (1965, 1970a, b) (but cf., McCawley (1968b, 1970, 1971), Lakoff (1970a, 1971a), Postal (1971b)), and case grammar in the sense of Fillmore (1968a, 1969, 1971) (but cf., Anderson (1971)). It is an attempt to integrate Fillmore's and Chomsky's hypotheses about the underlying structure of natural languages for the purpose of describing Bengali. Such an attempt at integrating Fillmore's and Chomsky's hypotheses has already been made in Stockwell et al (1973), which analyses the syntax of English. In the past few years transformational generative theory has seen the major development which is now known as 'Generative Semantics'. And Anderson (1971) has developed his localistic theory of case, which has much in common with 'Generative Semantics' in spirit. Although we share many ideas with the generative semanticists, we have not adopted that model for two main reasons: (a) its very powerful transformational component can be easily misused, and (b) no well-defined 'Generative Semantics' model has been built up yet. Anderson's (1971) localistic case grammar is also undergoing development, by Anderson and his followers. We have adopted Fillmore's and Chomsky's hypotheses because (a) these two hypotheses go together comfortably; and (b) they are much more well-defined than other competing models.

The model we have adopted and developed is justified for various reasons. A case grammar model is best suited for Bengali

for both language universal and language particular reasons. A case grammar model can express deeper generalizations than a Chomskyan model (cf., Fillmore (1968a), Anderson (1971)). It can take care of the case markers systematically without any ad-hoc device, and its unordered base rules allow a simpler description of Bengali, a language having an almost-free constituent order in a simplex sentence structure. The Chomskyan lexicalist hypothesis is helpful for practical reasons. The syntax of Bengali has not been extensively described in any framework, contemporary or otherwise. So we have wanted to use a more modest theory for an adequate description of Bengali. The model used is rather eclectic in nature. We have borrowed from different competing hypotheses in order to develop a composite and consistent model which is well suited for Bengali.

This model is not a faithful imitation of any of the current models, though it has many similarities with those of Chomsky (1965, 1970a, b), Fillmore (1968a, 1971) and Stockwell et al (1973). This model has the closest similarity to the framework developed in Stockwell et al (1973), which is an almost unique attempt to put a theory into practice. The base of our grammar is more abstract than that of a late Chomskyan grammar (1965, 1970a, b). We have maintained the Katz-Postal hypothesis (1964) that transformations are meaning preserving. The model used in this work can be shown diagrammatically as (1) (see below).

A Chomskyan grammar has three major components (cf., Chomsky (1965, 16-17), Lyons (1970b), Bach (1971): (a) a syntactic component, (b) a semantic component, and (c) a phonological component. The syntactic rules of the syntactic component generate the sentences of the language and assign to each both an

(1) The Model.

underlying phrase marker and a derived phrase marker.

The underlying phrase marker represents the deep or underlying structure, and the derived phrase marker represents the syntactic surface structure of the sentence. The semantic interpretation of the sentence is derived from its deep structure by interpretive semantic rules, and the phonetic structure is derived from the surface structure by interpretive phonological rules. The Chomskyan semantic and phonological components are interpretive as against the syntactic component, which is generative.

In a Chomskyan grammar the main component is the syntactic

component, which is further divided into two subcomponents:

(a) the base subcomponent and (b) the transformational subcomponent. The base subcomponent generates the deep (underlying) structures and the transformational subcomponent maps the deep structures onto surface structures through intermediate structures. The base subcomponent itself is divided into two parts: (a) the categorial subcomponent and (b) the lexicon. The categorial subcomponent consists of context-free phrase structure rules and complex symbol forming rules; and the lexicon contains the lexical items of the language.

Our model is Chomskyan in organization. Its main difference with a Chomskyan grammar lies in its categorial rules, which make a different claim about the underlying structure of natural languages. As can be seen in (1) our model has a base component and a transformational component. The base of the grammar is divided into two parts: (a) a categorial component consisting of phrase structure rules and complex symbol forming rules, which constitute the base rules of the grammar, and (b) a lexicon, called the first lexicon. We have another lexicon, called the second lexicon. The first lexicon contains those lexical items that can be inserted to a preterminal string, and the second lexicon contains those items that are derived by the application of the transformational rules. The semantic and phonological components are interpretive as in a Chomskyan grammar. The transformational rules of the transformational component generate derived structures through intermediate structures.

Let us show how the grammar works. The base rules of the grammar generate preterminal strings with dummy symbols, Δ , and feature complexes, and sometimes, abstract markers like Q , NEG etc., (cf., § 1.2). A terminal string is generated by first lexical insertion to the preterminal string, and this constitutes the deep

or underlying structure of a sentence. Transformational rules operate on this structure generating post-transformational strings via intermediate structures. The second lexical insertion rule inserts lexical items derived by transformational rules to the post-transformational string, and a syntactic surface structure generates. The surface structure passes through the morphophonological component in order to generate a phonetic structure.

1.1 Types of Rules.

This grammar contains three major types of rules: (a) base rules, (b) transformational rules, and (c) lexical redundancy rules. The base rules and the first lexicon form the base of the grammar.

A. Base Rules.

Base rules are context-free phrase structure (rewrite) rules of the form $X \rightarrow Y$, where X is a single non-null symbol and Y is a non-null string of symbols, and $Y \neq X$. After the initial symbol S is rewritten any rule applicable may be applied until all symbols are terminal. We have also used complex symbol forming rules in the base for the AUX (cf., Chomsky (1965, 82-90)). A tree or P(hrase)-marker that is formed by the base rules and the first lexical insertion is called a deep or underlying P-marker or structure.

B. Transformational Rules.

Transformational rules (T-rules) transform underlying P-markers into derived P-markers. These rules have the power of restructuring trees. Each T-rule consists of (a) a structure index (SI), (b) a structure change (SC), and often, (c) a set of conditions.

C. Lexical Redundancy Rules.

These rules operate on lexical items before they are introduced into the tree. They help to reduce the complexity of the feature

specification in the complex symbol associated with a lexical item (cf., § 11.1.4; § 11.1.5).

1.1.1 Deep and Surface Structure Constraints.

Chomsky (1965, 138-139) wants to use transformations as filters, that is, transformations will rule out ill-formed underlying trees. Perlmutter (1971) has shown that transformational filtering of ill-formed deep structures is not possible in many cases. He claims that the grammar needs deep and surface structure constraints to rule out unacceptable structures. Deep structure constraints disallow ill-formed deep structures, and the surface structure constraints are surface well-formedness conditions. We assume that both types of constraints will be needed in this grammar, but will not go into the details of such constraints in this work. Here we will mention one deep structure constraint that is required in infinitival complementation (cf., § 1.3.1.B.1; § 10.3). The verbs like thāk: 'Continue' and cā: 'Want' require coreferential matrix and constituent subjects for te-infinitivalization.

1.1.2 Morphophonological Component.

All syntactic surface structures pass through this component, whose rules generate the phonetic structure of sentences. Although it is beyond our scope to discuss these rules in detail, we will try to show, when necessary, how such rules produce phonetic structures (cf., § 11.3).

1.2 THE BASE RULES.

$$\text{RULE 1} \quad S \text{ ----} \rightarrow \neq \left\{ \begin{array}{lll} \text{CONJCO} & S & S^* \\ \text{CONJSUB} & S & S \\ & M & P \end{array} \right\} \neq$$

RULE 2 M -----> ({ ^Q }) (NEG) AUX (ADV)
 IMP

RULE 3 P -----> V (ESS)(INS)(LOC)(NEUT)(DAT)(AGT)

RULE 4 V -----> VB

RULE 5 a. AGT -----> NP CAM

b. Similarly DAT, NEUT, LOC, INS and ESS.

RULE 6 NP -----> { CONJCO NP NP* }
 S
 (D) N (S)

RULE 7 D -----> (PART)({ DEF })(SPEC)
 DEIC

RULE 8 DEIC -----> (DEM)(ORD)

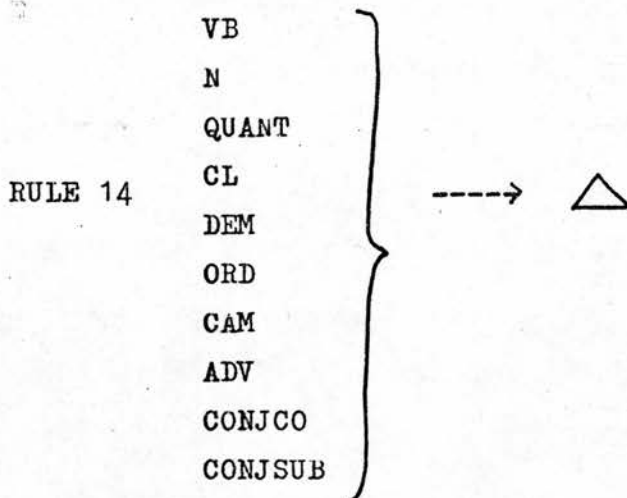
RULE 9 SPEC -----> { QUANT (CL) }
 PL

RULE 10 AUX -----> CS

RULE 11 [+AUX] -----> { [+PRES] }
 [+PAS]
 [+FUT]
 [+HAB P]

RULE 12 { [+PRES] } -----> { [+sim] }
 [+PAS] { [+prog] }
 [+perf]

RULE 13 { [+FUT] } -----> [+sim]
 [+HAB P]



1.2.1 Notes to the Base Rules.

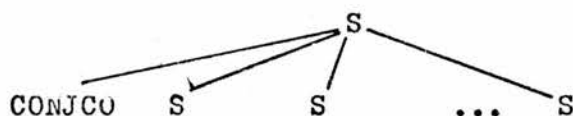
BASE RULE 1:

A. The symbols M and P are taken from Fillmore (1968a). He describes the M (Modality) component as the bearer of 'such modalities on the sentence-as-a-whole as negation, mood and aspect', and the P (Proposition) as 'a tenseless set of relationships involving verbs and nouns.'

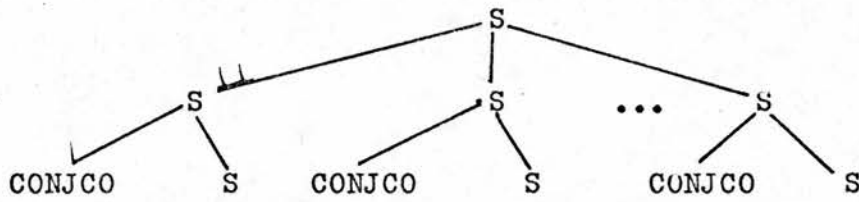
B. '#/' is the sentence boundary symbol.

C. The rule CONJCO S S* handles coordinating conjunction, and the rule CONJSUB S S handles subordinating conjunction. CONJCO may be filled by any of the following items each having the feature [+CONJCO] : o, ār, eban: 'And', bā, athabā: 'Or', and kintu: 'But'. We will adopt the conjunction spreading schema of Lakoff and Peters (1966), which distributes CONJCO. But we differ from them in respect of derived structures of coordinate conjoined structures. For example, they will derive the structure (b) from the underlying structure given below in (a), but we will derive the structure (c) by the schema.

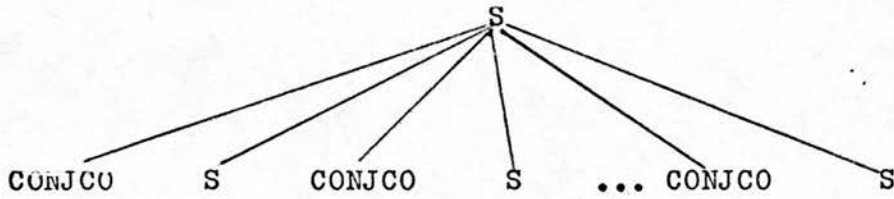
(a)



(b)



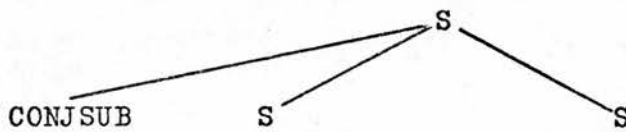
(c)



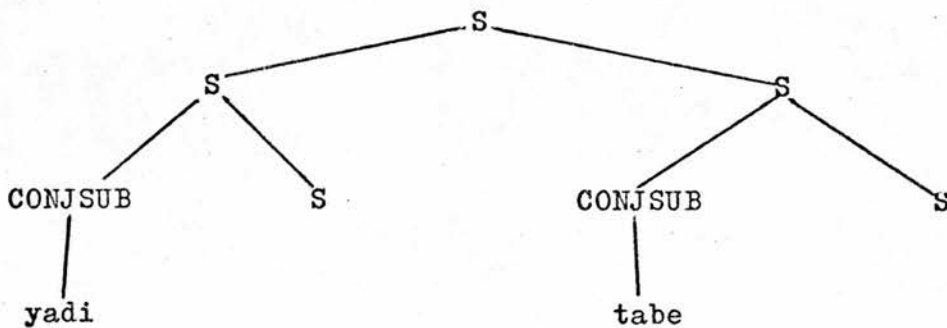
D. The asterisk '*' in the rule $\text{CONJCO } S \ S^*$ is an iteration symbol indicating two or more occurrences of S. This rule says that for a coordinating conjunction at least two sentences must be chosen, and if necessary, an indefinite number of sentences may be chosen.

E. The rule $\text{CONJSUB } S \ S$ is for subordinating conjunction. Only two sentences are conjoinable by this rule. The CONJSUB of the first conjunct will be filled by yadi: 'If', and that of the second conjunct by tabe: 'Then'. This rule will generate the structure (b) from (a) by a conjunction spreading schema.

(a)



(b)



BASE RULE 2:

A. Q, IMP and NEG in this rule are interrogative, imperative and negative markers respectively, which trigger the respective transformations. The AUX, which is subsequently developed as a complex symbol, is the carrier of tense and aspect. There is a transformation which later copies the features of person and grade from the subject to the complex symbol for the AUX (cf., § 1.6). ADV stands for adverb, which we do not consider in this work.

BASE RULE 3:

A. This rule develops the P as a verb and an array of six cases. The cases are ESS(ive), INS(trumental), LOC(ative), NEUT(ral), DAT(ive), and AGT(Agent). We have discussed the problems with cases in Chapter 3.

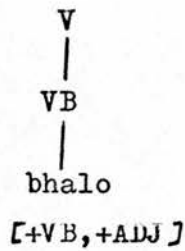
BASE RULE 4:

A. This rule develops the V as VB (verb root). Two types of lexical item can be inserted under VB. They are the verb roots having the features [+VB,-ADJ] , and the adjectives having the features [+VB,+ADJ] . This is in the line of Lakoff (1970a) and Stockwell et al (1973).

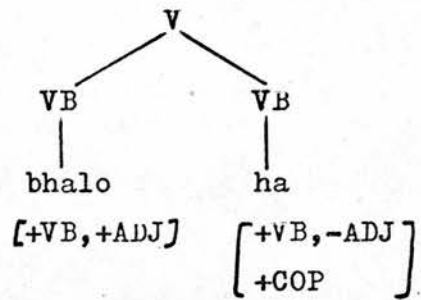
B. The copulative verbs ha, āch, thāk, ra etc., (cf., Ferguson (1972)) arise in two ways: (a) they are lexically inserted as underlying verb roots when ESS, and sometimes LOC and NEUT occur in the proposition, and (b) they are transformationally inserted when an adjective occurs as the head of the proposition. When an adjective occurs as the head of a proposition, an appropriate copulative verb will be inserted to the V as the right sister of the adjective. But in certain cases the copulative verb can be later deleted. Thus after the insertion of a copulative verb a structure

like (a) will be transformed into one like (b).

(a)



(b)



C. A later transformation, AUX MOVEMENT will move the AUX out of its underlying position and adjoin it as the right daughter of the V in order to derive the surface verb forms.

BASE RULE 5:

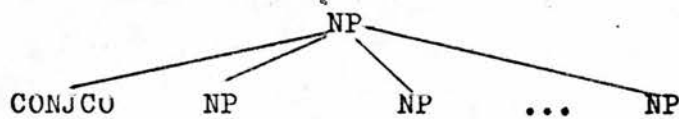
A. This rule develops each case as NP CAM. NP stands for a noun phrase and CAM for a case marker.

B. we have dealt with the cases and the case markers in chapter 3.

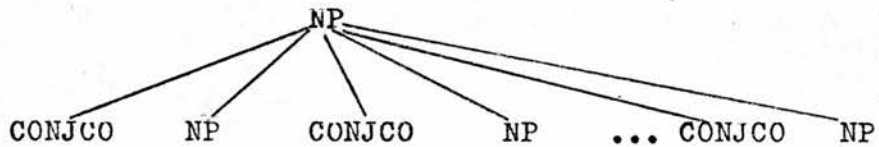
BASE RULE 6:

A. The rule CONJCO NP NP* accounts for underlying phrasal conjunction. By a conjunction spreading schema a structure like (a) will transform into (b).

(a)



(b)



B. CONJCO here can be filled by o, ār, eban̄: 'And', and bā, athabā: 'Or', but not by kintu: 'But', which is exclusively used as a sentence conjunction.

C. The rules NP → S and NP → D N S are for complementation

(cf., § 10.3). When an NP is rewritten as D N or N, a simple noun phrase structure generates (cf., Chapter 2). In this rule N stands for a noun and D for a determiner.

BASE RULE 7-9:

A. We have discussed D(eterminer), DEF(inite), DEIC(tic), SPEC(ifier), DEM(onstrative), ORD(inal), QUANT(ifier), CL(assifier) and PL(ural) in Chapter 2.

B. PART has been left unexpanded as we will not be considering its syntax exhaustively (but cf., § 2.4.7).

BASE RULE 10-14:

A. The rule 10 rewrites the AUX as a complex symbol, and the rules 11-13 form the complex symbols. PRES, PAS, FUT and HAB P in these rules stand respectively for present, past, future and habitual past tense, and sim, prog and perf stand respectively for simple, progressive and perfective aspect. The divisions of tense and aspect are formal (cf., § 1.3.2; § 1.3.3).

B. The rule 14 inserts a dummy symbol, Δ , to the categories listed in the rule. Thus we arrive at the preterminal string, where the first lexical insertion takes place.

1.3 VERB.

Our base rule 3 expands the P as a V and an array of cases, and the base rule 4 rewrites the V as VB. The Bengali verb forms, symbolized as V, can be easily segmented into two parts: the first part is the verb root, symbolized as VB, and the second part is the AUX, in our terminology, which carries tense, aspect and concord with the subject. Consider the verb forms in (2).

- | | |
|--------------------------|-------------------------|
| (2) kar-i: Do-PRES-sim-1 | kar-chi: Do-PRES-prog-1 |
| kar-echi: Do-PRES-perf-1 | kar-lām: Do-PAS-sim-1 |

kar-chilām: Do-PAS-prog-1

kar-echilām: Do-PAS-perf-1

kar-bo: Do-FUT-sim-1

kar-tām: Do-HAB P-sim-1

We see in (2) that each verb form (V) is segmentable into two parts as kar-i, kar-chi, kar-echi and so on. The first part in each verb form is a VB and the second part is an AUX. Bengali verb roots generally do not undergo any change for tense, aspect and concord (except that some verb roots undergo morphophonological change occasionally when an AUX is suffixed to them), it is the AUX that changes for tense, aspect and concord. The forms of the AUX are inflectional-fusional, and we will assume in this work that they are not further divisible (but cf., § 1.3.3), and they will be supplied from the second lexicon. We will list the verb roots (VB) in the first lexicon, from where they will be inserted to the trees. So we consider that the verb forms (V) are derived forms, which consist of VB and AUX. In the first lexicon, adjectives are also listed as verb roots, but they differ from the verb roots in specification for the feature [ADJ]. The verb roots have the features [+VB,-ADJ] and the adjectives have the features [+VB,+ADJ] (cf., Lakoff (1970a); Stockwell et al (1973)).

1.3.1 VERB ROOT (VB).

Traditional Bengali grammarians (cf., Chatterji (1939)) classify the Bengali verb roots into three major groups: (a) Simple verb roots, (b) Compound verb roots, and (c) Conjunct verb roots.¹

¹ There is another type of verb root, Causative verb roots.

Traditionally they are morphologically derived from the above mentioned verb roots. We will not deal with them here, but it is understood that they require a transformational treatment (cf., Shibatani (1973)).

Although an exhaustive analysis of the Bengali verb roots is not possible here, we will try to show that the traditional classification needs modification. We will show that the so-called conjunct verb roots cannot be considered as verb roots, and that not all so-called compound verb roots are underlying verb roots.

A. Simple Verb Roots.

The verb roots that are not further analysable are traditionally called simple verb roots. Examples of such verb roots are dekh: 'See', bal: 'Say, Tell', nāc: 'Dance', cā: 'Want', kar: 'Do' etc. These are listed in the first lexicon with the features [+VB, -ADJ]. Each of them has a case frame associated with it in the lexicon.

B. Compound Verb Roots.

The verb roots of the form $VB_1 - \left\{ \begin{matrix} e \\ te \end{matrix} \right\} VB_2$ such as kare phel: 'Finish doing', likhte thāk: 'Continue to write' are traditionally referred to as the compound verb roots, because they are composed of two verb roots. It is said that in a compound verb root the first verb root carries the dominant meaning, and the second verb root indicates completion, continuation, inception etc. When a simple verb root is used as the second element of a compound verb root, its basic meaning changes. Only a limited number of simple verb roots can function as the second element of a compound verb root. And there are cooccurrence restrictions between the VB_1 and VB_2 in a compound verb root.

B.1 So-called Compound Verb Roots of the Form $VB_1 - te VB_2$.

Let us see now whether the so-called compound verb roots of the form $VB_1 - te VB_2$ such as karte lāg: 'Start to do', dite cā: 'Want to give', calte pār: 'Be able to move', hāste thāk: 'Continue

to laugh' etc., (cf., Chatterji (1939, 419)) can be considered as underlying verb roots. Consider the following examples.

- (3) a. $\bar{a}mi$ $baste$ $c\bar{a}i$: I want to sit.
 1 2 3 1 3 2
- b. $\bar{a}mi$ $likhte$ $l\bar{a}gl\bar{a}m$: I started (and continued) writing.
 1 2 3 1 3 2
- c. $\bar{a}mi$ $h\ddot{a}tte$ $th\bar{a}kl\bar{a}m$: I continued walking.
 2* 3 3 2

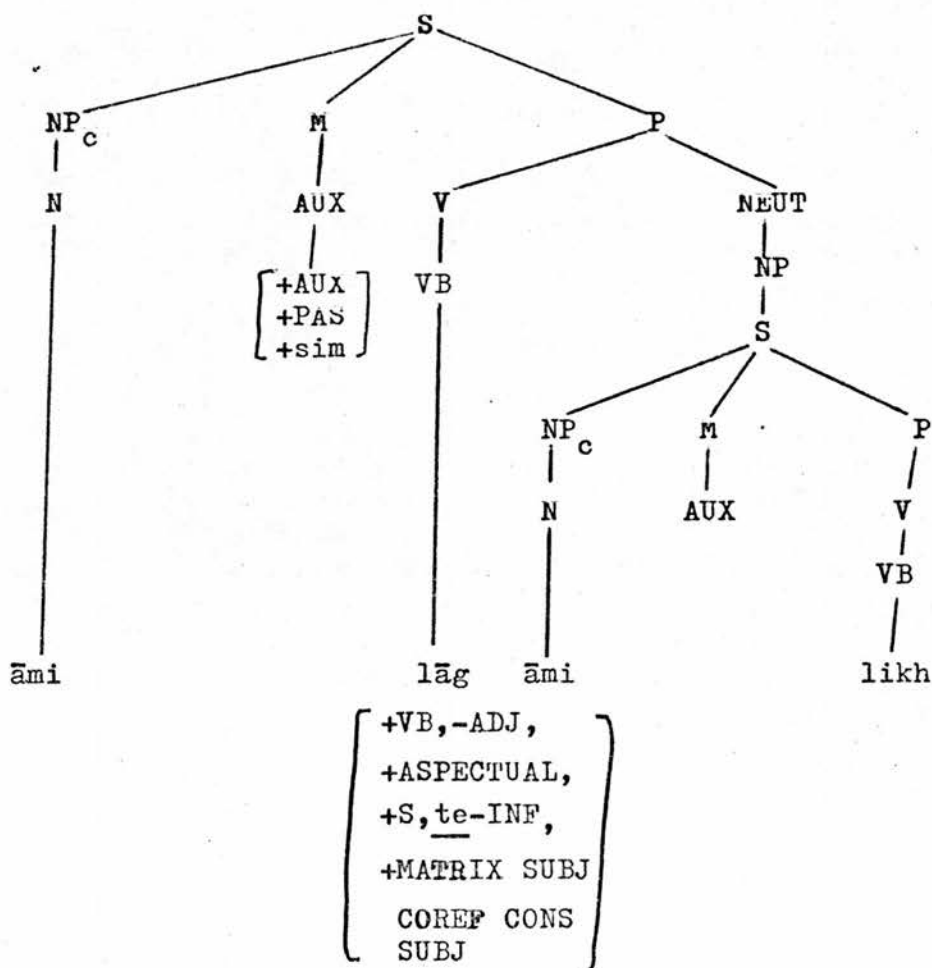
Traditionally baste cā in (3a), likhte lāg in (3b) and hātte thāk in (3c) are considered as single compound verb roots. But we will claim that they are not single verb roots in the underlying structure, and they will not be included in the lexicon in their compound form. The sentences in (3) are not simple, but are complex structures. They involve complementation (cf., § 10.3), and they are related to the intermediate structures in (4) respectively.

- (4) a. $_S [\bar{a}mi\ c\bar{a}i\ _S [\bar{a}mi\ basi]]$: [I want [I sit]].
- b. $_S [\bar{a}mi\ l\bar{a}gl\bar{a}m\ _S [\bar{a}mi\ likhi]]$: [I started [I write]].
- c. $_S [\bar{a}mi\ th\bar{a}kl\bar{a}m\ _S [\bar{a}mi\ h\ddot{a}ṭi]]$: [I continued [I walk]].

In (4) the AUX i in the constituent sentences has been supplied for the sake of readability, but it is not present in the underlying or intermediate structure. In (4) cā and bas, lāg and likh and thāk and hāṭ are elements of different simplex structures. (4a) is an instance of NP-Complementation. We can derive $\bar{a}mi\ c\bar{a}i\ ye\ \bar{a}mi\ basi$: 'I want that I sit' from (4a) by complementizer placement, and this sentence, although stylistically infelicitous, is acceptable and synonymous with (3a). In (4a) the matrix verb is cā, which is marked for optional te-infinitivalization of the constituent sentence, subject to the condition that the matrix and the constituent subjects are coreferential. This will delete the constituent subject under coreference with the matrix subject, and the Subject-AUX

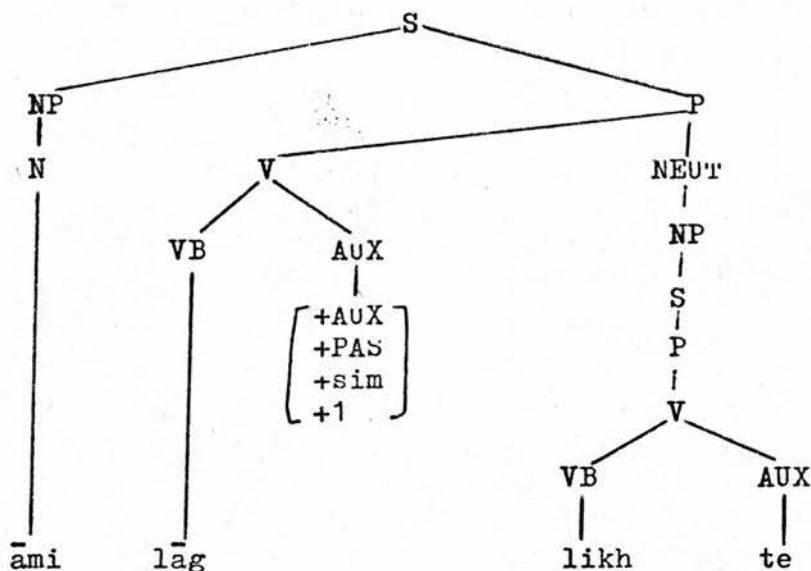
concord rule will not apply in the constituent S; instead te will be inserted to the constituent AUX. Thus we derive āmi cāi baste, which is synonymous with (3a). Subsequently, by placing the finite verb form sentence-finally we will get (3a). This shows that baste cā is not a single verb root. The verbs lāg and thāk are marked for obligatory te-infinitivalization of the constituent S. These are also marked with the deep structure constraint that they require the matrix and the constituent subjects to be coreferential. By obligatory te-infinitivalization we will derive āmi lāglām likhte and āmi thāklām hāṭte respectively from (4b) and (4c). Subsequently, by placing the finite verb forms sentence-finally we will get (3b) and (3c). This shows that likhte lāg and hāṭte thāk are not underlying verb roots. Let us show, for example, how we derive (3b). (3b) has the intermediate structure (5).

(5)



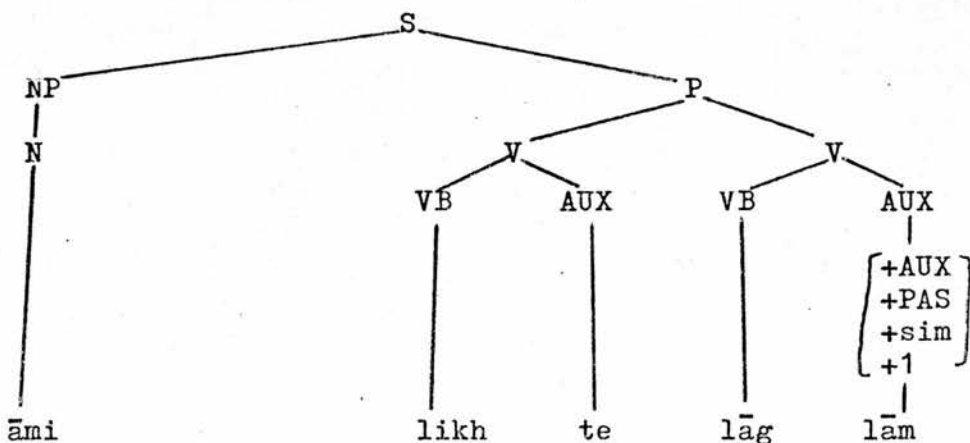
As lāg is specified for te-infinitivalization, the complementizer placement rule will not apply in (5), but the constituent subject will be deleted under coreference with the matrix subject. The Subject-AUX concord rule will not apply in the constituent S; instead te will be inserted to the constituent AUX (cf., @ 10.3). The Subject-AUX concord rule will apply in the matrix S. The rule AUX MOVEMENT, which places the AUX as the right daughter of V, will apply in the matrix and the constituent S. These operations will derive (6) from (5).

(6)



As Bengali usually places the finite verb form sentence-finally, the matrix V will be moved to the sentence-final position; all nodes that do not branch will be pruned, and by the second lexical insertion we will get (7) from (6).

(7)



(7) will generate āmi likhte lāglām, which is (3b). Thus we see that the so-called compound verb of the form VB_1 -te VB_2 is not a single verb root in the underlying structure. They are not, therefore, listed in the lexicon in such a form.

B.2 Compound Verb Roots of the Form VB_1 -e VB_2 .

Now we will consider the compound verb roots of form VB_1 -e VB_2 . Examples of such verb roots are base paṛ: 'Sit down', likhe phel: 'Finish writing', kēde uṭh: 'Cry suddenly' etc., (cf., Chatterji (1939, 419-20)). In these verb roots the basic meaning of the second element is lost. For example, phel in likhe phel indicates completion, but its meaning as a single simple verb root is 'Throw/Drop', and uṭh in kēde uṭh indicates suddenness, but its meaning as a simple verb root is 'Rise'. The second root in such a compound verb root is modal-aspectual in nature. Now consider the examples below.

- (8) a. nāju hese uṭhlo: Nazu laughed (suddenly).
 b. nāju likhe phello: Nazu wrote (quickly).
 c. nāju base paṛlo: Nazu sat down.
- (9) a. nāju hese ballo: Nazu smiled and said.
 b. nāju dekhe elo: Nazu saw and came.

We see above that hese uṭh, likhe phel, and base paṛ in (8) have formal similarity with hese bal, and dekhe ās in (9). But the former are considered as compound verb roots, while the latter are not. The reason is that the former express a single action and the latter do not express a single action. So the examples in (9) can be analysed as serial constructions, and can be paraphrased as (11) respectively, but the examples in (8) cannot be analysed as serial constructions, and cannot be paraphrased as (10) respectively.

- (10) a. *nāju hāslo o uḥlo: Nazu laughed and got up.
 b. *nāju likhlo o phello: Nazu wrote and threw/dropped.
 c. *nāju baslo o paḥlo: Nazu sat and fell.
- (11) a. nāju hāslo o ballo: Nazu smiled and said.
 b. nāju dekhlo o elo: Nazu saw and came.

As (8) and (10) are not related by paraphrase relations, we cannot consider the examples in (8) as serial constructions. So the VB's like hese uḥ, likhe phel etc., should be considered as single compound verb roots in the underlying structure. We will assume that the morphological component generates verb roots of this type, of which the second verb root is modal-aspectual.²

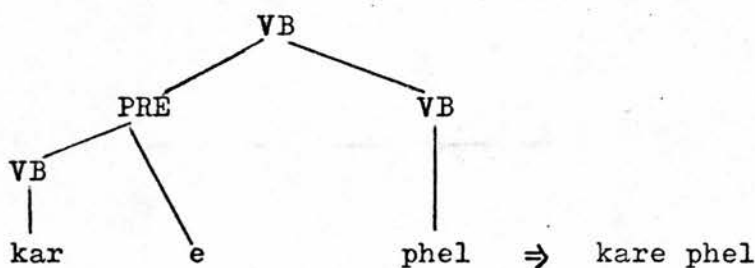
2 Our base rule 4 expands the V as VB. We have not gone any further than this. The lexicon lists the verb roots, which the grammar uses. We assume that the morphological component that produces the compound verb roots contains rules something like the following:

A. VB → (PRE) VB

B. PRE → VB e

So the internal structure of a compound verb root will be something like C:

C.



The rule will mention that there are selectional restrictions between the verb roots in a compound verb. When a compound verb root is inserted in a tree, the lexical insertion rule will take care of its internal structure.

C. Conjunct Verb Roots.

The so-called conjunct verb root root in Bengali has the form $\left\{ \begin{array}{l} \text{N} \\ \text{ADJ} \end{array} \right\} \text{VB}$. Examples of such verb roots are uttar de: 'Give answer', snān kar: 'Take a bath', kaṣṭa pā: 'Feel pain' etc., (cf., Chatterji (1939, 347-50)). It is said that either a noun or an adjective and a simple verb root combine together and form a conjunct verb root. In English verbs like 'joke', 'parise', 'attract' etc., convey an idea in a single lexical item, but the ideas in Bengali are expressed by a combination of nouns and verbs, such as tāmāsā kar: 'Joke', praśaṅsā kar: 'Praise', and ākaraṣaṅ kar: 'Attract'. Consider the examples below.

- (12) a. bādāl tāmāsā karche: Lit., Badal is doing fun.
 2 3 3 2
- b. bādāl snān karche: Lit., Badal is doing bath.
 2 2
- c. bādāl kaṣṭa pācche: Lit., Badal is getting pain.
 2 3 3 2
- (13) a. bādāl bai paṛche: Badal is reading (a) book.
 2 3 3 2
- b. bādāl chabi ākche: Badal is drawing (a) picture.
 2 3 3 2

Traditionally tāmāsā kar, snān kar and kaṣṭa pā in (12) are considered as conjunct verb roots, but bai paṛ and chabi āk are not considered so, although they have a formal similarity. Instead bai is considered as the object of paṛ in (13a), and chabi as the object of āk in (13b). But functionally tāmāsā is in the same relationship with kar in (12a) as bai is with paṛ in (13a). The only difference between them is that tāmāsā is an abstract noun, and bai is a concrete noun. In the so-called conjunct verb root in Bengali the noun is [+ABSTRACT], and that is why it is not considered as an object. But functionally it is an object of the verb with which it forms a conjunct verb root (cf., § 3.5). But we can consider them as

objects, and thus reject the notion of the conjunct verb root in Bengali.

But if we reject them as conjunct verb roots, we may come in conflict with Fillmore's (1968a) claim that each case can occur only once per simplex. Consider the examples below.

- (14) a. se anek $\dot{t}ākā$ $\bar{a}y$ kareche: Lit., He has done earning much
 1 2 3 4 5 1 5 4 2
 money (He has earned much money).
 3
- b. se anek $\dot{t}ākā$ byy kareche: Lit., He has done spending
 4 4
 much money (He has spent much money).

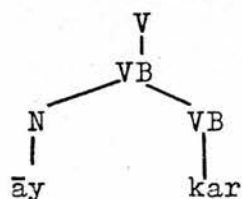
Fillmore (1968a) would consider anek $\dot{t}ākā$ in (14a, b) as Objective case. But what about $\bar{a}y$ and byy above? Each case once per simplex motto will force us either to consider $\bar{a}y$ kar and byy kar as single verb roots, or to postulate another case for $\bar{a}y$ and byy.³

1.3.2 Tense and Aspect.

Bengali has a four way distinction in the category of tense. This is expressed by systematic grammatical contrast. They are Present, Past, Future and Habitual Past. The category of aspect

³ In this work we cannot go into the intricacies of this problem and its solution, and will leave it here simply by pointing to the problem. For the ease of description, we will consider such noun and verb collocations as conjunct verb roots, and the first lexical insertion rule will take care of the internal structure of such verb roots. For example, while inserting $\bar{a}y$ kar to a tree, the lexical insertion rule will show its internal structure as below.

(A)

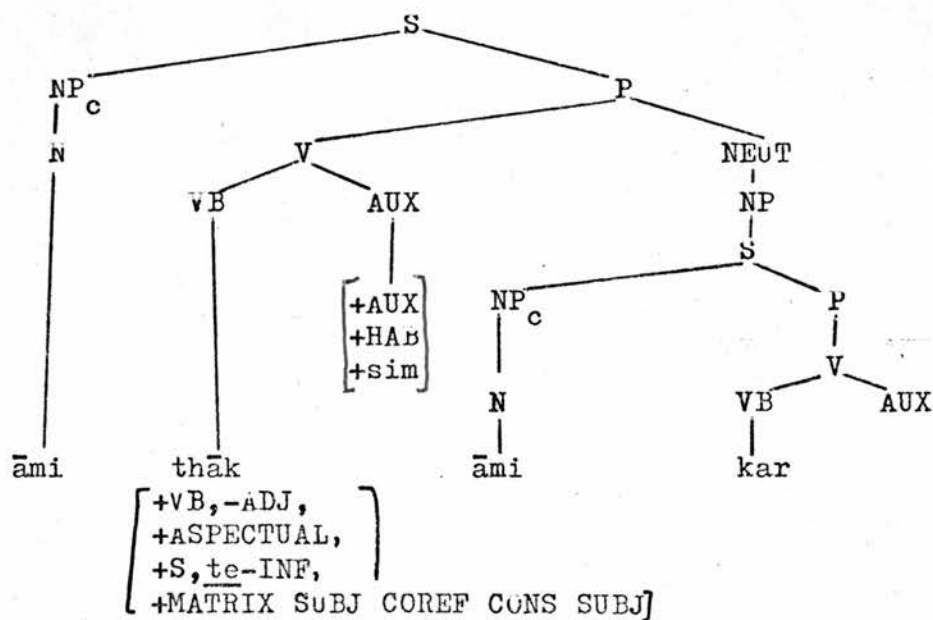


has a three way distinction: simple, progressive and perfective. The simple aspect can combine with any of the four tenses, but the other two can combine with only present and past tenses. Progressiveness and perfectiveness in the future and habitual past tenses are expressed differently from the case in present and past tenses. Consider the examples below.

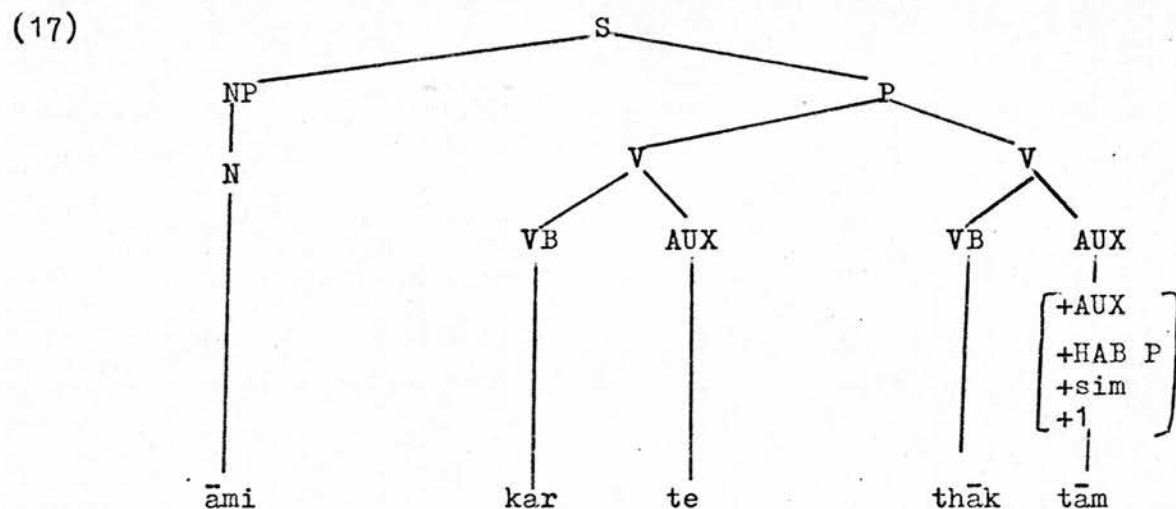
- (15) a. āmi karchi: I am doing.
 1 2 1 2
- b. āmi karchilām: I was doing.
- c. āmi karte thāktām: I used to do.
- d. āmi karte thākbo: I shall be doing.

The verb form karchi in (15a) indicates an action is progress in the present, and karchilām in (15b) indicates an action in progress in the past. In each case the form of the AUX such as chi and chilām, which carry tense, aspect and concord, is suffixed to the VB. But progressiveness (and perfectiveness) in future and habitual past tenses are not expressed similarly. In (15c) karte thāktām indicates progressiveness in the habitual past tense, and in (15d) karte thākbo indicates progressiveness in the future. These are similar to the sentences discussed in § 1.3.1.B.1. (15c, d) are not simple structures, but are complex structures. They are instances of infinitival complementation where obligatory te-infinitivalization has taken place. For example, (15c) has the intermediate structure (16).

(16)



In (16) the constituent subject noun phrase will be deleted under coreference with the matrix subject; and so the Subject-AUX concord rule will not apply in the constituent S. Instead te will be inserted to the constituent AUX. In this way we will get āmi thāktām karte, which is synonymous with (15c). Then by placing the finite verb form sentence-finally and pruning the non-branching nodes we will get (17).



(17) will generate āmi karte thāktām, which is (15c). This shows that progressiveness and perfectiveness in the future and habitual past should be treated differently from the case in the present and past tenses.

1.3.3 AUX.

The AUX is the carrier of tense, aspect, and concord with the subject. In this grammar AUX is a constituent of M in the underlying structure, but it is adjoined as the right daughter of V by a later transformation. The inflectional-fusional nature of the items that are inserted to the AUX from the second lexicon make any further significant segmentation of them difficult. Our grammar generates complex symbols for the AUX in the underlying structure, and a later concord rule copies the person and grade features of

the subject onto the AUX. The lexical forms of the AUX are inserted to the derived trees from the second lexicon. Consider the verb forms in (2), which we reproduce below as (18).

- (18) kar-i: Do-PRES-sim-1 kar-chi: Do-PRES-prog-1
kar-echi: Do-PRES-perf-1 kar-lām: Do-PAS-sim-1
kar-chilām: Do-PAS-prog-1 kar-echilām: Do-PAS-perf-1
kar-bo: Do-FUT-sim-1 kar-tām: Do-HAB P-sim-1

In (18) the VB is kar, and the forms of the AUX are i, chi, echi etc. i indicates [+PRES,+sim,+1], chi indicates [+PRES,+prog,+1], and so on. It is not impossible to decompose them into further segments and show the functions of those segments (see below), but by doing so we should not gain much. We have listed all the forms of the AUX in the second lexicon from where they are inserted to the trees by the second lexical insertion rule.

This method of inserting the forms of the AUX has been used for the sake of simplicity. But we could generate the forms of the AUX by the base rules of the grammar in combination with some morphophonological rules. The following fragment grammar of the AUX shows how it could be done.

(19) Fragment Grammar of the AUX.

A. Base Rules:

a. AUX \rightarrow AS T C

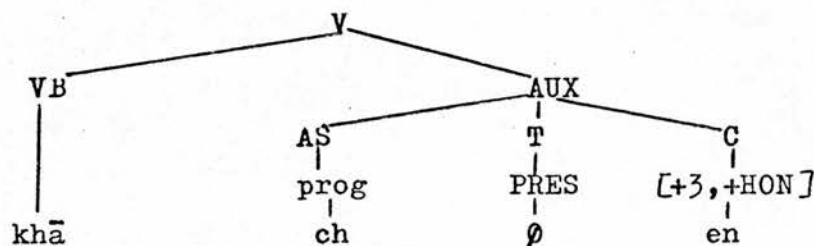
b. T \rightarrow $\left\{ \begin{array}{l} \text{PRES} \\ \text{PAS} \\ \text{FUT} \\ \text{HAB P} \end{array} \right\}$

c. AS \rightarrow $\left\{ \begin{array}{l} \left\{ \begin{array}{l} \text{prog} \\ \text{perf} \end{array} \right\} / \text{---} \left\{ \begin{array}{l} \text{PRES} \\ \text{PAS} \end{array} \right\} \\ \text{sim} \end{array} \right\}$

B. Morphophonological Rules:⁴

- a. $\text{sim} \rightarrow \emptyset$
- b. $\text{prog} \rightarrow \left\{ \begin{array}{l} \text{ch/} \text{---PRES} \\ \text{chi} \end{array} \right\}$
- c. $\text{perf} \rightarrow \left\{ \begin{array}{l} \text{ech/} \text{---PRES} \\ \text{echi} \end{array} \right\}$
- d. $\text{PRES} \rightarrow \emptyset$
- e. $\text{PAS} \rightarrow \text{l}$
- f. $\text{FUT} \rightarrow \text{b}$
- g. $\text{HAB P} \rightarrow \text{t}$

4 These rules will generate grammatical verb forms in non-imperative sentences when the verb root ends in a consonant. But for the verb roots that end in a vowel some special morphophonological rules will be required. Suppose, for example, in a V structure we have the VB khā: 'Eat', and AS(pect) is prog, T(ense) is PRES, and C(oncord) has the features [+3,+HON] copied from the subject; then we will get the following structure after the application of the morphophonological rules:



It will generate *khāchen, which is ungrammatical. So we will require a rule like $\text{V-ch-X} \rightarrow \text{V-cch-X}$, which will apply to the above structure and generate the grammatical form khācchen: 'Eating'.

$$h. [C,+1] \longrightarrow \left\{ \begin{array}{l} i/PRES \text{ ---} \\ \bar{a}m/ \left\{ \begin{array}{l} PAS \\ HAB P \end{array} \right\} \text{ ---} \\ o \end{array} \right\}$$

$$i. [C, \left\{ \begin{array}{l} +2, \\ +3, \end{array} \right\} +HON] \longrightarrow en$$

$$j. [C,+2,-HON] \longrightarrow \left\{ \begin{array}{l} o/PRES \text{ ---} \\ e \end{array} \right\}$$

$$k. [C,+2,+PEJ] \longrightarrow \left\{ \begin{array}{l} is/PRES \text{ ---} \\ i \end{array} \right\}$$

$$l. [C,+3,-HON] \longrightarrow \left\{ \begin{array}{l} e/ \left\{ \begin{array}{l} PRES \\ FUT \end{array} \right\} \text{ ---} \\ o \end{array} \right\}$$

These rules could have been included in the grammar, but have been avoided for the sake of simplicity in the base and in the morpho-phonological component. Our grammar generates complex symbols for the AUX, and the forms of the AUX are supplied from the second lexicon. They are listed in the second lexicon, and each has a CS associated with it. The entries look like this:

$$(20) \quad \begin{array}{ccc} \underline{i} & \underline{chi} & \underline{echi} \\ \left[\begin{array}{l} +AUX \\ +PRES \\ +sim \\ +1 \end{array} \right] & \left[\begin{array}{l} +AUX \\ +PRES \\ +prog \\ +1 \end{array} \right] & \left[\begin{array}{l} +AUX \\ +PRES \\ +perf \\ +1 \end{array} \right] \end{array}$$

1.4 INTERROGATIVE.

Both yes-no question and wh-question are possible in Bengali. We will try to show here, briefly, how this grammar will handle interrogative sentences. We have a dummy interrogative marker Q dominated by M in the underlying structure of yes-no question sentences. It is selected only if the sentence is a yes-no question. In the case of wh-question (ka-question in Bengali) the Q is not posited in the underlying structure. Let us first consider the yes-no question. Consider the following examples.

(21) a. āpani cā khāben?: Will you have tea?
 1 2 3 1 3 2

b. āpani cā khāben ki?: Will you have tea?

These are yes-no question sentences, which can be answered either by 'yes' or by 'no'. The Bengali yes-no question marker is ki. In (21b) the question marker ki is present, but absent from (21a). As no question marker is present in (21a), the verb form khāben will carry an interrogative intonation when (21a) is uttered. The presence of ki in (21b) shows that the sentence is a question. The question marker ki is usually placed sentence-finally in Bengali.⁵

⁵ ki is placed sentence-finally in sentences with the dominant constituent order (SOV) in case of sentential interrogation. If the yes-no question is a constituent interrogation ki may follow, even precede, the interrogated constituent, and can be placed sentence-finally too. Consider the examples below.

(A) a. āpani ki cā khāben?: Will you have tea?

b. āpani cā ki khāben?: Will you have tea?

These sentences are ambiguous in many ways. But when āpani is under the scope of interrogation in (Aa), the sentence means 'Is it you who will have tea?', and when cā is under the scope of interrogation

continued

We have the marker Q in the underlying structure of yes-no question sentences, and it will trigger the interrogative transformation. This transformation will place the marker sentence-finally, which will be later replaced by ki in order to generate a surface sentence. But we will deal with wh-question differently. In a wh-question sentence either some constituent contains an interrogative deictic or the head N is an interrogative pronoun (cf., § 4.2.2). All these elements are generated in the underlying structure in this grammar. Consider the examples below.

- (22) a. kon meyeṭi nācbe?: Which girl will dance?
 1 2 3 1 2 3
- b. ke gān gāibe?: Who will sing?
 1 2 1 2
- c. ke kāke bhālabāse?: Who loves whom?
 1 2 3 1 3 2

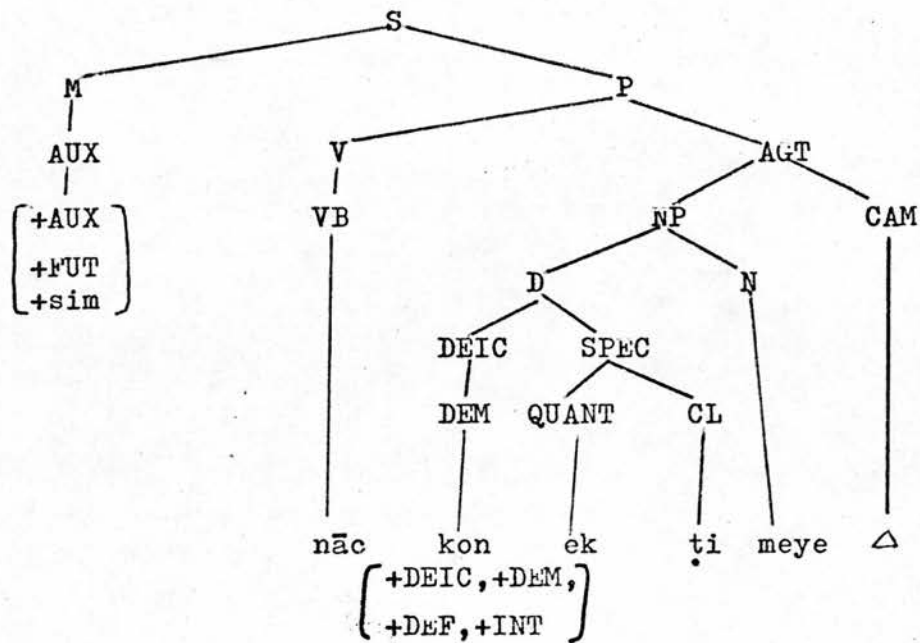
The 'questions' in these examples originate in the noun phrases, and the questions cannot be answered by 'yes/no'. In (22a) the question is due to the presence of the deictic kon, which is inherently interrogative, and in the other examples questions are due the presence of interrogative pronouns, which are also inherently interrogative. We generate all these elements, which indicate interrogation, in the underlying structure, and so it is redundant to postulate the Q in the underlying structure of these sentences.

in (Ab), the sentence means 'Is it tea that you will have?'

Here we will not deal with yes-no constituent interrogation, but assume that this can be handled by positing a constituent interrogation marker with noun phrases in the underlying structure.

For example, (22a) has the underlying structure (23).

(23)



In (23) kon has the feature [+INT(errogative)] , and this indicates that the sentence is a wh-question. There is no Q in the underlying structure (23). Similarly, in the underlying structure of (22b, c) no Q is postulated. These sentences are interrogative due to the inherent feature [+INT] of the underlying interrogative pronouns.

1.5 NEGATIVE.

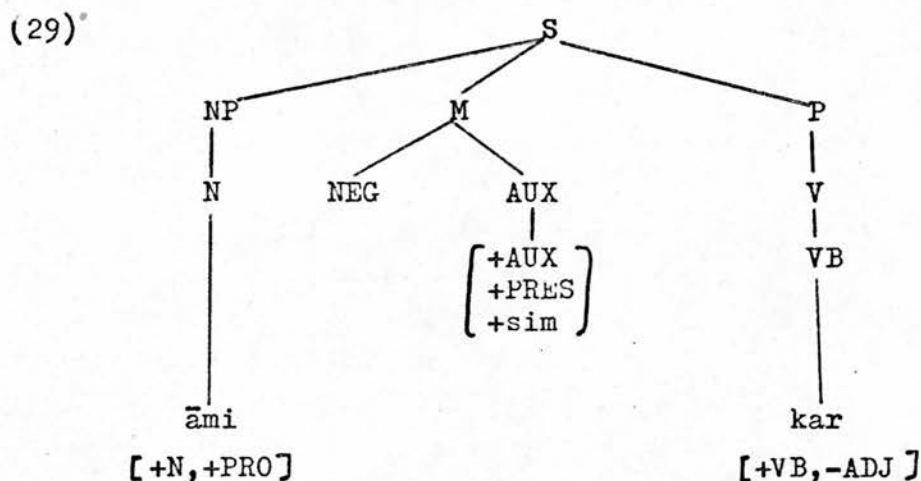
We cannot do justice to a complex process like negation in Bengali in such a short section, and so our discussion will be limited to a brief discussion of sentential negation in Bengali. There are four negative markers in Bengali: na, nā, ni, and nei. The negative marker and the verb of a sentence have a close affinity in both sentential and constituent negation in Bengali, and the negative markers usually follow the verb forms in negative sentences. In our grammar we have an underlying marker NEG for sentential negation. It acts as a conditioner for deep structure constraint and triggers the negative transformation. Consider the following

sentences.

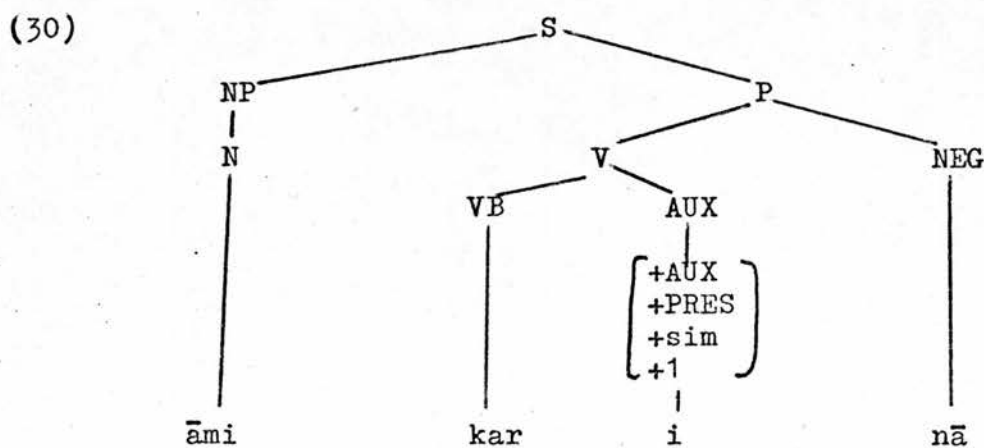
- (24) a. āmi kari nā: I do not.
 1 2 3 1 2 3
- b. āmi karchi nā: I am not doing.
- c. *āmi karechi nā: I have not done.
- d. āmi karlām nā: I did not (but could have done had I wanted).
- e. āmi karchilām nā: I was not doing.
- f. *āmi karechilām nā: I had not done.
- g. āmi kari ni: I did not/have not/had not done.
- h. āmi karbo nā: I shall not do.
- i. āmi kartām nā: I used not to do.
- (25) a. se bhālo nay: He is not good.
 1 2 3 1 3 2
- b. tini bhālo nan: He is not good.
- c. tumi bhālo nao: You are not good.
- d. āmi bhālo nai: I am not good.
- (26) a. {^{se}
 tini } bhālo nei: He is not well.
- b. tumi bhālo nei: You are not well.
- c. āmi bhālo nei: I am not well.
- (27) a. se bhālo chilo nā: He was not good/well.
- b. tini bhālo chilen nā: He was not good/well.
- c. tumi bhālo chile nā: You were not good/well.
- d. āmi bhālo chilām nā: I was not good/well.
- (28) a. bane bāgh nei: There is no tiger in the forest.
 1 2 3 3 2 1
- b. bane bāgh chilo nā: There was no tiger in the forest.

These examples show that nā is used as a negative marker when the sentence has an overt verb form. (24c, f) show that sentences having perfective aspect cannot be negated by nā, in which case the negative marker is ni. The negative marker ni carries tense and aspectual features in itself, and the AUX of such sentences

always have the features [+PRES,+sim]. We assume that this is a case where the underlying NEG acts as a conditioner of deep structure constraint; and neutralizes the features [+PRES,+perf] and [+PAS,+perf] of the AUX into [+PRES,+sim]; and incorporates the underlying aspectual feature [+perf] into itself. Now consider the derivation of the sentences in (24), all of which have the NEG in the underlying structure, but differ one from another in the specifications for tense and aspect. For example, (24a) has the intermediate structure (29).

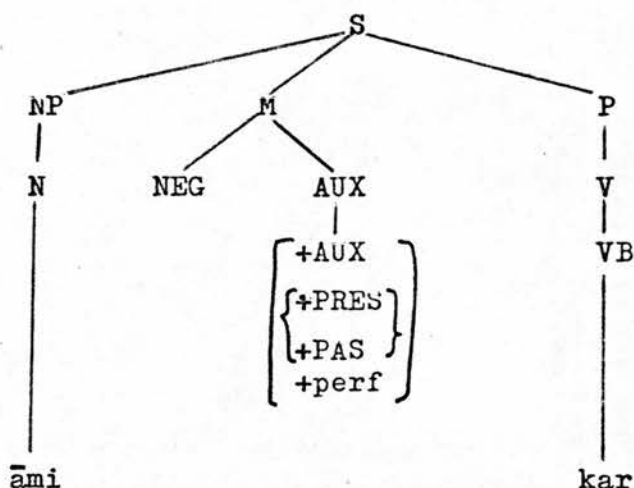


The presence of NEG in (29) indicates that the sentence is a negative one. As kar in (29) is a main verb, and the sentence is not a conditional one, the negative transformation will place the NEG immediately after the V. After AUX Movement and NEG placement, (29) will be transformed into (30).



The second lexical insertion rule has inserted i to the AUX and nā to the NEG. We will get āmi kari nā, which is (24), from (30). Now consider the derivation of (24g), which cannot be derived exactly like (24a), as the negative marker ni carries tense and aspectual features in itself. Should we consider ni as a verb because it bears tense and aspectual features? If we consider ni as an underlying verb, then we have to allow two verbs in a simplex structure, or we have to derive (24g) from a complex underlying structure. Instead we will consider ni as a negative marker with tense and aspectual features that is used to negate sentences in the perfective aspect. We will assume that (24g) has the intermediate structure (31).

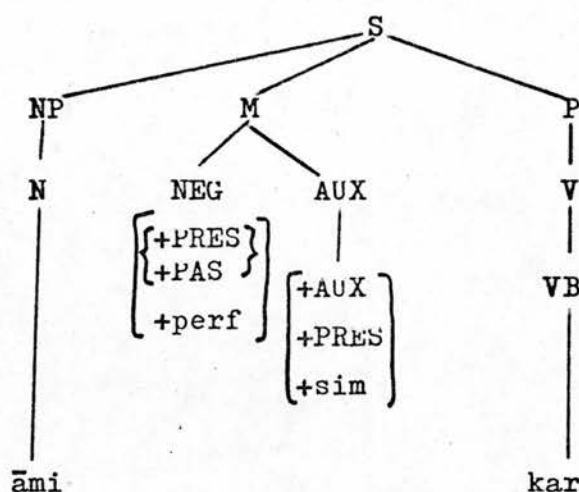
(31)



Now if the negative transformation applies to (31) in its general form (that is, the rule as applies in (29)), we will derive either (24c) or (24f), which are semantically all right, but syntactically unacceptable. The structure (31), where NEG and [{ +PRES } , +perf] cooccur, is a semantically well-formed structure, but it will generate unacceptable syntactic structures. We have said that NEG has a role as a conditioner for deep structure constraint, and as (31), although semantically acceptable, will give rise to ungrammatical syntactic structures, NEG will convert the tense and aspectual features of the AUX invariably into [+PRES, +sim] ; and will incorporate the previous tense and aspectual features into

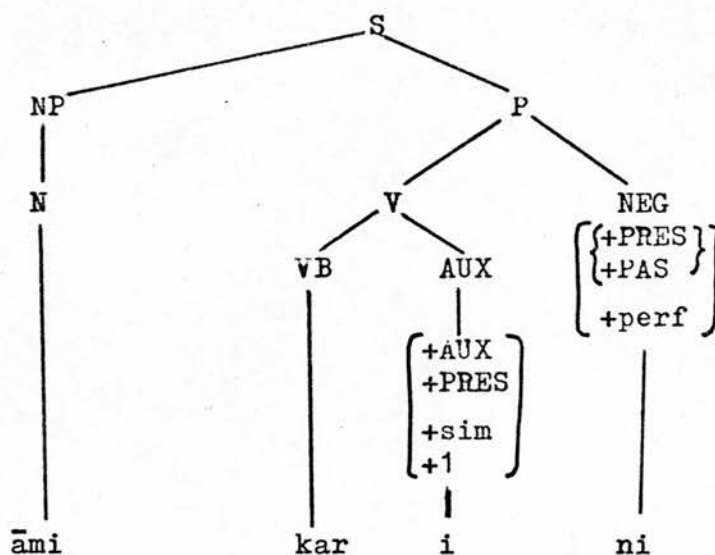
itself. this will convert (31) into (32).

(32)



The negative transformation will now apply to (32) in the usual form, which will place the NEG immediately after the V. After Subject-AUX concord, AUX Placement and NEG placement (32) will be transformed into (33).

(33)

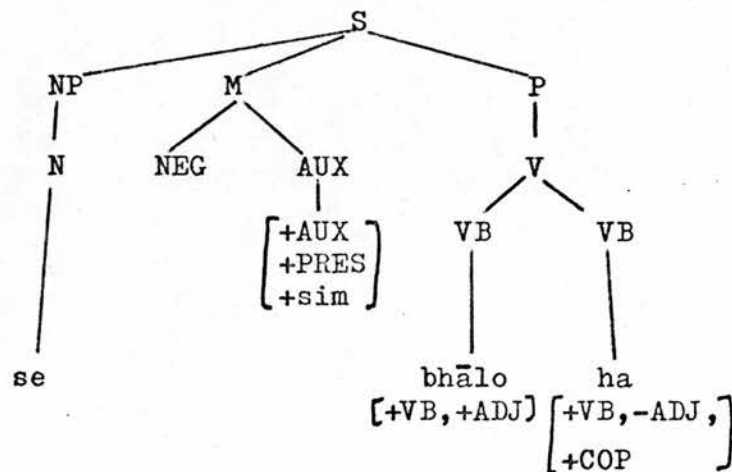


The second lexical insertion rule has inserted i to the AUX and ni to the NEG in (33), which will generate ami kari ni, which is (24g).

Now consider the sentences in (25), where the negative marker is na. This is usually considered to be a negative verb root (cf., Chatterji (1939, 416), Ferguson (1972, 93)) as the forms nay, nao, nan, nai etc., show superficial agreement with the subject in person and grade. In spite of the existence of such forms we

do not think that na is copulative negative verb root. The primitive nature of this negative marker produces this intuitive feeling that na is the basic negative marker in Bengali from which other negative markers have been derived. na is used in copulative sentences in the present tense and simple aspect as the negative marker. In Bengali, copulative verbs are not usually realized in surface in the present tense and simple aspect, and this has misled the grammarians to consider na as a copulative negative verb root. In sādhu Bengali we get forms like nahe, nahen, nahi etc., in copulative negative sentences that show that the verb root in such a sentence is ha, which is negated by the marker na. We assume that in Standard Colloquial Bengali the copulative verb ha is deleted or left unrealized in the present tense and simple aspect, and thus the negative marker na and the forms of the AUX come morphologically closer and take such forms as nai, nao, nan, nay etc. Thus na looks like a verb. Consider the sentence (25a). In this sentence the adjective bhālo is taken as the main verb, but by our general rule (cf., § 1.2.1, Note B to Base Rule 4) we can insert ha as the right sister of the adjective bhālo. Thus the intermediate structure for (25a) is (34).

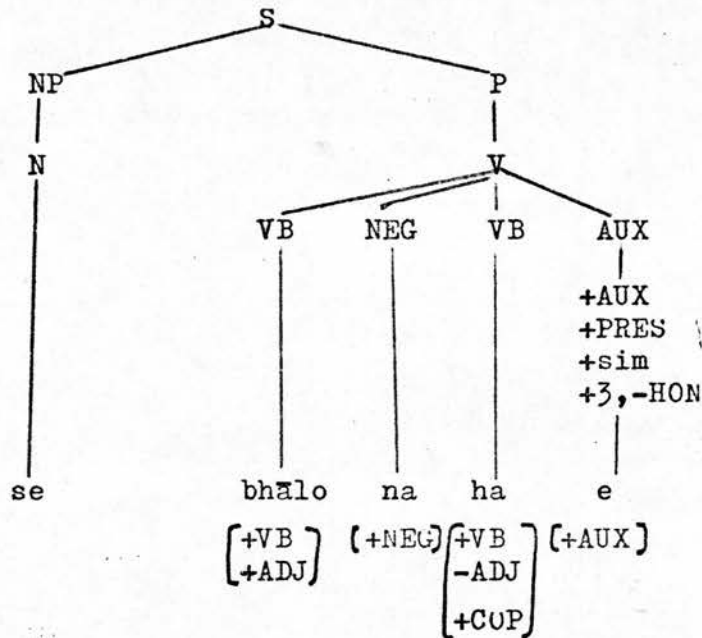
(34)



In modern Bengali the negative markers usually follow the verb form,

but in older Bengali they preceded the verb form (cf., Sen (1958)) and this phenomenon is still preserved in some dialects of Bengali. This is still at work in copulative negative sentences in SCB, which place the NEG preceding the VB ha, and the negative marker in this case is na. The NEG and AUX placement will transform (34) into (35).

(35)



The VB ha will be finally deleted, and the negative marker and the form of the AUX will form a word nay (\leftarrow na-e) by morphophonological rules. If ha in (35) is not deleted we will get na-ha-e \Rightarrow nahe, which is acceptable in sādhu Bengali, but not in SCB. Thus we see that na is a negative marker, but not a verb root, and is used in copulative negative sentences in the present tense and simple aspect. It is possible to generalize that the negative marker nā, which has taken over as the major negative marker in modern Bengali, is a suppletive form of na. We have seen that na is realized as a negative marker in a copulative sentence when the copulative verb is deleted, and it becomes nā when the verb root is realized in the surface structure. We find examples like the

following:

- (36) a. *yadi se bhālo nā-hay:* If he is not good.
 b. *yadi se nā-āse:* If he does not come.

In SCB, as we have seen before, nā usually follows the verb form, and in (35) we have seen that na precede the verb root. It is possible to achieve more generalization if we place NEG in both types of sentence preceding the VB, and give general rules like: (a) the negative marker is na if the verb root is not realized in surface, otherwise it is nā, and (b) nā, in SCB, moves to the right of the V unless they are elements of the first conjunct of a subordinate conjoined structure (that is, if they are not elements of the yadi: 'If' conjunct of a conditional sentence).

Now consider the sentences in (26)-(28). In (26) all the sentences have the negative marker nei, and in (28a) also, the negative marker is nei. If we compare the sentences in (25) with those in (26) we find that the sentences in (25) are negative copulative sentences, and those in (26) are negative existential sentences, and the sentences in each group differ semantically and syntactically. For example, (25a) means 'he is not good in character or appearance', but (26a) means 'he is not well physically'. (28a), which is negated similarly as the sentences in (26), indicates non-existence of tigers in the forest. The sentences in (26) and (28) are similar in that they are existential sentences in the negative. We see in the affirmative forms of these sentences that the existential verb āch is present. For example, the affirmative form of (26a) is (37a) and of (28a) is (37b).

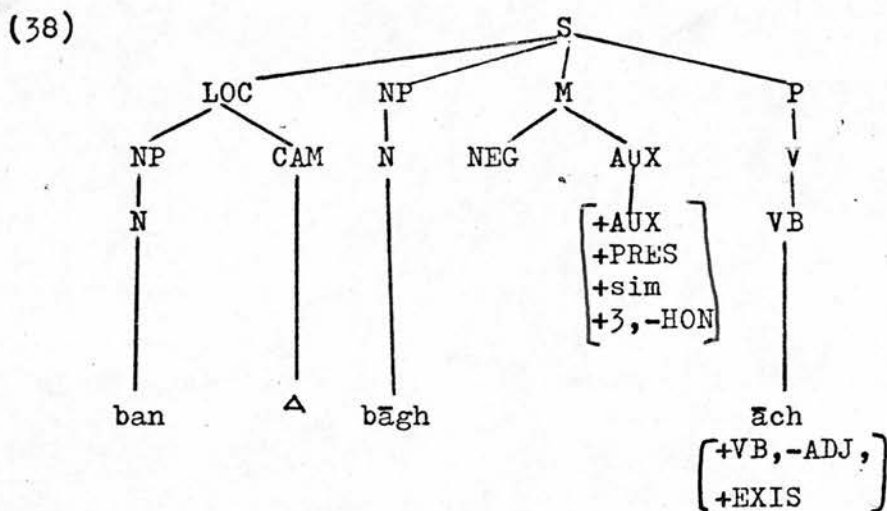
- (37) a. *se bhālo āche :* He is well.
 b. *bane bāgh āche:* There are tigers in the forest.

The verb form āche in (37a) is deletable, but this not possible

in (37b). We see in (26) and (28) that the sentences of the type (37a, b) are negated similarly in Bengali by the negative marker nei. Its form is always nei irrespective of the person and grade features of the subject. If we consider nei as the verb for non-existence, we first notice its limited use: it is used in the present tense and simple aspect only. We have to list nei in the first lexicon with the rule features that it occurs as the negative verb of existence in the present tense and simple aspect. This will complicate the negative transformation to a great extent. But we feel that semantically nei is equivalent to the negated forms of the verb āch: 'Exist' such as *āche nā, *āchen nā: 'Exist-not'. It is possible to give a general rule for the negative transformation in existential sentences, and consider nei as a surface realization of āch-AUX-NEG. For example, consider (28a),

[+PRES]
[+sim]

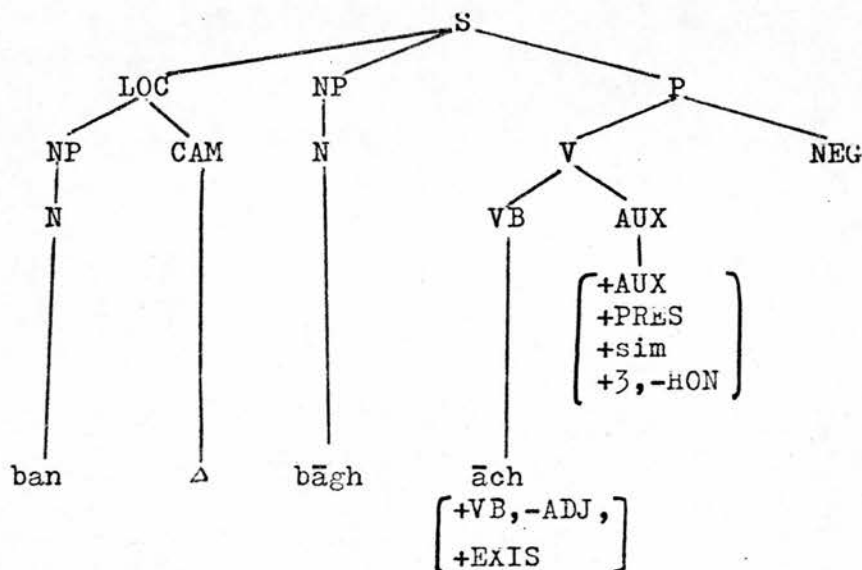
which has the intermediate underlying structure (38).



If the negative transformation applies to (38) in its general form, we will derive an ungrammatical structure *bane bāgh āche nā, which should have been grammatical in Bengali if Bengali allowed negation of the verb āch by nā (that is, if *āche nā, *āchi nā etc., were syntactically acceptable) as this is generally allowed with other

verbs (that is, kari nā: 'Do not', khāy nā: 'Do not eat' etc., are acceptable). So we will assume that the rule applies to (38) in its general form, but as *āche nā, *āchen nā etc., are unacceptable, they will be replaced by nei by the second lexical insertion rule. The AUX Placement and the Negative transformations will transform (38) into (39).

(39)



Now if the second lexical insertion rule inserts e to the AUX and nā to the NEG, we will derive an ungrammatical structure *bane bāgh āche nā. In order to block this we need a rule saying that āch-AUX-NEG should obligatorily be replaced by nei in the

[+PRES]
[+sim]

surface. This analysis shows that nei is a surface realization of the above mentioned structure, that is, nei is not an underlying verb (cf., Postal (1971b)). The sentences in (26) are also derived in this way. This derivation of the existential negative sentences in the present tense is satisfactory because this makes the negative transformation a general rule shared with other negative sentences, and the existential negative sentences in the past and the future tense. In the past tense the negative transformation in an existential sentence applies similarly, but no such surface verb

suppletion is necessary. In such a case (cf., (28b)) the NEG is realized as nā, and the existential verb forms are realized as chilo, chilen etc., depending on the features of the AUX, and by the deletion of the initial vowel of āch (āchilo ⇒ chilo).

The sentences in (27) are ambiguous between an existential and evaluative reading because both ha and āch are realized as āch in the past tense in stative and existential sense.

1.6 SUBJECT-AUX CONCORD.

The verb forms in Bengali agree with the subject noun phrase in person and grade (if any). In this work the AUX bears tense, aspect, and concord with the subject noun phrase. Subject-AUX concord is a transformational copying rule, which copies the person and grade (if any) features of the subject onto the CS of AUX, and the forms of the AUX are inserted to the derived trees from the second lexicon. After the application of the concord rule, a later transformation moves the AUX from its underlying position and adjoins it as the right daughter of V. The morphophonological rules merge the VB and the AuX into a single surface word.

We have said that the AUX agrees with the subject in person and grade (if any) in Bengali. Consider some sentences.

- (40) a. $\begin{matrix} \text{tini} & \text{yāben} & & \text{He (honorific)} & \text{will go.} \\ 1 & 2 & 1 & & \underline{\quad 2 \quad} \end{matrix}$
- b. se yābe : He (nonhonorific) will go.
- c. āmi yābo : I shall go.
- (41) a. mātin o minu yābe : Matin and Minu will go.
- b. $\text{se o āpani} \left\{ \begin{matrix} \text{yāben} \\ *yābe \end{matrix} \right\}$: He (non-hon) and you (hon) will go.
- c. $\text{mātin, tumi o āmi} \left\{ \begin{matrix} \text{yābo} \\ *yābe \end{matrix} \right\}$: Matin, you and I shall go.

In (40) the forms of AUX vary in different sentences (as ben, be, and bo) due to the varying person and grade features of the respective subjects. The subjects in (41) are coordinate noun phrases and in such cases the concord rule has some peculiarity. In (41a) each of the noun phrase conjuncts has the features [+3,-HON] and so the AUX agrees with the subject in these features. But in (41b) āpani has the features [+2,+HON] and se has the features [+3,-HON]. Here the AUX copies the features of the noun phrase conjunct that has the features [+2,+HON]. In (41c) mātin has the features [+3,-HON], tumi has the features [+2,-HON], and āmi has the features [+1]. Here the AUX agrees with the conjunct that has the [+1] feature. So we see that if the subject noun phrase is a coordinate conjunct the Subject-AUX concord rule works on the principle of dominance in the person and grade hierarchy. The AUX agrees with that noun phrase conjunct which is higher in the hierarchy of person and grade. The concord rule is given below.

(42) SUBJECT-AUX CONCORD RULE.

SI: S [X NP [X N X] X AUX X]

$$\begin{array}{c} \left[\begin{array}{l} +... \\ \alpha \text{ PERSON} \\ \beta \text{ GRADE} \\ +\text{SUBJ} \end{array} \right] \\ \begin{array}{ccc} 1 & 3 & 5 \end{array} \end{array}$$

$$\left[\begin{array}{l} +\text{AUX} \\ +... \end{array} \right] \\ 8$$

SC: a. Copy the $\left[\begin{array}{l} \alpha \text{ PERSON} \\ \beta \text{ GRADE (if any)} \end{array} \right]$ features of 5 onto 8.

b. If 3 is a coordinate noun phrase apply (42a) so that

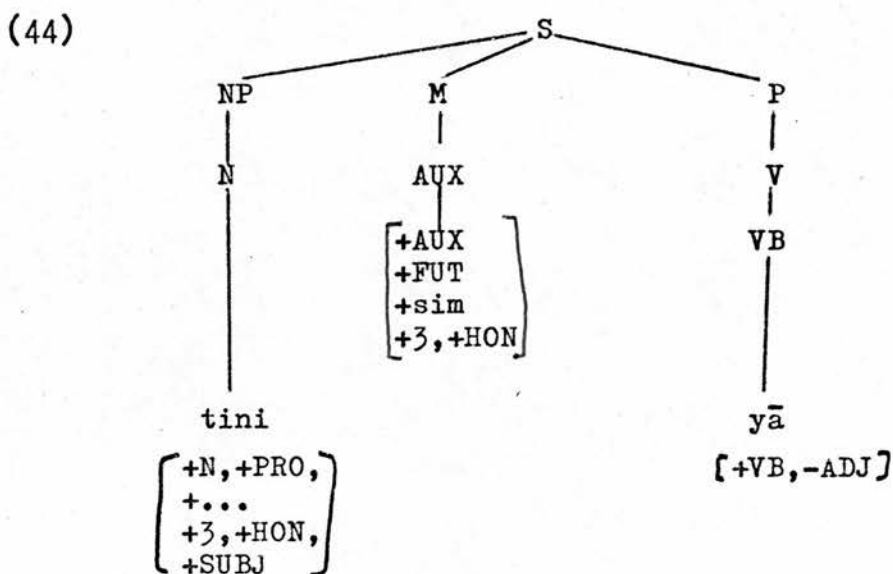
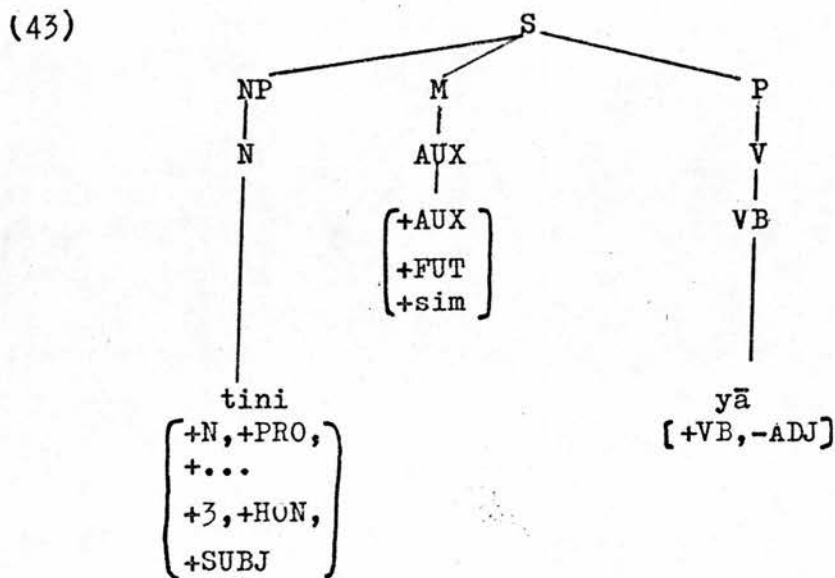
5 is higher in the hierarchy: $\left[\begin{array}{l} +1 \\ +2,+HON \\ , -HON \\ , +PEJ \\ +3,+HON \\ , -HON \end{array} \right]$

among the conjuncts.

Conditions: a. 3 is the Subject of 1.

b. 1 dominates 3 and 8.

This rule will derive (44) from (43).



1.7 AUX-MOVEMENT.

The AUX is an element of M in the underlying structure, but it is moved, after the application of the Subject-AUX concord rule, out of its underlying position, and is adjoined as the right daughter of the V in order to derive acceptable surface verb forms. This rule is given below.

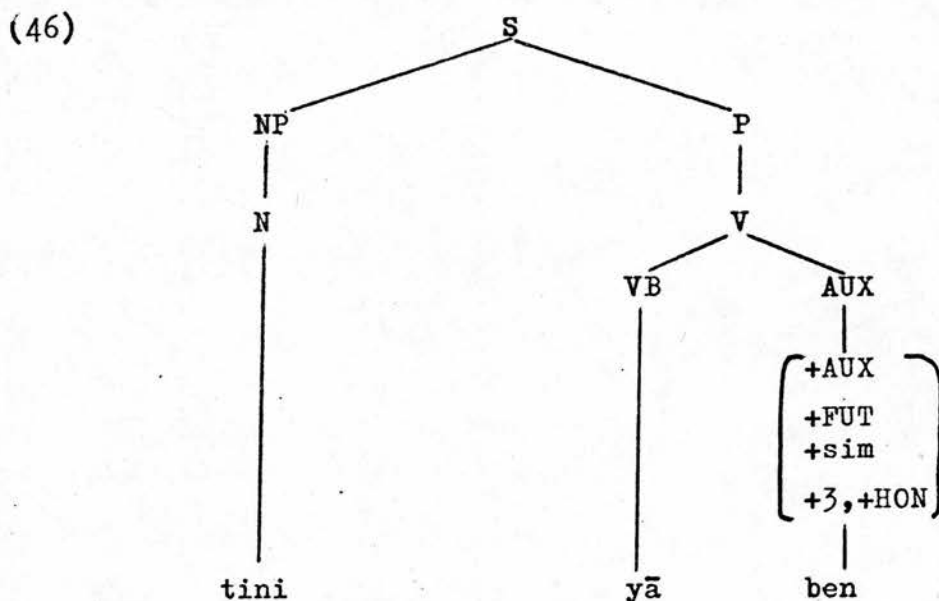
(45) AUX-MOVEMENT RULE.

SI: S [X AUX X V X]
 1 3 5

SC: Adjoin 3 as the right daughter of 5.

Condition: 3 and 5 are the constituents of 1, which is
 a simplex.

This rule will transform (44) into (46).



The form of the AUX ben is inserted by the second lexical insertion rule after the last rule of the transformational component (cf., § 11.2). (46) will generate tini yāben, which is (40a).

NOUN PHRASES2.0 Introduction.

As pronominalization involves noun phrases, we will discuss the structures and general characteristics of the Bengali noun phrase in this chapter. The head of a noun phrase is a noun (N). The node N of a noun phrase can be filled by two types of lexical item: Nouns, which have the inherent features [+N,-PRO], and pronouns, which have the inherent features [+N,+PRO]. They are considered as categories of the same type in the underlying structure, differing in the value for the feature [PRO], which is positively specified for the pronoun and negatively for the noun. We have discussed the pronouns of Bengali elsewhere (cf., § 4.1- § 4.1.4; § 4.2.1- § 4.2.3; § 4.3.1).

2.1 Subcategorization of Nouns.

This grammar does not generate nouns or complex symbols for nouns (cf., § 1.2), it simply utilizes them from the lexicon in order to generate sentences. The lexicon contains all the lexical items of Bengali with proper feature specification. So this section is related primarily to the lexicon (cf., Chapter 11), but has been included here on the assumption that it will help us to get a clear understanding of the noun phrase structures of Bengali.

When one comes to subcategorize the nouns of Bengali one will feel that at least the following inherent features require consideration.

- (1) [COMMON], [COUNT], [HUMAN], [ANIMATE], [HONORIFIC],
[PEJORATIVE], [MALE], [ABSTRACT], [LOCATION]

The features listed above are not exhaustive; some nouns might have some inherent features not included here, but no noun will

have all these features positively or negatively specified for it. We have discussed the method of feature specification elsewhere (cf., § 11.1.3), but it may be mentioned here that according to the inherent property of nouns some feature will be positively specified for a noun (for example, bhadralok: 'Gentleman' is [+MALE]) and this feature will be negatively specified for another noun (for example, mahilā: 'Woman' is [-MALE]), and the feature may be irrelevant for another noun (for example, [MALE] is irrelevant for kabitā: 'Poem'). Consider the nouns in (2) with their inherent feature specifications (redundant features omitted):

(2) bhadralok: 'Gentleman': [+N,-PRO,+COM,+COUNT,+HON,+MALE]

mahilā: 'Woman': [+N,-PRO,+COM,+COUNT,+HON,-MALE]

chele: 'Boy': [+N,-PRO,+COM,+COUNT,-HON,+MALE]

meye: 'Girl': [+N,-PRO,+COM,+COUNT,-HON,-MALE]

These nouns have a number of common features as well as some features that differentiate them. For example, bhadralok and mahilā differ only in the value for the feature [MALE], which is positive for bhadralok and negative for mahilā. The distinctive features common to all these nouns are [+HUM,+COM,+COUNT] , and thus they form a class by themselves. The feature [MALE] plays a minor role in the syntax of Bengali, but it should be included for the specification at least of [+HUM] nouns, because in some syntactic structures this feature is pivotal for selectional restrictions. Consider the examples below.

(3) a. mahilā sundarī: The woman is beautiful.

b. *bhadralok sundarī: The gentleman is handsome.

(4) a. mahilā adhyāpikā: The woman is a teacher.

b. *bhadralok adhyāpikā: The gentleman is a teacher.

(3b) and (4b) are ungrammatical due to the reason that the adjective

sundarī and the predicate nominal adhyāpikā are incompatible with [+MALE] nouns. The features [+HON] are necessary for semantic as well as syntactic reason that the AUX agrees with the subject in person and grade features in Bengali (cf., § 1.6). Now consider the nouns in (5):

- (5) pākhī: 'Bird': [+N, -PRO, +COM, +COUNT, -HUM]
 bai: 'Book': [+N, -PRO, +COM, +COUNT, -ANI, +ABS]
 jal: 'Water': [+N, -PRO, +COM, -COUNT, -ANI, -ABS]
 svapna: 'Dream': [+N, -PRO, +COM, +COUNT, -ANI, +ABS]
 path: 'Road': [+N, -PRO, +COM, +COUNT, -ANI, -ABS, +LOC]

The feature common to all the nouns of this group is [+COM], which is shared by the nouns in (2); but they differ from those in (2) for the value of [HUM] and [ANI]. In (5) all but pākhī are [-ANI], and by a general rule in Bengali they are [-HON]. Now we come to the class of proper nouns given in (6).

- (6) rabīndranāth: [+N, -PRO, -COM, -COUNT, +HUM, +MALE]
 suphiā khān: [+N, -PRO, -COM, -COUNT, +HUM, -MALE]
 hāsān: [+N, -PRO, -COM, -COUNT, +HUM, +MALE]
 ketakī: [+N, -PRO, -COM, -COUNT, +HUM, -MALE]
 dhākā: [+N, -PRO, -COM, -COUNT, -ANI, -ABS, +LOC]
 puṣi: [+N, -PRO, -COM, -COUNT, -HUM]

The most distinctive feature of these nouns is [-COM], that is, they are names of persons, places and animals etc. The [+HUM] proper nouns have many features common with [+HUM] common nouns, and similarly [-HUM] and [-ANI] proper nouns have many features common with their [+COM] counterparts. The feature [HON] is relevant for [+HUM] proper nouns; for example, rabīndranāth will be considered as [+HON] and ketakī will be considered as [-HON], but we assume that [+HON] is not an inherent feature of any proper noun. Whether a proper noun will have the feature [HON] positively or negatively

specified for it depends on social judgement. So two men with the same name may have entirely different status. The lexicon, if it lists proper nouns at all, may ignore these features for proper nouns. But the [+HON] feature of the human proper nouns must be taken into account when they are used in a sentence because the AUX concords with the subject in person and grade features. The place and animal names have the features [-COM, -COUNT] in common with the human proper nouns, and like the [-HUM] and [-ANI] common nouns, they are [-HON] by a general rule in Bengali (cf., § 11.1.4).

The features discussed above are semantic as well as syntactic. Their syntactic relevance in the generation of the grammatical noun phrase structures of Bengali will be seen in the following sections.

2.2 General Characteristics of Noun Phrases.

Some noun phrases of Bengali consist simply of a head noun, as can be seen in (7).

- (7) a. pākhī gān gāy: Birds sing songs.
 1 2 3 1 3 2
- b. pākhīrā gān gāy: Birds sing songs.
- c. nāju bai parche: Nazu is reading (a) book(s).
 2 3 3 2

The noun phrases pākhī and gān in (7a), pākhīrā and gān in (7b), and nāju and bai in (7c) are head nouns without any determiner. The noun phrase pākhī in (7a) is a head noun in the base form and is understood as generic. The noun phrase pākhīrā, which is plural and morphologically marked for plurality, is also generic, as in (7a, b) the tense and aspect are generic [+PRES, +sim]. The noun pākhī is [+COUNT], and it may take a determiner; but the noun nāju in (7c) is [-COM], and does not take any determiner. The noun phrase

gān in (7a, b) is also undetermined, simply a head noun in the base form, and understood as generic. The noun phrase bai in (7c) is also a head noun in the base form, but it is not understood as generic (due to non-generic tense and aspect). It can be understood either as singular or plural. In other words, it is unspecified for number. Thus we see that a Bengali noun phrase is simply a head noun when: (a) the N is a proper noun (and a pronoun), (b) the NP is generic (but not all generic noun phrases are undetermined), and (c) the noun is unspecified for number. The above noun phrases are undetermined in that they have not taken any determiner (but technically pākhīrā has taken a determiner (cf., § 2.4.4)). Now consider the following noun phrases.

- (8) a. ekjan bhadralok: A gentleman.
 b. duṭi pākhī: Two birds.
 c. pākhī duṭi: The two birds.

These NP's are determined, and of these (8a, b) are indefinite, and (8c) is definite. The elements ekjan and duṭi above are SPEC(ifier)s (cf., § 2.4.3), which are composed of quantifiers such as ek: 'One', and du: 'Two', and classifiers such as jan and ṭi. The SPEC's in (8a, b) are at the left of their respective head nouns, and these noun phrases are understood as indefinite; but the SPEC duṭi is at the right of its head noun in (8c), which is understood as definite. This shows that Bengali has no article (cf., § 2.4.3). Now consider the following noun phrases.

- (9) a. o(i) cheleṭi: That boy.
 1 2 1 2
 b. se(i) mahilā: That woman.
 c. e(i) meyerā: These girls.

These noun phrases are definite. In (9a) the deictic e(i) is at the left of the head noun and the classifier ṭi is suffixed to

the head noun, and in (9b), similarly, the deictic se(i) is at the left of the head noun, but no classifier is suffixed to it. (9c) is a definite unquantified plural noun phrase, which has the deictic e(i) at the left of the head noun, and no classifier is suffixed to it. This NP looks similar to (9b) on surface, but we will show later that (9a, b) are much more alike than (9b, c) are (cf., § 2.4.3).

The Bengali noun phrases can be recursive, that is, coordinately conjoined as in (10).

- (10) a. minu o mātin: Minu and Matin.
 b. ekṭi chele o ekṭi meye: A boy and a girl.
 1 2 3 4 5 1 2 3 4 5

In (10) the NP's are conjoined by the conjunction o: 'And'. Noun phrases may have attributive adjectives, which are usually placed immediately preceding the head noun. Consider the examples in (11).

- (11) a. ekṭi nīl pākhī: A blue bird.
 2 3 2 3
 b. nīl pākhīṭi: The blue bird.
 c. nīl ekṭi pākhī: A blue bird.

In (11a, b) the adjective nīl immediately precedes the head noun, but in (11c) the SPEC ekṭi intervenes between the adjective and the head noun. (11a, c) are cognitively synonymous, but the topicalization of the adjective in (11c) makes a subtle semantic distinction between (11a, c).

In this section we have introduced some NP structures in outline without any analysis. In the following sections we will consider how these and other NP's are handled in this grammar.

2.3 Number.

Lyons (1968, 281) considers number as a category of nouns. Bengali has a two-way distinction in the category of number, that is, nouns can be either singular or plural; and this applies to [+COUNT] nouns only. But in some noun phrases with [+COUNT] nouns the category of number may remain unspecified in Bengali, as we have seen in the noun phrase bai in (7c). Plurality of nouns in Bengali is expressed in two ways¹: (a) by suffixing a plural marker to the base form of a noun, and (b) by quantifiers. Let us see how is this done in Bengali.

1 A third way of indicating plurality is by the reduplication of the attributive adjectives as in the following examples.

- (A) a. choṭa choṭa bārī: Small small homesteads.
 b. lāl lāl phul: Red red flowers.

In (A) the adjectives are reduplicated and the head nouns are in the base form, but they are understood as plural. We derive the attributive adjectives from the predicate position, and so these noun phrases should be related to full, probably conjoined, underlying sentences. In this type of NP the collective quantifiers and the reduplicated adjectives can cooccur:

- (B) a. anek choṭa choṭa bārī: Many small small homesteads.
 b. bahu lāl lāl phul: Many red red flowers.

This sort of reduplication of adjectives indicates unspecificity and innumerability, but specificity of NP's with reduplicated adjectives can also be expressed by the use of the specific plural marker gulo with the head noun: lāl lāl phulgulo: 'Red red flowers'. As these noun phrases are presumably related to underlying full sentences, we will not consider them here.

2.3.1 Plurality by Plural Markers.

Plurality in such a case is expressed by the plural forms of nouns, which are formed by the suffixation of a plural marker to the base form of a noun as shown in (12).

(12)	<u>Base Form</u>	<u>Plural Marker</u>	<u>Plural Form</u>
	bālak: 'Boy'	erā	bālakerā
	meye: 'Girl'	rā	meyerā
	pākhī: 'Bird'	gulo	pākhīgulo
	pandit: 'Scholar'	gaṇ	panditgaṇ

We see above that the plural forms of nouns are formed by the suffixation of the plural markers to the base forms of nouns.

The plural markers that are commonly used in SCB are rā, erā, der and gulo. But there are some plural markers such as gaṇ, barga, samuha etc., which are now antiquated, and are usually used with Sanskritic nouns in formal writing and oratory. Here we will mainly deal with the plural markers rā, erā, der and gulo.

The plural markers rā, erā and der have two functions: (a) they are used if the head noun is [+HUM], and (b) if the NP is intended generically. gulo has two similar functions: (a) it is used if the head noun is [-ANI] or [-HUM], and (b) if the NP is intended as specific. rā, erā and der can be used with [-HUM] nouns if the noun phrase is generic or unspecific. gulo can be used with [+HUM, -HON] nouns if the noun phrase is specific, but it is not used with [+HON] nouns. Consider the examples below.

- (13) a. bhadralokerā kaphi khān: Gentlemen drink coffee.
 b. *bhadralokgulo kaphi khān:
 c. e-bhadralokerā kaphi khān: These gentlemen drink coffee.
- (14) a. pākhīrā gān gāy: Birds sing songs.
 b. pākhīgulo gān gāy: (Some) birds sing songs.

In (13a) bhadralokerā is [+HON] and generic, and the plural marker is rā; but (13b) is unacceptable as gulo is incompatible with a [+HON] noun. In (13c) e-bhadralokerā is specific, but the plural marker is rā as the noun is [+HON]. In (14a) pākhī is [-HUM], but as the NP is generic, the plural marker is rā. In (14b) the plural marker is gulo, because the noun is [-HUM] and the NP is specific.

The plural markers rā, erā and der are syntactically conditioned: (a) rā and erā are used if the noun phrase is the subject, and (b) der is used otherwise. But gulo has no such condition. Consider the examples below.

- (15) a. mahilārā elen: (The) women came.
 1 2 1 2
 b. *mahilāder elen:
- (16) a. āmi mahilāder(ke) cini: I know the women.
 1 2 3 1 3 2
 b. *āmi mahilārāke cini:
- (17) a. pākhīgulo uṛbe: The birds will fly.
 1 2 1 2
 b. āmi pākhīguloke dharbo: I shall catch the birds.

In (15a) mahilārā is the subject, and is grammatical, but (15b) is ungrammatical, because the subject NP has taken the plural marker der. In (16a) mahilāder is the object, and is grammatical; but (16b) is ungrammatical, because the object noun phrase has taken rā. In (17) we find that gulo can be used with a subject as well with an object noun phrase.

The plural markers rā and erā are phonologically conditioned: rā is used when a noun ends in a vowel, and erā elsewhere. Consider the examples below.



- (18) a. $\left\{ \begin{array}{l} \text{bālak: 'Boy'} \\ \text{rājñitibid: 'Politician'} \end{array} \right\} \left\{ \begin{array}{l} \text{erā} \\ \text{*rā} \end{array} \right\}$
- b. $\left\{ \begin{array}{l} \text{mahilā: 'Woman'} \\ \text{meye: 'Girl'} \end{array} \right\} \left\{ \begin{array}{l} \text{rā} \\ \text{*erā} \end{array} \right\}$

In (18a) we see that a noun that ends in a consonant takes erā as a plural marker and in (18b) we see that a noun that ends in a vowel takes rā as a plural marker.

Bengali has a number of plural markers like gaṇ, barga, samuha etc., which are now obsolete. They are used with Sanskritic nouns in formal writing and speech. Of these gaṇ and barga are used with [+HUM] nouns, and samuha is used with [-HUM] and [-ANI] nouns. So the formal Bengali has plural forms like panditgaṇ, panditbarga: 'Scholars', and pustaksamuha: 'Books', but not *panditsamuha, or *pustakgaṇ.

The plural markers rā, erā, der and gulo have entries in the second lexicon in the following manner:

- (19) $\begin{array}{c} \underline{rā} \\ \left[\begin{array}{l} +\text{PL MARKER} \\ +\text{HUM} \\ +\text{V}_- \\ +\text{SUBJ} \end{array} \right] \end{array}$ $\begin{array}{c} \underline{erā} \\ \left[\begin{array}{l} +\text{PL MARKER} \\ +\text{HUM} \\ +\text{C}_- \\ +\text{SUBJ} \end{array} \right] \end{array}$ $\begin{array}{c} \underline{der} \\ \left[\begin{array}{l} +\text{PL MARKER} \\ +\text{HUM} \\ -\text{SUBJ} \end{array} \right] \end{array}$ $\begin{array}{c} \underline{gulo} \\ \left[\begin{array}{l} +\text{PL MARKER} \\ -\text{HON} \end{array} \right] \end{array}$

2.3.2 Plurality by Quantifiers.

Numerical quantifiers such as du(i): 'Two', tin: 'Three', cār: 'Four' and so on, and the collective quantifiers such as bahu: 'Many', sakal: 'All' etc., are also used in Bengali in order to indicate plurality. In such a case the quantified noun does not take any plural marker. Consider the examples below.

- (20) a. dujan mahilā: Two women.
b. *dujan mahilārā:

(21) a. *tinṭi pākḥī*: Three birds.

b. **tinṭi pākḥīgulo*:

(22) a. *bahu lok*: Many men.

b. **bahu lokerā*:

Plurality of the nouns in (20, 21, 22a) are respectively indicated by the quantifiers *du*, *tin* and *bahu*; and the nouns in these examples remain in their base form. The ungrammaticality of the (b) examples above is due to the fact that the nouns have taken plural markers.

The category of number does not apply to [-COUNT] nouns. It is true that [+COM,-COUNT] nouns cannot be counted and cannot be pluralized, but they can be measured and quantified. Consider the examples below.

(23) a. **dudhgulo*: *Milks.

b. **ekṭi dudh*: *A milk.

c. *dumaṅ dudh*: Two maunds of milk.

(23a) shows that a [+COM,-COUNT] like *dudh* cannot be pluralized,

and (23b) shows that it cannot be counted either. But (23c)

shows that it can be measured and quantified. Like the [+COM,-COUNT] nouns, [-COM,-COUNT] nouns cannot be pluralized or counted.

Consider the examples below.

(24) a. **ekṭi ketakī*: *A Ketaki.

b. **ketakīrā*: *Ketaki's.

c. *ketakīrā āsbe*: The Ketaki's will come.

(24a) shows that a proper noun cannot be counted, and (24b) shows that it cannot be pluralized either. But in (24c) we see that a proper noun has been morphologically pluralized, which is a common practice in Bengali. *ketakīrā* in (24c) is not semantically plural because it does not refer to several Ketaki's, instead it means 'Ketaki and her family or friends'.

2.4 Underlying Structures of Noun Phrases.

Here we repeat our base rules relating to noun phrases for the sake of convenience of discussion.

(25) Base Rules Repeated:

$$\text{RULE 6} \quad \text{NP} \text{ ----} \rightarrow \left\{ \begin{array}{lll} \text{CONJCO} & \text{NP} & \text{NP*} \\ & \text{S} & \\ (\text{D}) & \text{N} & (\text{S}) \end{array} \right\}$$

$$\text{RULE 7} \quad \text{D} \text{ ----} \rightarrow (\text{PART}) \left(\left\{ \begin{array}{l} \text{DEF} \\ \text{DEIC} \end{array} \right\} \right) (\text{SPEC})$$

$$\text{RULE 8} \quad \text{DEIC} \text{ ----} \rightarrow (\text{DEM})(\text{ORD})$$

$$\text{RULE 9} \quad \text{SPEC} \text{ ----} \rightarrow \left\{ \begin{array}{ll} \text{QUANT} & (\text{CL}) \\ \text{PL} & \end{array} \right\}$$

These rules will account for all noun phrase structures of Bengali. The only constituent that is obligatory (but can be contextually deleted) for a noun phrase is the head noun (N). It plays an important role in the selection of other elements of a noun phrase. In the following sections we will describe the noun phrase constituents in some detail, and show how they combine with one another in order to generate grammatical noun phrases of Bengali.

2.4.1 NP --> N.

This rule accounts for the noun phrases which consist only of head nouns. Consider the examples below.

- (26) a. pākhī gān gāy: Birds sing songs.
 b. nāju bai paṛe: Nazu reads (a) book(s).
 c. āmi yābo: I shall go.
 1 2 1 2

The NP pākhī in (26a) is undetermined, simply a head noun in the

base form, and is understood as generic. It is not that all generic NP's in Bengali are undetermined, but NP's such as this are understood generically if the sentence is in the present tense and simple aspect. The rule under discussion allows this noun phrase to be generated. In Bengali [-COM] nouns and singular pronouns do not take any determiner. nāju in (26b) and āmi in (26c) are respectively a [-COM] noun and a pronoun. This rule accounts for these NP's. The NP bai in (26b) has no determiner either. It is unspecified for number, but its number is contextually understood, and the cooccurring verb also helps to ascertain its number. The rule under discussion allows the generation of NP's (a) which are undetermined and generic, (b) which are undetermined and unspecified for number, and (c) which have singular pronominal head nouns, and proper nouns as head.

2.4.2 D(eterminer).

The D is an auxiliary node, which dominates all constituents but the head noun, of a determined NP. All its constituents are optional, but if D is chosen at least one its constituents must be chosen. The D itself is an optional element of an NP.

2.4.3 SPEC --> QUANT (CL).

Consider the noun phrases below.

- (27) a. ekjan bhadralok: A gentleman.
 b. duṭi meye: Two girls.
 c. tinkhāni sāri: Three saris.

All the NP's above are indefinite. First consider the elements ekjan, duṭi and tinkhāni in these NP's. These elements are composed of quantifiers (QUANT) and classifiers (CL). In ekjan the QUANT is ek: 'One' and the CL is jan; in duṭi QUANT is du: 'Two'

and the CL is ṭi, and in tinkhāni the QUANT is tin: 'Three' and the CL is khāni. These elements, which we have termed SPEC(ifier), specify the head nouns in respect of quantity, definiteness and indefiniteness, specificity and unspecificity and so on. The SPEC is composed of a QUANT and a CL. The CL's are optional in some situations, and are extremely idiosyncratic.

We assume that the notion of QUANT is quite straightforward. It includes the numerals ek: 'One', du: 'Two', and so on, and the collectives like bahu, anek: 'Many', sab, sakal: 'All', kayek: 'Some' etc. We will come back to the QUANT later in this section, but first we will consider the CL's. The bound morphs ṭā, ṭi, khānā, khāni, jan etc., are traditionally called 'articles' or 'definitives' (cf., Chatterji (1939, 253; 1926, 780)). These are, specially ṭi and ṭā (as these two are frequently used), generally viewed as the Bengali counterpart of the English definite article 'The'. This view is unacceptable because they are used not only to indicate definiteness, but are used in indefinite noun phrases too. Furthermore, they have other functions, which are quite complicated, such as expressing the attitude of the speaker towards the noun concerned, the inherent properties of the noun etc. It is widely believed that these CL's, specially ṭi and ṭā, are suffixed to the nouns in order to indicate definiteness. This assumption is only superficially adequate; if we go deeper into the facts it cannot be maintained. To begin with, consider the NP's in (28).

- (28) a. ekṭi bai: A book.
 b. baiṭi: The book.

The NP (28b) is definite, and here the CL ṭi is suffixed to the noun bai. As (28b) is definite, it is traditionally assumed that

ṭi is a definite article like the English 'The'. But we see in (28a), which is an indefinite NP, the CL ṭi is suffixed to the QUANT ek. This NP being indefinite shows ṭi cannot be a definite article. CL's can be constituents of both definite and indefinite noun phrases.

Here we will study the syntax and semantics of ṭā, ṭi, jan, khānā and khāni, which are frequently used. We have said that CL's are optional elements, but if a specific quantifier is selected the selection of a CL becomes obligatory, and the inherent features of the head noun and the attitude of the speaker towards the head noun restricts the selection of a CL. For example, ṭā can be used if the head noun is [-HUM], but if it is used with a [+HUM] noun it will indicate that the speaker has no respect for the person concerned. The CL ṭi can be used with [-HON] nouns, which include human, non-human and inanimate nouns. The CL jan is used when the head noun is [+HUM], preferably [+HON]. The CL's khānā and khāni are used with [-ANI] nouns which are small in size, and preferably square or rectangular in shape. All these CL's are used with [+COUNT] nouns, not with [-COUNT] nouns (except khāni). Consider the examples below.

- (29) a. ek $\left\{ \begin{array}{l} \text{ṭi} \\ \text{ṭā} \\ *jan \\ *khānā \\ *khāni \end{array} \right\} \left\{ \begin{array}{l} \text{pākhī} \\ \text{nadī} \end{array} \right\} : \text{A } \left\{ \begin{array}{l} \text{bird} \\ \text{river} \end{array} \right\}.$
- b. du $\left\{ \begin{array}{l} \text{jan} \\ *ṭi \\ *ṭā \\ *khānā \\ *khāni \end{array} \right\} \left\{ \begin{array}{l} \text{mahilā} \\ \text{bhadralok} \end{array} \right\} : \text{Two } \left\{ \begin{array}{l} \text{women} \\ \text{gentlemen} \end{array} \right\}.$

c. tin $\left\{ \begin{array}{l} \dot{\text{t}}\text{i} \\ \dot{\text{t}}\text{ā} \\ \text{k}\text{h}\text{ā}\text{n}\text{ā} \\ \text{k}\text{h}\text{ā}\text{n}\text{i} \\ *\text{j}\text{an} \end{array} \right\} \left\{ \begin{array}{l} \text{bai} \\ \text{'s}\text{ā}\text{r}\text{i} \end{array} \right\} : \text{Three} \left\{ \begin{array}{l} \text{book} \\ \text{sari} \end{array} \right\} \text{s} .$

In (29a) we see that ekṭi pākḥī, ekṭā pākḥī, ekṭi nadī and ekṭā nadī are grammatical, and *ekkhānā pākḥī, *ekkhāni pākḥī, *ekjan pākḥī and *ekjan nadī are ungrammatical. The ungrammaticality of these NP's is due to an unacceptable selection of CL's. pākḥī is [-HUM], and so the CL should be ṭi or ṭā, but not jan; and it is [+ANI], so khānā or khāni cannot be used with it. Similarly, with nadī, which is [-ANI], the CL's ṭā or ṭi can be used, but not jan. The CL's khānā and khāni do not go with nadī, if it is understood as big or vast; but in emotional expressions like ekkhānā choṭa nadī: 'A tiny river', where the river is almost reduced to a picture, the use of khānā and khāni is acceptable. This shows how a CL expresses the attitude of the speaker towards the noun concerned. In (29b) we find that with [+HON]nouns only jan can be used as a CL. The use of ṭā or ṭi with [+HON] nouns is contemptuous, and the use of khānā or khāni indicates inanimateness. So a noun phrase like *ekṭi bhadralok: 'A gentleman', although syntactically well-formed, is unacceptable on social grounds, and a noun phrase like *ekkhānā mahilā: 'A woman' will be socially rejected, but can be used satirically to indicate some sort of inanimateness in a woman. (29c) shows that with nouns like bai and 'sāri, which are [-ANI] and not large in size, any of ṭā, ṭi, khānā and khāni can be used. But when ṭi is used the speaker expresses no emotion; and when ṭā is used the speaker expresses his indifference towards the object; and when khānā or khāni is used, the speaker expresses his tenderness towards the object.

We have said that CL is an optional element of the SPEC, which dominates QUANT and CL. Selection of a CL is mainly dependent on the QUANT, and the selection of particular classifier is restricted by the head noun. But it should be mentioned here that the behaviour of the CL's is so idiosyncratic that no single constituent or property of noun phrase can be thought of as pivotal behind their selection or omission. Their selection or omission depends on the entire property of the noun phrase, but it seems that the QUANT plays a major part in their selection. The following syntactic-semantic features are relevant for an analysis of the QUANT's in Bengali:

(30) [+DEFINITE] , [+SPECIFIC]

Some QUANT's are [+DEF] , some are [-DEF, +SPECIFIC] , some are [-DEF, -SPECIFIC] , and some are [-DEF, +SPECIFIC] . The QUANT's pratyek, prati: 'Each' are [+DEF], the numeral QUANT's are [-DEF, +SPECIFIC], kayek: 'Some' [-DEF, +SPECIFIC] , bahu: 'Many' and sakal: 'All' are [-DEF, -SPECIFIC], and anek is [-DEF, +SPECIFIC]. The general rule for a CL-selection is that if the chosen quantifier is [+DEF] or [+SPECIFIC], a CL should be selected, and if the quantifier is [-SPECIFIC], a CL should not be selected (although there are some exceptions to this general rule).

Consider the examples below.

(31) a. ekṭi pākhī uṛche: A bird is flying.
 1 2 3 1 2 3

b. *ek pākhī uṛche:

(32) a. ekṭi meye āmāke balechilo..: A girl had told me..
 1 2 3 4 1 2 4 3

b. ek meye āmāke balechilo:

(33) a. tinjan rūpasīr gopan kathā: Secret stories of three beauties.
 1 2 3 4 3 4 1 2

b. tin rūpasīr gopan kathā:

(34) a. $\begin{matrix} \text{daśjan} & \text{adhyāpak} & & \\ 1 & 2 & 1 & 2 \end{matrix}$: Ten teachers.

b. ?daś^1 adhyāpak:

(35) a. $\begin{matrix} \text{anek} & \text{lok} & & \\ 1 & 2 & 1 & 2 \end{matrix}$: Many men.

b. *anekjan lok:

(36) a. $\begin{matrix} \text{sab} & \text{pākhī} & & \\ 1 & 2 & 1 & 2 \end{matrix}$: All birds.

b. *sabṭi pākhī:

(31a) shows that the noun phrase ekṭi pākhī is grammatical, and (31b) shows that *ek pākhī is ungrammatical. It may be said here that as the quantifier is [+SPECIFIC] there should be a CL in this noun phrase. But our opinion is that this noun phrase seems unacceptable for some phonological reason: it sounds, literally, odd without a CL. In the NP ekṭi meye in (32a) the CL ṭi is present and the noun phrase is indefinite, but specific. In ek meye in (32b) there is no CL, and the noun phrase seems to be unspecific, but the semantic difference between these two NP's is so subtle that the NP in (32b) can be considered as a stylistic variant of the one in (32a). In the genitive noun phrase tinjan rūpasī in (33a) the SPEC contains the CL jan, but in tin rūpasī in (33b) there is no CL. When the CL is used the noun phrase is understood as specific, but without a CL it seems unspecific. Thus the noun phrase in (33a) means 'secret stories of three particular beauties', and (33b) means 'secret stories of three unspecific beauties'. Usually noun phrases of the type (33b) are used as titles, captions or head lines (such as pāc corer galpa: 'Stories of five thieves', sāt rājār dhan: 'Treasures of seven kings', tin kanyā: 'Three daughters') which deal with unspecific nouns. These are unspecific and stylistically catchy. Now consider the noun phrase daśjan adhyāpak in (34a). The noun adhyāpak is [+HON],

and it allows only jan as a CL. The noun phrase (34b), where there is no CL, is objectionable due to the omission of the CL. Here the absence of the CL does indicate the noun phrase is unspecific, but the noun phrase is objectionable because the deletion of the CL would dishonour the noun phrase, which is [+HON].

The QUANT's anek and sab are [-SPECIFIC] when used with a [+COUNT] noun and thus do not take a CL. (35, 36b) are ungrammatical due to the selection of CL's. Consider some more examples, this time with temporal nouns.

(37) a. tin bachar keṭe gelo: 'Three years passed by.
 1 2 3 1 2 3

b. tinṭi bachar keṭe gelo:

(38) a. āmi du-din dhare ekhāne āchi: I have been here for two days.
 1 2 3 4 5 6 1 6 5 4 2 3

b. āmi duṭi din dhare ekhāne āchi:

There is no CL in tin bachar in (37a), but ṭi is there in tinṭi bachar in (37b). Similarly in (38a) there is no CL in du-din, but ṭi is there in duṭi din in (38b). CL's are not generally used in temporal noun phrases like those in (37, 38a), but (37, 38b) show that a CL can be used in such a noun phrase. In the NP's tin bachar and du-din, it seems that the period is taken as a single stretch of time without any emphasis, but in tinṭi bachar and duṭi din the period is counted year by year and day by day, and thus emphasized to seem to be a long period of time. It is possible to take the temporal NP's without a CL as unspecific, and those with a CL as specific. Consider the examples in (39).

(39) a. ?se āmār jībaner sundar tin bachar naṣṭa kareche: She
 has wasted three beautiful years of my life.

b. se āmār jībaner sundar tinṭi bachar naṣṭa kareche:

The sentence (39a) is odd as the noun phrase sundar tin bachar

seems to indicate three specific years, but as there is no CL in this noun phrase it indicates unspecificity, and this contradiction makes this sentence odd (but in a different reading the sentence is natural: if the speaker's whole life is taken as beautiful, then the wasted years are unspecific, in which case deletion of the CL is expected.). In (39b) the noun phrase sundar tin̄i bachar contains the CL ti, and the noun phrase is understood as three specific years.

We have not said anything so far about the quantifiers that we consider as [+DEF]. When a definite quantifier is chosen a classifier should be chosen. Consider the examples below.

- (40) a. ?pratyek chele cād dekheche: Each boy has seen the moon.
 b. pratyek̄ti chele cād dekheche:
- (41) a. *prati chele cād dekheche: Each boy has seen the moon.
 b. pratīti chele cād dekheche:

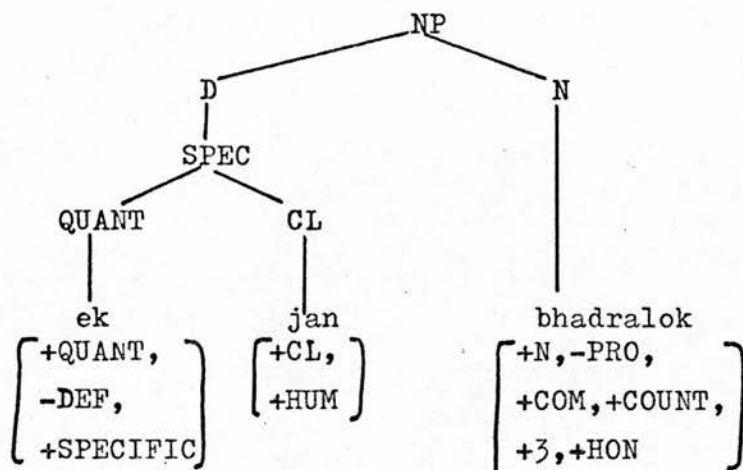
There is no CL in pratyek chele in (40a), and accordingly it is not as forceful as pratyek̄ti chele in (40b), where definiteness is mixed up with emphasis. In (41a) there is no CL with prati chele, and this makes the noun phrase objectionable, but the CL ti occurs with the QUANT in (41b), which is impeccable.

The CL's in Bengali are interesting syntactically as well as semantically and need a thorough investigation. This, however, is outwith our scope here.

In this section we have restricted ourselves to a discussion of indefinite noun phrases, and tried to show how the constituents of the SPEC combine with one another. We have mentioned before (cf., § 2.2) that the indefinite noun phrases that choose SPEC generate it at the left of the head noun. The rule under discussion generates specific and unspecific indefinite noun phrases. For

example, (27a) has the underlying structure (42).

(42)



(42) is indefinite (cf., § 2.4.5). The first lexical insertion rule will first insert the head noun, secondly the QUANT, and finally CL, as its selection is dependent both on N and QUANT. As the QUANT ek has been taken here as [+SPECIFIC], a CL should be selected. The CL here must be jan because the head noun is [+HON]. As (42) is indefinite, this structure will not undergo any change, and it will generate ekjan bhadralok, which is (27a).

The CL's and the QUANT's have the lexical entries in the first lexicon in the following manner (cf., § 11.1.5):

(43)	<u>ti</u>	<u>jan</u>	<u>khānā</u>	<u>ek</u> : 'One'
	[+CL, +COUNT, -HON,]	[+CL, +COUNT, +HUM,]	[+CL, +COUNT, -ANI, +EMOT,]	[+QUANT, -DEF, ±SPECIFIC, -PL,]

2.4.4 SPEC → PL.

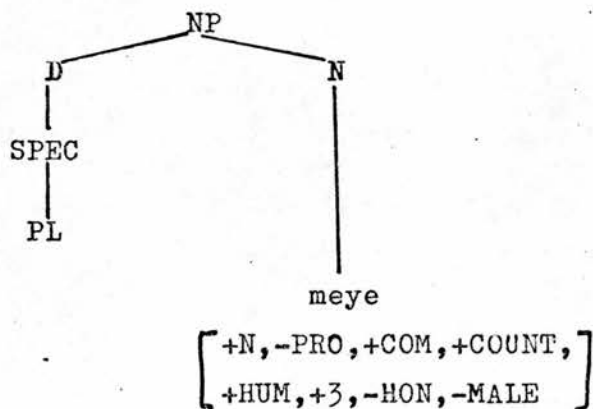
We have discussed two ways of indicating plurality of nouns in Bengali before (cf., § 2.3; § 2.3.1; § 2.3.2). Plurality in one way is indicated by (say, plural) quantifiers such as du: 'Two', tin: 'Three', bahu: 'Many' etc., and in the other way by the plural markers. Consider the examples below.

(44) a. duṭi meye: Two girls.

b. meyerā: Girls.

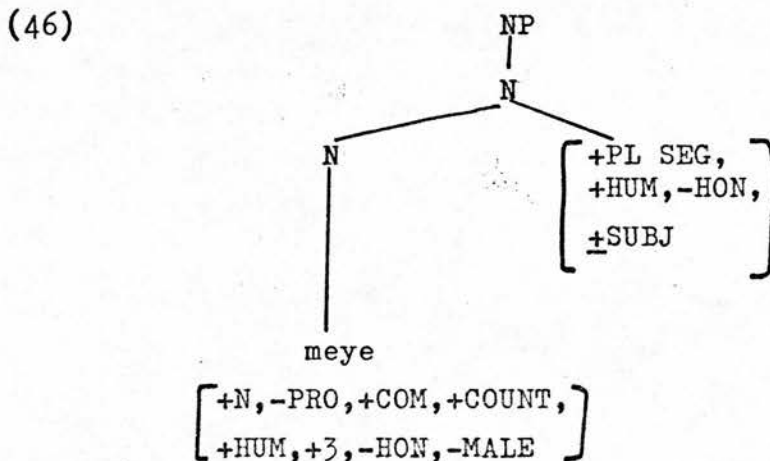
In (44a) the quantifier du indicates the plurality of the noun, although the noun meye is in the base form; and in (44b) there is no quantifier, but the plural marker rā indicates the noun as plural. In our grammar the QUANT is an element of the SPEC, and if the QUANT is plural the head noun will be understood as plural; and as a rule no plural marker will be suffixed to the noun. Let us show how the grammar handles NP's like (44b). We have seen (cf., § 2.2) that NP's unspecified for number have no SPEC or D. In (44a) number is indicated by a QUANT, which is an element of the SPEC. So it is assumed that plurality of NP's like (44b) too, in the underlying structure is indicated by the SPEC. This is intuitively satisfactory in that the same constituent (SPEC) accounts for number in all sorts of noun phrase. Singularity is indicated by the quantifier ek: 'One', and plurality by other quantifiers and by plural markers. But plural quantifiers and plural markers never cooccur in Bengali, which we have accounted for by a disjunctive rule in the base (cf., Base Rule 9). The rule under discussion accounts for morphologically marked plural nouns like (44b). The marker PL is an abstract marker in the underlying structure, and it indicates the plurality of the head noun. For example, (44b) has the underlying structure (45).

(45)

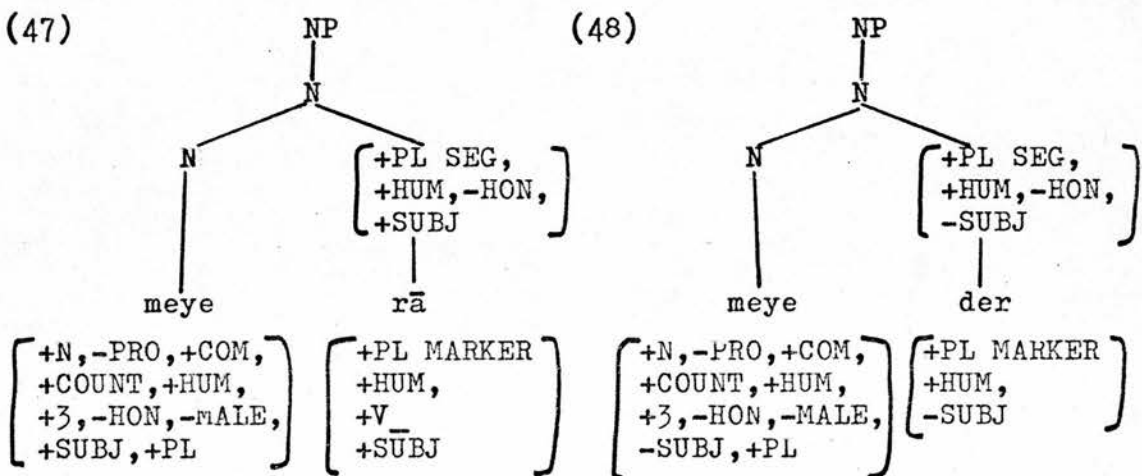


PL in (45) indicates plurality of the head noun, but no plural marker will be attached to it in the underlying structure, because

plural markers in Bengali depend on the functional relation of the noun phrase, and sometimes, on the phonological form of the noun. A later plural segment ($[+PL\ SEG]$) transformation will copy $[\alpha\ HUM, \beta\ HON]$ features of the noun and the functional feature of the noun phrase, and it will be adjoined to the N as its right daughter by Chomsky-adjunction. This transformation will delete the marker PL and its dominating nodes, and will mark the head noun as $[+PL]$. This will transform (45) into (46).



The second lexical insertion rule will attach the appropriate plural marker to the plural segment. For example, if the head noun in (46) has the feature $[+SUBJ]$, we will get (47), and if it is $[-SUBJ]$, we will get (48) after the second lexical insertion.



(47) will give rise to meyerā, which is (44a), and (48) will give rise to meyeder: 'Girls'. The plural segment transformation will

take place when PL is present in the underlying structure.

The base rule 9 shows that PL and QUANT (CL) do not cooccur, but the other constituents of D can be selected with PL. This is shown in the following examples.

- (49) a. oi pākhīgulo: Those birds.
 b. ei chelerā: These boys.

In (49) DEIC(tic) and PL cooccur, and the plural markers are derived by the plural segment transformation.

2.4.5 NP --> (DEF)(SPEC) N.

Broadly, noun phrases are either definite or indefinite. In this section we will deal with such NP's in their underlying and surface forms, and show how they can be handled. To begin with consider the examples below.

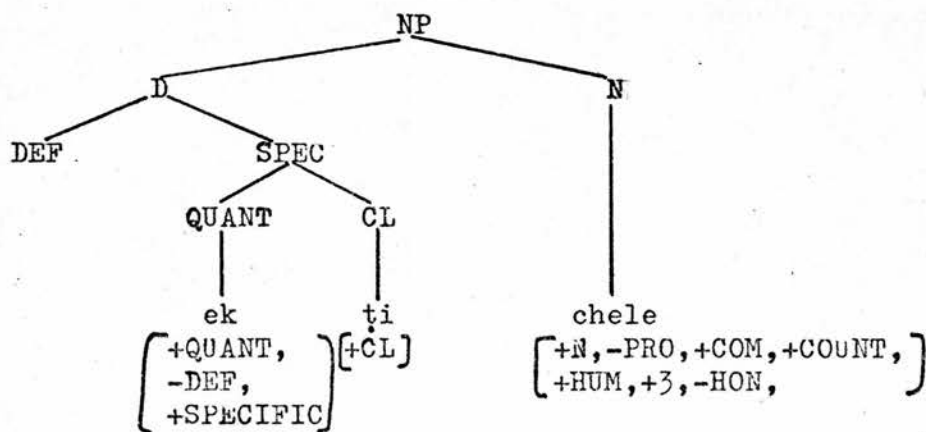
- (50) a. ekṭi chele: A boy.
 1 2 1 2
 b. cheleṭi: The boy.
 (51) a. duṭi chele: Two boys.
 1 2 1 2
 b. chele duṭi: The two boys.
 (52) a. *e ekṭi chele: This boy.
 1 2 1 2
 b. e-cheleṭi: This boy.
 1 2 1 2
 (53) a. ?e duṭi chele: These two boys.
 1 2 3 1 3 2
 b. e-chele duṭi: These two boys.
 1 2 3 1 3 2

The NP's (50, 51a) are indefinite, and (50, 51b) are definite. The SPEC's ekṭi and duṭi in (50, 51a) respectively, are at the left of the head nouns, and these NP's are understood as indefinite. In (50b) the CL ṭi is suffixed to the head noun, and in (51b) the SPEC is at the right of the head noun, and these NP's are definite. This shows that indefinite NP's in Bengali have the SPEC at the

left and the definite NP's have the SPEC at the right of the head noun.

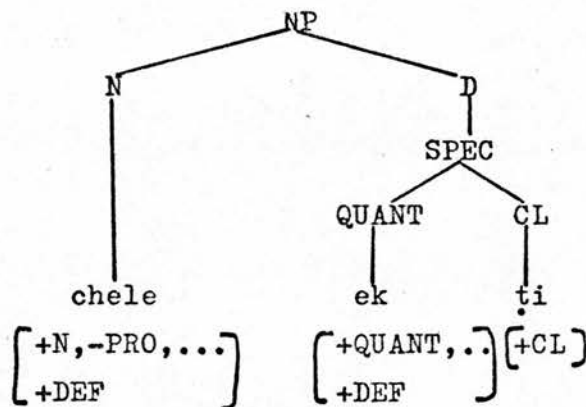
There is nothing in Bengali that can be called a definite or indefinite article. Determined noun phrases are usually understood as definite or indefinite according to the properties of the QUANT (for example, pratyek, prati: 'Each' are inherently definite), or DEIC, and according to the position of the SPEC in relation to the head noun. We will assume that both indefinite and definite NP's have the SPEC at the left of the head noun in the underlying structure. Whether an NP is definite or indefinite will be indicated by (a) the presence or absence of the marker DEF, and (b) by the inherent features of the deictics and quantifiers. The presence of DEF in the underlying structure of SPEC N type noun phrases will indicate definiteness, its absence will indicate indefiniteness. And if the DEF is present in the underlying structure the SPEC will be moved to the right of the head noun. In this way we can simplify the base rules, and show the relation between an indefinite NP and its definite counterpart. Accordingly we will propose a rule called SPEC MOVEMENT for those NP's which are definite and have the SPEC at the left of the head noun in the underlying structure. (50b), for example, has the underlying structure (54).

(54)



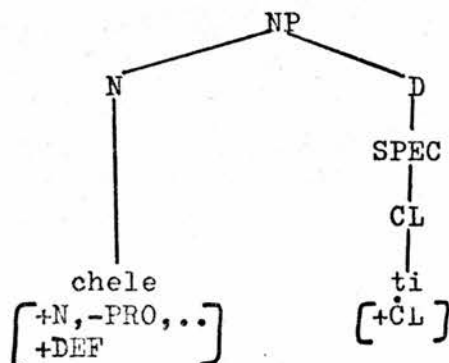
Suppose there were no DEF in (54), then it would be understood as indefinite and would underlie (50a). Because DEF is selected it is understood as definite. We have said earlier (cf., § 2.4.3) that quantifiers can be [+DEF], or [-DEF,+SPECIFIC]. The DEF restricts the choice of a quantifier, that is, if DEF is chosen, a [+DEF] or [+SPECIFIC] quantifier should be chosen. In (54) the quantifier is [+SPECIFIC]. The SPEC MOVEMENT rule will apply to (54). The rule will move the SPEC to the right of the N, change the feature [-DEF] of the quantifier to [+DEF], mark the head noun as [+DEF], and delete the marker DEF. The application of this rule will transform (54) into (55).

(55)



(55) will generate *chele ekti, which is ungrammatical due to the presence of the quantifier ek. The general rule is that the QUANT must be deleted after the SPEC MOVEMENT rule, if the quantifier is ek. The quantifier ek is deleted because it is recoverable. Other quantifiers are not deleted. The deletion of ek will generate (56) from (55).

(56)



(56) will generate cheleṭi, which is (50b). Thus we see that the definite and indefinite noun phrases in (50, 51) are related, and this also shows that CL's are constituents of the SPEC. We have said that if SPEC dominates a quantifier which is ek, it will be deleted after SPEC MOVEMENT. Now consider the examples below.

(57) a. ekjan bhadralok: A gentleman.
 1 2 1 2

b. *bhadralok ekjan:

c. *bhadralokjan:

d. bhadralok: The gentleman.

(58) a. dujan bhadralok: Two gentlemen.

b. bhadralok dujan: The two gentlemen.

(57a) is an indefinite NP, where the SPEC is at the left of the head noun. In (57b) the SPEC is at the right of the head noun, and the NP is ungrammatical. In (57c) the quantifier has been deleted after SPEC MOVEMENT, but still the noun phrase is ungrammatical. In (57d) the NP has no SPEC. It is grammatical, and is definite. In (58a) the SPEC is at the left of the head noun and this NP is indefinite. In (58b) the NP, which has the SPEC at the right of the head noun, is definite. In the examples (57) the head noun is [+HON], QUANT is ek and CL is jan. In (58) the head noun is [+HON], QUANT is du and CL is jan. As we see that SPEC MOVEMENT operates naturally in (58b), it is assumed that it operates in (58d) too; but the entire SPEC is deleted when the head noun is [+HON], QUANT is ek and CL is jan. The entire SPEC deletion takes place when the QUANT is ek, CL is jan and the head noun is [+HON]; but with other CL's and head nouns only the QUANT is deleted if it is ek. This can be seen in the following examples.

(59) a. ekkhānā śāri: A sari.

b. śārikkhānā: The sari.

- (60) a. dukhānā 'sāri: Two saris.
 b. 'sāri dukhānā: The two saris.

In (59, 60a) , which are indefinite noun phrases, the SPEC is at the left of the head noun. In (59b) ek has been deleted after SPEC MOVEMENT, leaving the CL khānā behind; but in (60b) the QUANT, being du, is left as it is.

We have said that if DEF is selected, the SPEC should select a [+SPECIFIC] quantifier. Consider the examples below.

- (61) a. bahu lok: Many men.
 b. *lok bahu:

In (61a) the quantifier is bahu, which is [-SPECIFIC]. If the SPEC MOVEMENT rule operates here, we will get (61b), which is ungrammatical. We have specified pratyek, prati: 'Each' as [+DEF]. These quantifiers can be selected with DEF, but as these are inherently definite, the SPEC MOVEMENT rule will not apply when they are selected. Consider the examples below.

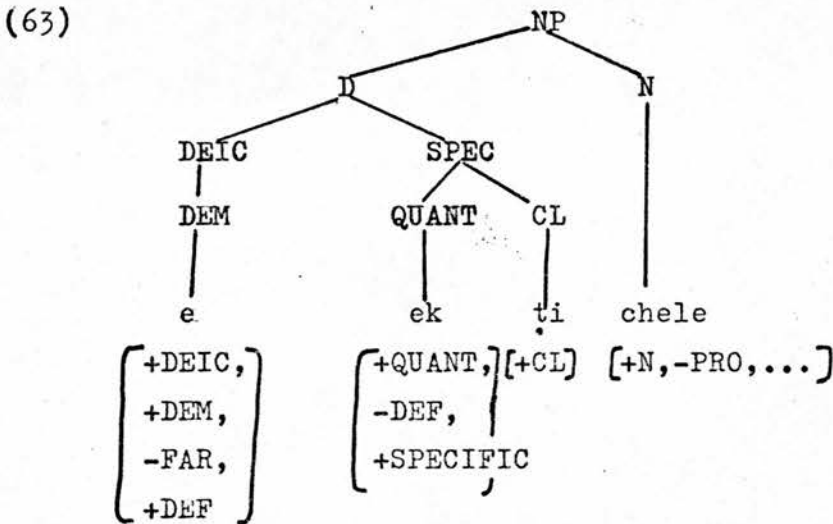
- (62) a. pratiṭi meye: Each girl.
 b. *meye pratiṭi:

In (62a) pratiṭi is at the left of the head noun, but the noun phrase is definite due to the inherent definiteness of prati. In such a case the SPEC MOVEMENT rule does not apply because this will derive ungrammatical noun phrases like (62b).

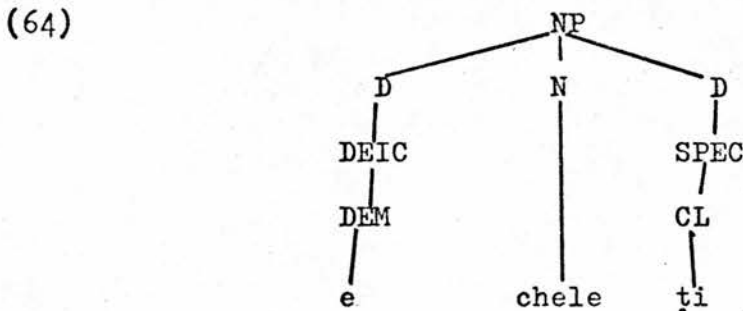
Now we come to the examples (52, 53). The NP's *e ekṭi chele and *e duṭi chele in (52, 53a) respectively are unacceptable.² In these NP's the definite demonstrative deictic e has been selected,

2 The NP e duṭi chele is marginally acceptable. Here the SPEC should be placed at the right of the head noun, but in careless speech it is not moved if the quantifier is anything other than ek and the deictic is definite.

which indicates the NP to be definite. As these NP's are definite the SPEC in these noun phrases should be moved to the right of the head noun. The SPEC MOVEMENT rule operates here exactly as we have shown above. We posit no DEF in the underlying structure of these noun phrases because the inherent feature [+DEF] of the deictic serves this purpose. For example, (52b) has the underlying structure (63).



The SPEC MOVEMENT and the subsequent ek deletion rule will apply to (63), and we will get (64).



(64) will generate e-cheleti, which is (52b). We will come back to this type of noun phrase in the next section.

2.4.6 NP --> (DEIC)(SPEC) N.

The constituent DEIC includes the deictic elements such as e(i): 'This', o(i): 'That, in sight', se(i): 'That, out of sight' and the ordinals such as pratham: 'First', dvitiya: 'Second' and so on.

The DEIC usually cooccurs with SPEC in Bengali. Consider the examples below.

(65) a. oi cheleṭi: That boy.

b. ?oi chele:

In (65a) DEIC and SPEC cooccur, but in (65b) the DEIC occurs with the head noun without any SPEC. (65a) is impeccable, (65b) is not. The deictics e(i), o(i) and se(i) are inherently definite, but they usually do not occur without a SPEC. That is why (65b), although semantically all right, is syntactically objectionable. Definiteness of NP's with definite deictics is indicated by the inherent feature [+DEF] of the deictics, but if a [-DEF] deictic is chosen the NP will be understood as indefinite. And if the chosen deictic is [+DEF], this feature will perform all the functions of the abstract morpheme DEF, which is not posited in the underlying structure of the NP's with [+DEF] deictics. That is, the feature [+DEF] of a deictic will restrict the selection of quantifiers and trigger SPEC MOVEMENT transformation. We have already shown how the [+DEF] feature of the definite deictics trigger SPEC MOVEMENT transformation (cf., § 2.4.5). The items that are considered as demonstrative deictics are listed below.

(66) e(i): 'This' o(i): 'That' se(i): 'That' kon: 'Which'

$\left[\begin{array}{l} +\text{DEIC,} \\ +\text{DEM,} \\ -\text{FAR,} \\ +\text{DEF} \end{array} \right]$	$\left[\begin{array}{l} +\text{DEIC,} \\ +\text{DEM,} \\ +\text{FAR,} \\ \text{IN SIGHT} \\ +\text{DEF} \end{array} \right]$	$\left[\begin{array}{l} +\text{DEIC,} \\ +\text{DEM,} \\ +\text{FAR,} \\ \text{OUT OF SIGHT} \\ +\text{DEF} \end{array} \right]$	$\left[\begin{array}{l} +\text{DEIC,} \\ +\text{DEM,} \\ +\text{DEF,} \\ +\text{INT} \end{array} \right]$
------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------

The demonstrative deictics can cooccur with PL (cf., § 2.4.4) as can be seen in (67).

(67) a. e-meyerā: These girls.

b. oi pākhīgulo: Those birds.

We have taken the view that the head noun is an obligatory constituent of a noun phrase, but there are some noun phrases which apparently lack any head noun. Consider the examples in (68).

- (68) a. e-ṭi bai: This is a book.
 b. e-gulo bai: These are books.
 c. o-ṭi kukur: That is a dog.

The noun phrases e-ṭi, e-gulo and o-ṭi in (68a, b, c) respectively superficially lack any head noun, and they show that the deictics can have a suffixed CL; and they can be pluralized. But our view is that they have deleted head nouns. The head nouns are omitted here as the predicate nominals spell out the objects that the deictics point to. We would claim that the examples in (68) are related to those in (69) respectively.

- (69) a. e-bastuṭi bai: This thing is a book.
 b. e-bastugulo bai: These things are books.
 c. o-jantuṭi kukur: That animal is a dog.

The head nouns in (69) have been deleted and the morphophonological rules have contracted the CL's and the plural marker with the deictics, generating the noun phrases in (68). This deletion is similar to pure contextual deletion of head nouns and the deletion of head nouns due to lexical identity - as can be seen below.

- (70) a. āmi e-ṭi cāi: I want this.
 1 2 3 1 3 2

- b. āmi se-ṭi cāi nā: I do not want that.

- (71) a. minur ekṭi kukur āche ār mātiner-o ekṭi kukur āche:
 2 3 4 5 7 8 9 10

Minu has a dog and Matin has a dog, too.
 4 2 3 5 10 8 9 7

- b. minur ekṭi kukur āche ār mātiner-o ekṭi āche:
 2 3 4 5 7 8 9

Minu has a dog and Matin has one, too.
 4 2 3 5 9 8 7

The NP's e-ti and se-ti in (70a, b) respectively, have no head noun as they are contextually understood, but nouns should be supplied if any misunderstanding develops. Similarly, we get (71b) by deletion of the head noun of the forward noun phrase due to lexical identity (cf., Chapter 9). Thus we see that the NP's that superficially lack any head noun have some deleted or understood head noun.

We have taken the ordinals as deictics because they are used as pointers like demonstrative deictics. The DEM and ORD, although rarely, can cooccur as shown below.

- (72) a. oi pratham meyeṭi: That first girl.
 b. sei pratham ciṭhiṭi: That first letter.

In (72) DEM, ORD and SPEC cooccur, and so the SPEC MOVEMENT rule applies here as usual. The following examples, where ORD and SPEC cooccur, are more natural than (72).

- (73) a. pratham meyeṭi: The first girl.
 b. pratham ciṭhiṭi: The first letter.

The ordinals are listed in the first lexicon with feature specifications as shown below.

- | | |
|-----------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------|
| (74) <u>pratham</u> : 'First' | <u>dvitiya</u> : 'Second' |
| $\left[\begin{array}{l} +\text{DEIC}, +\text{ORD}, \\ +\text{DEF} \end{array} \right]$ | $\left[\begin{array}{l} +\text{DEIC}, +\text{ORD}, \\ +\text{DEF} \end{array} \right]$ |

2.4.7 PART.

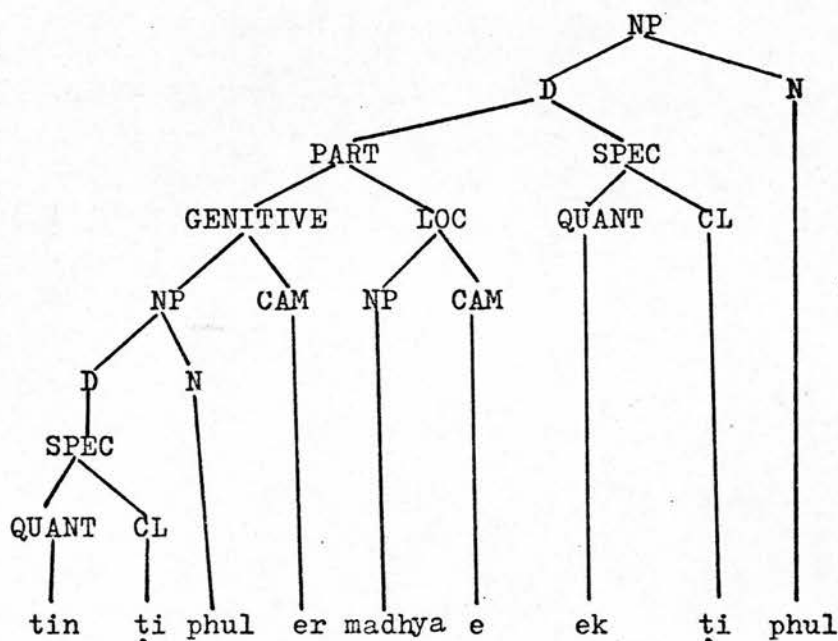
The PART(itive) accounts for the NP's of the following type:

- (75) a. tinṭi phuler madhye ekṭi: One of three flowers.
 b. tinṭi phuler madhye ekṭi phul: A flower of three flowers.
 (76) a. meyeder madhye dujan: Two among the girls.
 b. meyeder madhye dujan meye: Two girls among the girls.

The sentences in each set above are synonymous. The only

difference between the NP's in each set is that in the (a) examples the head nouns are deleted, but in the (b) examples they are present. We assume that the (a) examples in (75, 76) are derived from the (b) examples by deletion of the head nouns. A possible shallow structure for (75b) is given in (77).

(77)



In (77) the head noun of the superordinate NP and that of the genitive NP are identical, and each has a SPEC. In such a structure the head noun deletion can operate both forwards and backwards. If the deletion operates forwards, we will get (75a), and if it operates backwards, we will get tin̄ti madhya e ek̄ti phul̄: 'A flower among three'. This deletion operates when the head noun of the superordinate noun phrase and that of the genitive noun phrase are identical and the deletable noun phrase has a SPEC consisting of QUANT and CL.

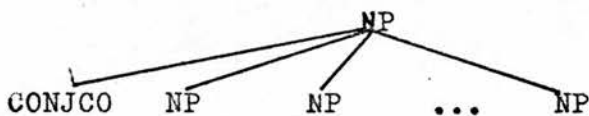
2.4.8 NP --> CONJCO NP NP*.

This rule is for the conjunction of the NP's in the base. Transformational grammarians hold differing views about conjunction. Some claim that all conjunctions are underlying sentence

conjunctions (cf., Gleitman (1965), Bellart (1966), Schane (1966) and Stockwell et al (1973)), while some others claim that all conjunctions are underlying phrasal conjunction (cf., Wierzbicka (1967), Dougherty (1970, 1971) and McCawley (1968b)). Still another group considers both phrasal and sentence conjunctions as basic (cf., Smith (1965), Lakoff and Peters (1966) and Ross (1967a, 1970)). Lakoff and Peters (1966) have shown that there are some 'symmetrical predicates' which cannot be explained if phrasal conjunction in the base is not taken into account. We also have taken both phrasal and sentence conjunctions as basic. The rule under discussion is a notational variant of Lakoff and Peters' (1966, 114) rule.

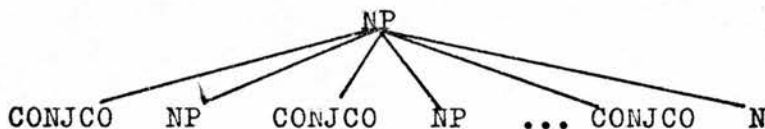
This rule asserts that a noun phrase in the underlying structure can be a coordinate conjunction of two or more noun phrases. The CONJCO of the rule may be filled by o, ar, eban: 'And', and bā, athabā: 'Or', each having the inherent feature [+CONJCO]. The asterisk used in this rule is an iteration symbol asserting occurrence of two or more noun phrases. This rule generates base structures of the form (78).

(78)



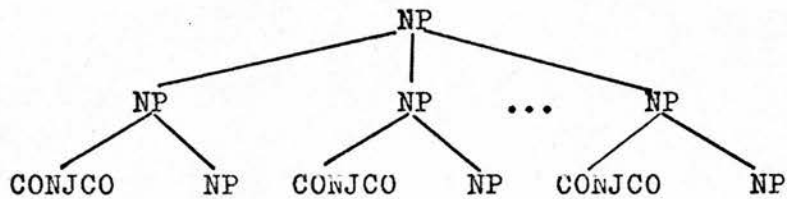
By an universal principle structures of the form (78) will convert into the form (79).

(79)



Our principle differs in this respect from that of Lakoff and Peters' (1966, 114), who will derive (80) from (78).

(80)



The difference is due to the fact that we consider that the conjunctions o, ār, ebaṅ and bā, athabā used in coordinate NP's are directly dominated by the superordinate NP as in (79) without any intermediary NP as in (80).

In (79) the initial conjunction is present, which if left undeleted, will give rise to ungrammatical structures like (81).

- (81) a. *ebaṅ ketakī ebaṅ hāsān: And Ketaki and Hasan.
 b. *athabā cheleṭi athabā meyeṭi: Or the boy or the girl.

The ungrammaticality of (81a, b) is due to the presence of the initial conjunctions. So we have a rule (which can be violated stylistically) that will delete the initial conjunction. We have another (optional) rule that will delete all, but the last, conjunctions when more than two NP's are conjoined. Thus we will arrive at the following sentences from the same underlying structure.

- (82) a. ekṭi chele o ekṭi meye o ekṭi beṛāl: A boy and a girl and a cat.
 b. ekṭi chele, ekṭi meye o ekṭi beṛāl: A boy, a girl and a cat.

In Bengali o, ār, ebaṅ, bā and athabā are used both as sentence conjunctions and noun phrase conjunctions, but the conjunction kintu: 'But' is exclusively used as a sentence conjunction. Consider the examples in (83).

- (83) a. bādal $\left\{ \begin{array}{l} o \\ \bar{a}r \\ ebaṅ \\ b\bar{a} \\ athab\bar{a} \\ *kintu \end{array} \right\}$ nāju yābe: Badal $\left\{ \begin{array}{l} \text{and} \\ \text{or} \\ *but \end{array} \right\}$ Nazu will go.

b. bādal yābe $\left\{ \begin{array}{l} \text{o} \\ \text{ār} \\ \text{ebaṅ} \\ \text{bā} \\ \text{athabā} \\ \text{*kintu} \end{array} \right\}$ nāju yābe: Badal will go

$\left\{ \begin{array}{l} \text{and} \\ \text{or} \\ \text{*but} \end{array} \right\}$ Nazu will go.

c. bādal yābe, kintu nāju yābe nā: Badal will go but
Nazu will not go.

(83a) shows that NP's can be conjoined by all but kintu conjunctions, and (83b) shows that non-contrasting conjuncts can be conjoined by all the conjunctions bar kintu. (83c) shows that contrasting conjuncts can be conjoined by kintu.

The main reason for postulating NP conjunction in the base is that of explaining the sentences with 'symmetrical predicates' (cf., Lakoff and Peters (1966)). The sentence given in (84a) cannot be derived from the underlying conjoined sentence (84b).

(84) a. keyā o ketakī ekrakam: Keya and Ketaki are similar.

b. *keyā ekrakam o ketakī ekrakam: Keya is (of) a kind
and Ketaki is (of) a kind.

We cannot derive (84a) from (84b) as the latter does not make sense. The predicate ekrakam: 'Similar, same kind' requires a plural subject noun phrase in the underlying structure. This shows that keyā o ketakī is a coordinate noun phrase in the base.

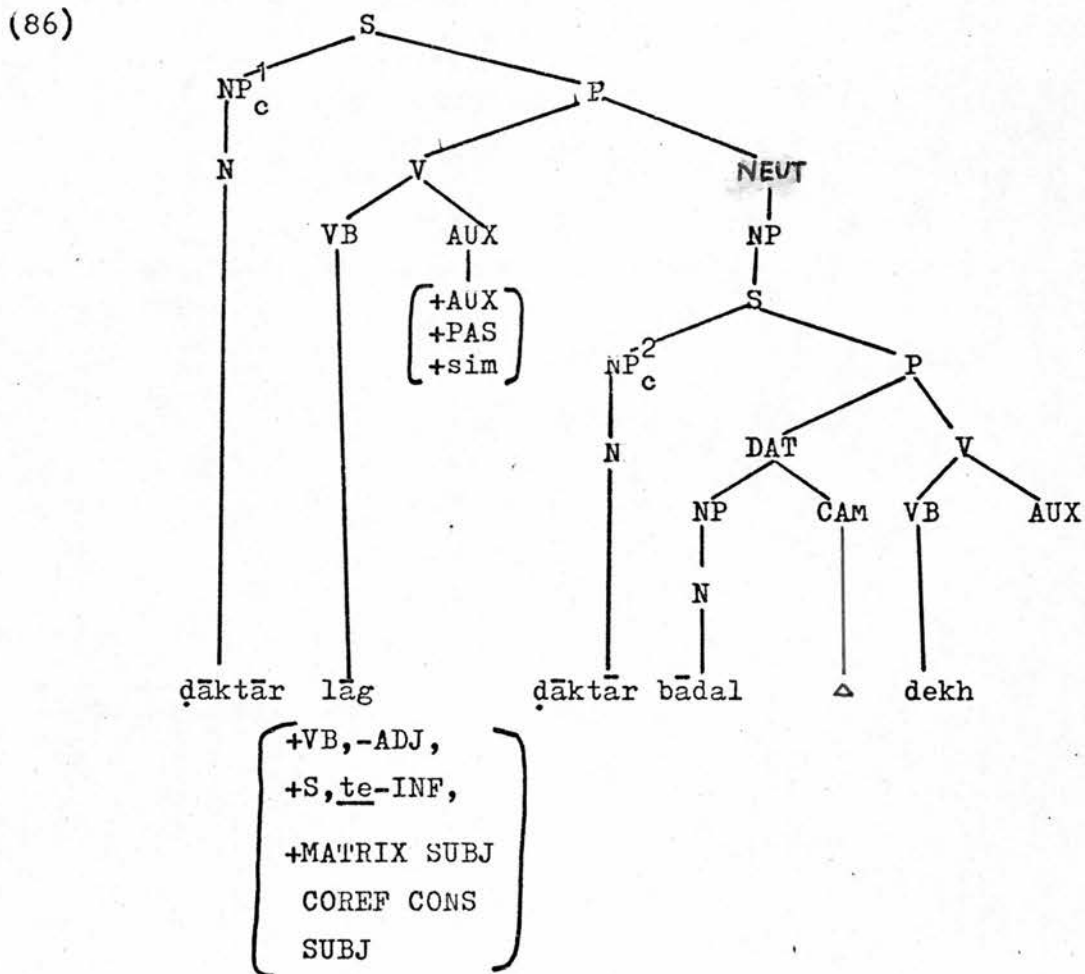
2.4.9 NP --> S.

This rule is for infinitival complementation (cf., § 10.3). All underlying embeddings in this grammar are NP embeddings, and the dominating NP's themselves are dominated by some case according to their relation to the verb. In infinitival

complementation the complement clauses are reduced due to te-infinitivalization of the constituent verb. Consider the examples below.

- (85) a. $\underset{1}{\text{dāktār}} \underset{3}{\text{bādalke}} \underset{4}{\text{dekhte}} \underset{1}{\text{lāglen}}$: 'The doctor began to
 examine Badal.
 $\underset{3}{\text{3}}$
- b. $\left[\underset{1}{\text{dāktār}} \underset{2}{\text{lāglen}} \left[\underset{3}{\text{dāktār}} \underset{5}{\text{bādalke}} \underset{1}{\text{dekh}} \right] \right]$: $\left[\left[\underset{1}{\text{The doctor}} \right] \right]$
 began $\left[\underset{2}{\text{the doctor}} \underset{3}{\text{examine}} \underset{5}{\text{Badal}} \right]$.

(85a) is derived from the underlying structure (85b), which has the intermediate structure (86).



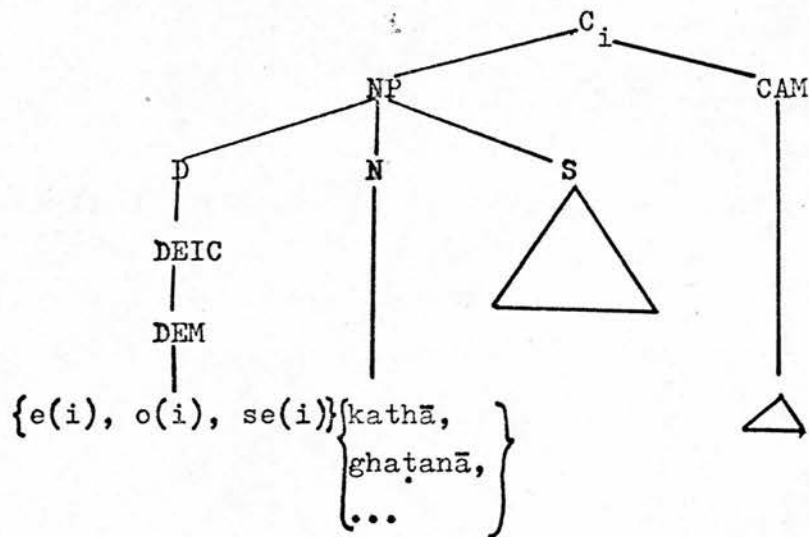
In (86) the matrix verb is marked for te-infinitivalization of the constituent S. So the constituent AUX is empty (that is, the constituent S is tenseless and aspectless). The constituent

subject in (86) will be deleted under coreference (NP^1 will be used to delete NP^2). So the Subject-AUX concord rule will not apply in the constituent sentence. Instead te will be inserted to the constituent AUX, and we will get ḍāktār lāḡlen bādalke dekhte, which is synonymous with (85a). Subsequently, by placing the finite verb form sentence-finally, we will get ḍāktār bādalke dekhte lāḡlen, which is (85a).

2.4.10 NP --> D N S.

This rule is for NP-Complementation (cf., § 10.3). The underlying structure for such complementation is something like (87).

(87)



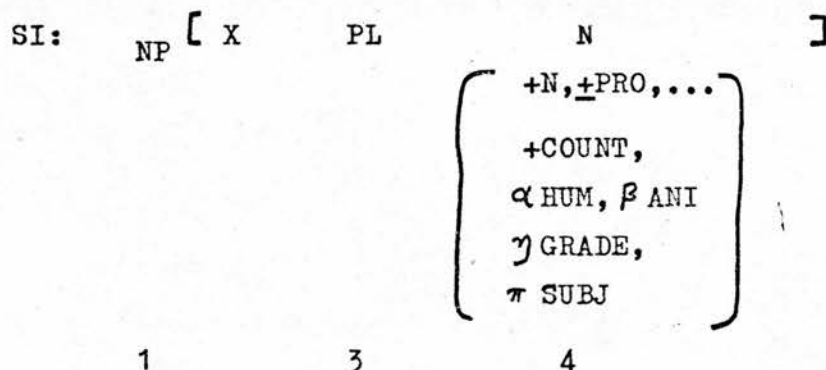
This is similar to Kiparsky and Kiparsky's (1971) structure for factive complementation in English. The place occupied by 'the fact' in their treatment is filled here by abstract noun phrases like e-kathā: 'This proposition', e-ghaṭanā: 'This event' etc.,. Consider the examples below.

- (88) a. minu mane kare ye āḡāmīkāl briṣṭi habe: Minu thinks that it will rain tomorrow.
- b. minu mane kare e-kathā ye āḡāmīkāl briṣṭi habe: Minu thinks this proposition that it will rain tomorrow.

The sentences (88a, b) are synonymous, but while the abstract noun phrase e-kathā is present in (88b), it is absent from (88a). We assume that the complement sentence āgāmīkāl bṛiṣṭi habe is a complementation of the abstract noun phrase e-kathā, which has been deleted in (88a) (cf., § 10.3).

2.5 Summary of the Major Rules Discussed.

(89) PLURAL SEGMENT TRANSFORMATION.

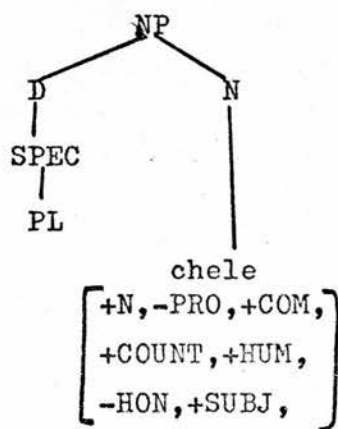


- SC: a. Form a node [+PL SEG] and adjoin it as the right daughter of 4 by Chomsky-adjunction.
- b. Copy the features [α HUM], [β ANI], [γ GRADE], [π SUBJ] of 4 onto the [+PL SEG].
- c. Add the feature [+PL] to 4 and delete 3.

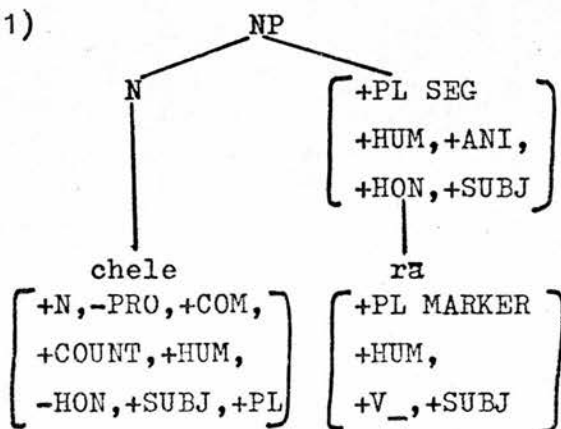
Condition: 3 and 4 are constituents of 1.

This rule will transform (90) into (91).

(90)



(91)



The plural marker rā is attached to the plural segment in (91) by the second lexical insertion rule.

(92) SPEC MOVEMENT.

SI: NP [X { DEF } SPEC [QUANT CL] N]
 1 3 4 5 6 7

- SC: a. Attach 4 as the right sister of 7.
 b. Add the feature [+DEF] to the CS of 7, and change the [-DEF] feature of 5 to [+DEF].
 c. Delete 3 if it is DEF.

- Conditions: a. If 3 is DEIC, it must be [+DEF].
 b. 5 is [+SPECIFIC].
 c. This rule does not apply if 5 is [+DEF].

(93) ek. and jan DELETION.

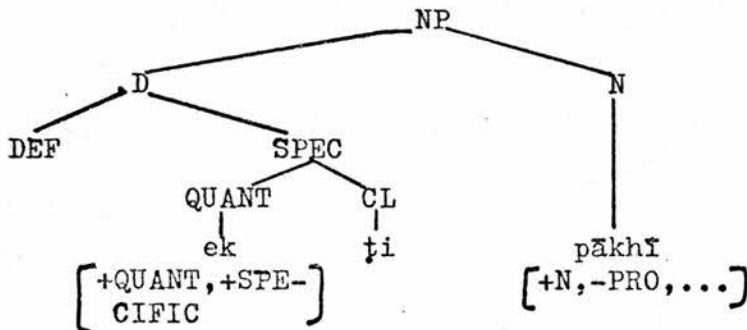
SI: NP [X N SPEC [QUANT CL]]
 1 3 4 5 6

- SC: a. Delete 5 if it dominates ek: 'One'.
 b. Delete 4 with all its constituents if 5 is ek, and 6 is jan.

Condition: 3 and 5 have the feature [+DEF].

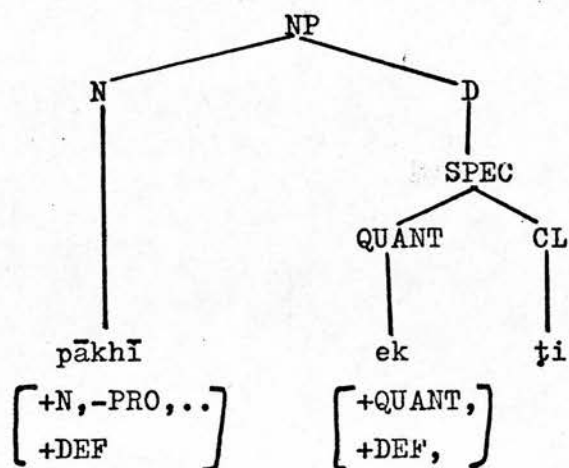
The rule (92) will apply to an underlying noun phrase like (94).

(94)



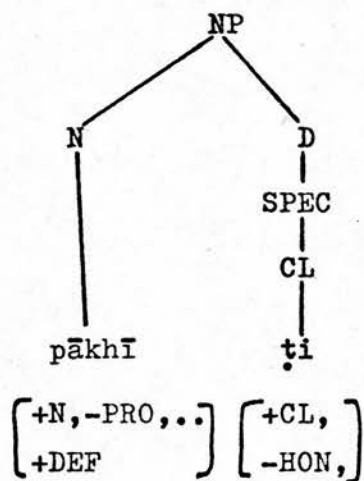
(94) will be transformed into (95) by the application of the rule (92).

(95)



The structure change (93a) will apply to (95), and (95) will be transformed into (96).

(96)



(96) will generate pākhīṭi: 'The bird'.

CASES AND CASE MARKERS3.0 Introduction.

In this work we have adopted the case grammar framework of Fillmore (1968a), and developed it in the line of Stockwell et al (1973) in order to describe the syntax of Bengali. Our main purpose is the investigation of the process of pronominalization in Bengali, and not an exhaustive investigation into its case system. A case grammar framework has been adopted in the base of the present work not because it handles pronominalization more adequately than an interpretive or generative constituent structure model, but because it shows deeper insight into the nature and structure of natural languages in general (cf., Fillmore (1968a), Anderson (1971), Stockwell et al (1973) and Starosta (1974)). Although the underlying cases have minor part to play in our main topic pronominalization, we thought it wise to adopt the case grammar framework considering Bengali in its total perspective. The direct profit the present work derives from the case grammar framework is that we do not have to handle the case markers in an ad hoc manner as is done in a constituent structure model, and we can account for the relatively free constituent order of Bengali easily, as the case grammar allows much more freedom than a rigid Chomskyan framework.

There are at least two models of case grammars available now. Fillmore (1966b, 1968a, 1968b, 1969, 1971) proposed his theory of case grammar as an alternative to the Chomskyan model of transformational generative grammar. A few years later Anderson (1971) proposed his localistic theory of underlying cases.

Of these two competing models, we have adopted the fillmorean model because it sets comfortably with Chomsky's theory of transformational grammar, and poses no problem for our main topic: pronominalization. The theoretical reason apart, practical reasons also worked behind the the adoption of a Fillmorean framework. It has been widely used, consequently a workable model has been built up. Anderson's (1971) localistic model is not yet fully developed, and is passing through a period of development by the author and his colleagues. Fillmore's framework has come under attack recently for its lack of any well-defined criteria for limiting the number of cases. Fillmore considers cases as atomic concepts, which proliferates the number of cases with every new role played by nominals. But we hope that the model can be saved if some cases, which are considered as distinct, are considered as non-distinct. For this the cases should be understood more abstractly than they have been so far in this model. We shall not attempt any such innovation here, as it is beyond our scope. The few cases that we have taken will serve our immediate purpose: pronominalization.

3.1 Case Grammars.

The case is not a morphological, but a semantic notion in a case grammar. The case grammars do not ignore the morphological realizations of different underlying cases, but they mean by the term 'Case' 'underlying case' or 'deep case', not merely their superficial realizations. The case is a semantic primitive term, which indicates 'certain semantically relevant syntactic relationships involving nouns and the structure that contains them' (cf., Fillmore (1968a, 5)). The base of a case grammar is

semantic rather than syntactic. Fillmore (1968a, 21) says:

The sentence in its basic structure consists of a verb and one or more noun phrases each associated with the verb in a particular case relationship. The 'explanatory' use of this framework resides in the necessary claim that, although there can be compound instances of a single case (through noun phrase conjunction), each case relationship occurs only once in a simple sentence.

The case is an abstract relational notion between a noun phrase and the verb of a sentence. Cases are not uniquely associated with any noun phrase. It is the verb which decides the case relationships of the noun phrases that cooccur with it.

The Fillmorean case grammar has a different claim from that of any constituent structure model about the underlying structure of natural languages. Fillmore (1968a, 23) claims:

In the basic structure of sentences, then, we find what might be called the 'proposition', a tenseless set of relationships involving verbs and nouns (and embedded sentences, if there are any), separated from what might be called the 'modality' constituent. This latter will include such modalities on the sentence-as-a-whole as negation, tense, mood and aspect.

The first rule in a Fillmorean case grammar is (1):

(1) $S \rightarrow M + P$

where S stands for 'sentence', M for 'modality' and P for 'proposition'. The proposition is again rewritten as a verb and an array of cases (Fillmore (1968a) took six cases and felt the need for some 'additional cases'.). The notional nature of cases is clear in the following passage quoted from Fillmore (1968a, 24):

The case notions comprise a set of universal, presumably innate, concepts which identify certain types of judgements human beings are capable of making about the events that are going on around them, judgements about such matters as who did it, who it happened to, and what got changed.

With the adoption of underlying cases, the traditional functional notions subject and object become superficial or intermediate underlying notions, not deep structure notions. The subject and object are syntactic notions which cannot indicate the roles played by the noun phrases holding those functions. Consider the sentences in (2).

- (2) a. bādal jānālāṭi khuleche: Badal has opened the window.
 b. jānālāṭi khule geche: The window has opened.

bādal is the subject in (2a), and jānālāṭi in (2b). They behave similarly syntactically in that the verb form agrees with bādal in (2a), and with jānālāṭi in (2b). But bādal in (2a) acts as an Agent, who opened the window volitionally, jānālāṭi in (2b) does not act volitionally. The roles played by these noun phrases are distinct, and this can be accounted for if we posit two distinct cases for bādal and jānālāṭi. Cases are semantic notions, but in many instances they have syntactic consequences.

Fillmore (1968a) posited six cases, and felt the need for some 'additional cases'; but the total inventory of cases in a Fillmorean case grammar is still uncertain. Fillmore (1968a, 24-25) posited six cases: Agentive, Instrumental, Dative, Factive, Locative and Objective. Later on Fillmore (1969) talked about cases like 'Counter-Agent' and 'Source', and replaced his Dative by 'Experiencer' and Factive by 'Result'. His main six cases

are the following:

(3) Agentive (A), the case of the typically animate perceived instigator of the action identified by the verb.

Instrumental (I), the case of the inanimate force or object causally involved in the action or state identified by the verb.

Dative (D), the case of the animate being affected by the state or action identified by the verb.

Factive (F), the case of the object or being resulting from the action or state identified by the verb, or understood as a part of the meaning of the verb.

Locative (L), the case which identifies the location or spatial orientation of the state or action identified by the verb.

Objective (O), the semantically most neutral case, the case of anything representable by a noun whose role in the action or state identified by the verb is identified by the semantic interpretation of the verb itself; conceivably the concept should be limited to things which are affected by the action or state identified by the verb. The term is not to be confused with the notion of direct object, nor with the name of the surface case synonymous with accusative.

In the last few years the Fillmorean case theory has come under attack from various corners (cf., Huddleston (1970)). It has been alleged that:

- a. this theory has no well-defined criteria for restricting the number of cases,
- b. the atomic concept of case differentiates between them and proliferates them in number, but fails to show

similarities among cases;

- c. the distinction between some cases such as Agentive and Instrumental, and Dative and Objective depends on non-casual features animateness and inanimateness.

The allegations (a) and (b) above point to the real flaw of the theory, as it cannot generate all the sentences of a language with a small number of cases. The atomic concept of case is simple for perception, but verbs allow the arguments to play a huge range of roles, and if all these roles are considered as instances of distinct cases, the case inventory will be vast indeed. This points to the fact that Fillmorean cases, although abstract notions, are not abstract enough to make an overall generalization. The third objection is minor. Fillmore, of course, considers that Agents and Datives are 'typically' animate, but it is not the sole criterion that distinguishes an Agentive from an Instrumental, and a Dative from an Objective. It is expected that an Agentive will act volitionally. It is seen that generally animate beings have the power of volitional activity. The animateness of Agents and Datives can be considered as a generalization that works in most of the cases, if not all. The real allegation against this theory is that it needs a vast number of cases to explain the facts of a language. For example, we can refer to Stockwell et al (1973, 743), who used Fillmorean case theory and posited six cases, but felt a need for cases like Benefactive, Comitative, Degree, Manner, Means, Referential, Resultative, Source and Time.

In Anderson's (1971) localistic theory of case, case is understood much more abstractly than in a Fillmorean grammar. His grammar is a localistic dependency grammar, where every role

played by an argument is viewed from the local point view.

This has enabled him to posit only two pairs of cases:

(a) Locative and Ablative, and (b) Ergative and Nominative.

It has been claimed that that these two pairs exhaust all the possible cases in natural languages. The cases taken here are not atomic. They can be broken into complexes of features which enable one to show similarities and dissimilarities among the cases. But this grammar is too complicated, due to continuous subcategorization of verbs and cases, and due to the complex symbols of cases. It depends on too much abstraction, and does not always satisfy the native speakers' intuition about the language they speak. The problem here is opposite in kind to that in Fillmore: Fillmore is not abstract enough, and has too many cases; Anderson is too abstract and has too few cases. One suffers for under-generalization, and the other for over-generalization.

3.2 CASES.

We have taken six cases for our immediate purposes. The base rule 3 develops the P as a V and an array of six cases: ESS(ive), INS(trumental), LOC(ative), NEUT(ral), DAT(ive) and AGT (Agent). The selection of a case depends on the VB that occurs as the head of the proposition. The underlying order of the cases is not a direct reflection of the surface, and so rules will be required to place the underlying elements in an acceptable order in the surface. In the following sections we will deal briefly with the cases that have been taken in this work. Readers should note that cases in Bengali are not our main point of investigation, and we will not attempt any exhaustive investigation there.

Furthermore we shall confess that the six cases that have been taken for our immediate purposes are not enough for a total description of the syntax of Bengali.

3.2.1 AGT and INS.

Fillmore (1968a) considers his Agentive as 'the case of the typically animate perceived instigator of an action', and Instrumental as 'the case of the force or object causally involved in the action.' Our AGT and INS are identical to Fillmore's Agentive and Instrumental, respectively. Fillmore's requirement for an Agentive to be animate and Instrumental to be an inanimate force or object has created much controversy, although this requirement seems basically justified. We will consider that although AGT is typically animate, it is the case of any instigator that acts or is believed to act volitionally; and INS is the case of any force or object involved in the action causally, but does not act volitionally.

Huddleston (1970) has raised an objection against the distinction between the Agentive and the Instrumental, especially when the INS is a natural force. He has proposed that these two should be merged into a single case. Bengali shows evidence that AGT and INS can be distinguished semantically as well as syntactically. Natural forces and inanimate objects behave alike in Bengali and they make a class by themselves in respect of participant role as against the role played by animate beings. Consider the examples below.

- (4) a. minu darajāṭi khulche: Minu is opening the door.
 2 3 -3- -2-

- b. $\text{minu} \left\{ \begin{array}{l} \text{cābi} \\ *jhar\dot{a} \end{array} \right\} \begin{array}{l} \text{diye darajāṭi khulche:} \\ \text{Minu is opening} \end{array} \begin{array}{l} 3 \quad 4 \quad 5 \\ \underline{\quad} \quad \underline{\quad} \quad \underline{\quad} \end{array} \begin{array}{l} \underline{\quad} \quad \underline{\quad} \quad \underline{\quad} \\ \underline{\quad} \quad \underline{\quad} \quad \underline{\quad} \end{array}$
the door with $\left\{ \begin{array}{l} \text{(a) key} \\ \text{gusty wind} \end{array} \right\}$.
 $\underline{\quad} \quad \underline{\quad} \quad \underline{\quad}$ $\underline{\quad}$
- c. $* \left\{ \begin{array}{l} \text{cābi} \\ jhar\dot{a} \end{array} \right\} \begin{array}{l} \text{darajāṭi khulche:} \\ \text{The key} \\ \text{Gusty wind} \end{array} \begin{array}{l} 2 \quad 3 \\ \underline{\quad} \quad \underline{\quad} \end{array} \begin{array}{l} \left\{ \begin{array}{l} \text{The key} \\ \text{Gusty wind} \end{array} \right\} \\ \text{is opening} \end{array} \begin{array}{l} \underline{\quad} \quad \underline{\quad} \\ \underline{\quad} \quad \underline{\quad} \end{array}$
the door.
 $\underline{\quad}$
- d. $\left\{ \begin{array}{l} \text{cābite} \\ *jhar\dot{a} \end{array} \right\} \begin{array}{l} \text{darajāṭi khulche:} \\ \text{The door is opening by} \end{array} \begin{array}{l} 2 \quad 3 \\ \underline{\quad} \quad \underline{\quad} \end{array} \begin{array}{l} \underline{\quad} \quad \underline{\quad} \\ \underline{\quad} \quad \underline{\quad} \end{array}$
 $\left\{ \begin{array}{l} \text{the key} \\ \text{gusty wind} \end{array} \right\}$.
 $\underline{\quad}$
- e. $\text{darajāṭi} \left\{ \begin{array}{l} \text{khulche} \\ \text{khule yācche} \end{array} \right\} : \begin{array}{l} \text{The door is opening.} \\ \underline{\quad} \quad \underline{\quad} \end{array}$
 $\underline{\quad}$ $\underline{\quad}$

The VB khul: 'Open' has the case frame ___(INS)NEUT AGT. It takes AGT and NEUT obligatorily, and INS optionally, and disallows other cases. There is a syntactic rule feature attached to khul, which allows the deletion of the unspecified AGT; but semantically the AGT is presupposed in all instances when khul is used. In (4a) minu is the AGT, who performs the act of opening the door, and darajāṭi is NEUT. By a general rule the AGT has become the subject in (4a). In (4b) minu is the AGT, cābi/jhar are INS and darajāṭi is NEUT. Again, the AGT is the subject in (4b). In minu cābi diye darajāṭi khulche, minu is the instigator of the action of opening which involves the manipulation of cābi. The sentence *minu jhar diye darajāṭi khulche is unacceptable only because minu has no power over jhar, so she cannot manipulate it. If we substitute minu in this sentence by some mythical god, who is

believed to be powerful over wind, this sentence becomes perfectly acceptable. The deviation of this sentence is not because both minu and jhar are AGT's, but because of the fact that the AGT is not powerful enough to utilize the INS. The sentences in (4c) are unacceptable because the INS noun phrases have become the subjects, which Bengali disallows. In the sentence cābite darajāṭi khulche in (4d), there are INS and NEUT, no AGT; but it is understood that the door is opening due to the manipulation of cābi by some AGT. In this sentence NEUT is the subject. The sentence *jhare darajāṭi khulche is unacceptable due to the reason that no immediate unspecified AGT can be thought of. In (4e) we come across a verb khule yā: 'Open (by itself)', which is lexically related to but semantically distinct from khul. khule yā has the case frame __ (INS) NEUT. In (4e) only NEUT is present, but the sentences here are not synonymous. The sentence darajāṭi khulche suggests that the door is opening due to the action of some AGT, which is unspecified. The sentence darajāṭi khule yācche does not presuppose any AGT: the door opens mysteriously by itself. These examples show that the natural force and the inanimate objects behave similarly as INS, and their role is distinct from the role of an AGT.

Bengali does not allow passivization when a sentence contains an overt AGT, but this is possible when an INS is present. Consider the following examples.

- (5) a. cābi diye darajāṭi kholā halo: The door was opened
by (a) key.
- b. *minu kartrik cābi diye darajāṭi kholā halo: The door
was opened by Minu with a key.

(6) a. lāṭhi diye glāṣṭi bhāṅā halo: The glass was broken
by (a) stick.

b. *minu kartrik lāṭhi diye glāṣṭi bhāṅā halo: The glass was
broken by Minu with (a) stick.

(5, 6a) are passive sentences, where there is no overt AGT; but it is understood. But each sentence has an INS. These are natural sentences in Bengali. (5, 6b) are unacceptable due to the presence of an overt AGT, which Bengali disallows.

We have seen above that natural forces and inanimate objects behave alike as INS, and are distinguished from AGT semantically as well as syntactically. This can be seen also when the INS is an abstract object. Consider the examples below.

(7) a. tār rupe āmarā mugdha halām: We became charmed by her
beauty.

b. *tār rup āmāder mugdha karlo: Her beauty charmed us.

In (7a) āmarā is DAT and subject of the sentence, and tār rup is abstract INS. This is a natural sentence in Bengali.

(7b) is unacceptable as the abstract INS has become the subject.

(7b) is acceptable only figuratively.

In English in a sentence containing 'a possessed noun' as INS, either the entire instrumental noun phrase becomes the subject, or the possessor becomes the subject and the instrumental noun phrase appears with the case marker 'with' (cf., Fillmore (1968a, examples 23-27)). In Bengali the 'possessor' can be the subject if it is an underlying AGT, DAT or NEUT. Otherwise the entire instrumental noun phrase appears with the INS case marker, if it is not objectivalized. But it cannot be the subject.

Consider the examples in (8, 9).

- (8) a. banduker gulite jānālāṭi bheṇeche: The window has broken by a gun's shot.
 b. *banduk gulite jānālāṭi bheṇeche: The gun has broken the window with a shot.
- (9) a. anuṣṭhāntir baicitrye āmarā mugdha halām: We became charmed by the varieties of the show.
 b. *anuṣṭhānti baicitrya diye āmāder mugdha karlo: The show charmed us with varieties.

In (8a) the INS is banduker guli. In this sentence the NEUT jānālāṭi is the subject and the INS noun phrase appears in the surface with the case marker te. In (8a) the genitive banduker is an underlying INS, which cannot be the subject. (8b) is ungrammatical because an INS noun phrase is the subject. In (9a) the DAT āmarā is the subject, and the whole instrumental noun phrase anuṣṭhāntir baicitrya appears with the case marker e. (9b) is ungrammatical because an abstract instrumental anuṣṭhānti has appeared as the subject.

All this shows that Bengali distinguishes between an AGT and an INS.

3.2.2 DAT and NEUT.

Fillmore (1968a) distinguishes his Dative and Objective mainly by the features animate and inanimate, and this has been questioned by different linguists (cf., Huddleston (1970)). Fillmore, however, allows both animate and inanimate noun phrases to occupy the Objective case, but allows only animate noun phrases to occupy the Dative case. In our grammar DAT and NEUT are taken as similar to Fillmore's Dative and Objective respectively. Fillmore (1971) changed his mind about these two cases, and postulated an Experiencer case for psychologically affected

nominals cooccurring with psychological verbs, and distributed his Dative among Experiencer, Objective and Goal cases. The definitions of Dative and Objective (our NEUT) as given in Fillmore (1968a) show the close similarity of these two cases, at the same time we need to distinguish them. It has been proposed (cf., Stockwell et al (1973, 733)) that DAT and NEUT can be semantically distinguished by considering the former as the case of affected persons and things, and the latter as the case of unaffected persons and things. But as this semantic distinction is not always reflected on the syntactic structure, we have to depend heavily on intuition, which may vary from person to person. Consider the examples in (10).

- (10) a. se darajāṭi bheṅche: He has broken the door.
 b. se darajāṭi khuleche: He has opened the door.

In (10a) darajāṭi is obviously affected and can be considered as DAT, but darajāṭi in (10b) is not affected similarly. But opening a door does affect it in some way. Shall we then consider darajāṭi as DAT or NEUT? Stockwell et al (1973, 733) discussed the problems with the potential inanimate Datives, and showed that they cannot be systematically dealt with as are animate Datives. So we will stick to the notion that DAT is the case of animate affected beings, and NEUT is the most neutral case which can be interpreted by an interpretation of the cooccurring verb. Thus darajāṭi in (10a, b) will be considered as NEUT.

We consider that each of the sentences in (11) contains a DAT.

- (11) a. hāsān ketakīke bhālabāse: Hasan loves Ketaki.
 b. āmi hāsānke ekṭi bai diyechi: I have given a book to Hasan.
 c. hāsān lokṭike khun kareche: Hasan has killed the man.

In (11a) hāsān is psychologically affected, in (11b) is the recipient of a book, and lokṭi in (11c) mortally affected.

All these noun phrases are taken as DAT. The verb bhālabās: 'Love' has the case frame __ (NEUT) DAT, de: 'Give' has the frame __ (INS) (NEUT) (DAT) AGT, and khun kar: 'Kill' has the frame __ (INS) DAT AGT.

The AGT has a priority over DAT to be the subject of a sentence and the DAT has a priority over NEUT to be the subject of a sentence. And with verbs that take NEUT, DAT and AGT in their case frame, the AGT is realized as the subject, DAT as the indirect object and NEUT as the direct object. But with verbs that take AGT and DAT, or DAT and NEUT, if the NEUT is also a human noun, the underlying distinction of DAT and NEUT is neutralized sometimes in the surface. Consider the examples below.

- (12) a. hāsān lokṭike khun kareche: Hasan has killed the man.
 b. hāsān lokṭike cene: Hasan knows the man.

lokṭi in (12a) is DAT and in (12b) is NEUT, but in the surface both the noun phrases are realized as objects without any formal distinction.

3.2.3 LOC and ESS.

Fillmore (1968a) considers that LOC is 'the case which identifies the location or spatial orientation of the state or action identified by the verb.' This 'spatial orientation' can be of various types depending on the head verb. But many of the locational expressions which apparently show relation with the verb of the sentence can be considered as proposition-external elements. They are usually treated as 'adjuncts' (cf., Lyons (1968, 334), and in our grammar they should be generated in the ADV constituent or in a superordinate S (cf., Fillmore (1971)).

(13c) is related to the sentence hāsān ketakīke grībāy cumu kheyechē: 'Hasan has kissed Ketaki on (her) neck', where ketakī and grībā appear as independent actants. In (13d) the location of the event (kalābhaban) and location of the action (ketakīr grībā) are present. In (13d), the location of the event kalābhaban appears sentence-initially. (13d) is a good and natural sentence. In (13e) the location of the action ketakīr grībā appears sentence-initially. (13e) although grammatical is not a natural sentence. In (13f) the location of the event kalābhaban is present, but no DAT or location of cumu kha is present. This makes the sentence objectionable unless we appreciate that these are understood. In (13g) the location of cumu khā is present, but no location of event is present. This is a natural sentence in Bengali. Now consider what is the relation of kalābhaban to the verb cumu khā in (13b). The surface sentence pretends that the verb and the locational expression have some relation, and so it should be considered as a proposition-internal LOC. But in actuality kalābhaban identifies the location of Hasan and Ketaki's existence. (13b) can be paraphrased as below.

- (14) a. hāsān o ketakī kalābhabane chilo o hāsān ketakīke cumu kheyechē: Hasan and Ketaki were in the Kalabhaban and Hasan has kissed Ketaki.
- b. hāsān o ketakī yakhan kalābhane chilo, takhan hāsān ketakīke cumu kheyechē: When Hasan and Ketaki were in the Kalabhaban, then Hasan has kissed Ketaki.

These examples show that kalābhaban is an existential locative, and originates in a separate sentence, and so it has no relation to the verb cumu khā. It is an actant of the existential locative verb āch: 'Exist'. The verb āch may be at the source of many

locative expressions which superficially show relations with other verbs in the surface. This is because of the fact that āch is not realized in the surface in many cases. In Bengali the proposition-external locative expressions usually come at the sentence-initial position, and the proposition-internal LOC's appear sentence-medially. That is why, (13d), where the proposition-external locative expression is placed sentence-initially, is a natural sentence; but (13e), where the proposition-internal locative is placed sentence-initially, is unnatural.

Although we cannot go into the details here, we assume that most of the 'strictly locative' verbs are stative. We have said that LOC can be of various types. Consider the examples below.

- (15) a. tini bārīte āchen: He is (exists) at home.
 1 2 3 1 3 2
- b. tini ḍhākā theke esechen: He has come from Dacca.
 1 3 4 1 4 3
- c. tini ḍhākā yāben: He will go to Dacca.
 1 3 1 3
- d. tini path diye yācchilen: He was going along the road.
 1 2 3 4 1 4 3 2

The verb āch is both an absolute existential and a locative existential verb. By absolute existence we mean existence of someone or something which cannot be located in space, but can be located in time. Consider the examples below.

- (16) a. tini āchen: He is (alive).
 b. āllā āche: Allah exists.
 c. *tini ḍhākā āchen: He is (alive) in Dacca.

In (16a, b) āch indicates existence of tini and āllā without locating them in space. (16c) is ungrammatical because absolute existence has been located in space in this sentence. In (15a)

āch locates tini in a location. Here āch is an existential locative verb, which obligatorily takes a LOC that can be overt or covert. Usually, āch in the locational existential sense does not allow the omission of LOC, if it is not contextually understood. So we find that there is no sentence tini āchen: 'He is (exists)' in the locational existential sense. āch and another stative verb thāk: 'Reside' are strictly locative, but the verbs of motion like yā: 'Go' and ās: 'Come' take a LOC optionally. In (15a) the locative expression is stative, in (15b) dhākā indicates the LOC as a source, in (15c) the LOC is a goal, and in (15d) the LOC is a path. Fillmore (1971) has considered the cases like 'Source', 'Goal' and 'Path' in order to capture the facts of the sentences like those in (15). We will consider that all these locational expressions are subcategories of the LOC, and they can be thought of as bundles of the features like [+LOC,+STAT], [+LOC,+SOURCE], [+LOC,+GOAL] and [+LOC,+PATH]. Fillmore usually does not take cases as complexes of features, but he has, of late, proceeded in this direction (cf., Fillmore (1969)). We assume that it is the right way to proceed if one wants to limit the number of cases in a grammar.

ESS(ive) is the case of predicate nominals in sentences like the following.

- (17) a. tini kabi: He is (a) poet.
 b. tini kamyunist: He is (a) communist.

The predicate nominals kabi and kamyunist in (17) are considered as ESS (cf., Stockwell et al (1973, 29); see also Fillmore (1968a, 84)). This type of predicate nominal is considered as LOC in a localistic case grammar, because the nominals above indicate class membership (cf., Lyons (1968, 389), Anderson (1971)).

The predicate nominals in sentences like (17) behave like adjectives in Bengali, that is, instead of indicating class membership they indicate the 'quality/nature' of the subject on which they are predicated. In (17) the predicate nominals have no SPEC, which parallels the use of predicate adjectives as shown in (18).

- (18) a. tini buddhimān: He is clever.
 b. se sundar: He is handsome.

But it can be argued that the predicate nominals in (17) have deleted SPEC's, and thus they are related to (19), respectively.

- (19) a. tini ekjan kabi: He is a poet.
 b. tini ekjan kamyunist: He is a communist.

But the predicate adjectives never take a SPEC. The following examples, where the predicate adjectives are specified, are ungrammatical:

- (20) a. *tini ekjan buddhiman: *He is a clever.
 b. *se ekjan sundar: *He is a handsome.

Now consider the following examples.

- (21) a. tāṛā kamyunist: They are communists.
 b. *tāṛā kamyunistṛā: They are communists.

In (21a) the subject noun phrase is plural, but the predicate nominal is unspecified singular. The sentence (21b) is ungrammatical because the predicate nominal is plural. This indicates that the predicate nominals are adjectival in Bengali. But counter arguments can be given against such an assumption. The unspecified generic noun phrases in Bengali usually appear in the base form as can be seen in (22).

- (22) a. se yubatī pachanda kare : He likes maidens.
 b. se phul bhālabāse: He loves flowers.

In (22) yubatī and phul are unspecified, and are similar to the predicate nominal in (21a); but here we cannot say that they behave like adjectives.

A detailed investigation is necessary in order to discover the syntax of predicate nominals and adjectives in Bengali. Tentatively, we will consider the predicate nominals as ESS, which occurs with stative verbs that have the frame ESS NEUT.

3.3 An Outline of Constituent Order in Bengali.

Bengali is a post-positional 'case language' with a dominant SOV order in the surface, but allows fairly free order of constituents in a simple sentence structure. Greenberg (1963^b) does not include Bengali in his paper on language universals, but many of his generalizations about post-positional languages apply well to Bengali. The dominant order of constituents in a simple sentence structure is SOV, but this is not rigid; other possible orders are also allowed. Ordering of constituents in Bengali is determined by emphasis and secondary topicalization, which have a bias to place the topicalized element at or towards to the sentence-initial position. Subject marking, which comes under primary topicalization, has nothing to do with placing the subject sentence-initially; but in a prosaic speech the subject is the topic, and is placed sentence-initially. Consider the examples in (23).

- (23) a. SOV: hāsān ketakīke bhālabāse: Hasan loves Ketaki.
 b. OSV: ketakīke hāsān bhālabāse:
 c. SVO: hāsān bhālabāse ketakīke:
 d. OVS: ketakīke bhālabāse hāsān:
 e. VOS: bhālabāse ketakīke hāsān:
 f. VSO: bhālabāse hāsān ketakīke:

The sentences in (23) show that structures with the three constituents subject, object and verb form, allow all the permutable orders. But the order SOV in (23a) is the dominant and prosaic order in Bengali. The sentences in (23) are synonymous, but differ from one another in secondary topicalization, which places the topic sentence-initially. (23a) is a colourless statement about a fact, and if none of its constituents is emphasized, it will sound like an objective report. But in other sentences emotion and stylistic preferences have changed the prosaic order of (23a) in different directions.

This relatively free order of constituents in Bengali is possible due to (a) the case markers, and (b) Subject-AUX concord in person and grade features (cf., § 1.6). But in those sentences where the objectivalization rule (cf., § 3.5) deletes the case marker of the objectivalized actant, the order of constituents becomes rigid. In such a structure the subject is always placed sentence-initially, and the order of the constituents is either SOV or SVO. Consider the examples below.

- (24) a. SOV: mānuṣ bāgh māre: Men kill tigers.
 b. SVO: mānuṣ māre bāgh:
 c. *OSV: *bāgh mānuṣ māre:
- (25) a. SOV: bāgh mānuṣ māre: Tigers kill men.
 b. SVO: bāgh māre mānuṣ:
 c. *OSV: *mānuṣ māre bāgh:

In (24a) the object has no case marker (and subjects in modern Bengali usually do not take any marker). It follows the subject and precedes the verb form. In (24b) the object follows the subject and the verb form. (24c) is ungrammatical as the object precedes the subject. This sentence is grammatical if bāgh is

taken as subject, in which case (24c) becomes (25a), which is converse of (24a, b). In (25a, b) the object noun phrases have no case marker, so they always come after the subject. (25c) is ungrammatical like (24c), where the object precedes the subject.

In sentences with both direct and indirect objects, the direct object follows the indirect object; but other orders are also allowed. Consider (26).

(26) a. cheleṭi meyeṭike ekṭi phul diyeche: The boy has given
 1 2 3 4 5 -1- -5-
 a flower to the girl.
 3 4 _____2_____

b. cheleṭi ekṭi phul meyeṭike diyeche:
 1 3 4 2 5

In (26a) the order is subject, indirect object, direct object and verb form, and this is an usual order in Bengali. In (26b) the direct object precedes the indirect object. (26a, b) are synonymous, but they differ in secondary topicalization.

With proposition-internal LOC, the order is S LOC V:

(27) a. tini jele āchen: He is in gaol.

b. ?jele tini āchen:

In (27a) the LOC appears between the subject and the verb form, but (27b), where the LOC is a sentence-initial constituent, is odd.

In sentences with temporal and locative expressions, the order is S TEMP LOC V:

(28) a. āmi bikele bārī yābo: I shall go home in the afternoon.
 1 2 3 4 1 -4- 3 -2-

b. bikele āmi bārī yābo:
 2 1 3 4

c. ?bārī āmi bikele yābo:
 3 1 2 4

In (28a) the order is S TEMP LOC V, in (28b) TEMP S LOC V and in

(28c) LOC S TEMP V. (28a) is a natural sentence, and (28b) is not as natural as (28a), but is fairly common. (28c) is awkward.

The major type of restrictive relative clause is placed sentence-initially (cf., § 8.2.1):

- (29) a. ye-meyeṭi nāclo, se sundar: Lit., Who-the girl danced,
 1 2 3 4 5 1 2 3
 she is beautiful (The girl who danced is beautiful).
 4 5
- b. ye-cheleṭi kabitā likhto, se mare geche: The boy who
 used to write poems has died.

The complement clauses are usually placed sentence-finally unless they are extraposed to the sentence-initial position (cf., § 10.3). Consider the examples in (30).

- (30) a. e-kathā satyi ye rehānā rūpasī: This proposition is
 1 2 3 4 6 1 2
 true that Rehana is beautiful.
 3 4 6
- b. rehānā ye rūpasī, tā satyi: That Rehana is beautiful
 2 3 4 5 2 3
 is true.
 5

In (30a) the complement clause rehānā rūpasī has been extraposed to the sentence-final position, and in (30b) it has been extraposed to the sentence-initial position after placing the complementizer inside the complement clause. In (30b) the abstract noun phrase e-kathā: 'This proposition' has been reduced to the abstract pronoun tā: 'That (thing)' (cf., § 10.3; § 10.4). Bengali does not allow sentential subjects. That is why, there is no sentence *rehānā ye rūpasī satyi: 'That Rehana is beautiful is true'; instead Bengali has (30b).

The SPEC precedes the head noun in an indefinite noun phrase in Bengali (cf., § 2.4.3), and it follows the head noun in a definite noun phrase (cf., § 2.4.5). Consider the examples below.

- (31) a. $\begin{matrix} \text{duti} & \text{meye:} & & \text{Two girls.} \\ 1 & 2 & 1 & 2 \end{matrix}$
- b. $\begin{matrix} \text{meye} & \text{duti:} & & \text{The two girls.} \\ 1 & 2 & \underline{2} & 1 \end{matrix}$

In (31a) the SPEC duti precedes the head noun and in (31b) follows the head noun.

The demonstrative deictics always precede the head noun:

- (32) a. $\begin{matrix} \text{e-meye} & \text{ti:} & & \text{This girl.} \\ 1 & 2 & 1 & 2 \end{matrix}$
- b. * $\begin{matrix} \text{meye} & \text{ti} & \text{e:} \\ & & \end{matrix}$

In (32a) the deictic e precedes the head noun meye, on which the CL ti is suffixed. (32b) is ungrammatical as the deictic e follows the noun.

The attributive adjective usually follows the SPEC and precedes the head noun:

- (33) a. $\begin{matrix} \text{ek} & \text{ti} & \text{sundari} & \text{meye:} & & \text{A beautiful girl.} \\ 1 & 2 & 3 & 1 & 2 & 3 \end{matrix}$
- b. $\begin{matrix} \text{sundari} & \text{ek} & \text{ti} & \text{meye:} \\ 2 & 1 & 3 & \end{matrix}$
- c. * $\begin{matrix} \text{ek} & \text{ti} & \text{meye} & \text{sundari:} \\ 1 & 3 & 2 & \end{matrix}$

In (33a) the adjective immediately precedes the head noun, and in (33b) it is the initial element of the noun phrase. In (33b) the adjective emphasized and so it is placed initially. (33c), where the adjective follows the head noun, is unacceptable.

3.4 Subject Marking and Subject Placement.

Functional notions such as subject and object are not considered to be underlying, but surface structure relations in case grammar (cf., Fillmore (1968a, 17-21)). But we will argue that these relations, although not underlying relations, are not merely surface structure relations. These relations should be

established in the intermediate underlying structure. There are many transformations which apply after subject and object marking, and certainly they cannot apply if subject and object are not considered to be intermediate structure relations. For example, Subject-AUX concord, subject and object raising and reflexivization rule apply after subject and object marking. They show that subject and object are not surface structure relations only. Let us consider the reflexive rule. This rule can be given as a general rule saying that the subject noun phrase always reflexivizes the object noun phrase when they are coreferential and are in the same simplex structure. But this rule becomes complicated if one wants to apply it before subject and object marking. Then we have to say that the AGT and DAT respectively reflexivize DAT and NEUT when they are coreferential and in the same simplex structure.

We have said that the underlying order of elements does not reflect the surface order in this grammar. This makes necessary some operations to place the underlying elements in an acceptable surface order. Here we will be concerned only with subject and object relations. The cases in this grammar have been placed in a particular order, but it is not a direct reflection of the surface. Fillmore (1971) prefers to place the cases in the underlying structure in a manner that sets up a hierarchy of subject selection in active sentences. He places the cases in a manner such that the left most actant has priority over other actants to be the subject in an active sentence. We have taken the opposite order: the cases are placed in the underlying structure in a manner such that the right most actant has priority over other actants, to be the active subject. There is no

theoretical issue related to this. This has been adopted as a matter of convenience. The right most actant in an underlying tree will be marked as subject, and will usually be moved to the front of the sentence. The other cooccurring actants can be left as they are if no other transformation moves them out of their underlying position. In certain situations the actants will be rearranged in order to derive acceptable surface structure. For example, in sentences where NEUT DAT AGT occur, the AGT will be the subject and will be moved to the sentence-initial position, and NEUT DAT will swap position in order to derive a natural surface order.

We have two operations: (a) Subject marking, and (b) Subject placement. The subject marking operation will mark the right most actant in an underlying tree as [+SUBJ], will add this feature to the CS associated with the head noun of the actant, and will delete the case node and case marker of the actant. The subject placement operation will place the subject either at the sentence-initial position (we will usually place the subject sentence-initially as this is the common practice in Bengali), or in any other position which will be decided by secondary topicalization.

We have already suggested that AGT, DAT, NEUT and LOC can be a subject in Bengali, INS can be a subject if the sentence is understood figuratively, and ESS can never be a subject. Here we are concerned only with active sentences. The AGT always becomes the subject with verbs which have the case frame X AGT (X is a variable for one or more cases, and it can be null) . Considering the examples below.

- (34) a. minu mātinke ekṭi bai diyeche: Minu has given a book to Matin.
 b. minu kalam diye likhche: Minu is writing with (a) pen.
 c. minu mātinke mereche: Minu has hit Matin.
 d. minu parche: Minu is reading.

In (34a)-(34d) minu is AGT and is the subject.

The DAT becomes the subject with verbs which have the case frame __X DAT. This can be seen in (35).

- (35) a. minu tāke cene: Minu knows him.
 b. minu tāke bhālabāse: Minu loves him.

The NEUT becomes the subject with verbs which have the case frame __X NEUT. This can be seen in (36).

- (36) a. cābite darajāṭi khule gelo: The door opened by a key.
 b. darajāṭi khule yācche: The door is opening.

In (36a) INS and NEUT occur, and in (36b) only NEUT occurs. In both sentences the NEUT is the subject.

INS can be a subject if the sentence is understood figuratively. Consider the examples in (37).

- (37) a. ?briṣṭi ṣasya naṣṭa kareche: The rain has destroyed the crop.
 b. briṣṭite ṣasya naṣṭa hayeche: The crop was (became) destroyed by rain.

In (37a) the INS briṣṭi is the subject, and the sentence is understood figuratively. Its natural counterpart is (37b), where the NEUT is the subject.

The LOC can be a subject with weather verbs, and with the verbs which have the frame __INS LOC. Consider (38).

- (38) a. skaṭlyāṇḍ ṭhāṇḍā: Scotland is cold.
 b. pukurṭi jale bhare geche: The pond has filled with water.

In (38a) LOC is the only actant and has become the subject.

In (38b) INS and LOC cooccur, but the LOC has become the subject and the INS appears in the surface with the case marker e.

3.4.1 Active Subject Marking and Placement.

The rules of active subject marking and placement (in the dominant constituent order) are given below.

(39) SUBJECT MARKING AND SUBJECT PLACEMENT.

SI: s [M V X C_i [NP CAM]]
 1 2 3 5 6 7

SC: a. Mark 6 as [+SUBJ] and add this feature to the CS associated with the head N of 6. Prune 5 and 7.

b. Attach 6 as the left sister of 2, optionally but preferably.

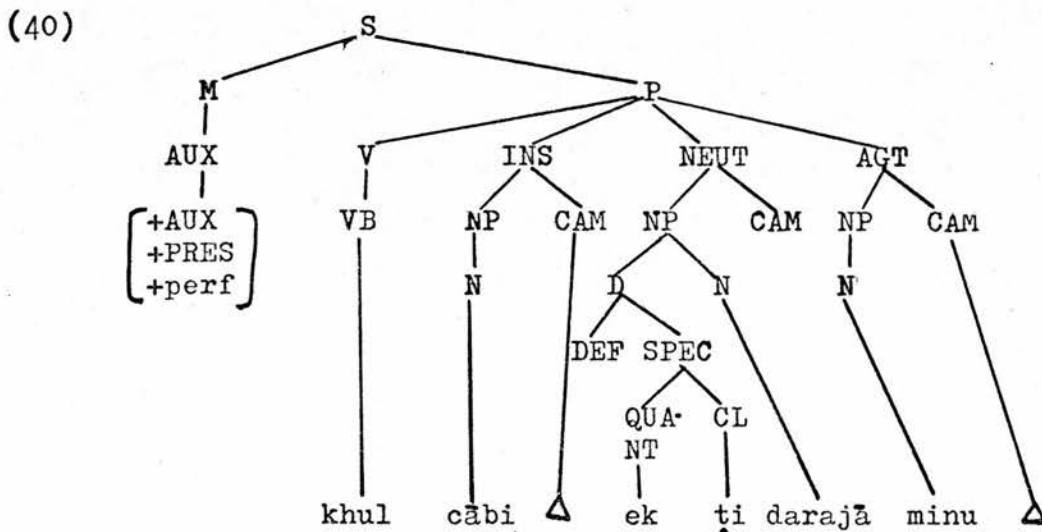
Conditions: a. 3-7 is a constituent.

b. 5 \neq ESS, INS.

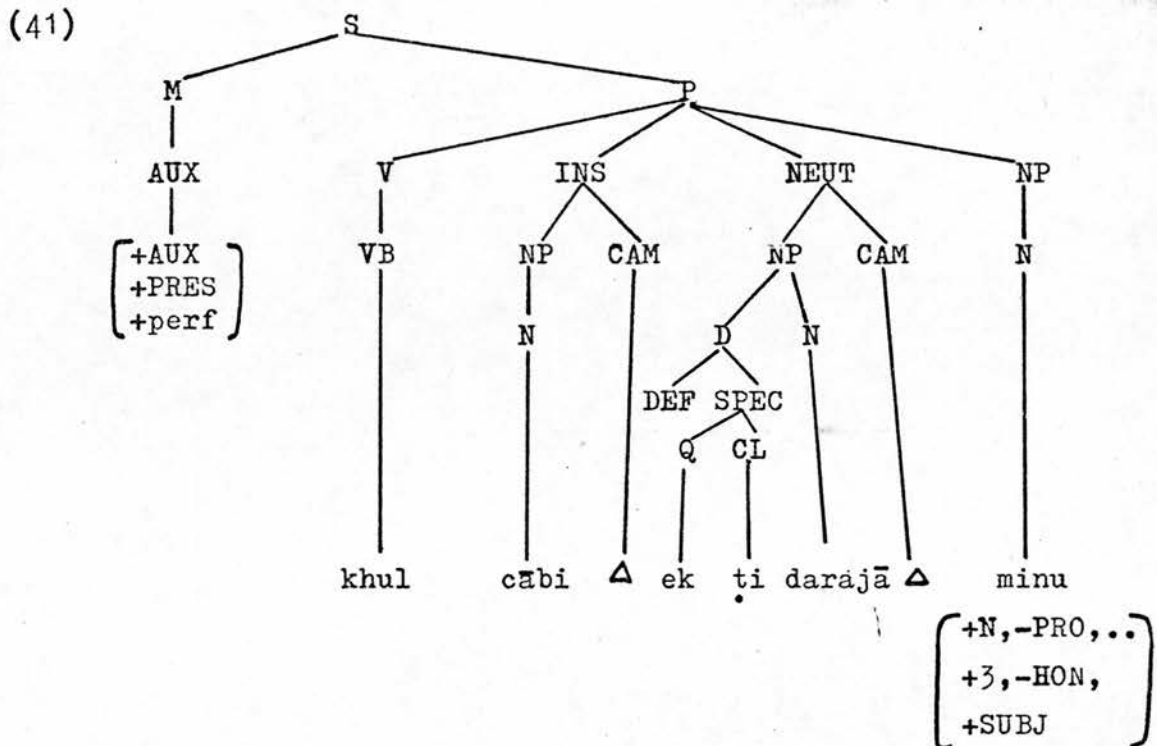
c. 3 and 5 are sisters.

d. 5 is the right most actant in the P.

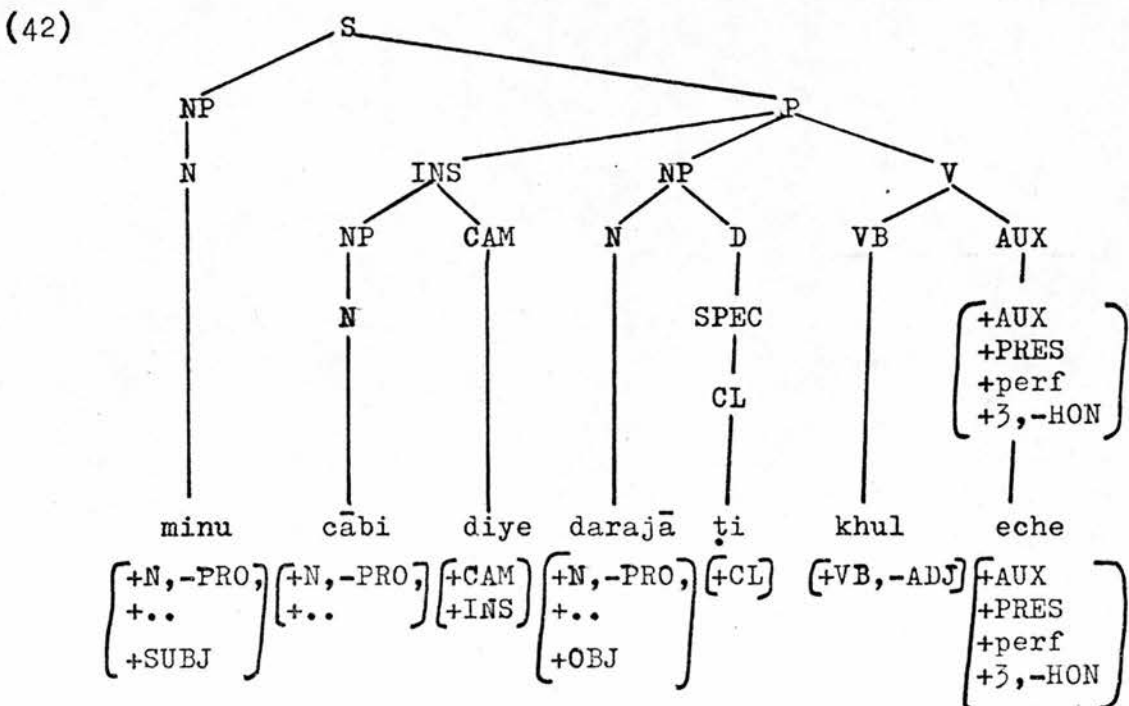
Consider the underlying phrase marker (40).



The structure change (39a) will apply to (40), and (40) will be transformed into (41).



The structure change (39b) will apply to (41). This will move the subject NP to the sentence-initial position. After the application of (39b) and other relevant transformations, (41) will be converted into the final derived structure (42).



As we have arrived at the post-transformational string in (42) the second lexical insertion rule has inserted the INS case marker diye to the case marker node dominated by INS, and the form of the AUX eche to the AUX. (42) will generate minu cābi diye darajāṭi khuleche: 'Minu has opened the door with (a) key'.

3.5 Object Marking and Objectivalization.

We will make a distinction between object marking and objectivalization: (a) object marking is the operation that marks some actant as an object of the verb, and (b) objectivalization is the operation which deletes the case and case marker nodes of the object noun phrase. The objectivalization operation brings the object noun phrase into closer association of the verb. The operation of object marking is quite simple. It usually marks NEUT and DAT noun phrases as [+OBJ]. Consider the examples below.

- (43) a. hāsān ketakīke bhālabāse: Hasan loves Ketaki.
 b. ketakī bai parche: Ketaki is reading (a) book.

In (43a) ketakī is an underlying NEUT and the object. It appears with the NEUT case marker ke. In (43b) bai is an underlying NEUT and the object. It has no case marker. We will assume that although ketakī is the object in (43a), it has not been 'objectivalized', and so it retains the case marker. But bai in (43b) is an object, and has undergone the objectivalization operation, which has deleted its case marker. Some object noun phrases in Bengali do not retain the case marker. Our notion of objectivalization will be able to explain the fact why some object noun phrases retain the case marker while some others do not.

Usually nonhuman and inanimate objects do not retain any case marker in Bengali, but it is not a general rule. We will see that in many instances human objects do not retain any case marker, but nonhuman and inanimate objects retain a case marker.

Consider the examples in (44, 45).

- (44) a. bhadralok ekṭi kiṣorīke pachanda karen: The gentleman likes a teen-ager.
 b. bhadralok e-kiṣorīṭike pachanda karen: The gentleman likes this teen-ager.
 c. bhadralok kiṣorī pachanda karen: The gentleman likes teen-agers.
- (45) a. tini ekṭi beṛālke pachanda karen: He likes a cat.
 b. tini e-beṛālṭike pachanda karen: He likes this cat.
 c. tini beṛāl pachanda karen: He likes cats.

The object noun phrases in (44a, b) are [+HUM] and they retain the case marker ke, but the object in (44c), which is also [+HUM], has no case marker. The case marker ke cannot be deleted from the object noun phrases in (44a, b), and conversely the case marker ke cannot be used with the object noun phrase in (44c). The object noun phrases in (45) are [-HUM]. The object noun phrases in (45a, b) retain the case marker, but it could have been deleted. The object noun phrase in (45c) has no case marker, because no case marker can be used with it. In (44, 45) we see that certain human object noun phrases do not retain any case marker and certain non-human object noun phrases can retain a case marker. In Bengali a human object noun phrase usually retains the case marker unless the noun phrase is unspecified and generic. And the case marker with non-human and inanimate object noun phrases is retained if they are definite.

The case marker ke or re with a non-human, an inanimate or an abstract object noun phrase is used in Bengali to indicate extreme definiteness. We will look into the fact from another angle, and consider why the case marker is not used rather than why it is used. This will help us to generalize all sorts of object noun phrases which do not retain any case marker. We have seen that human, non-human, and inanimate all sorts of objects need not retain the case markers in certain situations. We will assume that this is due to objectivalization, which deletes the case marker of the object noun phrase and brings it into a closer association with the verb.

The following types of DAT and NEUT noun phrases are susceptible of objectivalization:

- (46) a. [-HUM] and [-ANI] noun phrases in general.
 b. Unspecific indefinite noun phrases.
 c. Unquantified generic noun phrases.

Human noun phrases are not susceptible of objectivalization if they are not used indefinitely and unspecifically; and non-human and inanimate NEUT noun phrases are always susceptible of objectivalization unless they are intended particularly definitely. Consider the examples in (47).

- (47) a. hāsān ekjan rūpasīke khūjche: Hasan is looking for a beauty.
 b. hāsān ekjan rūpasī khūjche:

In (47a) the object noun phrase has the case marker ke, but the object in (47b) has no case marker. These two sentences are not semantically equivalent. (47a) means that Hasan is looking for a particular beauty, and (47b) means that Hasan is looking for any beauty. In (47a) ekjan rūpasī is [+HUM, -DEF, +SPECIFIC] and

ekjan rūpasī in (47b) is [+HUM,-DEF,-SPECIFIC]. The object noun phrase in (47a) does not undergo objectivalization, because it is specific; but the object noun phrase in (47b) undergoes objectivalization, because it is unspecific. Non-human and inanimate object noun phrases are generally susceptible of objectivalization unless they are intended particularly definitely. Consider the examples below.

- (48) a. āmi chabiṭi dekhchi: I am looking at the picture.
 b. āmi chabiṭike dekhchi:
 c. āmi chabi(*ke) dekhchi: I am looking on (some) picture(s).

The object noun phrases in (48a, b) are inanimate and definite, and the case marker ke is retained in (48b), but not in (48a). These sentences are synonymous, but objectivalization has applied in (48a), not in (48b). Although both the objects are definite, the object noun phrase in (48b) is understood more definitely than the one in (48a). We have said that non-human and inanimate NEUT noun phrases are always susceptible of objectivalization unless they are intended particularly definitely, what Rabīndranāth Thākur has called 'an special occasion' (cf., Chatterji, (1972, 86)). In (48c) the object noun phrase is bound to be objectivalized as it is [-ANI,-DEF,-SPECIFIC].

Objectivalization is a semantic as well as a syntactic process: semantically it brings the object in a closer association with the verb, and syntactically it causes the case marker of the object to be deleted. For example, bai paṛ, rūpasī khūj and chabi dekh in (43b), (47b) and (48c) respectively can be compared to the so-called conjunct verb roots such as praśna kar: 'Ask (a) question', snān kar: 'Take (Lit., Do) (a) bath' structurally and semantically. The conjunct verb roots are basically closely

collocated object noun phrases and verb roots. That the deletion of case markers of DAT and NEUT object noun phrases is due to objectivalization can be seen clearly where marked objectivalization takes place. The LOC and INS usually retain the case markers in the surface, but with some verbs objectivalization of these actants becomes obligatory. Consider the examples in (49).

- (49) a. minu glās̄tite jal bharlo: Minu filled water in the glass.
 2 3 4 4 3 2
- b. minu jal diye glās̄ti bharlo: Minu filled the glass
 2 3 4 5 5 4
 with water.
 3 2
- c. *minu glās̄tite jal diye bharlo:

The VB bhar has the case frame (INS)(LOC) AGT. There is a rule feature attached to this verb that when both INS and LOC occur in a proposition, either of these two actants will be objectivalized (better if INS is objectivalized), and if only one of the optional cases occurs in a proposition, it will be objectivalized. In (49a) INS has been objectivalized, and the case marker has been deleted; in (49b) the LOC has been objectivalized, and the case marker has been deleted. (49c) is ungrammatical, because none of the INS and LOC has been objectivalized, and they both have appeared in the surface with the case markers. So we see that the deletion of the case markers of the object noun phrases is caused by the objectivalization of the cases.

We have said that we have two operations: object marking, and objectivalization. These rules are given below.

(50) UNMARKED OBJECT MARKING.

SI: _S [X V X (NEUT)_i [NP CAM] (DAT) [NP CAM]]
 1 2 3 5 6 7 8 9 10

SC: Mark 5 and 8 as objects, and add the feature [+OBJ]
 to the head noun of 6 and 9.

Condition: 3-10 is a constituent.

(51) UNMARKED OBJECTIVALIZATION.

SI: _S [X V X C_i [NP CAM] X]
 1 2 3 4 5 6 7 8

SC: a. Erase 5 and 7.

b. Specify 6 as objectivalized, and add the feature
 [+OBJECTIVALIZED] to the head noun of 6.

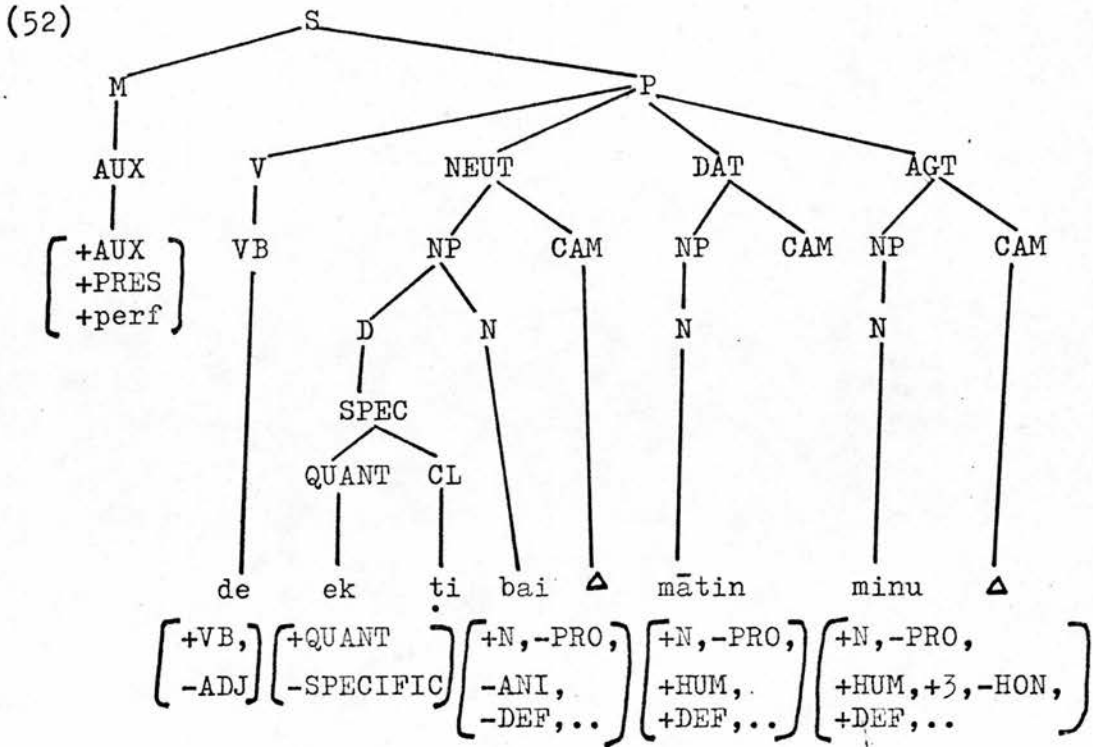
Conditions: a. 5 is either NEUT or DAT.

b. 6 is [-DEF, -SPECIFIC] , or unquantified
 generic.

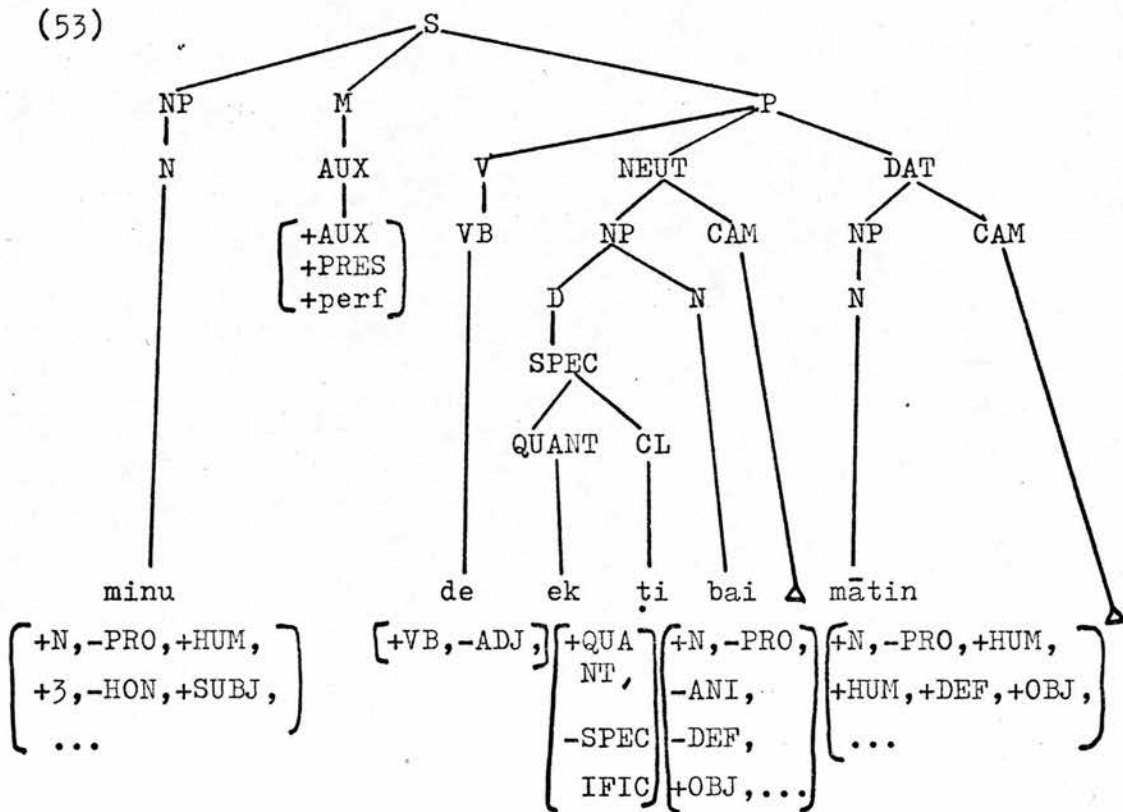
c. If 6 is [-HUM] or [-ANI] and [+DEF] ,
 the rule may apply if the noun phrase
 is not intended particularly definitely.

d. 3-8 is a constituent.

Consider the underlying phrase marker (52) (see page 119).

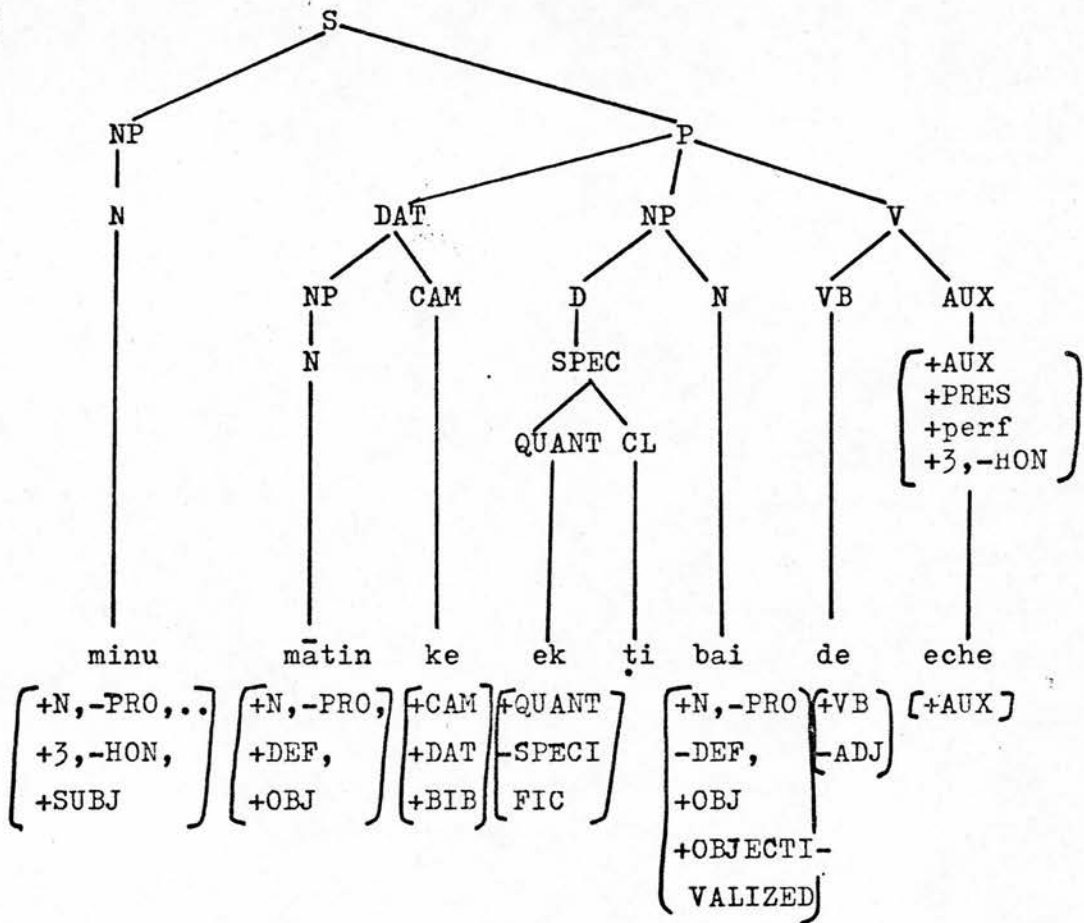


The structure change (39a, b) and (50) will apply to (52); and (52) will be transformed into (53).



The rule (51) is applicable to (53). This rule will objectivize the NEUT, and delete the case node and the case marker node of NEUT. Then DAT and the objectivized noun phrase will swap position with each other. Then the Subject-AUX concord rule will apply, and the AUX will be adjoined to the right of the V. Subsequently, the V will be placed sentence-finally. All these operations will transform (53) into (54).

(54)



The second lexical insertion rule has attached the case marker ke and the form of the AUX eche to the respective nodes in (54). The terminal string of (54) will pass through the morphophonological component, and we will derive the sentence minu mātinke ekṭi bai diyeche: 'Minu has given a book to Matin'.

3.6 CASE MARKERS.

Bengali has two types of lexical item which are used as case markers. The one type consists of the so-called 'semantically empty' bound forms that are suffixed to the noun phrase, and the other type consists of the independent items that follow the noun phrase, although not all of them are 'meaningful'. The first type is traditionally called bibhakti: 'Primary case markers', and the second type is called anusarga: 'Post-positional words' (cf., Chatterji (1939, 257-71)). The difference between these two types of case marker is that the bibhakti's such as ke, re, e, te etc., are considered as semantically empty, and are suffixed to the noun phrase; and the anusarga's such as dvārā, diye, theke, hote etc., are considered as meaningful, and are not suffixed to, although they follow, the noun phrase. Some of the anusarga's are lexically related to lexical categories of other type. For example, theke is related to the verb thāk: 'Stay, reside', diye is related to the verb de: 'Give', but they are semantically distinct. Although the anusarga's are considered as meaningful, in fact they have no lexical meaning. When they are used with noun phrases as case markers (consider dvārā, diye), they give some illusion of lexical meaning, which presumably comes from the case role of the noun phrases with which they occur. We assume that the traditional idea that the anusarga's are meaningful grew from the lexical relatedness of some of the anusarga's to verbs, and from the abstract locatives such as upar: 'Above, On', bhitar: 'Inside, In', pās: 'Side, beside, by', which are considered as case markers in traditional grammars (cf., Chatterji (1939, 259-60)). We will not consider these abstract locational

nouns as case markers. They will be considered as nouns having, usually, the LOC role. Here we will not go any further into the semantic and morphological details of the case markers, and the both types of item will be considered as case markers. The morphophonological component will decide whether a case marker will be suffixed to the noun phrase or not. We can help this component with the morphological features [+BIB] (bibhakti). That is, a case marker having the feature [+BIB] will be suffixed to the preceding item, and one having the feature [-BIB] will appear as a free lexical item in the surface.

In Bengali no case has an unique case marker. Some case markers are used for a number of cases, and each case has a number of case markers. In our grammar the base rule 5 generates a case marker node (CAM) for each case, but no case marker is inserted in the underlying structure. The case markers are listed in the second lexicon with a complex symbol associated with each, and they are inserted to the derived trees by the second lexical insertion rule. As each case in Bengali has a number of case markers, and some case markers are used for a number of cases, no general rule for case marker placement can be given. We cannot here go into the details of all the case markers used in Bengali. In the following sections we will consider the case markers that are frequently used for the cases that have been included in this grammar.

3.6.1 CAM for AGT.

It is difficult to decide on the case marker for AGT. The case marker for AGT is not realized in the surface structure of the active sentences because the AGT becomes the subject.

In old Bengali the marker e used to be used with the subject noun phrase, but it is seldom used in SCB. Moreover, it can be used with AGT as well as non-AGT subjects. Passive sentences are of no help in deciding the case marker for AGT, because Bengali disallows passivization of sentences with an overt AGT. In traditional grammars examples like the following are given in order to show that Bengali has passive constructions.

- (55) a. $\overset{2}{r\bar{a}m}$ $\overset{3}{kartrik}$ $\overset{4}{pustak}$ $\overset{5}{pa\ddot{t}hita}$ $\overset{3}{hailo}$: (A) book was read
by Ram.
- b. $\overset{2}{rahim}$ $\overset{3}{kartrik}$ $\overset{4}{candra}$ $\overset{5}{dri\ddot{s}ta}$ $\overset{3}{haiteche}$: The moon is
being seen by Rahim.

The sentences in (55) are instances of Sanskritized sādhu (Chaste) Bengali. Although the above style is obsolete now, the marker kartrik is still used in Bengali. The marker kartrik does suggest some sort of agentivity, but it is not strictly agentive. For example, rām in (55a) is AGT, but rahim in (55b) cannot be considered as an AGT. (Bengali, of course, suggests that many mental and stative processes are also some sort of 'action', which can be seen in conjunct verbs like kāmanā kar: Lit., 'Do desire', biśrām kar: Lit., 'Do rest'.) So we will tentatively assume that Bengali has no case marker for AGT.

3.6.2 CAM for DAT and NEUT.

The case marker for DAT and NEUT is ke (and re, which is now obsolete or rarely used). Consider the examples in (56).

- (56) a. se cheleṭike mereche: He has hit the boy.
b. se cheleṭike cene: He knows the boy.

cheleṭi in (56a) is an underlying DAT, and in (56) a NEUT.

Each of these noun phrases takes the case marker ke. We have already dealt with the situations in which the case markers of

DAT and NEUT are erased due to objectivalization (cf., §3.5).

We have shown that the NEUT noun phrases that are anything other than [+HUM] undergo objectivalization in general; unspecific and indefinite, and unquantified generic [+HUM] NEUT noun phrases undergo objectivalization, and this causes the case marker to be erased. Consider the examples in (57).

- (57) a. se glāṣṭi bheṇeche: He has broken the glass.
 b. se glāṣṭike bheṇeche:

The NEUT noun phrase glāṣṭi in (57a) has no case marker, but the one in (57b) has the case marker ke. These sentences are synonymous, and both the NEUT noun phrases are definite. But the NEUT noun phrase in (57b) is understood more definitely than the one in (57a). We will consider that the NEUT noun phrase in (57a) has undergone objectivalization, and consequently the case marker has been erased; but the one in (57b) has not been objectivalized as it is intended particularly definitely. Objectivalization of non-human and inanimate NEUT noun phrases is fairly common in Bengali, and objectivalization of DAT noun phrases and human NEUT noun phrases is not common, but not rare. Consider the examples in (58, 59).

- (58) a. bāghṭi dujan lokke mereche: The tiger has killed two men.
 b. bāghṭi dujan lok mereche:
 (59) a. āmi śisuderke pachanda kari: I like children.
 b. āmi śisuder pachanda kari:

The DAT noun phrase dujan lok in (58a) is [+HUM, -DEF, +SPECIFIC] and so it has not been objectivalized, and it retains the case marker ke. But the DAT noun phrase in (58b) is unspecific, and so it has been objectivalized, which has caused its case marker to be erased. Similarly, the NEUT noun phrase in (59a) has not

- c. kuṭhār { dvārā } gāchṭi kāṭā halo: The tree was
 1 { diye } 3 4 5 3 5
 { ?e }
 2
 chopped with an axe.
 4 2 1

In (61a) we see that the INS takes dvārā, diye and te as a case marker. The sentences in (61a) differ semantically. When dvārā or diye is used, (61a) means that Matin has cut his finger intentionally, and acted as an AGT; but when te is used (61a) means that Matin has cut his finger unintentionally, and so he is a DAT. In this sense, (61a) is synonymous with (61b), where no AGT is present or presupposed. There is no AGT in (61b), and so dvārā or diye cannot be used. (61c) is a passive sentence, where there is no AGT, but it is presupposed. For this reason the INS case marker should be either dvārā or diye in (61c). This sentence sounds odd when e is used as the INS case marker.

The case markers dvārā and diye maintain selectional restrictions with the INS noun phrase, and with the cooccurring verb. When the INS noun is concrete and rigid, and the verb suggests body-contact between the INS and DAT, and INS and NEUT either of dvārā or diye can be used as a case marker; otherwise dvārā cannot be used. Moreover, dvārā cannot be used if the INS noun has the feature [+ABS]. Consider (62).

- (62) a. mātin banduk { diye } pākhī mārlo: Matin killed
 2 { *dvārā } 4 5 5
 { }
 3
 birds with a gun.
 4 3 2
- b. mātin tāp { diye } mom galāy: Matin melts
 2 { *dvārā } 4 5 5
 { }
 3
 wax with heat.
 4 3 2

In (62a) diye can be used as the INS case marker, but dvārā cannot be used if the sentence is meant to be understood in the usual sense of using a gun to kill something. However, dvārā can be used if the sentence is understood to mean that Martin used the gun like a stick to kill birds. (62b) shows that dvārā is incompatible with an abstract instrumental like tāp.

We have said that dvārā requires the INS noun to be concrete and rigid. This can be seen in (63).

- (63) a. darī $\left\{ \begin{array}{l} \text{diye} \\ *dvārā \end{array} \right\}$ bādhō: Tie it with a rope.
 b. suto $\left\{ \begin{array}{l} \text{diye} \\ *dvārā \end{array} \right\}$ selāi karo: Sew with cotton.

(63) shows that dvārā is incompatible with non-rigid substances like darī and suto. It seems that dvārā is being used less and less in SCB.

The lexical entries for the case markers are like the following:

- (64) dvārā diye e te
- | | | | |
|-----------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------|
| $\left[\begin{array}{l} +\text{CAM}, +\text{INS} \\ +\text{__AGT} \\ +\text{INS RIGID} \\ -\text{BIB} \end{array} \right]$ | $\left[\begin{array}{l} +\text{CAM}, +\text{INS} \\ +\text{__AGT} \\ -\text{BIB} \end{array} \right]$ | $\left[\begin{array}{l} +\text{CAM}, +\text{INS} \\ -/i/_ \\ +\text{BIB} \end{array} \right]$ | $\left[\begin{array}{l} +\text{CAM}, +\text{INS} \\ +\text{V}_ \\ +\text{BIB} \end{array} \right]$ |
|-----------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------|

3.6.4 CAM for LOC and ESS.

The case markers that are frequently used for LOC are lexically identical with the INS case markers. They are e, te and diye, but not dvārā. The case markers e and te are phonologically conditioned exactly as they are conditioned when used as INS case markers. We will consider five case markers e, te, diye, theke and hote, which are frequently used for LOC.

diye or e can be used as the case marker in (66b). But the sentences in (66b) are not synonymous. When the verb suggests movement without any direction, the case marker is e, and path is understood as [+LOC,+STAT]; and when the verb suggests movement in a certain direction, path is understood as [+LOC,+PATH], and the case marker is diye.

We have seen that e, te and diye are used as case markers for INS as well as for LOC. So it is expected that some surface sentences will be ambiguous. This ambiguity can be seen in (67).

(67) a. āmarā gārīte elām: We came $\left\{ \begin{array}{l} \text{by a} \\ \text{to the} \end{array} \right\}$ car.

b. phulṭi briṣṭite naṣṭa hacche: The flower is being spoiled $\left\{ \begin{array}{l} \text{by} \\ \text{in} \end{array} \right\}$ the rain.

In (67) gārīte and briṣṭite can be taken either as INS or LOC. Traditional Bengali grammarians take cases as semantic notions, but in most cases they determine cases by the case markers. So they occasionally consider LOC as INS, although their examples suggest that the noun phrases which they consider as INS cannot possibly be used as INS. For example, Chatterji (1939, 287) considers that the following sentences each has an INS noun phrase.

(68) a. sojā pathe calo nā keno?: Why do not you walk along
 1 2 3 4 5 6 6 5 4 3
 a straight way?

b. kālikātā diyā āsibo: (I) shall come through Calcutta.

Chatterji considers that pathe in (68a) and kalikātā in (68b) are instrumentals. Presumably, Chatterji thinks that people use a road as an instrument, and so he considers path as INS in (68a). We assume that he takes kalikātā as an instrumental because of the case marker diyā (\Rightarrow diye in SCB).

The case markers for [+LOC,+SOURCE] are theke and hote.

Consider the examples in (69).

- (69) a. tini dhākā theke esechen: He has come from Dacca.
 b. tini bārī hote esechen: He has come from home.

In (69) either of theke or hote can be used as a case marker.

Some verbs allow the objectivalization of LOC, and thus the case marker is erased. In (70a) the LOC retains the case marker, but in (70b) it has been erased due to objectivalization:

- (70) a. āmi dhākāy thāki: I live in Dacca.
 b. āmi dhākā thāki:

Some abstract nouns like bhitar: 'Inside, In', bāir: 'Outside', kāch, nikat, pās: 'Side, Beside', upar: 'Over, Above, On', nīc: 'Below, Under' etc., are traditionally considered as locative case markers. These nouns occur in complex locative expressions like those given below.

- (71) a. baiṭi ṭebiler $\left\{ \begin{array}{l} \text{upar} \\ \text{nīc} \\ \text{pās} \\ \text{kāch} \end{array} \right\}$ e āche: The book is

$\left\{ \begin{array}{l} \text{on} \\ \text{under} \\ \text{beside} \\ \text{near} \end{array} \right\}$ the table (Lit., The book exists
 table's $\left\{ \begin{array}{l} \text{on} \\ \text{under} \\ \text{side} \\ \text{near} \end{array} \right\}$).

- b. baiṭi ṭebile āche: The book is on the table.

The LOC tebil in (71b) is a simple noun phrase, which indicates the location of the book. But in (71a) the abstract relational nouns upar, nīc, pās and kāch indicate the location of the book in relation to the table. They are abstract nouns and so they cannot point to the location specifically without being in relation to some concrete location or object. So it is semantically

satisfactory to consider these abstract noun phrases as LOC, and the concrete noun phrases as relational noun phrases. In (71a) the relational noun tebil has the genitive marker er, and the locative case marker is attached to the abstract locational noun phrase. The concrete relational noun phrase in a complex locational expression is an indication mark which points to the exact location. We will consider the abstract nouns in the complex locational expressions in (71a) as the head of the LOC. This will help us to reduce the number of the locative case markers.

That tebil is not the head of the LOC in (71a) can be better seen in the following examples, where the relational noun phrase cannot be understood as LOC.

- (72) a. baiṭi kalamṭir pāse āche: The book is beside the pen.
 b. ciruniṭi baiṭir kāche rākho: Keep the comb near the book.

In (72) kalamṭi and baiṭi cannot be considered as LOC; they are underlying nEUT. The locations of the abstract locatives pās and kāch in (72a, b) respectively are decided in relation to the genitive noun phrases kalamṭir and baiṭir. We cannot consider the genitive noun phrases in (72) as LOC, and so the genitive noun phrases in (71a) cannot be taken as LOC. We will consider expressions like tebiler unar: Lit., 'Teble's over', bārīr pāse: Lit., 'House's side' as complex locatives, where the abstract nouns are the head of the LOC, and the relational nouns as their modifiers.

The lexical entries for the locative case markers are like the following:

(73)	<u>e</u>	<u>te</u>	<u>diye</u>	<u>theke</u>	<u>hote</u>
	$\left[\begin{array}{l} +\text{CAM}, +\text{LOC}, \\ \{ +\text{STAT} \\ +\text{GOAL} \} \\ -/i/_ \\ +\text{BIB} \end{array} \right]$	$\left[\begin{array}{l} +\text{CAM}, +\text{LOC}, \\ \{ +\text{STAT} \\ +\text{GOAL} \} \\ +\text{V}_ \\ +\text{BIB} \end{array} \right]$	$\left[\begin{array}{l} +\text{CAM}, +\text{LOC}, \\ +\text{PATH} \\ -\text{BIB} \end{array} \right]$	$\left[\begin{array}{l} +\text{CAM}, +\text{LOC}, \\ +\text{SOURCE} \\ -\text{BIB} \end{array} \right]$	$\left[\begin{array}{l} +\text{CAM}, +\text{LOC}, \\ +\text{SOURCE} \\ -\text{BIB} \end{array} \right]$

Bengali has no case marker for ESS. The second lexical insertion rule will not insert any item to the node dominated by ESS, and so the case node, by a general rule, will be erased. This technical problem can be solved in other ways as well. For example, we can think of an empty case marker like ' \emptyset ', which will be attached to the case node dominated by ESS. Later the morphophonological rules will delete it.

PRONOUNS AND PRONOMINALIZATION4.0 Introduction.

This chapter will be devoted to a discussion of Bengali pronouns and pronominalization. In the first few sections we will discuss the semantics and lexical forms of different types of pronoun, and in later sections we will enter into theoretical discussion of their derivation. We have mentioned earlier that this work makes use of a lexicon divided into two parts: (a) The first lexicon, and (b) The second lexicon. The first lexicon contains those pronouns which are used deictically; the second lexicon contains those used anaphorically. Both lexicons list pronouns (and all other lexical items) in their basic form, such as āmi: 'I', tumi: 'You', tini: 'He' etc., but plural and other derived forms such as āmarā: 'We', tomarā: 'You', āmāke: 'Me', tomār: 'Your' etc., are absent from both the lexicons (cf., § 11.1.5; § 11.2.1). They are morphophonologically derived (cf., § 11.3).

Let us start with a discussion of what we mean by a pronoun, and pronominalization. By pronouns we mean the limited number of lexical items to be found in a natural language, that are traditionally said to 'stand for' or 'replace' some previously mentioned noun phrase. These lexical items function deictically, and some of them function anaphorically. By pronominalization we mean the derivation of pronouns in syntactic structures. This can be (a) an underlying derivation (that is, generating pronouns in the underlying structure, and (b) a transformational derivation (that is, transformational reduction of a noun phrase into a pronoun).

In this work we are mainly concerned with the transformational derivation of anaphoric pronouns under certain conditions.

Pronouns are limited in number in any language. Not everything that can 'stand for' some other noun phrase can be taken as a pronoun. Consider the examples below.

- (1) a. mātin elo o se baslo: Matin came and he sat.
 2 3 4 5 2 3 4 5
- b. mātin elo o duṣṭuṭi baslo: Matin came and the naughty
 2 3 4 5 2 3 4
 boy sat.

⁵
se and duṣṭuṭi in (1a, b) respectively refer to mātin in one reading, but here se, but not duṣṭuṭi is a pronoun. Expressions like duṣṭuṭi in (1b), which can be used to refer to some antecedent, are epithets, not pronouns (cf., § 5.5). Pronouns comprise that closed set of lexical items used only for referential purposes (cf., § 4.1.5; § 5.5).

4.1 Types of Bengali Pronoun.

Bengali pronouns can be classified into the following groups: (a) Personal pronouns, (b) Abstract pronoun, (c) Interrogative pronouns, (d) Locative pronouns, (e) Temporal pronouns, (f) Reflexive pronouns, and (g) Relative pronouns. We will discuss the pronouns (a)-(e) in this chapter. Other pronouns are discussed in relevant chapters (cf., § 6.2; § 7.3; § 8.1; § 10.1).

4.1.1 Personal Pronouns.

These pronouns are used to refer to human beings. We give a list of these pronouns below (see page 135).

(2) <u>Singular Form</u>	<u>Plural Form</u>
āmi: 'I'	āmarā ¹ : 'We'
āpani: 'You (honorific)'	āpanārā: 'You'
tumi: 'You (nonhonorific)'	tomarā: 'You'
tui: 'You (pejorative, affectionate)'	torā: 'You'
tini: 'He (honorific)'	tārā: 'They'
se: 'He (nonhonorific)'	tārā: 'They'

These are the regular personal pronouns in SCB.

Bengali has a tendency to derive (lexically) third person personal pronouns from the demonstrative deictics, and this has given rise to a number of third person pronouns, which are used in spoken Bengali. They are listed below.

(3) <u>Deictic</u>	<u>Derived Singular Form</u> <u>as a Pronoun</u>	<u>Derived Plural Form</u> <u>as a Pronoun</u>
e(i): 'This'	$\left\{ \begin{array}{l} e: 'He (-HON, -FAR)' \\ \tilde{e}: 'He (+HON, -FAR)' \\ ini: 'He (+HON, -FAR)' \end{array} \right\}$	$\left\{ \begin{array}{l} erā: 'They' \\ \tilde{erā}: 'They' \end{array} \right\}$
o(i): 'That, in sight'	$\left\{ \begin{array}{l} o: 'He (-HON, +FAR, \\ \quad \quad \quad \text{IN SIGHT})' \\ \tilde{o}: 'He (+HON, +FAR, \\ \quad \quad \quad \text{IN SIGHT})' \\ uni: 'He (+HON, +FAR, \\ \quad \quad \quad \text{IN SIGHT})' \end{array} \right\}$	$\left\{ \begin{array}{l} orā: 'They' \\ \tilde{orā}: 'They' \end{array} \right\}$

1 The forms āmarā, āpani, āpanārā and tomarā are phonologically basic forms. In colloquial speech they are often realized as āmrā, āpni, āpnārā and tomrā respectively, due to the unstressed vowel a-deletion.

In this way it can be shown that the pronoun se: 'He' is also a derivative of the deictic se(i): 'That, out of sight'.

Of the pronouns listed in (3) ini and uni are frequently used in Bangladesh; e and o are used almost pejoratively, and ẽ and õ are rarely used. We have discussed the relation between the deictics and the pronouns elsewhere (cf., § 4.1.5).

4.1.2 Abstract Pronoun.

There is one abstract pronoun in Bengali. It is tā, which has no plural form (cf., § 5.10. § 10.1).

4.1.3 Interrogative Pronouns.

Bengali has the following interrogative pronouns.

- | | | |
|-----|----------------|------------------|
| (4) | ke: 'Who' | ki: 'What' |
| | kakhan: 'When' | kab: 'When, day' |
| | kothā: 'Where' | keno: 'Why' |

There are some other interrogative items in Bengali, such as:

keman: 'How, What type, What quality'

kato: 'How much'

Of these keman should be considered as an interrogative adjective, and kato as an interrogative quantifier.

Bengali has a few indefinite-unspecific pronouns, which are lexically related to the interrogative pronouns. But they are not interrogative. They seem to arise due to suppletion of the interrogative pronouns. They are listed below with the interrogative pronouns that they are related to at their right.

- | | | | |
|-----|--------------------|----|----------------|
| (5) | keu: 'Anyone' | ←- | ke: 'Who' |
| | kichu: 'Anything' | ←- | ki: 'What' |
| | kakhano: 'Ever' | ←- | kakhan: 'When' |
| | kothāo: 'Anywhere' | ←- | kothā: 'Where' |

The deictic ei is used to refer to someone or something near to the speaker, oi is used to refer someone or something far from but in view of the speaker, and sei is used to refer someone or something out of sight of the speaker. In this sort of reference a deictic occurs with a noun, which may be deleted contextually. We will make a distinction between a deictic and a pronoun for the following reasons:

- (8) a. Deictics and pronouns belong to categories of different types. The personal pronouns have the features like [+HUM] , [+HON] , [+PEJ] , [+PL] etc., which they share with nouns. Deictics have no such features.
- b. Deictics are constituents of the D(eterminer) in a noun phrase, but pronouns are head nouns (and noun phrases).
- c. Deictics are used with a head noun in a noun phrase, but pronouns cannot be used with any head noun.
- d. Bengali disallows (socially) the use of a deictic with a [+HON] noun, but there is no such restriction on a pronoun.

Consider the following examples.

- (9) a. ?sei mahilā: That woman.
 b. tini: { He, She }
 c. *tini mahilā: *She woman.
- (10) a. kon cheleṭi?: Which boy?
 b. ke?: Who?
 c. *ke cheleṭi?: *Who boy?

The deictic sei occurs with a [+HON] noun in (9a). This noun phrase would be considered objectionable on social grounds.

In this noun phrase the deictic is [+FAR]. That is, the person

referred to by this noun phrase is absent from the situation of utterance, and so it may pass without objection (subject to the condition that the hearer is careless about the status of the person referred to); but the noun phrase will certainly be objectionable if we use a [-FAR] deictic in place of sei in (9a). In (9b) we have the pronoun tini, which can be used to refer to a definite person excluding the speaker and the hearer without any objection. (9c) shows that a pronoun cannot occur with a head noun. But we know that deictics are always used with a head noun. In (10a) the interrogative deictic kon occurs with a [-HON] head noun. The pronoun ke in (10b) can be used for the same purpose served by the noun phrase (10a), although not so specifically. (10c) shows that a pronoun and a head noun cannot cooccur. Pronouns are lexical items which carry nominal and deictic features inherently, but deictics carry only the deictic features. Bengali makes a distinction between a deictic and a pronoun in respect of inherent features, in that the deictics are [+DEIC, +DEM, +FAR, +INT, +DEF]; but the personal pronouns are noun-like in respect of the inherent features [+ANI, +HUM, +HON, +PL] etc.,.

We have seen that Bengali has a tendency to derive third person personal pronouns from the non-interrogative demonstrative deictics (cf., @ 4.1.1); but in doing so they are recategorized as pronouns. So they cannot be considered as deictics any more. How are they recategorized? Sometimes they assume a derived form as a pronoun, and sometimes, if the form is not distinct from the source deictic, it gets a special intonation when uttered. For example, consider the item o, which is lexically related to o(i): 'That'. When o is used as a nonhonorific pronoun, it does

not contain the i of the deictic, and gets a special intonation accompanied by lengthening of the vowel. When it is used as an honorific pronoun it is nasalized and the pronoun is õ. Thus we see that Bengali has a tendency to derive pronouns from the deictics, and at the same time recategorizes them as pronouns. This has given rise to a number of third person personal pronouns parallel to the deictics. This parallelism is shown in (11).

(11) e(i): 'This' \Rightarrow $\left\{ \begin{array}{l} \underline{e}: [+N, +PRO, -INT, +DEF, +COUNT, +HUM, +\text{3}, \\ \quad -HON, -FAR] \\ \underline{\tilde{e}}: [+N, +PRO, -INT, +DEF, +COUNT, +HUM, +\text{3}, \\ \quad +HON, -FAR] \\ \underline{ini}: [+N, +PRO, -INT, +DEF, +COUNT, +HUM, \\ \quad +\text{3}, +HON, -FAR] \end{array} \right.$

o(i): 'That,
in sight' \Rightarrow $\left\{ \begin{array}{l} \underline{o}: [+N, +PRO, -INT, +DEF, +COUNT, +HUM, +\text{3}, \\ \quad -HON, +FAR, IN SIGHT] \\ \underline{\tilde{o}}: [+N, +PRO, -INT, +DEF, +COUNT, +HUM, +\text{3}, \\ \quad +HON, +FAR, IN SIGHT] \\ \underline{uni}: [+N, +PRO, -INT, +DEF, +COUNT, +HUM, \\ \quad +\text{3}, +HON, +FAR, IN SIGHT] \end{array} \right.$

se(i): 'That,
out of sight' \Rightarrow $se: [+N, +PRO, -INT, +DEF, +COUNT, +HUM, \\ +\text{3}, -HON, +FAR]$

We see in (11) that the pronouns e and ẽ, and o and õ, and se are lexically related to the deictics e(i), o(i) and se(i) respectively. But the lexical relation ini and uni with the deictics is not so obvious. In sādhū Bengali there are three abstract pronouns: ihā: 'This (thing)', uhā: 'That (thing)' and tāhā: 'That (thing)'; but in SCB only tāhā survives as tā. Among the lexically derived pronouns, e, ẽ, and õ have fairly limited use, and o is used in Bangladesh with some pejorative hint.

The pronouns se and tini used in SCB are basically [+FAR, OUT OF SIGHT] pronouns, but now they are neutral to this feature. We cannot relate tini, and āmi, āpani, tui and tumi lexically to any deictic. And the pronouns derived from the deictics behave clearly distinctly from the deictics in respect of matters discussed above. So they should be considered as categories of two distinct type.

Bengali has so many third person personal pronouns, but there is no [-HUM] or [-ANI, -ABS] pronoun. That is, Bengali has no pronoun equivalent to the English 'It'. Bengali solves this problem with the help of the deictics. Consider the following examples.

- (12) a. āmi e-ṭi cāi: I want this.
 1 2 3 1 3 2
 b. āmi e-gulo cāi: I want these.
- (13) a. āmi o-ṭi cāi: I want that.
 b. āmi o-gulo cāi: I want those.
- (14) a. āmi se-ṭi cāi: I want that.
 b. āmi se-gulo cāi: I want those.

In the above examples the residual noun phrases e-ṭi, e-gulo, o-ṭi, o-gulo, se-ṭi and se-gulo superficially show that the deictics can take a CL, and they can be pluralized like nouns and pronouns (cf., § 2.4.6). But we will claim that the deictics take neither CL's, nor plural markers; nor are they pronouns. In (12)-(14) the residual noun phrases have been used deictically to refer to things which are contextually understood, and so the head nouns have been deleted from these noun phrases. Suppose that the deleted noun is bai: 'Book'. If we insert the noun to the residual noun phrases in (12, 13, 14), we will get the examples in (12', 13', 14') respectively.

- (12') a. āmi e-baiṭi cāi: I want this book.
 b. āmi e-baigulo cāi: I want these books.
- (13') a. āmi o-baiṭi cāi: I want that book.
 b. āmi o-baigulo cāi: I want those books.
- (14') a. āmi se-baiṭi cāi: I want that book.
 b. āmi se-baigulo cāi: I want those books.

But pronouns do not allow such insertion as they do not have any understood or deleted noun. Pronouns are inherently deictic nouns.

Now consider the interrogative deictics and the interrogative pronouns. Consider the examples in (15).

- (15) a. kon bhadralok?: Which gentleman?
 b. ke?: Who?
 c. *ke bhadralok?: *Who gentleman?

As we have seen above, we see here that an interrogative deictic occurs with a head noun, but an interrogative pronoun cannot occur with a head noun. We will make a distinction between an interrogative deictic and an interrogative pronoun.

But there are some lexical items which can be used both as a quantifier or a pronoun. kichu: 'Any, Some (thing)' is such an item. Consider (16).

- (16) a. āmi kichu bai cāi: I want some books.
 b. tumi kichu cāo ki?: Do you want anything?

In (16a) kichu appears as an indefinite, non-interrogative quantifier in the noun phrase kichu bai, but in (16b) kichu is an indefinite pronoun.

Similarly, consider the quantifier kato: 'How much' in (17).

- (17) a. kato cāo?: How much do you want?
 b. kato dām cāo?: How much price do you want?

In (17a) there is no head noun with the quantifier kato, and it looks like a pronoun. But it is an interrogative quantifier which allows deletion of the head noun, if it is contextually understood. However in (17b) the quantifier occurs with a head noun dām, and shows that it is a quantifier.

We will make a distinction among deictics, quantifiers and pronouns according to their inherent properties, and this distinction will be shown in the underlying structure. But this distinction may not be so clear in the surface structure, as surface structures are often ambiguous.

4.2.0 Forms and Feature Specifications of Pronouns.

4.2.1 Personal Pronouns.

The pronouns listed in (2) are traditionally called personal or human pronouns as they 'stand for' and refer to human beings. In our analysis we have taken Postal's (1970a) line that pronouns and nouns are the same type of category in the underlying structure, and differ only in the feature [+PRO]. That is, the pronouns have the features [+N,+PRO], and the nouns have the features [+N,-PRO]. Bengali personal pronouns can be classified into three groups in respect of person. Lyons (1968, 276) defines person as:

The category of person is clearly definable with reference to the notion of participant-roles: the 'first' person is used by the speaker to refer to himself as a subject of discourse; the 'second' person is used to refer to the hearer; the 'third' person is used to persons or things other than the speaker and the hearer.

According to this definition āmi is first person; āpani, tumi

and tui are second person; and se, tini, ini and uni are third person. Besides these person-divisions, the Bengali personal pronouns have class or grade distinctions. The first person pronoun has no such distinction, but the second person pronoun has three grades: (a) honorific, (b) nonhonorific, and (c) pejorative/affectionate. The third person pronoun has two grades: (a) honorific, and (b) nonhonorific. All these pronouns have singular and plural forms, but they do not have any distinction for gender.

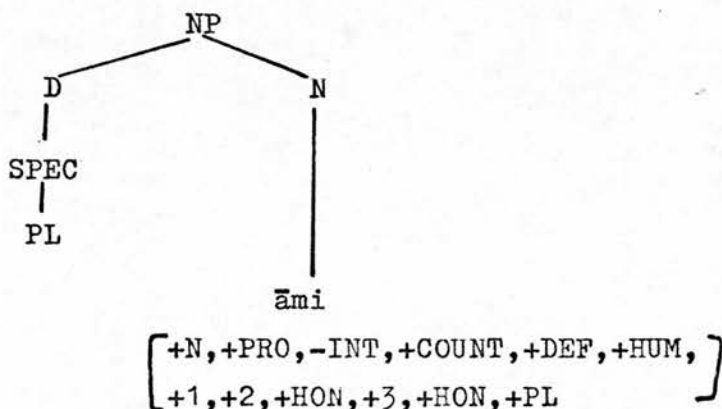
The first and second person plural pronouns may be either 'inclusive' or 'exclusive' (cf., Lyons (1968, 277)). Although morphologically āmarā: 'We', tomarā: 'You ([-HON,+PI])' and āpanārā: 'You ([+HON,+PI])' stand in the same relation to āmi, tumi, and āpani respectively as bālikārā: 'Girls' and cāṣīrā: 'Farmers' stand to bālikā and cāṣī respectively, semantically the relations are not the same. That is, āmarā cannot be simply considered to be the plural of āmi as bālikārā can be considered as the plural of bālikā. In Bengali the 'inclusiveness' or 'exclusiveness' of the first and second person pronoun is not morphologically marked on them, but it is possible to make a semantic distinction in 'inclusive'/'exclusive' use of first and second person pronouns. (cf., § 5.6.2). In the lexicon we will list pronouns only in their basic forms with a complex symbol of features.

In order to subcategorize personal pronouns in Bengali we will have to consider the following inherent and subcategorial features:

- (18) [N(oun)], [PRO(noun)], [INT(errogative)], [HUM(an)] ,
 [DEF(inite)], [HON(orific)], [PEJ(orative)], [COUNT] ,
 [PL(ural)], [1 (first person)], [2 (second person)] ,
 [3 (third person)] .

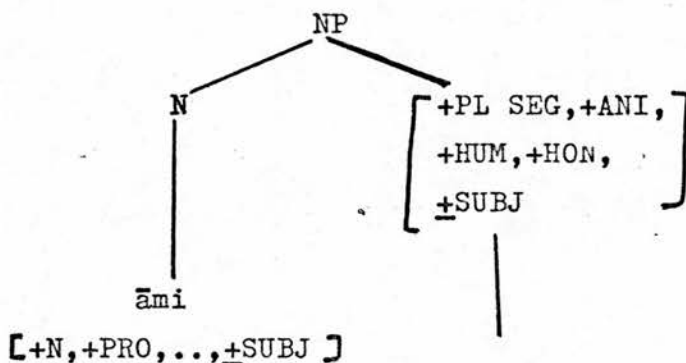
A personal pronoun can be inserted on the node N in (20) as both nouns and pronouns have the feature [+N]. The pronoun which will be inserted to (20) will be understood as plural as the noun phrase in (20) contains the abstract marker PL (cf., § 2.4.4). The first lexical insertion rule convert (20) into (21).

(21)



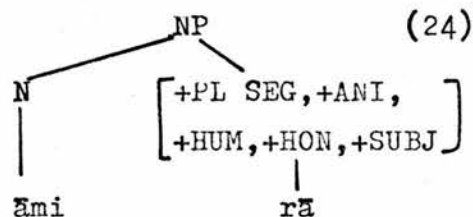
A later plural segment transformation will transform (21) into (22) (cf., § 2.4.4; § 2.5).

(22)

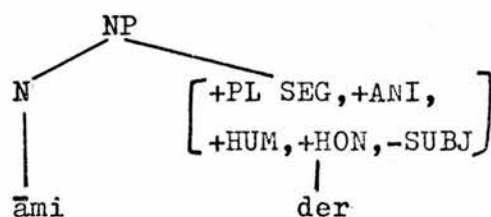


The second lexical insertion rule will insert rā to the plural segment if the noun phrase in (22) is a subject, and will insert der if it is not a subject. Thus we will arrive at (23) if the noun phrase is a subject, and at (24) if the noun phrase is not a subject after the second lexical attachment.

(23)



(24)



(23, 24) will generate forms *āmirā and *āmider respectively. These are ungrammatical. So these surface strings will pass through the morphophonological component which will derive āmarā from *āmi-rā, and āmāder from *āmi-der. The pronouns behave exactly like nouns in respect of plural segment transformation, and plural marker attachment. The plural segment transformation shows that the plural forms of Bengali personal pronouns need not be listed in the lexicon as we never list the plural forms of nouns.

Not only the plural forms of pronouns, but also other forms which are derivable by morphophonological rules, have been omitted from the lexicon. A Bengali pronoun changes morphologically when a plural marker or a case marker is suffixed to it. This change can be captured by morphophonological rules, and so it is not necessary to list these forms (cf., § 11.3). For example, āmi changes to āmā when the case marker ke is suffixed to it, and changes to āma when the plural marker rā is suffixed to it. Accordingly, the nonbasic forms of pronouns such as āmā, āma, tomā, tā, āpanā etc., have not been listed in the lexicon in this work.

Of the personal pronouns first and second person pronouns are used deictically (except in inclusive use, in which case they are always plural), and third person pronouns can be used both deictically and anaphorically. Pronouns, unlike nouns, do not take any constituent of the determiner except the PL, which is generated as a constituent of D in this work (cf., base rule 9, § 1.2). Consider the examples below.

- (25) a. *ekti tumi: *A you.
 b. *tumiṭi: *The you.
 c. tomarā: You (+PL).

(25a, b) are ungrammatical as the pronoun tumi occurs here with a SPEC. (25c) is a plural pronoun. The marker PL is generated in the underlying structure as a constituent of D in this work. This marker triggers the plural segment transformation and the plural markers are attached to derived nodes by the second lexical insertion rule.

A personal pronoun, when subject of a sentence, is contextually deletable in Bengali because the AUX agrees with the subject in person and grade (if any). Consider the examples in (26).

- (26) a. āmi tāke balechi: I have told him.
 b. āmārā tāke balechi: We have told him.
 c. tāke balechi: (I/we) have told him.

(26a, b) are unambiguous as these sentences have overt subjects. But (26c), where the subject is deleted, is ambiguous in that the deleted pronoun may be either āmi or āmārā. The AUX in Bengali agrees with the subject in person and grade, not in number.

4.2.2 Interrogative and Indefinite Pronouns.

We have given a list of the Bengali interrogative pronouns (all are ka-words) in (4). These are inherently interrogative. They are positively specified for the feature [INT], which distinguishes them from the personal and other pronouns. The interrogative and the indefinite pronouns which are lexically related to the interrogative pronouns have feature specifications in the following manner.

- (27) ke: [+N,+PRO,+INT,+COUNT,+DEF,+SPECIFIC,+HUM,-1,-2,+3,+HON]
keu: [+N,+PRO,-INT,-COUNT,-DEF,-SPECIFIC,+HUM,-1,-2,+3,+HON]

ki: [+N,+PRO,+INT,-COUNT,+DEF,+SPECIFIC,-HUM/+ABS,
-1,-2,+3,-HON]

kichu: [+N,+PRO,-INT,-COUNT,-DEF,-SPECIFIC,-ANI,+ABS,
-1,-2,+3,-HON]

kab: [+N,+PRO,+INT,-COUNT,+DEF,+TEMP, DAY]

kakhan: [+N,+PRO,+INT,-COUNT,+DEF,+TEMP, TIME]

kakhano: [+N,+PRO,-INT,-COUNT,-DEF,-SPECIFIC,+TEMP, TIME]
keno: [+N,+PRO,+INT,-COUNT,-DEF,+ABS]

kothā: [+N,+PRO,+INT,-COUNT,+DEF,+SPECIFIC,+LOC]

kothāo: [+N,+PRO,-INT,-COUNT,-DEF,-SPECIFIC,+LOC]

These pronouns are used deictically only, and so they are listed only in the first lexicon. Among these pronouns only ke has the feature [+COUNT], and it is the only pronoun that undergoes the plural segment transformation when inserted in a noun phrase with the marker PL. By a general morphological rule it becomes kā when a case marker or plural marker is suffixed to it.

Traditional Bengali grammars often include a pronoun kā usually found in an object position as in the sentence tumi kāke khūjcho?: 'Whom are you looking for?'. But our rule shows that kā is a suppletive form of ke.

We have specified the pronouns ki, kab and kothā with the feature [-COUNT]. These pronouns are usually used as singular, but some sentences show that they can be used as plural. In the plural sense they do not take any plural marker, instead they are reduplicated. Consider the examples below.

- (28) a. tumi ki ki cāo?: What do you want?
b. tumi kabe kabe sekhāne giyechile?: Which days had you been there?
c. tumi kothāy kothāy yābe?: Where will you go?

Here ki ki, kabe kabe and kothāy kothāy indicate plurality not as a set, but distributively. This sort of plurality which refers to things individually can be expressed by ke, too. This is shown in (29).

- (29) a. ke ke esechen?: Who have come?
 b. tumi kāke kāke ceno?: Whom do you know?

The way the reduplicated pronouns such as ki ki, kabe kabe and kothāy kothāy in (28), and ke ke and kāke kāke in (29) are understood suggests that they should not be considered as single noun phrases. They should be considered as conjoined noun phrases for the following reasons: (a) each of the reduplicated pronouns takes a case marker, and (b) they are understood as distinct noun phrases rather than a single plural noun phrase. We can relate the examples in (28) to those in (30), and the examples in (29) to those in (31).

- (30) a. tumi ki ebaṅ ki cāo?: What and what do you want?
 b. tumi kabe ebaṅ kabe sekhāne giyechile?: Which day and which day had you been there?
 c. tumi kothāy ebaṅ kothāy yābe?: Where and where will you go?
- (31) a. ke ebaṅ ke esechen?: Who and who have come?
 b. tumi kāke ebaṅ kāke ceno?: Whom and whom do you know?

Although (30) and (31) are stylistically infelicitous, they are semantically equivalent to (28) and (29) respectively. We assume that the sentences in (28, 29) are derived respectively from (30, 31), by deleting the conjunctions. Furthermore (30, 31) themselves are derived from coordinate conjoined structures. So we can consider that ki ki, kabe kabe, kothāy kothāy, ke ke and kāke kāke are not plural forms, but are conjoined pronouns.

For this reason we have specified these pronouns as [-COUNT], which means that they cannot be pluralized.

Transformational studies make a distinction between a yes-no question and wh-question. The scope of a yes-no question is usually an entire sentence, and the scope of a wh-question is usually a noun phrase. In our grammar we generate an abstract morpheme Q in case of a yes-no question, but no such Q is generated in case of a wh-question. We consider that a wh-question is an interrogation about some noun phrase, and this arises from the inherent feature [+INT] of the interrogative pronouns and other interrogative lexical items. Consider the examples below.

- (32) a. ke esechen?: Who has come?
 b. tumi ki bhālabāso?: What do you love?
 c. tumi kon baiṭi cāo?: Which book do you want?

All questions in (32) are ^{wh-} (ka-question in Bengali) questions, and they arise from the inherent interrogative property of the pronouns ke and ki, and that of the interrogative deictic kon. All these items are generated in the underlying structure in this grammar. So it is unnecessary to posit the Q in the underlying structure of a wh-question sentence (cf., § 1.4).

4.2.3 Abstract, Locative and Temporal Pronouns³

The pronoun tā, which is mainly an abstract pronoun, has the following features:

- (33) tā: [+N, +PRO, -INT, +DEF, -ANI, +ABS, -COUNT, -1, -2, +3, -HON]

We have discussed this pronoun in elsewhere (cf., § 5.10; § 10.1).

³ The locative and temporal pronouns have been taken as pronouns because of their pronominal function. But structurally they are full noun phrases composed of demonstrative deictics and locational and temporal nouns (cf., § 5.9).

The locative pronouns can be used deictically as well as anaphorically. Of the temporal pronouns, ekhan: 'Now' is used deictically, and okhan: 'Then' and takhan: 'Then' can be used both deictically and anaphorically. These pronouns have the features as shown below.

(34) ekhān: [+N,+PRO,-INT,-COUNT,+DEF,+LOC,-FAR]

okhān: [+N,+PRO,-INT,-COUNT,+DEF,+LOC,+FAR, IN SIGHT]

sekhān: [+N,+PRO,-INT,-COUNT,+DEF,+LOC,+FAR, OUT OF SIGHT]

(35) ekhan: [+N,+PRO,-INT,-COUNT,+DEF,+TEMP,-FAR]

okhan } : [+N,+PRO,-INT,-COUNT,+DEF,+TEMP,+FAR]
takhan }

4.3 Pronominalization: Some Approaches and Problems.

We have said that pronominalization is the process which derives pronouns in syntactic structures. This can be an underlying derivation or a transformational derivation. In transformational studies the term 'Pronominalization' is mainly used to refer to transformational derivation of anaphoric pronouns. This process reduces some noun phrase into a pronoun or pronominal under certain conditions. Pronominalization, in a broad sense, will be taken as the process which derives pronouns either in the underlying structure or by the transformational reduction of some noun phrase under certain conditions. But in a narrow sense pronominalization will mean the transformational derivation of pronouns or pronominals. In this work we will be concerned mainly with the transformational derivation of anaphoric pronouns under coreference and certain other conditions. So we will use

the term pronominalization mainly to refer to the process which derives pronouns transformationally. This process specifies some noun phrase as [+PRO] under certain conditions. These noun phrases are [-PRO] in the underlying structure, but acquire the feature [+PRO] transformationally when certain conditions are satisfied. The pronominalized noun phrases are usually realized as pronouns in the surface structure.

Pronominalization can be of various types, such as:

- (a) Pronominalization Proper, (b) Reflexivization,
- (c) Relativization, and perhaps, (d) Sentence Pronominalization.

The process of identical head noun deletion in Bengali is a sort of pronominalization (cf., Chapter 9). All these processes have one thing in common: they reduce an underlying noun phrase into a pronoun or a pronominal. It is generally assumed that reciprocal structures involve pronominalization. But we have shown that a reciprocal structure in Bengali does not involve pronominalization (cf., Chapter 7).

In this grammar pronouns are derived in two ways. A deictic pronoun is generated in the underlying structure as a member of the category noun with the feature [+PRO]. But anaphoric pronouns are derived transformationally under certain conditions. Here we will review some approaches to pronominalization that have been taken by transformational grammarians since 1963.

During the past few years, pronominalization in English has attracted many linguists, who made many aspects of the process clear, and at the same time discovered many problems that still remain unsolved. Let us start with Lees and Klima's (1963) classic treatment of the process. They took pronouns in the traditional sense, that pronouns are words used in place of nouns.

They went further to show that 'this replacement is subject to rigid grammatical rules.' Their rule is reproduced below.

(36) Pronoun Rule:

$$X - \text{Nom} - Y - \text{Nom}' - Z \quad \text{---}\rightarrow \quad X - \text{Nom} - Y - \text{Nom}' + \text{Pron} - Z$$

Where $\text{Nom} = \text{Nom}'$, and where Nom is in a matrix sentence while Nom' is in a constituent sentence embedded within the matrix sentence.

In this rule the matrix Nom reduces the constituent Nom' into a pronoun under lexical identity. The phonological shape of the pronominalized Nom is given later by morphophonological rules.

Chomsky (1965) took this approach and added some refinements. He considers the pronominalization transformation as an erasure transformation, which uses a term X to delete another term Y when X and Y are identical. Although this is basically similar to Lees and Klima's approach, Chomsky made the operation more subtle by introducing the notion of coreferentiality, indicated in the underlying structure by identical indices. In his approach two nouns are coreferential if they have identical integers in the underlying structure. In Lees and Klima's approach lexical identity is enough for the application of the rule, but Chomsky's rule requires strict identity. That is, the nouns involved must be identical both lexically and referentially. Chomsky (1965, 145-46) says:

Suppose that certain lexical items are designated as "referential" and that by a general convention each occurrence of a coreferential item is assigned a marker, say, an integer, as a feature. The reflexivization rule can be formulated as an erasure operation that uses one Noun Phrase

to delete another. As in the case of relativization, the erasure leaves a residue, in particular, the feature [+HUM], and it introduces the new phonetic element self. Thus when applied to "I hurt I," the first Noun Phrase is used to delete the second, finally giving, "I hurt myself." But by the recoverability condition on deletion, the reflexivization rule (similarly, the pronominalization rule) will apply only when the integers assigned to the two items are the same. The semantic component will then interpret two referential items as having the same reference just in case they are strictly identical - in particular, in case they have been assigned the same integer in the deep structure.

We see that Chomsky requires strict identity for the application of any sort of pronominalization rule. There are problems with with the condition of strict identity as well as with the use of integers. Chomsky wants to attach the integers to the head nouns, not to noun phrases. But they should be assigned to noun phrases as coreferentiality exists between whole noun phrases (cf., McCawley (1968b, 137)).

The condition of strict lexical identity between the noun phrases involved in a pronominalization rule has given rise to many problems. Bach (1970) has shown that if pronominalization operates on full noun phrases including relative clauses, then some English sentences cannot be given any deep structure. This problem has now come to be known as 'Bach's paradox'. Bach even suggested that there may be no such operation as pronominalization. This problem has attracted many linguists such as Dougherty (1969), McCawley (1970), Karttunen (1971) and Jackendoff (1972).

Postal (1970a) takes a transformational approach to pronominalization, and claims that pronouns are present in the underlying structure as features of nouns and are realized as articles in the surface structure. Our approach to pronominalization has some similarity to that of Postal, but we do not consider the Bengali pronouns as articles in the surface structure.⁴ Postal (1970a, 58) says:

I mention all these only because it is fundamental to my basic claim which is that the so-called pronouns I, our, they etc., are really "articles", in fact types of "definite" articles. However, article elements are introduced only as "segments" in the intermediate syntactic structures. In the deepest structures they are, I shall suggest, not present segmentally but are represented as syntactic features of nouns, features analogous to animate, human, countable, etc.,.

The way we derive deictic pronouns has some similarity to Postal's process, but has some dissimilarities, too. We consider that nouns and pronouns differ in the underlying structure only in respect of the feature [PRO], which is negatively specified for a noun and positively for a pronoun. So a noun has the features [+N, -PRO], and a pronoun has the feature [+N, +PRO]. A deictic pronoun is inserted in the underlying structure to a node N just as we insert a noun to it. We do not segmentalize a deictic pronoun from a complex symbol of a noun as Postal does. Postal

4 In our grammar there is nothing which can be called an 'article' as that is generally understood. We have taken the view that pronouns are deictic nouns (cf., § 4.1.5). The pronouns have both deictic and nominal features inherently in them, and so they function deictically as well as nominally.

derives anaphoric pronouns transformationally. The pronominalization transformation specifies a noun stem as [+PRO] if it is identical to some other noun in the sentence. Postal (1970a, 61-62) says:

The process of pronominalization is, I assume, a rule which specifies a noun stem as [+PRO] if it is identical to some other noun in the same sentence, subject to appropriate and not entirely understood conditions. The rule of reflexivization is one which specifies a noun stem as [+reflexive] and [+PRO] subject to its identity to another noun stem in the same simple structure (at the point of reflexivization). All nouns start out in the deep structure forms [-reflexive], i.e., the specification [+reflexive] is only introduced transformationally. However, this is, as we have seen, not true of the feature specification [+PRO] which will be present in some noun bundles in the base, namely, in those underlying such surface NP as someone, he, I etc., in sentences like:

- (37) a. Someone saw Bill.
 b. He is clever.
 c. I don't believe that.

Similarly, Definitization involves specifying a noun stem as [+definite] (and generally but not always [-demonstrative] as well) subject to certain conditions including previous transformational specification of [+PRO]. Under these assumptions, the overall process of reflexivization which occurs in sentences like:

- (38) A boy hurt himself.

and pronominalization which occurs in sentences like:

(39) A boy said he would help.

are considered to be quite similar. Both involve specification of the repeated noun as [+PRO,+definite , -demonstrative]. The difference is whether or not the specification [+reflexive] is also assigned.

Our approach to pronominalization is similar in spirit to the above approach in that we derive anaphoric pronouns transformationally under certain conditions. But we differ from him in many technical aspects of the rule. In his approach an NP node dominates only an N associated with a bundle of features, and he specifies only the N as [+PRO] when the conditions for the rule are satisfied. In our grammar an NP has several underlying constituents, and the feature [+PRO] is specified on the head noun of the noun phrase, but this should be understood as that the whole noun phrase has been pronominalized.

Another approach to pronominalization is the interpretive approach taken by Dougherty (1969) and Jackendoff (1972). In this approach all pronouns are generated in the underlying structure, and coreferentiality or non-coreferentiality is marked in the surface structure by interpretive rules. As Jackendoff (1972) has discussed this hypothesis more fully, we will look into his approach here. Jackendoff (1972) calls his hypothesis 'Interpretive Theory', and generates both anaphoric and nonanaphoric pronouns in the underlying structure. He does not mark them for coreference in the underlying structure. This is done on the surface structure by interpretive rules.

Jackendoff (1972, 108) says:

In this approach, which I will call Interpretive Theory, noun phrases are unmarked for coreference relations in

deep structure. Rules of semantic interpretation establish relations between pairs of noun phrases, marking them coreferential or noncoreferential with each other.

He considers coreference to be an exclusively semantic property that cannot be referred to by transformations. In order to establish coreference or noncoreference relations between noun phrases in the surface structure he needs at least three devices: (a) a table of coreference, (b) consistency condition, and (c) well-formedness condition. He says (1972, 111):

We will express coreference relations explicitly in a table of coreference. Each entry in the table will consist of a pair of noun phrases and one of the relations coreferential or noncoreferential. In a complete semantic interpretation, the table will contain an entry for each pair of noun phrases in the sentence. The table will be built up one entry at a time by the application of semantic rules of coreference. Some of the semantic rules will make reference to already existing entries in the table, but transformations will never refer to it.

Jackendoff's table of reference will look like (40), which we reproduce from him:

(40) (Reflexivization, First approximation)

Enter the table:

$$NP^1 \quad \alpha \text{ coref} \quad \left[\begin{array}{c} NP^2 \\ \alpha \text{ reflexive} \end{array} \right] \text{ in the environment...}$$

OBLIGATORY

His grammar will generate sentences like those in (41), and his reflexive rule will interpret the subject and object noun phrases

in them as coreferential.

(41) a. *The boy shot herself.

b. *Finkelstein shot yourself.

He will rule out these sentences as ungrammatical in the long run by his consistency condition. Jackendoff claims that his theory is superior to the transformational theory of pronominalization, and it has no problem with sentences involving the so-called 'Bach's paradox'. Jackendoff wants to handle pronouns and epithets by the same rule. We will not go into the theoretical controversies here (cf., Postal (1971a), Stockwell et al (1973)).

4.3.1 Our Approach in Outline.

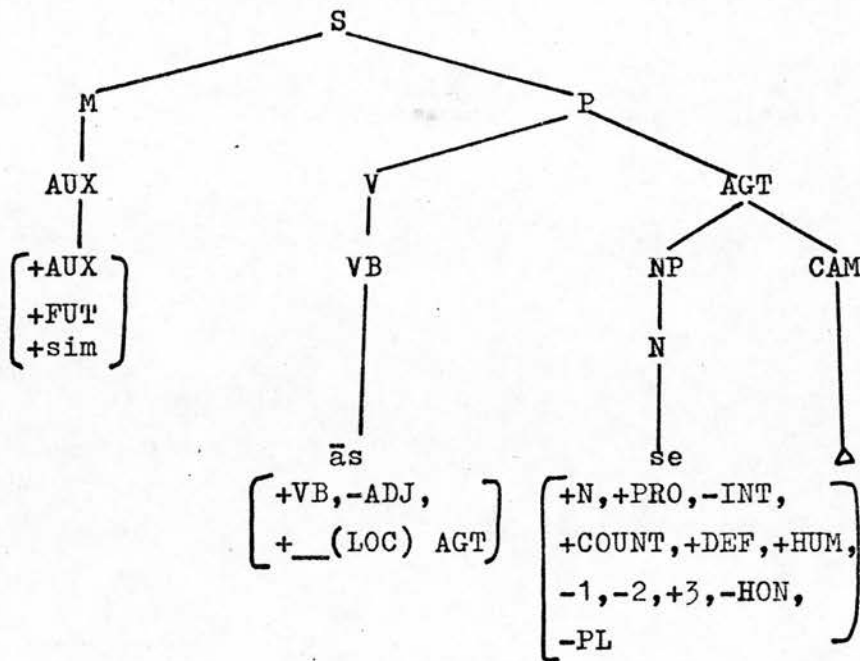
We generate deictic pronouns in the underlying structure, and derive anaphoric pronouns transformationally. We have taken the view that pronouns and nouns, in the underlying structure, differ only in the feature [PRO], which is positively specified for pronouns and negatively for nouns. Pronouns have the features [+N,+PRO], and nouns have the features [+N,-PRO]. Consider the following examples.

(42) a. se āsbe: He will come.
 1 2 1 2

b. tini dhākā yāben: He will go to Dacca.
 1 3 1 3

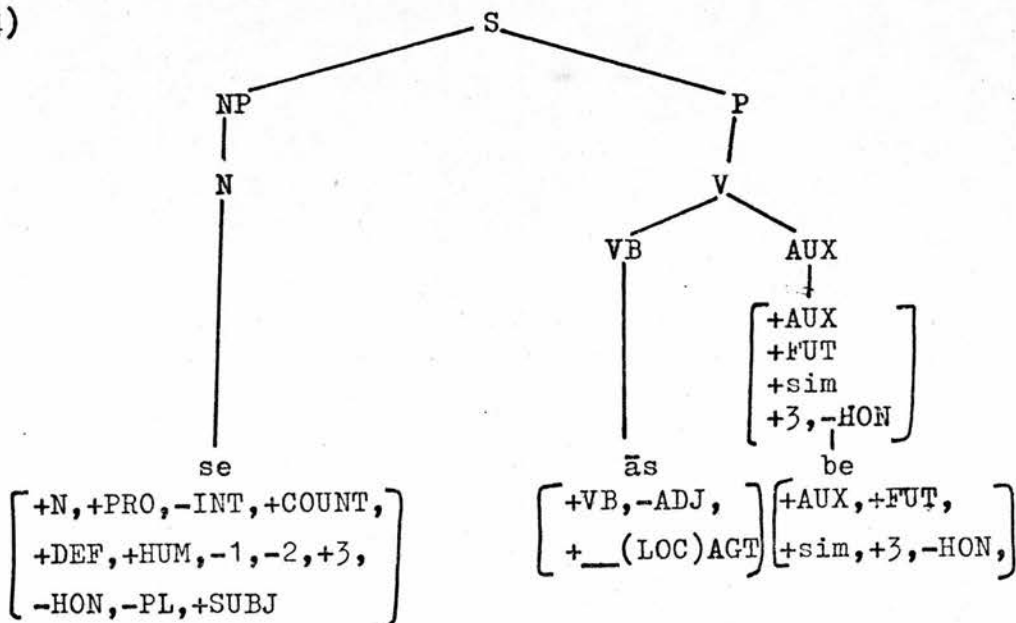
In (42) the pronouns se and tini are used deictically: they are not understood coreferentially with any noun phrase in these sentences. In such a case pronouns are generated in the underlying structure. For example, (42a) has the underlying structure (43)(see page 161). The pronoun se has been attached to the node N dominated by the AGT NP in (43) because this tree satisfies all conditions for its insertion on this node. As this

(43)



pronoun is generated in the underlying structure, it will be understood deictically. By the rule of subject marking, the AGT in (43) will be marked as subject, and will be placed at the sentence-initial position (optionally). Then the Subject-AUX concord rule will copy the person and grade features of the subject onto the AUX, and the AUX-MOVEMENT rule will adjoin the AUX as the right daughter of V. Then the V will be placed at the sentence-final position. After the last rule of the transformational component the second lexical insertion rule will attach be to the AUX. We will get the surface structure tree (44) after all these operations on (43).

(44)



The terminal string of (44) will generate se āsbe, which is (42a).

Now consider the following sentences, where the pronouns, in one reading, are understood coreferentially with some noun phrase in their respective sentences.

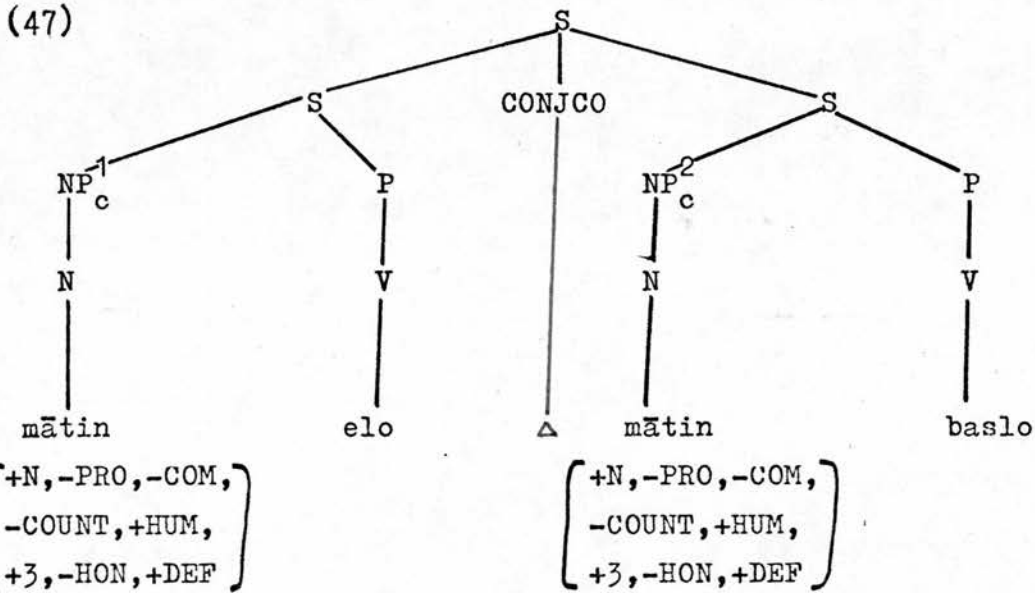
- (45) a. mātin_c elo o se_c baslo: Matin came and he sat.
 2 3 4 2 3 4 5
- b. yadi mātin_c āse, tabe āmi t̄ake_c e-biṣaye balbo: If
 1 3 4 5 6 7 8 9 1
 matin comes, then I shall tell him about this matter.
 3 4 5 9 6 7 8

In (45a) se refers anaphorically to mātin, and in (45b) t̄ake refers anaphorically to mātin, in one reading. We will derive such anaphoric pronouns by a transformational pronominalization rule, which reduces an underlying noun phrase into a pronoun when certain conditions are satisfied. The pronouns se and t̄ake in (45a, b) respectively are ambiguous in that they can refer either to mātin in these sentences or to someone else extrasententially. That is, these pronouns may be understood anaphorically as well as deictically. So when a pronoun is used anaphorically it will be derived transformationally, and when used deictically it will be derived in the underlying structure. This will explain the ambiguity of the pronouns in sentences like (45). When the pronouns are used anaphorically to mātin in (45), they are derived transformationally from the sentences in (46), which underlie (45), respectively.

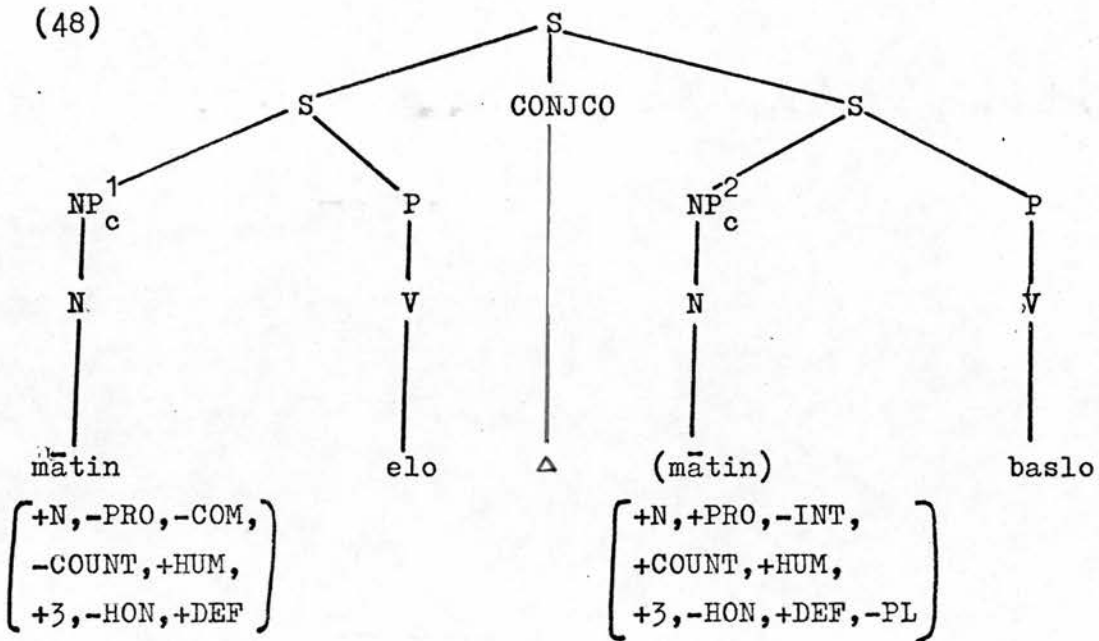
- (46) a. mātin_c elo o mātin_c baslo: Matin came and Matin sat.
- b. yadi mātin_c āse, tabe āmi mātinke_c e-biṣaye balbo: If
 Matin comes, then I shall tell Matin about this matter.

In (46) each sentence has two noun phrases which are understood

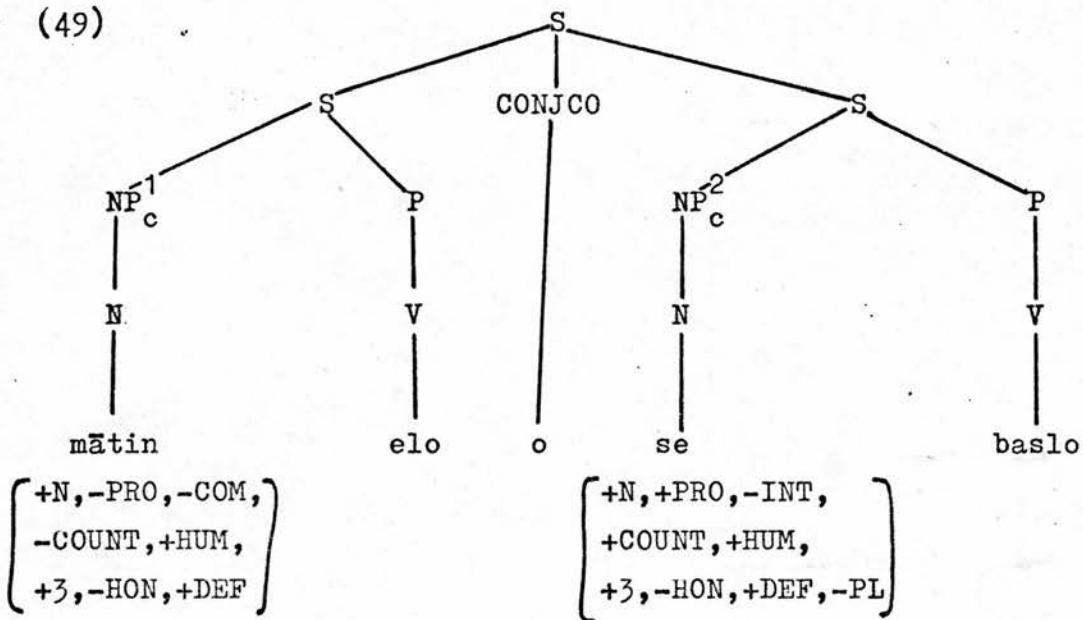
coreferentially. These noun phrases have been marked by identical indices. The pronominalization rule can apply in these structures as they satisfy all the conditions for the pronominalization rule (cf., § 4.3.2). For example, (46a) has the structure (47) immediately underlying it.



(47) is a coordinate conjoined structure where each conjunct has a noun phrase which is coreferential and identical in head noun with one in the other conjunct. This structure satisfies the conditions for the pronominalization rule. The rule will apply here forwards. That is, NP^1 will be used to specify NP^2 as $[+PRO, -INT]$. These features will be added to the CS associated with the head noun of the pronominalized noun phrase (NP^2). After the last rule of the transformational component the second lexical insertion rule will replace the first lexical item under NP^2 by the pronoun se by the matching condition of the second lexical insertion procedure. (cf., § 11.2). The pronominalization transformation will transform (47) into (48) (see page 164).



After the last rule of the transformational component the second lexical insertion rule will insert the pronoun se to the pronominalized noun phrase, and the surface structure (49) will be derived.



(49) will generate mātin_c elo o se_c baslo, which is (45a).

We have taken coreferentiality as the major condition for the pronominalization rule, but the rule does not apply between any coreferential noun phrases. The rule can apply only if the

head nouns of the coreferential noun phrases are identical. Coreferentiality between noun phrases is marked in the underlying structure by identical indices in the Chomskyan fashion (cf., Chomsky (1965, 145-46)).

We have taken a transformational, not an interpretive, approach to pronominalization. The decision is motivated by linguistic as well as practical reasons. The interpretive theory as a theory of language is unsatisfactory. And from the practical point of view a transformational approach to pronominalization, which has been extensively used in transformational studies, makes things simpler when one comes to study the syntax of a language, which has not been described in any recent framework.

Jackendoff (1972) claims that pronouns and epithets should be dealt with by the same rule, as they behave similarly. He (1972, 111) wants to 'mark epithets as special lexical items which may function as pronouns in certain contexts of the pronominalization rule, adding the lexical meaning to the intended attributes of the person they refer to.' But epithets are full noun phrases, not 'lexical items'. We consider that epithets are full noun phrases used coreferentially with some noun phrase in a sentence. Consider the examples below.

- (50) a. $m\bar{a}tin_c\text{elo } o\text{ } m\bar{a}tin_c\text{h}\bar{a}slo$: Matin came and Matin smiled.
 b. $m\bar{a}tin_c\text{elo } o\text{ } du\text{ṣ}\text{ṭu}\text{ṭi}_c\text{h}\bar{a}slo$: Matin came and the naughty boy smiled.
 c. $m\bar{a}tin_c\text{elo } o\text{ } se_c\text{h}\bar{a}slo$: Matin came and he smiled.

In (50a) the proper noun phrases $m\bar{a}tin$ and $m\bar{a}tin$ are coreferential, and have identical head nouns. In (50b) $m\bar{a}tin$ and $du\text{ṣ}\text{ṭu}\text{ṭi}$, although coreferential, are not identical. In (50c) the noun

phrase duṣṭuṭi has its own semantic content, and so it is generated in the underlying structure and is marked as coreferential with mātin. In (50c) se refers anaphorically to mātin, but unlike duṣṭuṭi, it has no semantic content.

Therefore the pronominalization rule can reduce the forward noun phrase mātin into a pronoun and generate (50c); but duṣṭuṭi, although coreferential with mātin in (50b), cannot be reduced into a pronoun. Epithets are underlying noun phrases used coreferentially with some noun phrase in a sentence. Moreover, pronouns are always determinate, but epithets are not. For example, only se can be used in (50c) in order to refer to mātin anaphorically, but in (50b) any epithet such as jubarāj: 'Crown prince', jātir pitā: 'Father of the nation', and so forth could have been used in place of duṣṭuṭi. We assume that epithets should be treated like underlying coreferential noun phrase, but not like anaphoric pronouns.

4.3.2 What is Pronominalizable?

The pronominalization rule involves noun phrases, and applies when the following basic conditions are satisfied.

(51) Conditions for the Pronominalization Rule.

- a. In a structure there must be two noun phrases which are coreferential.
- b. The relevant noun phrases must have identical head nouns.
- c. At the point the rule applies, the pronominalizable noun phrase must be definite.
- d. The rule will obey the directionality constraints.

4.3.3 Pronominalization and Cases.

We will consider here whether cases play any role in the pronominalization proper transformation. The pronominalization proper rule takes place between coreferential noun phrases if the conditions for the rule are satisfied (cf., § 4.3.2). The rule can apply between coreferential noun phrases holding identical as well as non-identical case relations. That is, a noun phrase holding a case role C_i can be used to pronominalize another noun phrase holding the case relation C_i or C_n if the noun phrases are coreferential and identical in head noun. Consider the examples below.

- (52) a. $m\bar{a}tin_c par\dot{c}he$ o $m\bar{a}tin_c likhche$: Matin is reading and Matin is writing.
 b. $m\bar{a}tin_c par\dot{c}he$ o $se_c likhche$: Matin is reading and Matin is writing.
- (53) a. $yadio m\bar{a}tin_c minuke$ mereche, $tabuo m\bar{a}tin_c minuke$ bhālabāse:
 Although Matin has hit Minu, yet Matin loves Minu.
 b. $yadio m\bar{a}tin_c minuke$ mereche, $tabuo se_c minuke$ bhālabāse:
 Although Matin has hit Minu, yet he loves Minu.

(52a) has two noun phrases which are coreferential and identical in head nouns. Both these noun phrases are AGT's. Pronominalization can apply here forwards, and we derive (52b) by the application of the pronominalization rule forwards in (52a). As (52b) is grammatical, we see that pronominalization can apply between noun phrases holding identical case relations. (53a) has two coreferential noun phrases, which hold distinct case relations. The noun phrase $m\bar{a}tin$ in the left conjunct is an AGT, and $m\bar{a}tin$ in the right conjunct is a DAT. As these noun phrases

are coreferential and identical in head noun we can apply the pronominalization rule on the underlying structure of (53a) in the forward direction and derive (53b). As (53b) is grammatical, we see that the pronominalization rule can apply between noun phrases holding non-identical case relations.

The pronominalization rule applies without any complication between AGT, DAT, NEUT and INS noun phrases (of course, by definition of different cases, noun phrases holding some case relations cannot be understood as coreferential with noun phrases holding some other case relation. For example, AGT noun phrases cannot be understood as coreferential with INS noun phrases because AGT is typically animate and INS is inanimate.), but some complications arise in the pronominalization of noun phrases holding LOC case role. Consider the examples below.

- (54) a. $\text{ḍhākā}_c \text{ekṭi choṭa 'sahar o ḍhākā}_c \text{ekṭi sundar 'sahar:}$
 Dacca is a small city and Dacca is a beautiful city.
- b. $?ḍhākā}_c \text{ekṭi choṭa 'sahar o se-ṭi}_c \text{ekṭi sundar 'sahar:}$
 Dacca is a small city and Dacca is a beautiful city.
- c. $\text{ḍhākā}_c \text{ekṭi choṭa o sundar 'sahar:}$ Dacca is a small and beautiful city.
- (55) a. $\text{tini ḍhākāy}_c \text{thāken ār āmio ḍhākāy}_c \text{thāki:}$ He lives in Dacca and I live in Dacca too.
- b. $\text{tini dhākāy}_c \text{thāken ār āmio sekhāne}_c \text{thāki:}$ He lives in Dacca and I live there too.
- c. $*\text{tini ḍhākāy}_c \text{thāken ār āmio se-ṭite}_c \text{thāki:}$ He lives in Dacca and I live in that too.
- (56) a. $\text{se-bāgānti}_c \text{choṭa ār se-bāgānti}_c \text{sundar:}$ That garden is small and that garden is beautiful.

- b. se-bāgānti_c choṭa ār se-ṭi_c sundar: That garden is small and that is beautiful.
- c. se-bāgānti_c choṭa ār sundar: That garden is small and beautiful.
- (57) a. se-bāgānti_c choṭa, tabe se-bāgāntite_c anek phul āche: That garden is beautiful, but there are many flowers in that garden.
- b. se-bāgānti_c choṭa, tabe se-ṭite_c anek phul āche: That garden is small, but there are many flowers in that.
- c. se-bāgānti_c choṭa, tabe sekhāne_c anek phul āche: That garden is small, but there are many flowers there.

In (54a) dhākā in the left conjunct and dhākā in the right conjunct are coreferential. dhākā is a [-COM, -COUNT, +LOCATION] noun and big in size. These noun phrases in (54a) hold NEUT case role, although they are inherently locational. As dhākā holds NEUT case role here, this noun phrase cannot be reduced to a pronoun. If we reduce the noun phrase dhākā into a pronominal by the application of the pronominalization rule forwards, we will get (54b), which is ungrammatical. In order to derive se-ṭi in (54b) we must have the noun phrase *se-dhākāṭi in the right conjunct in (54a), and the pronominalization rule will delete the head noun of this noun phrase leaving DEIC and CL behind. But we cannot have *se-dhākāṭi: 'That Dacca' in the underlying structure, as this noun, being [-COM, -COUNT], disallows any determiner constituent. But if (54b) is derived at all, it would be derived not from the underlying structure of (54a), but from dhākā ekṭi choṭa sahar o se-saharṭi ekṭi sundar sahar: 'Dacca is a small city and that city is a beautiful city'.

In this sentence the second conjunct contains redundant material; and coreferentiality exists here between ekṭi choṭa sahar and se-saharṭi, but not between dhāka and se-saharṭi. In (54a) the pronominalization rule does not reduce the forward noun phrase into a pronoun; here the pronominalization rule deletes the coreferential noun phrase in the right conjunct, and later by conjunction reduction we get (54c). This is similar to the application of the pronominalization rule involving [-HUM] and [-ANI] noun phrases where the NP^P has an underlying structure of the form DEF SPEC N (cf., § 5.3.1). In (55a) dhākā in the first conjunct and dhākā in the second conjunct are identical and coreferential, and they both are LOC's. The pronominalization rule applies here in the forward direction, and the pronominalized noun phrase is realized as the locative pronoun sekhān. (55c) is ungrammatical because it is derived from the unacceptable underlying structure *tini dhākāy thāken ār āmio se-dhākāṭite thāki: 'He lives in Dacca and I live in that Dacca, too'. Thus we see that a noun phrase having a [-COM, -COUNT, +LOCATION] head noun cannot be reduced to a pronoun if it holds any case relation other than LOC; in such a case the whole noun phrase is deleted under coreferentiality. In (56a) the indexed noun phrases are coreferential and identical in head noun. The noun in these noun phrases is [+COM, +COUNT, +LOCATION]. This noun can occur with DEIC SPEC in a noun phrase. In (56a) the coreferential noun phrases are NEUT's. (56a) is a structure where pronominalization rule can apply forwards. But as Bengali has no [-HUM] or [-ANI] pronoun, the pronominalization rule, instead of reducing the NP^P into a pronoun, will delete the head noun of the NP^P (as it has DEIC N SPEC intermediate structure).

In this way we will get (56b). Subsequently, by deleting the residual noun phrase se-ti we will get (56c). In (57a) se-bagānti in the left conjunct is NEUT and in the right conjunct is LOC. The pronominalization rule can apply forwards in (57a). The NP^P (bāgāntite in the right conjunct) has a [+COM,+COUNT,+LOCATION] noun as its head, and the location is understood as small. The pronominalization rule can delete the head noun of the NP^P (although it is a LOC) leaving the DEIC SPEC behind, and derive (57b). se-ti in (57b) is not a locative pronoun, although se-tite is a locative expression. The pronominalization rule, in the other way, can reduce the NP^P in (57a) into a locative pronoun because the pronominalized noun phrase is a LOC. In this way we can derive (57c), where the pronominalized noun phrase has been realized as a locative pronoun.

We see, in Bengali, that a LOC noun phrase is pronominalized into a locative pronoun if the head noun is [-COM,-COUNT,+LOCATION] or a [+COUNT,+LOCATION] noun indicating a vast location. But if the head noun of a locative expression is a [+COUNT] noun which indicates a small location, it can be pronominalized into a locative pronoun, or the pronominalization rule can delete the head noun of the NP^P if it has an intermediate structure DEIC N SPEC. We know that this sort of head noun deletion is a general operation for the pronominalization rule when the noun phrases involved have [-HUM] or [-ANI] head nouns (cf., § 5.3.1). This does not depend on the case relation of the noun phrase. But a noun phrase can be pronominalized into a locative pronoun only when the noun phrase is a LOC. This shows that the case relations of noun phrases have some effect on pronominalization. This is generally seen in LOC case, which

includes both locational and temporal expressions. But we have seen that a LOC noun phrase need not be reduced into a locative pronoun by the pronominalization rule if the head noun of the pronominalizable noun phrase is a count noun and indicates a small location.

The noun phrases which generally take part in pronominalization usually hold AGT, DAT, NEUT, LOC and INS, but ESS noun phrases usually do not take part in anaphoric pronominalization.

Consider the examples below.

- (58) a. hāsān sāheḅ lekhak ārahmān sāhebo lekhak: Mr Hasan is an author and Mr Rahman is an author, too.
- b. ?hāsān sāheḅ lekhak ārahmān sāhebo tā: Mr Hasan is an author and Mr Rahman is that, too.

In (58a) the predicate nominal lekhak in the first conjunct is identical with the predicate nominal in the second conjunct. They are not coreferential, although there is some sense-identity between them. If the pronominalization rule applies here the pronominalized noun phrase will be realized as the pronoun tā, which has the features [-HUM, -ANI, +ABS]. That is, lekhak is considered here as a mass or an abstract object. But this sort of nonanaphoric pronominalization due to lexical and sense-identity between predicate nominals can take place in [-ANI, +ABS] nouns. Consider the examples below.

- (59) a. e-ṭi bai āro-ṭio bai: This is a book and that is a book too.
- b. e-ṭi bai āro-ṭi tāi: This is a book and so is that.

The predicate nominal in the right conjunct in (59a) has been reduced into the pronoun tā in (59b) due to lexical and sense-identity.

4.3.4 Coreferentiality versus Definitization.

We have taken coreferentiality between the noun phrases as the major condition for the application of the pronominalization rule. But Kuroda (1968, 277) says:

.. the grammatical process which has hitherto been understood under the name of pronominalization is divided into two transformations, one Definitization and the other, we shall call Pronominalization.

Kuroda, of course, did not avoid coreferentiality as a condition. In his approach the coreferential noun phrases are indefinite in the underlying structure and the definitization rule makes the forward noun phrase definite in order to apply the pronominalization rule. We reproduce Kuroda's ^{rule} (1968, 278) below:

(60) $N_1 \quad X \quad \text{Det} \quad N_2 \quad \text{---} \rightarrow \quad N_1 \quad X \quad \text{THAT} \quad N_2$

if $N_1 = N_2$ (co-referential)

The problem with this rule is that it applies only forwards. But pronominalization, in certain cases, can apply backwards. Kuroda's rule is formulated in order to establish syntactic coreferentiality between noun phrases which are notionally assumed to be coreferential. Consider the examples in (61).

- (61) a. eḳti meye_celo o eḳti meye_cballo: A girl came and a girl said.
- b. eḳti meye_celo o meyeṭi_cballo: A girl came and the girl said.
- c. eḳti meye_celo o se_cballo: A girl came and she said.

In (61a) two indefinite noun phrases are assumed to be coreferential, which is indicated by the identical indices. But syntactically they cannot be taken as coreferential. As these noun

phrases are assumed to be coreferential, the definitization rule will apply in (61a) and will make the forward noun phrase definite. In this way we will get (61b), where the indexed noun phrases are understood as coreferential syntactically. The pronominalization rule will then apply in (61b) and derive (61c).

But some noun phrases are nonanaphorically definite (that is, they should be definite in the underlying structure), and the definitization rule cannot apply backwards; so we will make the definite/indefinite choice of noun phrases in the underlying structure in all types of pronominalization rule except in relativization. We adopt the definitization rule in restrictive relativization for the reason that restrictive relativization in Bengali is a kind of definitization rule, and we avoid it elsewhere because definitization by a rule in most cases is technically complex.

Postal (1970a) also talks of a definitization rule in relation to pronominalization. Postal (1970a, 62) says:

Definitization involves specifying a noun stem as
 [+definite] (and generally but not always [-demonstrative]
 as well) subject to certain conditions including previous
 transformational specification [+PRO] .

In his approach pronominalization precedes definitization, which is unacceptable. This rule, if adopted in a grammar, should apply prior to the pronominalization rule in order to make pronominalizable noun phrases definite. This rule makes noun phrases definite so that the pronominalization rule can apply between them.

PRONOMINALIZATION PROPER5.0 Introduction.

In the previous chapter we have defined and exemplified what we mean by pronominalization. (cf., § 4.0; § 4.3; § 4.3.1). The term pronominalization covers a number of basically similar processes which use one noun phrase to pronominalize another noun phrase under certain conditions. In this chapter we will deal with Pronominalization Proper, which is known in transformational literature as 'Simple Pronominalization'. The term 'Simple Pronominalization' was first used by Lees and Klima (1963) in their paper 'Rules for English Pronominalization' for the process that utilizes one noun phrase (Nom, in their language) to pronominalize another noun phrase, and the pronominalized noun phrase is realized as a 'simple personal pronoun' in the surface structure. They called this rule 'Pronoun Rule'. For the same process we use the term 'Pronominalization Proper' borrowed from Stockwell et al (1973), because it is more appropriate for this phenomenon. Pronominalization Proper is the rule that utilizes one noun phrase to reduce another noun phrase into a personal pronoun under certain conditions (cf., § 4.3.2). Some people refer to this process simply by the term 'Pronominalization'. (cf., Ross (1967b), Langacker (1969)). We will also sometimes refer to this process simply by the term Pronominalization, for brevity.

Now consider some examples where pronominalization proper takes place:

- (1) a. $\bar{b}\bar{a}d\bar{a}l_c$ $\bar{e}l\bar{o}$ o $\bar{b}\bar{a}d\bar{a}l_c$ $\bar{b}a\bar{s}l\bar{o}$: Badal came and Badal sat.
 2 3 5 2 3 5
- b. $\bar{b}\bar{a}d\bar{a}l_c$ $\bar{e}l\bar{o}$ o $\bar{s}e_c$ $\bar{b}a\bar{s}l\bar{o}$: Badal came and he sat.
 2 3 4 5 2 3 4 5
- (2) a. $y\bar{a}d\bar{i}$ $\bar{m}\bar{a}t\bar{i}n_c$ $\bar{a}se$, $\bar{t}a\bar{b}e$ $\bar{m}\bar{a}t\bar{i}n_c$ $\bar{p}u\bar{r}a\bar{s}k\bar{a}r$ $\bar{p}\bar{a}b\bar{e}$: If Matin
 1 3 4 6 7 1
 comes, then Matin will get (a) prize.
 3 4 -7- 6
- b. $y\bar{a}d\bar{i}$ $\bar{m}\bar{a}t\bar{i}n_c$ $\bar{a}se$, $\bar{t}a\bar{b}e$ $\bar{s}e_c$ $\bar{p}u\bar{r}a\bar{s}k\bar{a}r$ $\bar{p}\bar{a}b\bar{e}$: If Matin
 1 3 4 5 6 7 1
 comes, then he will get (a) prize.
 3 4 5 -7- 6

There are two coreferential noun phrases in (1a) and (2a), as indicated by the identical indices. These structures satisfy the conditions for the pronominalization proper transformation. This rule utilizes the left noun phrase to pronominalize the right noun phrase in (1a) and (2a), and in the surface structure the pronominalized noun phrases are realized as personal pronouns generating (1b) and (2b) respectively.

We have said above that pronominalization proper is the transformation which derives anaphoric personal pronouns; but its boundary should be extended to accommodate the derivation of anaphoric locative and temporal pronouns. (cf., § 5.9). And due to an accidental gap there is no neuter pronoun (that is, a pronoun parallel to the English 'It') in SCB. But the pronominalization rule will apply between coreferential [-HUM] and [-ANI] noun phrases. In such a case, the pronominalization rule, instead of reducing a noun phrase into a pronoun, will delete either the head noun of the pronominalizable noun phrase or the entire pronominalizable noun phrase (cf., § 5.3.1).

Transformationally derived pronouns must be understood as anaphoric to their antecedents because the rule applies under coreference. Nonanaphoric (deictic) pronouns are not derived

transformationally: they are generated in the underlying structure (cf., § 4.3.1). Although anaphoric pronouns can refer to antecedents beyond a single sentence boundary, here we shall deal with the derivation of anaphoric pronouns within a single sentence boundary. The pronominalization proper rule will apply within a sentence boundary ($\# S \#$).

5.1 Antecedent and Anaphora: Rules of Coreference.

Pronominalization is a syntactic process which depends on the notion of coreference, and can apply in a structure only if it contains two coreferential noun phrases which are identical in head noun. Coreferentiality and identity of head noun between the noun phrases involved in the rule are two major conditions for its application. The rule depends on coreference as it derives anaphoric pronouns, and depends on identity of the head nouns of the noun phrases involved in the rule in order to restrict it from applying between coreferential but non-synonymous noun phrases (cf., § 4.3.1). In this work noun phrases are marked for coreferentiality in the underlying structure. We have not used the definitization rule for establishing syntactic coreferentiality between noun phrases (cf., § 4.3.4). As we have avoided the definitization rule, the definite/indefinite choice of noun phrases will be made in the underlying structure. The motivation behind this decision is similar to that of Stockwell et al (1973), who say (1973, 192):

In our grammar, we have had to assume that the definite/indefinite choice is made entirely at the deep structure level, since the problems connected with definitization by rule are so complex.

In the following sections we will consider what types of noun phrase are understood as coreferential in different types of Bengali sentence structure.

5.1.1 Coreference in Coordinate Conjoined Structure.

Here we will consider coreference relations of those noun phrases which have identical head nouns, as pronominalization applies only in such noun phrases. Identity of head nouns in the pronominalization rule is required for semantic reasons. Without this the rule will affect the meaning of a structure, by the application of the rule between coreferential but non-synonymous noun phrases. Consider the examples below.

- (3) a. *ekṭi chele_c elo o ekṭi chele_c ballo: A boy came and
 1 2 3 4 5 6 7 1 2 3 4
 a boy said.
 5 6 7
- b. ekṭi chele_c elo o cheleṭi_c ballo: A boy came and
 1 2 3 4 5 6 1 2 3 4
 the boy said.
 5 6
- c. ekṭi chele_c elo o se-cheleṭi_c ballo: A boy came and
 that boy said.
- d. cheleṭi_c elo o cheleṭi_c ballo: The boy came and
 the boy said.
- e. *cheleṭi_c elo o ekṭi chele_c ballo: The boy came and
 a boy said.
- f. *se-cheleṭi_c elo o ekṭi chele_c ballo: That boy came
 and a boy said.
- g. mātin_c elo o mātin_c ballo: Matin came and Matin said.

In (3a) both the indexed noun phrases are [-DEF], and they are noncoreferential. In (3e, f) the left noun phrases are [+DEF] and the right noun phrases are [-DEF], and they are noncoreferential. In (3b) the left noun phrase is [-DEF] while the right

noun phrase is [+DEF]. They are coreferential. (3c) is similar. But the definite noun phrases in (3b, c) differ structurally in that the noun phrase cheleṭi in (3b) has the intermediate structure N SPEC, and the noun phrase se-cheleṭi in (3c) has the intermediate structure DEIC N SPEC (cf., § 2.3.6; § 2.4.5). In (3d) both indexed noun phrases are definite and coreferential. In (3g) the noun phrases are proper nouns, which are definite. They are coreferential. From these data we can arrive at the rules of coreference given in (4) for coordinate conjoined structure.

(4) Rules of Coreference in Coordinate Conjoined structure.

SI:	NP	CONJCO	NP
	1	2	3

Rules: 1 and 3 are coreferential, if

- a. 1 is [-DEF] and 3 is [+DEF],
- b. both 1 and 3 are [+DEF],

Noncoreferential otherwise.

Conditions: a. 1 and 3 are identical in head noun,
 b. They are elements of two conjuncts
 in a coordinate conjoined structure.

5.1.2 Coreference in Subordinate Conjoined Structure.

Now we will consider coreference relations of noun phrases with identical head nouns in subordinate conjoined structure. Consider the examples below.

- (5) a. yadi mātin_cāse, tabe mātin_cpuraskār pābe: If Matin comes, then Matin will get (a) prize.
- b. *yadi ekṭi chele_cāse, tabe ekṭi chele_cpuraskār pābe: If a boy comes, then a boy will get (a) prize.

c. yadi $\left\{ \begin{array}{l} \text{kono} \\ \text{ekṭi} \end{array} \right\}$ chele_cāse, tabe cheleṭi_cpuraskār pābe:

If some/any boy comes, then the boy will get (a) prize.

d. *yadi cheleṭi_cāse, tabe ekṭi chele_cpuraskār pābe: If the boy comes, then a boy will get (a) prize.

e. yadi cheleṭi_cāse, tabe cheleṭi_cpuraskār pābe: If the boy comes, then the boy will get (a) prize.

These examples show that two noun phrases which are elements of two conjuncts in a subordinate conjoined structure are understood as coreferential if they both are definite or if the left noun phrase is indefinite and the right noun phrase is definite.

From these data we can arrive at the rules of coreference given in (6) for subordinate conjoined structures.

(6) Rules of Coreference in Subordinate Conjoined Structure.

SI:	NP	CONJSUB	NP
	1	2	3

Rules: 1 and 3 are coreferential, if

a. 1 is [-DEF] and 3 is [+DEF] ,

b. both 1 and 3 are [+DEF] ,

Noncoreferential otherwise.

Conditions: a. 1 and 3 have identical head noun,

b. 1 is an element of the first conjunct and 3 of the second conjunct in a subordinate conjoined structure.

We find that the rules of coreference in subordinate conjoined structure and coordinate conjoined structure are identical.

They can be telescoped into a single rule given in (7).

(7) Rules of Coreference in Conjoined Structure.

SI:	NP	$\left\{ \begin{array}{l} \text{CONJCO} \\ \text{CONJSUB} \end{array} \right\}$	NP
	1	2	3

Rules: 1 and 3 are coreferential, if

a. 1 is [-DEF] and 3 is [+DEF],

b. both 1 and 3 are [+DEF],

Noncoreferential otherwise.

Conditions: a. 1 and 3 have identical head nouns,

b. They are elements of two conjuncts
in a conjoined structure.

5.1.3 Coreference in Complex Structures with Direct Discourse Constituent Sentence.

We have seen that (synonymous) coreferential noun phrases in coordinate and subordinate structures have identical head nouns. We will see now that coreferential noun phrases in complex structures with direct discourse constituent sentences cannot have identical head nouns. Consider the examples below.

- (8) a. ekṭi meye_c ballo, 'āmi_c yābo': A girl said, 'I shall go.'
1 2 3 4 5 1 2 3 4 5
- b. meyeṭi_c ballo, 'āmi_c yābo': The girl said, 'I shall go.'
- c. mātin_c ballo, 'āmi_c yābo': Matin said, 'I shall go.'
- d. *ekṭi meye_c ballo, 'ekṭi meye_c yābe': A girl said, 'A girl will go.'
- e. *meyeṭi_c ballo, 'meyeṭi_c yābe': The girl said, 'The girl will go.'
- f. *mātin_c ballo, 'mātin_c yābe': Matin said, 'Matin will go.'
- (9) a. hāsān_c ketakīke_a ballo, 'āmi_c tomāke_a bhālabāsi': Hasan told Ketaki, 'I love you.'
- b. ekṭi chele_c ekṭi meyeke_a ballo, 'āmi_c tomāke_a bhālabāsi':
A boy told a girl, 'I love you.'

- c. cheleṭi_c meyeṭike_a ballo, 'āmi_c tomāke_a bhālabāsi': The boy told the girl, 'I love you.'
- d. ekṭi chele_c meyeṭike_a ballo, 'āmi_c tomāke_a bhālabāsi': A boy told the girl, 'I love you.'
- e. cheleṭi_c ekṭi meyeke_a ballo, 'āmi_c tomāke_a bhālabāsi': The boy told a girl, 'I love you.'
- f. *hāsān_c ketakīke_a ballo, 'hāsān_c ketakīke_a bhālabāse': Hasan told Ketaki, 'Hasan loves Ketaki.'
- g. *ekṭi chele_c ekṭi meyeke_a ballo, 'ekṭi chele_c ekṭi meyeke_a bhālabāse': A boy told a girl, 'A boy loves a girl.'
- h. *cheleṭi_c meyeṭike_a ballo, 'cheleṭi_c meyeṭike_a bhālabāse': The boy told the girl, 'The boy loves the girl.'

There are verbs like mane kar: 'Assume, Think', bal: 'Say, Tell', dābī kar: 'Claim' in all languages; they can be classified as direct discourse verbs because they are usually used in direct discourse formation. These verbs take complement clauses which express the direct discourse or feeling of the matrix subject (cf., Kuno (1972)). In order to form a direct discourse sentences we require a 'narrator', who forms a direct discourse sentence. The function of the narrator is to reproduce the exact discourse or feeling of someone. For example, consider (8a). The narrator hears some girl saying āmi yābo, and forms (8a) by making the girl the subject of the sentence with the verb bal.

In (8a) the matrix subject has the features [+N, -PRO, +HUM, +3, -HON, -DEF, -PL], and the constituent subject, which is a pronoun, has the features [+N, +PRO, +HUM, +1, +DEF, -PL] (only the relevant features are mentioned). These two noun phrases are coreferential. In a direct discourse sentence it does not matter whether the matrix subject is definite or indefinite. If the

matrix subject is $[...+\beta, ...]$, and the constituent subject is is a pronoun with the features $[...+1, ...]$, and agrees with the matrix subject in number, then they are understood as coreferential. This can be seen in (8a, b, c). In a direct discourse sentence the matrix subject and the constituent subject cannot be understood as coreferential if they they are lexically identical and have the features $[+N, -PRO]$. Consider the examples (8d, e, f). In these sentences the matrix and constituent subject noun phrases are lexically identical and both have the features $[+N, -PRO]$. Although they agree in number, they cannot be understood coreferentially.

Now we come to the examples in (9), which are similar to those in (8) except that these have object noun phrases both in matrix and constituent sentences. In these sentences too, definiteness and indefiniteness of matrix subject noun phrases are irrelevant for coreference relations with constituent noun phrases. In (9a)-(9e) the constituent subject and the constituent object noun phrases respectively are coreferential with the matrix subject and matrix object noun phrases. But it is not the case that the constituent subject noun phrase should always be coreferential with the matrix subject noun phrase, and the constituent object noun phrase should always be coreferential with the matrix object noun phrase. The coreferential relation may exist between matrix subject and constituent object noun phrases, and with matrix object and the constituent subject noun phrases. The general rule is that the matrix subject noun phrase with the features $[+N, +\underline{PRO}, \alpha PERSON, \beta PL]$ is always understood as coreferential with the constituent noun phrase which has the features $[+N, +PRO, +1, \beta PL]$, and the matrix object noun phrase

with the features $[+N, \pm PRO, \alpha PERSON, \beta PL]$ is always understood as coreferential with the constituent noun phrase which has the features $[+N, +PRO, +2, \beta PL]$. Although the notions of matrix subject and object are relevant in establishing coreferential relations in direct discourse sentences, the notions of constituent subject and object are not relevant. Thus we see in (10) the matrix subject noun phrase is coreferential with the constituent noun phrase which has the features $[...+1,...]$, and the matrix object is coreferential with the constituent noun phrase which has the features $[...+2,...]$.

- (10) a. $h\bar{a}s\bar{a}n_c$ $ketak\bar{i}ke_a$ $ballo$, 'tumi $\bar{a}m\bar{a}ke_c$ $bh\bar{a}lab\bar{a}so!$: Hasan told Ketaki, 'You love me.'
- b. $chele\bar{t}i_c$ $meye\bar{t}ike_a$ $ballo$, 'tumi $\bar{a}m\bar{a}ke_c$ $j\bar{a}no$ $n\bar{a}!$: The boy told the girl, 'You do not know me.'

We come back to the examples (9f, g, h). These are similar to (8d, e, f). In (9f) both the matrix and the constituent subject noun phrases have the features $[+N, -PRO, ...]$ and they are lexically identical. Similarly the matrix and constituent object noun phrases have the features $[+N, -PRO, ...]$ and they are lexically identical. As this is a direct discourse sentence the lexically identical matrix and constituent noun phrases cannot be understood as coreferential. Lexically identical matrix and constituent noun phrases can be understood as coreferential if they are pronouns. In (9g, h) the noun phrases with identical indices are noncoreferential as they both have the features $[+N, -PRO]$ and are lexically identical.

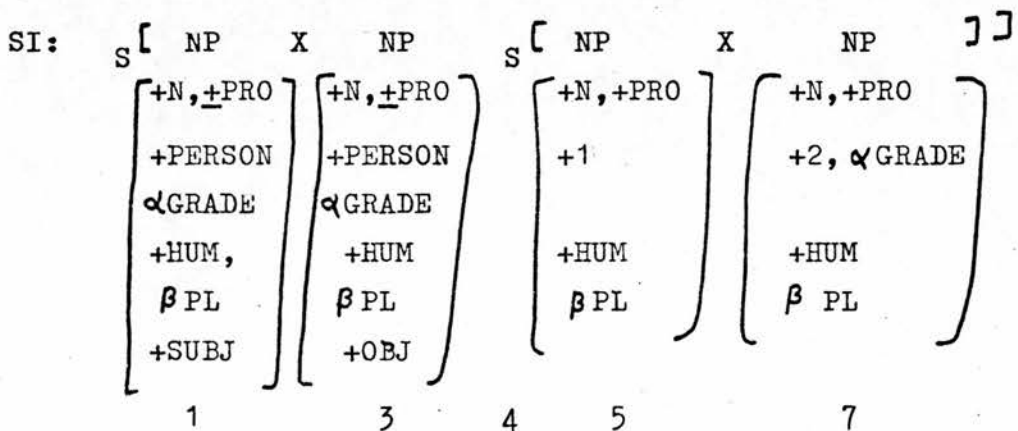
Our main claim is that the constituent noun phrases in direct discourse sentences must be pronouns ($[+N, +PRO]$) in order to be understood coreferentially with matrix subject

or object noun phrases. But we are aware of situations where the constituent noun phrase in a direct discourse structure can be a full noun phrase and can be understood coreferentially with a matrix subject noun phrase. Extralinguistic knowledge plays an important part in such coreference. Consider (11).

- (11) a. $\underline{d\bar{a}li}_c$ ghoṣaṇā kareche, 'dāli_csarbakāler 'sreṣṭha 'silpī':
Dali has declared, 'Dali is the greatest painter of all times.'
- b. $\underline{ketak\bar{i}}_c$ likheche, 'ketakīke_cbhulo nā': Ketaki has written, 'Do not forget Ketaki.'

In (11) dāli and ketakī in the constituent sentences are not pronouns, but they can be understood as coreferential with their respective matrix subject noun phrases by extralinguistic knowledge. The constituent noun phrases dāli and ketakīke in (11a, b) respectively should have been āmi: 'I', and āmāke: 'Me' in a natural sentence. We will not consider such stylistic coreference here, and will consider that in examples like (8d, e, f), (9f, g, h) and (11a, b), where the matrix and constituent noun phrases are nominals and lexically identical, they are noncoreferential. We can arrive at the rules given in (12) for direct discourse sentence structures.

(12) Rules of Coreference in Direct Discourse Structure.



- Rules: a. 1 and 5 are coreferential, and
 b. 3 and 7 are coreferential.

- Conditions: a. 1 and 3 are in a matrix sentence, and
 5 and 7 are in a constituent sentence.
 b. 4 is a direct discourse of the
 matrix subject.

5.2 Antecedent and Anaphora: Hierarchy of Antecedents.

The notions of antecedent and anaphora are important in pronominalization. We derive anaphoric pronouns by the application of a pronominalization rule which utilizes one noun phrase to reduce another into a pronoun or pronominal under certain conditions. The noun phrase which is used to reduce another noun phrase into a pronoun is the antecedent of the derived pronoun. We will use the symbol NP^a for the antecedent noun phrase, and NP^D for the pronominalizable noun phrase (cf., Langacker (1969)). Lakoff (1968) has set up a hierarchy of antecedents, which is given below.

(13) Hierarchy of Antecedents.

- a. Proper names.
- b. Definite descriptions.
- c. Epithets.
- d. Pronouns.

The hierarchy in (13) says that an NP class which is higher in the hierarchy may be the antecedent of an NP class which is lower in the hierarchy, but not vice versa. That is, a proper name can be the antecedent of a definite description and so forth, but a definite description or an epithet or a pronoun cannot be the antecedent of a proper name. Consider the examples below.

(14) a. jībanānanda_c elen o jībanānanda_c hāslen: Jibanananda
came and Jibanananda smiled.

b. jībanānanda_c elen o $\left\{ \begin{array}{l} \text{'rūpasī bāṅlā'r kabi} \\ \text{bāṅlār bisanṅatama kabi} \\ \text{tini} \end{array} \right\}$ hāslen:

Jibanananda came and $\left\{ \begin{array}{l} \text{the poet of 'Rupasi Bangla'}$
 $\left\{ \begin{array}{l} \text{the most melancholic poet of Bengal} \\ \text{he} \end{array} \right\}$

smiled.

c. * $\left\{ \begin{array}{l} \text{'rūpasī bāṅlā'r kabi} \\ \text{bāṅlār bisanṅatama kabi} \\ \text{tini} \end{array} \right\}$ c elen o jībanānanda_c

hāslen:

d. * $\left\{ \begin{array}{l} \text{bāṅlār bisanṅatama kabi} \\ \text{tini} \end{array} \right\}$ c elen o 'rūpasī bāṅlā'r

kabi_c hāslen:

e. 'rūpasī bāṅlā'r kabi_c elen o $\left\{ \begin{array}{l} \text{bāṅlār bisanṅatama kabi} \\ \text{tini} \end{array} \right\}$ c

hāslen:

f. bāṅlār bisanṅatama_c kabi elen o tini_c hāslen: The most
melancholic poet of Bengal came and he smiled.

g. *tini_c elen o bāṅlār bisanṅatama kabi_c hāslen: He came
and the most melancholic poet of Bengal smiled.

Each of the sentences in (14) is a coordinate conjoined structure,
where the antecedent must be in the left conjunct and the anaphoric

noun phrase or pronoun in the right conjunct. The above examples show that the hierarchy in (13) is correct. For example, (14a) has two coreferential noun phrases, which are proper nouns, and so the question of antecedent and anaphora does not arise here. In (14b) left noun phrase is a proper name, and the right noun phrase is a definite description, an epithet and a pronoun. They refer to the antecedent in the first conjunct. The situation is the reverse in (14c), and thus the indexed noun phrases are not understood coreferentially. (14d) shows that an epithet or a pronoun cannot be the antecedent of a definite description. (14e) shows that a definite description can be the antecedent of an epithet or pronoun. (14g) shows that a pronoun cannot be the antecedent of an epithet, but that the reverse is possible can be seen in (14f).

5.3 Application of the Pronominalization Proper Rule.

The pronominalization proper transformation applies in nonsimplex, that is, complex and compound, structures. The rule requires for its application two coreferential noun phrases identical in head nouns, in a complex or conjoined structure. The coreferentiality of noun phrases will be marked, as above, by identical indices in the underlying structure, in the Chomskyan manner (cf., Chomsky (1965, 145-46)). But we differ somewhat from Chomsky in this matter, in that Chomsky marks head nouns as coreferential; we shall mark entire noun phrases as coreferential. The widely used convention is that if two noun phrases bear the same index in the underlying structure they are understood as coreferential. Although we are aware of the problems of the indexing approach we are in no position

to avoid it. One way to avoid indexing the noun phrases would have been the application of the rules of coreference worked out, in earlier sections, in the underlying structure of sentences. But we shall not avoid the indexing method, as it is helpful for practical purpose.

Now let us consider how the rule applies in a coordinate conjoined structure. Consider the examples below.

- (15) a. bhadrālok dujan_c elen o bhadrālok dujan_c ballen: The
 1 2 3 4 5 6 7
 two gentlemen came and the two gentlemen said.
 2 1 3 4 6 5 7
- b. bhadrālok dujan_c elen o tārā_c ballen: The two
 1 2 3 4 5 6 2
 gentlemen came and they said.
 1 3 4 5 6
- c. *tārā_c elen o bhadrālok dujan ballen: They came and
 1 2 3 4 5 6 1 2 3
 the two gentlemen said.
 5 4 6

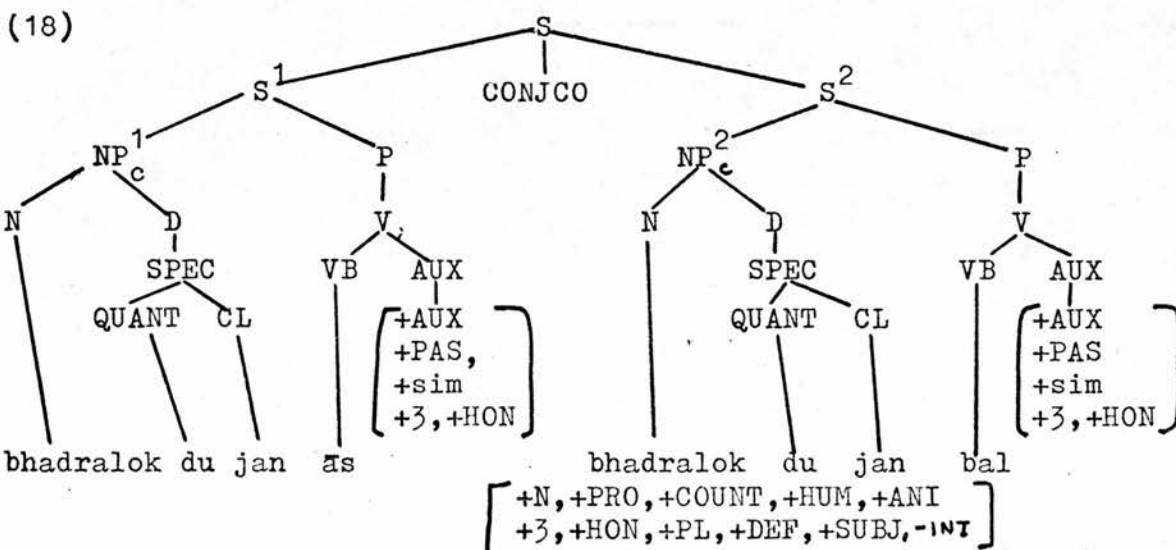
(15a) is a coordinate conjoined structure with two coreferential noun phrases and they have identical head nouns. As these noun phrases are coreferential, identical in head noun and synonymous, the pronominalization rule can apply in (15a). As (15a) is a coordinate conjoined structure, the rule will apply forwards. Almost all the transformational grammarians claim that the pronominalization rule applies forwards in a coordinate conjoined structure, and cannot apply backwards in such a structure (cf., Ross (1967b), Langacker (1969), Stockwell et al (1973)).¹

¹ Recently Hinds and Okada (1975) have claimed that pronominalization can apply backwards in a coordinate conjoined structure if the pronoun is unstressed. Like Langacker (1969) and others we will not take stress into consideration, and will study its application in neutral environments. (15c) is unacceptable in a normal neutral environment.

(16) is a coordinate conjoined structure where pronominalization can apply, as it satisfies the following conditions:

- (17) a. It has two noun phrases which are coreferential and identical in head nouns (NP^1 in S^1 and NP^2 in S^2).
- b. NP^2 is [+DEF].

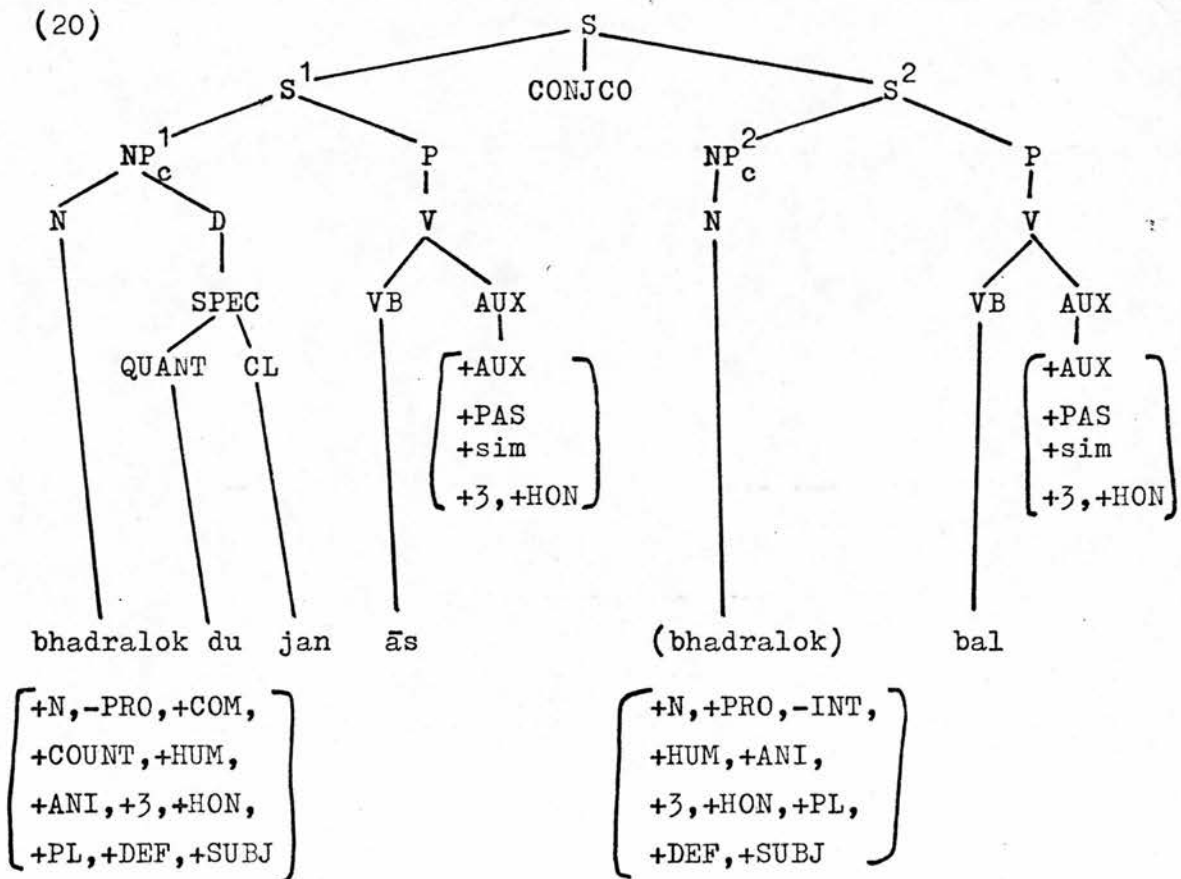
As (16) is a coordinate conjoined structure, the pronominalization rule will apply forwards. The rule will use NP^1 as NP^a (antecedent) and NP^2 as NP^D (pronominalizable noun phrase). The rule will utilize NP^1 to specify NP^2 as [+PRO,-INT] due to its coreferentiality and identity (in head noun) with NP^1 . In (16) we have gathered the features of the noun phrases under the head nouns, but they are the features of the entire noun phrase. In (16) the head nouns of the noun phrases involved have the feature [+HUM]. The pronominalization rule will change the feature [-PRO] of NP^2 into [+PRO,-INT], and will delete [+COM] feature of the pronominalized noun phrase. The features [+PRO,-INT] will be specified on the complex symbol associated with the head noun of the pronominalized noun phrase; but it will be understood that the entire noun phrase has been pronominalized. In this way, the pronominalization rule will transform (16) into (18).



The pronominalization rule has applied in (18) only partially, in that it has specified NP² as [+PRO, -INT]; but some more operations will take place in (18). This rule, besides specifying NP² as [+PRO, -INT], will delete all the constituents of the D, except PL, if any, of the pronominalized noun phrase. The deletion of the constituents of D is necessary because the Bengali pronouns cannot occur with PART, DEIC and QUANT CL. The following noun phrases are ungrammatical because of this reason:

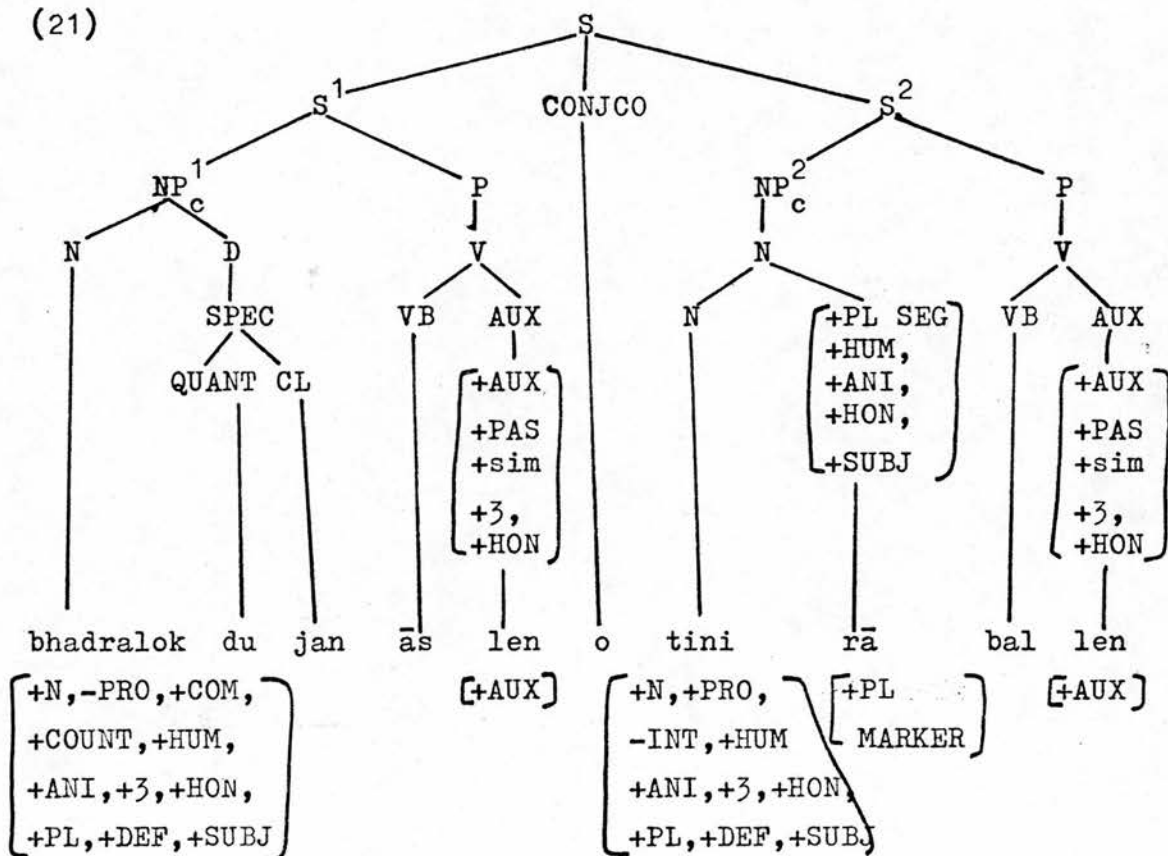
- (19) a. *dujan tārā: *Two they.
 b. *ekjan tini: *A he.
 c. *dujan tāder madhye ekjan se: *One he among two them.

By the deletion of the D of the pronominalized noun phrase (NP²) in (18) we will get (20).



NP² in (20) has been pronominalized, but the first lexical item bhadralok is still on this node. This item will be replaced by an appropriate pronoun by the second lexical attachment rule after the last rule of the transformational component. This noun phrase has the feature [+PL], and so a plural segment transformation will take place in this noun phrase. The plural segment transformation rule will create a node [+PL SEG] and will copy the features [α HUM, β ANI, γ HON, π SUBJ] of the head noun of NP² onto the plural segment. The plural segment will be adjoined as the right daughter of the N by Chomsky-adjunction (cf., § 2.4.4; § 2.5). The plural segment transformation will transform (20) into (21).

(21)



All relevant transformations have applied in (21). Thus we have arrived at the post-transformational string in (21), where the second lexical insertion rule can apply. In (21) the second

lexical insertion rule has replaced the first lexical item bhadralok attached on the pronominalized noun phrase by the third person honorific pronoun tini, as its features match with those associated with the head noun of NP². The second lexical insertion rule has inserted len to the AUX's, the conjunction o to the CONJCO, and the plural marker rā to the plural segment in (21) by the matching condition of second lexical insertion (cf., @ 11.2). The terminal string of (21) will generate the syntactic surface structure (22).

(22) $S [S [[[\neq \text{bhadralok} \neq] [\neq \text{du+jan} \neq]] [\neq \text{ās+len} \neq]] [\neq \text{o} \neq]$
 $S [[\neq \text{tini+rā} \neq] [\neq \text{bal+len} \neq]]]_S$

The surface structure string (22) will pass through the morphophonological component and the phonetic structure ≠ bhadralok dujan elen o tārā ballen ≠, which is (15b), will be derived. Thus we see that the pronominalization rule applies on the intermediate structure of (15a) forwards, and derives (15b). The rule cannot apply here backwards. If the rule applies backwards, we will get the ungrammatical sentence (15c).

Now consider whether pronominalization is an obligatory or optional rule in Bengali. Consider the following examples.

(23) a. ekṭi meye_c nāclo o meyeṭi_c gān gāilo: A girl danced
 1 2 3 4 5 6 7 1 2 3
 and the girl sang.
 4 5 7

b. ekṭi meye_c nāclo o se_c gān gāilo: A girl danced and
 she sang.

(24) a. yadi minu_c āse, tabe minu_c nācbe: If Minu comes,
 then Minu will dance.

b. yadi minu_c āse, tabe se_c nācbe: If Minu comes, then
 she will dance.

- c. $yadi\ se_{\text{c}}\bar{a}se,$ $tabe\ minu_{\text{c}}\bar{n}acbe:$ If she comes, then
Minu will dance.
- (25) a. $minu_{\text{c}}\text{ballo},$ $'\bar{a}mi_{\text{c}}\text{phul}\ \bar{b}h\bar{a}lab\bar{a}si':$ Minu said,
'I love flowers.'
- b. $minu_{\text{c}}\text{ballo}\ ye\ se_{\text{c}}\text{phul}\ \bar{b}h\bar{a}lab\bar{a}se:$ Minu said that she
loved flowers.
- c. $*minu_{\text{c}}\text{ballo}\ ye\ \bar{a}mi_{\text{c}}\text{phul}\ \bar{b}h\bar{a}lab\bar{a}si:$ Minu said that
I loved flowers.

(23a) is a coordinate conjoined structure with two coreferential noun phrases where pronominalization can take place. If the rule applies here we get (23b), and if it does not apply we get (23a). Both (23a, b) are grammatical. The difference between them is that (23a) is elaborate and contains the repetition of the same lexical items, but (23b) is economical and devoid of any repetition. (23b) will be preferred by most of the speakers. But as both (23a, b) are grammatical, we can say that pronominalization is an optional but preferred rule in Bengali. (24a) is a subordinate conjoined structure which satisfies the structure index for the pronominalization rule; if the rule applies forwards we get (24b), and if backwards we get (24c). All the sentences in (24) are grammatical, but (24b, c) will be preferred by most of the speakers for economy. These also show that pronominalization is an optional but preferred rule in Bengali. (25a) is a direct discourse complex structure, where the matrix subject noun phrase and the constituent subject, which is a pronoun, are coreferential. By the application of the indirect discourse formation rule in (25a) we derive (25b). This rule makes the constituent subject symmetrical in person and grade features to its antecedent. (25c) is ungrammatical

because the indirect discourse formation rule has applied here, but the constituent subject has not been made symmetrical to its antecedent in person and grade features. Thus we can arrive at the following conclusions:

- (26) a. Pronominalization proper is an optional rule in all the Bengali sentence structures except in indirect discourse formation from direct discourse complex structures.
- b. Pronominalization rule makes sentences economical. Accordingly this rule is preferred by most speakers.

5.3.1 Application of the Pronominalization Proper Rule in [-HUM] and [-ANI] Noun Phrases.

We have seen that Bengali has no nonhuman and/or inanimate pronoun, and for this reason Bengali uses deictics in order to refer to nonhuman and inanimate objects (cf., § 4.1.5). Consider the following examples.

- (27) a. āmi e-ṭi kinbo: I shall buy this.
- b. āmi e-duṭi kinbo: I shall buy these two.

In (27a) e-ṭi refers to some object deictically. e-ṭi is not a pronoun, but the residue of a noun phrase of the structure DEIC N SPEC from which the head noun has been deleted because the noun is contextually understood. That the elements e-ṭi: 'This', e-gulo: 'These' etc., are not pronouns can be seen in the residual noun phrase e-duṭi: 'These two' in (27b). Bengali pronouns are either singular or plural, they cannot indicate the numerical number or quantity of the persons or things they refer to. The residual noun phrase e-duṭi deictically refers to two objects, which are understood contextually. We have discussed

the similarity and dissimilarity of the deictics and the pronouns elsewhere (cf., § 4.1.5).

As Bengali has no nonhuman and inanimate pronoun, the pronominalization rule cannot reduce [-HUM] and [-ANI] noun phrases into pronouns when they are in a structure which satisfies all conditions for the pronominalization rule. In such a situation the pronominalization transformation prefers either of the two following operations depending on the structure of the pronominalizable noun phrase: (a) the pronominalization rule utilizes the NP^a to delete the NP^p if the NP^p has the intermediate structure N SPEC (that is, the underlying structure DEF SPEC N); (b) the rule utilizes the head noun of the NP^a to delete the head noun of the NP^p if the NP^p has the intermediate structure DEIC N SPEC (that is, the underlying structure DEIC SPEC N). In this section we will show how these operations take place in Bengali. Consider the examples below.

- (28) a. mātin ekṭi chabi_c āklo o chabiṭi_c āmāke dekhālo: Matin
 2 3 4 5 6 7 8
 drew a picture and showed the picture to me.
 4 2 3 5 8 6 7
- b. mātin ekṭi chabi_c āklo o āmāke dekhālo: Matin drew
 2 3 4 5 6 7 4
 a picture and showed me.
 2 3 5 7 6
- c. *mātin ekṭi_c chabi āklo o ṭi_c āmāke dekhālo:

In the coordinate conjoined structure (28a) the coreferential noun phrases have the feature [-ANI]. This is a structure in which pronominalization can apply. But we know that Bengali has no [-ANI] pronoun, and for this reason the pronominalization rule cannot reduce the forward noun phrase into a pronoun. In (28a) the forward noun phrase chabiṭi is the NP^p. It has the structure

N SPEC at the point the pronominalization rule applies. The pronominalization rule will apply here and delete the forward noun phrase entirely. In this way we get (28b) from (28a). As the pronominalizable noun phrase in (28a) has the structure N SPEC, the rule cannot delete its head noun under lexical identity with its antecedent. If we do so, we will get the ungrammatical structure (28c).

Now consider the operation of head noun deletion by the pronominalization rule involving [-HUM] and [-ANI] nouns. This is an operation frequently used in Bengali, and it leaves the DEIC SPEC of a noun phrase behind, due to the contextual and anaphoric deletion of head nouns. In the surface structure the residual DEIC SPEC structure looks like a pronoun. Consider the examples below.

- (29) a. mātin ekṭi pākhī_c dharlo o se-pākhīṭike_c khācāy rākhlo:
 2 3 4 5 6 7 8 9
 Matin caught a bird and kept that bird in a cage.
 4 2 3 5 9 6 7 8
- b. mātin ekṭi pākhī_c dharlo o se-ṭike_c khācāy rākhlo:
 2 3 4 5 6 7 8
 Matin caught a bird and kept that in a cage.
 4 2 3 5 8 6 7

(29a) is not an elaborate structure for we have already deleted the subject noun phrase from the forward conjunct because it is coreferential with the subject of the backward conjunct. (29a) is sufficient for our purpose because we are interested in the application of the pronominalization rule involving [-HUM] noun phrases.³

³ We have said that the pronouns āmi, āpani, tumi, tui, tini and se are human pronouns because they are used to refer to human beings. But, besides fairy tales, in some situations the third

In the coordinate conjoined structure (29a) the noun phrase ekṭi pākhī in the left conjunct and the noun phrase se-pākhīṭi in the right conjunct are coreferential and identical in head noun. This structure satisfies all the conditions for the pronominalization rule, which will apply here forwards. But we know that these noun phrases are [-HUM], and as Bengali has no nonhuman pronoun we cannot reduce the right noun phrase into a pronoun by the pronominalization rule. However, the forward noun phrase in (29a) has the structure DEIC N SPEC. We can use the left noun phrase to delete the head noun of the right noun phrase by the pronominalization rule, and derive (29b). Here we will

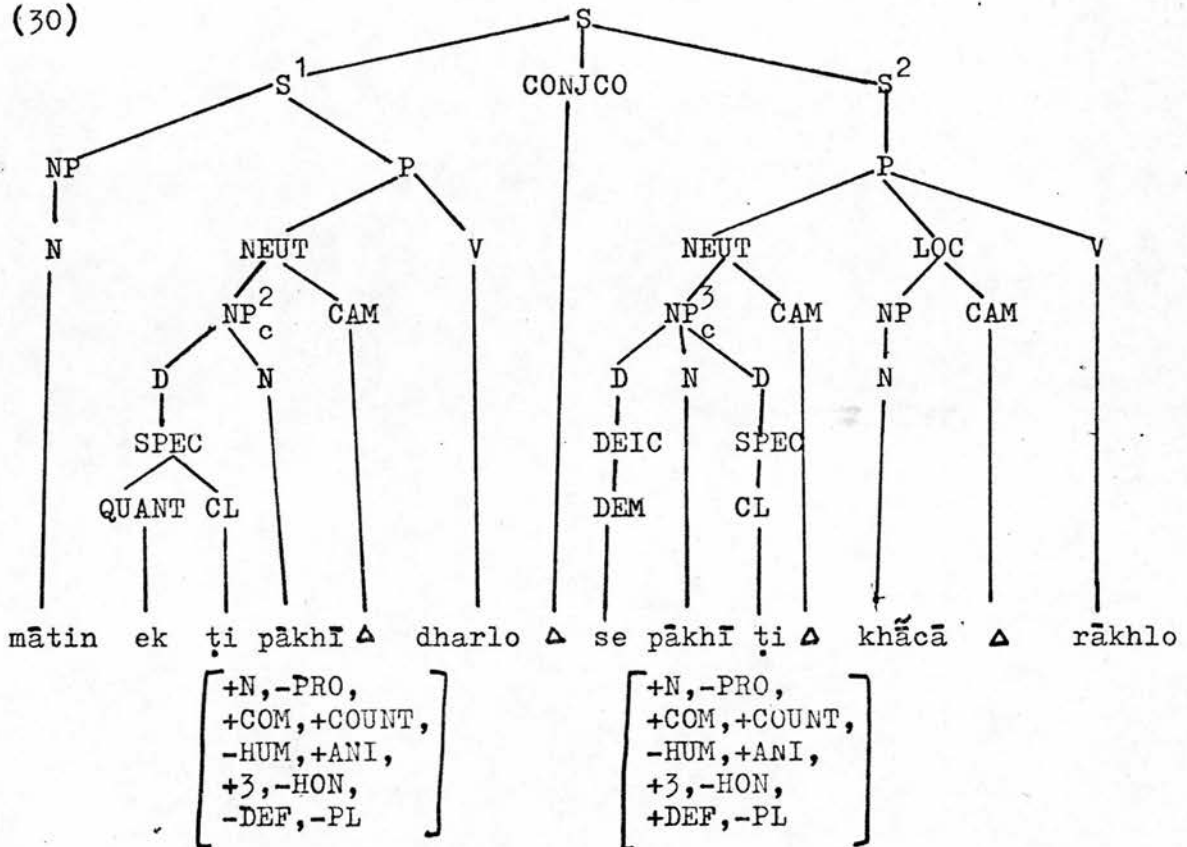
person nonhonorific pronoun can be used to refer to [-HUM] beings, especially if the being referred to is a dear or lovely one and is not the subject of the sentence. Consider the sentences below.

- (A) a. mātinēr ekṭi kukur_cāche. {^{*se}se-ṭi_c} rātbhar gheu-gheu
 kare: Matin has a dog. {^{He}That (one)} barks all night.
- b. mātinēr ekṭi kukur_cāche. mātin tāke_c khub bhālabāse:
 Matin has a dog. Matin loves him very much.

In (Aa) se, which is the subject, will refer to mātin, but not to the dog, and se-ṭi will refer to the dog. In (Ab) tāke refers to the dog which Matin owns. In place of tāke we could have used se-ṭike (\Leftarrow se-kukurṭike), which is the residue of a noun phrase with the structure DEIC N SPEC. The third person nonhonorific pronoun is not usually used to refer to a nonhuman being when it is singular and the subject of a sentence, but its derivative plural form tārā, objective forms such as tāke, tāder, and the genitive forms tār and tāder are often used to refer to nonhuman beings.

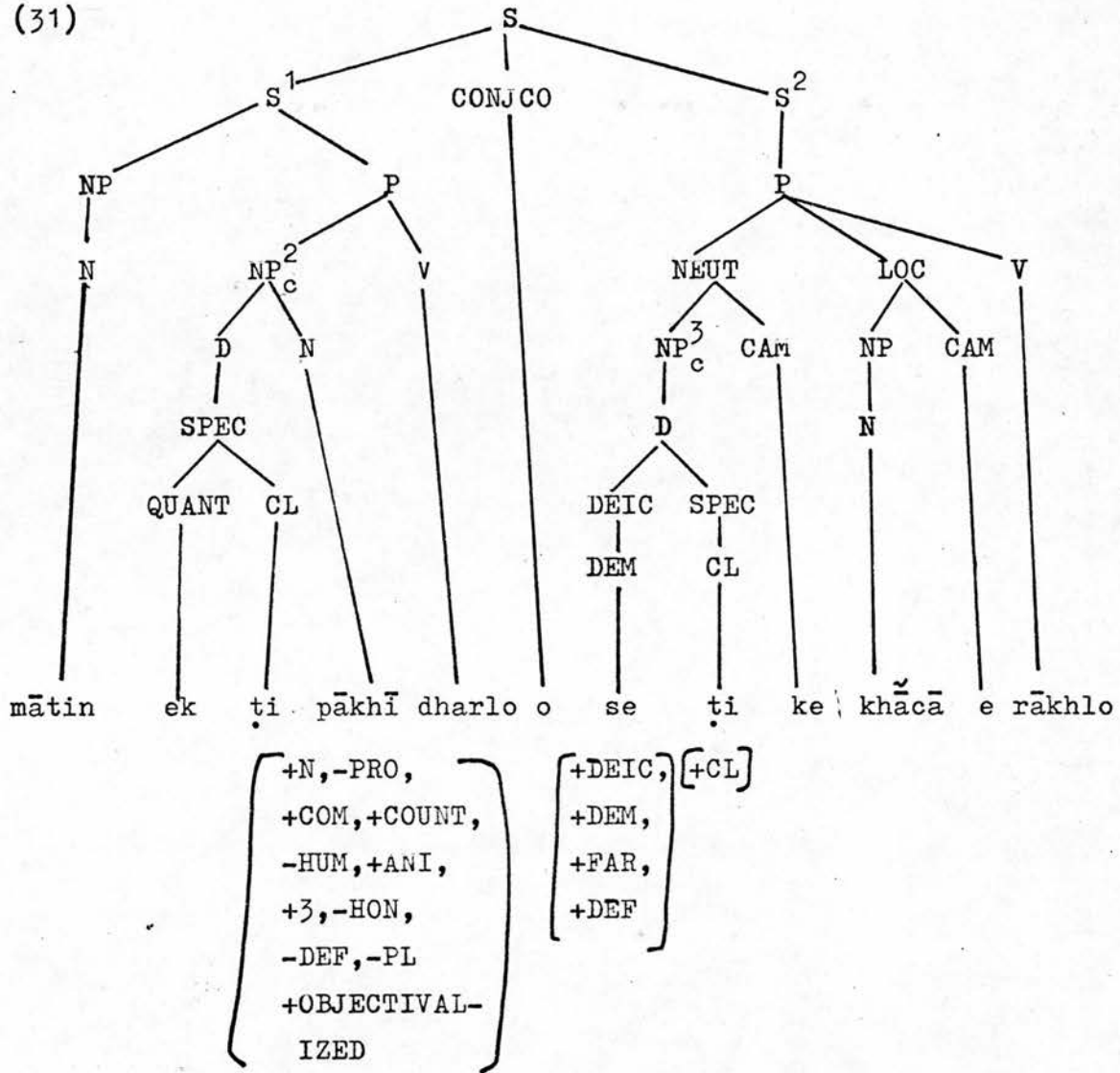
show how (29b) is derived from from an intermediate structure of (29a). (29a) has the intermediate structure (30).

(30)



In (30) NP^2 and NP^3 are coreferential and identical in head noun. These noun phrases are $[-HUM, +ANI]$. As (30) is a coordinate conjoined structure, the pronominalization rule will apply here forwards. But the pronominalization rule cannot reduce the NP^D (NP^3) in (30) into a pronoun. As the NP^D has the structure DEIC N SPEC the pronominalization rule will use the antecedent noun phrase to delete the head noun of the pronominalizable noun phrase. That is, NP^2 will be used to delete the head noun of NP^3 in (30). After the last rule of the transformational component (30) will be transformed into the derived structure (31), where the second lexical insertion rule will attach the appropriate lexical items to the derived nodes.

(31)



The terminal string of (31) will pass through the morphophonological component and will generate the phonetic structure

mātin ekṭi pākhī dharlo o se-ṭi ke khācāy rākhlo, which is (29b).

The residual noun phrase se-ṭi in (29b), which refers anaphorically to its antecedent ekṭi pākhī, is simply a combination of the deictic se(i) and the classifier ṭi. It is not a pronoun, but a pronominal. It is a residual noun phrase whose head noun has been deleted under coreference.

5.4 Forward and Backward Pronominalization.

The terms forward and backward pronominalization are quite common in the transformational literature dealing with pronominalization. These terms are self-explanatory in that when a pronominalization rule applies forwards (that is, when the antecedent precedes or comes first) it is called forward pronominalization; and when a pronominalization rule applies backwards (that is, when the anaphora precedes or comes first) it is called backward pronominalization. For example, (24b) is an instance of forward pronominalization; and (24c) is an instance of backward pronominalization. Forward pronominalization is more natural than backward pronominalization, and we assume that forward pronominalization is permissible in all Bengali structure. But backward pronominalization is not permissible in some sentence structures in Bengali. Backward pronominalization has more constraints than forward pronominalization. Forward pronominalization can be considered as an instance of unmarked pronominalization, and backward pronominalization as marked pronominalization. Consider the examples below.

- (32) a. hāsān_c tākā dhār kare, kintu se_c kakhano tā 'sodh kare nā:
 2 3 4 5 6 7 8 9
 Hasan borrows money, but he never pays it back.
 3 2 4 5 6,9 8 7
- b. *se_c tākā dhār kare, kintu hāsān_c kakhano tā 'sodh kare nā:
 1 2 3 4 6 7 8 9
 He borrows money, but Hasan never pays it back.
 1 3 2 4 6,9 8 7
- (33) a. hāsān_c ketakīke bhālabāse ār se_c nīlimāke ghrinā kare:
 Hasan loves Ketaki and he hates Nilima.
- b. *se_c ketakīke bhālabāse ār hāsān_c nīlimāke ghrinā kare:
 He loves Ketaki and Hasan hates Nilima.

- (34) a. yadi minu_cnāce, tabe se_cpuraskār pābe: If Minu dances,
then she will get (a) prize.
- b. yadi se_cnāce, tabe minu_cpuraskār pābe: If she dances,
then Minu will get (a) prize.
- (35) a. hāsān_cye-meyeṭike cumu kheyechē, se-meyeti tāke_c
ghrinā kare: The girl whom Hasan has kissed hates him.
- b. se_cye-meyeṭike cumu kheyechē, se-meyeṭi hāsānke_c
ghrinā kare: The girl whom he has kissed hates Hasan.
- (36) a. meyeṭi_cbhābe ye se_csundar: The girl thinks that
she is beautiful.
- b. *se_cbhābe ye meyeṭi_csundar: She thinks that the
girl is beautiful.

The sentences in (32, 33) are coordinate conjoined structures. The grammaticality of (32, 33a) shows that pronominalization can apply forwards in a coordinate conjoined structure, and the ungrammaticality of (32, 33b) shows that pronominalization cannot apply backwards in a coordinate conjoined structure. The examples of (34) are of subordinate conjoined structure, and they show that pronominalization is permissible both backwards and forwards in a subordinate conjoined structure. The examples in (35) are complex structures with restrictive relative clauses. In such a structure pronominalization can apply both forwards and backwards as can be seen in (35a, b). (36a) is an indirect discourse structure derived from a direct discourse structure by pronoun symmetry. In such a structure the pronominalization rule always applies forwards, but never backwards, as the coreferential noun phrase in the complement clause is an underlying pronoun, which has no power to reduce a full noun phrase into a pronoun. Thus we see that forward pronominalization is

permissible in all Bengali structures, but backward pronominalization is permissible only in a subordinate conjoined structure and in a complex structure with a relative (subordinate) clause.

5.4.1 Constraints on Pronominalization.

We have seen in the last section that pronominalization cannot apply backwards in some sentence structures in Bengali. In this section we will try to discover the directionality constraints on pronominalization. Langacker (1969) tried to determine the constraints on pronominalization and developed the notions of 'Command' and 'Precedes', which he put together as 'Primacy Relations'. Langacker's constraints on pronominalization are reproduced below (Langacker (1969, 167)):

NP^a may be used to pronominalize NP^D unless (1) NP^D precedes NP^a ; and (2) either (a) NP^D commands NP^a , or NP^a and NP^D are elements of separate conjoined structures.

Langacker's notions are based on a particular tree-structure. Of these 'Precedes' is simpler: an element X precedes another element Y if X is at the left side of Y in a tree-structure. This shows that 'Precedes' is based on a linear order of elements. But the notion 'Command' is based on a particular type of tree-structure, and dominance relations. It becomes useless if a different type of tree-structure is proposed.

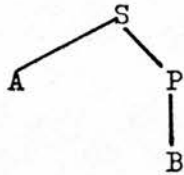
Langacker (1969, 167) defines 'Command' as follows:

.. a node A "commands" another node B if (1) neither A nor B dominates the other; and (2) the S-node that immediately dominates A also dominates B.

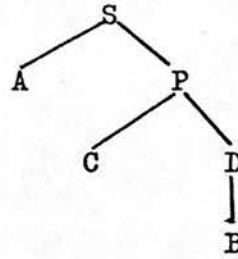
In each of the following trees in (37), the node A commands

the node B, but B does not command A.

(37) a.



b.



Langacker's constraints say that A can be used to pronominalize B, but B cannot be used to pronominalize A in each structure in (37).

Here we will examine Langacker's 'Command' and 'Precedes' notions and see whether they can be included in our grammar for stating constraints on pronominalization. To begin with consider the following simplex structures.

(38) a. $\text{minu}_c \text{minuke}_c \text{bhālabāse}$: Minu loves Minu

b. $\text{minu}_c \text{nijke}_c \text{bhālabāse}$: Minu loves herself.

c. $*\text{nij}_c \text{minuke}_c \text{bhālabāse}$: *Herself loves Minu.

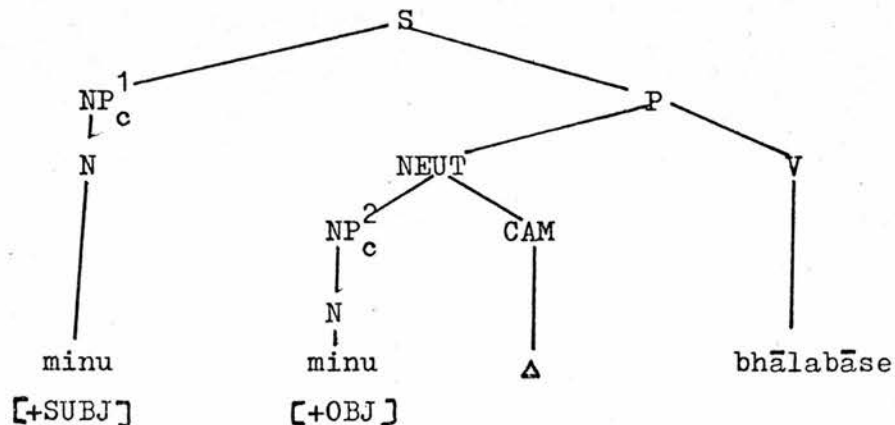
(39) a. $\text{minuke}_c \text{minu}_c \text{bhālabāse}$: Minu loves Minu.

b. $\text{nijke}_c \text{minu}_c \text{bhālabāse}$: Minu loves herself.

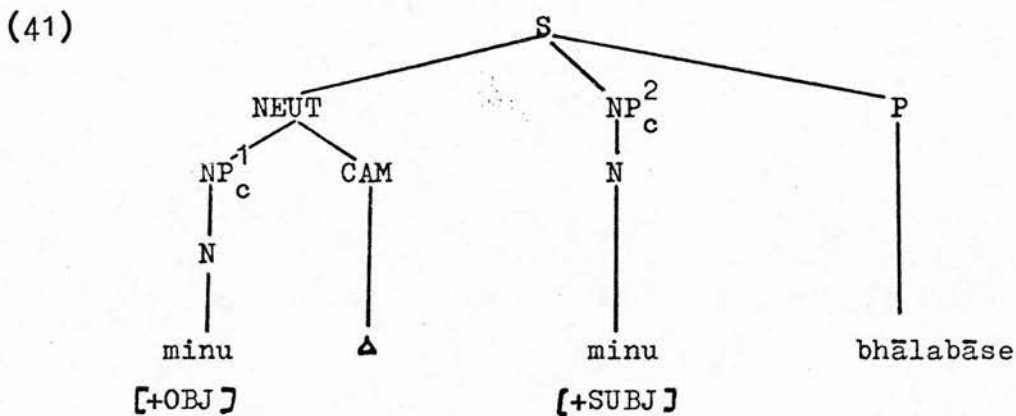
c. $*\text{minuke}_c \text{nij}_c \text{bhalabase}$: *Herself loves Minu.

(38a) is a simplex structure with two coreferential noun phrases, of which the left noun phrase is the subject and the right one is the object. The structure that immediately underlies (38a) is given in (40).

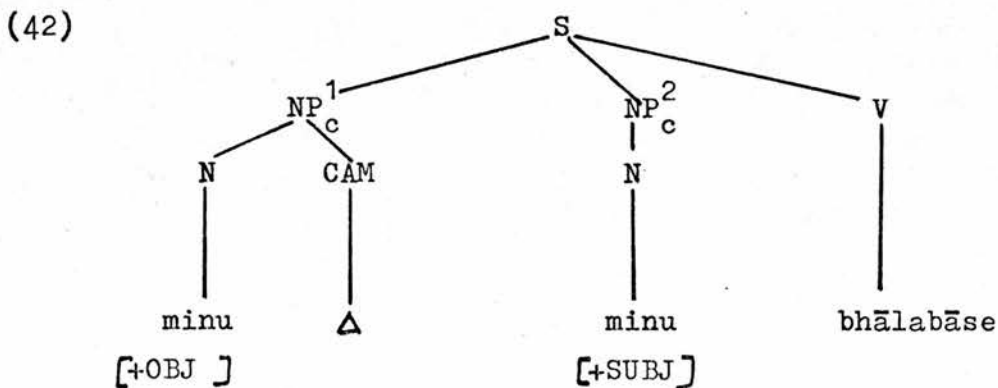
(40)



In (40) NP¹ both precedes and commands NP², and therefore NP¹ can be used to reflexivize NP², but not vice versa. In this way Langacker's constraints allow the derivation (38b) from (40), and blocks the derivation of (38c) from (40). This shows that his constraints give the correct result. Now consider (39a), which is synonymous with (38a) but differs in constituent order in that the object noun phrase precedes the subject noun phrase in (39a). We assume that (39b) is derived from (39a), which has the immediate underlying structure (41).



(41) is a simplex structure with two coreferential noun phrases. The reflexive rule can apply here. In (41) NP¹ precedes but does not command NP². But Langacker does not work with this type of tree-structure: in his constituent structure model NP¹ will be immediately dominated by the S-node. In his framework (41) will be (42), where NP¹ is immediately dominated by the S-node.

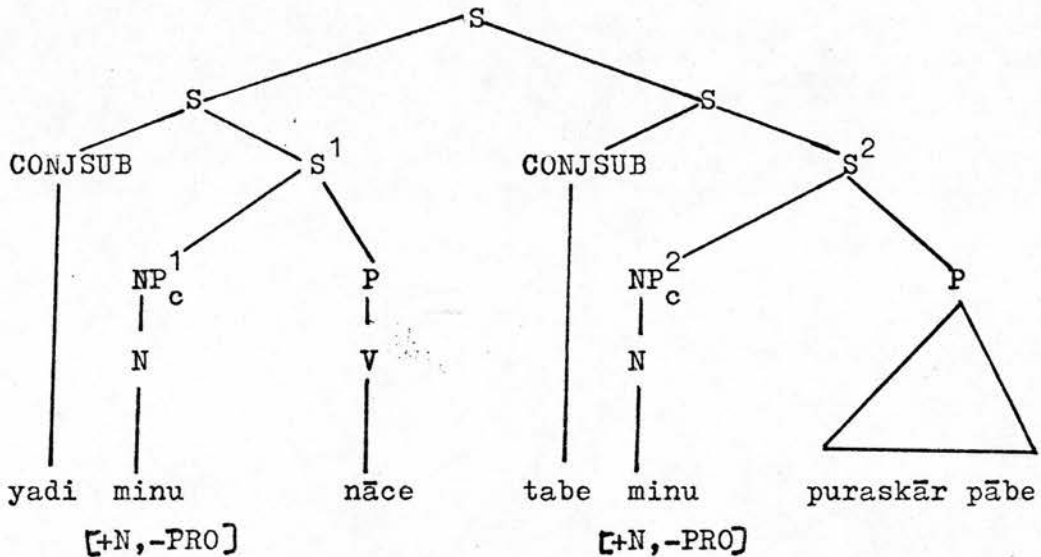


In (42) NP¹ both precedes and commands NP². So Langacker's constraints will allow NP¹, which is the object, to reflexivize NP², which is the subject, and generate the ungrammatical sentence (39c); and will block the grammatical sentence (39b). It might be argued that sentences like (39b) are derived from the underlying structure of (38a) by constituent reordering after the application of the reflexive rule. This argument is absurd because if movement of constituents is permissible after reflexivization, it is also permissible prior to the reflexive rule. Langacker's 'Command' notion is developed with a view to explaining languages like English, which have relatively fixed constituent order. But it seems that languages like Bengali, which allow a relatively free order of constituents, cannot be explained completely satisfactorily with 'Command'. We shall accordingly abandon the notion of 'Command', but retain the notion 'Precedes' in order to explain linear order of elements.

Our question is how to account for the facts of the examples (38, 39) when the notion 'Command' is gone? The facts of these sentences can be explained if we accept the assistance of the functional notions of subject and object. We have seen that Bengali allows a relatively free order of constituents in a simple sentence structure (cf., § 3.3). In (38a) the subject noun phrase precedes the object noun phrase, and in (39a) the object noun phrase precedes the subject noun phrase; but it is always the subject noun phrase which reflexivizes the object noun phrase when they are coreferential, identical in head noun and in the same simplex structure. We see that the subject noun phrase reflexivizes the object noun phrase in (38a) and generates (38b); but the object noun phrase cannot reflexivize the subject

pronominalize NP^D unless NP^D precedes NP^A , is correct. This shows that 'Precedes' is relevant in pronominalization in a coordinate conjoined structure. Now consider the sentences in (34), which are subordinate conjoined structures. The sentences in (34) have the intermediate structure (44).

(44)

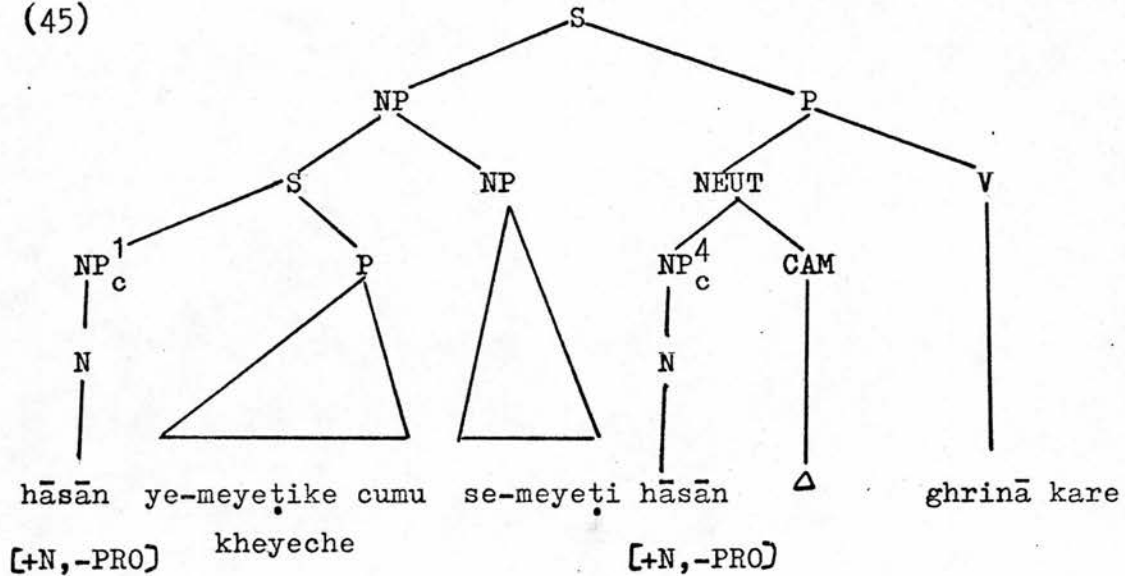


In (44) NP^1 and NP^2 are coreferential, and neither of them commands the other; but NP^1 precedes NP^2 . Langacker's constraints will allow only NP^1 to pronominalize NP^2 , but not vice versa.

We notice, however, that pronominalization can apply both forwards and backwards in (44). If the rule applies forwards in (44), we will derive (34a); and if applies backwards, we will derive (34b). We see that 'Precedes' fails to state the constraint on pronominalization in a subordinate conjoined structure. The facts can be stated simply by saying that pronominalization is permissible both forwards and backwards in a subordinate conjoined structure.

Now consider the examples in (35), which are instances of complex structures with restrictive relative clauses. These sentences have the intermediate structure (45).

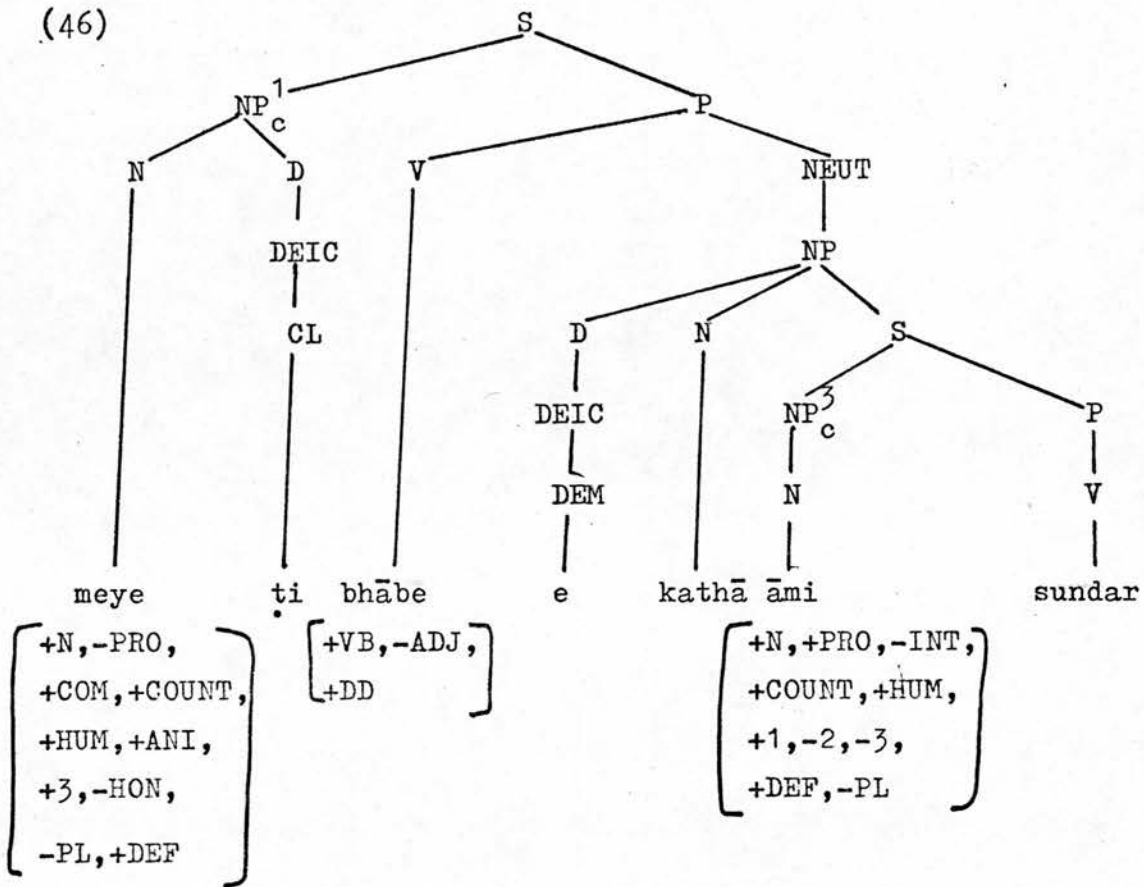
(45)



In (45) NP¹ (in the subordinate S) and NP⁴ (in the matrix S) are coreferential. Neither commands the other but NP¹ precedes NP⁴. Langacker's constraints will allow the application of the pronominalization rule only forwards; but the rule can apply here in both directions. If the rule applies forwards in (45) we derive (35a), and if applies backwards we derive (35b). Again we find that the constraints on pronominalization cannot be stated in terms of the 'Precedes' relation. We have seen above that backward pronominalization is possible in a subordinate conjoined structure, and here we see that it is possible in a complex structure when the pronominalizable noun phrase is in an subordinate clause.

Finally consider the sentences in (36). (36a) is an indirect discourse sentence derived from an underlying direct discourse sentence structure. Our rule of coreference in a direct discourse sentence structure says that any noun phrase in the constituent sentence which is coreferential with the matrix subject or object noun phrase must be a pronoun in the underlying structure (cf., § 5.1.3). (36a) has the intermediate

structure (46).



In (46) NP^1 and NP^3 are coreferential. NP^1 is a full noun phrase, but NP^3 is a pronoun. Our hypothesis is that a pronoun has no power to reduce a full noun phrase into a pronoun, and so NP^3 cannot reduce NP^1 into a pronoun. That is, pronominalization cannot apply backwards in a direct discourse structure. The rule, therefore, applies only forwards in such a structure. The pronominalization rule in collaboration with the indirect discourse formation rule will make NP^3 in (46) symmetrical to its antecedent in the features of person and grade (cf., § 5.8.1). The abstract noun phrase e-kathā in (46) will be deleted (cf., § 10.3). In this way we will derive (36a) from (46). Our principle that a pronoun cannot reduce a full noun phrase into a pronoun blocks the application of the pronominalization rule backwards in (46), and thus rules out (36b) as

ungrammatical.

We can arrive at the following constraints on pronominalization:

(47) Directional Constraints on Pronominalization Proper.

- a. Pronominalization rule applies forwards in all nonsimplex structures in Bengali,
- b. Pronominalization rule can apply backwards in a subordinate conjoined structure, or in a complex structure where the NP^D is an element of a subordinate clause.

5.5 Pronouns, Definite Descriptions and Epithets.

We have said before that not anything which can refer deictically or anaphorically is a pronoun (cf., § 4.1.5; § 4.3.1; § 5.2). Besides pronouns, full noun phrases can be used to refer anaphorically to some antecedent. They are broadly of two types: (a) Definite descriptions, and (b) Epithets. Jackendoff (1972) wants to deal with pronouns and epithets by the same rule. Our view is that although definite descriptions and epithets can occur in some subset of environments where pronominalization is possible, they should not be dealt with like pronouns because they are not pronouns. In the following sentences pronouns, definite descriptions and epithets occur in the same environments.

- (48) a. kerāmat ālī sāheb_c elen o $\left\{ \begin{array}{l} \text{pānjābi parihiṭa eknāyak} \\ \text{dayār sāgar} \\ \text{tini} \end{array} \right\}_c$
- hāslen: Mr Keramat Ali came and $\left\{ \begin{array}{l} \text{the dictator in} \\ \text{punjabi} \\ \text{the ocean of kindness} \\ \text{he} \end{array} \right\}$
- smiled.

b. * $\left. \begin{array}{l} \text{pānjābi parihita eknāyak} \\ \text{dayār sāgar} \\ \text{tini} \end{array} \right\} \text{c} \quad \text{elen o kerāmat}$

ālī sāheb_c hāslen:

In (48) we see that a definite description, an epithet and a pronoun can occur in the same environment, and they all cannot occur in another environment. Jackendoff (1972, 110) says:

These "pronominal epithets" can occur in some subset of the environments in which pronominalization is possible, and they function semantically more or less as specialized pronouns. We would obviously miss a generalization if we do not handle them by a rule of the same kind as pronominalization, hopefully a rule that could collapse with pronominalization.

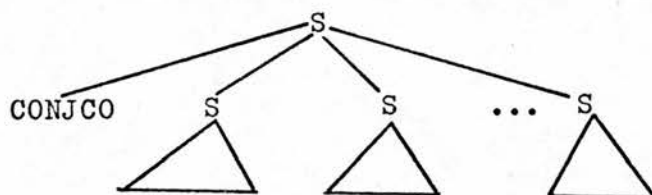
Our views differ from Jackendoff's by stating that definite descriptions and epithets, unlike pronouns, bear heavy semantic content. Definite descriptions and epithets are full noun phrases with heavy semantic content used anaphorically with some other noun phrase. Pronouns have no such semantic content of their own. That is why the sentences in (48a) are not synonymous. Jackendoff (1972, 111) wants to mark the epithets as 'special lexical items which may function as pronouns in certain contexts of the pronominalization rule, adding their lexical meaning to the intended attributes of the person they refer to.' We know that pronouns in any language must be limited in number, but there is no limit to number of definite descriptions and epithets. For example, the definite description and the epithet in (48a) could have been replaced by a hundred other definite descriptions

and epithets, but only one pronoun tini: 'He' can be used to refer to kerāmat ālī sāheb. Therefore, the pronominalization rule will not handle them. They are generated in the underlying structure and used coreferentially with some other noun phrase. But the pronominalization rule can apply when the coreferential noun phrases are synonymous. We will generate them in the underlying structure. The anaphoric pronouns are, however, derived transformationally, and inserted to a derived structure by the second lexical insertion rule.

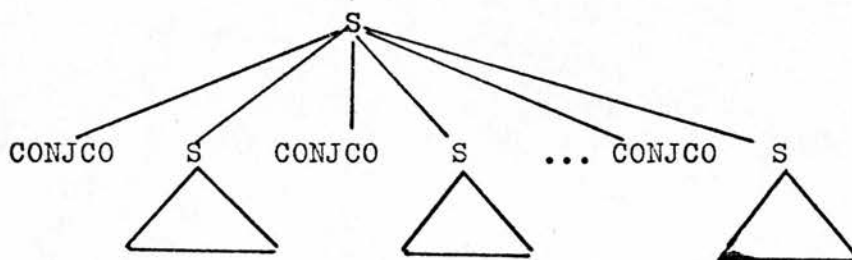
5.6 Pronominalization in Coordinate Conjoined Structure.

The base rule 1 of our grammar generates structures like (49a) and the conjunction spreading schema transforms them into structures like (49b)(cf., § 1.2.1: note to the base rule 1).

(49) a.



b.



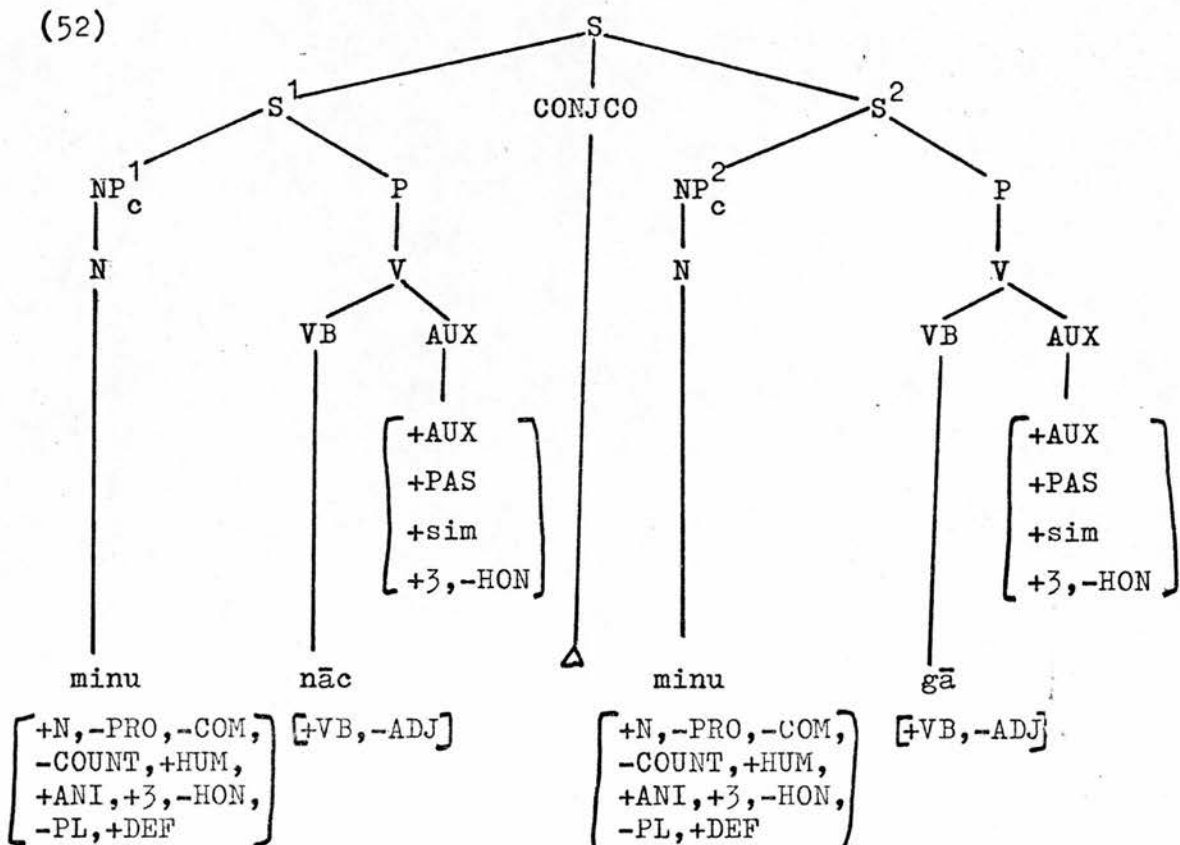
We have a rule which obligatorily deletes the initial conjunction, and we have another optional rule which deletes all but the last conjunction from a structure like (49b). We now come to the application of the pronominalization rule in such structures. Consider the examples below.

(50) a. minu_c nāclo o minu_c gāilo: Minu danced and Minu sang.
 2 3 5 2 3 5

b. minu_c nāclo o se_c gāilo: Minu danced and she sang.
 2 3 4 5 2 3 4 5

- (51) a. meyeṭi_c ṣahare gelo o meyeṭi_c gān gāilo o meyeṭi_c sukh
 1 2 3 4 5 6 7 8 9 10
 hārālo: The girl went to (a) town and the girl sang
 11 1 3 2 4 5 7
 songs and the girl lost happiness.
 6 8 9 11 10
- b. meyeṭi ṣahare gelo o se gan gailo o se sukh haralo:
 1 5 9
 The girl went to (a) town and she sang songs and she
 1 5 9
 lost happiness.

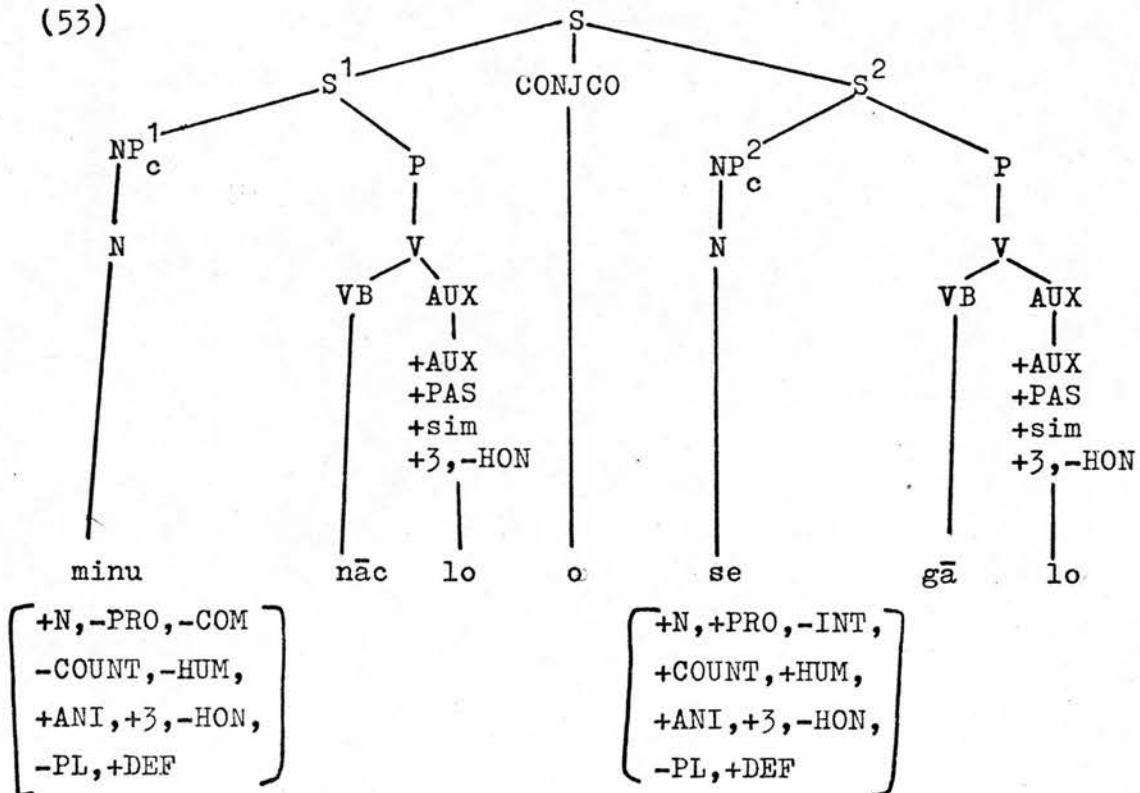
(50a) is a coordinate conjoined structure with two conjuncts and (51a) is one with three conjuncts. Our base rule 1 allows an indefinite number of sentences to be conjoined together; here we will deal with coordinate conjoined structures with a small number of conjuncts, which will serve our requirements. In (50a) each conjunct has a noun phrase which is coreferential and identical in head noun with a noun phrase in the other conjunct. This structure satisfies the conditions for the pronominalization rule. (50a) has the immediate underlying structure (52).



In (52) NP¹ and NP² are coreferential and identical in head noun. This structure satisfies all conditions for the pronominalization rule, which will apply here forwards.

As it is a coordinate conjoined structure, the pronominalization rule cannot apply here backwards. NP¹ in S¹ will be used as the antecedent to reduce NP² into a pronoun. The pronominalization rule will specify NP² as [+PRO,-INT] due to its coreferentiality and identity in head noun with NP¹. As NP² consists simply of a head noun, the pronominalization rule has no other operation to perform. The pronominalization rule will change the feature [-PRO] in the complex symbol associated with the head noun of NP² into [+PRO,-INT], and will change the [-COUNT] feature into [+COUNT]. This will transform (52) into (53).

(53)



In (53) all the relevant transformations have applied, and so we have arrived at the post-transformational string, where the second lexical insertion can take place. The second lexical

inserted rule has already attached the pronoun se to the pronominalized noun phrase by the matching condition of second lexical attachment. This rule has attached other appropriate lexical items to the derived nodes. The terminal string of (53) will now pass through the morphophonological component, and will generate the phonetic structure minu_c nāclo o se_c gāilo, which is (50b).

Now consider the application of the rule in (51).

(51a) is a coordinate conjoined structure with three conjuncts each having a noun phrase coreferential with a noun phrase in the other conjunct. The pronominalization rule can apply here forwards. When the rule reduces the coreferential noun phrases in the second and third conjuncts, we get (51b). Besides (51b), the following sentences are derivable from the underlying structure of (51a):

- (54) a. meyeṭi_c 'sahare gelo o meyeṭi_c gān gāilo o se_c sukh hārālo:
The girl went to (a) town and the girl sang songs
and she lost happiness.
- b. meyeṭi_c 'sahare gelo o se_c gān gāilo o meyeṭi_c sukh hārālo:
The girl went to (a) town and she sang songs and
the girl lost happiness.

In (51a) in each application of the pronominalization rule two coreferential noun phrases will take part. If the pronominalization rule reduces the coreferential noun phrase in the second conjunct into a pronoun due to its coreferentiality with the noun phrase in the first conjunct, we will get (54b). We can derive (54a) using either the coreferential noun phrase in the first conjunct or the one in the second conjunct to reduce the coreferential noun phrase in the third conjunct into a pronoun.

Now consider the application of the pronominalization rule in coordinate conjoined structures which contain more than one set of coreferential noun phrases. Consider the examples in (55).

- (55) a. $\text{minu}_c \text{mātinke}_a \text{dekhlo o minu}_c \text{mātinke}_a \text{dāklo}$: Minu saw
Matin and Minu called Matin.
- b. $\text{minu}_c \text{mātinke}_a \text{dekhlo o minu}_c \text{tāke}_a \text{dāklo}$: Minu saw
Matin and Minu called him.
- c. $\text{minu}_c \text{mātinke}_a \text{dekhlo o se}_c \text{mātinke}_a \text{dāklo}$: Minu saw
Matin and she called Matin.
- d. $\text{minu}_c \text{mātinke}_a \text{dekhlo o se}_{c,a} \text{tāke}_{a,c} \text{dāklo}$: Minu saw
Matin and he/she called her/him.

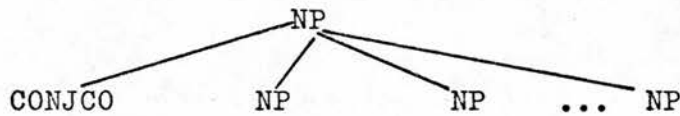
The coordinate conjoined structure (55a) has two sets of coreferential noun phrases, and each set differs from the other only in the value for the feature [MALE]. Let us refer to minu in the left conjunct and minu in the right conjunct as set-1, and mātin in the left conjunct and mātin in the right conjunct as set-2. The rule will apply in (55a) forwards because it is a coordinate conjoined structure. If the rule applies in set-2 we will derive (55b), and if applies in set-1 we will derive (55c); and if the rule applies in both sets we will derive the ambiguous sentence (55d). That is, se in (55d) can refer either to minu or mātin, and tāke can refer either to minu or mātin. This is because of these two sets differ only in the feature specifications [+MALE], the features which are not morphologically marked on the Bengali personal pronouns. But if these two sets would differ one from another in features of number or in person and grade, the application of the

pronominalization rule in both sets would not produce such ambiguous structures because the Bengali personal pronouns are morphologically marked for person, grade and number.

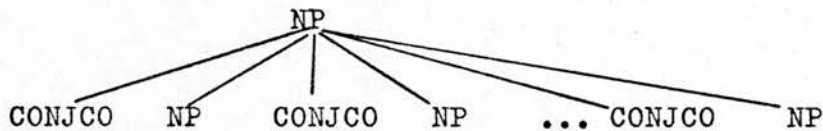
5.6.1 Pronominalization in Coordinate Conjoined Structures with Coordinate Noun Phrases.

A coordinate noun phrase is one which conjoins two or more noun phrases by the conjunctions o, ār, ebaṅ: 'And', and bā, athabā: 'Or'. Our base rule 6 allows noun phrases to be coordinately conjoined in the underlying structure. This rule generates coordinate noun phrase structures like (56a), which transform into structures like (56b) by a conjunction spreading convention. (cf., § 1.2.1: note to the base rule 6; § 2.4.8).

(56) a.



b.



Here we will not go into the controversy whether all conjunctions are phrasal conjunctions or sentence conjunctions in the underlying structure, but we will assume that both are possible in the underlying structure of natural languages (cf., Lakoff and Peters (1966); § 2.4.8).

In this section we will consider the application of the pronominalization rule in coordinate conjoined structures involving underlying coordinate noun phrases. Consider the examples below.

coordinate conjoined structure pronominalization will apply here forwards. One interesting fact about (58) is that, although the superordinate noun phrases NP¹ and NP² are coreferential as a set, pronominalization does not apply in them as a set. Instead the lower noun phrases dominated by NP¹ individually pronominalize the lower noun phrases dominated by NP². If the rule applies forwards in NP³ and NP⁵, we derive (57b); and if the rule applies forwards in NP⁴ and NP⁶, we derive (57c). But if the rule applies in both sets of coreferential noun phrases, we derive the ungrammatical structure (57d). This sentence is ungrammatical because of its right conjunct conjoins two identical pronouns (*se o se). In order to block (57d) we require a pronoun conjunction rule which will automatically reduce conjoined identical pronouns into a plural pronoun. When the conjoined identical pronouns in (57d) are reduced to a single plural pronoun tārā, we derive (57e), which is grammatical. We will consider pronoun conjunction in more detail in the next section.

5.6.2 Pronoun Conjunction.

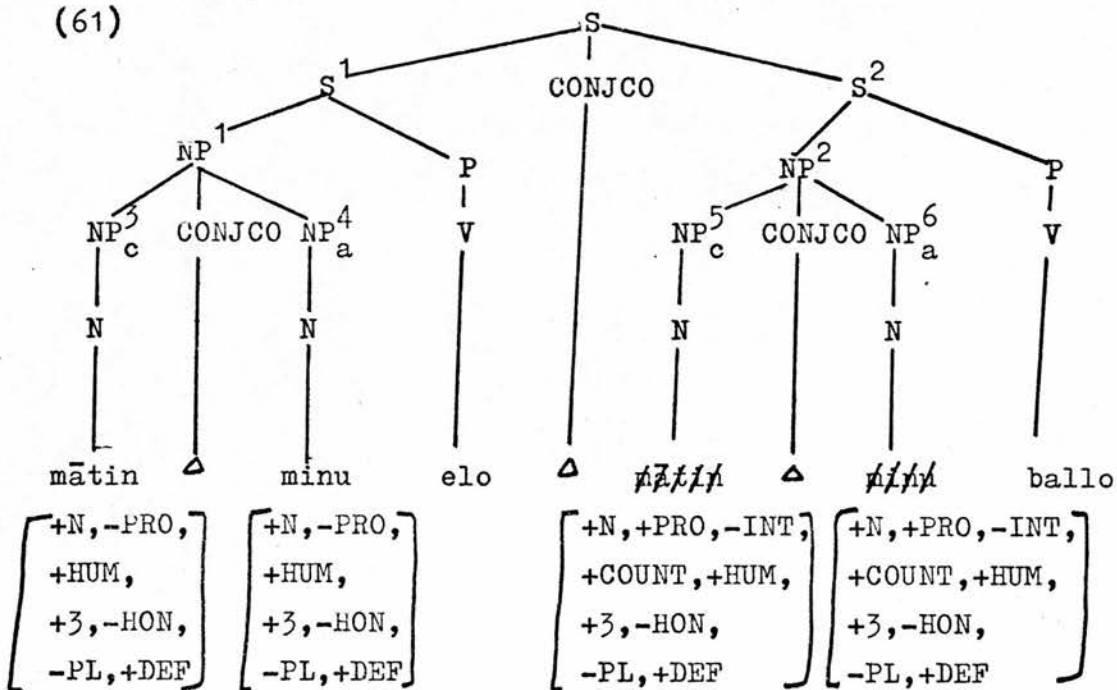
Consider the following examples.

- (59) a. *se o se yābe: He and he will go.
 b. *tini o tini o tini yāben: He and he and he will go.
- (60) a. tumi o āmi yābo: You and I shall go.
 b. ini, uni o tini yāben: He, he and he will go.
 c. se o tumi yābe: He and you will go.

(59a, b) are ungrammatical because they conjoin identical pronouns, but (60a, b, c) are grammatical where non-identical pronouns are conjoined. So we see that Bengali disallows coordination of

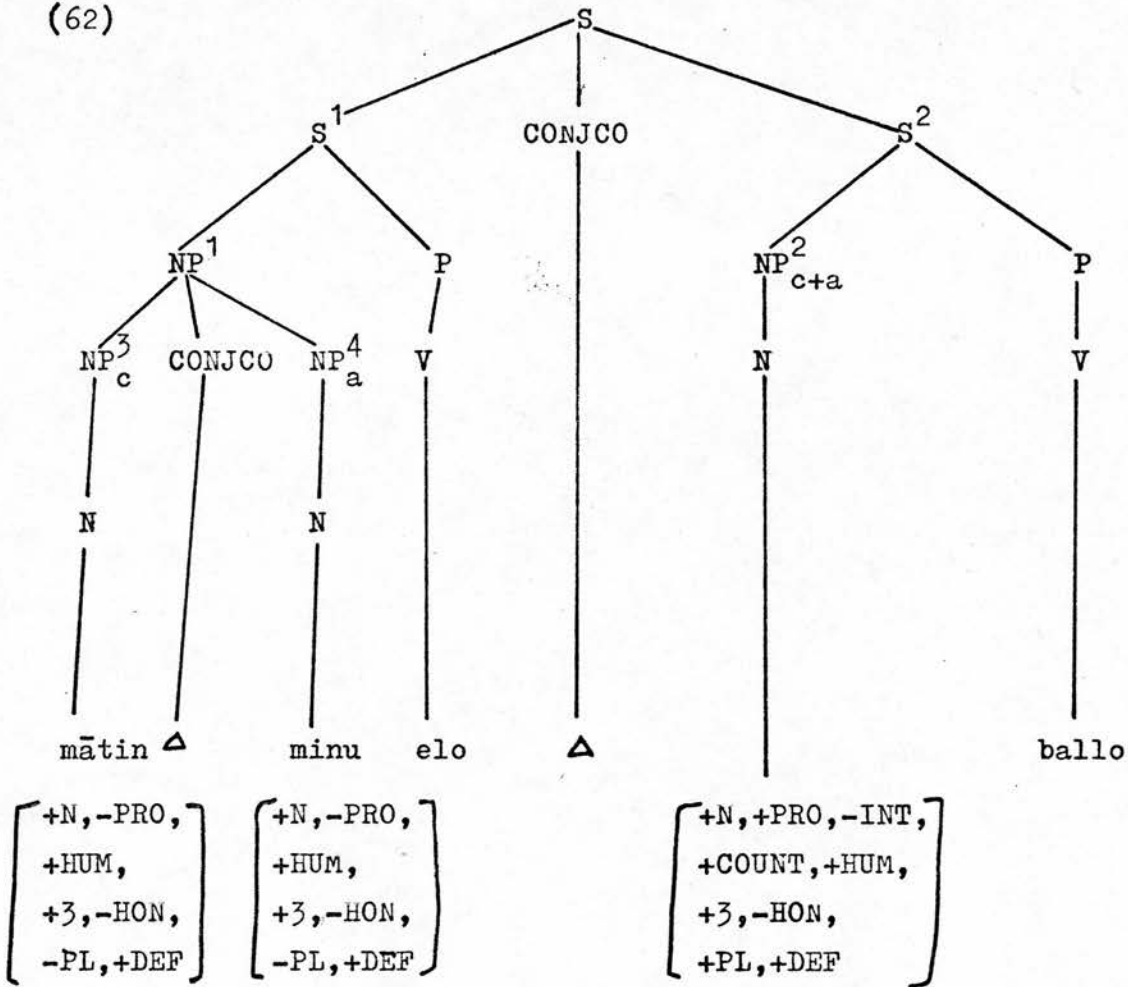
identical pronouns, but allows coordination of non-identical pronouns. But we have seen in (57e) that a structure with a coordination of identical pronouns becomes grammatical if the conjoined identical pronouns are reduced to a single plural pronoun. So we need a pronoun conjunction rule which will apply after the pronominalization rule, and will automatically reduce conjoined identical pronouns into a single plural pronoun. The pronoun conjunction rule will apply obligatorily if the pronominalized noun phrases are the elements of a coordinate noun phrase and have non-distinct values for the features of person, grade, number, and in certain cases [FAR]. This rule will collapse the pronominalized noun phrase conjuncts into a single noun phrase and will specify it as [+PL]. For example, consider (58). If the pronominalization rule applies in both the lower noun phrases in (58) forwards, it will be transformed into (61).

(61)

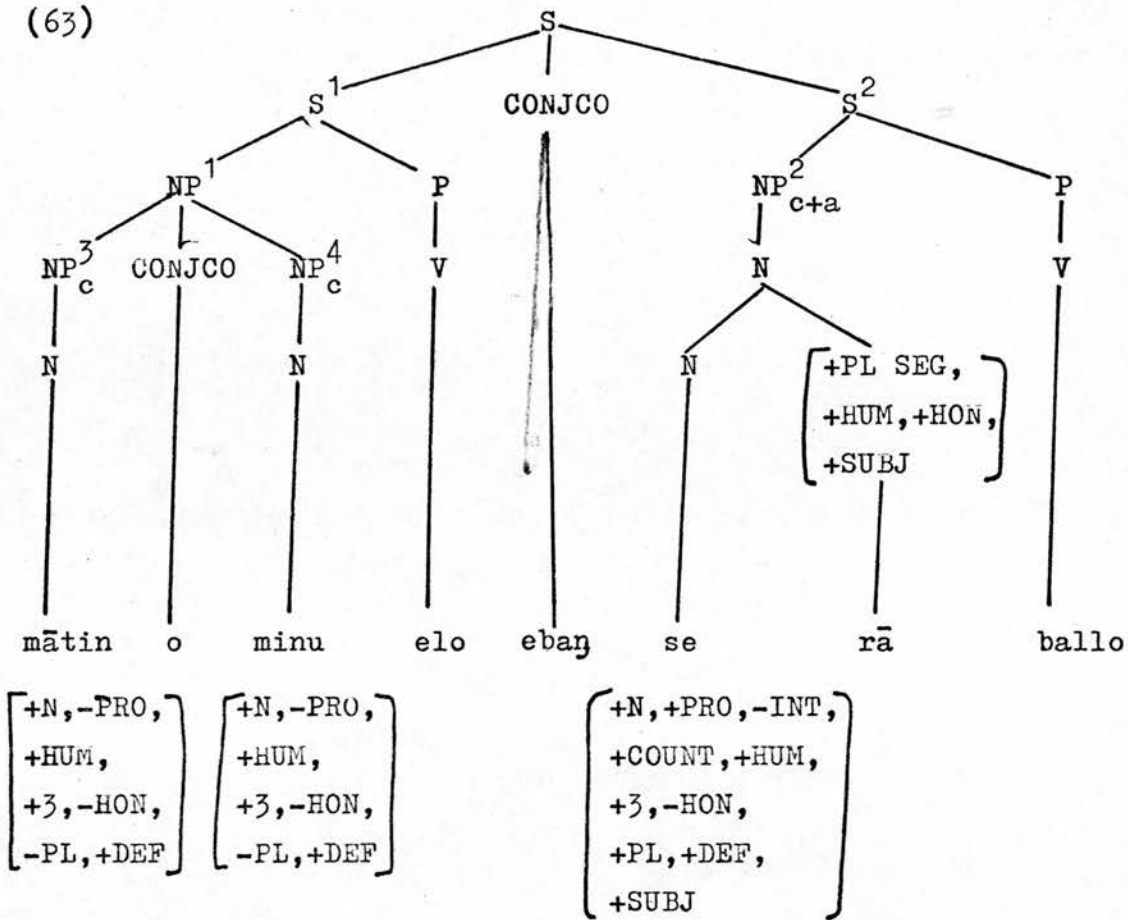


In (61) the pronominalized noun phrases NP⁵ and NP⁶, which are elements of a coordinate noun phrase, have non-distinct features and so the pronoun conjunction rule will apply on them obligatorily. This rule will collapse the pronominalized noun phrases into a single noun phrase, and will specify the noun phrase as [+PL]. This rule will transform (61) into (62).

(62)



NP² in (62) has been positively specified for the feature [PL] by the pronoun conjunction rule. So a plural segment transformation will take place in NP², which will transform (62) into (63).



In (63) we have arrived at the post-transformational string, where the second lexical insertion rule has inserted appropriate lexical items to the derived nodes (we have attached derived verb forms in these trees for readability, although they should not present in the underlying structures). (63) will generate the syntactic surface structure \neq mātin_c o minu_a elo ebaṅ se-rā_{c+a} ballo \neq , which will pass through the morphophonological component, and will generate the phonetic structure mātin_c o minu_a elo ebaṅ tārā_{c+a} ballo, which is (57e).

The pronoun conjunction rule will apply obligatorily to pronominalized conjoined noun phrases which have non-distinct features (especially in features of person, grade, number and sometimes in [FAR], if there is any), and will apply optionally

to pronominalized conjoined noun phrases which have distinct features of person, grade, number, and in some cases, [FAR]. Consider the examples below.

(64) a. *se o se o se \Rightarrow tārā : He and he and he \Rightarrow They.

b. *tini o tini \Rightarrow tārā : He and he \Rightarrow They.

c. se o āpani \Rightarrow āpanārā : He and you \Rightarrow You.

d. tini o āpani \Rightarrow āpanārā : He and you \Rightarrow You.

e. $\left\{ \begin{array}{l} \text{āpani} \\ \text{tumi} \\ \text{tui} \end{array} \right\}$ o āmi \Rightarrow āmarā : You and I \Rightarrow We.

f. $\left\{ \begin{array}{l} \text{tini} \\ \text{se} \end{array} \right\}$ o āmi \Rightarrow āmarā : He and I \Rightarrow We.

g. $\left\{ \begin{array}{l} \text{tini} \\ \text{se} \end{array} \right\}$ o $\left\{ \begin{array}{l} \text{āpani} \\ \text{tumi} \\ \text{tui} \end{array} \right\}$ o āmi \Rightarrow āmarā :

He and you and I \Rightarrow We.

i. *āmi o āmi \neq āmarā : I and I \neq We.

In (64a, b) identical pronouns have been conjoined, which is unacceptable; they will be obligatorily reduced to a plural pronoun by the pronoun conjunction rule.

In (64a) each of the conjoined pronouns has the features [+N,+PRO,+3,-HON,-PL] (irrelevant features omitted), and so the pronoun conjunction rule will collapse them into single noun phrase which will be specified as [+PL] for the number feature. In the surface the pronoun will be realized as tārā. In (64b) the conjoined pronouns are honorific, and the pronoun conjunction rule will similarly collapse them into a single noun phrase with the features [+N,+PRO,+3,+HON,+PL]. In the surface the pronoun will be realized as tārā. In (64c) the conjoined pronouns differ in

the features of person and grade: se has the features [+3,-HON,-PL], and āpani has the features [+2,+HON,-PL]. The pronoun conjunction rule will apply here optionally, and will collapse these two pronouns into a single one with the features [+N,+PRO,+2,+HON,+3,-HON,+PL], and in the surface the pronoun will realize as āpanārā, but not as tārā or tomarā. If we had taken a dominance approach to the conjunction of pronouns, then we could collapse these pronouns into a single one with the features [+N,+PRO,+2,+HON,+PL], and would get the same result. We have avoided this because this approach is counter-intuitive. Instead, we have taken a feature bundling approach to pronoun conjunction, because it can show the inclusive and exclusive use of the first and second person pronouns. In this approach the pronoun conjunction rule will reduce the conjoined pronouns into a single noun phrase and their features will be bundled up (excluding the common features), and the second lexical insertion rule will attach that pronoun to the derived node which is higher in the person and grade hierarchy of the pronouns. The hierarchy is given in (65).

(65) Person and Grade Hierarchy of Pronouns.

$$\left[\begin{array}{l} +1, \\ +2,+HON, \\ +2,-HON, \\ +2,+PEJ \\ +3,+HON \\ +3,-HON \end{array} \right]$$

In this hierarchy the first person is higher than the second

person, and the second person is higher than the third person; and the honorific grade is higher than the nonhonorific grade, and nonhonorific grade is higher than the pejorative grade. That is why, the second lexical insertion rule will attach the pronoun āpani to a node having the feature bundle [+2,+HON,+3,-HON,+PL], but not se, as āpani is higher in the person and grade hierarchy. The pronouns are listed in the lexicons with feature specifications according to this principle. The pronoun conjunction rule will apply optionally to the examples (64d)-(64g), and in the surface they will be realized as pronouns given at the right side of the arrows. In this way we can show that the first and second person plural pronouns in Bengali can be either exclusive or inclusive. Thus we see in (64i) that āmarā is not a conjunction of āmi o āmi, neither simply a plural āmi; but a conjunction either of second and first, or third and first, or third and second and first person pronouns (cf., 64e, f, g).

When several distinct pronouns are coordinately conjoined in Bengali, they are usually conjoined in a certain order. The principle is that the pronoun that is lower in the hierarchy (65) is placed preceding another which is higher in the hierarchy. This is presumably a surface structure constraint in Bengali required for linguistic as well as extra-linguistic reasons. The extra-linguistic reason is that this makes an utterance polite, and the linguistic reason is that when such a coordinate noun phrase is the subject of a sentence, the pronoun with which the verb form agrees occurs closer to the verb form and makes the sentence comfortable. Consider the examples in (66).

(66) a. se, tumi o āmi sinemāy yābo: He, you and I shall
a cinema.

b. ?āmi, tumi o se sinemāy yābo:

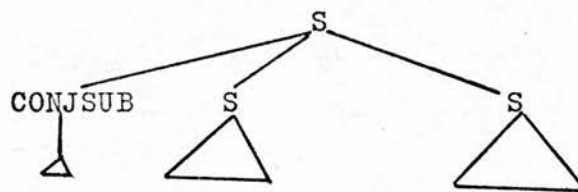
c. ?tumi, āmi o se sinemāy yābo:

In (66a) the third, second and first person pronouns are conjoined in that order. (66a) is a good and natural sentence in Bengali, but (66b, c) are objectionable because the pronouns have been placed in an unacceptable order in these examples.

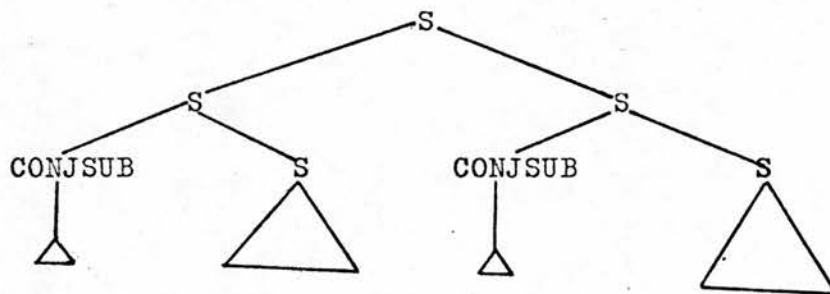
5.7 Pronominalization in Subordinate Conjoined Structure.

Our base rule 1 generates structures like (67a) which transform into structures like (67b) by a conjunction spreading schema.

(67) a.



b.



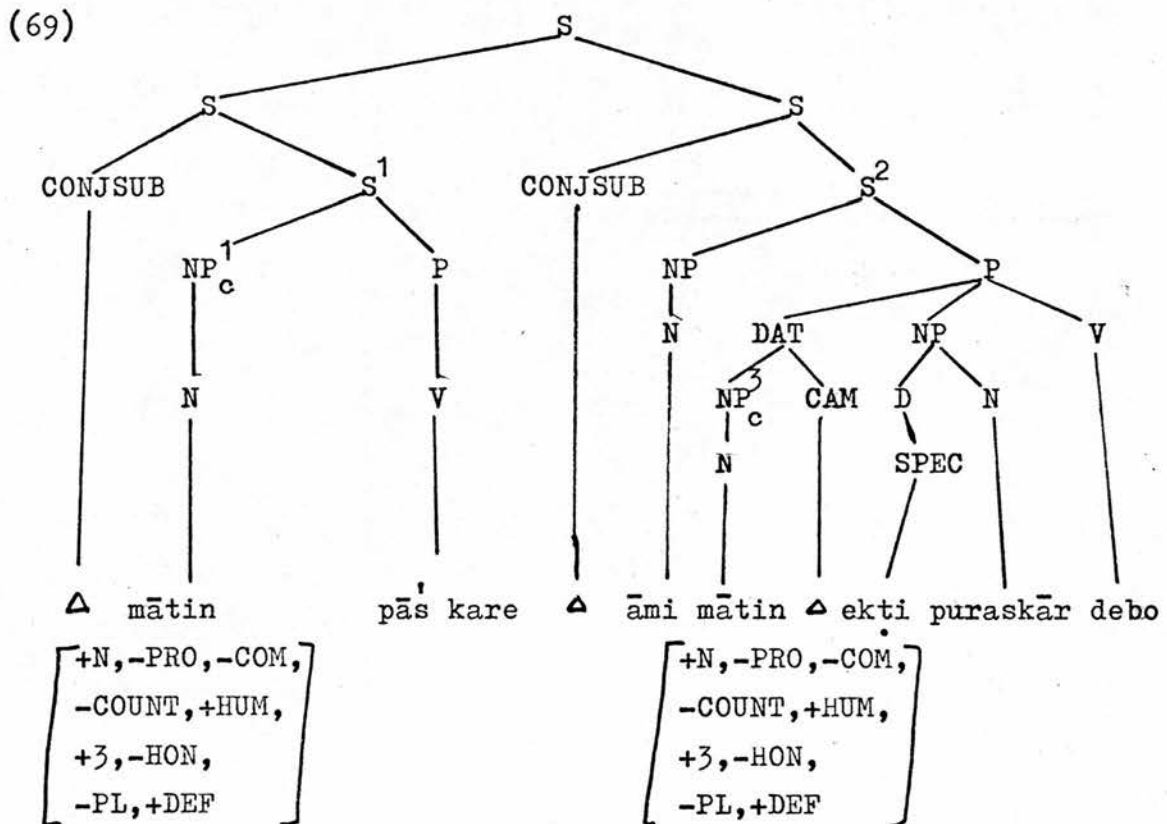
An underlying structure like (67a) which conjoins two conjuncts is a subordinate conjoined structure. In a subordinate conjoined structure the first conjunct is always preceded by the conjunction yadi: 'If', and the second conjunct is preceded by the conjunction tabe. We have seen that this type of structure allows both forward and backward pronominalization (cf., § 5.4; § 5.4.1).

Consider the sentences below.

- (68) a. yadi mātin_c pās¹ kare, tabe āmi tāke_c ekṭi puraskār debo:
 1 3 4 5 6 7 8 9
 If Matin passes, then I shall give him a prize.
- b. yadi se_c pās¹ kare, tabe āmi mātinke_c ekṭi puraskār debo:
 1 2 3 4 5 7 8 9
 If he passes, then I shall give a prize to Matin.

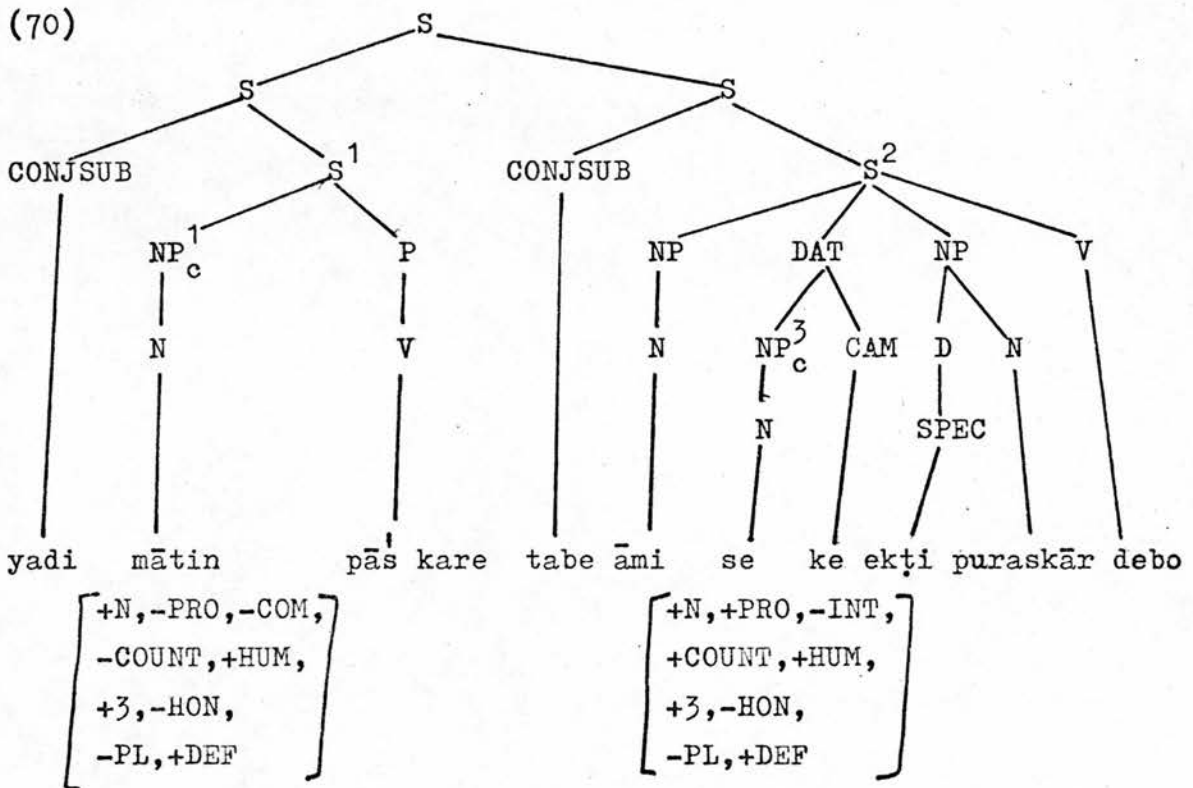
In (68a) tāke refers to mātin, and in (68b) se refers to mātin.

In (68a) pronominalization has applied forwards and in (68b) backwards. These two sentences are derived from an identical underlying structure by the application of the pronominalization rule forwards and backwards. The intermediate structure which underlies the sentences in (68) is given in (69).



(69) has two coreferential noun phrases in two conjuncts, and the noun phrases are identical. The pronominalization rule

can apply in (69) in both directions. We can use NP¹ to reduce NP³ into a pronoun, and similarly we can use NP³ to reduce NP¹ into a pronoun. If the rule applies forwards in (69), we will derive (70) from (69).



In (70) NP³ has been pronominalized under coreference with NP¹. As we have arrived at the post-transformational string in (70), the second lexical insertion rule has attached the pronoun se to the pronominalized noun phrase, and other items to the derived nodes. The terminal string of (70) will generate the syntactic surface structure \neq yadi mātin_c pās¹ kare, tabe āmi se-ke_c ekṭi puraskār debo \neq , which after the application of the morphophonological rules will generate the phonetic structure yadi mātin_c pās¹ kare, tabe āmi tāke_c ekṭi puraskār debo, which is (68a). The pronominalization rule can apply backwards in (69). If the rule reduces NP¹ into a pronoun due to its coreference with NP³, we will derive (68b).

5.8 Pronominalization in Complex Structures with Direct Discourse Complement Clause.

Some complement clauses contain the direct discourse or feeling of the matrix subject. Below we quote from Kuno (1972, 163), who first took direct discourse into consideration in order to explain certain facts of pronominalization in complex structures.

One of the differences between verbs such as expect, claim, know, think, request on the one hand, and such verbs as deny, forget, be aware (of) on the other hand, is that the content of the complement clause of the former represents "more or less" the direct discourse of the matrix subject, while this is not the case for the latter.

Bengali has some verbs which take complement clauses that represent the direct discourse or feeling of the matrix subject. Consider the examples below.

- (71) a. ketakī_c baleche ye se_c gān gāibe: Ketaki has said that
 2 3 4 5 2 3
 she would sing.
- 4 5
 b. ketakī_c mane kare ye se_c rūpasī: Ketaki thinks that
 2 3 4 5 2 3
 she is beautiful.
- 4 5
 c. ketakī_c dabī kareche ye se_c cāde giyechilo: Ketaki
 2 3 4 5 6
 claimed that she had been to the moon.
 2 3 4 6 5
- (72) a. ketakī_c asvikār karlo ye se_c asustha: Ketaki denied
 2 3 4 5 2
 that she was sick.
 3 4 5
- 4 5
 b. ketakī_c bhule gelo ye se_c rūpasī: Ketaki forgot that
 2 3 4 5 2 3
 she was beautiful.
 4 5

The complement clause in each of the sentences in (71) is understood as what ketakī actually said, thought or claimed, but the complement clauses in (72) are not understood as such. In (71) the narrator objectively reproduces Ketaki's direct discourse or feeling, and converts the sentences into indirect discourse, but in (72) the narrator is subjectively involved and gives his own impression of a situation. As the sentences (71) contain complement clauses which represent the direct discourse or feeling of the matrix subject, they are related to their direct discourse counterparts given respectively in (73).

- (73) a. $\text{ketakī}_c \text{ baleche, 'āmi}_c \text{ gān gāibo'}$: Ketaki has said,
'I shall sing.'
- b. $\text{ketakī}_c \text{ mane kare, 'āmi}_c \text{ rūpasī'}$: Ketaki thinks,
'I am beautiful.'
- c. $\text{ketakī}_c \text{ dābī kareche, 'āmi}_c \text{ cāde giyechilam'}$: Ketaki
has claimed, 'I had been to the moon.'

But as the complement clauses in (72) are not the direct discourse of the matrix subjects, the sentences in (72) cannot be related to any direct discourse structure. The sentences given in (74) are unacceptable, and they cannot be related to the sentences in (72).

- (74) a. $*\text{ketakī}_c \text{ asvīkār karlo, 'āmi}_c \text{ asustha'}$: Ketaki denied,
'I am sick.'
- b. $*\text{ketakī}_c \text{ bhule gelo, 'āmi}_c \text{ rūpasī'}$: Ketaki forgot,
'I am beautiful.'

The sentences in (72) give the narrator's impression of some situations, but not the matrix subjects'. These sentences are related to the sentences given respectively in (75).

(75) a. ketakī_c asvīkār karlo e-kathā ye ketakī_c asustha:

Ketaki denied this proposition that Ketaki was sick.

b. ketakī_c bhule gelo e-kathā ye ketakī_c rūpasī:

Ketaki forgot this proposition that Ketaki was beautiful.

In (72, 75a) what Ketaki denies is a proposition or rumour spread by others, and in (72, 75b) the narrator gives his own impression of a situation.

So the verbs can be marked for the features [+DD] (direct discourse) depending on whether they take direct discourse complement clauses or not. The verbs which take strictly direct discourse complement clauses are verbs of uttering like bal: 'Say', dābī kar: 'Claim', ghoṣaṇā kar: 'Declare' etc.,. These verbs can be specified with the feature [+DD], because they take complement clauses which represent the direct discourse or feeling of the matrix subject. The verbs like asvīkār kar: 'Deny' can be marked with the feature [-DD], because their complement clauses do not represent the direct discourse or feeling of the matrix subject. The verbs of feeling such as bhāb, mane kar: 'Think', jān: 'Know' can be marked with the features [+DD], because their complement clauses may or may not represent the direct discourse of the matrix subject. We have said before that a direct discourse sentence is formed by a narrator (cf., § 5.1.3). Consider the direct discourse sentence (73a). We understand that Ketaki simply uttered the sentence āmi gān gaibo, and the narrator has objectively reproduced her utterance in (73a) by making ketakī the subject of the verb bal. The narrator is not involved in any way with

the content of this utterance. When the narrator reports or reproduces someone's discourse or feeling without being personally involved as in (73), we will call such a structure a direct discourse or uninvolved representation; but if the narrator involves himself with the content of the complement clause, we will call such a structure a non-direct discourse or involved representation. In a non-direct discourse the complement clause is not a direct discourse of the matrix subject. Instead, it represents the feeling, judgement or impression of the narrator, as in (72).

5.8.1 Indirect Discourse Formation and Pronoun Symmetry.

We have shown before what sorts of noun phrase can be understood coreferentially in a direct discourse structure (§ 5.1.3). In other types of structure we have seen that coreferential noun phrases must have identical head noun for the application of the pronominalization rule; but identity of head nouns, when the matrix noun phrases have the features [+N, -PRO], is disallowed in a direct discourse structure. We have seen that the noun phrases in a direct discourse complement clause must be pronouns in the underlying structure in order to refer anaphorically to the matrix subject and object. Consider the examples in (76).

- (76) a. meyeṭi_cballo, 'āmi_cyābo': The girl said, 'I shall go.'
 b. meyeṭi_cballo ye se_cyābe: The girl said that she would go.

In (76a) the constituent subject āmi, which is a pronoun, refers to the matrix subject meyeṭi. We claim that āmi is an underlying pronoun. It was never a full noun phrase, because

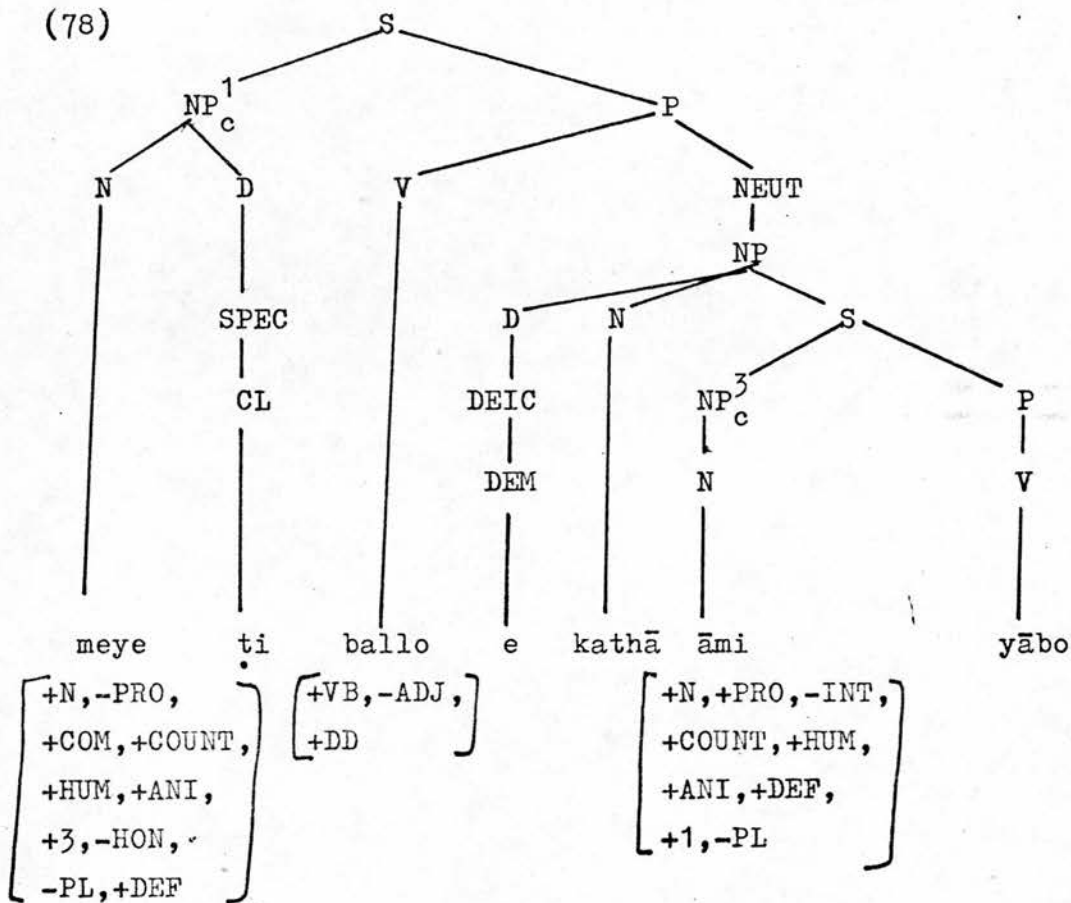
a speaker cannot refer to himself with a full noun phrase in a natural environment. (76b) is derived from (76a) by the indirect discourse formation rule. As āmi in (76a) is an underlying pronoun, it can never be realized as a full noun phrase, nor can it reduce a full noun phrase into a pronoun. Any noun phrase which has the features [+N,+PRO] in the underlying structure will always be realized in the surface as a pronoun; and it has no power to reduce a full noun phrase into a pronoun. So we cannot derive the sentences in (77) from (76a).

- (77) a. *se_cballo ye meyeṭi_cyābe: She said that the girl would go.
- b. *meyeṭi_cballo ye meyeṭi_cyābe: The girl said that the girl would go.

In a direct discourse structure the complement clause contains an underlying pronoun as in (76a); so the pronominalization rule does not derive any pronoun transformationally, but it generates pronouns in such a structure in the underlying structure. The indirect discourse formation rule, which derives sentences like (76b) from (76a), makes the anaphoric pronouns in a direct discourse complement clause symmetrical to their antecedents in the features of person and grade. Consider (76a). In this sentence the matrix subject has the features [+N,-PRO,+3,-HON,-PL] (irrelevant features omitted), and the constituent subject has the features [+N,+PRO,+1,-PL]. They are asymmetrical in the feature of person (the grade feature is irrelevant here because the first person pronoun has no grade distinction). The indirect discourse formation rule when applied to (76a) makes the anaphoric pronoun symmetrical to its antecedent in the features of person and

grade. In this way āmi is converted into se.

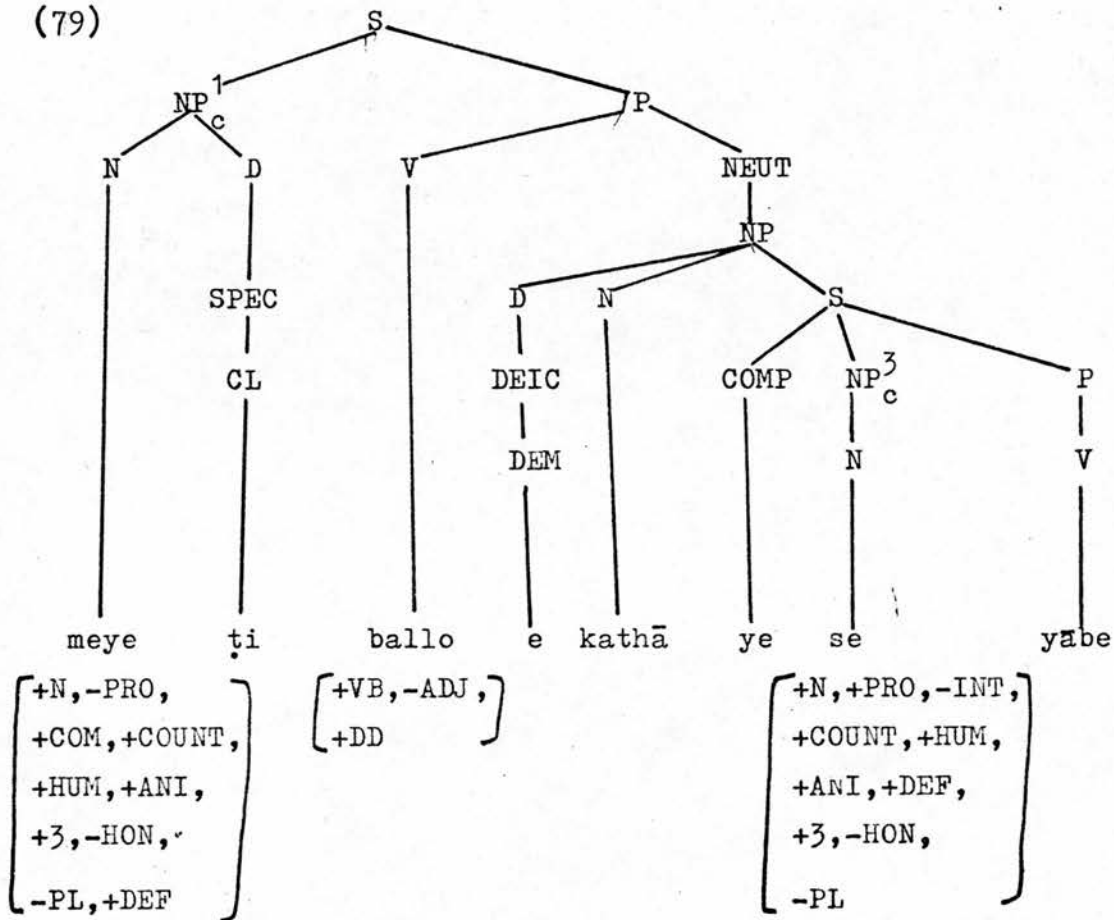
(76a) has the intermediate structure (78).



In (78) NP³, which is an underlying pronoun, refers anaphorically to NP¹. In this structure the pronominalization rule generates the pronoun in the underlying structure. In this structure the indirect discourse formation rule can apply. As this structure involves NP-Complementation, the complementizer ye will be placed as the left daughter of the complement clause, and the indirect discourse formation rule will then apply in (78) obligatorily. The indirect discourse formation rule will make the features of person and grade of the constituent subject in (78) symmetrical to its antecedent NP¹. That is, this rule will change the feature [+1] of NP³ to [+3, -HON].

The complementizer placement and the indirect discourse formation rule rules will transform (78) into (79).

(79)



All the relevant transformations have applied in (79), and we have arrived at the post-transformational string. The second lexical insertion rule has replaced the first lexical item āmi attached with NP³ by se. (79) will generate meyeti ballo e-kathā ye se yābe: 'The girl said this proposition that she would go'. We will derive meyeti ballo ye se yābe, which is (76b), by deleting the abstract noun phrase e-kathā in (79) (cf., § 10.3).

We have said that pronominalization is possible in the forward direction in all sorts of nonsimplex structure in Bengali. But there are certain sentences which show that an anaphoric pronoun can precede its antecedent but the antecedent cannot precede its anaphoric pronoun. These structures show

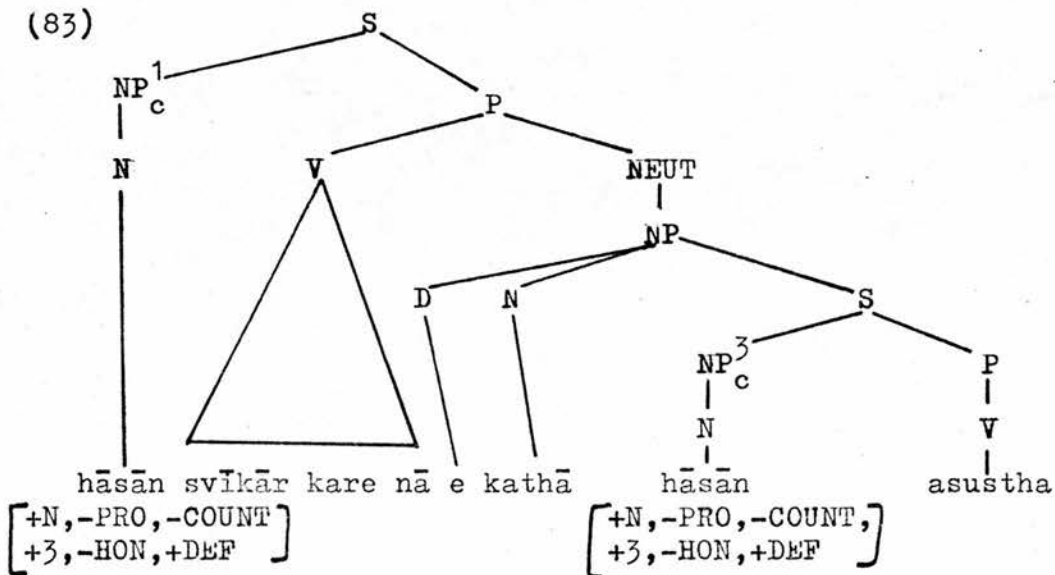
that pronominalization can apply backwards, but not forwards. This astonishing fact can be explained by taking direct and indirect discourse into consideration. Consider the examples below.

- (80) a. $\text{hāsān}_{c_2} \text{e-kathā}_{c_3} \text{svīkār}_{c_4} \text{kare nā ye se}_{c_5} \text{asustha}_{c_6} \text{: Hasan}$
 does not admit this proposition that he is sick.
 5 4 2 3 6 7 8
- b. $*\text{se}_{c_5} \text{e-kathā}_{c_3} \text{svīkār}_{c_4} \text{kare nā ye hāsān}_{c_6} \text{asustha}_{c_7} \text{: He does}$
 not admit this proposition that Hasan is sick.
- c. $\text{se}_{c_5} \text{ye asustha}_{c_6}, \text{e-kathā}_{c_3} \text{hāsān}_{c_2} \text{svīkār}_{c_4} \text{kare nā} \text{: That he}$
 is sick, Hasan does not admit this proposition.
- d. $\text{hāsān}_{c_2} \text{ye asustha}_{c_6}, \text{e-kathā}_{c_3} \text{se}_{c_5} \text{svīkār}_{c_4} \text{kare nā} \text{: That}$
 Hasan is sick, he does not admit this proposition.
- (81) a. $\text{e-bodh}_{c_1} \text{hāsānke}_{c_2} \text{kaṣṭa}_{c_4} \text{dey ye se}_{c_5} \text{gopane}_{c_6} \text{ketaḱīke}_{c_7}$
 $\text{bhālabāse} \text{: This feeling pains Hasan that he}$
 secretly loves Ketaki.
 9 1 2 4 5 6 7 9
- b. $\text{se}_{c_5} \text{ye gopane}_{c_6} \text{ketaḱīke}_{c_7} \text{bhālabāse}_{c_8}, \text{e-bodh}_{c_1} \text{hāsānke}_{c_2}$
 $\text{kaṣṭa}_{c_4} \text{dey} \text{: That he loves Ketaki secretly, this}$
 feeling pains Hasan.
- c. $*\text{hāsān}_{c_2} \text{ye gopane}_{c_6} \text{ketaḱīke}_{c_7} \text{bhālabāse}_{c_8}, \text{e-bodh}_{c_1} \text{tāke}_{c_9}$
 $\text{kaṣṭa}_{c_4} \text{dey} \text{: That Hasan secretly loves Ketaki,}$
 this feeling pains him.

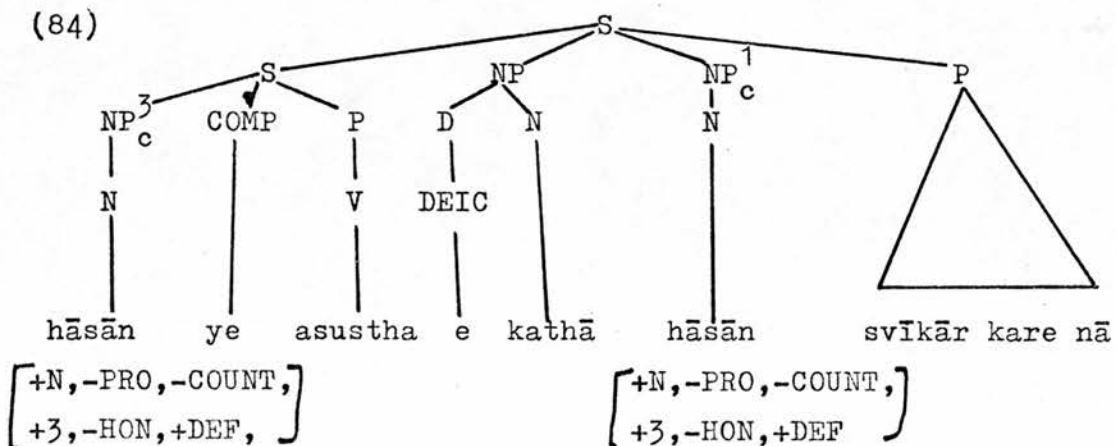
The complement clause in (80a) is not a direct discourse of the matrix subject hāsān; it is understood as some abstract fact about hāsān believed by people other than hāsān. But the complement clause in (81a) is a direct feeling of hāsān himself. We assume that the sentences in (80) and (81) are derived respectively from (82a) and (82b).

- (82) a. $hāsān_c$ svīkār kare nā e-kathā ye $hāsān_c$ asustha:
 Hasan does not admit this proposition that Hasan
 is sick.
- b. e-bodh, 'āmi_c gopane ketakīke bhālabāsi', $hāsān_c$
 kaṣṭa dey: This feeling, 'I love Ketaki secretly',
 pains Hasan.

In (82a) the complement clause $hāsān$ asustha is an abstract fact believed by others, and this clause contains a full noun phrase $hāsān$, which is coreferential with the matrix subject. But in (82b) the complement clause āmi gopane ketakīke bhālabāsi is a secret and direct feeling of $hāsān$; and this clause contains a first person pronoun in order to refer to $hāsān$. The pronoun āmi in the complement clause in (82b) is an underlying pronoun, and so it can never be realized as a full noun phrase in the surface, nor can it reduce another noun phrase into a pronoun. The complement clause in (82a) contains a nominal noun phrase, so it can be realized either as a full noun phrase or as a pronoun in the surface, depending on the application of the pronominalization rule. The structure (83) immediately underlies (82a).

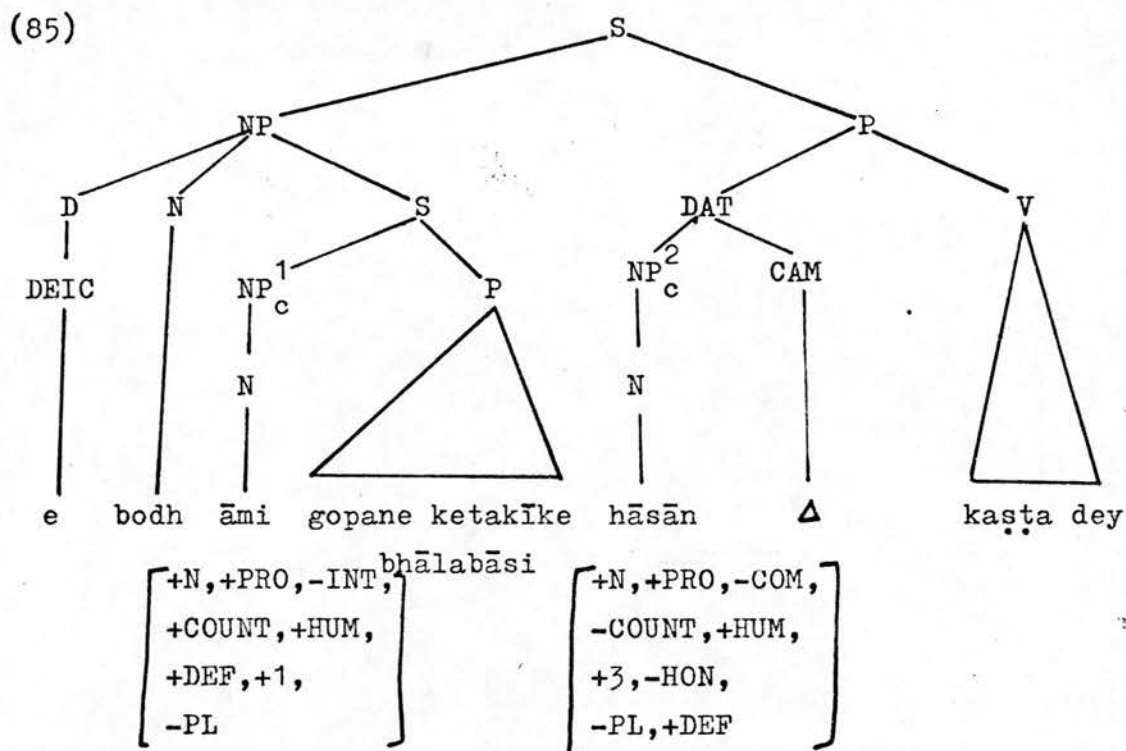


In (83) the complement clause is not a direct discourse of the matrix subject, and so it contains a nominal noun phrase hāsān, which is coreferential with the matrix subject. In (83) NP¹ and NP³ are coreferential, but as NP¹ is in the matrix sentence and NP³ is in the constituent sentence, and NP¹ precedes NP³, so we cannot use NP³ to pronominalize NP¹. This structure disallows the application of the pronominalization rule backwards. But the rule can apply forwards: NP¹ can be used to pronominalize NP³. If the pronominalization rule applies backwards in (83) we will derive the unacceptable sentence se svīkār kare nā e-kathā ye hāsān asustha: 'He does not admit this proposition that Hasan is sick', which becomes the equally unacceptable sentence (80b) after constituent reordering. If the rule applies forwards in (83) we will derive (80a). The complement clause in (83) can be extraposed to the sentence-initial position. This operation is preceded by the movement of the complementizer from the left side of the constituent sentence to a medial position of the complement sentence (the complementizer is moved inside the complement clause if the complement clause is factive, and emphasized (cf., § 10.4)). These operations will transform (83) into (84).



In (84) NP³, which is in the subordinate clause, precedes the matrix subject NP¹. In this structure pronominalization can apply both forwards and backwards. If the rule reduces NP¹ into a pronoun due to its coreference with NP³, we will derive (80d); and if it reduces NP³ into a pronoun due to its coreference with NP¹, we will derive (80c).

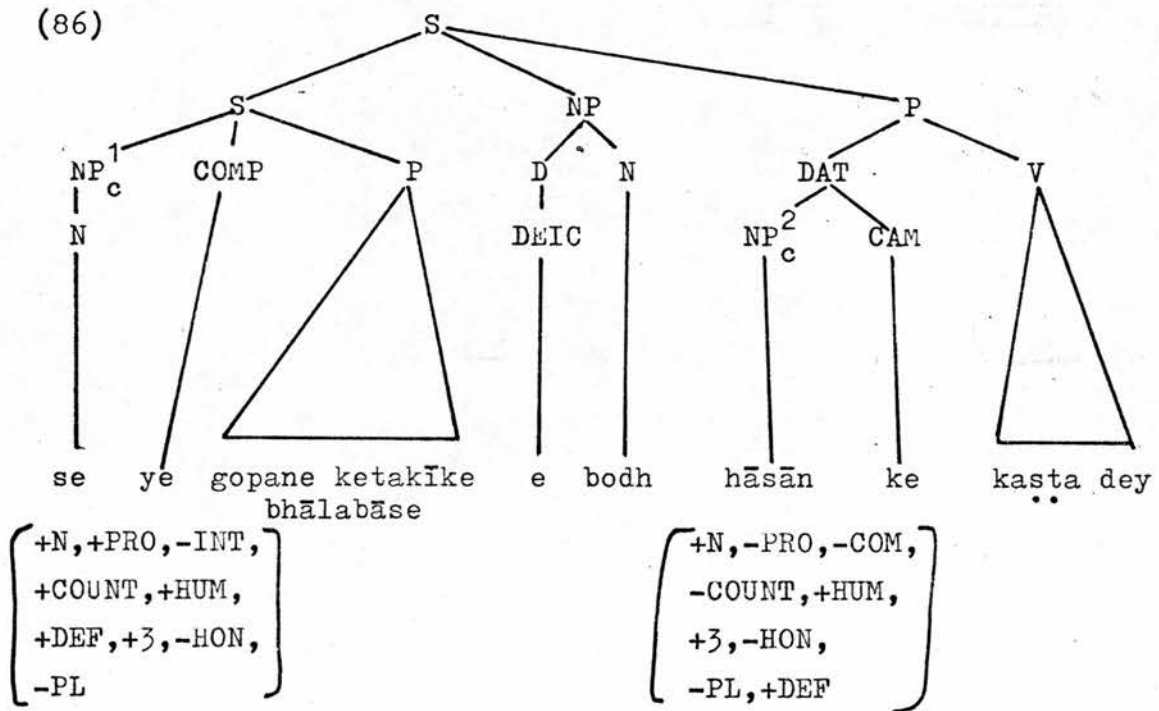
Now consider the structure (85), which immediately underlies (82b).



In (85) the complement clause is a direct discourse of NP², which is in the matrix sentence. Here the constituent subject (NP¹) is an underlying pronoun, which refers to hāsān in the matrix sentence. As NP¹ is a pronoun, it cannot be realized as a full noun phrase in the surface, nor can it reduce another noun phrase into a pronoun. So the transformational pronominalization rule has nothing to do in (85). But the complementizer placement and the indirect discourse formation rules can apply to (85). The complementizer placement rule will place the

complementizer ye as the left daughter of the complement clause, and the indirect discourse formation rule will change the feature of person of NP¹ to [+3,-HON], because its antecedent has the features [+3,-HON]. In the surface structure NP¹ will be realized as se. In this way will derive e-bodh ye se_c gopane ketakīke bhālabāse hāsānke kaṣṭa dey: 'This feeling that he secretly loves Ketaki pains Hasan'. This sentence becomes (81a) when we extrapose the complement clause to the sentence-final position.

Instead, we can carry out the following operations on (85): (a) complementizer placement as the left daughter of the constituent sentence, (b) indirect discourse formation in the constituent sentence, (c) complementizer movement and re-placement inside the complement clause following NP¹, and (d) extraposition of the complement sentence to the sentence-initial position. These operations will transform (85) into (86).



The indirect discourse formation rule has changed the features

of person and grade associated with NP¹ into [+3,-HON].

So it will be realized as se in the surface structure. The second lexical insertion rule has attached appropriate lexical items to the derived nodes in (86). The terminal string of (86) will generate se_c ye gopane ketakīke bhālabāse, e-bodh hāsānke_c kaṣṭa dey, which is (81b).

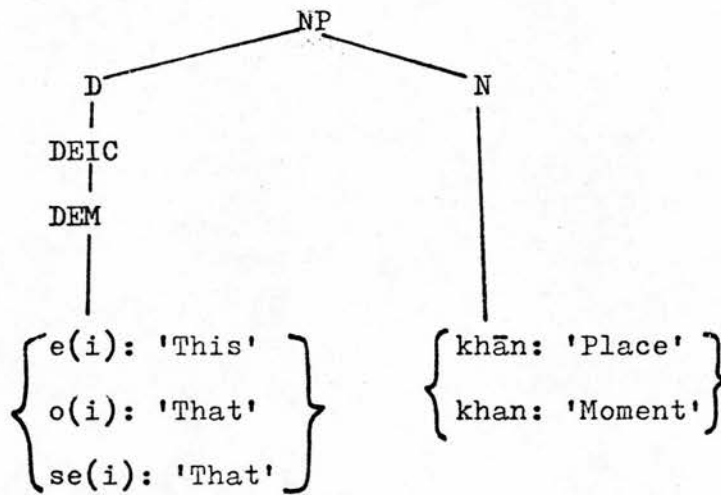
As NP¹ is an underlying pronoun, it cannot reduce NP² in (85) into a pronoun, and realize itself as a full noun phrase. So there is no way to derive the ungrammatical structure (81c) from (85). Thus we see that the application of the pronominalization rule in the forward direction in a complex structure is disallowed only when the structure contains a direct discourse complement clause, which contains no underlying full noun phrase to reduce another into a pronoun. Ross (1968b) faced difficulties with sentences like 'Realizing that he_c was unpopular did not disturb Oscar_c', where the pronominalization rule applies backwards, but does not apply forwards. But we can explain this fact when we take direct and indirect discourse into consideration. The verb 'realize' takes a direct discourse complement clause, and the complement clause contains an underlying pronoun 'I', which converts into 'He' due to indirect discourse formation. The application of the transformational pronominalization rule is blocked here because the structure has no full noun phrase which can be used as an antecedent.

5.9 The Derivation of Anaphoric Locative Pronouns.⁴

The locative pronoun ekhān: 'Here' is used to refer to a place in the vicinity of the speaker, okhān: 'There' is used to refer to a place far from but in sight of the speaker, and

4 The locative and temporal pronouns in Bengali are not 'genuine' pronouns. Structurally they are full noun phrases consisting of the demonstrative deictics and the locational noun khān: 'Place', and the temporal noun khan: 'Moment' (cf., Chatterji (1926, 857-59)). They have a structure like (A) given below.

(A)



In this work they have been taken as pronouns, because (a) they are used for the purpose of reference, (b) the nouns khān and khan are rarely used as free forms in SCB, and (c) they are limited in number. In this section we will consider the derivation of the anaphoric locative pronouns briefly. We will derive the anaphoric locative pronouns transformationally, but an alternative treatment of them is possible if they are considered as full noun phrases. We will not deal with temporal pronominalization in this work. But it may be mentioned that ekhan: 'Now' is usually used deictically; takhan: 'Then' is used deictically as well as anaphorically; and okhan: 'Then' is rarely used in Bengali.

sekhān: 'There' is used to refer to a place far from and out of sight of the speaker. The locative pronouns can be used deictically as well as anaphorically. When used deictically the locative pronouns are generated in the underlying structure in this grammar, and when they are used anaphorically, they are derived transformationally.

We have seen that a pronominalized noun phrase is realized as a locative pronoun in the surface when the noun phrase is dominated by LOC; but it is not the case that a pronominalized noun phrase dominated by LOC should always be realized as a locative pronoun (cf., § 4.3.3). Here we will consider the application of the pronominalization rule in noun phrases dominated by LOC, and will limit this to inherently locational nouns. Consider the examples below.

(87) a. bādāl ḍhākāy_c āche ār mātino $\left\{ \begin{array}{l} \text{sekhān} \\ *ekhān \end{array} \right\}_c$ e āche:

Badal is in Dacca and Matin is $\left\{ \begin{array}{l} \text{there} \\ *here \end{array} \right\}$, too.

b. *bādāl $\left\{ \begin{array}{l} \text{sekhān} \\ *ekhān \end{array} \right\}_c$ e āche ār mātino ḍhākāy āche:

The speaker of the above sentences is not in Dacca. In the underlying structure of (87a) there are two coreferential noun phrases dominated by LOC, and their head nouns are identical.

They have the inherent features [-COM, -COUNT, +LOCATION].

The pronominalization rule can apply in this underlying structure.

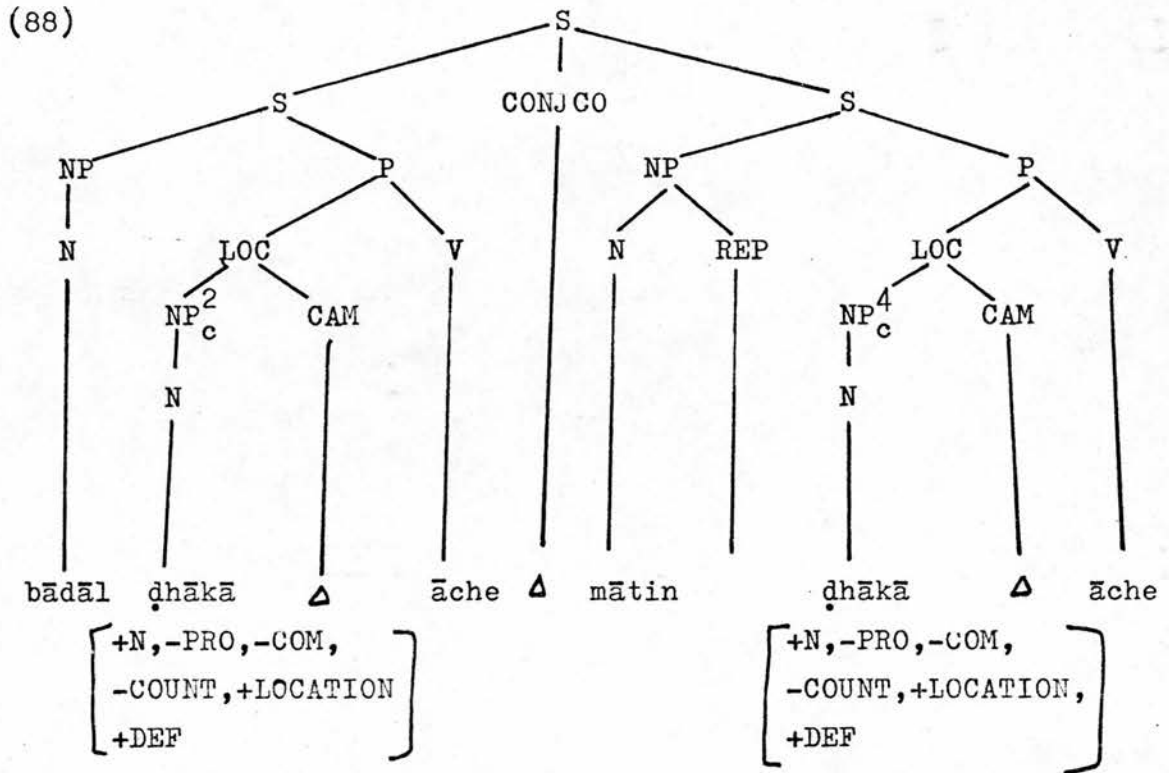
The rule will apply here forwards as the structure that contains these noun phrases is a coordinate conjoined structure.

The rule cannot apply backwards in such a structure, and if it

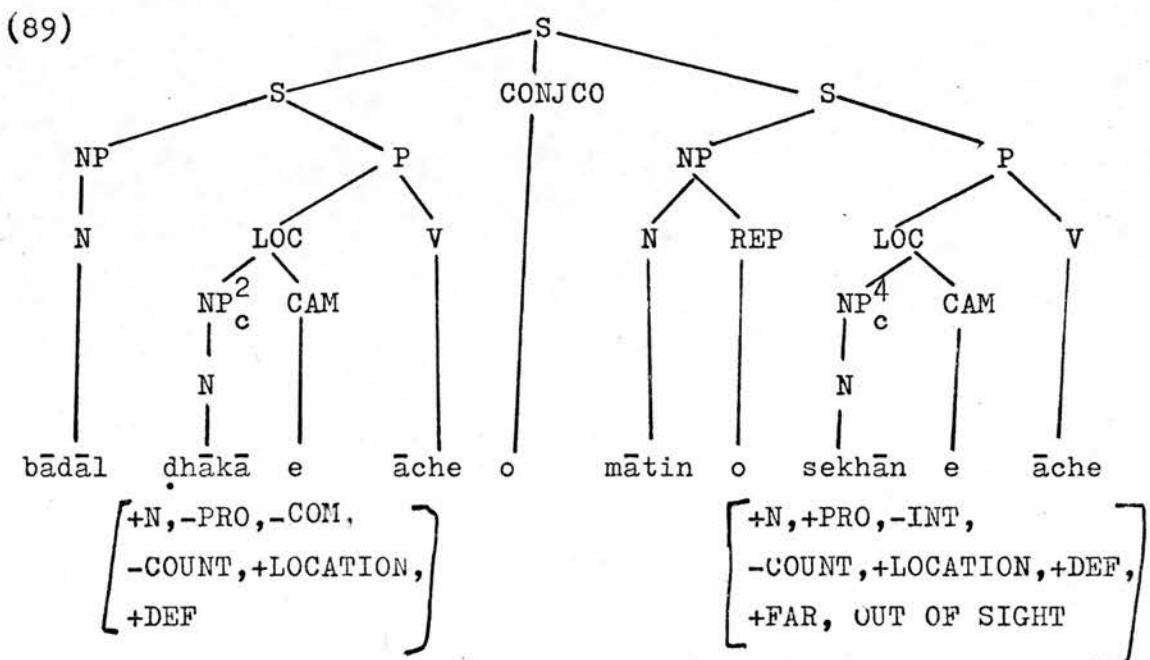
applies backwards, we will derive the unacceptable sentences in (87b). In (87a) we see that the pronominalized noun phrase has been realized as the pronoun sekhān, because the place referred to is far from and out of sight of the speaker. The pronominalized noun phrase in (87a) cannot be realized as ekhān. But the pronoun ekhān would be acceptable if the place referred to were in the vicinity of the speaker. This shows that the pronominalization rule when applied to a locative noun phrase must indicate the [+FAR] features of the pronominalized noun phrase in relation to the speaker, in order to realize the appropriate locative pronoun in the surface structure. This problem arises with [-COM, -COUNT, +LOCATION] nouns, which disallow cooccurrence of any deictic element with them. But it would be counter-intuitive to consider that [+FAR] is an inherent feature of inherently locative nouns. There are verbs like ās: 'Come' and yā: 'Go', which can be considered as deictic verbs in that ās indicates movement towards the speaker, and yā indicates movement away from the speaker. It is possible to specify them for deictic features as should help in the realization of the appropriate locative pronoun. (cf., Fillmore (1966a)). But this is not possible with non-motional verbs like āch: 'Exist, Reside', and thāk: 'Reside'. We will solve this problem tentatively by relegating the power of determining the [+FAR, +IN SIGHT] features of the pronominalized locative noun phrases to the pronominalization rule. The pronominalization rule, besides specifying the NP^D as [+PRO, -INT], will specify it for the deictic features mentioned above.

We will now consider the derivation of the sentence in (87a). The structure which immediately underlies (87a) is

given in (88).



(88) is a coordinate conjoined structure which have two coreferential noun phrases with identical head nouns. This structure satisfies all the conditions for the pronominalization rule. The rule will apply here forwards. The pronominalization rule will specify NP⁴ as [+PRO, -INT, +FAR, OUT OF SIGHT] due to its coreference with NP². This will transform (88) into (89).



We have arrived at the post-transformational string in (89); and so the second lexical insertion rule has inserted the appropriate lexical items to the derived nodes. The second lexical insertion rule has attached the pronoun sekhān to the pronominalized noun phrase because its feature complex matches with that of the pronoun sekhān.

5.10 The Derivation of the Abstract Pronoun tā.

Bengali has a pronoun tā, which we refer to as the abstract pronoun. It is mainly used to refer to abstract objects, but can be used to refer to concrete objects when they are understood as 'things'. It is used deictically, anaphorically, and under identity of sense. It is used deictically in (90).

- (90) a. āmi tā jāni nā: I do not know that.
 1 2 3 4 1 4 3 2
- b. tumi ki tā cāo?: Do you want that?
 1 2 3 4 1 4 3

In the above examples tā refers to some abstract objects deictically, and so the pronoun is generated in the underlying structure of these sentences.

When tā is used to refer to concrete (count or uncount) objects deictically, anaphorically or under identity of sense, its referents are understood not as what they are, but simply as 'things'. Consider the examples in (91).

- (91) a. sekhāne sabāi mad khāy ār āmio tā kheyechi: Everybody
 drinks wine there and I have drunk it (that thing), too.
- b. mātinke āmi ekṭi bai_c diyechilām, kintu se tā_c kakhano
 khuleo dekhe ni: I had given Matin a book, but he
 never even opened it (that thing) to have a look.

In (91a) tā refers to mad under identity of sense, but in (91b) tā refers anaphorically to ekṭi bai. In each sentence the

antecedent of the pronoun is understood as a 'thing'. Its peculiarity becomes transparent if we replace tā in (91b) by se-ti (\leftarrow se-baiṭi: 'That book'), which refers to the book as a book, but not as a 'thing'.

In (92) tā refers to an abstract noun phrase under identity of sense.

(92) a. takhan āsā chilo, kintu ekhan tā nei: There was hope then (at that time), but it is no more now.

b. *takhan tā chilo, kintu ekhan āsā nei:

(92a) shows that pronominalization can reduce an abstract noun phrase into the abstract pronoun tā under identity of sense.

(92b) is ungrammatical because the rule cannot apply backwards in a coordinate conjoined structure. We have dealt with the syntax of this pronoun in more detail in Chapter 10.

5.11 Summary of the Major Rules Discussed.

(93) PRONOMINALIZATION PROPER.

We give this rule in two parts: one for the [+HUM] noun phrases, and the other for the [-HUM] and [-ANI] noun phrases, for the sake of simplicity.

A. Application of the pronominalization rule in [+HUM] noun phrases.

$$\text{SI: } \underset{1}{S} \left[\underset{2}{X} \underset{3}{NP} \left[\underset{4}{X} \underset{5}{N} \underset{6}{X} \right] \underset{7}{X} \underset{8}{NP} \left[\underset{9}{X} \underset{10}{N} \underset{11}{X} \right] \underset{12}{X} \right]$$

$$\left[\begin{array}{c} +N, -PRO, \\ \dots \\ +HUM \end{array} \right] \quad \left[\begin{array}{c} +N, -PRO, \\ \dots \\ +HUM \end{array} \right]$$

1 2 3 4 5 6 7 8 9 10 11 12

SC: (1) Forward Pronominalization.

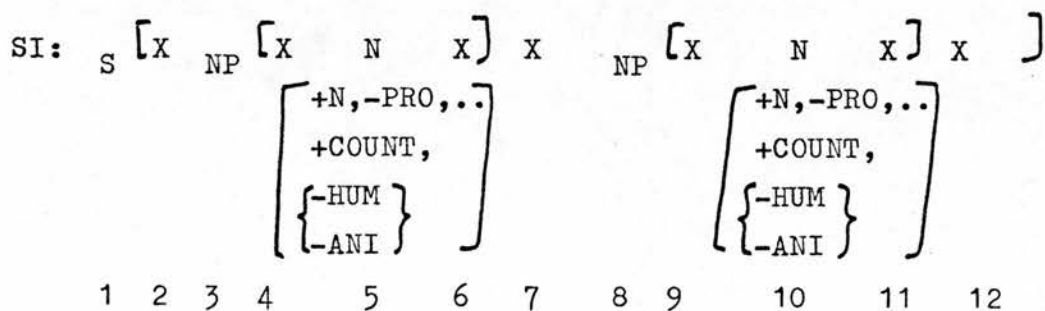
- a. Replace the feature [-PRO] of 10 by [+PRO,-INT] ; delete [+COM] features of 10; and if 10 has the feature [-COUNT] , change it to [+COUNT] .
- b. Delete 9 and 11 if none is PL.
- c. This rule applies in all nonsimplex structures.

SC: (2) Backward Pronominalization.

- a. Replace the feature [-PRO] of 5 by [+PRO,-INT] ; delete [+COM] features of 5; and if 5 has the feature [-COUNT] , change it to [+COUNT]
- b. Delete 4 and 6 if none is PL.
- c. This rule applies in a subordinate conjoined structure, or in a complex structure where 3 is in a subordinate clause.

- Conditions:
- a. 3 and 8 are coreferential, and 5=10.
 - b. For the SC (1) 8 must be [+DEF] , and for SC (2) 3 must be [+DEF] .
 - c. 3 and 8 are not in the same simplex.

B. Application of the pronominalization rule in [-HUM] and [-ANI] noun phrases.



SC: a. Delete 10 if 9 is DEIC and 11 is SPEC.

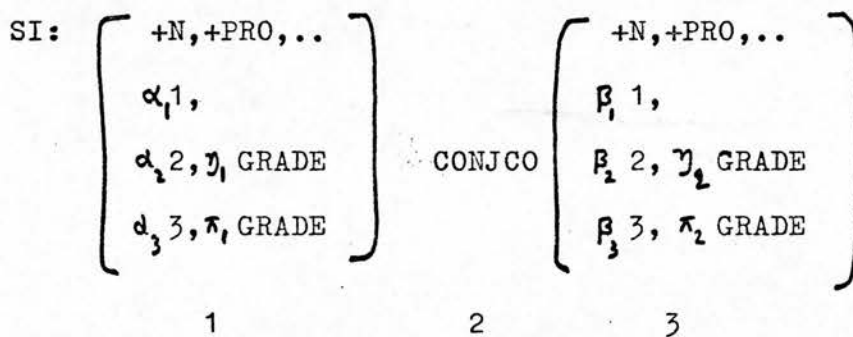
b. Otherwise delete 8.

Conditions: a. 3 and 8 are coreferential, and 5=10.

b. 8 is [+DEF].

c. 3 and 8 are not in the same simplex.

(94) PRONOUN CONJUNCTION.



SC: a. Conjoin the features of 1 and 3 into a complex symbol.

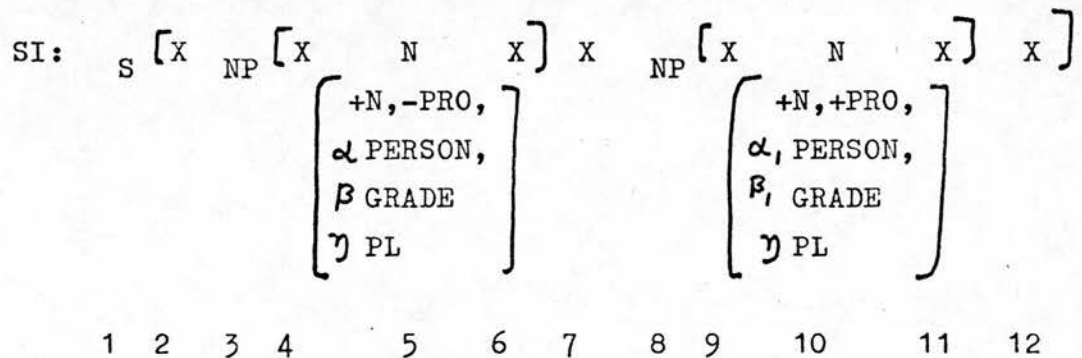
b. Specify the CS with [+PL].

c. If values for any feature are negative (-) both in 1 and 3, then the value for that feature will be negative in the CS; if one value for any feature is positive (+) then the value for this feature will be positive in the CS.

d. If the value for a feature is positive both in 1 and 3, then the rule is obligatory; otherwise optional.

The pronoun conjunction rule given in (94) will obligatorily change tumi o tumi \Rightarrow tomarā, se o se \Rightarrow tārā, and so forth; and this rule will optionally change tini o āpani \Rightarrow āpanārā, se o āpani \Rightarrow āpanārā, tumi o āmi \Rightarrow āmarā and so forth.

(95) INDIRECT DISCOURSE FORMATION.



SC: Make 10 symmetrical in person and grade with 5.

Conditions: a. 3 and 8 are coreferential,

b. 3 is in a matrix sentence, and 8 is
in a direct discourse complement sentence.

This rule will transform (96a) into (96b).

- (96) a. $nāju_c \text{ baleche}$, 'āmi_c yābo': Nazu has said, 'I shall go.'
 b. $nāju_c \text{ baleche ye se}_c \text{ yābe}$: Nazu has said that she
 would go.

REFLEXIVIZATION6.0 Introduction.

Lyons (1968, 361) defines a reflexive construction as 'one in which the subject and object refer to the same person (or thing).' Moyne (1971, 155) assumes that 'reflexivization is a process between the subject and object of the verb.' We shall also consider that reflexivization is a process between the subject and object of a sentence (but cf., § 6.6.2 for some residual problems). Reflexivization is a process which specifies the object noun phrase as [+PRO,+REFL] if it is coreferential with the subject noun phrase. The rule operates in a simplex structure when the subject and object noun phrases are coreferential and identical in head noun. In the surface the reflexivized noun phrase is realized as either of the reflexive pronouns nij or āpan, both of which can be glossed as 'Self' in English. Some reflexive sentences are exemplified in (1).

- (1) a. ketakī_c nijke_c bhālabāse: Ketaki loves herself.
 2 3 3 2
- b. āmi_c nijke_c jāni nā: I do not know myself.
 1 2 3 4 1 4 3 2
- c. meyeti_c nijke_c sundarī mane kare: The girl thinks
 1 2 3 4 1 4
 herself beautiful.
 2 3

In each of the sentences in (1) the object noun phrase is reflexivized and refers anaphorically to the subject noun phrase.

In this chapter we will consider the process of reflexivization in Bengali, and show how the reflexive pronouns in Bengali can be derived. We will also try to show that not all surface

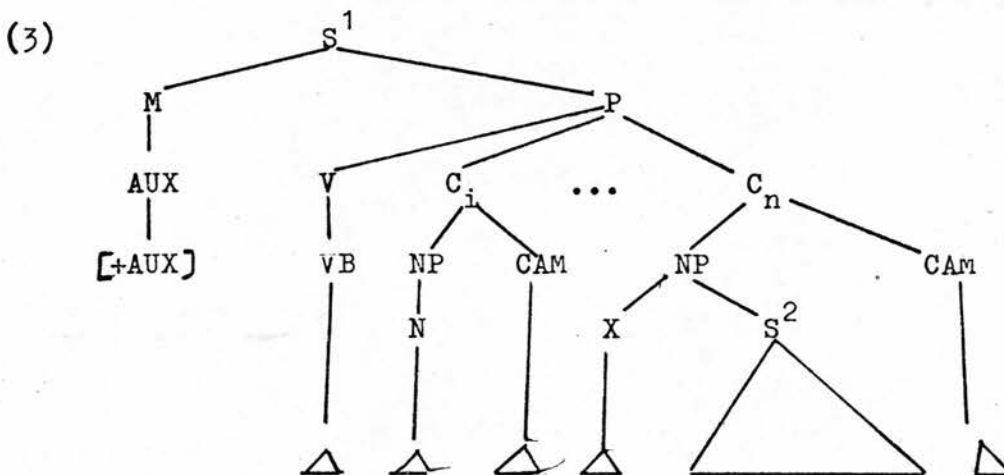
structures with the items nij and āpan can be considered as reflexive structures. They can be used for the purpose of emphasis as shown in (2).

- (2) a. mātin nije giyechilo: Matin had gone himself.
 b. e-ṭi mātiner nijer bai: This is Matin's own book.
 c. tumi āpan kāje man dāo: You mind your own business.

We will consider that nij and āpan in (2) are used as emphatic morphemes (cf., § 6.6; § 6.6.1).

6.1 Reflexivization and the Simplex Structure Constraint.

The major condition for the 'Reflexive Rule' of Lees and Klima (1963) is that the Nom and Nom' involved in the rule must be 'within the same simplex sentence'. A simplex structure is one which has no other embedded S'. But a simplex structure may be embedded in another simplex structure and another sentence may be embedded in it. However, when the rule applies it takes only one S into account. Consider the structure below.



In (3) a simplex structure S^2 is embedded in another simplex structure S^1 : the reflexive rule can operate between the subject and object of S^2 , or between the subject and object of S^1 (excluding S^2), but not between the subject of S^1 and

and the object of S^2 and vice versa.

We will now consider some reflexive sentences, and show that the same simplex structure constraint is necessary for the reflexive rule in Bengali. Consider the examples below.

- (4) a. $\text{ketakī}_c \text{ketakīke}_c \text{bhālabāse}$: Ketaki loves Ketaki.
 b. $\text{ketakī}_c \text{nijke}_c \text{bhālabāse}$: Ketaki loves herself.
- (5) a. $\text{tumi}_c \text{tomāke}_c \text{bhālabāso}$: You love you.
 b. $\text{tumi}_c \text{nijke}_c \text{bhālabāso}$: You love yourself.

In each of the (a) sentences above, the subject and object noun phrases are coreferential with identical head noun, and are in the same simplex structure. So the reflexive rule reduces the object noun phrase into a reflexive pronoun, and in this way (4, 5b) are derived respectively from the underlying structure of (4, 5a). As (4, 5a) are simplex structures, we see that the reflexive rule operates well in a simplex structure if other conditions for the rule are satisfied.

We have said that a subject noun phrase reflexivizes an object noun phrase under coreferentiality when they are in the same simplex structure, but we have not said anything about the direction of the rule. This is because directions such as forwards and backwards are not relevant for the reflexive rule in Bengali. In a Bengali simplex structure the order of constituents is almost free: the subject can precede or follow the object in a simplex. But in a reflexive rule it is always the subject which reflexivizes the object. So what is important in reflexivization is not the direction, but the functional relations held by the noun phrases involved. Consider the examples below.

- (6) a. ketakī₁ke_c ketakī₂ bhālabāse₃: Ketaki loves Ketaki.
 1 2 3 2 3 1
- b. nijke₁ ketakī_c bhālabāse₃: Ketaki loves herself.
 1 2 3 2 3 1
- (7) a. tomāke₁ tumi_c bhālabāso₃: You love you.
 1 2 3 2 3 1
- b. nijke₁ tumi_c bhālabāso₃: You love yourself.
 1 2 3 2 3 1

We assume that (6, 7b) are derived from (6, 7a) respectively by application of the reflexive rule. In each example above, the object noun phrase precedes the subject noun phrase; but in each case the subject noun phrase reflexivizes the object noun phrase. So we see that the rule can apply in either direction subject to the condition that the subject noun phrase reflexivizes the object noun phrase. But an object noun phrase cannot be used to reflexivize a subject noun phrase in any circumstances. If we use an object noun phrase to reflexivize a subject noun phrase, we will derive sentences like those in (8), which are ungrammatical.

- (8) a. *nij_c ketakīke_c bhālabāse: *Herself loves Ketaki.
 b. *ketakīke_c nij_c bhālabāse:

We will state the reflexive rule in terms of the functional relations held by the noun phrases, not in terms of Langacker's (1969) 'Primacy Relations'. We have discussed the inadequacy of these relations elsewhere (cf., § 5.4.1).

Now we will consider whether the reflexive rule can apply beyond a simplex structure boundary. The following examples show that it does not.

- (9) a. ketakī_c mane kare ye hāsān tāke_c pachanda kare: Ketaki
 -2- 3 5 -6-
 thinks that Hasan likes her.
 2 3 6 5

- b. *ketakī_c mane kare ye hāsān nijke_c pachanda kare:
 *Ketaki thinks that Hasan likes herself.
- (10) a. ketakī_c hāsānke pachanda kare ā₃r hāsān ketakī_cke_c₄
 pachanda kare: Ketaki likes Hasan and Hasan likes
 Ketaki. ₇ ₃ ₄ ₇
- b. *ketakī_c hāsānke pachanda kare ā₃r hāsān nijke_c pachanda
 kare: *Ketaki likes Hasan and Hasan likes himself.
- (11) a. yadi mātin_c ā₃se, tabe nāju mātin_cke_c puraskār debe: If
₁ ₃ ₄ ₇ ₈ ₁
 Matin comes, then Nazu will give a prize to Matin.
₃ ₄ ₈ ₇
- b. *yadi mātin_c ā₃se, tabe nāju nijke_c puraskār debe: *If
 Matin comes, then Nazu will give a prize to herself.
- (12) a. ye-meyeti_c hāsānke ciṭhi likheche, hāsān tāke_c pachanda
₁ ₃ ₄ ₆ _—
 kare: Hasan likes the girl who has written a letter
₇ ₇ _{—1—} _{—4—} ₃
 to Hasan.
- b. *ye-meyeti_c hāsānke ciṭhi likheche, hāsān nijke_c pachanda
 kare: *Hasan likes herself who (the girl) has written
 a letter to Hasan.

In the examples (9)-(12) the (b) sentences, which are ungrammatical, are derived from the corresponding (a) sentences. In each of the (a) sentences above, there are two coreferential noun phrases which are not the constituents of the same simplex structure; and so the application of the reflexive rule generates ungrammatical sentences.

The reflexive rule does not apply at the deepest structure. This rule follows rules like subject and object marking, and 'Raising' (cf., Kiparsky and Kiparsky (1971), Postal (1974)). Postal (1971)⁹ notices that the reflexive rule applies neither at the deepest structure nor at the surface structure, but at

some intermediate structure. Postal (1971^a, 14) says:

It is important to emphasize that the clause mate constraint on reflexivization is applicable at neither the deep structure nor surface structure levels but rather at some point between them, this being defined by exactly that class of transformational rules which can apply in derivations before reflexivization. That the constraint is not relevant at the level of deep structure is shown by such examples as:

- 2.(4) a. I believe myself to be correct about that.
 b. Margaret found herself unable to move.

These surely must have deep structures fundamentally similar if not wholly identical to sentences respectively:

- 2.(5) a. I believe that I am correct about that.
 b. margaret found that she was unable to move.

But in these the coreferents are in different clauses.

That the clause mate condition is inapplicable to surface structures is trivially shown by a variety of examples like:

- 2.(6) a. Wash yourself.
 b. Who do you think Mary saw praising himself.

In Bengali we have sentences like those in (13):

- (13) a. meyeṭi_c nijke_c rūpasī mane kare: The girl thinks herself beautiful.
 b. mātin_c nijke_c buddhimān bhābe: Matin thinks himself to be clever.

The sentences in (13) are related respectively to those in (14), which are again related respectively to those in (15).

- (14) a. meyeṭi_c mane kare ye se_c rūpasī: The girl thinks that she is beautiful.

- b. mātin_c bhābe ye se_c buddhimān: Matin thinks that he is clever.
- (15) a. meyeṭi_c mane kare e-kathā, 'āmi_c rūpasī': The girl thinks this proposition, 'I am beautiful.'
- b. mātin_c bhābe e-kathā, 'āmi_c buddhimān': Matin thinks this proposition, 'I am clever.'

We consider that the examples in (13, 14, 15) are related: the sentences in (13) are derived from the underlying structure of the sentences in (15) through the intermediate structures (14). The sentences in (15) are direct discourse structures, which are transformed into the structures in (14) by the application of the indirect discourse formation rule (cf. § 5.8.1). In (15) and (14) the coreferential noun phrases are not in the same simplex structures, and so the reflexive rule cannot apply in these structures. If the reflexive rule applies to the sentences in (15), we will derive the ungrammatical sentences in (16); and if the rule applies to the sentences in (14), we will derive the ungrammatical sentences in (17).

- (16) a. *meyeṭi_c mane kare e-kathā, 'nij_c rūpasī': *The girl thinks this proposition, '(My)self is beautiful.'
- b. *mātin_c bhābe e-kathā, 'nij_c buddhimān': *Matin thinks this proposition, '(My)self is clever.'
- (17) a. *meyeṭi_c mane kare ye nij_c rūpasī: *The girl thinks that (her)self is beautiful.
- b. *mātin_c bhābe ye nij_c buddhimān: *Matin thinks that (him)self is clever.

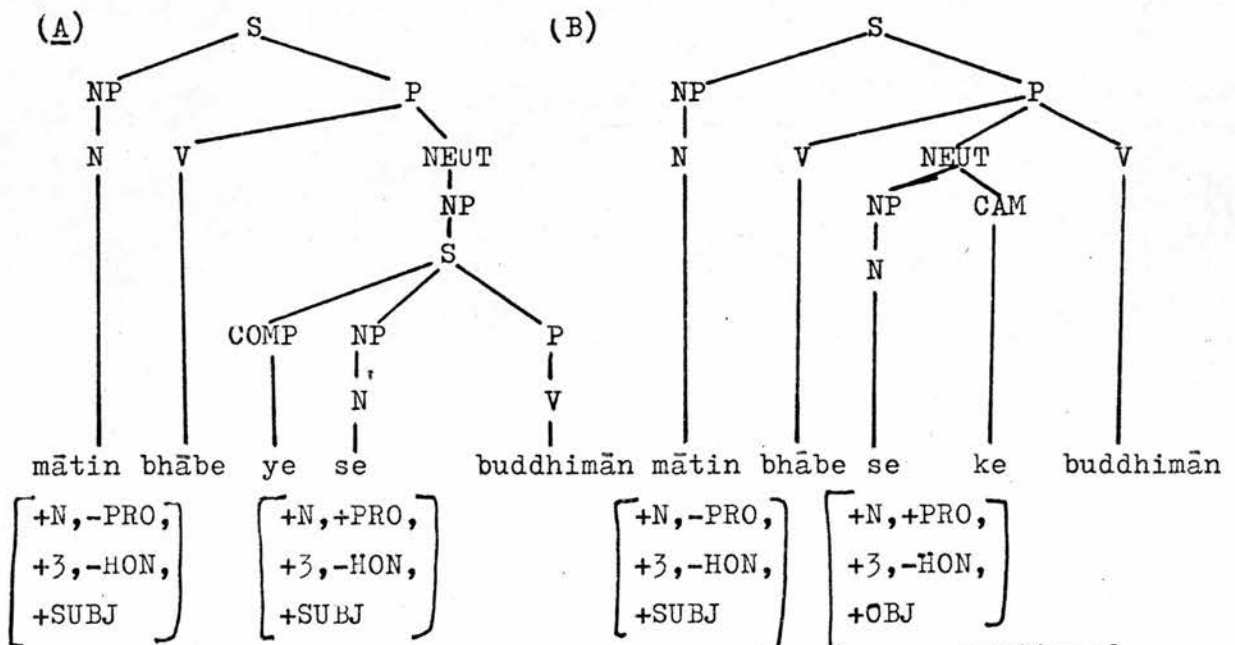
These sentences are ungrammatical because the reflexive rule has violated the simplex structure constraint. The reflexive rule has applied in these structures in two coreferential

noun phrases which are elements of two different structures. Then how do we derive the sentences in (13)? We derive the sentences in (18) by the application of the rule known as 'Raising'¹ on the structures that immediately underlie the sentences in (14).

(18) a. meyeṭi_c tāke_c rūpasī mane kare: The girl thinks her to be beautiful.

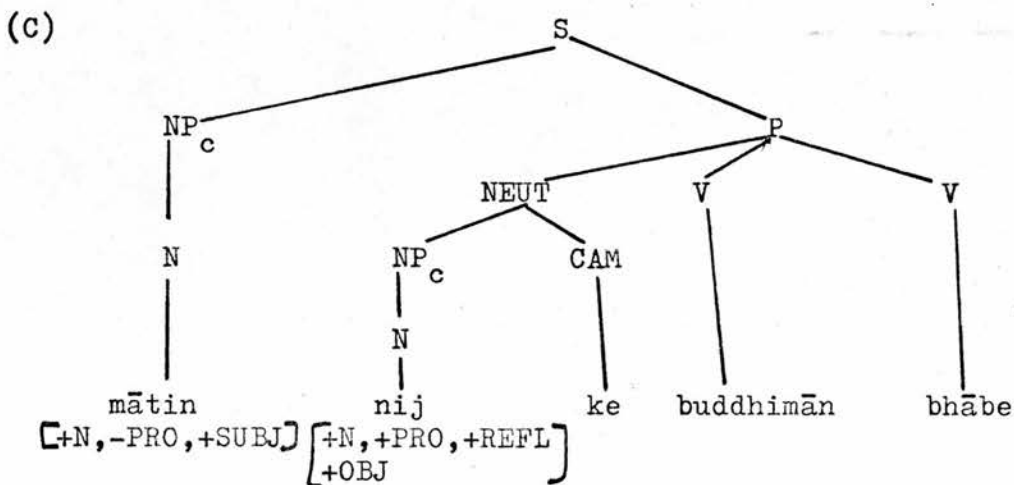
b. mātin_c tāke_c buddhimān bhābe: Matin thinks him to be clever.

1 The rule 'Raising' was first proposed by Kiparsky and Kiparsky (1971), and was later elaborated by Postal (1974). This rule raises the constituent subject and object noun phrases to the matrix sentence under certain conditions. For example, (14b) has the intermediate structure (A) (irrelevant details omitted) on which 'Raising' operates and raises the constituent subject noun phrase to the matrix sentence as an object, and derives the structure (B). The reflexive rule operates on (B), because the subject and object noun phrases in (B) are coreferential. Thus we derive (C), which generates (13b).



The rule 'Raising' has raised the constituent subject noun phrases of the sentences in (14) to the matrix sentences as objects, and the sentences in (18) have been derived. The subject and object noun phrases in each sentence in (18) are in the same simplex; and as they are coreferential the reflexive rule will specify the object noun phrases as [+PRO,+REFL], and the sentences in (13) will be produced. Thus we see that the coreferential subject and object noun phrases need not be in the same simplex structure in the underlying structure for the reflexive rule to operate; but they must be in the same simplex structure at the point the reflexive rule operates. This shows that the reflexive rule does not operate at the deepest structure. We can come to the following conclusions about the reflexive rule:

- (19) a. The subject noun phrase reflexivizes the object noun phrase if the head nouns are identical and coreferential.
- b. The subject and object noun phrases must be in the same simplex structure at the point the rule applies.
- c. The subject and object noun phrases need not be in the same simplex structure at the deepest level.



6.2 Reflexive Pronouns in Bengali.

There are two reflexive pronouns in Bengali: nij and āpan.² In modern Bengali āpan as a reflexive pronoun has a very limited use. So nij is commonly used as a reflexive pronoun. The reason for a limited use of āpan as a reflexive pronoun may be that it is lexically almost identical with the second person honorific pronoun āpani: 'You'. When case markers or plural markers are suffixed to āpan and āpani, they generate identical derived forms and create ambiguity. But there is no such problem with nij.

2 It should be mentioned here that there is a third reflexive item ātma: 'Self' in Bengali, but it is not used as a reflexive pronoun. It is used usually as a component in compound words like ātmakathā: 'One's own (self) story', ātmajībanī: 'One's own (self) biography': 'Autobiography', ātmahatyā: 'Self-killing': 'Commit suicide', etc., but it cannot be used as an object noun phrase. Consider the examples below.

- (A) a. se_c nijke_c hatyā kareche: He has killed himself.
 b. *se_c ātmake_c hatyā kareche:
 c. se ātmahatyā kareche: He has committed suicide (Lit., He has done self-killing).

In (Aa) the reflexive pronoun nij refers to the subject se, but (Ab) shows that ātma cannot be used as an independent noun phrase. In (Ac) ātma and hatyā have formed a compound word ātmahatyā: 'Self-killing'. Although (Aa) and (Ac) are semantically equivalent, they differ syntactically. In modern Bengali ātma has a fairly limited use, and it is predominantly used as the first component in reflexive compound words.

Consider the examples in (20).

- (20) a. $\bar{a}pani \bar{a}pan\bar{a}ke j\bar{a}nen n\bar{a}$: You do not $\left\{ \begin{array}{l} \text{you} \\ \text{yourself} \end{array} \right\}$.
- b. $\bar{a}mi \bar{a}pan\bar{a}ke j\bar{a}ni n\bar{a}$: I do not know $\left\{ \begin{array}{l} \text{you} \\ \text{myself} \end{array} \right\}$.

In (20) $\bar{a}pan\bar{a}ke$ may be taken either pronominally or reflexively; but in everyday use it will be taken as a pronoun in the above sentences. In order to avoid this problem $\bar{a}pan$ is rarely used as a reflexive pronoun in Bengali.

The Bengali reflexive pronouns have the feature [+ANI] in their feature bundle, and so the noun phrases involved in the reflexive rule should be animate. Reflexivization cannot take place in a structure where the coreferential subject and object noun phrases are inanimate. Although the pronouns have the feature [+ANI], reflexivization involving [-HUM,+ANI] noun phrases is not common. So the sentences in (21) are grammatical, but uncommon.

- (21) a. $ber\bar{a}l\bar{t}i_c nijke_c k\bar{a}m\bar{r}\bar{a}cche$: The cat is biting itself.
- b. $kukur\bar{t}i_c nijke_c dekhche$: The dog is looking at itself.

The 'uncommonness' of the above examples becomes odd, if we use $\bar{a}pan$ in place of nij , as in (22).

- (22) a. $ber\bar{a}l\bar{t}i_c \bar{a}pan\bar{a}ke_c k\bar{a}m\bar{r}\bar{a}cche$: The cat is biting $\left\{ \begin{array}{l} \text{you} \\ \text{itself} \end{array} \right\}$.
- b. $kukur\bar{t}i_c \bar{a}pan\bar{a}ke_c dekhche$: The dog is looking
at $\left\{ \begin{array}{l} \text{you} \\ \text{itself} \end{array} \right\}$.

The Bengali reflexive pronouns remain the same lexically for all persons, grades and gender. In English the reflexive pronouns agree with their antecedents in person, gender and number. In Bengali the reflexive pronouns change morphologically only for number. Consider the examples in (23).

(23) a. $\begin{bmatrix} \bar{a}mi \\ tumi \\ tini \end{bmatrix}_c$ nijke_c $\begin{bmatrix} j\bar{a}ni \\ j\bar{a}no \\ j\bar{a}nen \end{bmatrix}$ nā: $\begin{bmatrix} I \\ You \\ He \end{bmatrix}$ do(es) not know

$\begin{bmatrix} my \\ your \\ him \end{bmatrix}$ -self.

b. $\begin{bmatrix} \bar{a}mar\bar{a} \\ tomar\bar{a} \\ t\bar{a}r\bar{a} \end{bmatrix}_c$ nijder_c $\begin{bmatrix} j\bar{a}ni \\ j\bar{a}no \\ j\bar{a}nen \end{bmatrix}$ nā:

$\begin{bmatrix} We \\ You \\ They \end{bmatrix}$ do not know $\begin{bmatrix} our \\ your \\ them \end{bmatrix}$ -selves.

In (23a) the same form of the reflexive pronoun nij refers to singular first, second and third person antecedents, and in (23b) the same form nijder refers to plural first, second and third person antecedents. The Bengali reflexive pronouns agree with their antecedents in number only. In this respect they are similar to the Japanese reflexive pronoun zibun (cf., Kuno (1972)), and the Mohawk reflexive pronoun atat (cf., Postal (1970^b)).

6.3 Reflexivization is a Partly Optional Rule.

Reflexivization is a partly optional rule in Bengali.

Consider the examples in (24, 25).

(24) a. $\bar{a}mi$ $\bar{a}m\bar{a}ke$ $bh\bar{a}lab\bar{a}si$: I love me.
 1 2 3 1 3 2

b. $\bar{a}mi$ $nijke$ $bh\bar{a}lab\bar{a}si$:³ I love myself.
 1 2 3 1 3 2

3 The vowel /e/ is frequently inserted between the pronoun nij and anything that is suffixed to it. This phonological realization has not been taken into account in this work. The pronoun nij is pronounced as nija in some dialects of Bengali.

- (25) a. $\text{tumi}_1 \text{c} \text{tomāke}_2 \text{c} \text{bhālabāso}_3$: You love you.
 1 2 3 1 3 2
- b. $\text{tumi}_c \text{nijke}_c \text{bhālabāso}$: You love yourself.

In (24, 25a) the subject and object noun phrases are coreferential, identical in head noun, and are in the same simplex structures. The reflexive rule applies in these structures and derives (24, 25b) respectively from (24, 25a). The sentences in each set above are synonymous and equally acceptable. That they are synonymous can be seen from the anomaly of the following sentences.

- (26) a. $?āmi_c \text{āmāke}_c \text{bhālabāsi}$, $\text{tabe } āmi_c \text{nijke}_c \text{bhālabāsi nā}$: I love me, but I do not love myself.
- b. $?āmi_c \text{nijke}_c \text{bhālabāsi}$, $\text{tabe } āmi_c \text{āmāke}_c \text{bhālabāsi nā}$: I love myself, but I do not love me.

The anomaly of the sentences in (26) shows that the sentences in each set in (24, 25) are synonymous; and as each of them is acceptable, we can say that reflexivization is an optional rule in Bengali. But the rule is optional only when the noun phrases involved have either first or second person feature. If the noun phrases involved in the rule have the third person feature, the reflexive rule becomes obligatory. Consider the examples in (27).

- (27) a. $\text{ketakī}_c \left\{ \begin{array}{l} \text{ketakī} \\ \text{nij} \end{array} \right\}_c \text{ke bhālabāse}$: Ketaki loves

$\left\{ \begin{array}{l} \text{Ketaki} \\ \text{herself} \end{array} \right\}$.

- b. $\text{se}_c \left\{ \begin{array}{l} \text{tā} \\ \text{nij} \end{array} \right\}_c \text{ke bhālabāse}$: He loves $\left\{ \begin{array}{l} \text{him} \\ \text{himself} \end{array} \right\}$.

In (27) the coreferential noun phrases have the third person feature. In each example, however, the non-reflexivized sentences are ambiguous, but the reflexivized sentences have no ambiguity. In the reflexivized sentences the subject and object are understood coreferentially without any ambiguity, but in the non-reflexivized sentences the subject and object noun phrases may or may not be understood as coreferential in everyday use. So if the coreferentiality between the subject and object noun phrases in the sentences in (27) is intended in the surface, one would use the reflexive structures, but not the non-reflexive structures. This makes the reflexive rule obligatory in cases where the coreferential subject and object noun phrases have the third person feature.

6.4 Application of the Reflexive Rule.

The reflexive rule applies in a simplex structure when its subject and object noun phrases are coreferential and have an identical head noun. This rule requires that the noun phrase which will be reflexivized should be definite. This rule specifies the reflexivizable noun phrase as [+PRO,+REFL], and add these features to the feature bundle associated with the head noun of the reflexivized noun phrase. Besides these, the reflexive rule deletes all but the PL (if any) constituents of the D(eterminer) of the reflexivized noun phrase. This is because the reflexive pronouns do not take any determiner (except PL). This can be seen in (28).

- (28) a. *cheleṭi_c nijṭike_c bhālabāse: *The boy loves the himself.
 b. *meyeṭi_c sei nijṭike_c bhālabāse: *The girl loves that herself.

The sentences in (28) are ungrammatical because the reflexive pronoun occurs in these sentences with a CL, which Bengali does not allow.

The reflexive rule can be formalized as (29).

(29) The Reflexive Rule.

SI: $S [X \quad NP [X \quad N \quad X] \quad X \quad NP [X \quad N \quad X] \quad X]$
 1 2 3 4 5 6 7 8 9 10 11 12

SC: a. Specify 10 as [+PRO,+REFL].

b. Delete 9 and 11 if none is PL.

Conditions: a. 3 is the subject and 8 is the object.

They are coreferential, and 5=10.

b. 3 and 8 are in the same simplex structure
1.

c. 5 and 10 have the feature [+ANI] ,
preferably [+HUM] .

d. The rule is obligatory if 5 and 10 have
the features [-1,-2,+3] , optional
otherwise.

The rule will generate sentences like those in (30), and will disallow sentences like those in (31).

(30) a. mātin_c nijke_c praśaṅsā kare: Matin praises himself.

b. ekjan biplabī_c nijke_c bhālabesechilen: A rebel had
loved himself.

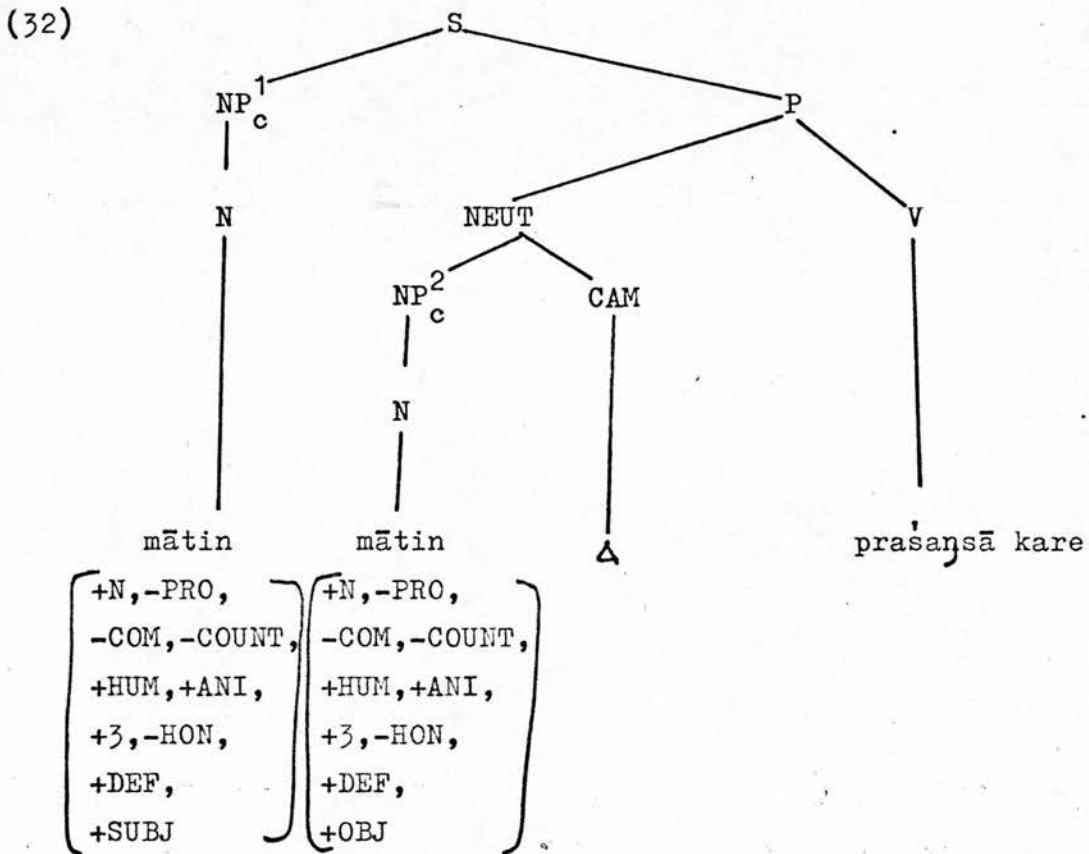
c. ketakī_c nijke_c rūpasī mane kare: Ketaki thinks herself
beautiful.

d. tārā_c nijder_c bhālabāse: They love themselves.

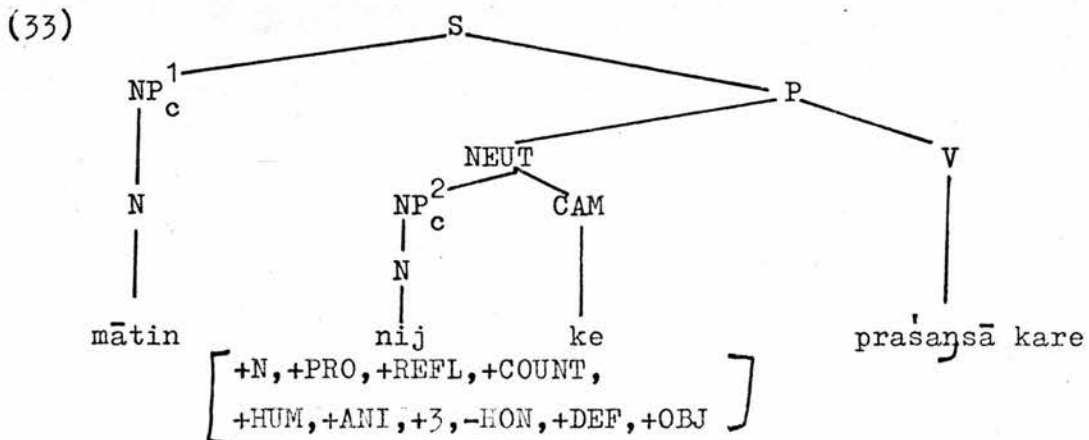
- (31) a. *āmarā nijke bhālabāsi: *We love myself.
 b. *āmi nijder bhālabāsi: *I love ourselves.
 c. *ekti bai nijke....: *A book... itself.

Now we will consider the application of the reflexive rule.

For example, (30a) has the immediate underlying structure (32).



In the simplex structure (32) the subject (NP¹) and object (NP²) are coreferential and identical in head noun. As (32) satisfies all the conditions for the reflexive rule, the rule will transform it into (33).



All the relevant transformations have applied to (33), and so we have arrived at the post-transformational string. The second lexical insertion rule has attached the reflexive pronoun nij to the reflexivized noun phrase. The terminal string of (33) will generate mātin nijke prasānsā kare, which is (30a).

6.5 Reflexivization and Coordination.

There is no ambiguity in a reflexive structure with singular noun phrases involved in the rule, but some ambiguity crops up in reflexive structures with plural and conjoined noun phrases. Consider the sentences in (34).

- (34) a. meyerā_c nijder_c bhālabāse: Girls love themselves.
 b. hāsān o ketakī_c nijder_c bhālabāse: Hasan and Ketaki love themselves.

The subject and object noun phrases involved in the reflexive rule are plural in (34a) and coordinate in (34b). These sentences are ambiguous. In one reading (34a) means that meyerā love nijder distributively, that is, each girl loves herself; and in another reading (34a) means that meyerā love nijder as a set, that is, each girl loves all other girls including herself. A similar sort of ambiguity can be seen in (34b), where the antecedent of the reflexive pronoun is a coordinate noun phrase. As the reflexive sentences with plural and coordinate noun phrases are ambiguous, they must be surface realizations of two distinct underlying sources. For example, (34b) should be related to hāsān nijke bhālabāse ebaṅ ketakī nijke bhālabāse: 'Hasan loves himself and Ketaki loves herself' for its distributive reading. In this sentence there is no

ambiguity. So we will assume that the noun phrases involved in reflexivization in sentences like (34) are not coordinate noun phrases in the underlying structure when they are understood distributively; and when they are understood in the set reading they are plural or phrasally conjoined in the underlying structure (cf., Lakoff and Peters (1966); ϕ 2.4.8; ϕ 5.6.1). In the next section we will consider this problem in more detail.

6.5.1 Reflexivization in Coordinate Noun Phrases and Derived Reflexives: Pronoun Conjunction.

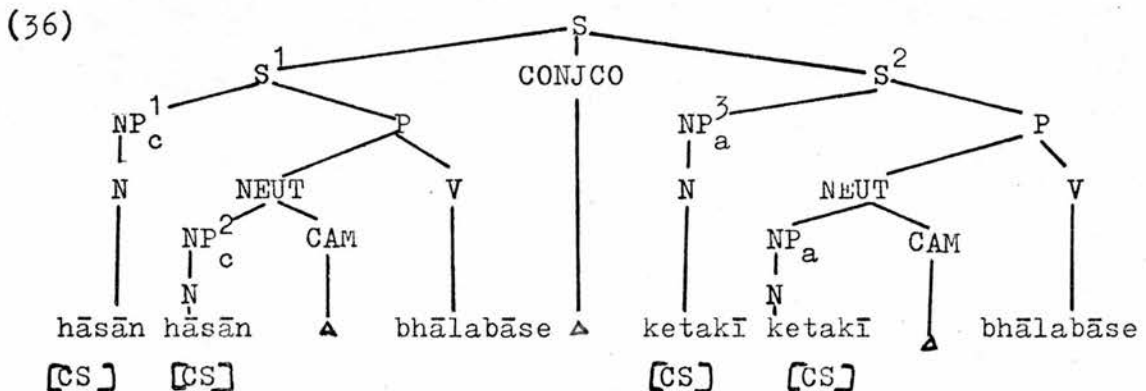
We have seen in the previous section that the surface reflexive sentences in which the noun phrases involved in the rule are plural or conjoined are ambiguous. It is possible that reflexivization operates between singular noun phrases in the intermediate structure, which undergo derived conjunction schema, and the noun phrases are realized as plural or conjoined noun phrases in the surface. It is also possible that the rule operates between underlying plural or conjoined noun phrases. So we assume that sentences like (34) are derived by derived conjunction for their distributive reading, and the noun phrases are plural or conjoined, as the case may be, in the underlying structure for the set reading of these sentences. We will consider reflexives sentences with conjoined noun phrases first for their distributive reading. Consider the examples below.

(35) a. hāsān_c hāsānke_c bhālabāse o ketakī_a ketakīke_a bhālabāse:

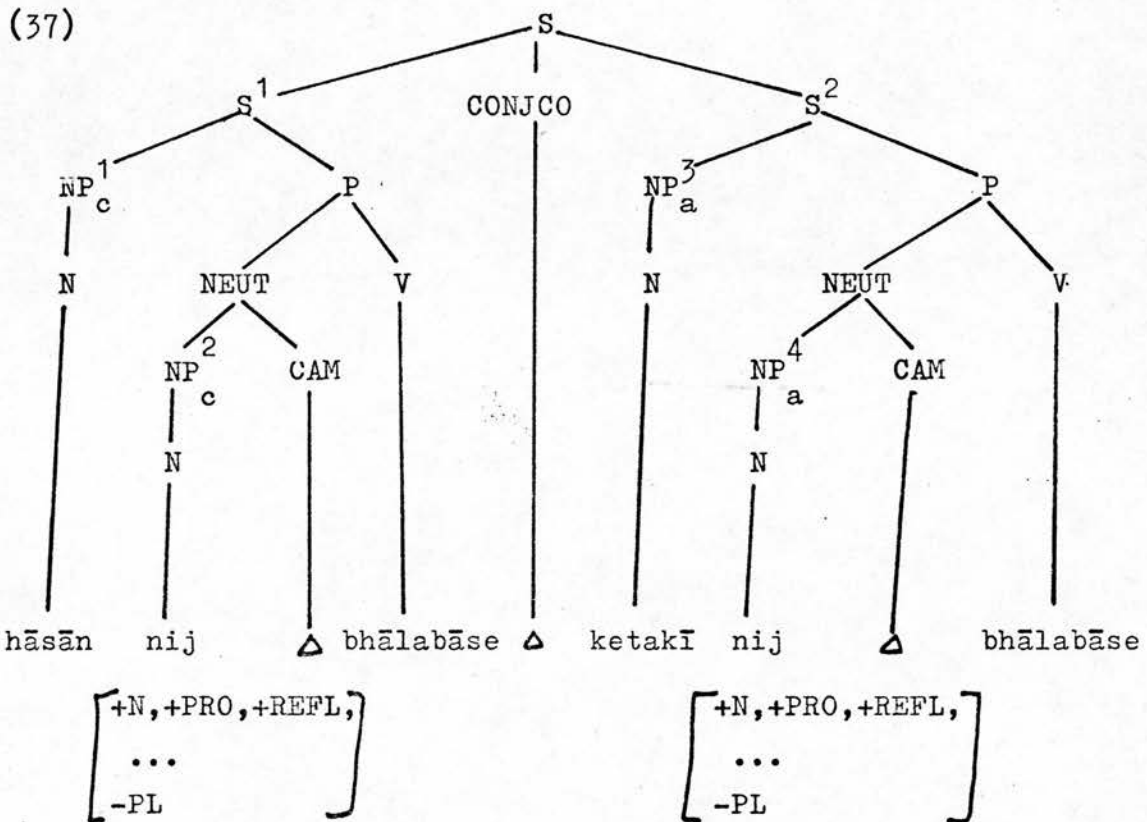
Hasan loves Hasan and Ketaki loves Ketaki.

- b. $\text{hāsān}_c \text{ nijke}_c \text{ bhālabāse}$ o $\text{ketaḱī}_a \text{ nijke}_a \text{ bhālabāse}$:
Hasan loves himself and Ketaki loves herself.
- c. $\text{hāsān}_c \text{ nijke}_c$ o $\text{ketaḱī}_a \text{ nijke}_a \text{ bhālabāse}$: Hasan loves himself and Ketaki herself.
- d. * hāsān_c o $\text{ketaḱī}_a \text{ yathākrame}$ nijke_c o $\text{nijke}_a \text{ bhālabāse}$:
Hasan and Ketaki love himself and herself respectively.
- e. hāsān_c o $\text{ketaḱī}_a \text{ nijder(ke)}$ bhālabāse : Hasan and Ketaki love themselves.

The surface sentence (35e) is ambiguous between a set reading and a distributive reading. We are concerned here with the distributive reading of the sentence. In this reading (35e) is derived from an underlying coordinate conjoined structure (35a), where there is no phrasal conjunction. (35a) is a coordinate conjoined structure, where each conjunct has a coreferential subject and object noun phrase. So the reflexive rule can apply in each conjunct generating (35b). (35c) is derived from (35b) by coordination deletion (cf., Koutsoudas (1971)). The structure (35d) is derived from (35c) by a derived conjunction schema which inserts yathākrame: 'Respectively' into the structure after conjoining the subject noun phrases (cf., Stockwell et al (1973, 396)). (35e) is derived from (35d) by yathākrame-deletion, and obligatory reflexive pronoun conjunction. (35a) has the immediate underlying structure (36).



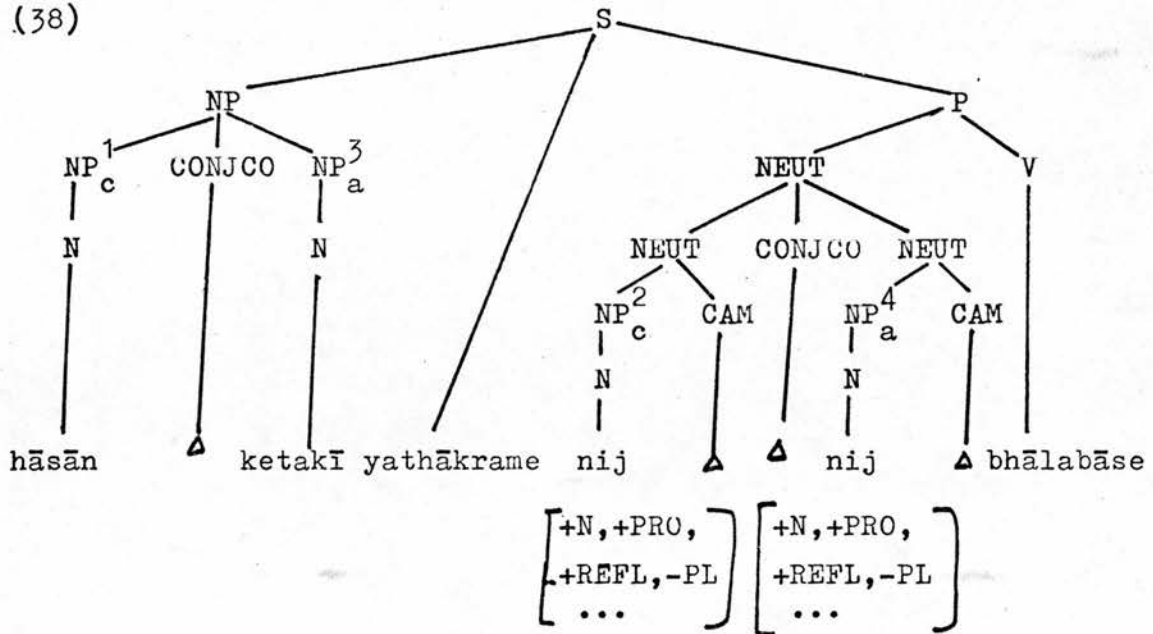
In (36) the subject and object noun phrases in each conjunct are coreferential and identical in head noun. So the reflexive rule can apply in each conjunct here, and its application will transform (36) into (37).



We have attached the reflexive pronoun to the reflexivized noun phrases in (37). This structure will generate (35a) after the second lexical insertion. But (37) is susceptible to further transformations: as the V's in both conjuncts are identical, we can delete the V of the first conjunct by coordination deletion and derive (35c). Instead the derived conjunction schema can be applied to (37). This will conjoin the antecedents into a single noun phrase, and the reflexivized noun phrases into a single noun phrase; and will insert yathākrame to the structure. These operations will convert

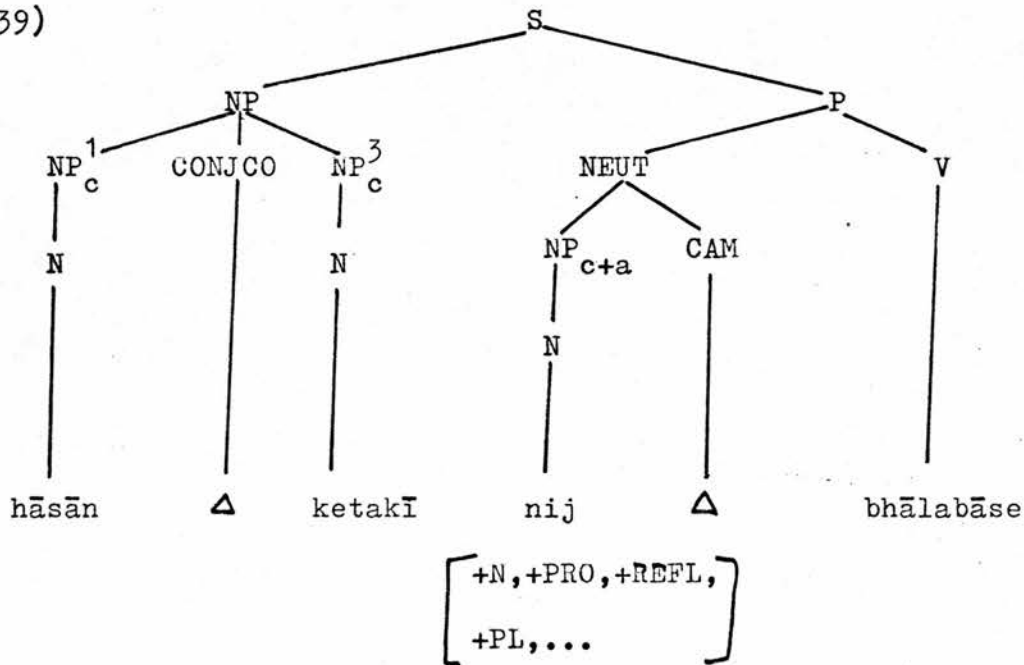
(37) into (38).

(38)



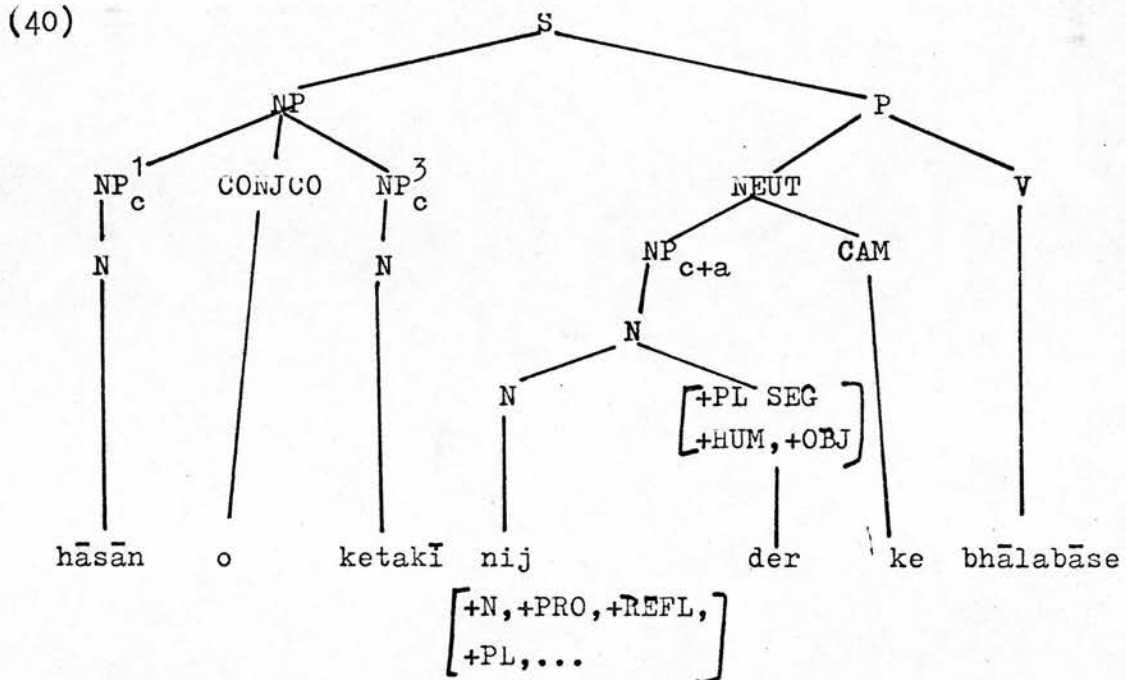
(38) will generate the strange sentence (35d) after second lexical insertion. So the pronoun conjunction rule will apply to (38). This rule will collapse the reflexivized noun phrases into a single noun phrase and will specify it with the feature $[+PL]$. It will delete yathākrame. These operations will derive (39) from (38).

(39)



In (39) the reflexivized noun phrases have been collapsed into

a single noun phrase which has been specified for the feature [+PL]. So the plural segment transformation will apply to this noun phrase, and this will transform (39) into (40).

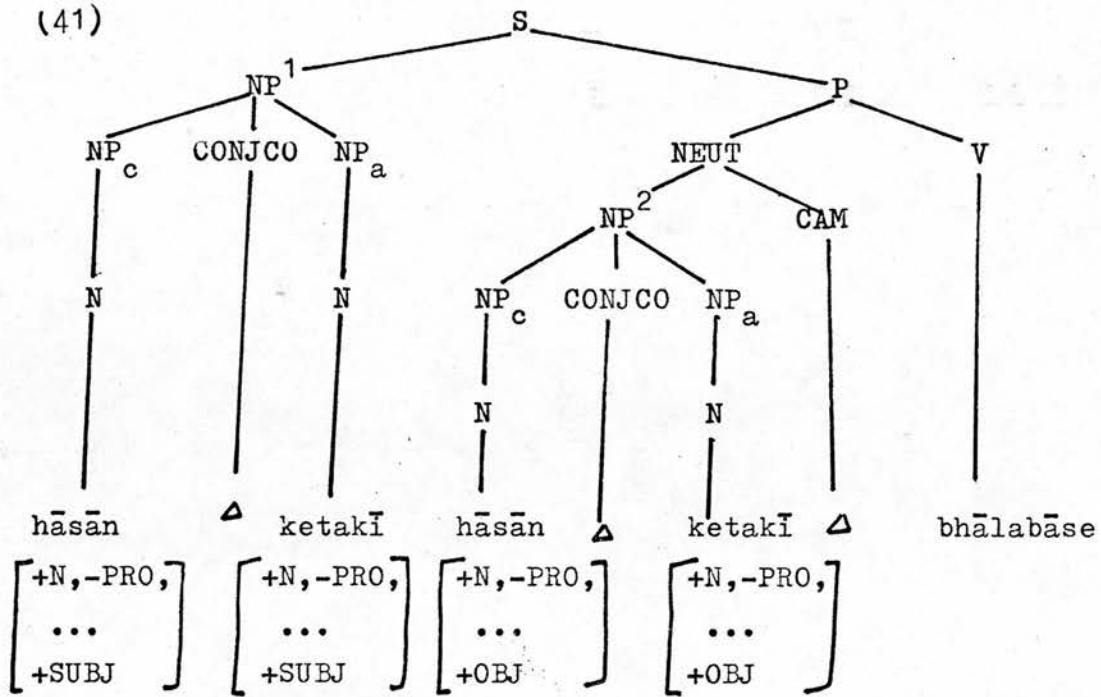


As we have arrived at the post-transformational string in (40), the second lexical insertion rule has attached appropriate lexical items to the derived nodes in (40). (40) will generate hāsān_c o ketakī_a nijderke_{c+a} bhālabāse, which is (36e).

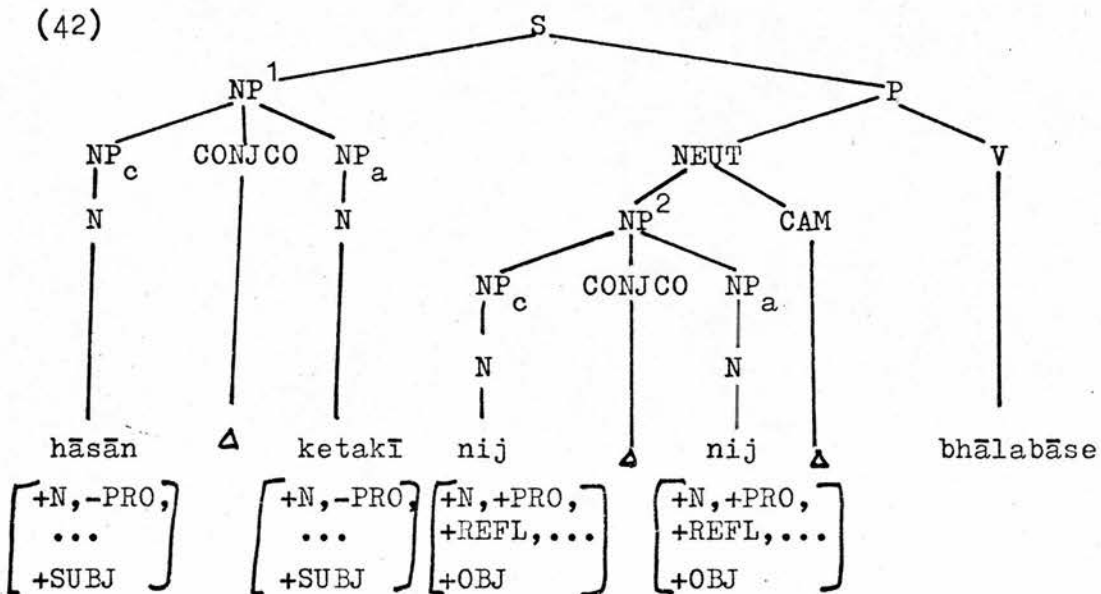
The object noun phrase in (40) can be objectivalized (cf., § 3.5).

This will delete the case marker from the reflexivized noun phrase, and will generate hāsān_c o ketakī_a nijder_{c+a} bhālabāse, which is a variant of (35e).

Now we will consider the derivation of (35e) in its set reading. For this reading of (35e), we assume that the noun phrases involved in the reflexive rule are phrasally conjoined in the underlying structure. (35e) has the intermediate structure (41) for its set reading.

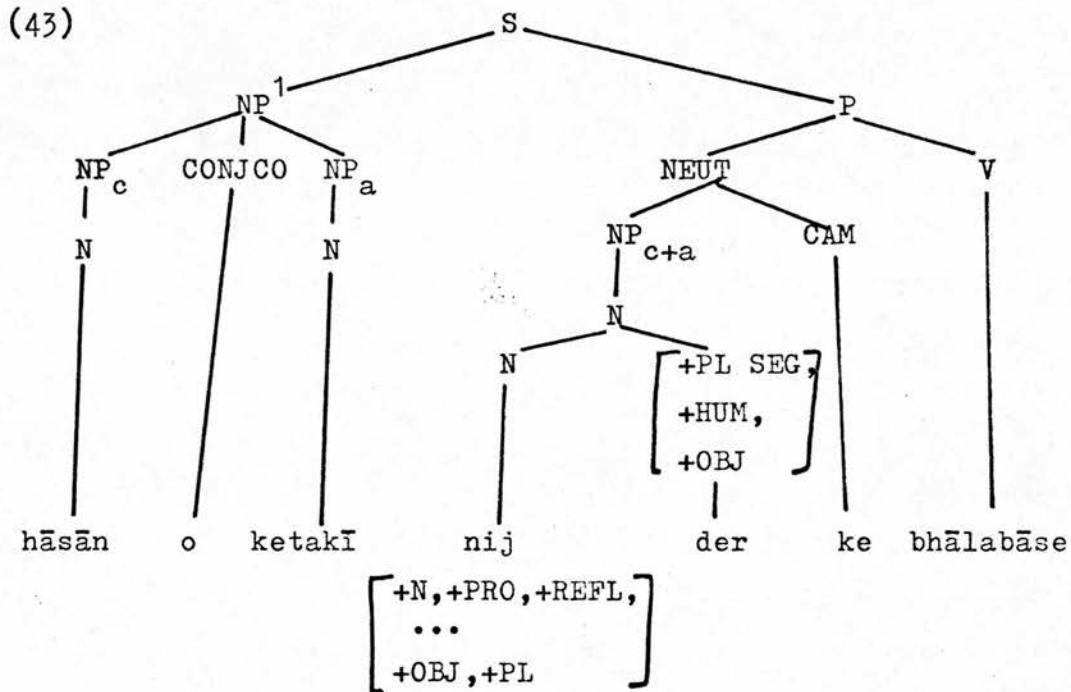


In (41) the subject and object noun phrases are conjoined noun phrases, and they are coreferential. As they are in the same simplex structure, the reflexive rule will specify the object noun phrase as [+PRO,+REFL] and will add these features to the complex symbol associated with the head noun of each noun phrase conjunct. The reflexive rule will transform (41) into (42).



(42) will generate the ungrammatical sentence hāsān_c ketakī_a nij_c nijke_a bhālabāse, after second lexical insertion. So the

pronoun conjunction rule will collapse the reflexivized noun phrase conjuncts dominated by NP² into a single noun phrase, and the noun phrase will be specified as [+PL]. After the conjunction of the pronouns, a plural segment transformation will take place in the reflexivized noun phrase; and we will derive (43).



(43) will generate (35e), after the second lexical attachment. Thus we see that the ambiguous sentence (35e) has two sources, and this explains the ambiguity of this sentence.

The pronoun conjunction rule is necessary because the coordination of identical pronouns is unacceptable in Bengali (cf., § 5.6.2). This rule applies after the derived conjunction schema and after reflexivization in a coordinate noun phrase. Consider the examples below.

- (44) a. *ketakī_c o hāsān_a nijke_c o nijke_a bhālabāse: Ketaki and Hasan love herself and himself.
- b. ketakī o hāsān nijder bhālabāse: Ketaki and Hasan love themselves.

- (45) a. *se_a tumi_bo āmi_c nijke_a nijke_bo nijke_c bhālabāsi: He,
you and I love himself, yourself and myself.
- b. se_a tumi_bo āmi_c nijder_{a+b+c} bhālabāsi: He, you and I
love ourselves.

In (44, 45a) the reflexive pronouns are coordinately conjoined, which makes these sentences ungrammatical; but in (44, 45b) they have been collapsed into a single plural reflexive pronoun. These sentences are grammatical. So when a noun phrase conjoins the reflexive pronouns, it will be obligatorily telescoped into a single plural reflexive pronoun.

6.6 Reflexivization versus Emphasis.

In some situations the reflexive pronoun nij (and in some situations āpan) is used as an emphatic morpheme. These sentences create problems for a clear cut analysis of reflexivization. Here we will consider the emphatic use of nij only. Consider the examples below.

- (46) a. bādāl nījē eseche: Badal has come himself.
 2 3 3 2
- b. *bādāl nij eseche:
- (47) a. āmi nījē tāke balechi: I have told him myself.
 1 2 3 4 1 4 3 2
- b. *āmi nij tāke balechi:

In (46, 47a) the item nij is used to emphasize the subject. Here it obligatorily takes the subject marker e, which is rarely used in modern Bengali. (46, 47b) are ungrammatical due to the absence of e. In these sentences nij has been used to emphasize the subject, and so we can call it the 'Subject Emphatic' (cf., Moyne (1971)). The subject emphatic nij emphasizes a subject, but the reflexive pronoun nij is used as an object which refers anaphorically to the subject. The emphatic nij

can be used to emphasize a subject, but not an object. This can be seen in (48).

(48) a. **mātin nājūke nījēke dekheche*: Matin has seen Nazu herself.

b. **āmi tāke nījēke balechi*: I have told him himself.

(48) are ungrammatical because nij has been used to emphasize object noun phrases.

Although the subject emphatic nij usually follows the subject, it can be removed away from the subject, and other constituents may intervene between them. Yet it is understood as emphasizing the subject as in (49).

(49) a. *āmi tomāke nījē balechi*: I have told you myself.
 1 2 3 4 1 4 2 3

b. *āmi tāke baiṭi nījē diyechi*: I have given him the book myself.
 1 2 3 4 5 1 5 2 3
 4

In (49) the emphatic nij has been removed away from the subject, yet it emphasizes the subject.

That nij has been used as a subject emphatic in the above examples will be clear from the following examples, where reflexivization has taken place and the subjects have been emphasized as well:

(50) a. *āpani_c nījē nījke_c praśna karun*: You yourself ask yourself.

b. *āmi_c nījē nījke_c bārbār praśna karechi*: I myself have asked myself once and again.

In (50) reflexivization has taken place and the subjects have been emphasized as well. So nij in the above examples should be considered as a subject emphatic. The reflexive pronouns are derived transformationally, but the emphatic nij cannot be

derived by transformational reduction of any noun phrase. It should be generated either in the underlying structure or should be inserted transformationally to a structure for the purpose of emphasis.

Another difference between the emphatic nij and the reflexive nij is that the emphatic nij can be used to emphasize subjects of transitive as well as intransitive verbs, but the reflexive pronoun occurs always with transitive verbs. So we find that not all structures with nij involve reflexivization.

6.6.1 Possessive Emphatic nij.

nij can be used as a possessive emphatic morpheme.

Consider the examples in (51, 52).

(51) a. āmār bai: My book.

b. āmār nijer bai: My own book.

(52) a. mātiner beṛāl: Matin's cat.

b. mātiner nijer beṛāl: Matin's own cat.

In the possessive noun phrases (51, 52a) the possessors have taken the genitive case form; and in (51, 52b) the possessors have been emphasized by the genitive form of the morpheme nij. The examples in each set are cognitively synonymous; the only difference between them is that in the (b) examples the possessors have been emphasized by nijer, but they are not emphasized in the (a) examples. As the examples in each set above differ only in respect of emphasis, we can consider nij as a possessive emphatic. Moyne (1971, 149) claims that the Persian kitab-e xodam: 'My own book', ketab-e xodas: 'His own book', and the English 'My own', 'Your own' etc., are instances of possessive emphatic. We will also consider

āmār nijer: 'My own', tār nijer: 'His own', mātiner nijer: 'Matin's own' etc., as emphatic variants of āmār: 'My', tār: 'His', and mātiner: 'Matin's' respectively. When a possessor is emphasized by nij, it takes a genitive form like the possessor, and suggests an extreme sense of possession.

6.6.2 Reflexivization, Emphasis and the 'Picture Nouns'.

In this section we will consider some constructions with nij which create real problems for a straight-forward derivation of the reflexive structures. Postal (1971^a), Jackendoff (1972) and Stockwell et al (1973) notice that the constructions with 'Picture Nouns' do not obey the simplex structure constraint. Such constructions are exemplified below.

- (53) a. Tom believes there is a picture of himself hanging in the post-office.
- b. Unflattering descriptions of himself have been banned by LBJ.

English constructions like 'picture of himself', 'a story about myself', etc., are known as 'picture reflexives' or 'picture noun nominalizations' (cf., Postal (1971^a, 12)). These constructions violate the same simplex structure constraint required for reflexivization. Although it is generally assumed that they differ from the genuine reflexive structures in many respects (cf., Stockwell et al (1973, 203-05)), Jackendoff (1972, 132-42) has tried to deal with them by the same rule that deals with the genuine reflexives. Postal (1971^a, 12) says:

Jackendoff's assumption that all instances of reflexive forms follow from a unitary operation is not only logically unsound but factually incorrect.

There are two types of structure with nij which seem to involve reflexivization, but are not handled by our reflexive rule. One type involves picture nouns like chabi: 'Picture', galpa: 'Story', kahinī: 'Tale' etc.,; and the other type involves 'benefactive' constructions. They are exemplified below.

- (54) a. e-ṭi āmār_cnij_cchabi: { This is a picture of myself.
This is my own picture.
- b. āmi_cāmār_cnij_cjanya ekṭi bāṛī bāniyechi: I have built a house for my (own)self.

(54a) is ambiguous between a reflexive and a possessive emphatic reading, but there is no ambiguity in (54b): it is understood in the possessive emphatic sense only. The picture nouns in Bengali create ambiguity when used in a possessive noun phrase. Consider the examples in (55).

- (55) a. e-ṭi ki āpanār_cnij_cchabi?: Is this
1 2 3 4 5 1
{ your own picture?
3 4 5 }
{ a picture of yourself? }
5 3,4
- b. hāsān_cketakīke pratidin tār_cnij_cgalpa 'sonāy: Hasan
3 4 5 6 7
tells his { own stories
7 4 { 5 6 } to Ketaki
{ stories about himself }
everyday. 6 —5—
3

Each sentence in (55) is ambiguous at least in two ways as shown in the gloss. This sort of ambiguity cannot be found in similar constructions with 'non-picture nouns'. Consider the

examples in (56).

- (56) a. e-ti ki āpanār nijer kukur?: Is this your own dog?
 b. nīlimār nijer sārī: Nilima's own sari.

The sentences in (56) have only one reading: the possessive emphatic reading. But ambiguity crops up with nouns like chabi: 'Picture', galpa: 'Story' etc., in similar constructions, because pictures, stories etc., can be based on the same person who owns and/or creates them.

Before we go into the derivations of the sentences in (55), we will look into some of their syntactic facts. In Bengali the picture reflexives such as āpanār nijer chabi in (55a), and the possessive emphatics such as āpanār nijer kukur in (56a) have identical structures in the surface. Their syntactic facts are summarized in (57).

- (57) a. They both have the structure GEN₁ GEN₂ NP (GEN stands for genitive).
 b. GEN₂ has the head item nij, which either refers to, or emphasizes GEN₁.
 c. GEN₁ is obligatory and GEN₂ is optional. But if GEN₂ is deleted when GEN₁ has third person feature and can be understood coreferentially with any noun phrase other than the subject of its own simplex, the sentence will be ambiguous. GEN₁, however, is deletable if (a) GEN₁ is coreferential with the subject of its own simplex, and (b) GEN₁ is not an element of a constituent sentence.

Now we come back to (55). If we delete GEN₂ (nijer) from these sentences, we will get (58) respectively.

- (58) a. e-ṭi ki āpanār chabi?: Is this $\left\{ \begin{array}{l} \text{your picture?} \\ \text{a picture of yourself?} \end{array} \right.$

- b. hāsān ketakīke pratidin tār galpa 'sonāy:

Hasan tells his stories to Ketaki everyday.

(58a) is still two-way ambiguous like (55a), but (58b) is ambiguous in at least four ways: (a) Hasan tells stories composed by him, (b) Hasan tells stories about Ketaki, (c) Hasan tells stories about someone other than Ketaki and Hasan, and (d) Hasan tells stories about himself.

In an analysis of such structures, GEN₁ demands particular attention because it is pivotal in the anaphoric relations in such structures. In these structures GEN₂ (nijer) refers to its antecedent through GEN₁, which is its immediate antecedent.

Consider the examples below.

- (59) a. āmi_{1c} tomāke₂ āmār_{3c} nijer_{4c} sambandhe ekṭi galpa balbo:
 1 2 3 4 5 6 7 8
 I shall tell you a story about myself (my own self).
 1 8 2 6 7 5 3 4

- b. āmi_c tomāke āmār_c sambandhe ekṭi galpa balbo: I shall tell you a story about myself (me).

- c. āmi_c tomāke nijer_c sambandhe ekṭi galpa balbo: I shall tell you a story about (my)self.

- (60) a. āmi tomāke_c tomār_c nijer_c sambandhe ekṭi galpa balbo:
 I shall tell you a story about yourself (your own self).

- b. āmi tomāke_c tomār_c sambandhe ekṭi galpa balbo: I shall tell you a story about yourself (you).

- c. *āmi tomāke_c nijer_c sambandhe ekṭi galpa balbo:

- (61) a. hāsān_c mane kare ye gaṇabhabane tār_c nijer_c ekṭi chabi āche:
 Hasan thinks that there is a picture of himself in the Ganabhaban.

- b. hāsān_c mane kare ye gaṇabhabane tār_c ekṭi chabi āche:
 Hasan thinks that there is a picture of himself
 (his) in the Ganabhaban.

c. * hāsān_c mane kare ye gaṇabhabane nijer_c ekṭi chabi āche:
 In (59a) GEN₂ refers to GEN₁, which is coreferential with the subject; and in (59b) GEN₂ is deleted, but it is understood synonymously with (59a). As GEN₁ in (59a) is coreferential with the subject, it can be deleted; and we get (59c) by deleting GEN₁. In (59c) GEN₂ (nijer) refers to the subject. In (60a) GEN₁ is non-coreferential with the subject, but GEN₂ refers to GEN₁. Here we can delete GEN₂ and derive (60b); but we cannot delete GEN₁. If we delete GEN₁ from (60a), we derive (60c), which is ungrammatical. In (60c) GEN₂ (nijer) cannot refer to the object, but it refers to the subject. In (61a) the picture reflexive elements are in the constituent sentence; and here GEN₁ refers to the matrix subject hāsān, and GEN₂ refers to GEN₁. Here, too, we can delete GEN₂, but not GEN₁. If we delete GEN₁ we will derive the ungrammatical sentence (61c); and if we delete GEN₂, we will derive (61b). In (61c) GEN₂ (nijer) does refer to the matrix subject, because there is no other noun phrase in this sentence which can be taken as its antecedent; but the sentence is syntactically unacceptable.

We have seen that the picture reflexives and the possessive emphatic constructions are identical in the surface. There is not much of a problem with the possessive emphatic constructions. We have claimed that possessive emphatic constructions like āmār nijer: 'My own', tomār nijer: 'Your own', etc., are emphatic variants of āmār: 'My' and tomār: 'Your'

respectively (cf., § 6.6.1). The real problem lies with the so-called picture reflexives. The noun phrases like āmār chabi: 'My picture'/'Picture of myself', where there is no reflexive pronoun, are equally ambiguous with the noun phrases like āmār nijer chabi: 'My own picture'/'Picture of myself', which contain the reflexive pronoun nij. The benefactive constructions like āmār nijer janya: 'For myself (my own)' are similar to the possessive emphatic constructions in that they can be considered as variants of forms like āmār janya: 'For myself (me)'. But the picture reflexives constitute a problem for our reflexive rule. It is clear that that the subject does not reflexivize the object in these constructions, although the subject has a major part to play in them. In order to deal with the picture reflexives, we can propose that reflexivization takes place in a 'simplex noun phrase structure' containing two coreferential noun phrases in these sentences (cf., Stockwell et al (1973, 169)). But this proposal cannot solve the problem satisfactorily, because of the peculiar inherent meaning and characteristics of the picture nouns. Consider the peculiarities of the two picture nouns galpa, and chabi in (55). Although a paraphrase in which two noun phrases are coreferential can be given for (55b), a similar paraphrase for (55a) will be unnatural. Consider the paraphrases of (55) given respectively in (62).

- (62) a. *e-ṭi ki āpanār_cāpanār_cchabi?: Is this your your picture?
- b. hāsān_cketakīke pratidin hāsāner_chāsān_c-samparkita galpa¹sonāy: Hasan tells Hasan's stories about Hasan to Ketaki everyday.

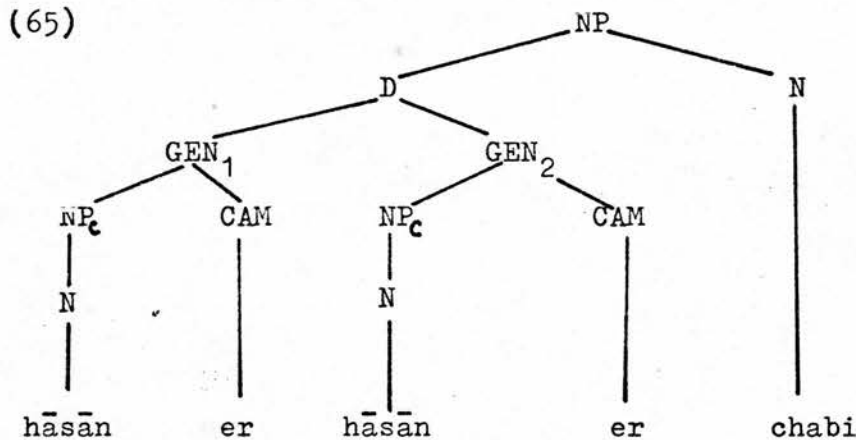
(62a) is unnatural. We can take GEN₁ in this sentence either reflexively or as a possessor noun phrase, but GEN₂ does not make any sense. The genitive noun phrases in Bengali are usually ambiguous, and the difficulty with them is that they are not paraphrasable into structures where the possessor is a non-genitive noun phrase and the structure is lexically related to the surface genitive noun phrase. Consider the examples below.

- (63) a. $\bar{a}m\bar{a}r$ $chabi:$ $\left\{ \begin{array}{l} \text{My picture.} \\ \text{Picture of myself.} \end{array} \right.$
- b. $chabi\dot{t}i, y\bar{a} \bar{a}m\bar{a}r:$ This picture that is $\left\{ \begin{array}{l} \text{mine} \\ \text{of myself.} \end{array} \right.$
- c. $\bar{a}mi$ $chabi\dot{t}ir$ $m\bar{a}lik:$ I am the owner of the picture.
- d. $?chabi\dot{t}i \bar{a}m\bar{a}r$ $ceh\bar{a}r\bar{a}r:$ The picture is of my face.
- (64) a. $\bar{a}m\bar{a}r$ $nijer$ $chabi:$ $\left\{ \begin{array}{l} \text{My own picture.} \\ \text{Picture of myself.} \end{array} \right.$
- b. $chabi\dot{t}i, y\bar{a} \bar{a}m\bar{a}r$ $nijer:$ The picture that is $\left\{ \begin{array}{l} \text{my own.} \\ \text{of myself.} \end{array} \right.$
- c. $\bar{a}mi$ $nije$ $cabi\dot{t}ir$ $m\bar{a}lik:$ I myself am the owner of the picture.
- d. $?cabi\dot{t}i \bar{a}m\bar{a}r$ $nijer$ $ceh\bar{a}r\bar{a}r:$ The picture is of my own face.

(63, 64a) are two-way ambiguous, and their relative clause paraphrases (63, 64b) are similarly ambiguous. They can be disambiguated by structures like (63c, d) and (64c, d), which are far removed from their surface structures in lexical

material. But notice that (64d) still contains the reflexive pronoun nij. This shows however deep we delve, we cannot avoid nij in a construction with the noun chabi. That is, nij in such a construction cannot be considered as a pronoun derived by the reflexive rule. But in a similar construction with the picture noun galpa, nij can be considered as a reflexive pronoun derived by the reflexive rule.

If we consider that the picture reflexives are derived transformationally, we assume that they have a structure something like (65) with nouns like chabi: 'Picture':

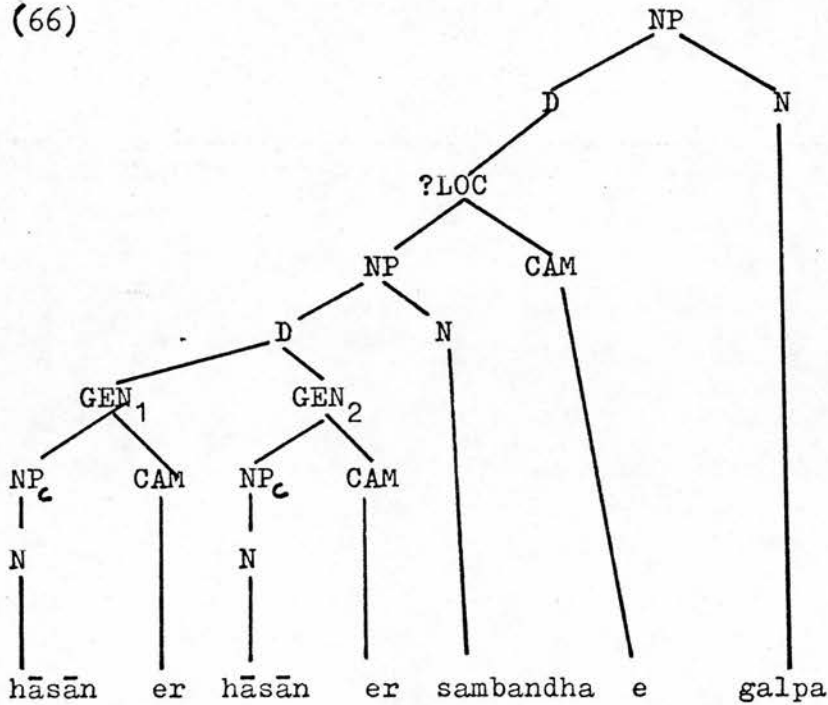


In (65) the noun phrases dominated by GEN_1 and GEN_2 are coreferential; and the reflexive rule will use the left noun phrase to reflexivize the right noun phrase. We have said that GEN_2 is optional; but when it dominates a noun phrase coreferential with the noun phrase dominated by GEN_1 , we will get picture reflexives. Instead, GEN_2 might have nij as a possessive emphatic morpheme, which will generate a possessive emphatic construction.

The intermediate structure (65) is appropriate for a picture reflexive construction with the noun chabi: 'Picture', but not with galpa: 'Story'. We assume that an intermediate structure for a picture reflexive noun phrase construction

with the noun galpa will be something like (66).

(66)



In (66) GEN₁ and GEN₂ are coreferential. The reflexive rule can apply here forwards and forwards only. The rule will derive a noun phrase structure like hāsāner nijer sambandhe galpa: 'Hasan's stories about himself' from (66). The structure allows the deletion of the ?LOC optionally. When it is deleted, we derive a structure like (65), which will produce a noun phrase like hāsāner nijer galpa: 'Hasan's own stories'/'Stories about Hasan'. The above is only a tentative solution. It is hoped that if it has failed to solve the problem, it has at least raised the problems of picture reflexives and their transformational derivation.

in (2) do not contain any such item, yet these sentences are understood as reciprocal. The items eke anya and ekjan anyajan superficially look like pronouns which refer to some antecedent involved in a reciprocal relation, but we will later show that they are not pronouns, and are not even a single constituent. The item paraspar, which is tatsama (Sanskritic) compound word, may be considered as a reciprocal item. In (2) reciprocity is expressed by the items mārāmāri, hātāhāti and kānākāni, which are morphologically derived by reduplication in order to express reciprocity.

7.1 The Semantics of Reciprocity.

The reciprocal structures indicate a reciprocal relation (active or stative) usually between two persons, but this relation can be established among an indefinite number of people in reciprocal structures. Consider the examples below.

- (3) a. tārā eke anyake bhālabāse: They love one another.
 b. tārā parasparke bhālabāse:
 (4) a. tārā mārāmāri karche: They are hitting (one another).
 b. tārā hātāhāti karche: They are striking (one another) by hand.

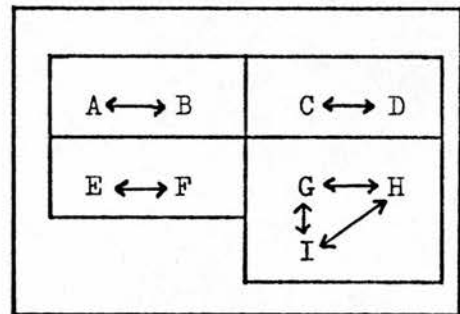
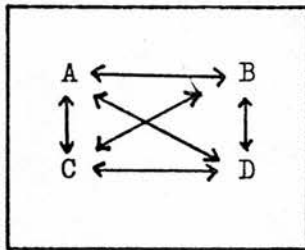
If tārā in (3) refers to two people, then these sentences are understood as that each one of them loves and is loved by the other one; and if tārā refers to three or an indefinite number of people, then these sentences are understood as that there is a mutual relation of love among all of them. But the situation is not so clear in (4). If tārā in (4) refers to two people, then these sentences are understood in similar fashion to those in (3) in that each one hits and is hit by the other one. But

if tārā refers to an indefinite number of people, and if these sentences are to be understood in a 'strict reciprocal sense' then we will assume that each member hits the other members and is hit by the other members. But tārā in (4) when refers to an indefinite number of people, these sentences are not usually understood in the 'strict reciprocal sense', but are understood, let us say, in the 'vague reciprocal sense'.

If a brawl is going on among a group of people, then that situation can be described by (4a, b) in the 'vague reciprocal sense'. A 'strict reciprocal sense' can be shown, for example, diagrammatically as (5a), and a 'vague reciprocal sense' by (5b).

(5) a. 'Strict Reciprocal'

b. 'Vague Reciprocal'



In (5a) A, B, C and D each maintains a reciprocal relation with the other(s) in the group, and so their relation is strictly reciprocal; but in (5b) A, B, C, D, E, F, G, H and I, each does not maintain a reciprocal relation with every other member in the whole group, although in each sub set in (5b) each member maintains a strict reciprocal relation with the other. The relation, if it is active, among all the members of (5b) can be vaguely expressed as a reciprocal relation by sentences like (4a, b). We will assume that a reciprocal sentence can have two interpretations: (a) a strict reciprocal

interpretation, and (b) a vague reciprocal interpretation. The interpretation is always strict if the relation holds between two members, but the relation can be either vague or strict if the relation holds among three or indefinitely more members. Bengali usually utilizes structures of the type exemplified in (3) if a reciprocal sentence is intended in the strict sense; and utilizes structures of the type exemplified in (4) if the sentence is intended in a vague reciprocal sense.

A reciprocal structure demands a reciprocal relation between /among the members involved; but if the relation is active and the antecedent refers to an indefinite number of people, then the relation cannot logically hold reciprocally, and hence a vague interpretation results. Consider the examples below.

- (6) a. rām, rahim, yadu o madhu parasparke khun kareche:
 Ram, Rahim, Jadu and Madhu have killed one another.
- b. bāṅālīrā parasparke khun karbe: The Bengalees will
 kill one another.

The sentences in (6) are syntactically well-formed, but they cannot be understood in the strict reciprocal sense. If we take (6a) in the strict reciprocal sense, then the sentence will mean that each member have been killed three times, which is logically impossible. Similarly, each Bengalee in (6b) cannot kill and be killed by all other Bengalees, and so this sentence will be understood in the vague reciprocal sense.

The interpretation which we call vague, can be called 'ambiguous', but it is better and in fact accurate to call it vague. The reciprocal relation, although it holds vaguely if the relation is active, can hold strictly among an indefinite number of people

if the relation is stative. Consider the examples in (7).

- (7) a. *tārā parasparke bhālabāse*: They love one another.
 b. *bāṅālīrā parasparke īrṣā kare*: The Bengalees envy one another.

The relations of love and jealousy are stative, and so the sentences above can be understood in the strict reciprocal sense. But they can also be understood in the vague reciprocal sense. Reciprocity is a symmetric relation, and so it cannot hold when the predicate indicates asymmetry. This can be seen in the comparative constructions below.

- (8) a. **mātin o minu eke anyer ceye buddhimān*: Matin and Minu each is cleverer than the other.
 b. **ketakī o nīlimā parasparer ceye rūpasī*: Ketaki and Nilima each is more beautiful than the other.

The sentences in (8) are syntactically well-formed, but logically impossible, and semantically anomalous. (8a) means that mātin is cleverer than minu, and simultaneously minu is cleverer than mātin, which is impossible. Similarly, (8b) means that ketakī is more beautiful than nīlimā, and at the same time, nīlimā is more beautiful than ketakī, which is entirely anomalous.

The reciprocal relation is usually a simultaneous relation (cf., Fiengo and Lasnik (1973, 450-01): if X has some relation with Y, Y must have the same relation with X at the same time. Consider the examples in (9).

- (9) a. *minu mātinke sanibār mereche ār mātin minuke budhbār mereche*: ?*minu o mātin parasparke mereche*: Minu (has) hit Matin on saturday and Matin (has) hit Minu on wednesday: Minu and Matin have hit one another.

- b. tārā parasparer samālocanā karten: They used to criticize one another.

In (9a) the action is not simultaneous, and so it cannot be rightly expressed by a reciprocal structure. But in (9b), we assume, the action is not a face-to-face affair, but still the sentence is understood as if they criticized one another over the same, long period of time.

7.2 Constraints on Reciprocity.

Lees and Klima (1963, 156) claim that 'the reciprocal pronominalization' operates in a simple sentence structure. In Bengali a remote reciprocal relation can be expressed beyond a simplex structure by the items like eke anya, ekjan anyajan and ekṭi anyati etc., but not with paraspar and the morphologically derived items like mārāmāri, hātāhāti etc. We will see later that the items eke anya, ekjan anyajan etc., are not pronouns (cf., § 7.3). And even paraspar cannot be considered as a pronoun. The item paraspar must be in the same simplex structure as its antecedent.

Another condition for reciprocity is that the antecedent in a reciprocal structure must be plural. First consider the condition of plurality of the antecedent. Consider the examples below.

- (10) a. tārā parasparke mereche: They have hit one another.
 b. *se parasparke mereche: *He has hit one another.

In (10a) the antecedent is plural and the sentence is grammatical; but (10b) is ungrammatical because the antecedent is singular.

Now we will consider whether the antecedent and the reciprocal items should always be in the same simplex structure. Consider the examples below.

(11) a. $\bar{a}mar\bar{a}$ $\left\{ \begin{array}{l} \text{paraspar} \\ \text{eke anya} \\ \text{ekjan anyajan} \end{array} \right\}$ ke cini: We know one another.
 1 2 3 1 3 2

b. $\bar{m}\bar{a}tin$ o $\bar{m}in$ u $\bar{t}hel\bar{a}t\bar{h}eli$ karche: Martin and Minu are
2 pushing one another.
2

In (11a) the antecedent and the reciprocal items are in the same simplex structure. In (11b) there is no reciprocal item, but the verb $\bar{t}hel\bar{a}t\bar{h}eli$ kar due to its inherent reciprocal meaning indicates reciprocal action between the members referred to by the antecedent. The antecedent and the verb in (11b) are in the same simplex structure. In the following examples the antecedents are in a matrix sentence and the reciprocal items are in a constituent sentence; and the sentences are ungrammatical.

(12) a. * $\bar{k}etak\bar{i}$ o $\bar{n}\bar{i}lim\bar{a}$ mane kare ye $\bar{h}\bar{a}s\bar{a}n$ $\left\{ \begin{array}{l} \text{paraspar} \\ \text{eke anya} \\ \text{ekjan anyajan} \end{array} \right\}$ ke

$\bar{b}h\bar{a}lab\bar{a}se$: Ketaki and Nilima think that Hasan loves one another.

b. * $\bar{k}etak\bar{i}$ o $\bar{n}\bar{i}lim\bar{a}$ $\bar{r}\bar{a}t\bar{i}yeche$ ye $\bar{h}\bar{a}s\bar{a}n$ $\left\{ \begin{array}{l} \text{paraspar} \\ \text{eke anya} \\ \text{ekjan anyajan} \end{array} \right\}$ er

$\bar{k}\bar{a}che$ $\bar{c}i\bar{t}hi$ $\bar{l}ikheche$: Ketaki and Nilima have spread the rumour that Hasan has written letters to one another.

These sentences are ungrammatical because the items paraspar,

eke anya, ekjan anyajan etc., are in a constituent sentence and their antecedent is in a matrix sentence. Among these, the sentences which select paraspar as the reciprocal item are irreparable; that is, no other arrangement of their constituents can make them grammatical. But the sentences which select eke anya and ekjan anyajan as the reciprocal items can be made grammatical by rearranging their constituents. Their constituents can be rearranged as follows respectively.¹

(13) a. ketakī o nīlimā { eke
ekjane } mane kare ye hāsān { anya
anyajan } ke

bhālabāse: Ketaki and Nilima each thinks that Hasan love the other.

b. ketakī o nīlimā { eke
enkajne } raṭiyecche ye hāsān

{ anya
anyajan } er kāche ciṭhi likheche: Ketaki and

Nilima each has spread the rumour that Hasan has written letters to the other.

In (13) the elements eke anya and ekjan anyajan have been split into two parts, and have been distributed between the matrix and the constituent sentence: their first parts, eke and ekjan are in the matrix sentences; and their second parts, anya and anyajan are in the constituent sentences. This sort of splitting is not possible with paraspar, although it is a compound word composed of par: 'Other' and par: 'Other'.

¹ The sentences in (13) are not 'directly' reciprocal, but they are involved in a reciprocal relation indirectly.

If we decompose it into its components, and distribute them between the matrix and the constituent sentences, we will derive ungrammatical sentences like *ketakī o nīlimā pare mane kare ye hāsān parke bhālabāse: 'Ketaki and Nilima other thinks that Hasan loves the other'. So we see that a reciprocal relation cannot be expressed beyond a simplex structure with paraspar; but an indirect reciprocal relation can be expressed beyond a simplex structure by the elements like eke anya and ekjan anyajan. They can express a direct reciprocal relation only in a simplex structure.

The antecedent of a reciprocal relation need not be the subject of the sentence, and the reciprocal elements may or may not be the object of the sentence. This can be seen in (14).

- (14) a. āmarā mātin o minuke $\left\{ \begin{array}{l} \text{paraspar} \\ \text{eke anya} \\ \text{ekjan anyajan} \end{array} \right\}$ er dike

ṭhele dilām: We pushed Matin and Minu towards one another.

- b. mātin o minu $\left\{ \begin{array}{l} \text{paraspar} \\ \text{eke anya} \\ \text{ekjan anyajan} \end{array} \right\}$ ke mereche:

Matin and Minu have hit one another.

The sentences in (14a) are ambiguous in that the reciprocal items may refer to āmarā, the subject; and may refer to mātin o minu, the object. This shows that the antecedent of a reciprocal element may be either the subject or the object of the sentence. In (14b) the antecedent is the subject of the sentence.

Although the reciprocal items can be an object or some other

type of noun phrase in a reciprocal structure, they cannot be the subject. This can be seen in the ungrammatical examples in (15).

- (15) a. * $\left\{ \begin{array}{l} \text{paraspar} \\ \text{eke anya} \\ \text{ekjan anyajan} \end{array} \right\}$ mātin o minuke mereche: Each other

has hit Matin and Minu.

The reciprocal elements can precede their antecedents in Bengali. This can be seen in (16).

- (16) a. $\left\{ \begin{array}{l} \text{paraspar} \\ \text{eke anya} \\ \text{ekjan anyajan} \end{array} \right\}$ ke tārā mereche: They have hit

one another.

-1-

7.3 Reciprocal Items in Bengali: Are They Pronouns?

We have seen that Bengali utilizes at least three types of item in order to express reciprocity: (a) the compound word paraspar (← par par: 'Other other'), (b) the Specifiers eke anya, ekjan anyajan, ekṭi anyāṭi, ekṭi aparṭi etc., which mean 'One (an)other', and (c) the reduplicated reciprocal verbs like mārāmāri kar: 'Hit one another', ṭānāṭānai kar: 'Pull one another' etc.,. These three types of reciprocal item give rise to three types of reciprocal structure, which we will consider below.

First we will consider the items like eke anya, ekjan anyajan etc.,. Consider the examples below.

- (17) a. mātin o minu $\left\{ \begin{array}{l} \text{eke anya} \\ \text{ekjan anyajan} \\ \text{eke apar} \\ \text{*ekṭi anyāṭi} \end{array} \right\}$ ke mereche: Matin and

Minu have hit one another.

b. chelegulo $\left\{ \begin{array}{l} \text{eke anya} \\ \text{ekjan aparjan} \\ \text{eke apar} \\ \text{ekṭi anyāṭi} \end{array} \right\}$ ke mārche:

The boys are hitting one another.

c. tārā $\left\{ \begin{array}{l} \text{eke anya} \\ \text{ekjan aparjan} \\ \text{ekjan anyajan} \\ \text{*ekṭi anyāṭi} \end{array} \right\}$ ke bhālabāse: $\left\{ \begin{array}{l} \text{They love} \\ \text{one another.} \end{array} \right.$

In (17a) the items eke anya, ekjan anyajan and eke apar indicate that mātin o minu have hit one another. The item ekṭi aparṭi is unacceptable here because its antecedent conjoins two human proper nouns. In (17b) all the above items have been used as reciprocal elements. In (17c) we find that all the other items but ekṭi anyāṭi can be used as the reciprocal elements. The item ekṭi aparṭi is unacceptable because it will make the sentence pejorative. In all these examples the antecedents superficially seem to be the subjects of the respective sentences. But we will show below that the reciprocal items are not pronouns, but are SPEC(ifier)s (cf., § 2.4.3); and the antecedents are not real subjects. They are PART(itive)s (cf., § 2.4.7). The sentences in (17) can be paraphrased respectively as those in (18).

(18) a. mātin o minur madhye $\left[\begin{array}{l} \text{eke} \\ \text{ekjan} \\ \text{eke} \end{array} \right] \left[\begin{array}{l} \text{anya} \\ \text{anyajan} \\ \text{apar} \end{array} \right]$ ke mereche:

Of Matin and minu one has hit the other.

b. chelegulor madhye $\left[\begin{array}{c} \text{eke} \\ \text{ekjan} \\ \text{eke} \\ \text{ekṭi} \end{array} \right] \left[\begin{array}{c} \text{anya} \\ \text{aparjan} \\ \text{apar} \\ \text{anyaṭi} \end{array} \right] \text{ ke mārche:}$

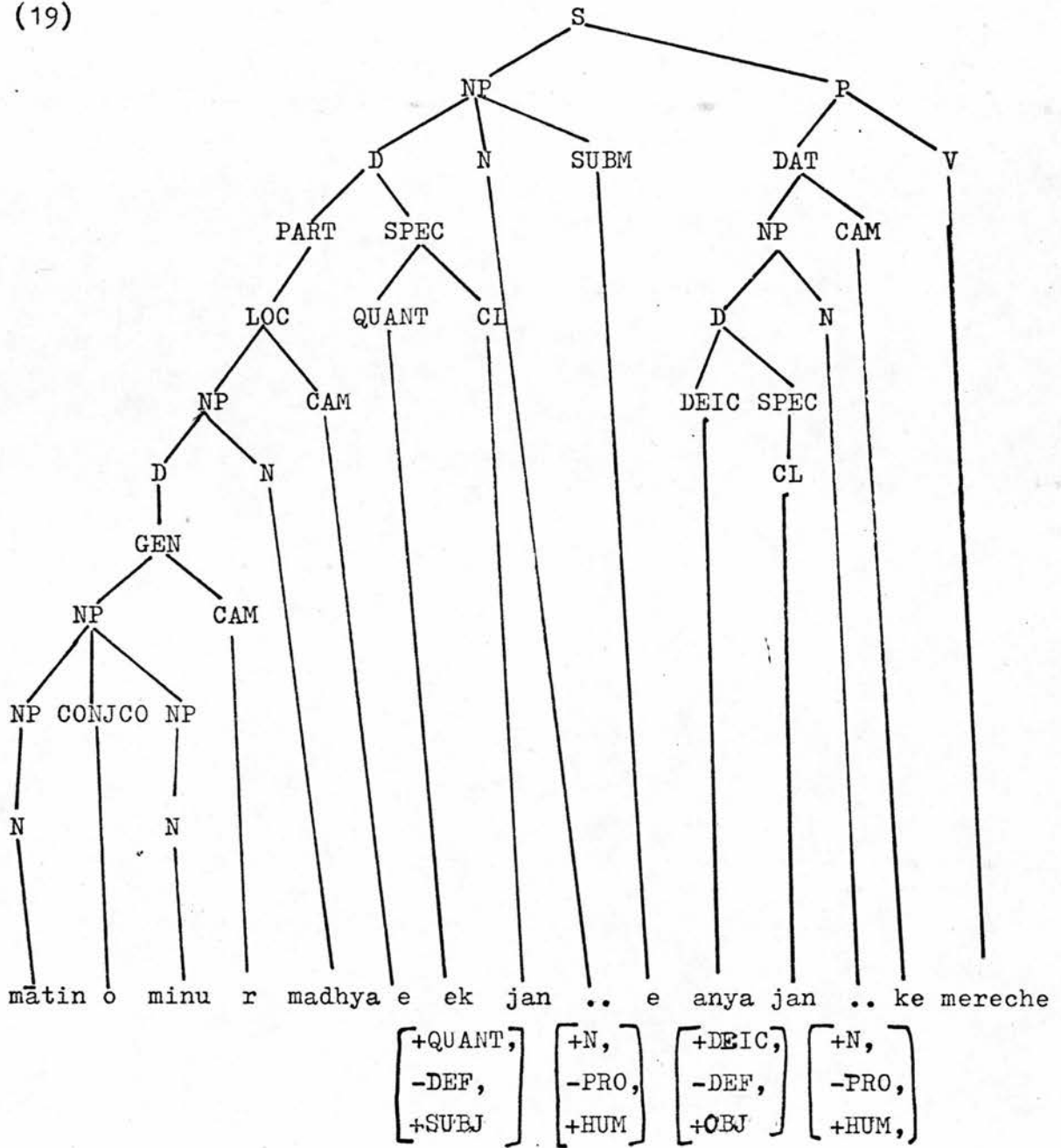
Of the boys one is hitting the other.

c. tāder madhye $\left[\begin{array}{c} \text{eke} \\ \text{ekjan} \\ \text{ekjan} \end{array} \right] \left[\begin{array}{c} \text{anya} \\ \text{aparjan} \\ \text{anyajan} \end{array} \right] \text{ ke bhālabāse:}$

Of them one loves the other.

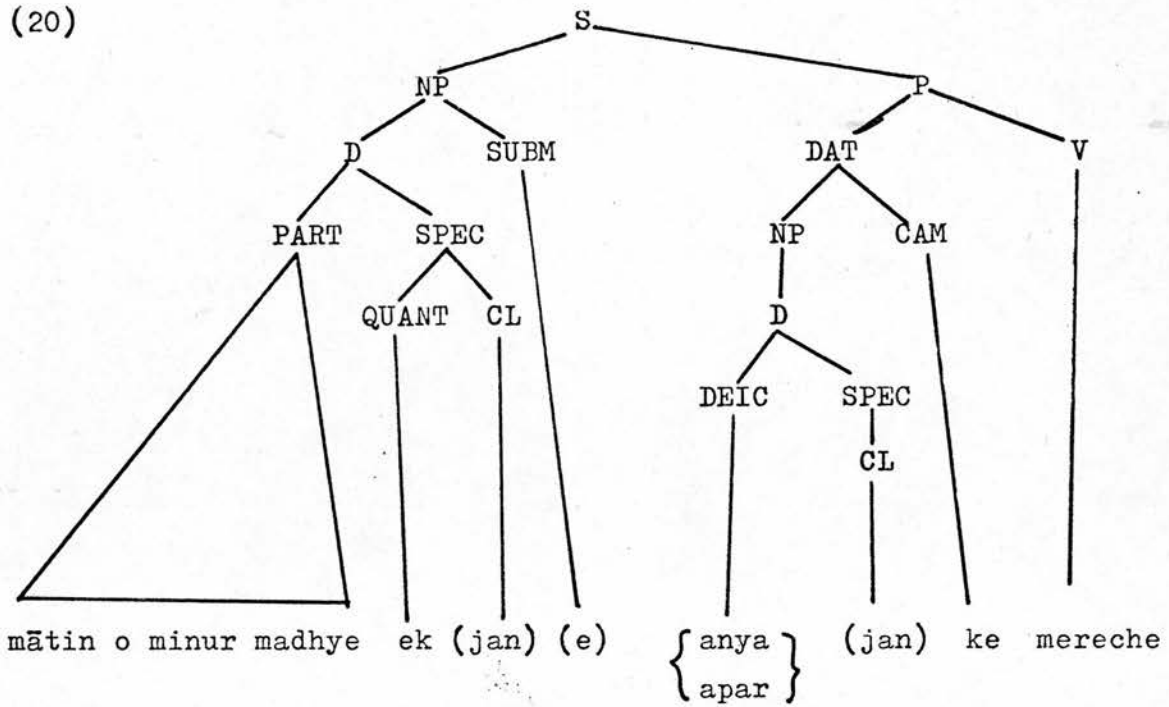
In (18) the antecedents are PART's, and we see that the elements eke anya, ekjan anyajan etc., are not single constituents. In (18) eke, ekjan and ekṭi etc., are subjects, and anya, anyajan and aparjan etc., are objects. In the surface these subjects and objects are placed in such a manner that they look like single constituents. The items eke, ekjan, ekṭi etc., are SPEC's consisting of the quantifier ek: 'One' and a classifier like jan or ṭi; and sometimes, the subject marker e. The elements anya, apar, anyaṭi etc., are determiners consisting of the deictics anya and apar: 'Other', and sometimes, a classifier like jan or ṭi. In (17) the antecedents are subjects in the surface, but we will assume that these sentences are derived from the underlying structure of the sentences in (18) respectively, by substantivization of the PART. We will propose that structures like (19) (see page 301) immediately underlie (18a). In (19) the head nouns of the subject and object noun phrases are null; and the classifiers are optional, and so is the subject marker. The semantic rules will derive a reciprocal reading from (19). In (19) we find that eke is an element of the subject noun phrase, and anya (apar, aparjan etc.,)

(19)



is an element of the object noun phrase. As the head nouns of the subject and object noun phrases in (19) are null, they will be deleted by a general rule. (19) will be transformed into (20) when these null head nouns are deleted. (see page 302).

(20)



(20) will generate the following structures depending on the selection of the optional elements (indicated by round brackets around the items):

(21) a. mātin o minur madhye eke {anya/apar} ke mereche:

Of Matin and Minu one has hit the other.

b. mātin o minur madhye ekjan {anya/apar} janke mereche:

Of Matin and Minu one has hit the other.

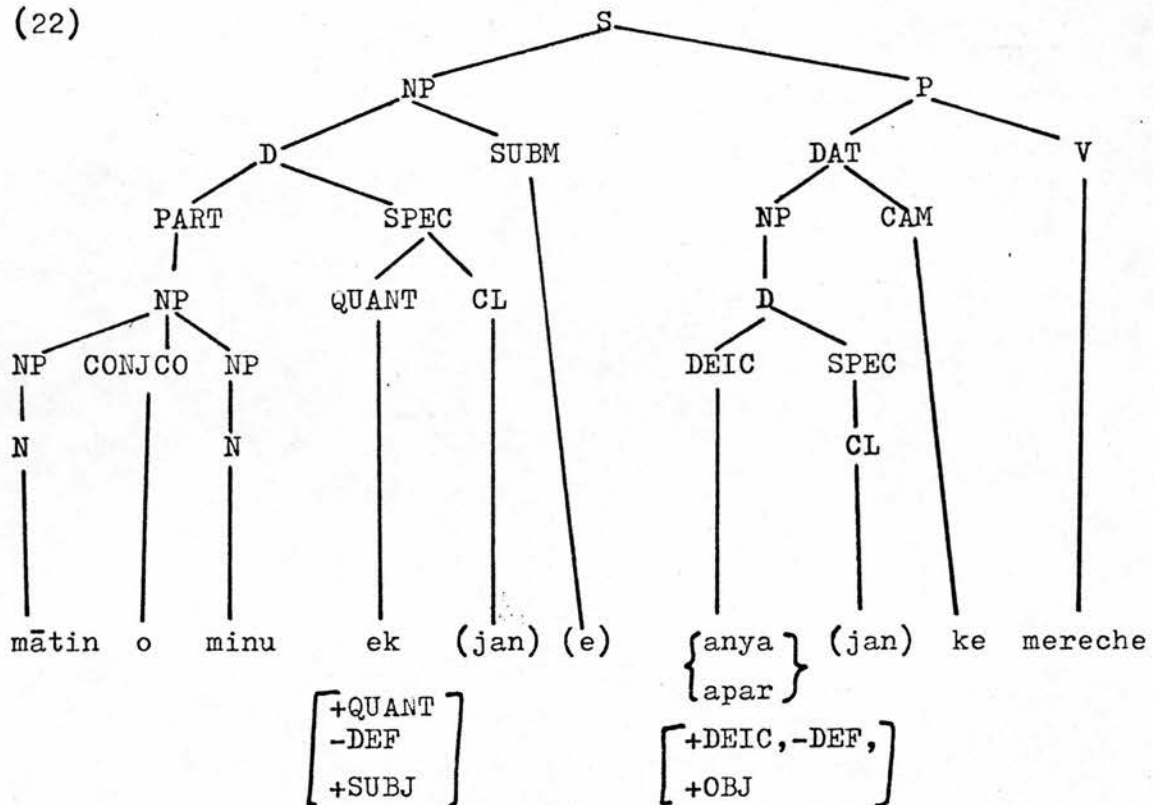
c. mātin o minur madhye ekjane {anya/aparjan} ke mereche:

Of Matin and Minu of has hit the other.

In (20) the antecedent is a PART constituent. The PART in (20) can be substantivized. This rule converts a PART constituent into a single noun phrase as shown in (21) by deleting all its constituents except the noun phrase dominated

by GEN. This will transform (20) into (22).

(22)



(22) will generate the sentences in (17a) as well as the sentences in (23) depending on the selection of the optional elements.

- (23) a. mātin o minu ekjane anyajanke mereche: Matin and Minu one has hit the other.
- b. mātin o minu ekjan aparjanke mereche:
- c. mātin o minu ekjane aparjanke mereche:

The derivation that we have proposed for these sentences shows that eke anya, ekjan anyajan, ekṭi aparṭi etc., are not pronouns, and they are not even single constituents. These elements appear in the surface in a order which creates the impression that they are single constituents. They do not replace any noun phrase, but by an interesting grammatical structure indicate reciprocal relation. We have seen before that a reciprocal relation can be expressed by these elements beyond a simplex structure. This is possible because they are not single

constituents, and are not pronouns; and so they can be distributed between the matrix and the constituent sentences. These sentences show that reciprocity can be expressed in Bengali without any reciprocal pronoun.

We will now consider whether the item paraspar is a pronoun. This item is used to express reciprocity in a simplex structure. A thorough investigation into its syntax and semantics will show that it is not a pronoun. We will prefer to call it a 'reciprocal item'. Semantically it does not refer to or replace some previously mentioned noun phrase as a pronoun does, but it interprets by virtue of its lexical meaning. It can be used as an object or a genitive noun phrase of an abstract locative expression (for example, parasparer dike: Lit., 'Towards each other's direction'), in order to express reciprocity. Moreover, it can be used as an extension of the antecedent at the same time as it appears as an object or a genitive noun phrase. Consider the examples in (24).

- (24) a. tārā parasparke bhālabāse: They love one another.
 b. tārā paraspare parasparke bhālabāse: They one another love one another.

In (24a) the object noun phrase paraspar interprets that the people referred to by its subject tārā have a reciprocal relation of love. (24b), which is synonymous with (24a), includes the item paraspar both as the extension of the subject and as the object.

The item paraspar, by its lexical meaning, denotes reciprocity. Like many nouns in Bengali, paraspar participates morphologically in the formation of an adjective pārasparik: 'Reciprocal'. So paraspar can be considered as an abstract noun

which interprets an antecedent appearing with it in the same simplex structure.

The antecedent in a reciprocal structure is always plural, but a reciprocal relation holds always with one person to another person. This is reflected in the morphological form of paraspar. This item is always singular in form. This can be seen in (25).

(25) a. *tārā parasparke cene*: They know one another.

b. **tārā parasparderke cene*:

(25b) is ungrammatical because the item paraspar has been pluralized, which is unacceptable syntactically as well as semantically. We will consider that paraspar is a reciprocal item with the categorial status of a noun. So we will derive it in the underlying structure.

Now we will consider the reduplicated reciprocal verbs like those in (26).

(26) a. *mātin o minu mārāmāri karche*: Matin and Minu are hitting one another.

b. *lokgulo hātāhāti karche*: Then men are striking one another (by hand).

Words like *mārāmāri*: 'Hit one another', *hātāhāti*: 'Strike one another by hand' etc., are traditionally considered as compound words derived by the rule of reciprocal compounding. This sort of compounding derives words from nouns and verbs by a rule something like $X \Rightarrow X\text{-}\bar{a}\text{-}X\text{-}i$, where X is either a noun or a verb. Subsequently, that derived word is used as an element of a conjunct verb as in (26). This sort of reduplicated verb denotes reciprocity if the components of the reduplicated word denote action that can be directed to others, as in (26). Otherwise it denotes action in a group or action overdone. Consider

the examples in (27, 28).

(27) a. $\begin{matrix} \text{tārā} & \text{hāsāhāsi} & \text{karche} \\ 1 & \underline{\quad 2 \quad} & 1 & \underline{\quad 2 \quad} \end{matrix}$ They are laughing.

b. *se hāsāhāsi karche: He is laughing.

(28) a. $\text{tārā bakābaki karche}$: They are rebuking one another.

b. $\text{ābbā sārādin bakābaki karen}$: Father rebukes all day long.

In (27a) hāsāhāsi is derived in a similar fashion to mārāmāri; but laughing is an action which cannot be directed towards others, and so the sentence (27a) is not understood reciprocally. It is understood as that a group of people laughing together. This verb takes a plural subject. (27a) is ungrammatical because its subject is singular. (28a) can be understood either in the strict or in the vague reciprocal sense, as rebuking is usually directed towards others. In (28b) the subject is singular, and so the sentence cannot be understood reciprocally. But it is understood as that father over-rebukes.

A structure containing such a verb derives its reciprocal interpretation due to the lexical meaning of the verb. These verbs cannot be derived transformationally; they must be derived by a morphological rule. Consider (26a). This sentence cannot be derived from a coordinate conjoined structure, as can be seen below.

(29) a. * $\text{mātin mārāmāri karche o minu mārāmāri karche}$: Matin is hitting one another and Minu is hitting one another.

b. * $\text{mātin mārā karche o minu māri karche}$:

(29a) is ungrammatical because the subject of each conjunct is singular, but the verb mārāmāri kar requires a plural subject.

In (29b) the components of mārāmāri have been split and

distributed between the conjuncts, and this has produced a peculiar and ungrammatical structure. These verbs are generated in the underlying structure. Thus we see that reciprocal structures in Bengali do not involve pronominalization, and none of the reciprocal items in Bengali is a pronoun.

7.4 Approaches to Reciprocity and Problems.

At least three approaches have been proposed so far to deal with reciprocal structures in English: (a) A moderate transformational approach by Lees and Klima (1963), (b) A conjunction reduction (extreme transformational) approach taken by Gleitman (1965) and Stockwell et al (1973), and (c) An interpretive approach taken by Jackendoff (1972), Fiengo and Lasnik (1973), and Dougherty (1974). The approach taken by Fiengo and Lasnik (1973) differs radically in detail from that of Jackendoff (1972) and Dougherty (1974); but their approaches can be grouped together as all of them generate reciprocal items in the underlying structure. There is also a vagueness about the process itself: is it a sort of pronominalization or something else? It is true that Lees and Klima (1963), and Stockwell et al (1973) clearly refers to the process as 'Reciprocal Pronominalization', but Fiengo and Lasnik (1973), and Dougherty (1974) do not refer to the process as pronominalization; they simply refer to the structures either as 'Reciprocal sentences' or as 'Each Other Constructions'. We have referred to these structures as 'Reciprocal Structures'. We have seen in the previous section that none of the reciprocal items in Bengali can be considered a pronoun, and so the process of denoting a reciprocal relation should not be considered a pronominalization.

Below, we will review, briefly, the approaches mentioned above. First consider Lees and Klima's (1963) approach. We quote the relevant passage from Lees and Klima (1963, 156-57):

We shall say then that in addition to the reflexive pronominalization transformation there is an optional rule of the following form:

(C) Reciprocal Rule (Optional):

$$X - N + PL - Y - N' + PL - Z \text{ --->}$$

$$X - N + PL - Y - N' + PL + \text{Recip} - Z$$

Where $N = N'$ and they are within the same simplex, and where N is a noun, PL is the plural morpheme, and Recip is the reciprocal morpheme.

Later morphophonemic rules will then yield the appropriate forms one another and each other from $N + PL + \text{Recip}$.

Lees and Klima's rule is inadequate, because it will derive, for example, (30a) from (30b).

(30) a. John and Mary frighten one another.

b. John and Mary frighten John and Mary.

(30b) has no reciprocal interpretation, but (30a) is a reciprocal structure. So the recent development of the transformational theory would not allow the derivation of (30a) from (30b).

An elaborate discussion of the conjunction reduction approach to reciprocal structures is available in Dougherty (1974), who refutes this approach for syntactic and semantic reasons. Stockwell et al (1973) propose a conjunction reduction approach for reciprocal structures. Their approach is based on Gleitman's (1965) approach. In this approach a reciprocal structure is derived from an underlying coordinate conjoined structure. In this approach the reciprocal items are

never present in the underlying structure. Stockwell et al (1973) do not deal with the reciprocal rule in any detail, but their view on this is to be found in their comments on conjunction. We quote them (1973, 312) below:

While we do not feel that symmetric predicates like agree belong to a single syntactic class, we would claim that this class is defined not by "symmetry" but rather, by susceptibility to a reciprocal-pronoun deletion transformation. Thus we would propose derivations like the following:

(31) Johnson agreed with Kosygin and Kosygin agreed with Johnson.

⇒ (by derived conjunction, etc)

Johnson and Kosygin agreed with Kosygin and Johnson respectively.

⇒ (by reciprocal pronominalization)

Johnson and Kosygin agreed with one another.

⇒ (by reciprocal pronoun deletion)

Johnson and Kosygin agreed.

So they derive reciprocal items from an underlying coordinate conjoined structure via a respectively-transformation. This transformation, although it solves some semantic problems with the help of peculiar underlying structures, cannot solve all the problems that arise due to a transformational approach to reciprocal structures. The problem with the respectively-transformation is that it cannot express a reciprocal relation involving more than two persons, and it faces great difficulty to deal with reciprocal structures containing a 'symmetric predicate'.

A third approach to reciprocal structures has been taken by Jackendoff (1972), Fiengo and Lasnik (1973), and Dougherty (1974). Their approaches have been taken here as the interpretive approach because all of them generate the reciprocal items in the underlying structure. But their treatments differ radically from one to the other's in detail. Jackendoff (1972, 173) says:

Thus the environments of each other seem to be virtually identical to those of reflexives, and any analysis which does not capture this fact is missing an important generalization... In the phrase structure theory, each other is generated within a single constituent. Thus the lexicon can list it as an idiom, with special semantic interpretation including the feature [+refl], which is not shared by the others.

Fiengo and Lasnik (1973) and Dougherty (1974) also generate the reciprocal items in the underlying structure, but not as Jackendoff does. Because their formulations are organised for the peculiarities of the reciprocal items in English, we need not go into their detail. What is important is that they generate the reciprocal items in the underlying structure.

7.5 The Derivation of Reciprocal Structures.

We have seen that none of the reciprocal items in Bengali is a pronoun, and so a reciprocal structure does not involve a pronominalization rule (cf., § 7.3). We have proposed that the reciprocal items should be generated in the underlying structure; and the rules of semantic interpretation will give a reciprocal reading to a structure in which they appear.

In the underlying structure of reciprocal sentences containing eke anya, ekjan anyajan etc., an antecedent is generated as a Partitive constituent, and eke and anya etc., are generated as distinct constituents. So the surface sentence (32a) will be derived from the underlying structure of (32b).

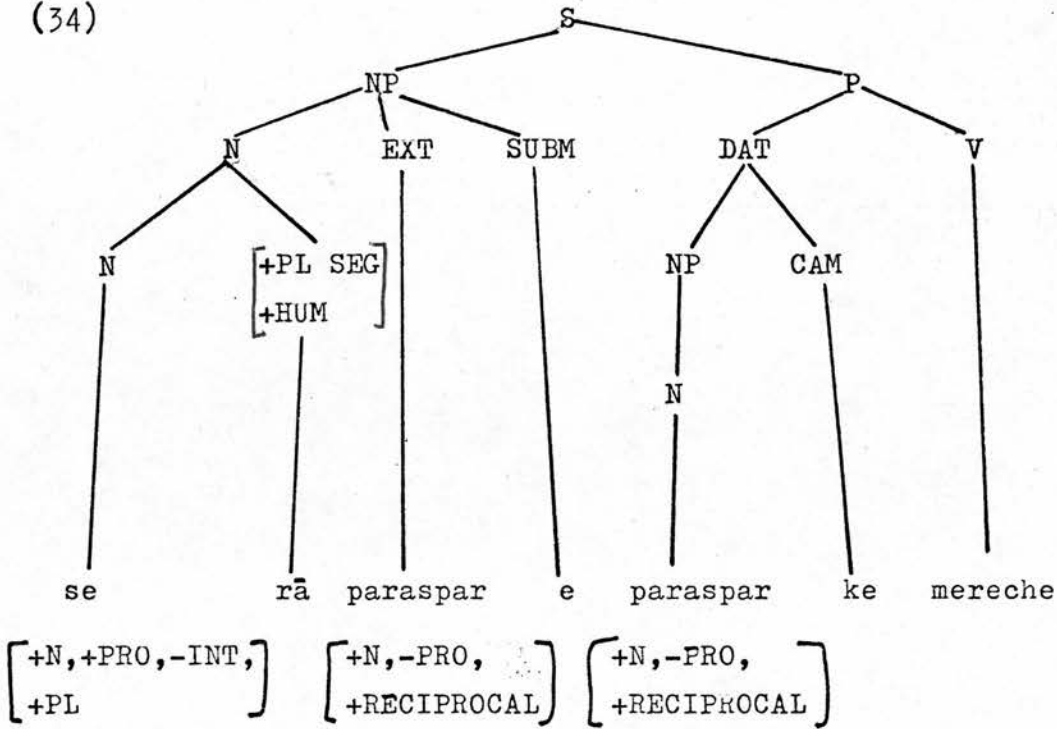
- (32) a. mātin o minu eke anyake mereche: Matin and Minu
one has hit the other.
- b. mātin o minur madhye eke anyake mereche: Of Matin
and Minu one has hit the other.

We will also generate the reciprocal item paraspar in the underlying structure, both as an extension of the antecedent and as an individual constituent. Consider the examples in (33).

- (33) a. tārā parasparke mereche: They have hit one another.
- b. tārā paraspare parasparke mereche: They one another
have hit one another.

The sentences in (33) are synonymous, but they differ in that (33a) has paraspar only as an object and (33b) has it both as an object and as an extension of the antecedent. We will posit an intermediate underlying structure (34) for both the sentences in (33); and the extension paraspar will be deleted by a general rule that any redundant material can be optionally deleted. (see page 312 for (34)).

(34)



The semantic rules will derive a reciprocal reading for (34). The extension of the antecedent and the subject marker are optional in (34); so they can be deleted. (34) will generate (33b); and after the deletion of the extension and the subject marker, we will derive (33a).

We have already seen that the reduplicated reciprocal verbs should be generated in the underlying structure (cf., § 7.3). These verbs are generated by the morphological component of the grammar, and the syntactic component utilizes them.

RELATIVIZATION8.0 Introduction.

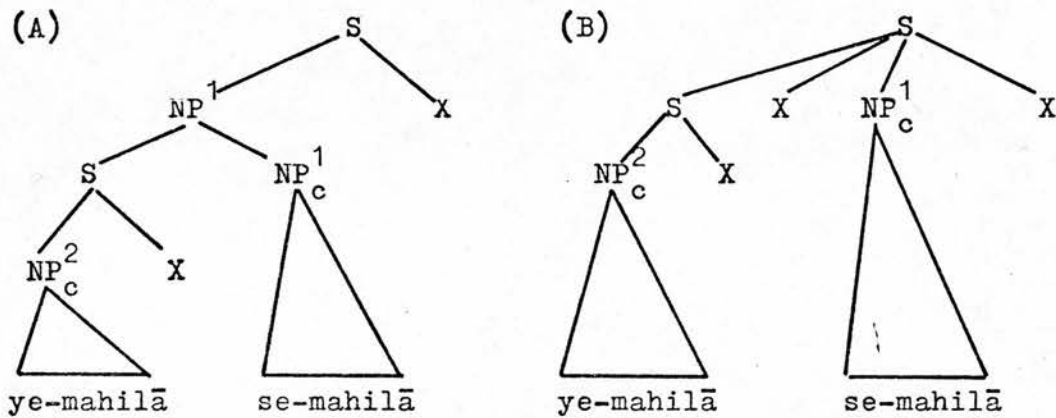
Relativization is the process which generates relative clauses. A relative clause in Bengali is an embedded sentence, in the surface, which acts as a modifier of a noun phrase. A relative clause contains a relativized noun phrase which is either a full noun phrase containing the relative deictic ye, or a relative pronoun such as ye, yini and yā. Bengali has restrictive as well as nonrestrictive relative clauses.

Restrictive relative clause sentences in Bengali are classifiable into three types in respect of their structure: (a) Pre-nominal restrictives, (b) Elaborate restrictives, and (c) Post-nominal (Anglicized) restrictives. We will consider all of them in a later section (cf., § 8.2.1), but in this introductory section we will take a close look into the pre-nominal restrictive relative clauses as they constitute the major class of the restrictive relative clauses in Bengali. A pre-nominal restrictive relative clause is one which precedes its containing noun phrase.¹ We assume that this type of restrictive

1 We will refer to the matrix noun phrase which is understood coreferentially with a relativized noun phrase (NP-Rel) as the containing noun phrase, irrespective of the fact that in some cases it may not contain the relative clause. We can have structures like (A) and (B) (see page 314):

continued

relative clause immediately precedes its containing noun phrase in the intermediate structure (not in the underlying structure because we derive restrictive relative clauses from underlying coordinate conjoined sentences); but it is moved out of that position to the sentence-initial position if the sentence



In (A) NP^1 is the containing noun phrase in which the relative clause is Chomsky-adjoined and which is understood coreferentially with the relativized noun phrase NP^2 . In (B) NP^1 is the containing noun phrase, with which the relativized noun phrase NP^2 is coreferential. In (B) the containing noun phrase does not contain the relative clause, because the relative clause has been extraposed to the sentence-initial position. In notional terms a containing noun phrase is one which is understood coreferentially with a relativized noun phrase. In transformational literature there are several terms like 'head noun', 'head noun phrase' etc., used for the purpose for which we use the term 'containing noun phrase'. It was first used by Smith (1964).

contains a single restrictive relative clause. Consider the examples below.

- (1) a. ye-cheleṭi_c kabitā likhto, se_c mare geche: The boy
 1 2 3 4 5 6 2
 who used to write poems has died.
 1 4 3 6
- b. āmi ye-cheleṭi_c kabitā likhto, tāke_c cintām: I knew the
 1 2 3 4 5 6 7 1 7
 boy who used to write poems.
 3 2 5 4
- c. ye-cheleṭi_c kabitā likhto, āmi tāke_c cintām: I knew
 1 2 3 4 5 6 7 5 7
 the boy who used to write poems.
 2 1 4 3

In (1a) the restrictive clause ye-cheleṭi kabitā likhto precedes the pronoun se: 'He', which is modified by the relative clause. The relative clause in this sentence comes automatically at the sentence-initial position, because its containing noun phrase, which has been pronominalized into se, is the initial constituent of its own simplex. In (1b) the relative clause precedes its containing noun phrase tāke, and is inside the sentence as its containing noun phrase is a medial constituent of the matrix sentence. As (1b) has a single relative clause, the relative clause can be extraposed to the sentence-initial position. If we move it from this position and place it sentence-initially, we will get (1c), which is synonymous with (1b). Although (1b) is a grammatical sentence, it creates performance problems as the relative clause is centre-embedded in this sentence (cf., Kuno (1974, 119)). In order to avoid this problem, the relative clause has been placed sentence-initially in (1c). This shows that Bengali has a tendency to place restrictive relative clauses sentence-initially if a sentence contains a single restrictive relative clause (cf., 8.3.3).

In Bengali the relativized noun phrase (NP-Rel) can be placed at the initial, medial or final position of the relative clause. Consider the examples below.

- (2) a. ye-cheleṭi kabitā likhto: Lit., Who-the boy used to write poems (The boy who used to write poems).
 b. kabitā likhto ye-cheleṭi:
 c. kabitā ye-cheleṭi likhto:

The examples in (2) are of a single restrictive relative clause in differing constituent order. In (2a) the NP-Rel ye-cheleṭi is an initial constituent, in (2b) a final constituent, and in (2c) a medial constituent of its own clause.

A relativized noun phrase in a pre-nominal restrictive relative clause need not be reduced to a relative pronoun in Bengali. Consider the examples in (1, 2). In these examples the NP-Rel contains the relative deictic ye, as in ye-cheleṭi; but they are not relative pronouns (cf., § 8.1; § 8.4). The pronominalization proper transformation follows the relativization rule in Bengali, and reduces either the containing noun phrase or the relativized noun phrase into a pronoun (cf., § 8.4).

A nonrestrictive relative clause in Bengali usually follows the containing noun phrase, but it can be placed sentence-finally. Consider the examples in (3).

- (3) a. cheleṭi_c, ye_c kabitā likhto, mare geche: The boy, who used to write poems, has died.
 b. cheleṭi_c mare geche, ye_c kabitā likhto:

The sentences in (3) are ambiguous in that the relative clause ye kabitā likhto in each sentence can be understood either restrictively or nonrestrictively. In (3a) the relative clause follows the containing noun phrase, and in (3b) it has been

placed sentence-finally. A post-nominal or sentence-final NP-Rel is usually realized as a relative pronoun in the surface as it creates an environment for the application of the pronominalization rule forwards. It may be mentioned here that nonrestrictive relative clause sentences are rare in Bengali.

Relativization in Bengali is a fairly complex process. It involves such notions as coreferentiality, definitization, presupposition and pronominalization proper. We will deal with relativization in Bengali in this chapter; and will show how relative clause sentences in Bengali can be derived.

8.1 The Relative Deictic *ye* and the Relative Pronouns.

Bengali has a relative marker ye, which is placed at the left of a relativized noun phrase as , for example, in ye-lok: 'ye-man', ye-cheleṭi: 'ye-the boy' etc., and three 'genuine' relative pronouns: ye, yini and yā (see (9) at the end of this section for their feature specifications). We assume that the relative pronouns are lexically derived from the relative marker ye. The relative marker functions deictically, and so we will refer to it as the relative deictic. It is homonymous with the complementizer ye and the nonhonorific relative pronoun ye. We will consider them as three distinct lexical items.

Consider the examples in (4).

- (4) a. ye-mahilā_cātmaḥatyā karechen, tini_c samudra bhālabāsten:
 1 2 3 4 5 6
 Lit., Who-the woman has committed suicide, she used to
 1 2 3 4
 love the sea (The woman who has committed suicide used
 6 5
 to love the sea).

b. ye-bhadralok_c āschen, tini_c kamyuniṣṭ: The gentleman who is coming is a communist.

c. āmi āj ye-kabitāṭi_c likhechi, se-ṭi_c āmār 'šeṣ kabitā:

The poem which I have written to-day is my last poem.

ye-mahilā ātmahatyā karechen, ye-bhadralok āschen and āmi āj ye-kabitāṭi likhechi are the relative clauses in (4a, b, c) respectively. We will refer to a noun phrase which is relativized or relativizable as an NP-Rel (relative noun phrase). The NP-Rel's are ye-mahilā, ye-bhadralok and ye-kabitāṭi in (4a, b, c) respectively. The presence of the relative deictic ye in each of these noun phrases indicates that they have been relativized. It can be seen in (4) that the NP-Rel's are not relative pronouns, but are full noun phrases with the relative deictic ye; but the containing noun phrases are pronouns. What is interesting in the sentences is that the relativized noun phrases do not refer to the containing noun phrases. Instead, the containing noun phrases, which have been pronominalized due to their coreference with the NP-Rel's, refer to the relativized noun phrases. The structures which immediately underlie the sentences in (4) are given in (5), respectively.

(5) a. ye-mahilā_c ātmahatyā karechen, se(i)-mahilā_c samudra bhālabāsten: Lit., Who-the woman has committed suicide, that woman used to love the sea.

b. ye-bhadralok_c āschen, se(i)-bhadralok_c kamyuniṣṭ: Lit., Who-the gentleman is coming, that gentleman is a communist.

c. āmi āj ye-kabitāṭi_c likhechi, se(i)-kabitāṭi_c āmār 'šeṣ kabitā: Lit., Which-the poem I have written to-day, that poem is my last poem.

In each example in (5) the NP-Rel contains the relative deictic ye and the containing noun phrase contains the demonstrative deictic se(i): 'That'. In pre-nominal restrictive relativization the relative deictic is placed at the left of the NP-Rel, and the demonstrative deictic se(i) is placed at the left of the containing noun phrase. This is a simultaneous process of relativization and definitization (but cf., § 8.4): when ye is placed at the left of a noun phrase it becomes definite, restricted and relativized; and the simultaneous placement of se(i) at the left of the containing noun phrase makes it definite, and a syntactic coreferential relation between these two noun phrases is established. In each example in (5) relativization has taken place, but the NP-Rel's have not been reduced to relative pronouns. In order to derive the relative pronouns in Bengali, the pronominalization proper rule must apply to the relativized noun phrases. We will come back to this later (cf., § 8.4). In (5) each structure has two noun phrases (the NP-Rel and the containing noun phrase) which are coreferential. They satisfy all the conditions for the pronominalization rule. In structures like (5) the pronominalization rule usually applies forwards and derives the sentences in (4) (cf., § 8.4). In (4, 5) we have relativized noun phrases, but no relative pronoun. The relative deictic ye is used to relativize noun phrases of all inherent features ([+HON], [+ABS], [+HUM] etc.), but the relative pronouns differ from one another for the features like [+HON], [-HON] and [-HUM, -ANI]. So we can conclude that ye in (4, 5) is the Bengali relative deictic (marker).

We will now consider some restrictive relative sentences where we usually come across relative pronouns. Consider the examples in (6).

(6) a. ye_{1c} $chāde$ $geche$, se_{4c} $cād$ $dekheche$: $\left\{ \begin{array}{l} \text{The one} \\ \text{Any one} \end{array} \right\}$ who has
 gone to the roof has seen the moon.
 $\underline{3}$ $\underline{2}$ $\underline{6}$ $\underline{5}$

b. $yini_c$ $chāde$ $gechen$, $tini_c$ $cād$ $dekhechen$:

c. $āmi$ $yā_c$ $cāi$, $tā_c$ $pāi$ $nā$: I do not get $\left\{ \begin{array}{l} \text{which} \\ \text{whatever} \end{array} \right\}$

I want.

In (6) the relativized noun phrases are relative pronouns ye, yini and yā in (6a, b, c) respectively, and their containing noun phrases are pronouns se, tini and tā in (6a, b, c) respectively. If we compare these sentences with those in (4, 5), we find that the relativized noun phrases in (4, 5) contain the relative deictic ye, but in (6) the relativized noun phrases are relative pronouns. The problem which arises is whether we will consider ye in the relativized noun phrases in (4, 5) and the ye in (6a) as the same lexical item. Our view is that they are distinct homonymous items: ye in (4, 5) is the relative deictic and ye in (6a) is the nonhonorific relative pronoun. The sentences in (6) are ambiguous due to the inherent ambiguity of the relative pronouns. ye in (6a) suggests that the one or any one who has gone to the roof is a nonhonorific human being, and yini in (6b) suggests that the one or any one who has gone to the roof is an honorific human being; and yā in (6c) suggests that I want some specific or unspecific thing, but does not mention what that thing is.

So there are two problems with the relative pronouns in (6): (a) they may be understood as definite, or as indefinite-unspecific; and (b) they do not spell out their referents (that is, they do not mention whether the referent is a man, or a woman, or a book etc.,; they only mention that the referent is a nonhonorific human being, or an honorific human being, or a thing etc.,). Although Bengali has a relative pronoun which can refer to the whole class of the nonhonorific human beings, there is no noun which includes all nonhonorific human beings; and similarly there is no noun which includes the whole class of the honorific human beings, but the relative pronoun yini can refer to the whole class of the honorific human beings. If we had such nouns, we could show that the examples in (5) and (6) are derived in the similar fashion, and ye in (6a) is a relative pronoun. We can take the view that, although there are concepts for a class of all nonhonorific human beings, for a class of all honorific human beings, and for a class of all abstract and concrete objects, due to an accidental gap there is no lexical item for them in Bengali. We can consider them as phonetically null lexical items. When the conception of phonetically null lexical items is taken, we can relate the sentences in (6) to the structures in (7) respectively as the sentences in (4) are related to those in (5). We will use dummy symbols (Δ) for null lexical items with proper feature specifications. So we assume that the structures in (7) immediately underlie those in (6) respectively.

(7) a. ye- Δ chāde geche, se- Δ cā̃d dekheche:

$\left[\begin{array}{l} +N, -PRO, +COM, \\ +COUNT, +HUM, \\ +3, -HON, \\ \underline{+DEF}, -PL \end{array} \right]$	$\left[\begin{array}{l} +N, -PRO, +COM, \\ +COUNT, +HUM, \\ +3, -HON, \\ \underline{+DEF}, -PL \end{array} \right]$
----------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------

Who- Δ has gone to the roof, that Δ has seen the moon.

b. ye- Δ chāde gehen, se- Δ cā̃d dekhechen:

$\left[\begin{array}{l} +N, -PRO, +COM, \\ +COUNT, +HUM, \\ +3, +HON, \\ \underline{+DEF}, -PL \end{array} \right]$	$\left[\begin{array}{l} +N, -PRO, +COM, \\ +COUNT, +HUM, \\ +3, +HON, \\ \underline{+DEF}, -PL \end{array} \right]$
----------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------

Who- Δ has gone to the roof, that Δ has seen the moon.

c. āmi ye- Δ cāi, āmi se- Δ pāi nā:

$\left[\begin{array}{l} +N, -PRO, +COM, \\ \underline{+COUNT}, \underline{+ABS} \\ +3, -HON, \\ \underline{+DEF}, -PL, \end{array} \right]$	$\left[\begin{array}{l} +N, -PRO, +COM, \\ \underline{+COUNT}, \underline{+ABS}, \\ +3, -HON, \\ \underline{+DEF}, -PL \end{array} \right]$
----------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------

I want which- Δ , I do not get that Δ .

For the sake of convenience, we have stacked all the features of the noun phrases involved in relativization above, under the dummies. The restrictive relative transformation has applied in all the structures above. This is indicated by the relative deictic ye in each NP-Rel. As the head nouns of the noun phrases involved in relativization in (7) have no phonetic shape, we cannot have ye-N..., se-N restrictive relative structures from these, as we get in (5).² So, the pronominalization rule will apply

² Those who are not happy with null lexical items can use jan meaning 'human being' for the dummies in (7a, b), and jinis: 'Thing' for the dummies in (7c). This will give rise to the structures below, which are semantically close to (6a, b, c).

continued

obligatorily in these structures in both directions, and will specify the NP-Rel's and the containing noun phrases as [+PRO]. Later the relative deictic ye will be deleted from the relativized noun phrases, and the feature [+REL] will be added to the feature bundles associated with the head nouns of the relativized noun phrases. The relative pronouns ye, yini and yā will be attached to the relativized noun phrases in (7a, b, c) respectively. We are not concerned with the formalities of the operations here, what we want to say is that each relative pronoun in (6) originates from a pronominal reduction of structures like ye-NP, where ye is the relative deictic. The nonhonorific relative pronoun ye, which we find in (6a), is not the same ye, which we find in (7) or (4, 5).

The relative pronoun ye can be used only if the relativized noun phrase is nonhonorific, but the relative deictic ye is used to relativize noun phrases of all types of inherent feature. In pre-nominal restrictive relativization in Bengali the relativized noun phrases are rarely realized as relative pronouns: only in cases like (6) a relativized noun phrase is realized

-
- (A) a. ye-jan_cchāde geche, se-jan_ccā̃d dekheche: Lit., Who-the one has gone to the roof, that one has seen the moon.
 b. ye-jan_cchāde gechen, se-jan_ccā̃d dekhechen:
 c. āmi ye-jinis_ccāi, āmi se-jinis_cpāi nā: Lit., I which-the thing want, I do not get that thing.

The structures in (A) are very close semantically to those in (6a, b, c) respectively.

as a relative pronoun. Although almost all types of noun phrase can be relativized in Bengali, there are only three 'genuine' relative pronouns in Bengali: the nonhonorific ye, the honorific yini, and the inanimate and abstract yā. Other types of noun phrase can be relativized, but cannot be reduced to a relative pronoun. Consider the examples below.

- (8) a. ye-khān: ye-place.
 b. ye-bhāb: ye-manner.
 c. ye-samay: ye-time.

The relativized noun phrases in (8) cannot be replaced by any relative pronoun. But in some relativized noun phrases morphophonological rules may contract ye-NP structures into single lexical items. For example, the relativized noun phrase ye-khan: 'ye-moment' will be obligatorily contracted into the single lexical item yakhan: 'When' by the morphophonological rules. It seems that the Bengali relative pronouns are morphologically derived from the relative deictic. This is not unique. In old English the relative clauses were introduced by what Jespersen calls 'demonstrative pronouns' (se, seo, pæt), or by the relative particle 'þe', or by both jointly (cf., Jespersen (1927, 80), Curme (1931, 204-38)).

In this work the relative deictic is introduced transformationally by the relativization rule, but the relative pronouns ye, yini and yā are introduced into the derived nodes by the second lexical insertion rule. We have taken these three relative pronouns as the 'genuine' relative pronouns in Bengali, because they cannot be derived by any morphophonological rule. But other relative items like yekhān: 'Where', yakhan: 'When', etc., are derived by morphophonological rules, and so they are not listed

in the lexicon. The lexical entries for the relative pronouns are as follows:

- (9) ye: 'Who' yini: 'Who' yā: 'Which'
- | | | |
|-------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------|
| $\left[\begin{array}{l} +N, +PRO, +REL, \\ +COUNT, +ANI, \\ +HUM, +\text{3}, -HON, \\ \underline{+DEF}, \end{array} \right]$ | $\left[\begin{array}{l} +N, +PRO, +REL, \\ +COUNT, +ANI, \\ +HUM, +\text{3}, +HON, \\ \underline{+DEF}, \end{array} \right]$ | $\left[\begin{array}{l} +N, +PRO, +REL, \\ -COUNT, -ANI, \\ \underline{+ABS}, +\text{3}, -HON, \\ \underline{+DEF}, -PL, \end{array} \right]$ |
|-------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------|

8.2 Restrictive and Nonrestrictive Relative Clauses.

Bengali has restrictive as well as nonrestrictive relative clauses, but while restrictive relative clause sentences are fairly common in Bengali, nonrestrictive relative clause sentences are rare. We should guess that nonrestrictive relative clauses are recent innovations in Bengali, under the influence of English. A restrictive relative clause is traditionally called a 'defining clause' (cf., Jespersen (1927, 82)), and a nonrestrictive relative clause is called a 'descriptive relative' (cf., Curme (1931, 223)). The English nonrestrictive relative clauses are also called 'appositive relatives' (cf., Smith (1964)). The main difference between them is that a restrictive relative clause 'restricts' a noun phrase, and a nonrestrictive relative clause is a parenthetical statement about its referent.

8.2.1 Restrictive Relative Clauses in Bengali.

Restrictive relative clauses in Bengali are classifiable into three types according to their structure: (a) Pre-nominal restrictives, (b) Elaborate restrictives, and (c) Post-nominal restrictives. Among these, the first two types are fairly common in Bengali, but the third type is rare. We give

examples of each type below.

Pre-nominal restrictives:³

- (10) a. ye-cheleṭi₁ c₂ kabitā₃ likhto, se₄ mare geche: Lit., who-
 the boy used to write poems, he has died (The boy who
 used to write poems has died).
 1 2 3 4 5 6 1
- b. ye-mahilā₁ c₂ ātmahatyā₃ karechen, tini₄ c₅ samudra bhālabāsten:
 Lit., Who-the woman has committed suicide, she used to
 love the sea (The woman who has committed suicide
 used to love the sea).
 1 2 3 4 5 6

Elaborate restrictives:

- (11) a. cheleṭi₁ c₂, ye₃ kabitā₄ likhto, se₅ mare geche: Lit., The
 boy, who used to write poems, he has died.
 1 2 3 4 5 6
- b. mahilā₁ c₂, yini₃ c₄ ātmahatyā₅ karechen, tini₆ c₇ samudra
 bhālabāsten: Lit., The woman, who has committed
 suicide, she used to love the sea.
 1 2 3 4 5 6

Post-nominal (Anglicized) restrictives:

- (12) a. cheleṭi₁ c₂, ye₃ kabitā₄ likhto, mare geche: The boy
 who used to write poems has died.
 1 2 3 4 5 1
- b. mahilā₁ c₂, yini₃ c₄ ātmahatyā₅ karechen, samudra bhālabāsten:
 The woman who has committed suicide used to love
 the sea.
 1 2 3 4 5

³ Some remarks on conventions of writing a relative clause sentence are in order here. We have said that a pre-nominal restrictive relative clause is usually placed, in the surface,

continued

In (10) the restrictive relative clauses have been placed immediately preceding their containing noun phrases, which have been pronominalized. So in these examples the relativized noun phrase does not refer to the containing noun phrase; instead, the containing noun phrase refers anaphorically to the relativized noun phrase. In (11) the referent is mentioned first, and is followed by a relative clause; then comes the matrix sentence.

sentence-initially if the sentence contains a single restrictive clause. Otherwise it is placed preceding its containing noun phrase (cf., § 8.0; § 8.3.3). In spoken Bengali there is always a short pause between the relative clause and the elements that follow or precede it. But written Bengali has no convention of using a comma to mark that pause. Nevertheless, many writers, nowadays, mark that pause by a comma. In all the examples that we have cited so far, we have used a dash (-) between the relative deictic and the relativized noun phrase . . . (e.g., ye-mahilā), and between the demonstrative deictic and the containing noun phrase (e.g., se-mahilā). This convention is not universally practised in Bengali. We will, however, follow this convention throughout, as it is technically accurate.

The Bengalees will write a relative clause sentence carelessly, for example, as (Aa), which in our convention will be written as (Ab):

(A) a. ye cheleṭi kādche se cheleṭi kṣudhārtha: The boy
who is crying is hungry.

b. ye-cheleṭi kādche, se-cheleṭi kṣudhārtha:

(Ab) is more accurate than (Aa) in the following ways: (a) the comma helps to sort out the relative clause easily, (b) the dash

continued

Consider (11a). Here the referent cheleṭi is mentioned first, and then comes the restrictive relative clause containing a relative pronoun, which refers to cheleṭi. And next comes the matrix sentence containing the pronoun se. This pronoun refers to the relative pronoun. In this way cheleṭi, ye and se are anaphorically related. This type of restrictive relative clause sentence is elaborate in that it uses redundant elements. The elaborate restrictives are related to the pre-nominal restrictives. We will later show that the elaborate restrictives are derived from the pre-nominal restrictives (cf., § 8.4.1). In (12) the restrictive relative clauses follow their containing noun phrases, and the relative clauses contain relative pronouns which refer to the containing noun phrases. The examples in (11) and (12) show that the elaborate and post-nominal restrictive relative clauses are related: that is, a post-nominal restrictive clause seems to be derived by the deletion of the matrix pronoun. But we consider that the post-nominal restrictives are direct imitations of the English restrictive relative clauses.

between ye and cheleṭi presently indicates that ye is the relative deictic, not the nonhonorific relative pronoun nor the complementizer, and (c) the dash between se and cheleṭi presently indicates that se is the demonstrative deictic, not the third person nonhonorific pronoun. (Aa) is a careless stacking of items, which may mislead the readers; but (Ab) is self-explanatory. Similar problems arise with elaborate and post-nominal restrictives, and with nonrestrictive clauses. We will follow the above mentioned convention in writing these relative clauses too.

In this chapter we will primarily be concerned with pre-nominal restrictive relative clauses, as they are the representative of the restrictive relative clauses in Bengali. A pre-nominal restrictive relative clause is placed pre-nominally (that is, immediately preceding its containing noun phrase) if a restrictive relative sentence contains more than one restrictive relative clause. Otherwise the restrictive clause is extraposed to the sentence-initial position. Consider the examples in (13, 14, 15).

- (13) a. ye-meyeṭi_c nācche, se_c sundar: The girl who is dancing is beautiful.
- b. ye-bhadralok_c āschen, tini_c kamyunist: The gentleman who is coming is a communist.
- (14) a. āmi ye-meyeṭi_c nācche, tāke_c cini: I know the girl who is dancing.
- b. ye-meyeṭi_c nācche, āmi tāke_c cini:
- (15) a. ye-meyeṭi_c gān gāiche, se_c ye-cheleṭi_a behālā bājacche, tāke_a bhālabāse: The girl who is singing loves the boy who is playing a violin.
- b. *ye-meyeṭi_c gān gāiche ye-cheleṭi_a behālā bājacche, se_c tāke_a bhālabāse:

In (13) the restrictive relative clauses immediately precede their containing noun phrases, and as the containing noun phrases are the initial constituents of their own simplexes, the relative clauses come automatically at the sentence-initial position. In (14a) the restrictive clause immediately precedes its containing noun phrase, and as the containing noun phrase is a sentence-medial constituent, the relative clause is inside the sentence. As it has a single restrictive clause, the relative

clause can be moved to the sentence-initial position. Bengali has a tendency to extrapose a restrictive clause sentence-initially if the sentence has a single restrictive relative clause. If we move the relative clause in (14a) to the sentence-initial position, we will get (14b). (14b) is more natural than (14a). (15a) has two restrictive relative clauses, each being placed preceding its containing noun phrase. As this sentence has two restrictive clauses, they cannot be stacked at the sentence-initial position. If we stack the relative clauses sentence-initially, we will derive the ungrammatical sentence (15b).

We assume that elaborate restrictives are derived from the pre-nominal restrictives (cf., § 8.4.1). In this type the antecedent is mentioned first and is followed by a relative clause. Then comes the matrix sentence containing a pronoun which refers to the antecedent through the relative pronoun in the relative clause. If we convert the pre-nominal restrictive relative sentences in (13) into elaborate restrictives, we will derive the sentences in (16), respectively.

- (16) a. meyeṭi_c, ye_c nācche, se_c sundar: Lit., The girl who is dancing, she is beautiful.
- b. bhadralok_c, yini_c āschen, tini_c kamyunistṭ: Lit., The gentleman who is coming, he is a communist.

These sentences are synonymous with those in (13), although they differ structurally.

The post-nominal (Anglicized) restrictives are exactly like the restrictive relative clauses in English, with one difference: that we enclose the relative clause in Bengali between commas. In this type, the relative clause follows the containing noun

phrase. The sentences in (13) will be those in (17) if post-nominal restrictive relative clauses are formed:

- (17) a. meyeṭi_c, ye_c nācche, sundar: The girl who is dancing is beautiful.
- b. bhadralok_c, yini_c āschen, kamyuniṣṭ: The gentleman who is coming is a communist.

The sentences in (17) are ambiguous in that the relative clauses can be taken either restrictively or nonrestrictively. Although they are grammatical, they do not sound like Bengali. They are imitations of English restrictive relative clause sentences.

We will now consider whether a restrictive relative clause can be placed sentence-finally. If we place the relative clauses in (17) sentence-finally, we will get the sentences in (18) respectively.

- (18) a. ?meyeṭi_c sundar, ye_c nācche:
- b. ?bhadralok_c kamyuniṣṭ, yini_c āschen:

These sentences are odd, if not ungrammatical.

In § 8.0 we have discussed the position of an NP-Rel in a relative clause briefly. There is no fixed position for a relativized noun phrase in Bengali. The question of its position is bound with the question of constituent-order in Bengali. It is beyond our scope to deal with the order of constituents here (cf., § 3.3), but some comments on the position of an NP-Rel in a relative clause are in order. Consider the examples in (2), which we reproduce below as (19) for convenience in discussion.

- (19) a. ye-cheleṭi kabitā likhto: The boy who used to write poems.

b. kabitā likhto ye-cheleṭi:

c. kabitā ye-cheleṭi likhto:

In (19) a restrictive relative clause has been given in three different orders of constituents. In each example the NP-Rel ye-cheleṭi is the subject of the clause, but it is an initial constituent in (19a), a final constituent in (19b) and a medial constituent in (19c). This shows that there is no fixed position for a relativized noun phrase in Bengali. Its position is determined by the same principles which determine the order of constituents in the surface structure of simple sentences. Many factors such as emphasis, topicalization, functional relation held by the constituent, stylistic preference etc., determine the position of a constituent in the surface, and all these apply to a relativized noun phrase, too.

Although Bengali allows a relativized noun phrase to be placed at any position in a relative clause, it seems that the relativized noun phrases have a preferred position, and that is the initial position of the relative clause. The relativized noun phrases are, in some sense, topicalized noun phrases, and accordingly they prefer the initial position of the relative clause.

8.2.2 Proper Nouns, Unique Reference and Restrictive Relativization.

A noun phrase which has a proper noun with unique reference as its head cannot be relativized restrictively. This can be seen in the examples in (20).

(20) a. *ye-jībanānanda_cbiṣaṅṅa, tini_cāmār priya kabi:

Jibanananda who is melancholic is my favourite poet.

b. *ye-dhākā_c bāṅlādeśer rājdhānī, āmi sekhāne_c thāki:

I live in that Dacca which is the capital of Bangladesh.

(20a) suggests the existence of several Jibanananda's, and (20b) suggests the existence of several Dacca's. These sentences are unacceptable if we want to maintain the unique references associated with these names. The purpose of restrictive relativization is to restrict the right person, object etc., from a set. So it is not possible if the noun phrase refers uniquely. Proper nouns do not always refer uniquely. It happens quite often that a set of persons (and sometimes a set of places) have the same name, and in such a case restrictive relativization is possible on noun phrases with proper names.

(21) a. ye-nīlimā_c padya likhe, se_c nīl sārī pareche: The Nilima who writes verse has put on a blue sari.

b. ye-hāsān_c kamyunist, se_c ekhan jele āche: The Hasan who is a communist is in jail now.

The sentences in (21) are grammatical, although the relativized noun phrases have proper nouns as their head elements. The proper nouns in these sentences do not refer uniquely. We can conclude that what disallows restrictive relativization is uniqueness of reference, not a proper name. Anything that has unique reference disallows restrictive relativization as this process destroys the uniqueness of reference. The following examples are ungrammatical because the noun phrases having unique reference have been restrictively relativized:

(22) a. *ye-cādti_c ālo dey, se-ṭi_c ekṭi upagraha: The moon which gives light is a satellite.

b. *nīlimār ye-svāmī_c dhākā thāken, tini_c ukil: Nilima's husband who lives in Dacca is a pleader.

Furthermore, speaking from impressions, these sentences do not sound like genuine Bengali sentences. A speaker of Bengali who is not influenced by the English relative clause sentence will not utter such sentences either restrictively or nonrestrictively. If he intends the restrictive, a speaker of Bengali will prefer, we assume, a pre-nominal restrictive structure, and produce the sentences in (26) for those in (23).

- (26) a. ye-ḍakṭar hāsān_c kamyuniṣṭ, tini_c baktritā diben: The Dr Hasan who is a communist will give a lecture.
 b. ye-ketakī_c nīl sārī pareche, se_c berāl bhālabāse: The Ketaki who has put on a blue sari loves cats.

Now consider the relative clauses in (23) in the non-restrictive sense. In this sense, the reference or identity of the containing noun phrases ḍakṭar hāsān and ketakī is well-established, and the relative clauses are parenthetical statements about them. A speaker of Bengali, in such a case, instead of forming a relative clause on the containing noun phrase, will make two separate statements on the containing noun phrase. The sentences in (23), in the nonrestrictive sense, will be uttered something like those in (27).

- (27) a. ḍakṭar hāsān_c baktritā diben. tini_c kamyuniṣṭ: Dr Hasan will give a lecture. He is a communist.
 b. ketakī_c berāl bhālabāse. se_c nīl sārī pareche: Ketaki loves cats. She has put on a blue sari.

Nonrestrictive relativization is not common in Bengali. But when a nonrestrictive relative clause is formed in Bengali, it is usually formed on a proper name with well-established identity. We see above that the nonrestrictive relative clause sentences in (23) have been converted into separate but anaphorically related sentences in (27).

8.3 Containing Noun Phrases and Restrictively Relativizable Noun Phrases (NP-Rel's).

In this section we will investigate such areas as

(a) functional relations of containing noun phrases (that is, what types of noun phrase can act as a containing noun phrase), (b) functional relations of NP-Rel's (that is, what types of noun phrase can be restrictively relativized), and (c) how much of a 'load' of restrictive relative clauses a Bengali relative clause sentence can bear. Keenan and Comrie (1972) investigated the functional relations of relativizable noun phrases in some forty languages, and developed their 'noun phrase accessibility hierarchy'. That is, they dealt only with our (b) above. But as they have not dealt with our (a) and (c), a total picture of restrictive relative clause sentences is not available there. Keenan and Comrie (1972) have proposed the following accessibility hierarchy:

(28) ACCESSIBILITY HIERARCHY (AH)

(i) Subj \gg DO \gg IO \gg OPrep \gg Poss-NP \gg O-Comp-Particle

(ii) if X \gg Y and Y dominates Z then X \gg Z

This hierarchy says that a subject noun phrase is easier to relativize than an object noun phrase, and so on (' \gg ' means "greater than or equal to in accessibility"); and so if a language allows relativization of a noun phrase that is at the right side of the arrow, then it will allow relativization of a noun phrase which is at the left side of the arrow.

We will modify this hierarchy for our purpose, and apply it to relativizable noun phrases as well as to containing noun phrases. Our hierarchy is given in (29).

(29) Accessibility Hierarchy in Bengali.

- a. SUBJ \gg DO \gg IO \gg LOC \gg TEMP \gg INS \gg GEN-NP \gg O-COMP-PARTICLE⁴
 b. If X \gg Y and Y dominates Z then X \gg Z

This hierarchy will apply both to the containing noun phrases and to the relativizable noun phrases. It says that a subject noun phrase can act as a containing noun phrase more easily than a direct object noun phrase, and so on; and that a subject noun phrase is easier to relativize than a direct object noun phrase, and so on. This hierarchy gives a generalization for containing noun phrases as well as for relativizable noun phrases. The hierarchy (29) does not only state the facts of Bengali relative clause sentence, but also implies some interesting consequences. The implication is that a relative clause sentence where the containing noun phrase and the relativized noun phrase both are subjects of their own clause, is more comfortable than one where the noun phrases involved in relativization are not subjects. Consider the examples in (30).

- (30) a. ye-cheleṭi_c gān gāiche, se_c ketakīke cene: The boy who is singing knows Ketaki.
 b. ketakī ye-cheleṭike_c bhālabāse, nīlimā tāke_c pachanda kare nā: Nilima does not like the boy who Ketaki loves.

In (30a) the relativized noun phrase ye-cheleṭi and its containing noun phrase se are subjects of their respective clauses, but the relativized noun phrase ye-cheleṭike and its containing noun phrase tāke in (30b) are objects of their own clauses. (30a) is more comfortable than (30b).

4 This list of noun phrase-types is not extensive, but it includes all major types of noun phrase.

In the accessibility hierarchy (29) we have indicated the functional relations of the noun phrases, but nothing has been said about their structure. In Bengali simple and compound noun phrases can be relativized, and they can act as a containing noun phrase, but complex noun phrases cannot be relativized. Consider the examples below.

- (31) a. ye-meyeṭi o ye-cheleṭi eseche: Lit., Who-the girl and who-the boy have come.
 b. ye-cheleṭi o meyeṭi eseche: The boy and the girl who have come.

In (31a) the conjoined noun phrases have been individually relativized, and in (31b) the entire conjoined noun phrase has been relativized. In (31a) the referents are understood distributively and they are understood as a set in (31b).

8.3.1 Containing Noun Phrases.

In (29) we have shown which noun phrases can act as containing noun phrases, and it also shows their relative accessibility. We give examples below to show that this hierarchy states the facts of Bengali relative clause correctly. Consider the examples below.

Containing noun phrase as SUBJ:

- (32) a. ye-meyeṭi_celo, se-meyeṭi_cnācbe: The girl who came will dance.
 b. ye-meyeṭi_cnācbe, se-meyeṭi_celo: The girl who will dance came.

Containing noun phrase as OBJ (DO and IO):

- (33) a. ye-meyeti_cnāclo, kerāmat ālī se-meyeṭike_cbhālabāse: Keramat Ali loves the girl who danced.

- b. ye-meyeti_c nāclo, kerāmat ālī se-meyetike_c ekṭi phul
dicche: Keramat Ali is giving a flower to the
girl who danced.

Containing noun phrase as a LOC, TEMP, INS, GEN and O-COMP-
PARTICLE etc., noun phrase:

- (34) a. āmarā ye-deṣe_c yābo, se-deṣe_c jal nei: There is no
water in that country where we will go.
- b. āmi ye-samay_c phirchilām, se-samay_c ākāṣe cād chilo:
The moon was in the sky at the time when I was
coming back.
- c. ye-churiṭi_c natun, mātin se-churiṭi_c diye hāt keteche:
Matin has cut his hand with the knife that is new.
- d. ye-bhadralok_c āschen, se-bhadraloker_c bau pāliyeche:
The wife of the gentleman who is coming has eloped.
- f. ye-cheleṭi_c āsche, mātin se-cheleṭir_c ceye choṭa:
Matin is shorter than the boy who is coming.

All the relativized noun phrases in (32)-(34) contain the relative deictic ye, and the containing noun phrases have the demonstrative deictic se. These examples show that our accessibility ^{hierarchy} states the facts correctly. The sentences above have been arranged in order of relative accessibility: among these sentences a sentence is more than, or equally comfortable to the other which follows it.

8.3.2 Relativizable Noun Phrases (NP-Rel's).

We have given the accessibility hierarchy of relativizable noun phrases in (29). It shows that the range from subject noun phrases to objects of comparative constructions can be restrictively relativized in Bengali. It also shows that a

subject noun phrase is easier to relativize than a direct object noun phrase, and a direct object is easier to relativize than an indirect object noun phrase, and so on. We give examples of different types of relativized noun phrase below (only relative clauses are given):

NP-Rel as SUBJ:

- (35) a. ye-mahilā ātmahatyā karechen: The woman who has committed suicide.
 b. ye-pākhīṭi nīl: The bird which is blue.

NP-Rel as OBJ (DO and IO):

- (36) a. tumi ye-meyēṭike ceno: The girl who you know.
 b. tumi ye-meyēṭike ciṭhi likhecho: The girl who you have written a letter to.

NP-Rel as LOC, TEMP, INS, GEN and O-COMP-PARTICLE etc., noun phrase:

- (37) a. āmarā ye-deṣe yābo: The country where we will go.
 b. āmarā ye-din yābo: The day we will go.
 c. mātīn ye-kalamṭi diye likhche: The pen which Matin is writing with.
 d. ye-phulṭir raṅ lāl: The flower whose colour is red.
 e. ye-cheleṭi mātiner ceye choṭa: The boy who is shorter than Matin.
 f. mātīn ye-cheleṭir ceye choṭa: The boy who Matin is shorter than.

In these examples all the NP-Rel's have the relative deictic ye, and they show that Bengali allows restrictive relativization through the whole hierarchy given in (29). The sentences in (35)-(37) are all grammatical, but they are not equally comfortable.

8.3.3 How much 'Load'?

We have so far considered restrictive relative clause sentences with a single relative clause. Bengali relative clause sentences (both restrictive and nonrestrictive), with more than one relative clause, create problems in performance and comprehension. The relative clauses in Bengali, speaking impressionistically, are something like 'loads', which make a sentence 'heavy'. Let us call a relative clause sentence having a single relative clause a single-loaded sentence or a sentence with a single load, and one with two relative clauses a double-loaded sentence and so on. In order to limit our scope we will deal with restrictive relative clause sentences only in this matter.

We have seen that a pre-nominal restrictive relative clause in Bengali is extraposed to the sentence-initial position if a sentence has a single restrictive clause. Otherwise, it is placed immediately preceding its containing noun phrase.

Consider the examples in (38).

- (38) a. ye-meyeti₁ nācche₂, se₃ camatkār₄: The girl who is
 1 2 3 4 5 2 1
 dancing is lovely.
 3 5
- b. kerāmat ālī ye-meyeti₂ nācche₃, tāke₄ cithi likheche₅:
 2 3 4 5 6 7
 Keramat Ali has written a letter to the girl
 7 6 2,3,
 who is dancing.
 5 4
- c. ye-meyeti₁ nācche₂, kerāmat ālī tāke₅ cithi likheche₆:
 1 2 3 5 6 7
 Lit., Who-the girl is dancing, Keramat Ali has
 1 2 3 —
 written a letter to her.
 7 6 5

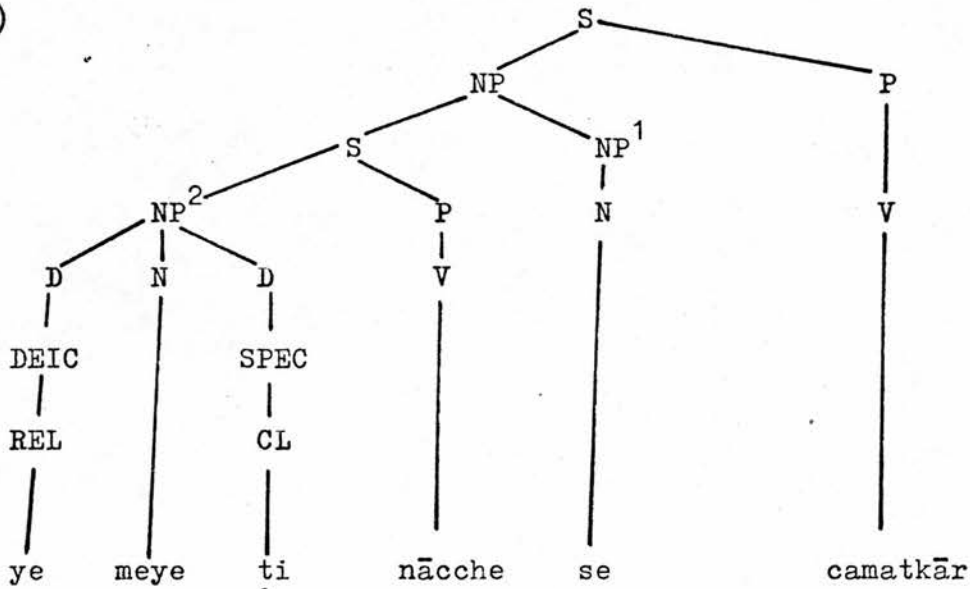
(38a) is a single-loaded sentence, where the relative clause is

placed immediately preceding its containing noun phrase se.

As the containing noun phrase is the initial constituent of its own simplex, the relative clause automatically comes at the initial position of the entire sentence. In (38b) the relative clause immediately precedes its containing noun phrase tāke, and as the containing noun phrase is a medial constituent of its own simplex, the relative clause appears inside the sentence.

This sentence has a single load; so it is preferable to extrapose the relative clause to the sentence-initial position. If we extrapose the relative clause to the sentence-initial position, we will derive (38c). (38c) is synonymous with, but structurally more comfortable than (38b). These sentences are single-loaded, and they do not create any problem in performance and in comprehension. For example, (38a) has the structure (39).

(39)

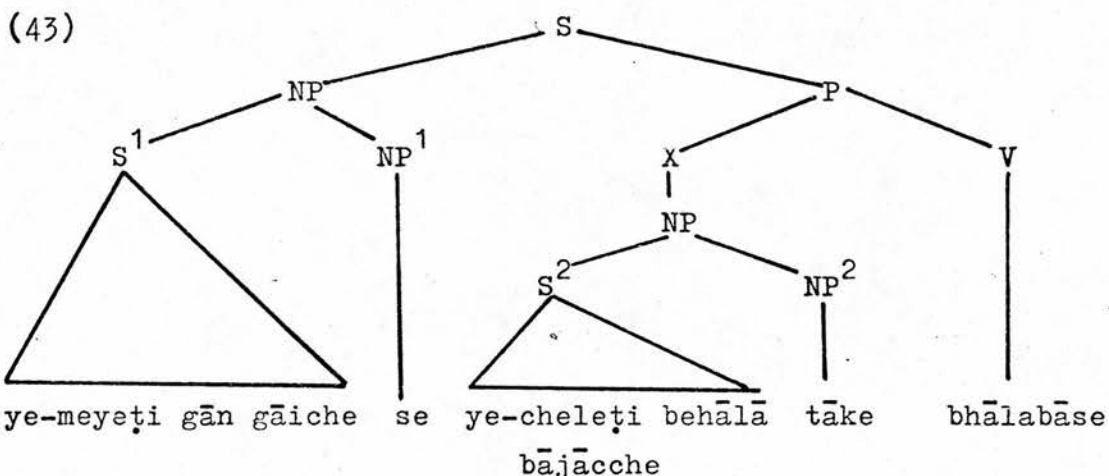


The structure (39) has a single restrictive relative clause embedded in its containing noun phrase (NP¹) by Chomsky-adjunction. Such a structure is natural in Bengali.

The pre-nominal restrictive relative clause sentences in (38) can be converted into elaborate restrictives. When

b. *ye-meyeṭi_c gān gāiche ye-cheleṭi_a behālā bājācche,
 se_c tāke_a bhālabāse:

(42a) is a double-loaded restrictive relative sentence, where the relative clauses are placed pre-nominally. It is a fairly heavy sentence. As this sentence has two relative clauses, they cannot be extraposed to the sentence-initial position together. If they are moved to the sentence-initial position, we will derive the ungrammatical sentence (42b). Thus we see that relative clauses must precede their containing noun phrases in a relative clause sentence which has more than one load. (42a), although grammatical, is heavy and poses a production and comprehension problem. (42a) has the structure (43) (irrelevant details omitted).



(43) has two restrictive relative clauses embedded in two distinct noun phrases. This structure is heavy, but not incomprehensible.

But the problem of production and comprehension multiplies in double-loaded elaborate, and anglicized restrictive relative clause sentences. The elaborate and anglicized versions of (42a) are given in (44a, b) respectively. They are exceedingly odd.

- (44) a. ?meyeti_c, ye_cgān gāiche, se_ccheleti_a, ye_abehālā
 bājācche, tāke_abhālabāse: Lit., The girl who is
 singing she, the boy who is playing a violin,
 loves him.
- b. ?meyeti_c, ye_cgān gāiche, cheleti_a, ye_abehālā
 bājācche, bhālabāse: The girl who is singing loves
 the boy who is playing a violin.

Pre-nominal restrictive relative clauses in Bengali are left-embedded elements in their containing noun phrases. In Bengali it is possible to embed a relative clause in a noun phrase; and then embed the entire structure in another noun phrase of another sentence; and then embed the whole structure in another noun phrase, ad infinitum. This sort of structure is quite simple to produce, but it does not make any sense. It is agreed nowadays that self-embedded structures are difficult for perception (cf., Chomsky (1965, 10-15), Lyons (1970c, 88-95), Kuno (1974)). Bengali disallows self-embedding or centre-embedding. But Bengali allows a long chain of left-embeddings, although such a structure becomes incomprehensible or nonsensical beyond a certain limit. Consider the examples below.

- (45) a. !ye-meyeti_cgān gāiche, se_cye-cheleti_abhālabāse, tāke_a
 ye-behālāti_iupahār diyechē, se-ṭi_icamatkār: The
 violin which the girl who is singing has presented to
 the boy who she loves is marvellous.
- b. ekṭi meye_cgān gāiche ebaṅ se_cekṭi cheleke_abhālabāse
 ebaṅ se_ccheleti_aekṭi behālā_iupahār diyechē ebaṅ
 se-behālāti_icamatkār: A girl is singing and she loves
 a boy and she has presented a violin to that boy
 and that violin is marvellous.

(46) a. !ye-meyeti_c gān gāiche, se_c ye-cheletike_a bhālabāse,
tāke_a ye-behālāti_i upahār diyeche, se-ti_i ye-lok_j
tairī kareche, se_j ye-meyetike_e biye karbe, se_e
kakhano behālā 'sone ni: (see gloss for (46b)).

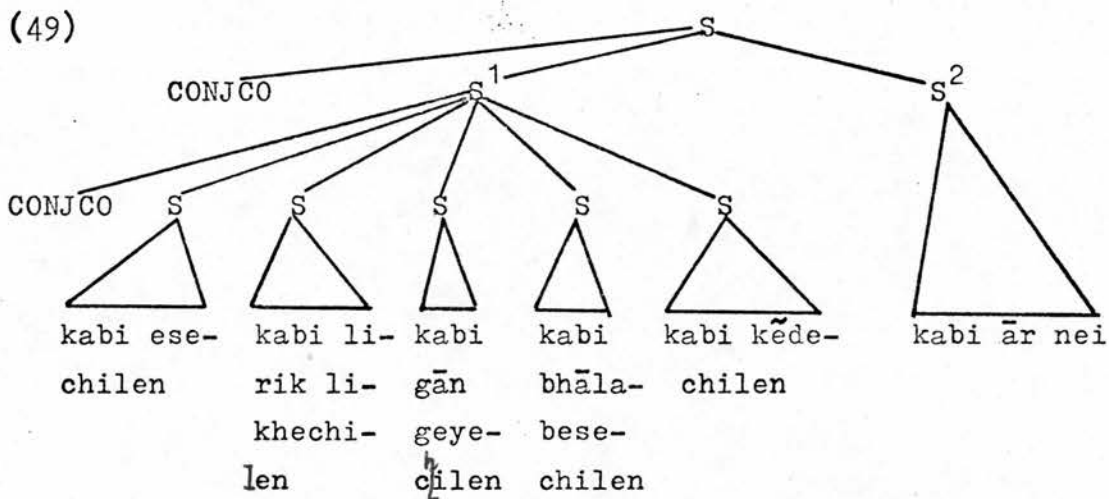
b. eḳti meye_c gān gāiche ebaṅ se_c eḳti cheleke_a bhālabāse
ebaṅ se_c tāke_a eḳti behālā_i upahār diyeche ebaṅ eḳti lok_j
se-behālāti_i tairī kareche ebaṅ se_j eḳti meyeke_e biye
karbe ebaṅ se-meyeti_e kakhano behālā 'sone ni:

A girl is singing and she loves a boy and she has
presented him a violin and a man has made that violin
and he will marry a girl and that girl has never
listened to a violin.

(45, 46a) are instances of multiple embeddings. These sentences are produced mechanically by quite simple rules: first make a restrictive relative clause; embed it in a noun phrase of a sentence; then embed the entire structure in another noun phrase; and so on. (45, 46a) can be paraphrased respectively as (45, 46b), which are stylistically infelicitous but semantically equivalent to (45, 46a). We assume that (45, 46a) are relativized versions of (45, 46b) respectively. We have seen that (42a) has two loads, but its difference from (45, 46a) is that while (42a) has two distinct loads, (45, 46a) have echelons of relative clauses one embedded in the other, and so on. For example, (45a) has the structure (47) (see page 347). The surface structure (47) has an echelon of relative clauses. This sort of structure generates nonsensical sentences.

Although relative clause structures with an echelon of restrictive relative clauses produce nonsensical sentences, Bengali restrictive relative clause sentences where the

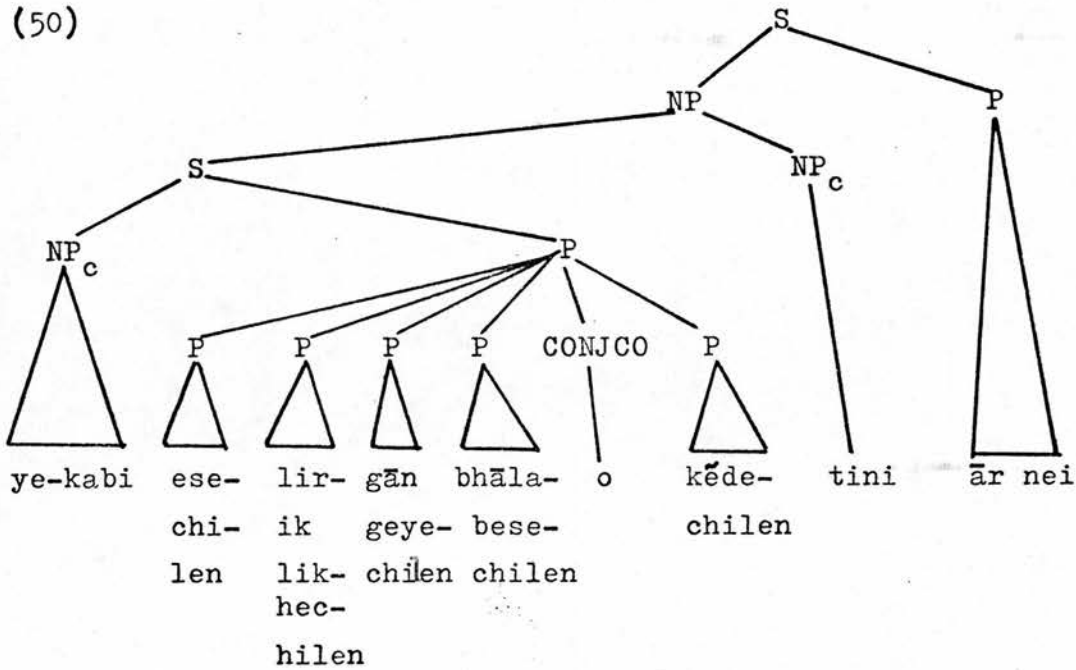
where the coordination deletion rule has applied (cf., Koutsoudas (1971)). The second way of dealing with these sentences offers the possibility that the noun phrases involved in relativization in one relative clause structure, may not be understood coreferentially with the noun phrases involved in relativization in another relative clause structure. But in (48) the relativized noun phrase and its containing noun phrase are coreferential in each example. So we assume that in sentences like (48) the relative clause is a coordination of several sentences. For example, (48a) has the underlying structure (49).



In (49) S^1 and S^2 are coordinately conjoined, and S^1 is a coordination of several sentences. In (49) each conjunct of S^1 has a noun phrase which is coreferential with one in the other conjunct; and those noun phrases are coreferential with one in S^2 . The anaphoric noun phrase deletion rule will apply in S^1 , and subsequently, the restrictive relativization rule will apply in (49). After the application of the restrictive relativization rule in (49), the pronominalization proper rule will reduce the containing noun phrase into a pronoun. In this way we will derive the surface structure (50) from

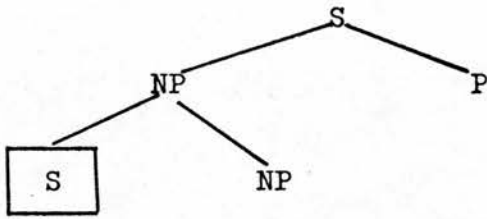
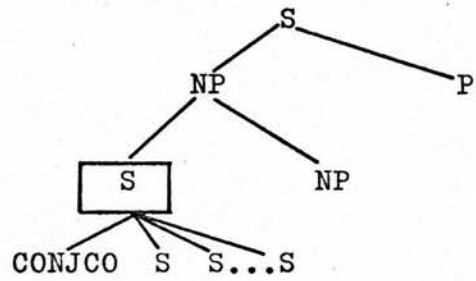
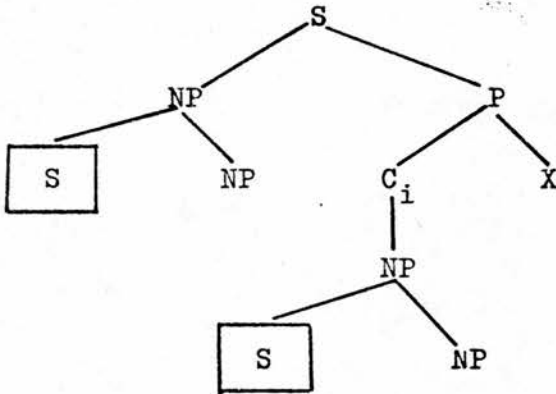
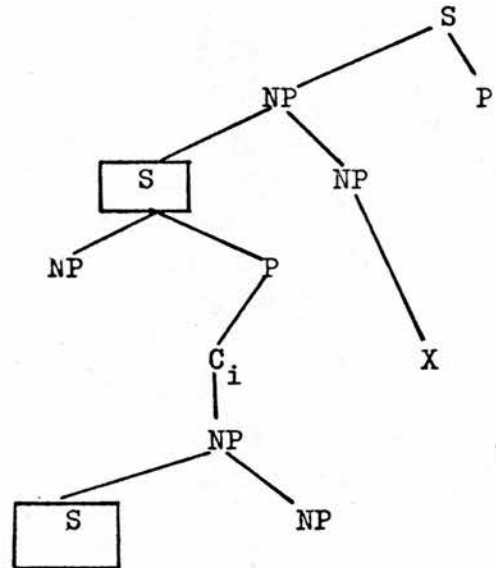
the structure (49).

(50)



(50) will generate (48a). (50) is a natural structure in Bengali. It is as comfortable as a single-loaded sentence. Thus we can arrive at the following conclusions about restrictive relative clause sentences in Bengali (these will apply equally to nonrestrictive relatives):

- (51) a. Left-embedded single-loaded sentences are natural in Bengali (cf., schema (52a)). Restrictive relative clause sentences where the relative clause is a coordination of several sentences are similarly natural (cf., schema (52b)).
- b. Double-loaded sentences where relative clauses are embedded in distinct noun phrases are heavy, but comprehensible (cf., schema (52c)).
- c. Echelons of relative clauses make a structure nonsensical (cf., schema (52d)).

(52) a. Single-load: 'Natural'b. Coordinate single-load: 'Natural'c. Double-load: 'Heavy'd. Echelon of loads: 'Nonsensical'

Relative clauses are indicated by rectangles in (52).

8.4 The Derivation of Restrictive Relative Clause Sentences.

Transformational grammarians are not unanimous about the underlying structure of restrictive relative clause sentences. There have been at least four analyses proposed so far for restrictive relative clauses: (a) the ART S analysis (cf., Smith (1964), Chomsky (1965)), (b) the NP S analysis (cf., Ross (1967a), Jacobs and Rosenbaum (1968)), (c) the NOM S analysis (cf.,

Stockwell et al (1973), and (d) the underlying conjunction analysis (cf., Postal (1967), Thompson (1971)). A brief survey of these analyses is available in Stockwell et al (1973). Of these analyses the underlying conjunction analysis has been tried least in transformational studies; but we think that it is the right way to proceed for an analysis of restrictive relative clause sentences. Here we will not argue against the ART S, NP S and NOM S analyses as acute arguments have already been made in Thompson (1971). But we shall make a point which has been ignored by all who have dealt with restrictive relative clause sentences.⁵ Consider the examples below.

- (53) a. ye-cheleṭi_ceseche, se_cgān gāibe: The boy who has come will sing.
 b. ye-cheleṭi_cgān gāibe, se_ceseche: The boy who will sing has come.

In ART S, NP S and NOM S analyses the semantic and syntactic differences between (53a, b) will be mechanically attributed to the interchange of matrix and constituent sentences. This analyses do not explain why the relative clause is ye-cheleṭi eseche in (53a), and ye-cheleṭi gān gāibe in (53b). What remains unexplained in these analyses is a major condition for restrictive relativization. We assume that restrictive relativization is dependent on the presupposition of the speaker. The underlying conjunction analysis takes presupposition into account, and shows deeper insight into restrictive relativization.

5 As it is practically impossible to deal with all aspects of relativization in this work, we must confine ourselves to the derivation of restrictive relative clause structures with a single relative clause.

Now we come back to the sentences in (53). We will suggest that the sentences in (53) have an underlying structure something like (54).

(54) S [CONJCO S [ekṭi chele_c eseche] S [ekṭi chele_c gān gāibe]]:

(A boy has come)(A boy will sing).

Each conjunct in (54) has a noun phrase which is coreferential with a noun phrase in the other conjunct. That is, (54) contains two propositions about the same argument ekṭi chele. If the speaker presupposes that the hearer knows the first proposition, but does not know the second one; ~~the~~ the hearer's supposedly known proposition will be the relative clause, and (53a) will be produced. Instead, if the speaker supposes that the hearer knows the second proposition, but does not know the first one; then (53b) will be generated by restrictive relativization. If the hearer knows none of the propositions, then relativization of either of the propositions will be puzzling to the hearer. In that situation any of the following coordinate conjoined sentences will be appropriate:

(55) a. ekṭi chele_c eseche ebaṅ se_c gān gāibe: A boy has come and he will sing.

b. ekṭi chele_c gān gāibe ebaṅ se_c eseche: A boy will sing and he has come.

So we see that the speaker's presupposition plays an important part in restrictive relativization. When there are two propositions on the same argument, the hearer's supposedly known proposition is converted into a restrictive relative clause. But the question will arise as to how we can account for the presupposition of the speaker in the underlying structure of

coordinate conjoined structures? This can be done if we assume that coordinate conjoined structures are ordered in the underlying structure in accordance with the speaker's presupposition.

That is, in a coordinate conjoined structure with two conjuncts, the first conjunct contains information which is known to the hearer, and so it will be converted into a restrictive clause. In such a formulation (53a, b) will be derived from (56a, b) respectively.

- (56) a. S [CONJCO S [ekṭi chele_c eseche] S [ekṭi chele_c gān gāibe]] :
 (A boy has come)(A boy will sing).
 b. S [CONJCO S [ekṭi chele_c gān gāibe] S [ekṭi chele_c eseche]] :
 (A boy will sing)(A boy has come).

In (56) the first conjuncts are supposed to be known to the hearer, and so they will be converted into restrictive relative clauses.

It can be seen from the above examples that the containing noun phrases and the relativized noun phrases have been taken as indefinite in the underlying structure, but they become definite by the application of the restrictive relative transformation. This is because restrictive relativization in Bengali is a definitization transformation (cf., Kuroda (1968), Browne (1970)). The noun phrases involved in restrictive relativization are marked for coreference in the underlying structure and are definitized by the restrictive relativization rule (but cf., below).

Now we will consider how pre-nominal restrictive relative clause sentences in Bengali are generated. Consider (57a), which has the underlying structure (57b).

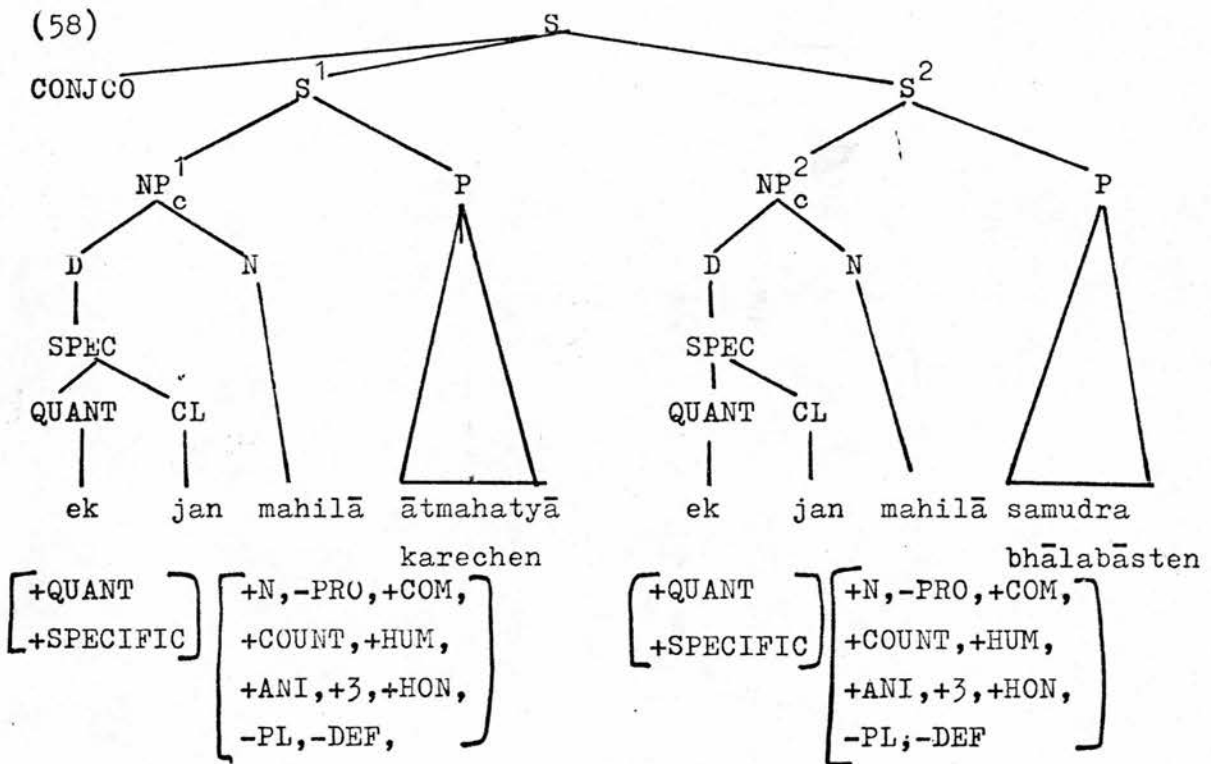
(57) a. ye-mahilā_cātmahtyā karechen, tini_csamudra bhālabāsten:

The woman who has committed suicide used to love
the sea.

b. S [CONJCO S [ekjan mahilā_cātmahtyā karechen]
 S [ekjan mahilā_csamudra bhālabāsten]]:

(A woman has committed suicide)(A woman used to love
the sea).

(57b) has the structure (58) when the restrictive relativization transformation applies.



(58) has two conjuncts, and each conjunct has a noun phrase which is coreferential with a noun phrase in the other conjunct. Coreferentiality is a major condition for relativization. This is indicated by identical indices in (58). We have already said that the noun phrases involved in restrictive relativization are indefinite in the underlying structure. Nevertheless they become definite due to application of the restrictive

relativization rule. The relative deictic ye and the demonstrative deictic se(i) are inherently definite. The restrictive relative transformation inserts ye to the NP-Rel, and se(i) to the containing noun phrase, and these noun phrases become definite. But when the noun phrases involved have phonetically null nouns as heads (cf., § 8.1), ye and se(i) cannot make these noun phrases strictly definite. In such a case the noun phrases may be understood either as definite or as indefinite. So we assume that the underlying noun phrases which have the features [-DEF,+SPECIFIC] become definite by the restrictive relativization rule; but those with the features [-DEF,-SPECIFIC] do not become definite.

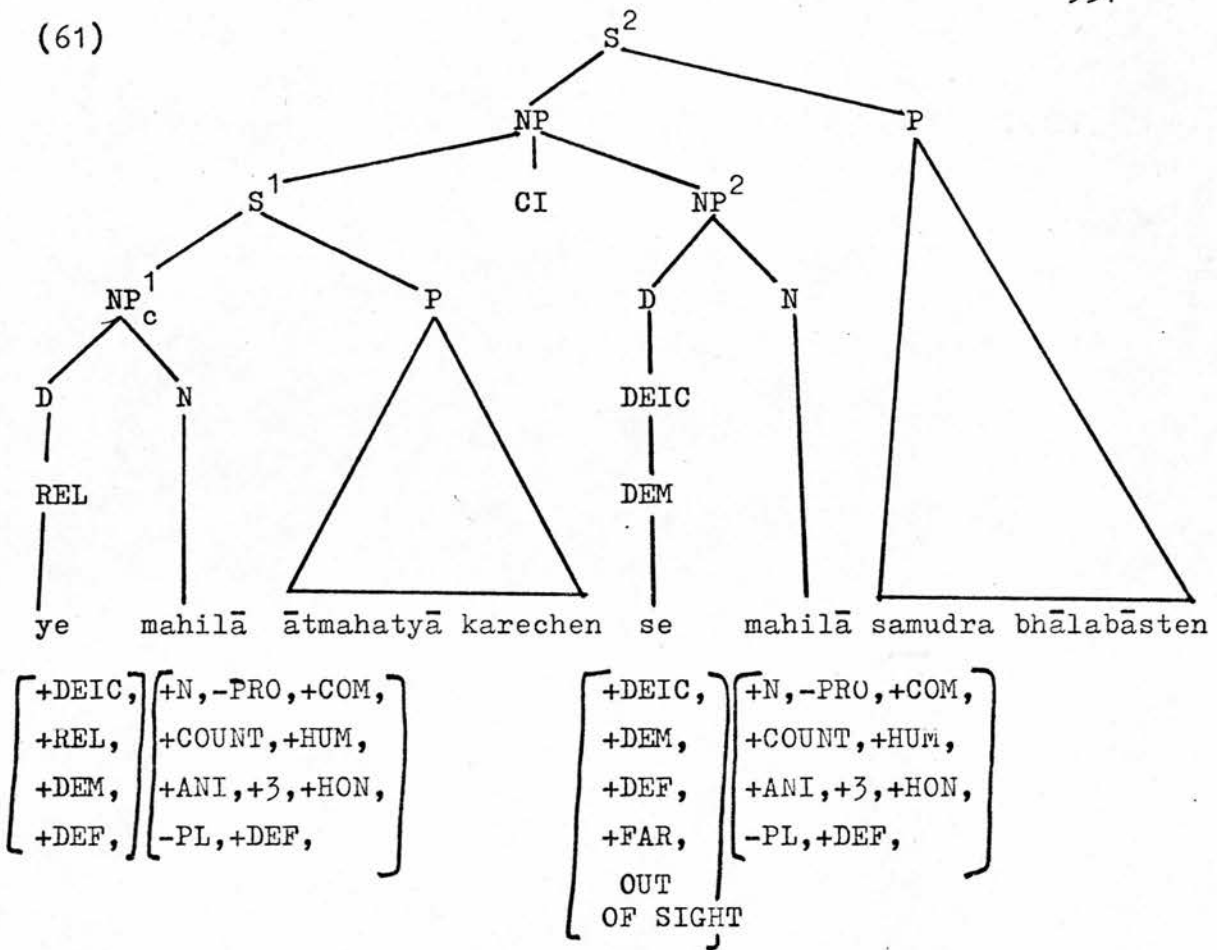
According to our formulation, S^1 in (58) contains the information which is known to the hearer; and so it will be the restrictive relative clause. The restrictive relative transformation will carry out the following operations in (58):

(59) Operations for Restrictive Relativization.

- a. Delete CONJCO.
- b. Insert the relative deictic ye in the NP-Rel as a DEIC, and the demonstrative deictic se(i) in the containing noun phrase as a DEIC.
- c. Adjoin the relative clause as the left daughter of the containing noun phrase by Chomsky-adjunction.
- d. Place comma intonation (CI) between the relative clause and the containing noun phrase.

The operations in (59) will delete the CONJCO from (58), insert ye to the NP-Rel (NP^1) and se(i) to the containing noun phrase (NP^2), and adjoin S^1 as the left daughter of the containing noun phrase by Chomsky-adjunction. In this way S^1 will lose its

(61)



(61) will generate ye-mahilā_c ātmahatyā karechen, se-mahilā_c samudra bhālabāsten: Lit., 'The woman who has committed suicide, that woman used to love the sea', which is a good and grammatical sentence (but may not be so good socially due to honorific feature of the noun phrases involved).

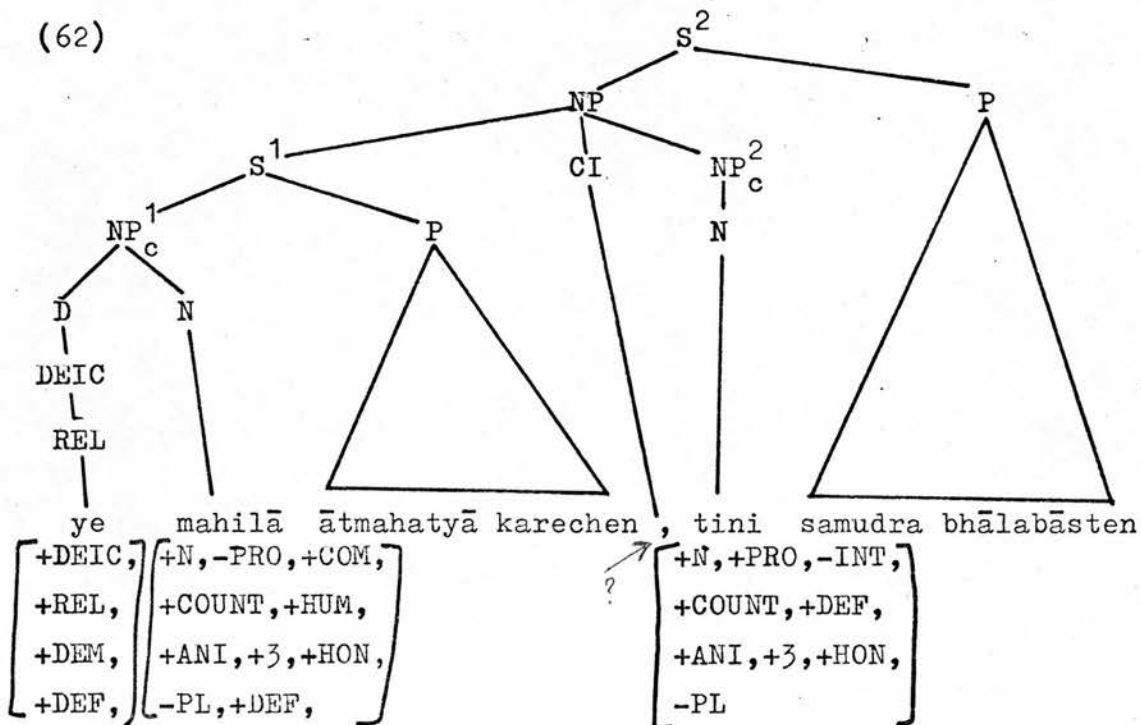
In Bengali the relativization transformation is followed by the pronominalization proper transformation, which derives personal and relative pronouns. Masica (1972) misunderstood this process. He said that after relative clause formation 'something odd happens: the noun remains in the relative clause, but is deleted from the main clause, leaving the correlative as a dummy to represent it.' What he misunderstood as a dummy is, in fact, a pronoun or pronominal. His misunderstanding

grew from the lexical identity of the third person nonhonorific pronoun se and the demonstrative deictic se(i). What happens in Bengali is that the containing noun phrase is pronominalized after the application of the restrictive relativization rule.

We will now consider how the pronominalization proper rule follows, and interacts with the restrictive relative rule. In (61) there are two coreferential noun phrases. This structure satisfies all the conditions for the pronominalization proper transformation. The NP-Rel (NP^1) and the containing noun phrase (NP^2) are coreferential in (61), and the pronominalization rule can apply in (61) both forwards and backwards. As forward pronominalization is an instance of unmarked pronominalization,

pronominalization usually applies forwards in a structure like (61). The pronominalization rule will reduce NP^2 into a pronoun. By the application of the pronominalization rule, and by second lexical insertion, we will derive (62) from (61).

(62)



The second lexical insertion rule has inserted the pronoun tini to the pronominalized noun phrase, and the comma intonation has also been placed in (62). (62) will generate ye-mahilā_cātmaḥatyā karechen, tini_csamudra bhālabāsten, which is (57a).

The relativized noun phrase (NP¹) in (62) is a full noun phrase, which contains the relative deictic ye. That is, NP¹ has been relativized, but has not been pronominalized. The restrictive relative rule in Bengali inserts the relative deictic ye as the left daughter of a relativized noun phrase, but it does not specify the NP-Rel as [+PRO]. In (62) the containing noun phrase has been pronominalized. Consider (61). We have seen that forward application of the pronominalization rule in (61) generates (57a), which is a natural restrictive relative sentence in Bengali. But the pronominalization rule can also apply backwards in (61). We can use the containing noun phrase (NP²) to reduce the relativized noun phrase (NP¹) into a pronoun. We know that NP¹ has already been affected by the relativization rule. If we apply the pronominalization rule backwards in (61), NP¹ will be affected by two rules: relativization and pronominalization.

Let us apply the pronominalization rule backwards in (61). This rule will specify NP¹ as [+PRO], and this feature will be added to the feature bundle associated with the head noun of NP¹. This rule will delete the relative deictic ye from NP¹, but will preserve the information that NP¹ has already been relativized. In order to preserve this information the pronominalization rule, while deleting the relative deictic ye from NP¹, will add the feature [+REL] to the feature bundle associated with the head noun of NP¹. These operations will

perfect, but socially objectionable). Consider the examples in (64, 65).

- (64) a. ye-bhadralok_cāschen, tini_cekjan adhyāpak: The gentleman who is coming is a professor.
 b. yini_cāschen, se-bhadralok_cekjan adhyāpak:
 (65) a. ye-cheleṭi_cāsche, se_ckhub duṣṭu: The boy who is coming is very naughty.
 b. ye_cāsche, se-cheleṭi_ckhub duṣṭu:

The noun phrases involved in relativization are honorific in (64) and nonhonorific in (65). In (64, 65a) the containing noun phrases have been pronominalized, and in (64, 65b) the relativized noun phrases have been pronominalized. (64, 65a) are more natural than (64, 65b) respectively. (65b), however, seems better than (64b). This is because the noun phrases involved are honorific in (64b) and nonhonorific in (65b).

We come now to relative clause sentences where the relativized noun phrase and the containing noun phrase are obligatorily pronominalized. We have seen that pronominalization of an NP-Rel and its containing noun phrase is obligatory in those cases where the noun phrases involved in restrictive relativization have phonetically null head nouns (cf., ϕ 8.1).

Such sentences are exemplified below.

- (66) a. ye_cājo āsābādī, se_cniścāy pāgal: $\left\{ \begin{array}{l} \text{The} \\ \text{Any} \end{array} \right\}$ one ([-HON])
 who is optimistic even to-day must be mad.
 b. yini_cājo āsābādī, tini_cniścāy pāgal: $\left\{ \begin{array}{l} \text{The} \\ \text{Any} \end{array} \right\}$ one ([+HON])
 who is optimistic even to-day must be mad.

In (66) pronominalization has applied obligatorily in the NP-Rel and in the containing noun phrase, because the noun phrases have head nouns which are phonetically null.

8.4.1 Elaborate Restrictives and 'Anglicized' Restrictives.

In an elaborate restrictive relative clause sentence the referent is mentioned first and is followed by a relative clause, and the matrix sentence. We have exemplified such sentences in (11), which we reproduce below as (67).

- (67) a. $\text{cheleṭi}_c, \text{ye}_c \text{kabitā} \text{likhto}, \text{se}_c \text{mare geche}$: See (11a) for gloss.
 b. $\text{mahilā}_c, \text{yini}_c \text{ātmahatyā} \text{karechen}, \text{tini}_c \text{samudra bhālabāsten}$: See (11b) for gloss.

In (67) the referents cheleṭi and mahilā are mentioned first in each sentence. They are followed by the relative clauses ye kabitā likhto and yini ātmahatyā karechen, and in turn the relative clauses are followed by the matrix sentences se mare geche, and tini samudra bhālabāsten in (67a, b) respectively. In these sentences the relative clauses each contains a relative pronoun which refers to the referent. But we have seen that in a pre-nominal restrictive relative sentence the relativized noun phrase contains the relative marker ye, and the containing noun phrase contains the demonstrative deictic se(i). In an elaborate restrictive relative sentence the referent noun phrase may or may not contain the demonstrative deictic se(i), but the NP-Rel cannot be a full noun phrase with the relative deictic ye. Consider the examples below.

- (68) a. $*\text{cheleṭi}_c, \text{ye-cheleṭi}_c \text{kabitā} \text{likhto}, \text{se-cheleṭi}_c \text{mare geche}$: = (67a)
 b. $*\text{mahilā}_c, \text{ye-mahilā}_c \text{ātmahatyā} \text{karechen}, \text{se-mahilā}_c \text{samudra bhālabāsten}$: = (67b)

These sentences are ungrammatical, because the referents have been mentioned first in each sentence, and the relativized noun

phrases are full noun phrases with ye, and the containing noun phrases are full sentences with se.

We assume that an elaborate restrictive relative sentence is derived from a pre-nominal relative clause sentence. Consider the sentences in (67). In these sentences the referents are mentioned first, which is a matter of topicalization and emphasis. Compare these sentences with their pre-nominal counterparts given below in that order.

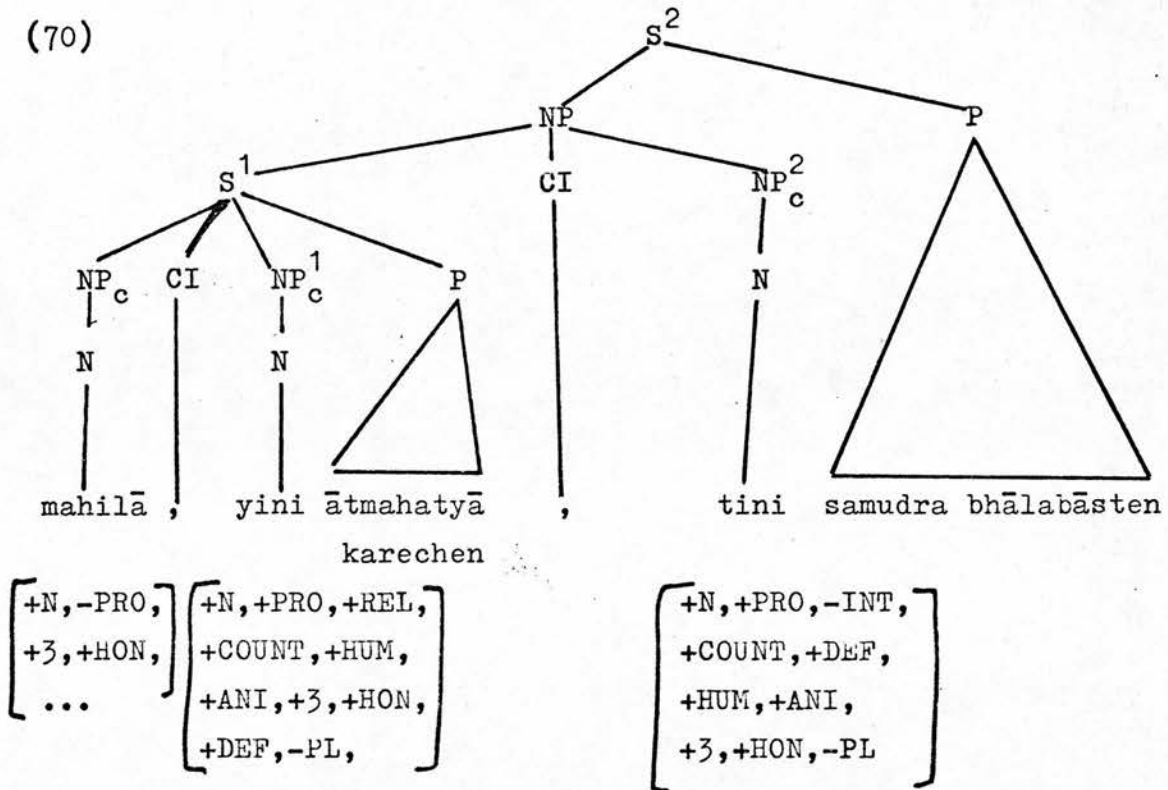
- (69) a. ye-cheleṭi_ckabitā likhto, se_cmare geche: The boy who used to write poems has died.
- b. ye-mahilā_cātmahtyā karechen, tini_csamudra bhālabāsten: The woman who has committed suicide used to love the sea.

In these sentences the relativized noun phrases contain the relative deictic ye, which is the initial element of the relativized noun phrase. In (67) the referents have been isolated from the relativized noun phrases, and have been placed sentence-initially; and the relativized noun phrases have been realized as relative pronouns. We assume that a pre-nominal restrictive relative clause is transformed into an elaborate restrictive by the topicalization of the referent. Let us show, for example, how (67b) is derived from (69b). The structure which immediately underlies (69b) is (62).

In (62) the NP-Rel (NP¹) consists of the relative deictic ye and the head noun mahilā. The process of topicalization moves the head noun including its definiteness feature to the sentence-initial position, and leaves other constituents including the features of the noun phrase behind. This process is followed by the application of the pronominalization rule

in the residual noun phrase. We assume that these operations derive a structure like (70) from (62).

(70)



The second lexical insertion rule has attached the relative pronoun yini to NP¹, as its features match with those of the relative pronoun. (70) will generate mahilā_c , yini_c ātmahatyā karechen , tini_c samudra bhālabāsten, which is (67b). Thus we see that an elaborate restrictive relative sentence is derived from a pre-nominal restrictive relative sentence.

Anglicized restrictive relative clauses are imitations of English restrictive relative clause sentences. They should be derived directly from an underlying coordinate conjoined structure by a rule which places the relative clause post-nominally. But a surface comparison between some elaborate restrictives and their anglicized counterparts shows that anglicized restrictives can be derived from elaborate restrictives. Consider the examples in (71).

- (71) a. mahilā_c, yini_c ātmahatyā karechen, tini_c samudra
bhālabāsten: = (67b)
- b. mahilā_c, yini_c ātmahatyā karechen, samudra bhalabasten:
The woman who has committed suicide used to love
the sea.

The sentences in (71) superficially show that the anglicized restrictive relative sentence (71b) can be derived from the elaborate restrictive relative sentence (71a) simply by deleting the matrix pronoun tini. But this analysis will be unacceptable because in (71a) mahilā is not the subject of any clause, but in (71b) mahilā is the subject of the matrix sentence. It is an accident that the deletion of tini in (71a) renders it identical to (71b). This sort of deletion is not possible in other elaborate structures. Consider the examples in (72).

- (72) a. meyeṭi_c, ye_c gān gāiche, āmi tāke_c cini: Lit., The girl
1 2 3 4 5 6 7 1
who is singing, I know her.
2 3,4 5 7 6
- b. *meyeṭi_c, ye_c gān gāiche, āmi cini:
- c. āmi meyeṭike_c, ye_c gān gāiche, cini: I know the girl
who is singing.

We derive the ungrammatical sentence (72b) by deleting the matrix pronoun tāke from (72a), which is an elaborate restrictive sentence. The anglicized counterpart of (72a) is (72c). They cannot be related to one another. We consider that anglicized restrictives are derived directly from underlying coordinate conjoined structure by a rule which places the relative clauses post-nominally.

8.4.2 Quantifiers, Generic Noun Phrases and Restrictive Relativization.

Noun phrases with numeral quantifiers can be restrictively relativized, but noun phrases with quantifiers such as anek: 'Many', bahu: 'Many' and pratyek: 'Each' etc., disallow pre-nominal restrictive relativization. Consider the examples below.

(73) a. *ye- $\left\{ \begin{array}{l} \text{anek} \\ \text{bahu} \end{array} \right\}$ chele_cphel kareche, tārā_csamasyā sriṣṭi karbe: Many boys who have failed will create problems.

b. *ye-pratyek chele_cphel kareche, se_csamasyā sriṣṭi karbe: Each boy who has failed will create problems.

These sentences are ungrammatical because noun phrases with quantifiers like anek, bahu and pratyek disallow pre-nominal restrictive relativization. We have to block these sentences, saying that noun phrases with such quantifiers cannot be restrictively (pre-nominally) relativized.

Noun phrases with generic quantifiers sab, sakal: 'All' allow pre-nominal restrictive relativization. Consider the examples in (74).

(74) a. ye-sab chele_cphel kareche, tārā_csamasyā sriṣṭi karbe: Those boys (Lit., who-all boys) who have failed will create problems.

b. ye-chelerā_cphel kareche, tārā_csamasyā sriṣṭi karbe: The boys who have failed will create problems.

The NP-Rel ye-sab chele in (74a) contains the quantifier sab; but this sentence does not mean that all boys have failed. Instead, it means that some boys have failed and all of them will create problems. Here the relative deictic ye restricts

the generic quantifier to a limited number. (74a) may be considered as a colloquial version of (74b), where the NP-Rel has no quantifier, but indicates plurality. (74a) cannot be derived from an underlying coordinate conjoined structure where the quantifier sab is present. For example, we cannot derive (74a) from (75).

(75) s [CONJCO s [sab chele_cphel kareche] s [sab chele_csamasyā sriṣṭi karbe]] : (All boys have failed)(All boys will create problems).

It has been suggested that this sort of sentence should be derived from a representation underlying an 'if-then' sentence in English (cf., Thompson (1971)). If we adopt this suggestion, we have to derive this type of sentence in Bengali from an underlying subordinate conjoined sentence (yadi-tabe sentence). Although some such sentences can be derived from an underlying subordinate conjoined sentence structure, many sentences will remain which cannot be derived from such a structure. The sentences in (76), which can be considered as 'true generics', can be derived from an underlying subordinate conjoined structure.

- (76) a. ye-sakal meye_ctip pare, tāder_csundar dekhāy: The girls who put on brow-marks look beautiful.
 b. ye-meyerā_ctip pare, tāder_csundar dekhāy: The girls who put on brow-marks look beautiful.
 c. yadi meyerā_ctip pare, tabe tāder_csundar dekhāy: If girls put on brow-marks, then they look beautiful.

It is possible to derive (76b, a) from (76c). But the sentences in (74) cannot be derived from an underlying subordinate conjoined structure, because this structure disallows some

tenses and aspects. For example, this hypothesis will postulate (77) as the underlying structure of (74a); but (77) is unacceptable because a subordinate conjoined structure disallows perfective aspect.

(77) *yadi chelerā_cphel kareche, tabe tārā_csamasyā sriṣṭi karbe: If boys have failed, then they will create problems.

Our suggestion is that sentences like (74a), which are not true generics, should be derived from an underlying coordinate conjoined structures like other pre-nominal restrictive relative clause sentences. Nevertheless the quantifiers occurring with the noun phrases involved are not generic in the underlying structure. For Bengali sentences of the type (74a) it can be considered that the quantifier in the underlying structure is kichu or kayek: 'Some', but not sab or sakal: 'All'. So (74a) can be derived from an underlying structure something like (78).

(78) _S [CONJCO _S [kayekjan chele_cphel kareche] _S [kayekjan chele_csamasyā sriṣṭi karbe]] : (Some boys have failed)(Some boys will create problems).

After the application of the restrictive relativization rule in (78), we will need a rule which will optionally replace the quantifier by sab or sakal, or will delete the underlying quantifiers while making the noun phrases plural. In this way we can derive the sentences in (79) from (78).

(79) a. ye-ka(yek)jan chele_cphel kareche, $\left. \begin{array}{l} \text{se-ka(yek)jan chele} \\ \text{tārā} \end{array} \right\} \sub{c}$
 samasyā sriṣṭi karbe: Those (some) boys who have failed will create problems.

b. ye- $\left\{ \begin{array}{l} \text{sab} \\ \text{sakal} \end{array} \right\}$ chele_cphel kareche, $\left\{ \begin{array}{l} \text{se-} \left\{ \begin{array}{l} \text{sab} \\ \text{sakal} \end{array} \right\} \text{chele} \\ \text{tārā} \end{array} \right\}$ c

samasyā sriṣṭi karbe: Those (all) boys who have failed will create problems.

c. ye-chelerā_cphel kareche, $\left\{ \begin{array}{l} \text{se-chelerā} \\ \text{tārā} \end{array} \right\}$ c samasyā

sriṣṭi karbe: The boys who have failed will create problems.

The sentences in (79) are cognitively synonymous, and so they can be derived from the same source, which¹⁵ [(78).

8.5 The Derivation of Nonrestrictive Relative Clause Sentences.

A nonrestrictive relative clause in Bengali behaves like a parenthetical statement on an argument. It is not an obligatory part of the containing noun phrase, but carries some independent information about the containing noun phrase. Ross (1967a) wants to derive English nonrestrictive relative clause sentences from a sequence of related sentences (cf., Thompson (1971), Stockwell et al (1973)). We have already suggested that Bengali nonrestrictive relative clause sentences should be derived from a sequence of independent sentences, each containing a noun phrase coreferential with one in the other sentence. Ross does not want to derive a nonrestrictive relative clause from an underlying coordinate conjoined sentence structure, because although a nonrestrictive relative sentence can be interrogative, the conjunction of a declarative and an interrogative sentence is not allowed in English. We want to derive a nonrestrictive relative clause from a sequence of related sentences, but not because of this fact. We have derived

restrictive relative clause sentences from an underlying coordinate conjoined structure, although the conjunction of a declarative and an interrogative sentence is not permissible. But a restrictive relative sentence can be interrogative as shown below.

(80) a. ye-mahilā_c ātmahatyā karechen, tini_c ki samudra bhālabāsten?:

Did the woman who has committed suicide love
the sea?

b. *ekjan mahilā_c ātmahatyā karechen ebaṅ tini_c ki samudra

bhālabāsten?: A woman has committed suicide and
did she love the sea?

Thompson (1971) has proposed that the connector is deleted between a declarative and an interrogative sentence, and the structure becomes a relative clause sentence.

We want to derive a nonrestrictive relative clause sentence from a sequence of related sentences for the following reasons:

- (a) it makes a difference between the underlying structure of restrictive and nonrestrictive relative clause sentences, and
(b) a nonrestrictive relative clause is superflously embedded in the matrix sentence, but can be better used as an independent sentence.

Now we will consider how a nonrestrictive relative clause sentence in Bengali can be derived. Consider the examples below.

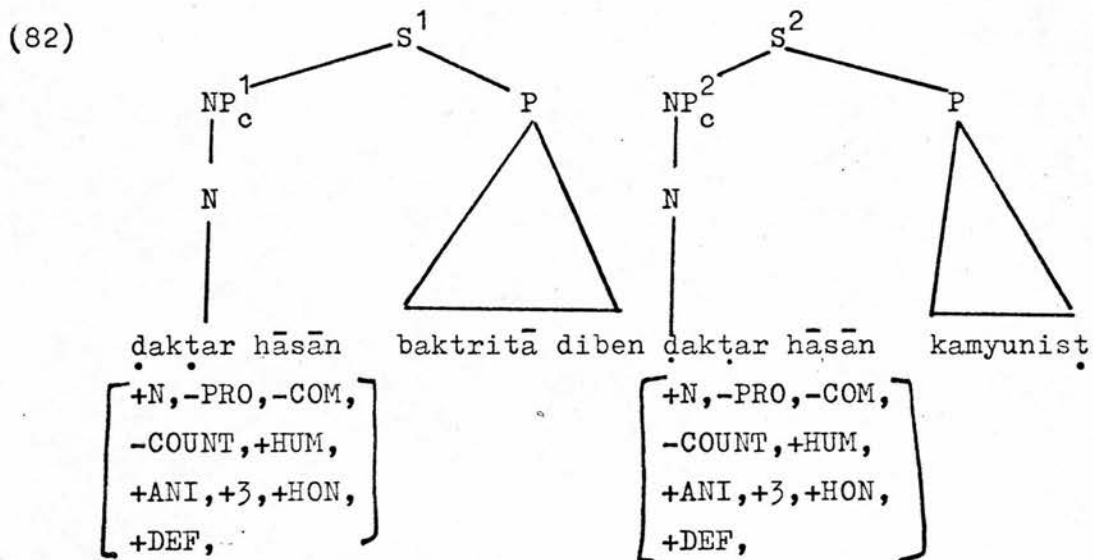
(81) a. $\overset{t}{\text{daktar}} \text{ hāsān}_{\text{c}}, \underset{2}{\text{yini}} \underset{3}{\text{kamyunist}}, \underset{4}{\text{baktritā}} \underset{5}{\text{diben}}:$

Dr Hasan, who is a communist, will give a lecture.

b. $\overset{t}{\text{daktar}} \text{ hāsān}_{\text{c}} \underset{5}{\text{baktritā}} \text{ diben. } \underset{4}{\text{daktar}} \text{ hāsān}_{\text{c}} \underset{3}{\text{kamyunist}}:$

Dr Hasan will give a lecture. Dr Hasan is a communist.

In a nonrestrictive relative clause the identity of the containing noun phrase is well-established beforehand. It is seen in Bengali that the containing noun phrase of a nonrestrictive relative clause is usually a proper name having a well-established identity. In (81a) the identity of ḍakṭar hāsān is well-established, and the relative clause gives some information about him. We will derive (81a) from (81b), which is a sequence of related sentences: each sentence here contains a noun phrase ḍakṭar hāsān. These sentences are not conjoined, but related, because they are propositions about the same person. (81b) can be sequentially ordered like (82) in the intermediate structure.



There are two independent but related sentences in (82). Of these two, which one will be the relative clause depends on the intention of the speaker. The sentence which the speaker wants to focus (the one which contains the main information) will be the matrix sentence; and the other one will be the nonrestrictive relative clause. We can formulate this by placing the focus sentence first in a sequence of

related sentences. When it is decided which of the sentences will be the focus, and which will be the nonrestrictive relative clause, the following operations will take place:

(83) Operations for Nonrestrictive Relativization.

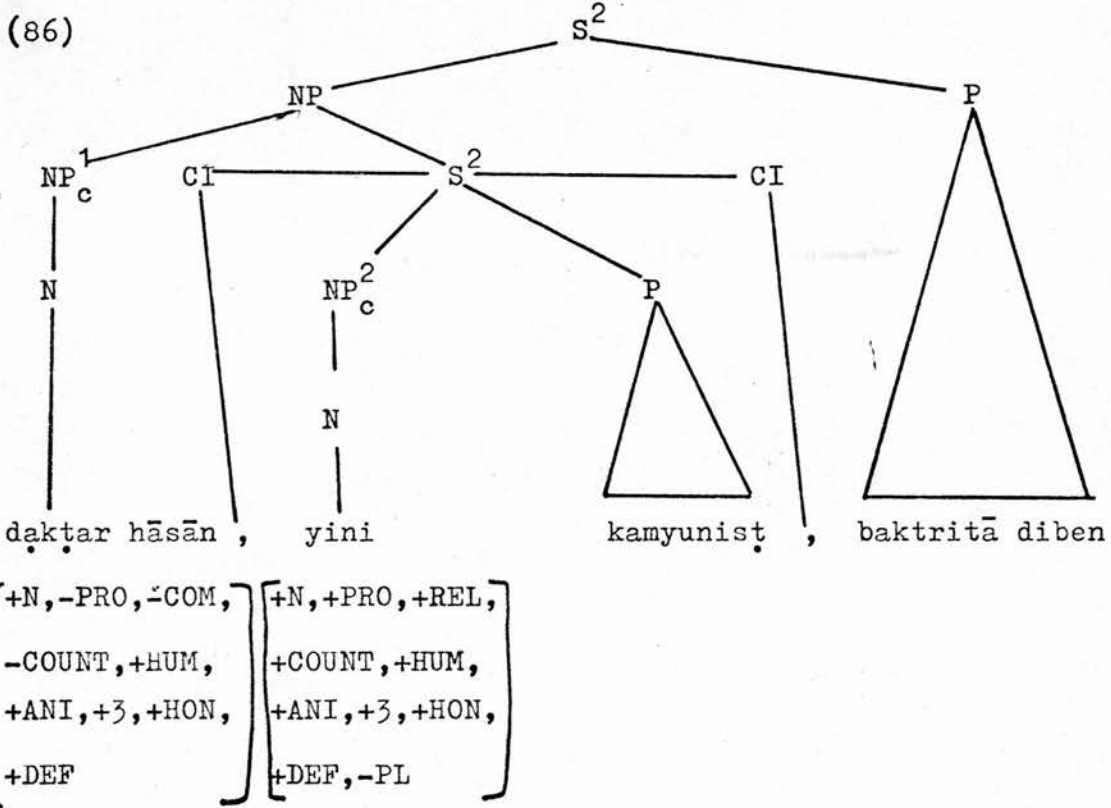
- a. Embed the relative clause as the right daughter of the containing noun phrase by Chomsky-adjunction.
- b. Insert the relative deictic ye to the NP-Rel as a DEIC.
- c. Specify the NP-Rel as [+PRO].
- d. Replace the sentence-boundary marks of the relative clause by comma-intonation (CI).

The operations (83b, c) need attention. We have not specified the NP-Rel as [+PRO,+REL] by a single operation. Instead, this has been done by two operations. This is because of the fact that although Bengali has human relative pronouns, it has no nonhuman and inanimate [+COUNT] relative pronoun. If a single operation would specify the NP-Rel as [+PRO,+REL], then we could not account for nonrestrictive relative sentences which involves nonhuman or inanimate [+COUNT] noun phrases. Consider the examples below.

- (84) a. jagadīś¹ basu_c, yini_c rāṛī khāle janmagrahaṇ kareṇ:
Jagadish Bose, who was born at Rari Khal.
- b. mātinēr pākhiṭi_c, ye-ṭi_c lāl: Matin's bird,
which is red.

In (84a) the NP-Rel has been realized as a relative pronoun, as the relativized noun phrase is human; but in (84b) the NP-Rel is not a relative pronoun, but a residual noun phrase consisting of the relative deictic ye and the classifier ṭi. In (84a) the pronominalization proper rule applies obligatorily to the

The operation (83c) will now apply to (85). It will specify the relativized noun phrase (NP²) as [+PRO]. The pronominalization rule will reduce the NP-Rel into a relative pronoun if it is [+HUM], or [-COUNT, +ABS]; otherwise this rule will simply delete the head noun of the NP-Rel. The application of (83c, d) will transform (85) into (86).



We have arrived at the post-transformational string in (86). The second lexical insertion rule has attached appropriate lexical items to the derived nodes in (86), and the comma-intonations have also been placed. (86) will generate daḡtar hāsān_c, yini_c kamyunist, baktritā diben, which is (81a).

IDENTICAL HEAD NOUN DELETION9.0 Introduction.

Bengali allows the deletion of a head noun if it is identical to the head noun of another noun phrase in the same sentence. This sort of head noun deletion, due to lexical identity, is common in Bengali. It is a kind of nonanaphoric pronominalization. Pronominalization proper utilizes one noun phrase to reduce another into a pronoun under coreference, and in certain cases this rule deletes the head noun of a noun phrase under coreference (cf., § 5.3.1). The identical head noun deletion transformation is similar to this process, except that it operates between noncoreferential noun phrases. Below we give some examples which will be discussed in this chapter.¹

- (1) a. mātin ekṭi chabi ēkeche ār minuo ekṭi chabi ēkeche:
 2 3 4 5 7 8 9 10
 Matin has drawn a picture and Minu has drawn a
 picture, 4 2 3 5 10 8
 9 7
- b. mātin ekṭi chabi ēkeche ār minuo ekṭi ēkeche: Matin
 2 3 4 5 7 8 9
 has drawn a picture and Minu has drawn one, too.
 4 2 3 5 9 8 7
- (2) a. yadi mātin ekṭi bai kine, tabe minuo duṭi bai kinbe:
 1 3 4 5 6 8 9 10
 If Matin buys a book, then Minu will buy two books.
 1 5 3 4 6 10 8 9
- b. yadi mātin ekṭi bai kine, tabe minuo duṭi kinbe:
 1 3 4 5 6 8 9
 If Matin buys a book, then Minu will buy two.
 1 5 3 4 6 9 8

¹ Noun phrases which have identical head nouns will be italicized in this chapter.

- (3) a. se nīl phulṭi tullo ār āmi lāl phulṭi tullām: He
 1 2 3 4 5 6 7 8 9 1
 plucked the blue flower and I plucked the red flower.
 4 2 3 5 6 9 7 8
- b. se nīl phulṭi tullo ār āmi lālṭi tullām: He plucked
 1 2 3 4 5 6 7 8 1 4
 the blue flower and I plucked the red one.
 2 3 5 6 8 7

(1a) has two noun phrases which are identical. (1b) is derived from (1a) by deletion of the head noun of the forward noun phrase, due to lexical identity. (2a) has two noun phrases of which only the head nouns are identical. (2b) is derived from (2a) by deletion of the head noun of the forward noun phrase. Similarly, (3b) is derived from (3a) by deletion of the head noun of the forward noun phrase. This sort of head noun deletion is possible when two noun phrases in a structure have identical head nouns.

9.1 Direction of Deletion.

Identical head noun deletion is permissible even in a simple sentence structure. This can be seen in the following examples.²

- (4) a. ekjan bhadrālok anya ekjan bhadrālokke ballen...: A gentleman told another gentleman..
- b. ekjan bhadrālok anya ekjanke ballen: A gentleman told another one.
- c. *ekjan anya ekjan bhadrālokke ballen:
- (5) a. pratham cheleṭi dviṭīya cheleṭike cine nā: The first boy does not know the second boy.
- b. pratham cheleṭi dviṭīyaṭike cine nā: The first boy does not know the second one.
- c. *prathamṭi dviṭīya cheleṭike cine nā:

2 We will consider simple sentence structures with SOV order.

(4a) has two noun phrases with identical head noun. (4b) is derived from (4a) by deletion of the head noun of the forward noun phrase. When the head noun of the forward noun phrase is deleted, the case marker ke is suffixed to the element which immediately precedes the deleted head noun. Here the deletion operates forwards. That is, the head noun of the left noun phrase is used to delete the head noun of the right noun phrase. (4c) is derived from (4a) by deletion of the head noun backwards. (4c), which is ungrammatical, shows that backwards deletion is not possible in a simple structure having an SOV order. The examples in (5) show that identical head noun deletion is possible only forwards. Accordingly we can arrive at the following conclusion:

- (6) In a simplex structure with the SOV order identical head noun deletion is allowed only forwards.

We will now consider the direction of identical head noun deletion in complex and compound structures. Consider the examples in (7, 8, 9).

- (7) a. ekjan bhadralok elen o dujan bhadralok cale gelen:
A gentleman came and two gentleman^e went away.
- b. ekjan bhadralok elen o dujan cale gelen: A gentleman came and two went away.
- c. ekjan elen o dujan bhadralok cale gelen: One came and two gentlemen went away.
- (8) a. yadi mātin ekṭi bai kine, tabe minu duṭi bai kinbe:
If Matin buys a book, then Minu will buy two books.
- b. yadi mātin ekṭi bai kine, tabe minu duṭi kinbe:
If Matin buys a book, then Minu will buy two.

- c. yadi mātin ekṭi kine, tabe minu duṭi bai kinbe: If
Matin buys one, then Minu will buy two books.
- (9) a. mātin bādaler baiṭi dekhe ballo ye seo ekṭi bai kinbe:
Matin saw Badal's book and said that he would buy a
book, too.
- b. mātin bādaler baiṭi dekhe ballo ye seo ekṭi kinbe:
Matin saw Badal's book and said that he would buy
one, too.
- c. mātin bādalerṭi dekhe ballo ye seo ekṭi bai kinbe:
Matin saw Badal's one and said that he would buy
a book, too.

(7a) is a coordinate conjoined structure where there are two noun phrases identical in head noun. (7b) is derived from (7a) by deleting the head noun of the forward noun phrase. and (7c) is derived from (7a) by deleting the head noun of the backward noun phrase. Although (7c) is not a very good sentence, it is acceptable. So we see that identical head noun deletion is allowed both forwards and backwards in a coordinate conjoined structure. (8a) is a subordinate conjoined structure with two noun phrases which are identical in head noun. We derive (8b) by forward deletion of identical head noun, and (8c) by backward deletion of head noun from (8a). We see that deletion can take place both forwards and backwards in a subordinate conjoined structure. (9a) is a complex structure which has two noun phrases with identical head noun. In such a structure also, both forwards and backwards deletion of identical head noun is possible. So we can give the following condition for identical head noun deletion:

(10) Direction of IHN Deletion.

The identical head noun deletion rule operates forwards in a simple structure with SOV order, and both forwards and backwards in conjoined and complex structures.

9.2 Identical Head Noun Deletion.

The identical head noun deletion rule operates between noncoreferential noun phrases which are identical in head noun. This is a kind of pronominalization in that it utilizes one noun to delete another under the identity condition, and the residual noun phrase functions pronominally. The noun phrases involved in such a deletion are noncoreferential, but the rule operates under lexical identity and sense identity.

Identical head noun deletion can take place even if the nouns are the only elements of the noun phrases involved. Consider the examples in (11).

- (11) a. bādal bai kinlo ār āmio bai kinlām: Badal bought
(a) book(s) and I bought (a) book(s), too.
- b. bādal bai kinlo ār āmio kinlam: Badal bought
(a) book(s) and I bought, too.

In (11a) each noun phrase involved has the noun bai as its only element. These noun phrases are noncoreferential. Here the deletion operates and generates (11b). But this sort of deletion is not acceptable if one noun phrase has other constituents besides the head noun, and the other one has only the head noun. Consider the examples in (12).

- (12) a. bādal ekṭi bai kinlo ār āmio bai kinlām: Badal
bought a book and I bought (a) book(s), too.
- b. *bādal ekṭi bai kinlo ār āmio kinlām: Badal bought
a book and I bought, too.

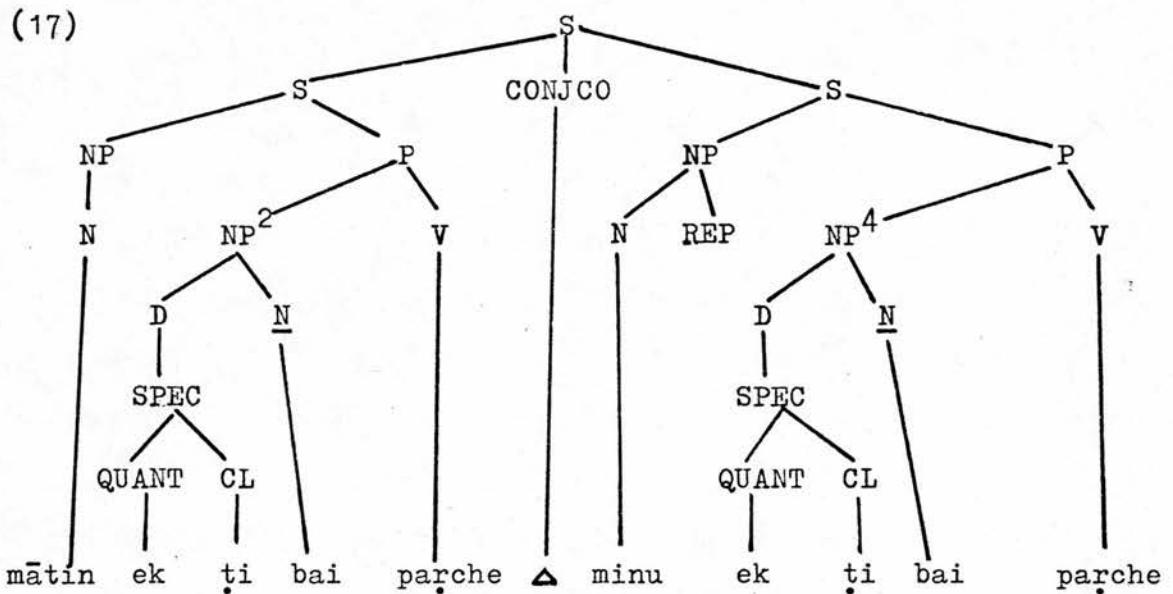
(12a) has two noun phrases with identical head nouns. Here the left noun phrase has two constituents SPEC and N, but the right noun phrase has only the head N. So deletion generates the unacceptable sentence (12b). In (12a) the noun phrases involved are noncoreferential, but in (12b) they are understood as coreferential.

But noun phrases usually contain other constituents besides head nouns. Consider the examples in (13)-(16).

- (13) a. mātin ekṭi bai paṛche āṛ minuo ekṭi bai paṛche:
 Matin is reading a book and Minu is reading a book, too.
- b. mātin ekṭi bai paṛche āṛ minuo ekṭi paṛche: Matin is reading a book and Minu is reading one, too.
- (14) a. minuo e-phulṭi cāy āṛ mātin o-phulṭi cāy: Minu wants this flower and Matin wants that flower.
- b. minuo e-phulṭi cāy āṛ mātin o-ṭi cāy: Minu wants this flower and Matin wants that one.
- (15) a. minuo lāl phulṭi cāy āṛ mātin nīl phulṭi cāy: Minu wants the red flower and Matin wants the blue flower.
- b. minuo lāl phulṭi cāy āṛ mātin nīlṭi cāy: Minu wants the red flower and Matin wants the blue one.
- (16) a. minuo nīl phulgulo cāy āṛ mātin lāl phulgulo cāy:
 Minu wants the blue flowers and Matin wants the red flowers.
- b. minuo nīl phulgulo cāy āṛ mātin lālgulo cāy: Minu wants the blue flowers and Matin wants the red ones.

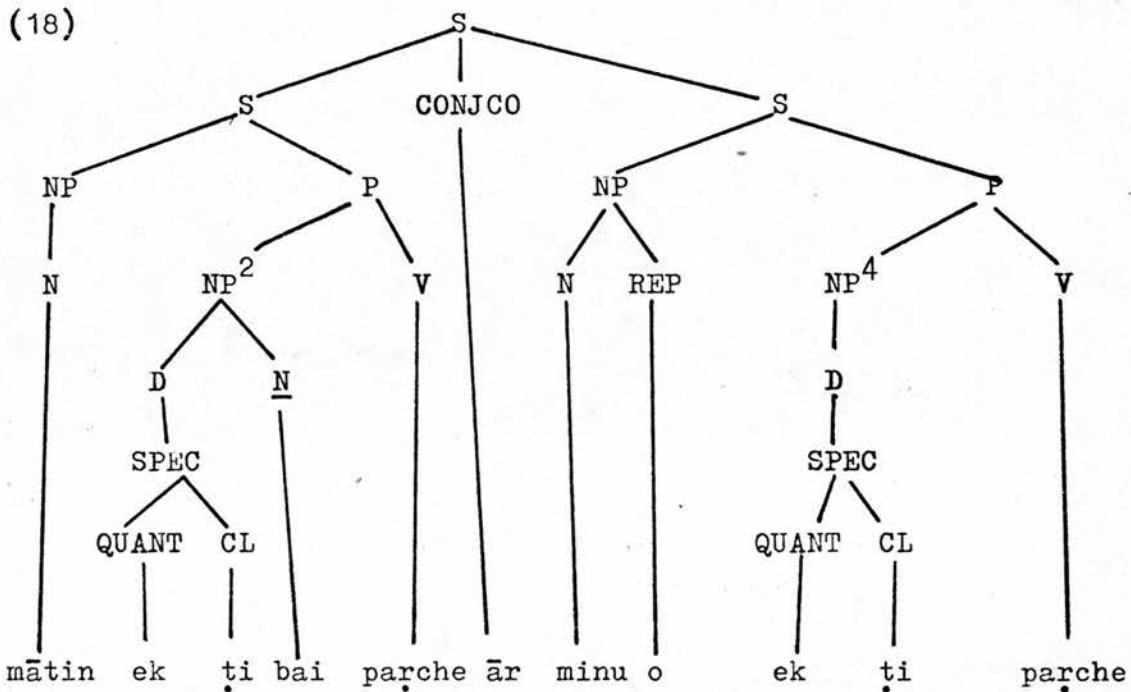
(13a) has two noun phrases which are identical both in SPEC and N. We derive (13b) from (13a) by deleting the head noun of the forward noun phrase, due to lexical identity. However, the SPEC of the deleted noun is left undeleted. (14a) has two noun

phrases which are identical in head noun and classifier. We derive (14b) from (14a) by deleting only the head noun of the forward noun phrase. Similarly, we derive (15b) from (15a) by deleting the head noun of the forward noun phrase. (16a) has two noun phrases each with the structure ADJ N PLMARKER. They are identical both in head noun and plural marker. (16b) is derived from (16a) by deletion of the head noun of the forward noun phrase; other elements, including the plural marker, are left behind. So we see that the identical head noun deletion rule deletes the head noun only, and suffixes of the head noun are suffixed to the element which precedes the deleted head noun. Let us show, for example, how the head noun deletion rule operates. (13a) has the intermediate structure (17).



In (17) the head nouns of NP² and NP⁴ are identical. As (17) is a coordinate conjoined structure, the identical head noun deletion rule can apply both forwards and backwards here. We can use the head noun of NP² to delete the head noun of NP⁴, and we can also apply the rule in the reverse order. If we

apply the deletion rule forwards in (17), we will derive (18), after the last rule of the transformational component.



In (18) the head noun has been deleted from NP⁴ due to its lexical identity with the head noun of NP². (18) will generate mātin ekṭi bai paṛche ār minuo ekṭi paṛche, which is (13b).

9.3 Identical Modifier and Head noun Deletion.

A head noun can be deleted with all its attributive adjectives if the head noun and its adjectives are identical with the head noun and its attributive adjectives of another noun phrase. Consider the examples below.

(19) a. minuo ekṭi lāl kalam kineche ār mātinuo ekṭi lāl kalam

kineche: Minu has bought a red pen and Matin has bought a red pen, too.

b. minuo ekṭi lāl kalam kineche ār mātinuo ekṭi kineche:

Minu has bought a red pen and Matin has bought one, too.

- (20) a. minu ekṭi dāmī lāl kalam kineche ār mātino ekṭi dāmī lāl kalam kineche: Minu has bought an expensive red pen and Matin has bought an expensive red pen, too.
- b. minu ekṭi dāmī lāl kalam kineche ār mātino ekṭi kineche: Minu has bought an expensive red pen and Matin has bought one, too.

In (19a) there are two noun phrases which are identical in every respect. They have the structure SPEC ADJ N. We can derive (19b) from (19a) by deleting the adjective and the head noun of the forward noun phrase under their identity with the adjective and the head noun of the backward noun phrase. The SPEC of the forward noun phrase, although it is identical with the SPEC of the backward noun phrase, cannot be deleted. The sentence (19b) is ambiguous. The ambiguity is due to the fact the adjective lal could have been present in or absent from the forward noun phrase in the underlying structure. (20b) is derived from (20a) by deletion of the identical adjectives and the head noun of the forward noun phrase due to their identity with those of the backward noun phrase. (20b), like (19b), is ambiguous. This ambiguity is also due to the deletion of the identical adjectives.

Bengali avoids this sort of ambiguity in two ways:

(a) only the identical head noun is deleted and the identical adjective is stressed, and (b) only the identical head noun is deleted, and the SPEC and ADJ are permuted. In these ways we will derive the sentences in (21) from (19a), and those in (22) from (20a).

- (21) a. minu ekṭi lāl kalam kineche ār mātino ekṭi lāl
 kineche: Minu has bought a red pen and Matin has
 bought a red (one), too.
- b. minu ekṭi lāl kalam kineche ār mātino lāl ekṭi
 kineche: Minu has bought a red pen and Matin has
 bought a red (one), too.
- (22) a. minu ekṭi dāmī lāl kalam kineche ār mātino ekṭi dāmī
lāl kineche: Minu has bought an expensive red pen
 and Matin has bought an expensive red (one), too.
- b. minu ekṭi dāmī lāl kalam kineche ār mātino dāmī lāl
ekṭi kineche: Minu has bought an expensive red pen
 and Matin has bought an expensive red (one), too.

(21a) is derived from (19a) by deletion of the head noun of the forward noun phrase under lexical identity. The adjective of the forward noun phrase has not been deleted, although it is identical with the adjective of the backward noun phrase. The adjective lāl in the forward noun phrase in (21a) is mildly stressed so that it is understood as an adjective, but not as a noun. (21b) is derived from (19a) by deletion of the head noun of the forward noun phrase, and by permutating its SPEC and ADJ. There is no ambiguity in (21a, b). The sentences in (22) are similarly derived from (20a). Sentences of the type (21, 22b) are more common in Bengali than those of the type (21, 22a).

9.4 Deletion of Identical Head Noun with Nonidentical Modifiers.

Consider the examples in (23).

- (23) a. minu ekṭi lāl phul tuleche ār mātin duṭi nīl phul
 tuleche: Minu has plucked a red flower and Matin
 has plucked two blue flowers.
- b. minu ekṭi lāl phul tuleche ār mātin duṭi nīl tuleche:
 Minu has plucked a red flower and Matin has plucked
 two blue (ones).
- c. minu ekṭi lāl phul tuleche ār mātin nīl duṭi tuleche:
 Minu has plucked a red flower and Matin has
 plucked two blue (ones).

(23a) has two noun phrases which are identical in head noun only. Each of these noun phrases has a SPEC and an adjective besides the head noun. The adjectives of these noun phrases are nonidentical. (23b) is derived from (23a) by deletion of the head noun of the forward noun phrase. The adjective in the forward noun phrase in (23b) is mildly stressed so that it is not understood as a noun. (23c) is derived from (23a) by deletion of the head noun, and permutation of the SPEC and ADJ of the forward noun phrase. It should be mentioned here that the residual noun phrase nīl duṭi in (23c) is ambiguous between a definite and an indefinite reading. These examples show that operations in structures with nonidentical modifiers are similar to those in structures with identical modifiers.

SENTENCE PRONOMINALIZATION10.0 Introduction.

It is generally assumed that a sentence can be pronominalized under certain conditions. In English 'It' and 'So' are usually used as sentence pronouns as can be seen in the examples in (1), taken from Cushing (1972).

(1) a. Noam said that deep structure exists, and I believe

{ it }
{ *so }.

b. George asked me whether deep structure exists;

I said that I believed { so }
{ *it } .

Bengali also shows that a sentence can be pronominalized into the abstract pronoun tā under certain conditions. We have sentences like those in (2).

(2) a. mātin mane kare ye prithibīṭā ekṭā mārbel ār
 2_ 3 4 5 6 7

minuo tā mane kare: Matin thinks that the earth is
9 10 11_ 2 3 4_

a marble and Minu thinks so, too.
5 6 7 11 10 9

b. kerāmat āli ghoṣanā kareche ye se jātir pitā kintu
 2_ 3 4 5 6 7

bānālīrā tā mane kare nā: Keramat Ali has declared that
8 9 10_ 11 2_ 3

he is the father of the nation, but the Bengalees
4 6 5_ 7 8

do not think so.
11 10 9

In (2a) prithibīṭā ekṭā mārbel and tā, and in (2b) se jātir pitā and tā are understood coreferentially; and so it can be assumed that the pronoun tā is derived by the 'sentence pronominalization'

transformation. In this chapter we will investigate this process in Bengali, and try to discover whether there is any process at all which can be called 'sentence pronominalization'.

10.1 The Abstract Pronoun tā.

Bengali has a pronoun tā with the features [+ABS, -COUNT], to which we refer as the abstract pronoun. It must not be confused with the morphophonologically derived forms or parts of forms of the pronouns such as tār: 'His', tāke: 'Him', tārā: 'They', tāder: 'Them, Their' etc.,. These forms are derived by morphophonological rules, which apply on a surface structure (cf., § 11.3). But the abstract pronoun tā is both a deep structure and surface structure pronoun. It can be used deictically, anaphorically and under identity of sense (cf., § 5.10). It is generated in the underlying structure when used deictically as in (3).

- (3) a. āmi tā bhuli ni: I have not forgotten that.
 b. āmi tā mane rākhbo: I shall remember that.

It is derived transformationally in (4).

- (4) a. yā_cpeyechi, tā_cabismaranīya: What (I) have got (that) is unforgettable.
 b. āmi yā_ccāi, tā_csiṅhāsan nay: What I want (that) is not a throne.

The pronoun tā is usually used to refer to abstract objects or entities, but when it is used to refer to concrete objects its referents are understood as 'things'. So it can be easily used to refer to [-COUNT] objects, but when it is used to refer to [+COUNT] concrete objects, the 'thingness' of the referents are taken into account. Consider the examples in (5).

- (5) a. āmi bādalkē ekṭi kārḍ_c pāṭhiyechilām, kintu se $\left\{ \begin{array}{l} \text{tā} \\ \text{se-ṭi} \end{array} \right\}_c$

pāy ni: I had sent a card to Badal, but he did not receive it.

- b. yadi tumi āj-rāte ekjan bāndhabī_c cāo, tabe tumi

$\left\{ \begin{array}{l} \text{tā} \\ *tāke \end{array} \right\}_c$ pābe: If you want a girl-friend to-night,

then you will get that.

In (5a) tā refers to a concrete object and in (5b) it refers to an indefinite and unspecific human being. But there is a semantic difference between the uses of tā and se-ṭi in order to refer to ekṭi kārḍ in (5a). when tā is used, it refers to its antecedent simply as a 'thing'; but when se-ṭi is used, the antecedent is understood as it is. In (5b) the pronoun tāke cannot be used to refer to an indefinite and unspecific antecedent; but tā can be used. The antecedent in (5b) is a human noun, but tā refers to her simply as a 'thing', as if she is an unexpected gift.

Bengali utilizes the same pronoun tā as a sentence pronoun as sentences are abstract entities. This pronoun has no plural form, but it can be used to refer to both singular and plural noun phrases. Consider the examples below.

- (6) a. āpanār siddhāntasamuha_c camatkār ār tā_c mānabik: Your decisions are excellent and they are humane.

- b. āpanār siddhāntasamuha camatkār ār $\left\{ \begin{array}{l} *tāgulo \\ *tasamuha \\ *tārā \end{array} \right\}$

mānabik:

In (6a) tā refers to a plural noun phrase. (6b) is ungrammatical because of the plural forms of tā, which are ungrammatical.

10.2 Constraints on Sentence Pronominalization.

We will first assume that there is a pronominalization rule which reduces a sentence into a pronoun under certain conditions. If such a rule exists, then we could derive the pronoun tā in (2) by that rule. We will come later to the question whether sentences like those in (2) are derived by the sentence pronominalization rule, or by any other process. Here we will take the existence of such a rule for granted and try to find out the directional constraint on 'sentence pronominalization'. Consider the sentences below.

- (7) a. tumi jāno ye rājnītibiderā asat ār āmio jāni ye rājnītibiderā asat: You know that politicians are dishonest and I, too, know that politicians are dishonest.
- b. tumi jāno ye rājnītibiderā asat ār āmio tā jāni: You know that politicians are dishonest and I know it, too.
- c. *tumi tā jāno ār āmio jāni ye rājnītibiderā asat: You know it and I, too, know that politicians are dishonest.
- (8) a. yadi bādāl jānto ye nāju āsbe, tabe se āmāke balto ye nāju āsbe: If Badal knew that Nazu would come, then he would tell me that Nazu would come.
- b. yadi bādāl jānto ye nāju āsbe, tabe se āmāke tā balto: If Badal knew that Nazu would come, then he would tell me so.
- c. yadi bādāl tā jānto, tabe se āmāke balto ye nāju āsbe: If Badal knew it, then he would tell me that Nazu would come.

(7a) is a coordinate conjoined structure. It has two complement clauses which are coreferential, that is, synonymous and

structurally identical. If we use the left complement sentence to pronominalize the right complement sentence under coreferentiality, we will derive (7b). But if the rule is applied backwards, we will derive the ungrammatical sentence (7c). Thus we see that the sentence pronominalization rule applies forwards in a coordinate conjoined structure. (8a) is a subordinate conjoined structure. It has two complement clauses which are coreferential. If we apply the sentence pronominalization rule forwards in (8a), we derive (8b); and if the rule is applied backwards, we derive (8c). As both of (8b, c) are grammatical, we see that the sentence pronominalization rule can apply both forwards and backwards in a subordinate conjoined structure.

The sentence pronominalization rule involves complement sentences, not the matrix sentences. We have seen above that this rule applies to complement sentences. Now consider the sentences in (9).

- (9) a. bādal jāne ye nāju āsbe ār bādal jāne ye mātin āsbe:
Badal knows that Nazu will come and Badal knows that
Matin will come.
- b. *bādal jāne ye nāju āsbe ār tā ye mātin āsbe: Badal
knows that Nazu will come and it that Matin will come.

The matrix sentences in (9a) are coreferential, but (9b) shows that the rule cannot apply between them.

The sentence pronominalization rule can apply in a structure if it has two coreferential (synonymous and structurally identical) complement sentences. Only when these conditions are satisfied can the sentence pronominalization rule apply.

The constraints for the rule can be given as (10).

(10) Constraints on S-Pronominalization.

If a structure contains two complement sentences which are (a) synonymous (coreferential) and (b) structurally identical, the sentence pronominalization rule can apply, optionally but preferably, only forwards if the structure is a coordinate conjoined one, and both forwards and backwards if the structure is a subordinate conjoined one.

10.3 An Outline of Complementation in Bengali.

Different types of pronominalization involve complement sentences. So we need to consider the process of complementation in Bengali to have a picture of the underlying structure of complement clauses. But as it is beyond our scope to tackle this process exhaustively, we will deal here with it in outline, as far as we deem necessary for our present purpose of sentence pronominalization and other types of pronominalization.

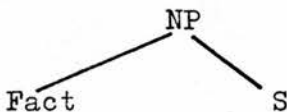
Complementation is an instance of embedding of one sentence into another. The embedded sentence may get transformed during the process of derivation so much that it may not look like a sentence in the surface. Traditional Bengali grammarians are silent about complementation (cf., Chatterji (1939)), and we have not come across anything which deals with complementation in Bengali. Before going into any details, let us consider some Bengali sentences which involve complementation. Consider the sentences in (11, 12, 13).

- (11) a. bādāl mane kare ye āgāmīkāl bṛiṣṭi habe: Badal thinks
 2 3 4 5 6 2
 that it will rain tomorrow.
 3 6 5 4

(13) b'. e-ghaṭanā tātparyapūrṇa: This event is significant. In (11b') e-kathā is the object, in (12b') e-janya is the object, and in (13b') e-ghaṭanā is the subject. But these abstract noun phrases are semantically almost empty and so need semantic content. This semantic content is supplied by the complement sentences. The function of the complement sentences in these examples is to supply semantic content to the abstract noun phrases of which they are extensions. We will call this sort of complementation 'NP-Complementation'. In this sort of complementation the complement sentence is an extension of an abstract noun phrase.

Rosenbaum's (1967) analysis of English complementation has been questioned by different linguists (cf., Stockwell et al (1973, 514-36)). Kiparsky and Kiparsky (1971) have shown that complementation in English can be better explained if the semantics of the matrix predicate is taken into account. They have subcategorized English predicates into factive and non-factive paradigms, and shown that [±Fact] feature of the predicates give rise to interesting syntactic consequences. So they have posited different underlying structures for factive and non-factive complement sentences. The claim that the factive complements have the underlying structure (14a), and the non-factive complements have the underlying structure (14b).

(14) a.



b.



Bengali has both factive and non-factive predicates as well predicates which are neutral to [±Fact]. For example,

tātparyapūrṇa: 'Significant' and duḥṣita: 'Sorry' are factive, and sambhabpar: 'Likely' and satya: 'True' are non-factive.

Kiparsky and Kiparsky (1971) claim that factive predicates can occur only when the speaker presupposes that the propositional object or subject of the predicate is factually true; and non-factive predicates can occur when the speaker merely asserts or believes the proposition to be true, but does not presuppose its factuality. They have given some criteria for the diagnosis of factive and non-factive predicates, of which the negation test is the most important. The presupposition remains constant in both the negative and positive forms of the sentence with a factive predicate, but the assertion of the proposition changes with a non-factive predicate when the matrix sentence is negated. Consider the examples in (15, 16).

- (15) a. āmi dukṣita ye bṛiṣṭi hacche: I am sorry that it is raining.
 b. āmi dukṣita nai ye bṛiṣṭi hacche: I am not sorry that it is raining.
- (16) a. āmi mane kari ye rehānā rūpasī: I think that Rehana is beautiful.
 b. āmi mane kari nā ye rehānā rūpasī: I do not think that Rehana is beautiful.

According to the above criterion duḥṣita is a factive and mane kar is a non-factive predicate.

Kiparsky and Kiparsky (1971) posits different underlying structures for these two types of complement, but Bengali does not show any immediate syntactic difference between these two types of complement (cf., § 10.4 for some remote consequences). Their underlying structure for the factive complements seems

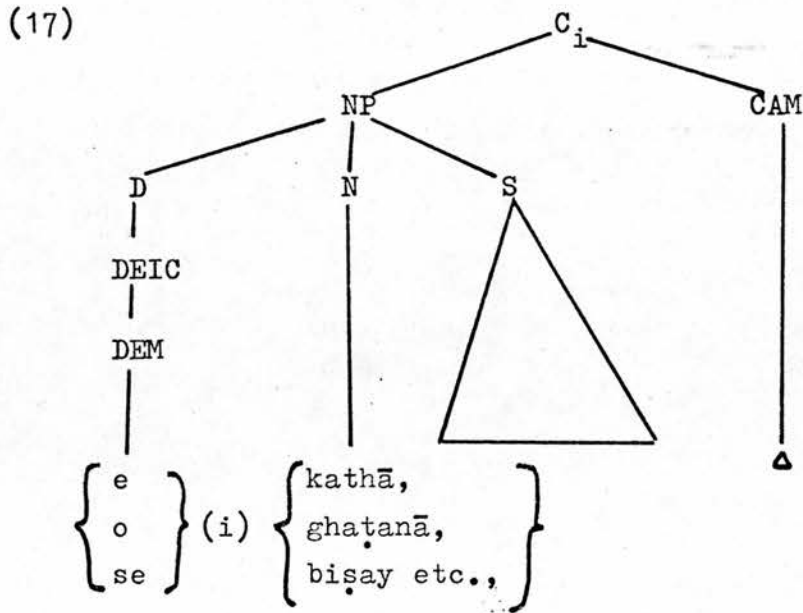
necessary for both factive and non-factive NP-Complementation in Bengali. That is, both factive and non-factive complement sentences are extensions of some abstract noun phrases. The place occupied by 'the fact' in their treatment of factive complements would be filled in by abstract noun phrases like e-kathā, e-ghaṭanā, e-biṣay, e-janya etc.,, or simply by the deictic e(i): 'This', where an appropriate head noun is missing, for both factive and non-factive complements in Bengali. We assume that these abstract noun phrases are the heads of the NP-Complement sentences. These abstract noun phrases are dominated by a case node in the underlying structure. The case relations of these abstract noun phrases are determined by their relationships with the verbs, which cooccur with them. They are usually dominated by the NEUT case, but can be dominated by other cases like INS and LOC etc.,.¹

1 Most of these abstract noun phrases are dominated by the NEUT case in the underlying structure, and they are realized as objects in the surface structure. But consider the sentences below.

- (A) a. āmi biśvās kari e-kathā ye se sat: I believe this proposition that he is honest.
 b. āmi biśvās kari e-kathāy ye se sat: I believe in this proposition that he is honest.
 c. āmi biśvās kari ye se sat: I believe that he is honest.
- (B) a. āmi dukṣita e-te ye se āse ni: I am sorry for this that he did not come.
 b. āmi dukṣita e-kāraṇe ye se āse ni: I am sorry for this reason that he did not come.

continued

The underlying structure for NP-Complementation is (17).



In (17) C_i indicates the case relationship of the noun phrase. The node D is usually filled in by the demonstrative deictic $e(i)$: 'This', but this position can be filled in by $o(i)$: 'That' or $se(i)$: 'That'. The selection of one or another of these

-
- c. āmi dukṣita ye se āse ni: I am sorry that he did not come.

There is subtle semantic distinction between (Aa) and (Ab). The abstract noun phrase e-kathā in (Aa) seems to be a NEUT, and e-kathāy in (Ab) seems to be a LOC, where the case marker e ($\Rightarrow y$) is present. (Ac) is the surface realization of (Aa, Ab) by deletion of the abstract noun phrases. In (Ba) we could not supply any appropriate noun to the noun phrase e-te, but, we assume, that nouns like karāṇ: 'Reason', kāj: 'Action' etc., could be used in this noun phrase. The noun phrases e-te and e-karāṇe seem to be abstract instrumentals in the underlying structure. (Bc) is the surface realization of (Ba, Bb).

deictics depends on the relative proximity of the proposition to the speaker. Consider the examples below.

- (18) a. āmi jāntām $\left\{ \begin{array}{c} e \\ ?o \\ ?se \end{array} \right\}$ -kathā ye tumi āsbe: I knew
 $\left\{ \begin{array}{c} \text{this} \\ \text{that} \end{array} \right\}$ proposition that you would come.
- b. āmi bali ni $\left\{ \begin{array}{c} e \\ o \\ se \end{array} \right\}$ -kathā ye tārā mithyābādī: I did
 not say $\left\{ \begin{array}{c} \text{this} \\ \text{that} \end{array} \right\}$ proposition that they were liars.

In (18a) the complement clause is understood as a direct discourse of the matrix subject, and so the abstract noun phrase contains the proximate deictic e. (18a) sound odd if we use the non-proximate deictics o or se, because the complement clause is a direct discourse of the matrix subject. In (18b) the matrix subject is denying a proposition which has been alleged to be his direct discourse. In such a case any of the deictics can be used.

The abstract noun phrases of the NP-Complements are optionally deletable under certain conditions. These noun phrases are optionally deletable if they are not the subjects of their sentences. Consider the examples in (19, 20).

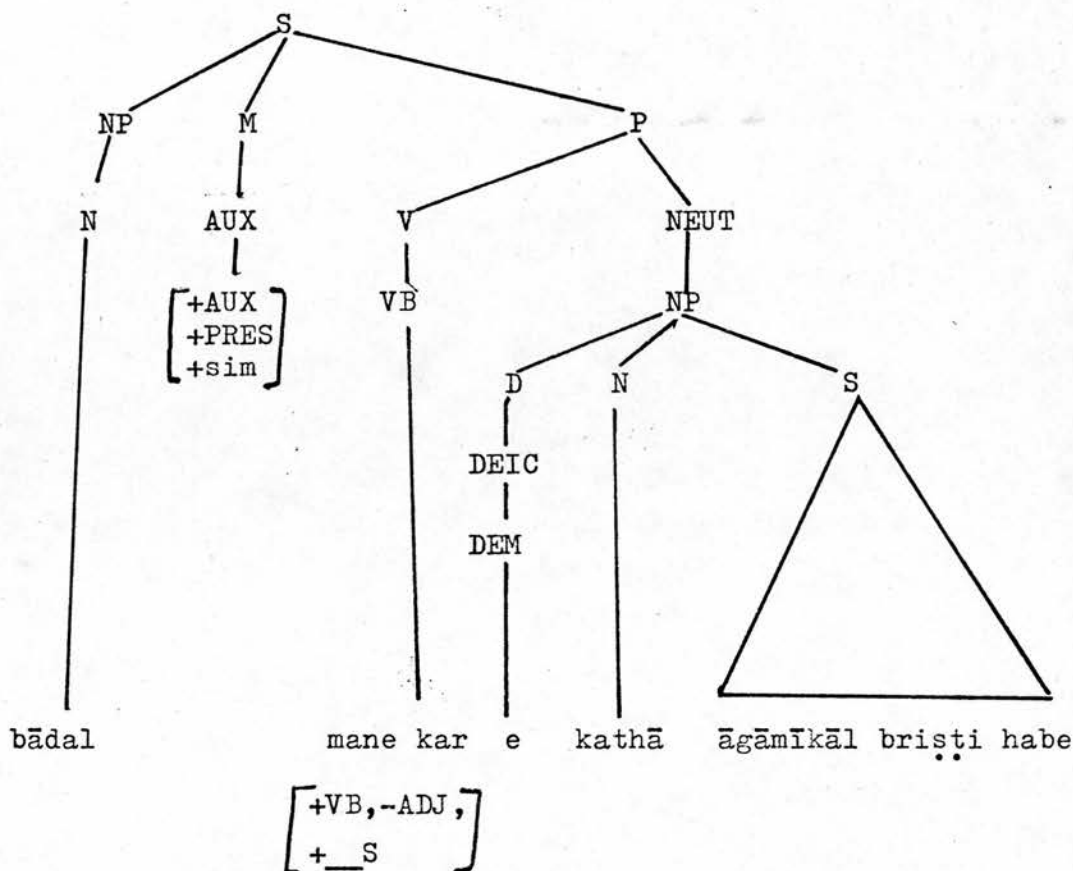
- (19) a. bādal likheche e-kathā ye se bārī yābe: Badal has
 written this proposition that he would go home.
- b. bādal likheche ye se bārī yābe: Badal has
 written that he would go home.
- (20) a. e-kathā ye se mantrī habe bismaykar: This
 proposition that he will be a minister is surprising.

- b. e-kathā bismaykar ye se mantrī habe: This proposition is surprising that he will be a minister.
- c. *ye se mantrī habe bismaykar: That he will be a minister is surprising.
- d. se ye mantrī habe, $\left\{ \begin{array}{l} \text{e-kathā} \\ \text{tā} \end{array} \right\}$ bismaykar: Lit.,
 That he will be a minister, $\left\{ \begin{array}{l} \text{this proposition} \\ \text{that} \end{array} \right\}$
 is surprising.

In (19a) e-kathā is the object and is deletable. We derive (19b) from (19a) by its deletion. In (20a) e-kathā is the subject, and so it cannot be deleted. If we delete it, we will derive the ungrammatical sentence (20c). If the complement sentence is an extension of the subject noun phrase, it is generally extraposed to the sentence-final position in Bengali. We derive (20b) from (20a) by extraposing the complement sentence to the sentence-final position. (20d) is derived by the extraposition of the complement clause to the sentence-initial position (cf., § 10.4 for a brief description of this process). Bengali generally does not allow sentential subjects. This can be seen in (20c), where the complement clause is the subject. This sentence is ungrammatical because the abstract noun phrase e-kathā has been deleted from it.

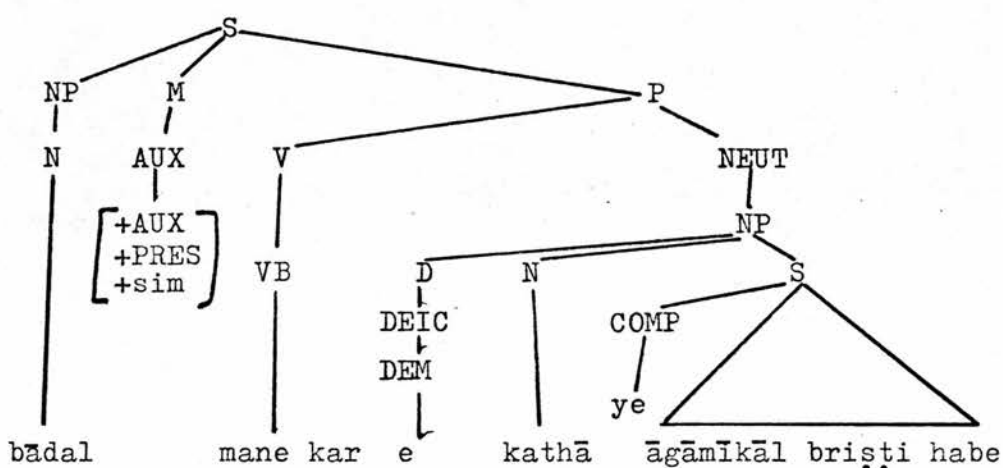
Now let us come back to the examples in (11)-(13), and show how they are derived. In these examples the (a) sentence in each set is derived from the underlying structure of the (b) sentence. For example, (11b) has the intermediate underlying structure (21).

(21)



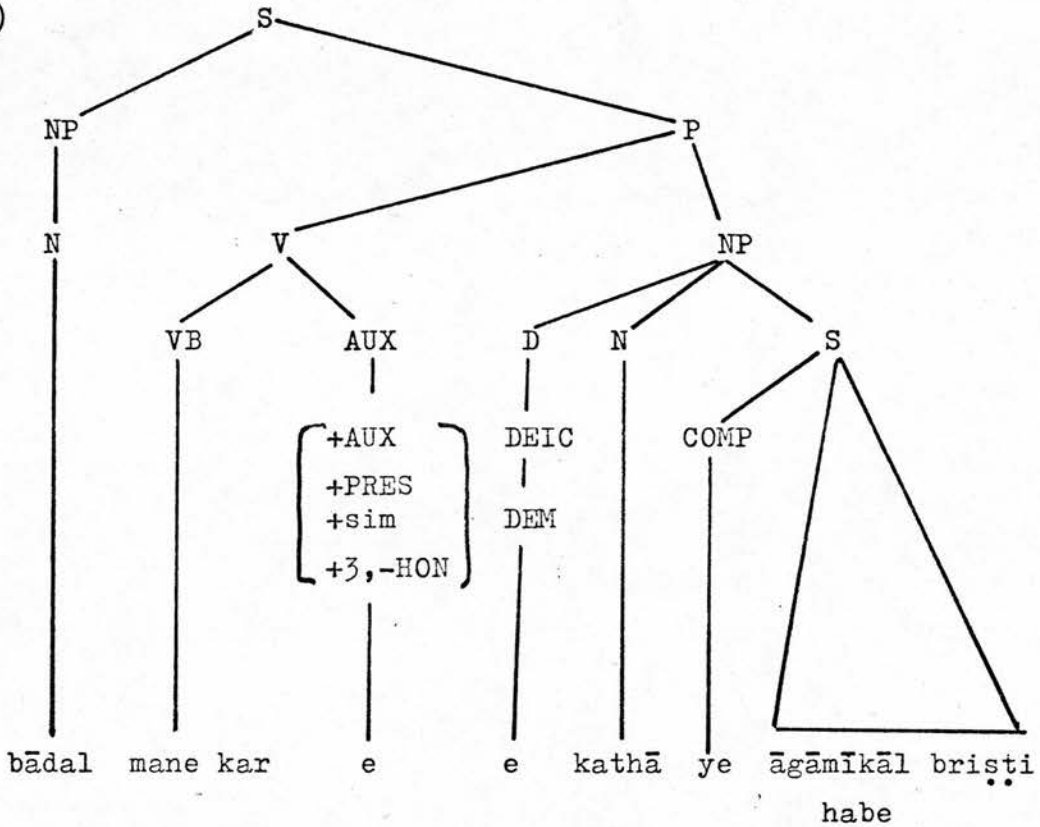
The complementizer in Bengali is ye, which is homonymous with the relative deictic and the nonhonorific relative pronoun. The complementizer is placed in the case of NP-Complementation only; and it is placed as the left daughter of the complement sentence. The complementizer placement rule will transform (21) into (22).

(22)



After the application of the last rule of the transformational component in (22), the second lexical insertion rule will attach appropriate lexical items to the nodes which require lexicalization. Thus we will derive the surface structure (23).

(23)

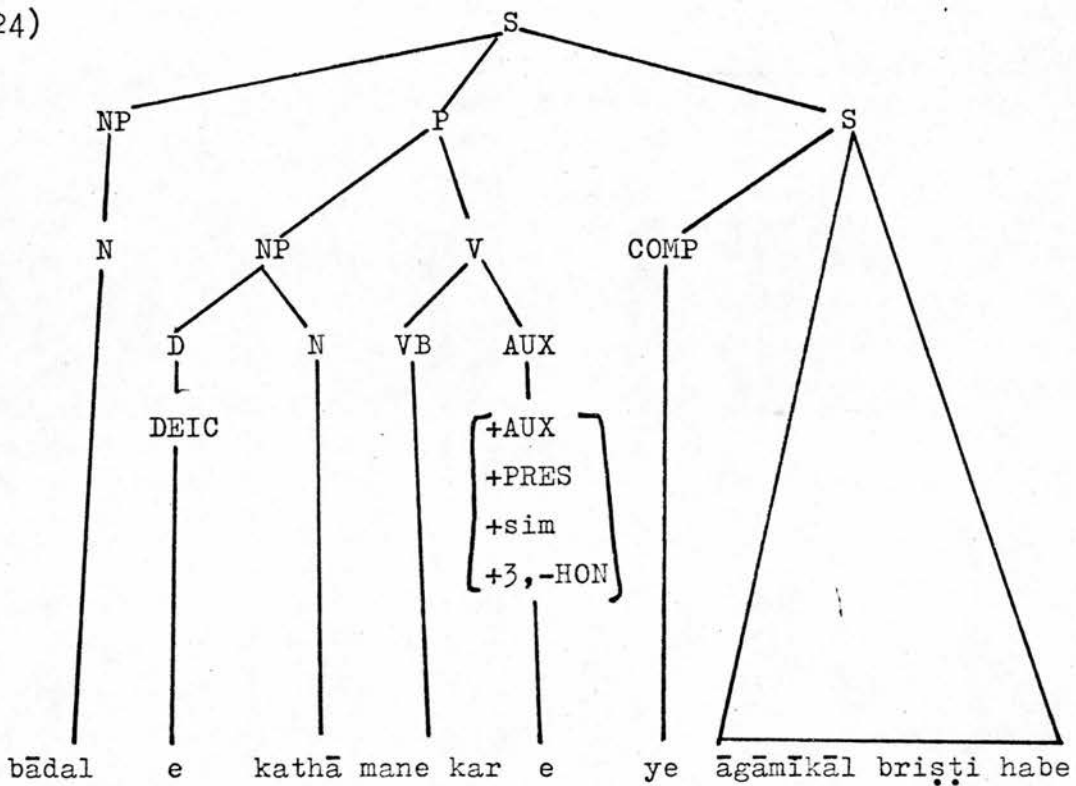


(23) will generate bādal mane kar e-kathā ye āgāmīkāl bṛiṣṭi habe, which is (11b). The abstract noun phrase e-kathā is optionally deletable from (23). If we delete it, we will derive bādal mane kar ye āgāmīkāl bṛiṣṭi habe, which is (11a). If we delete the abstract noun phrase as well as the complementizer from (23), we will derive bādal mane kar āgāmīkāl bṛiṣṭi habe: 'Badal thinks it will rain tomorrow'.

We can apply a transformation which we will call 'Object Extrapolation' to (22). This rule places the matrix verb at the end of the matrix proposition, and adjoins the complement

sentence as the right daughter of the matrix sentence. This rule will transform (22) into (24).

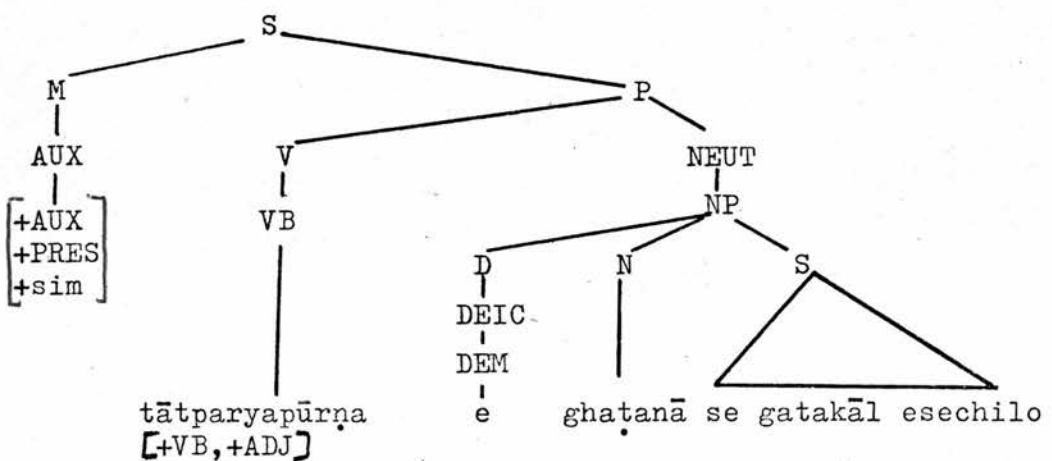
(24)



(24) will generate bādal e-kathā mane kar e ye āgāmīkāl bṛiṣṭi habe, which is synonymous with (11a, b). The structure (24) is more comfortable and natural than (23).

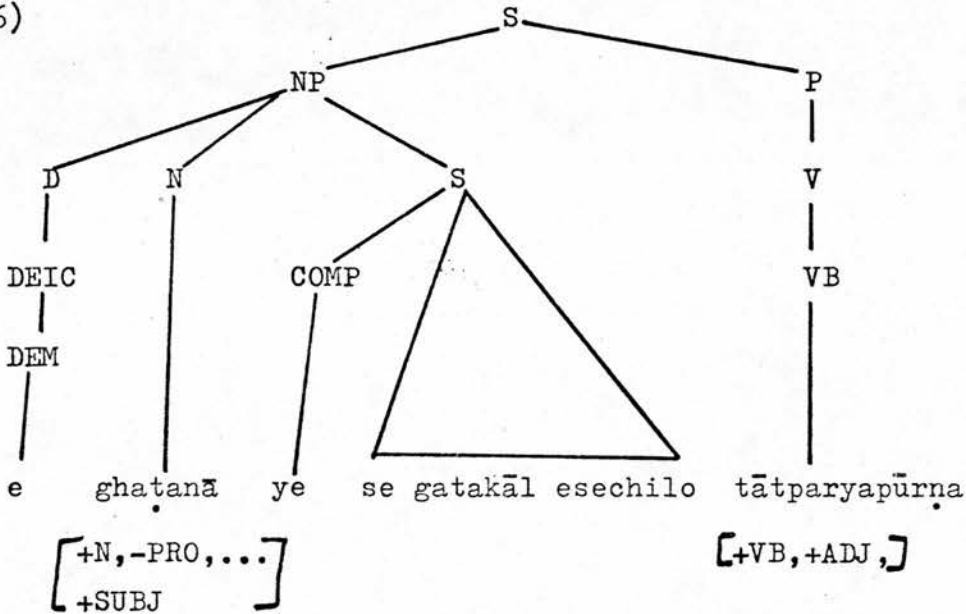
We will now consider a sentence in which the abstract noun phrase of an NP-Complement is realized as the subject. (13b) provides us with such a sentence. (13b) has the intermediate structure (25).

(25)



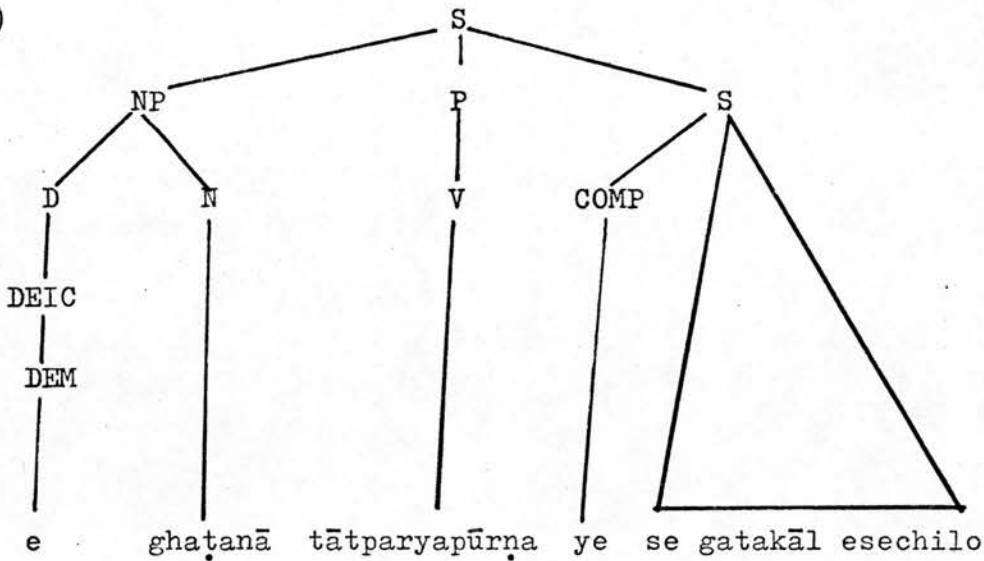
There is only one actant in (25), and so it will be the subject. The rule of subject marking will mark the NP dominated by NEUT as the subject, and the subject placement rule will move it to the sentence-initial position. The complementizer will be placed as the left daughter of the complement sentence. These operations will transform (25) into (26).

(26)



(26) will generate (13b). (13b) is a grammatical, but uncomfortable sentence. In a structure like (26), where the subject noun phrase contains a complement sentence, and is at the sentence-initial position, Bengali prefers an operation which can be called 'Extrapolation from Subject'. This operation moves the complement sentence to the sentence-final position and adjoins it as the right daughter of the matrix sentence. This operation will transform (26) into (27) (see page 403). The structure (27) is more comfortable and natural than the structure (26). Bengali has a tendency to extrapose subordinate clauses to the sentence-final or the sentence-initial position (cf., § 8.2.1).

(27)



(27) will generate e-ghaṭanā tātparyapūrṇa ye se gatakāl esechilo, which is (13a).

We will now consider another type of complementation, which differs from NP-Complementation in many ways. This type of complement sentence cannot be considered as an extension of an abstract noun phrase like e-kathā. This type disallows complementizer placement, and obligatory te-infinitivalization takes place in the complement sentence. We will call this type 'Infinitival Complementation'. Consider the examples below.

(28) a. ḍāktār bādalke dekhte lāglen: The doctor began to
 1 2 3 4 5 1 5 4
 examine Badal.
 3

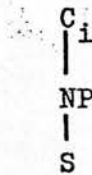
b. *ḍāktār lāglen e-kāje ye ḍāktār bādalke dekhlen: The
 doctor began in this action that the doctor
 examined Badal.

c. *ḍāktār lāglen ye ḍāktār bādalke dekhlen: The doctor
 began that the doctor examined Badal.

Traditionally a sentence like (28a) is considered as a simple sentence, and items dekhte lāg as a single compound verb root.

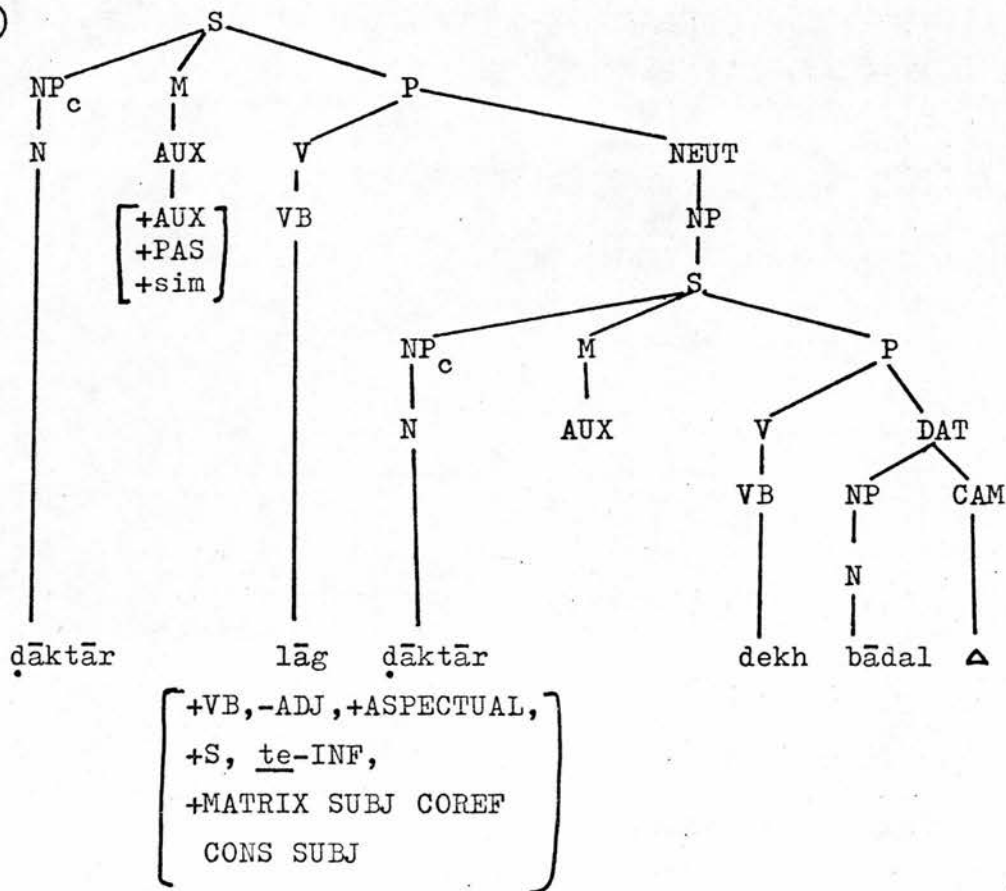
But a sentence of this type cannot be considered as a simple sentence, and items like karte lāg cannot be considered as single compound verb roots (cf., § 1.3.1). (28a) is a complex sentence derived by infinitival complementation. In such a complex sentence the matrix verb is usually a modal-aspectual verb like lāg: 'Begin and continue', pār: 'Be able', thāk: 'Continue' etc.,. These verbs are marked for obligatory te-infinitivalization of the complement sentence. Infinitival complementation differs syntactically from NP-Complementation. Infinitival complements have the underlying structure (29).

(29)

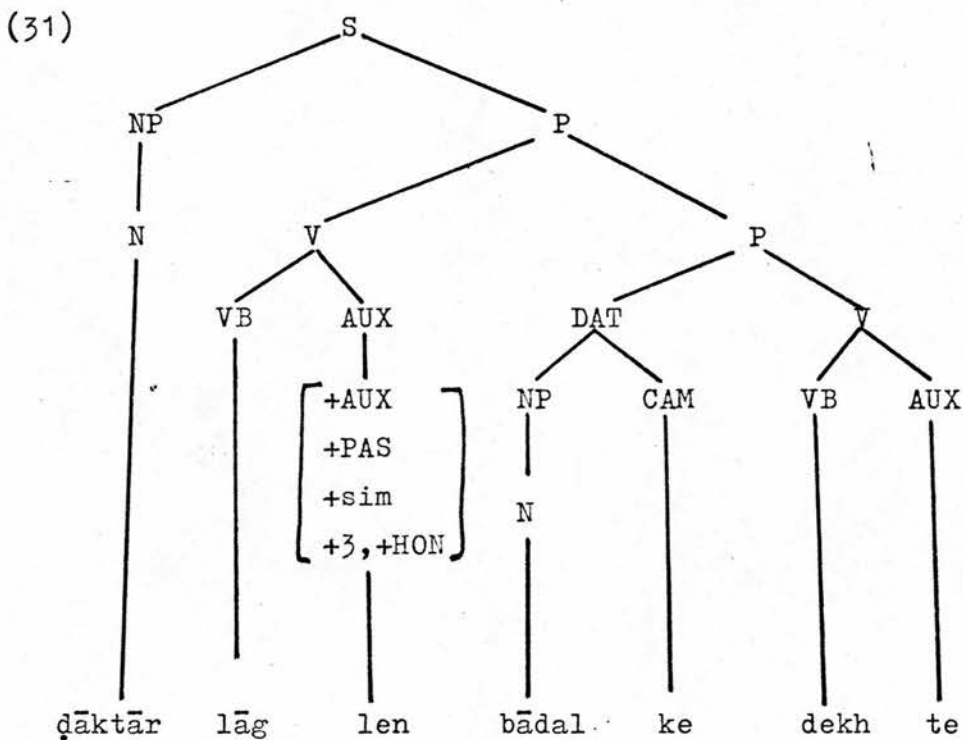


(28a) has the intermediate underlying structure (30).

(30)

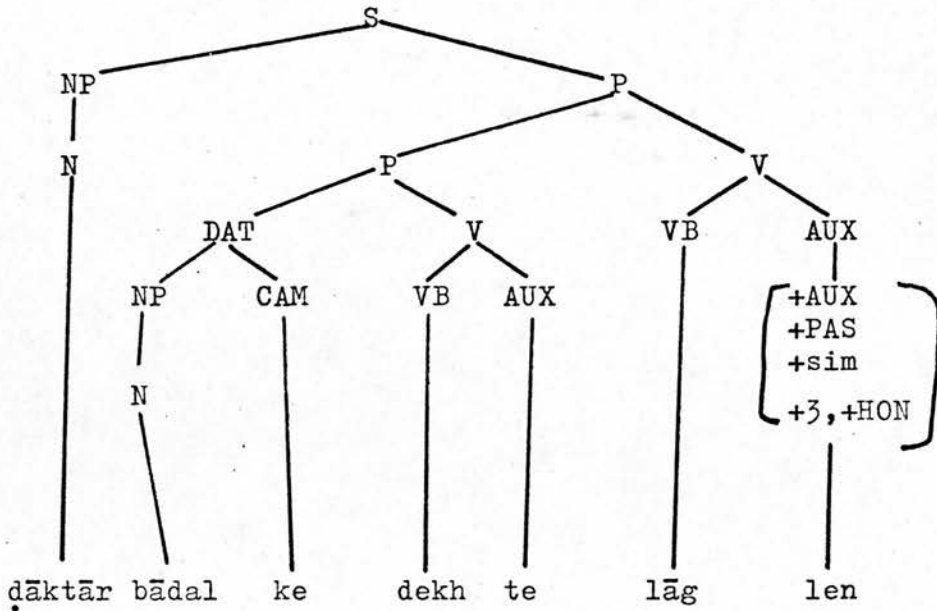


In (30) the AUX of the complement sentence has not been developed into a complex symbol because the matrix verb is marked for te-infinitivalization. The constituent subject will be deleted from (30) due to its coreference with the matrix subject. The complementizer placement rule will not apply in (30). The Subject-AUX concord rule will not apply in the complement sentence. Instead te will be attached to the AUX. By the application of all these and other relevant operations in (30), we will derive (31).



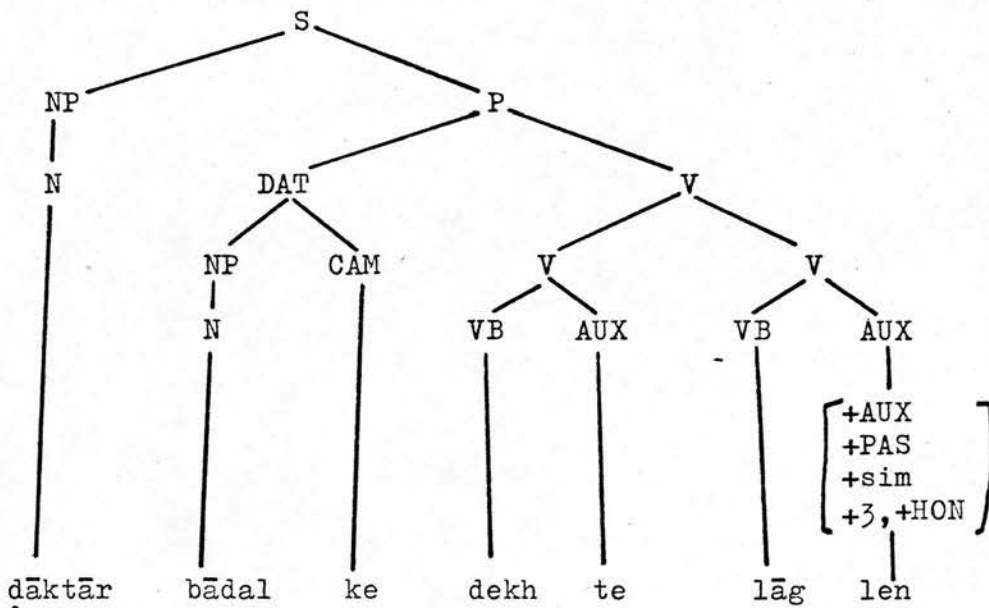
(31) will generate dāktār lāglen bādālke dekhte, which is (28a) in a different constituent order. The finite verb form is placed usually at the sentence-final position in Bengali. We can place the matrix V in (31) at the sentence-final position, and derive (32).

(32)



(32) will generate dāktār bādalke dekh te lāg len, which is (28a). (32) is the derived syntactic structure of (28a). The morphophonological rules can derive the ultimate phonetic structure (33) from the syntactic structure (32).

(33)



(33) differs structurally from (32), and looks like a simple sentence structure. This superficial appearance of (33) is behind the traditional analysis of verb forms like kāte lāg: 'Begin to do', balte thāk: 'Continue to speak' etc., as

single compound verb roots (cf., § 1.3.1). But our derivation shows that sentences of the type (28a) are complex sentences derived by infinitival complementation, and verb forms like those above are not single verb roots in the underlying structure.

10.4 Sentence Pronominalization.

We have so far assumed the existence of a sentence pronominalization transformation which uses one complement sentence to reduce another into a pronoun under the conditions given in (10). The complement sentences involved in this process must be synonymous and structurally identical. We will call two sentences coreferential when they are synonymous and structurally identical. We have used identical indices to indicate noun phrase coreferentiality, but coreferential sentences and their anaphoric pronouns have been italicized in this work. The rule of S-Pronominalization can be given informally as (34).

(34) S-Pronominalization Rule.

One complement sentence can be used to pronominalize another in a structure when they satisfy the conditions in (10). The pronominalized sentence is realized as the pronoun tā in the surface structure.

First we will consider the application of the sentence pronominalization rule in structures which involve NP-Complementation, and see whether sentences like those in (2) are derived by the sentence pronominalization rule, or by some other process. There are two ways in which these sentences can be derived. In one way, they are derived by the sentence pronominalization rule, which we will consider first.

Consider the following examples.

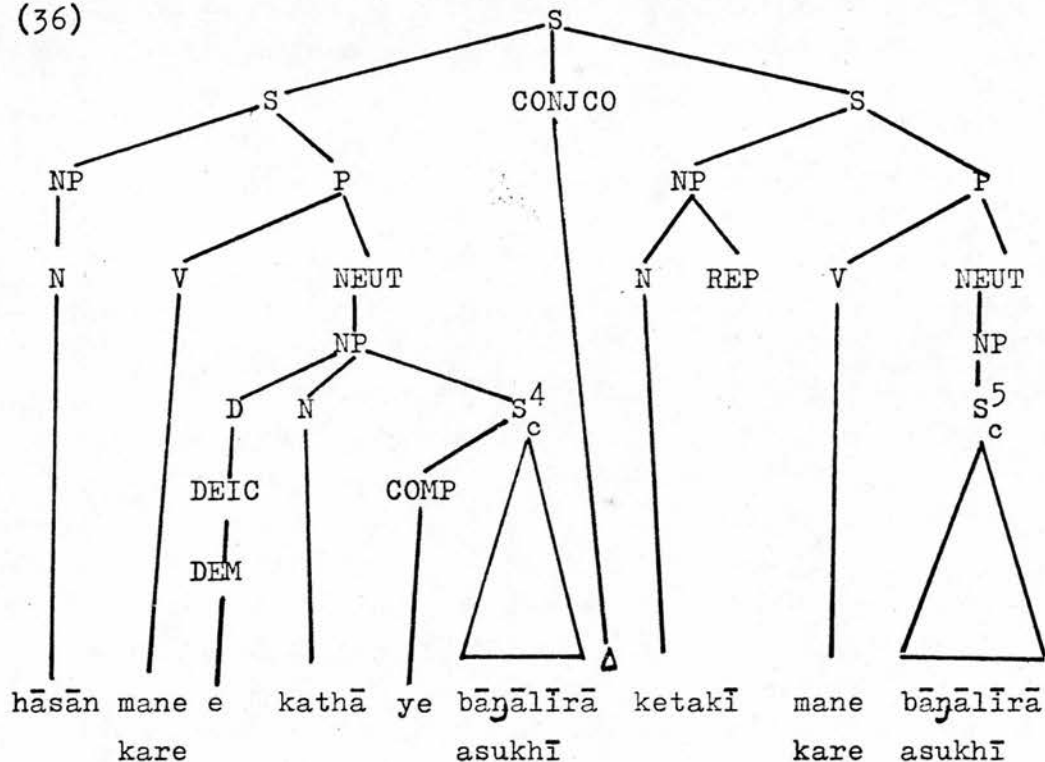
- (35) a. $\text{hāsān mane kare e-kathā ye } \underline{\text{bāṅālīrā asukhī}} \text{ ār}$
 2 3 4 5 6 7 8
 $\text{ketakīo mane kare e-kathā ye } \underline{\text{bāṅālīrā asukhī}} \text{: Hasan}$
 10 11 12 13 14 15 16
 thinks this proposition that the Bengalees are
 2 3 4 5 6
 unhappy and Ketaki, too, thinks this proposition that
 7 8 10 11 12 13 14
 the Bengalees are unhappy.
 15 16
- b. $\text{hāsān mane kare (e-kathā) ye } \underline{\text{bāṅālīrā asukhī}} \text{ ār}$
 6 7
 $\text{ketakīo } \underline{\text{tā}} \text{ mane kare: Hasan thinks (this Proposition)}$
 10
 that the Bengalees are unhappy and Ketaki thinks
 6 7
 so, too.
 10
- c. * $\text{hāsān } \underline{\text{tā}} \text{ mane kare ār ketakīo mane kare (e-kathā) ye}$
 2
 $\underline{\text{bāṅālīrā asukhī}} \text{: Hasan thinks it and Ketaki, too,}$
 2
 thinks (this proposition) that the Bengalees are
 unhappy.

The coordinate conjoined structure (35a) has two synonymous and structurally identical complement sentences. If the sentence pronominalization rule is applied forwards to (35a), we derive the grammatical sentence (35b); and if the rule is applied backwards, we derive the ungrammatical sentence (35c). This is what the rule (10) predicts. But the rule does not apply to the structure which immediately underlies (35a). If it applies to the structure which immediately underlies (35a), we will derive the ungrammatical sentence (35d).

- (35) d. * $\text{hāsān mane kare e-kathā ye } \underline{\text{bāṅālīrā asukhī}} \text{ ār}$
 $\text{ketakīo } \underline{\text{e-kathā}} \underline{\text{tā}} \text{ mane kare: Hasan thinks that the}$
 Bengalees are unhappy and Ketaki, too, thinks this
 proposition so.

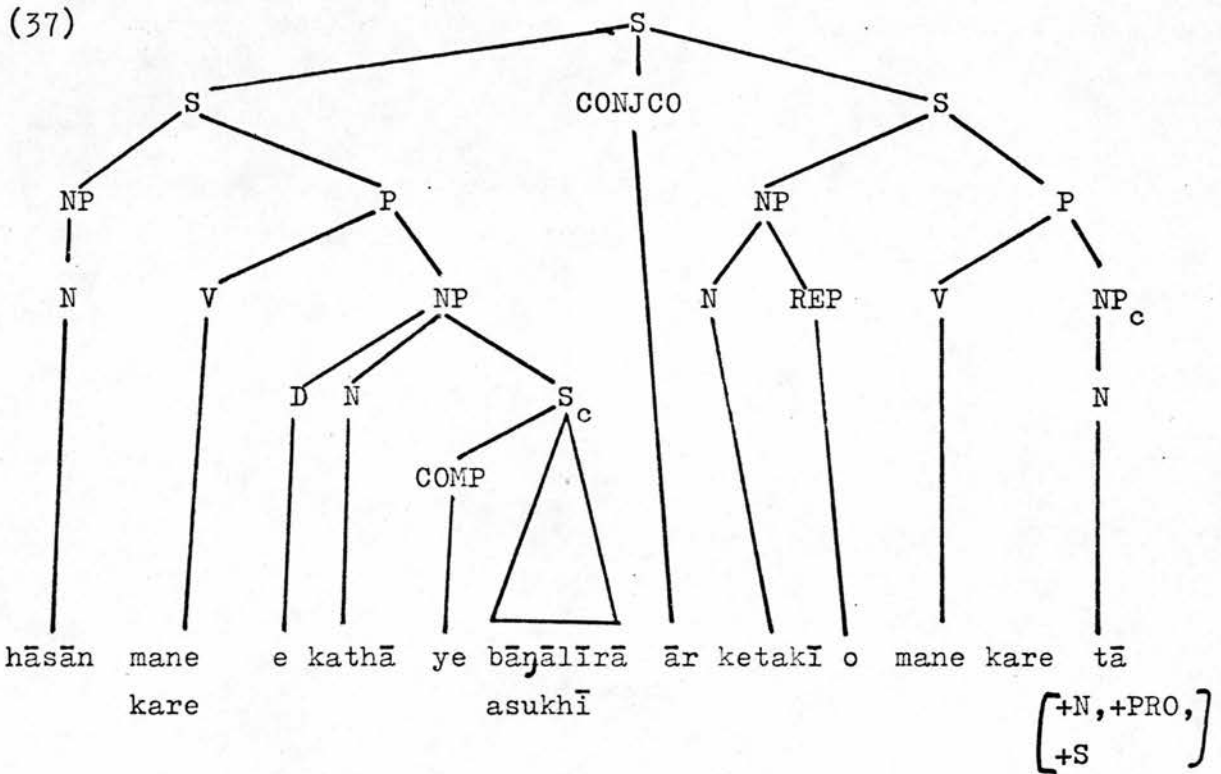
The ungrammaticality of (35d) is due to the presence of the abstract noun phrase e-kathā of the pronominalized sentence. The sentence pronominalization rule applies after the deletion of the abstract noun phrase and the complementizer from the pronominalizable complement sentence. So (35a) has the intermediate structure (36) at the point the sentence pronominalization rule applies.

(36)



As in most of our examples irrelevant details are omitted from (36), and second lexical items supplied for the sake of readability. In (36) the sentence pronominalization rule will apply forwards. This rule will use S^4 to reduce S^5 into a pronoun. The sentence pronominalization rule will throw away the S node of the pronominalized sentence, and accumulate the features [+PRO,+S] under the NP which dominates the pronominalizable sentence prior to the application of

the pronominalization rule. This is necessary for second lexical insertion rule, which will insert a single lexical item to the derived node. The application of the sentence pronominalization and other relevant rules will transform (36) into the derived structure (37).



We have attached the abstract pronoun tā to the pronominalized node, and other second lexical items have also been attached to the appropriate nodes in (37). (37) will generate hāsān mane kare e-kathā ye bāṅālīrā asukhī āṛ ketakī o mane kare tā, which is (35b) in a different constituent order. We will derive (35b) by placing the V of the forward conjunct in (37) sentence-finally. The sentences in (38) are also derivable from (37) by carrying out some more operations on it.

- (38) a. hāsān mane kare ye bāṅālīrā asukhī āṛ ketakī o tā
 mane kare: Hasan thinks that the Bengalees are
 unhappy and Ketaki thinks so, too.

- b. hāsān e-kathā mane kare ye bāṅālīrā asukhī ār
ketakīo tā mane kare: Hasan thinks this proposition
that the Bengalees are unhappy and Ketaki thinks
so, too.
- c. hāsān mane kare bāṅālīrā asukhī ār ketakīo tā mane kare:
Hasan thinks the Bengalees are unhappy and Ketaki
thinks so, too.

This analysis shows that there is a transformation which reduces one complement sentence into a pronoun by another sentence under certain conditions. Now we will consider the alternative way in which the pronoun tā can arise and refer anaphorically to a sentence. We will again consider the derivation of tā in (35b). In the above analysis we derived (35b) from (35a) first by deleting the abstract noun phrase and the complementizer from the pronominalizable complement sentence and then by the application of the sentence pronominalizable rule. In an alternative way, we can derive the sentences in (39) from (35a).

- (39) a. hāsān mane kare e-kathā ye bāṅālīrā asukhī ār
ketakīo mane kare e-kathā: Hasan thinks this proposition
that the Bengalees are unhappy and Ketaki
thinks this proposition, too.
- b. hāsān mane kare e-kathā ye bāṅālīrā asukhī ār
ketakīo mane kare tā: Hasan thinks this proposition
that the Bengalees are unhappy and Ketaki
thinks this/that, too.
- c. hāsān mane kare e-kathā ye bāṅālīrā asukhī ār ketakīo
tā mane kare: Hasan thinks this proposition that the
Bengalees are unhappy and Ketaki thinks this/that, too.

- (41) a. Bill resents it that people are always comparing him to Mozart.
- b. They did not mind it that a crowd was beginning to gether in the street.

Kiparsky and Kiparsky (1971) claim that 'it' in these examples arises due to optional pronominal reduction of 'the fact' in factive complements. Stockwell et al (1973, 551-53) argue against this claim. We will claim that tā in (40) arises due to optional pronominal reduction of abstract noun phrases in NP-Complementation. The sentences in (40) are related respectively to those in (42).

- (42) a. āmi jāni e-kathā ye ketakī rūpasī: I know this proposition that Ketaki is beautiful.
- b. āmi bali ni e-kathā ye tumi mithyābādī: I did not say this proposition that you were a liar.
- c. { e-ghaṭanā ye bādal bārī geche tātpar̥yapūr̥ṇa: This event that Badal has gone home is significant.
e-ghaṭana tātpar̥yapūr̥ṇa ye bādal bārī geche: This event is significant that Badal has gone home.

These sentences involve NP-Complementation on abstract noun phrases like e-kathā and e-ghaṭanā. Bengali has an operation which extraposes the NP-Complements to the sentence-initial position. This operation is sometimes accompanied by another operation which places the complementizer ye inside the complement sentence, usually preceding a major constituent. The operation which places the complementizer inside the complement sentence is usually carried out if the speaker believes in the factuality of the complement sentence.

Otherwise the complementizer is simply deleted, and the complement sentence is extraposed to the sentence-initial position. We assume that the complementizer placement inside the complement sentence is possible when the matrix predicate is factive. Here we cannot go into the details of this operation's conditions. These operations will transform the sentences in (42) into those in (43).

(43) a. $\left\{ \begin{array}{l} \text{ketakī ye rūpasī} \\ \text{ketakī rūpasī} \end{array} \right\}$, e-kathā āmi jāni: Lit., (That)

Ketaki is beautiful, I know this proposition.

b. $\left\{ \begin{array}{l} \text{?tumi ye mithyābādī} \\ \text{tumi mithyābādī} \end{array} \right\}$, e-kathā āmi bali ni: Lit.,

(That) you are liar, I did not say this proposition.

c. $\left\{ \begin{array}{l} \text{bādal ye bārī geche} \\ \text{bādal bārī geche} \end{array} \right\}$, e-ghaṭanā tātparyapūrṇa: Lit.,

(That) Badal has gone home, this event is significant.

The first sentence in each set in (43) contains the complementizer ye inside the complement sentence. This indicates that the speaker emphatically believes in the factuality of these propositions. The second sentence in each set above does not contain the complementizer inside, and so they are understood as mere assertions. The sentence ?tumi ye mithyābādī, e-kathā āmi bali ni in (43b) is odd, because it emphasizes the factuality of the complement sentence ; and at the same time the matrix sentence, being negative, contradicts this proposition. In (43) the abstract noun phrases e-kathā and e-ghaṭanā are present in the matrix sentences. They can be optionally reduced to the abstract pronoun tā.

If we reduce these abstract noun phrases into the abstract pronoun tā, we will derive the following sentences from (43), respectively.

- (44) a. $\left\{ \begin{array}{l} \text{ketakī ye rūpasī} \\ \text{ketakī rūpasī} \end{array} \right\}$, tā āmi jāni: Lit., (That)
Ketaki is beautiful, I know that.
- b. $\left\{ \begin{array}{l} \text{?tumi ye mithyābādī} \\ \text{tumi mithyābādī} \end{array} \right\}$, tā āmi bali ni: Lit., (That)
you are a liar, I did not say that.
- c. $\left\{ \begin{array}{l} \text{bādal ye bārī geche} \\ \text{bādal bārī geche} \end{array} \right\}$, tā tātparityapūrṇa: Lit.,
(That) Badal has gone home, that is significant.

Thus we see that the abstract pronoun tā in (40, 44) arises due to the pronominal reduction of the abstract noun phrases of NP-Complements. This also gives support to our claim that tā in (35b) and (39b, c) does not arise due to sentence pronominalization. It arises due to pronominalization of the abstract noun phrases of NP-Complements. So we can say that structures with coreferential noun phrase complement sentences do not involve sentence pronominalization. We can replace the rules (10) and (34) by (45).

(45) Complement S-Deletion and Pronominalization of Abstract Noun Phrases.

If a structure contains two NP-Complement sentences which are synonymous and structurally identical, then one of them can be deleted and its dominating noun phrase can be reduced to the abstract pronoun tā. This rule operates

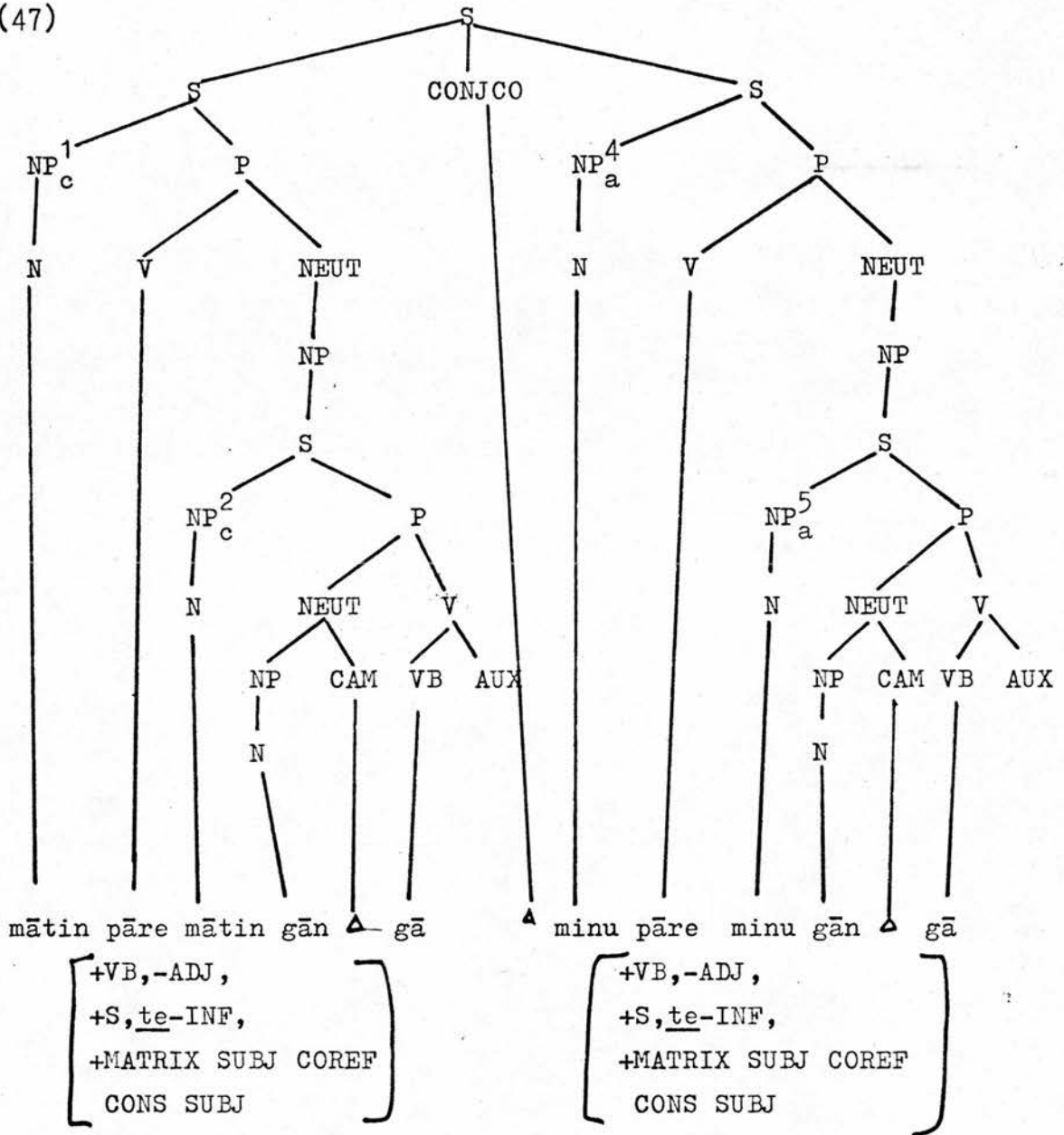
forwards in a coordinate conjoined structure, and both forwards and backwards in a subordinate conjoined structure.

We will now consider some structures which involve infinitival complementation, and see how the abstract pronoun arises in such structures. Consider the examples in (46).

- (46) a. mātin gān gāite pāre ār minuo tā pāre: Matin can
 2 3 4 5 7 8 9 4
 sing songs and so can Minu.
 3 2 5 8 9
- b. mātin gān gāite pāre ār minuo gān gāite pāre: Matin
 2 3 4 5 7 8 9 10
 can sing songs and Minu, too, can sing songs.
 4 3 2 5 7 10 9 8
- c. S [S [mātin pāre S [mātin gān gāy] CONJCO
 2 4 5
 S [minu pāre S [minu gān gāy]]: [[Matin
 7 9 10
 can [Matin sing songs]] and [Minu can [Minu sing
 2 5 4 7 10
 songs]].
 9

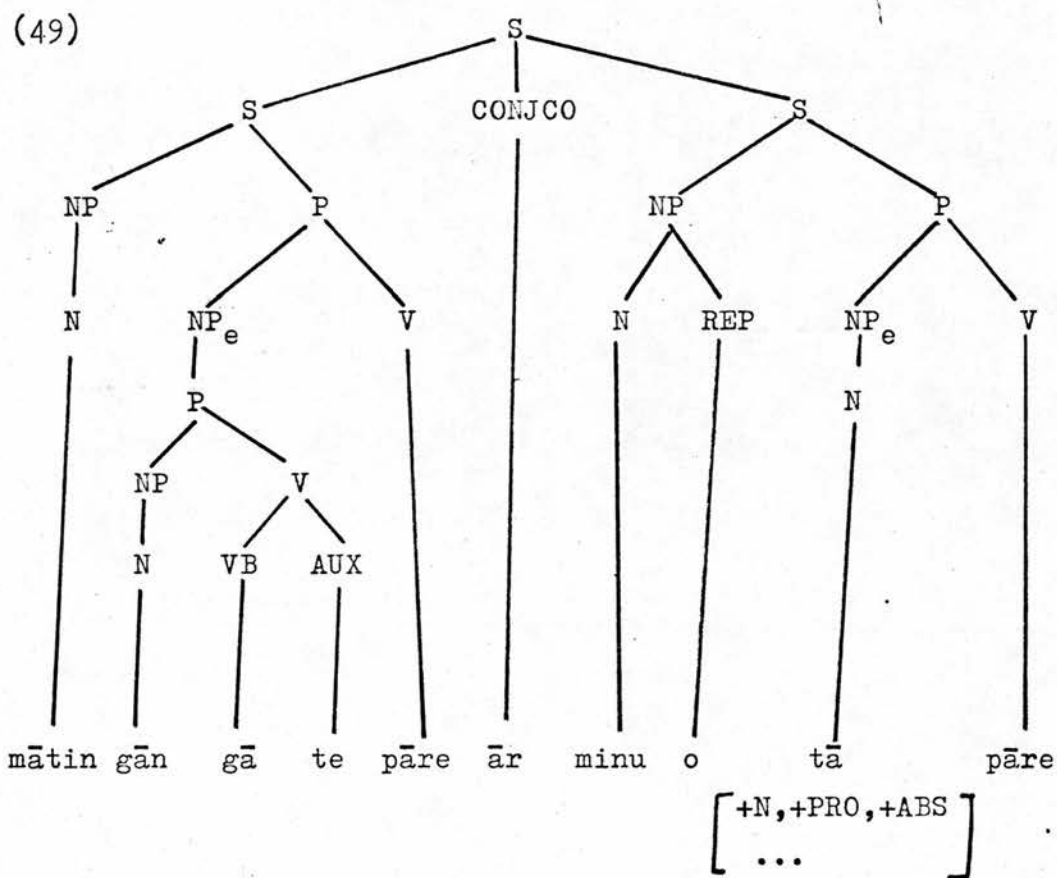
In (46a) the pronoun tā refers to gān gāite in the first conjunct. This sentence is related to (46b), where both the conjuncts contain the nominalizations gān gāite. The way (46a) is understood indicates that tā in this sentence arises due to pronominalization of the infinitival nominalization gān gāite of the second conjunct in (46b). (46a, b) involve infinitival complementation, and so abstract noun phrases like e-kathā, e-ghaṭanā etc., are not present in the underlying structure of these sentences. Accordingly, we cannot assume that tā arises through pronominal reduction of these noun phrases. These sentences have the intermediate structure (46c), which has the configuration (47).

(47)



In each conjunct in (47) the matrix and constituent subjects are coreferential. So the constituent subjects will be deleted under their coreference with the matrix subjects. That is, NP¹ will be used to delete NP², and NP⁴ will be used to delete NP⁵ in (47). The Subject-AUX concord rule will not apply in the constituent sentences; instead, the infinitival marker te will be attached to the constituent AUX's. In this way the constituent sentences will be reduced into infinitival

as pronominalization of infinitival nominalizations. The rule will apply forwards in (48) as it is a coordinate conjoined structure. We assume that the pronominalization rule will apply after the S-nodes of the infinitival nominalization have been pruned. So they will be taken as noun phrases. The noun phrase in the left conjunct will be used to reduce the noun phrase in the right conjunct into a pronoun. This rule will delete all the constituents dominated by the pronominalized noun phrase, and will gather the features $[+N,+PRO,+ABS]$ under this noun phrase. The application of this and other rules will transform (48) into the derived structure (49).



The second lexical items have been attached to the derived nodes in (49). The terminal string of (49), after the application of the morphophonological rules, will generate mātin gān gāite

pāre ār minuo t̄ā pāre, which is (46a).

Thus we see that the pronoun t̄ā does not arise due to the application of the sentence pronominalization rule in structures which involve infinitival complementation. Instead, it arises due to the pronominalization of an infinitivally nominalized noun phrase. Accordingly, we can conclude that Bengali does not have a sentence pronominalization rule. The pronoun t̄ā arises in structures which involve NP-Complementation by the pronominal reduction of abstract noun phrases like e-kathā, e-ghaṭanā etc.,. These noun phrases dominate the NP-Complements. In structures which involve infinitival complementation, the pronoun t̄ā arises due to pronominalization of infinitival nominalizations.

THE LEXICON11.0 Introduction.

The concept of a lexicon as a part of transformational generative grammar has its origin in 'Aspects of the Theory of Syntax' (1965), where Chomsky proposes a division of the base component of the grammar into a categorial component and a lexicon. In earlier transformational grammars lexical items were introduced by the terminal rewriting rules of the categorial component. In 'Aspects' (1965, 84-90) Chomsky proposes that lexical items, instead of being introduced by rewriting rules of the categorial component, should be listed in a lexicon from which they would be introduced into phrase markers by lexical insertion rules. Chomsky (1965, 84) says:

.. the base of the grammar will contain a lexicon, which is simply an unordered list of all lexical formatives. More precisely, the lexicon is a set of lexical entries, each lexical entry being a pair (D, C), where D is a phonological distinctive feature matrix "spelling" a certain lexical formative and C is a collection of specified syntactic features (a complex symbol).

This separation of the lexicon from the categorial component has many advantages. It allows the categorial component to be context-free and more general as much of the burden that the categorial component used to bear is now borne by the lexicon, which lists each entry with an associated feature matrix. We have mentioned earlier that this work makes use of a lexicon divided into two parts: (a) The first lexicon, and (b) The

second lexicon. The first lexicon contains all the content words of the language, which are introduceable into underlying phrase markers; and the second lexicon contains grammatical items, and those items which are derived by transformational operations. Besides these items, we have some lexical items, which are inserted by transformational rules. These items are not listed in any lexicon. These are, for example, the relative deictic ye, the complementizer ye, and the infinitive markers te, le and e. We could list these items in the lexicon and introduce them later by the second lexical insertion rule. Instead, we insert them by transformational operations simply because this is simpler.

This chapter, which is peripheral to our main topic, has been included only to demonstrate how the grammar is related to the lexicon(s). So we shall not attempt any serious investigation into the theory and practice of the lexicon, and in the sample lexical entries given at the end of each lexicon many features will be left unspecified, although they may be theoretically necessary (cf., Katz and Fodor (1963), Katz and Postal (1964), Chomsky (1965), Gruber (1965, 1967), Weinreich (1966), McCawley (1968a, b), Fillmore (1968a, b, 1969), Botha (1968), Stockwell et al (1973), and Hudson (1976)).

11.1 The First Lexicon.

This is the main lexicon and contains all the content words of the language. The lexical items listed in it are introduceable in the underlying phrase markers.

11.1.1 Lexical Insertion Rule.

Chomsky (1965, 84-90; 120-23) suggested two alternative ways of lexical insertion into an underlying phrase marker. In one way the rewriting rules of the categorial component will generate derivations terminating with strings which consist of grammatical formatives and complex symbols. He calls such a string a 'preterminal string'. A terminal string is formed from a preterminal string by insertion of lexical items by a lexical rule. Chomsky's (1965, 84) lexical insertion rule is quoted below:

If Q is a complex symbol of a preterminal string and (D, C) is a lexical entry, where C is not distinct from Q , then Q can be replaced by D .

In the second proposal the categorial component consists of a number of context-free phrase structure rules which generate a string of dummy symbols, Δ , and grammatical formatives (cf., Chomsky (1965, 121-22)). The dummies indicate the position of the lexical categories. The lexical items are then inserted by a substitution transformation where the complex symbol in the lexical entry is the structure index for the transformation, and the lexical item is appropriate for substitution if the tree meets the conditions of the structure index specified by the complex symbol. This method of lexical insertion in the underlying structure has been adopted in this work. The base rules of the grammar generate preterminal strings consisting of dummy symbols, Δ , and a complex symbol for the AUX (cf., §1.2). The first lexical insertion rule applies at this stage. The first lexical insertion rule inserts lexical items into phrase markers by a substitution transformation as described above.

The first lexical insertion rule is given in (1).

(1) The First Lexical Insertion Rule.

A lexical item listed in the first lexicon can be inserted in place of a dummy symbol, Δ , in an underlying phrase marker if it satisfies the conditions of the structure index specified by the complex symbol associated with the lexical item.

The second lexical insertion rule applies after the last rule of the transformational component. It is basically similar to the first lexical insertion rule, but has some characteristics of its own. The second lexical insertion rule has something in common with Chomsky's first method of lexical insertion. It introduces a lexical item to a derived node if the complex symbol associated with the lexical item matches with the one associated with the derived node (cf., § 11.2).

We assume that lexical items of different categories should be inserted to underlying phrase markers in a particular order. We presume that the principle should be something like this: a lexical category which is selectionally dominant should be inserted before one which is selectionally dependent. This principle will allow verbs to be inserted before nouns, and nouns before a determiner. Chomsky (1965) argues that a noun should be inserted before a verb. But his claim is untenable as verbs are really dominant over nouns (cf., Stockwell et al (1973, 721). A lexical item will bring all its features with it when it is inserted to a phrase marker (theoretically so, but in this grammar we have omitted many redundant features from the complex symbols of lexical items for economy).

11.1.2 Forms of Lexical Entries.

Chomsky (1965, 84) considers a lexical entry to be a pair (D, C), where D is a phonological feature matrix and C is a set of specified syntactic features. We have ignored the phonological features altogether, and so the "spelling" for each entry is given in the Roman alphabet with a gloss in English (if possible). We have considered only syntactic and semantic features for each entry. Each lexical entry has associated with it a complex symbol containing four types of feature: category features, inherent features, contextual features, and sometimes, rule features. Theoretically these four types of feature should be specified in a complex symbol, but in the sample entries we have specified contextual and rule features for a few items. A category feature denotes the category to which the item belongs such as verb, noun or deictic etc.,. Inherent features denote inherent properties like animate, human, abstract etc.,. The rule features indicate the rules which apply to the item. The contextual features refer to the environments in which the lexical item can appear. This type of feature has been considered specially for verbs and has been represented in the manner of Fillmore (1968a). In his case frames Fillmore (1968a) uses parentheses for optional cases, no parenthesis for obligatory cases, and linked parentheses for optional cases from which at least one must be chosen. We have also used parentheses for optional cases, no parenthesis for obligatory cases, and impossible cases have been omitted. We could not use linked parentheses because the cases are ordered in a particular manner in the underlying structure.

11.1.3 Feature Specification.

It was Chomsky (1965, 81-83) who first introduced the system of feature specification for lexical items similar to the form of a phonological distinctive feature matrix. He allowed three values for a feature: positive, negative and unspecified. Stockwell et al (1973, 726) indicate that three values for a feature may not be sufficient for the purpose, and they have proposed five values for a feature. The system of feature specification adopted in this work is similar to that used in Stockwell et al (1973). Stockwell et al (1973, 728) have proposed the following conventions for feature specification:

- (2) (1) + positive specification
- (2) (2) - negative specification
- (3) (3) * obligatory specification
- (4) (4) +/- optional specification
- (5) (5) absence of specification would mean
that the feature was irrelevant

These conventions have been used in this work. We have said earlier that the following inherent features need consideration in subcategorizing the nouns in Bengali (cf., §2.1):

- (3) [COMMON], [COUNT], [HUMAN], [ANIMATE], [HONORIFIC],
[PEJORATIVE], [MALE], [ABSTRACT], [LOCATION]

(3) is an incomplete list of features required to subcategorize the nouns in Bengali. Most of the features above have three values: a particular feature may be positively specified for an item, and negatively specified for another item; and may be irrelevant for a third item. But some features need upto five values. A lexical item is positively specified for a feature if the item has the positive value for the feature (for example,

yubak: 'Young man' is [+MALE]); negatively specified for a feature if the item has the negative value for the feature (for example, yubatī: 'Maiden' is [-MALE]). An item is unspecified for a feature if the feature is irrelevant for the item (for example, the feature [MALE] is irrelevant for bai: 'Book'). A feature is optionally specified if it is optional for the item (for example, [MALE] is optional for garu: 'Cow', and so it will have the specifications [+/-MALE]). Apart from these values for the features, a lexical item may be specified for 'obligatory specification' for a feature. This is indicated by an asterisk appearing with the feature. An obligatory specification means that the value for the feature is left unspecified in the complex symbol, but the value must be specified positively or negatively before the lexical item is introduced to a tree. For example, āmi: 'I' has the complex symbol

[+N,+PRO,-INT,+COUNT,+DEF,+HUM,+1,*2,+HON,*3,+HON,*PL] .

,-HON, ,-HON,
,+PEJ

The values for the features which have been specified for obligatory specification (*) must be positively or negatively specified before the pronoun āmi is inserted to a tree. If the value for * is taken as negative for all the features specified for obligatory specification, āmi will have the complex symbol

[+N,+PRO,-INT,+COUNT,+DEF,+HUM,+1,-2,+HON,-3,+HON,-PL], and

,-HON, ,-HON,
,+PEJ

will be understood as singular ([-PL]). But if the value for * for any feature is taken as positive, āmi will be understood as plural ([+PL]). Suppose that the value for * is taken as positive only for the features [2,HON], then āmi will be understood as

plural, and it will have the feature specifications $[+N, +PRO, -INT, +COUNT, +DEF, +HUM, +1, +2, +HON, -3, +HON, +PL]$. This complex symbol indicates inclusive use of $\bar{a}mi$ ($\bar{a}pani$: 'You ([+HON])' and $\bar{a}mi$: 'I'). As this complex symbol is specified as [+PL], a later plural segment transformation will derive $\bar{a}mar\bar{a}$.

Apart from the features for nouns mentioned in (3), the features of person, such as [1 (first person)], [2 (second person)] and [3 (third person)], must be included in the feature inventory, but need not be (except for pronouns) specified each time with each noun. As the category of person is defined with reference to the notion of participant-roles, noun items have the features $[-1, -2, +3]$; and this can be stated by a general rule. But the feature of person must be specified in the complex symbol for each personal pronoun, because they each differ from the other for the feature of person.

11.1.4 Redundancy Rules.

The grammar will make use of redundancy rules because they reduce the number of feature specifications in a complex symbol. These rules allow not to specify an item with a particular feature in the complex symbol if that feature is predictable from another feature. Thus we need not specify an item as [+ANI] if it has the specification [+HUM] or [-HUM]. In Bengali only [+HUM] nouns can be [+HON]. So the specification [+HON] in an entry predicts that it is also [+HUM, +ANI]. All [-HUM] and [-ANI] nouns are [-HON] in Bengali. So it is redundant to specify these nouns as [-HON], because the features [-HUM] and [-ANI] predict that these nouns are [-HON].

11.1.5 Sample First Lexicon.Redundancy Rules:

These rules will apply to each lexical item before lexicalization. The rules given below are not exhaustive. They are given as examples.

[+N]	⇒	[-PRO]
[+PRO]	⇒	[+N]
[+HON]	⇒	[+HUM,+ANI]
[+HUM]	⇒	[+ANI]
[-ANI]	⇒	[-HON]
[+COUNT]	=	{ *PL _____ }
[-COUNT]	⇒	[-PL]
[+VB]	⇒	[-ADJ]
[+ADJ]	⇒	[+VB]
[+ABS]	⇒	[-ANI]
[+N,-PRO]	=	[-1,-2,+3]

Some Nouns:bai: 'Book'+N,+COM,+COUNT,
-ANI,+ABS,bhadralok: 'Gentleman'+N,+COM,+COUNT,
+HON,+MALE,Jal: 'Water'+N,+COM,-COUNT,
-ANI,-ABS,path: 'Road'+N,+COM,+COUNT,
-ANI,+LOC,pākhī: 'Bird'+N,+COM,+COUNT,
-HUM,svapna: 'Dream'+N,+COM,+COUNT,
+ABS,Some Verb Roots:ās: 'Come'+VB,
+___(INS)(LOC)AGTbhālo: 'Good'+ADJ,
+___NEUTdekh: 'See'+VB,
+___(NEUT)DATkhul: 'Open'+VB,
+___(INS)NEUT AGTlāg: 'Begin and
continue'+VB,+ASPECTUAL,
+___NEUT AGT
+S, te-INF,
+MATRIX SUBJ COREF
CONS SUBJmār: 'Hit'+VB,
+___(INS)DAT AGTmane kar: 'Think'+VB,
+___NEUT DAT,
+DD,+S,śun: 'Hear'+VB,
+___NEUT DAT,
+S,tātparyapūrṇa:'Significant'
+VB,+FACT,
+___NEUT,
+S,

Pronouns and Other Items:āmi: 'I'

+PRO,-INT,+COUNT,

+DEF,+HUM,

+1,

*2,+HON,

,-HON,

,+PEJ

*3,+HON,

,-HON,

āpani: 'You'

+PRO,-INT,+COUNT,

+DEF,

-1,

+2,+HON,

*3,+HON,

,-HON,

tumi: 'You'

+PRO,-INT,+COUNT,

+DEF,+HUM,

-1,

+2,-HON,

*3,-HON,

,+HON,

tui: 'You'

+PRO,-INT,+COUNT,

+DEF,+HUM,

-1,

+2,+PEJ

*3,-HON,

,+HON,

tini: 'He'

+PRO,-INT,+COUNT,

+DEF,

-1,

-2,

+3,+HON

se: 'He'

+PRO,-INT,+COUNT,

+DEF,+HUM,

-1,

-2,

+3,-HON

ini, ẽ: 'He'

+PRO,-INT,+COUNT,

+DEF,

-FAR,

-1,

-2,

+3,+HON,

uni, õ: 'He'

+PRO,-INT,+COUNT,

+DEF,

+FAR, IN SIGHT,

-1,

-2,

+3,+HON,

e: 'He'

+PRO,-INT,+COUNT,

+DEF,+HUM,

-FAR,

-1,

-2,

+3,-HON,

<u>o</u> : 'He'	<u>tā</u> : 'That (thing)'	<u>paraspar</u> : 'One another'
+PRO,-INT,+COUNT,	+PRO,-INT,-COUNT,	+N,-COUNT,+HUM,
+DEF,+HUM,	+DEF,+ABS,	+RECIPROCAL,
-1,	-1,	
-2,	-2,	
+3,-HON,	+3,-	

<u>ke</u> : 'Who'	<u>ki</u> : 'What'	<u>kab</u> : 'When, day'
+PRO,+INT,+COUNT,	+PRO,+INT,+COUNT,	+PRO,+INT,-COUNT,
+DEF,+SPECIFIC,	+DEF,+SPECIFIC,	+DEF,
+HUM,	-HUM/+ABS,	+TEMP, DAY,
-1,	-1,	
-2,	-2,	
+3,+HON,	+3,	

<u>kakhan</u> : 'When'	<u>keno</u> : 'Why'	<u>kothā</u> : 'Where'
+PRO,+INT,-COUNT,	+PRO,+INT,-COUNT,	+PRO,+INT,-COUNT,
+DEF,+SPECIFIC	-DEF,	+DEF,+SPECIFIC,
+TEMP, TIME,	+ABS,	+LOC,

<u>kichu</u> : 'Anything'	<u>kakhano</u> : 'Ever'	<u>kothāo</u> : 'Anywhere'
+PRO,-INT,-COUNT,	+PRO,-INT,-COUNT,	+PRO,-INT,-COUNT,
-DEF,-SPECIFIC,	-DEF,-SPECIFIC,	-DEF,-SPECIFIC,
-ANI,+ABS,	+TEMP, TIME,	+LOC,
-1,		
-2,		
+3,		

keu: 'Anyone'

+PRO, -INT, -COUNT,

-DEF, -SPECIFIC,

+HUM,

-1,

-2,

+3, ±HON

ekhān: 'Here'

+PRO, -INT, -COUNT,

+DEF, +LOC, -FAR,

okhān: 'There'

+PRO, -INT, -COUNT,

+DEF, +LOC,

+FAR, IN SIGHT,

sekhān: 'There'

+PRO, -INT, -COUNT,

+DEF, +LOC,

+FAR, OUT OF SIGHT

ekhan: 'Now'

+PRO, -INT, -COUNT,

+DEF, +TEMP,

-FAR,

takhan, okhan: 'Then'

+PRO, -INT, -COUNT,

+DEF, +TEMP,

+FAR,

e(i): 'This'

+DEIC,

+DEM,

-FAR,

+DEF,

o(i): 'That'

+DEIC,

+DEM,

+FAR, IN SIGHT,

+DEF,

se(i): 'That'

+DEIC,

+DEM,

+FAR, OUT OF SIGHT

+DEF,

kon: 'Which'

+DEIC,

+DEM,

+DEF,

+INT,

pratham: 'First'

+DEIC,

+ORD,

+DEF,

ek: 'One'

+QUANT, -DEF,

±SPECIFIC,

-PL,

du(i): 'Two'

+QUANT,
-DEF,+SPECIFIC,
+PL,

anek: 'Many'

+QUANT,
-DEF,+SPECIFIC,
+PL,

bahu: 'Many'

+QUANT,
-DEF,-SPECIFIC,
+PL,

sab, sakal: 'All'

+QUANT,
-DEF,-SPECIFIC,
+PL,

pratyek, prati: 'Each' ka(yek): 'Some'

+QUANT,
+DEF,
+PL,+DISTRIBUTIVE,

+QUANT,
-DEF,+SPECIFIC,
+PL,

tā

+CL,+COUNT,
-HON,

ti

+CL,+COUNT,
-HON,

jan

+CL,+COUNT,
+HUM,
+HON (PREFERABLY)

khānā

+CL,+COUNT,
-ANI,+EMOT,
+SMALL,

khāni

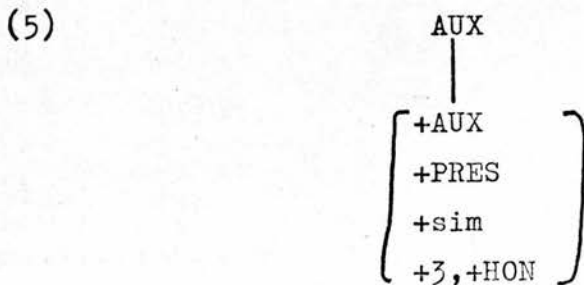
+CL,+COUNT,
-ANI,+EMOT,
+SMALL,

11.2 The Second Lexicon.

This work makes use of a second lexical insertion procedure which applies after the last rule of the transformational component. This second lexical insertion rule inserts lexical items to the sets of syntactic-semantic features which have been derived by the application of the transformational rules. This rule is generally used to insert transformationally derived pronouns, plural markers, forms of the AUX and the case markers. For example, the AUX is developed as a complex symbol of features by the base rules. No lexical item can be attached to the AUX in the underlying structure. A complex symbol for the AUX may have the structure (4) in the underlying structure.



No lexical item can be attached to this complex symbol because Bengali has no lexical item which has the features [+AUX,+PRES,+sim]. A later concord rule will apply to (4). This rule will copy the features of person and grade of the subject onto the AUX (cf., § 1.6). This concord rule will transform (4) into (5) if the subject has the features [+3,+HON].



After the last rule of the transformational component the second lexical insertion rule will attach the lexical item en to the node dominated by the AUX as en has the features [+AUX,+PRES,+sim,+3,+HON].

In this case we see that the second lexical insertion rule inserts a lexical item to a derived node dominating a complex symbol, where no lexical item could have been inserted in the underlying structure. Another function of the second lexical insertion rule is to insert those lexical items which could have been inserted by the first lexical insertion rule, but have not been inserted because of some peculiarities of those items. Case markers are such items. Case markers are not always realized in the surface. If we insert them in the underlying structure, we need to delete case markers of some actants at some stage of derivation. So we have used the second lexical insertion rule to introduce the case markers to the derived trees.

But the most important function of the second lexical insertion rule is to insert those lexical items which are derived by transformational operations. In such a case the second lexical insertion rule replaces a first lexical item by an appropriate item from the second lexicon. We face such situations after the pronominalization transformations such as pronominalization proper, reflexivization and relativization. For example, the pronominalization proper transformation derives an anaphoric pronoun by specifying an underlying noun phrase as [+PRO,-INT] under certain conditions. The pronominalization rule does not delete the first lexical item attached to the pronominalized

noun phrase. The function of the second lexical insertion rule is to replace that item by an appropriate item from the second lexicon. This rule inserts items by the matching condition of lexical insertion. But this rule does not require strict identity of features between the complex symbol under a derived node and the one associated with a second lexical item. The second lexical insertion rule is something like (6).

(6) The Second Lexical Insertion Rule.

A second lexical item will replace an item of a derived tree if the features associated with the second lexical item are a subset of features associated with the first lexical item in the derived tree.

The second lexical insertion rule does a quite simple job. Its function is to assign lexical items (phonological matrices) to the derived feature bundles. We have not given any power of deletion, addition or node-relabelling to this rule. But we feel that some operations which we perform by transformational rules could have been done by the second lexical insertion rule.

11.2.1 Sample Second Lexicon.

Derived Pronouns:

See § 11.1.5 for the complex symbols for the following pronouns:

āmi: 'I', āpani: 'You', tumi: 'You', tui: 'You', tini: 'He', se: 'He', tā: 'That (thing)', ekhān: 'Here', okhān: 'There', sekhān: 'There', and okhan, takhan: 'Then'.

The transformationally derived first and second person pronouns

are 'inclusive' pronouns, and so they are plural. The plural forms of these pronouns will be derived morphophonologically from the syntactic surface structure string. Other major second lexical items are listed below.

āpan, nij: 'Self'

+PRO,+REFL,
+COUNT,+DEF,

ye: 'Who'

+PRO,+REL,
+COUNT,+HUM,
+3,-HON,+DEF,

yini: 'Who'

+PRO,+REL,
+COUNT,
+3,+HON,+DEF,

yā: 'Which'

+PRO,+REL,-COUNT,
-ANI,+ABS,
+3,
+DEF

Conjunctions:

o, ār, eban: 'And'

+CONJCO,

ba, athabā: 'Or'

+CONJCO,

kintu: 'But'

+CONJOCO,

+S,

yadi: 'If'

+CONJSUB,
+S¹

tabe: 'Then'

+CONJSUB,
+S²

Plural Markers:

<u>rā</u>	<u>erā</u>	<u>der</u>	<u>gulo</u>
+PL MARKER,	+PL MARKER,	+PL MARKER,	+PL MARKER,
+HUM,	+HUM,	+HUM,	-HON,
+V__	+C__	-SUBJ	
+SUBJ	+SUBJ		

Case Markers:

<u>e</u>	<u>te</u>	<u>dvārā</u>	<u>diye</u>
+CAM,+BIB,	+CAM,+BIB,	+CAM,-BIB,	+CAM,-BIB,
-/i/___	+V__	+INS,	_____
_____	_____	+__AGT,	+INS,
+INS,-__AGT	+INS,-__AGT	+INS RIGID	+__AGT,
_____	_____	_____	_____
+LOC, { +STAT } { +GOAL }	+LOC, { +STAT } { +GOAL }		+LOC, +PATH,

<u>hote, theke</u>	<u>ke</u>	<u>re</u>
+CAM,-BIB,	+CAM,+BIB,	+CAM,+BIB,
+LOC,	{ +DAT } { +NEUT },	{ +DAT } { +NEUT },
+SOURCE,		+OBSOLETE

Forms of the AUX:

<u>i</u>	<u>o</u>	<u>en</u>	<u>is</u>	<u>e</u>
+AUX,	+AUX,	+AUX,	+AUX,	+AUX,
+PRES,	+PRES,	+PRES,	+PRES,	+PRES,
+sim,	+sim,	+sim,	+sim,	+sim,
+1,	+2,-HON,	{+2 +3}, +HON,	+2,+PEJ,	+3,-HON,
<u>bo</u>	<u>be</u>	<u>ben</u>	<u>bi</u>	<u>lām</u>
+AUX,	+AUX,	+AUX,	+AUX,	+AUX,
+FUT,	+FUT,	+FUT,	+FUT,	+PAS,
+sim,	+sim,	+sim,	+sim,	+sim,
+1,	{+2 +3}, -HON,	{+2 +3}, +HON,	+2,+PEJ,	+1,
<u>le</u>	<u>len</u>	<u>li</u>	<u>lo</u>	<u>tām</u>
+AUX,	+AUX,	+AUX,	+AUX,	+AUX,
+PAS,	+PAS,	+PAS,	+PAS,	+HAB P,
+sim,	+sim,	+sim,	+sim,	+sim,
+2,-HON,	{+2 +3}, +HON,	+2,+PEJ,	+3,-HON,	+1,
<u>te</u>	<u>ten</u>	<u>ti</u>	<u>tō</u>	<u>chi</u>
+AUX,	+AUX,	+AUX,	+AUX,	+AUX,
+HAB P,	+HAB P,	+HAB P,	+HAB P,	+PRES,
+sim,	+sim,	+sim,	+sim,	+prog,
+2,-HON,	{+2 +3}, +HON,	+2,+PEJ,	+3,-HON,	+1,

<u>cho</u>	<u>chen</u>	<u>chis</u>	<u>che</u>	<u>chilām</u>
+AUX,	+AUX,	+AUX,	+AUX,	+AUX,
+PRES,	+PRES,	+PRES,	+PRES,	+PAS,
+prog,	+prog,	+prog,	+prog,	+prog,
+2,-HON,	$\begin{Bmatrix} +2 \\ +3 \end{Bmatrix}$,+HON,	+2,+PEJ,	+3,-HON,	+1,

<u>chile</u>	<u>chilen</u>	<u>chili</u>	<u>chilo</u>	<u>echi</u>
+AUX,	+AUX,	+AUX,	+AUX,	+AUX,
+PAS,	+PAS,	+PAS,	+PAS,	+PRES,
+prog,	+prog,	+prog,	+prog,	+perf,
+2,-HON,	$\begin{Bmatrix} +2 \\ +3 \end{Bmatrix}$,+HON,	+2,+PEJ,	+3,-HON,	+1,

<u>echo</u>	<u>echen</u>	<u>echis</u>	<u>eche</u>	<u>echilām</u>
+AUX,	+AUX,	+AUX,	+AUX,	+AUX,
+PRES,	+PRES,	+PRES,	+PRES,	+PAS,
+perf,	+perf,	+perf,	+perf,	+perf,
+2,-HON,	$\begin{Bmatrix} +2 \\ +3 \end{Bmatrix}$,+HON,	+2,+PEJ,	+3,-HON,	+1,

<u>echile</u>	<u>echilen</u>	<u>echili</u>	<u>echilo</u>
+AUX,	+AUX,	+AUX,	+AUX,
+PAS,	+PAS,	+PAS,	+PAS,
+2,-HON,	$\begin{Bmatrix} +2 \\ +3 \end{Bmatrix}$,+HON,	+2,+PEJ,	+3,-HON,

11.3 The Morphophonological Component.

This grammar presupposes a morphophonological component. Its rules will apply to surface structure strings after second lexical insertion. These rules are used to derive acceptable phonetic forms of words. For example, no lexicon in this work contains the plural forms of pronouns, such as, āmarā: 'We', tomarā: 'You', tārā: 'They' etc., or other derived forms like āmāke: 'Me', tāke: 'Him', tomār: 'Your' etc. These forms are derived by the application of morphophonological rules on surface structure strings. We will come across surface structure strings like āmi-rā, āpani-rā, tumi-rā etc.,. So we will require morphophonological rules which will derive āmarā, āpanārā, tomarā from these strings. We can have further rules which will derive āmorā, āmrā, āpanārā, āpnārā, tomorā, tomrā from these strings. Although it is beyond our scope to formulate the morphophonological rules, we give below some rules informally which are frequently used in this grammar in order to arrive at acceptable phonetic forms of words.

(7) Some Morphophonological Rules:

$$a. \quad \bar{a}mi \quad \Rightarrow \quad \left\{ \begin{array}{l} \bar{a}ma/_ \neq r\bar{a} \\ \bar{a}m\bar{a}/_ \neq X \quad \text{SUFFIX} \end{array} \right\}$$

$$b. \quad tumi \quad \Rightarrow \quad \left\{ \begin{array}{l} toma/_ \neq r\bar{a} \\ tom\bar{a}/_ \neq X \quad \text{SUFFIX} \end{array} \right\}$$

$$c. \quad \left\{ \begin{array}{l} \bar{a}pani \\ \bar{a}pan \end{array} \right\} \Rightarrow \quad \bar{a}pan\bar{a}/_ \neq X \quad \text{SUFFIX}$$

- d. tui \Rightarrow to/___~~f~~X SUFFIX
- e. tini \Rightarrow t[~]ā/___~~f~~X SUFFIX
- f. se \Rightarrow tā/___~~f~~X SUFFIX
- g. ke \Rightarrow kā/___~~f~~X SUFFIX
- i. ye \Rightarrow yā/___~~f~~X SUFFIX
- j. yini \Rightarrow y[~]ā/___~~f~~X SUFFIX

The rule (7a) says that āmi must be rewritten as āma when the plural marker rā is suffixed to it; and it must be rewritten as āmā when any other suffix is attached to it. This rule will derive the phonetic forms given below:

- (8) āmi-rā \Rightarrow āmarā: 'We'
- āmi-ke \Rightarrow āmāke: 'Me'
- āmi-der \Rightarrow āmāder: 'Us, Our'
- āmi-r \Rightarrow āmār: 'My'

The rule (7b) will derive phonetic forms like those given below:

- (9) tumi-rā \Rightarrow tomarā: 'You (t+PL)'
- tumi-ke \Rightarrow tomāke: 'You, Objective'
- tumi-der \Rightarrow tomāder: 'You, Your'
- tumi-r \Rightarrow tomār: 'Your'

ON THE BENGALI LANGUAGEA.0 Introduction.

The language internationally known in English as 'Bengali' is called Bāṅlā by its native speakers. It also has several other antiquated names: e.g., Bāṅgālā, Bāṅālā and Baṅga-Bhāṣā: 'The language of Bengal'. Bengali is spoken in Bangladesh and in India in West Bengal, and some parts of Assam, Tripura, Orissa and Bihar. This is a densely populated area. At a rough estimate Bengali is spoken by over 120 million people, in several different dialects. In extreme cases these dialects vary so much morphophonologically that speakers of two dialects may find themselves mutually unintelligible. In such a situation the link between them has been Sādhu Bhāṣā: 'Chaste, Decent, Written Language' which until recently, as the medium used in Bengali literature, related the different dialects of Bengali one to another. Recently Sādhu Bhāṣā has been ousted from its previous position by Calit(a) or Calti Bhāṣā: 'Current Language, Standard Colloquial Bengali', which is now used predominantly in literature and as the language of communication among educated Bengalees. Thus Calit Bhāṣā has lately taken over the functions of Sādhu Bhāṣā in the Bengali speaking world.

Bengali is an Indo-Aryan language with much Dravidian influence. It emerged as a distinct language around the 10th century A.D., and is genetically related to Assamese, Oriya and Hindi (cf., Chatterji (1926), Shahidullah (1928, 1960, 1965), Pattanayak (1966)). Dr S. K. Chatterji (1926) divided the history of the Bengali language into three broad periods,

of which the middle one is again broken up into sub-periods.

Dr Chatterji's (1926, 129-36) divisions are as follows:

- (1) A. The Formative or Old Bengali Period: 950-1200.
- B. The Middle Bengali Period: 1200-1800.
 - a. Transitional Middle Bengali: 1200-1300.
 - b. Early Middle Bengali: 1300-1500.
 - c. Late Middle Bengali: 1500-1800.
- C. Modern or New Bengali: From 1800.

The above divisions are made on morphological, phonological and literary grounds, which are satisfactory in outline. The most important, and perhaps the only, surviving work of the old Bengali period is caryāpada, a collection of Buddhist mystic songs (cf., Shahidullah (1928, 1960), Mukherji (1963)). Like the old period, the middle period was also a period of verse literature. Prose was consciously developed in the early 19th century. It was in this period that Sādhu Bhāṣā was given a definite shape by various writers. Pramatha caudhuri (1868-1946) in his Sabuj Patra (1914) launched a movement against Sādhu Bhāṣā in literature, which has resulted in its abandonment in literature in recent times.

A.1 The Phonemes of Bengali, Orthography and Transliteration.

An intensive study of the phonemes of Bengali has been carried out by Ferguson and Chowdhury (1960) (cf., also Chatterji (1928)). According to Ferguson and Chowdhury (1960) Bengali has twenty eight consonants, four semivowels and fourteen vowels. Besides these, Bengali has seven marginal consonants and one marginal vowel. Of the vowels seven are oral and all of them have phonemically distinct nasal counterparts; and

this makes the vowel inventory in Bengali quite large. The following are the phonemes of Bengali (see Ferguson and Chowdhury(1960) for the features of the phonemes):

- (2) Oral Vowels: / i e æ a o o u /
 Nasal Vowels: / ãẽ ãã ã õ õ ã /
 Semivowels: / ɨ ɣ ɔ ʌ /
 Consonants: / p t ʈ c k b d ɖ
 j ɟ
 ph th ʈh ch kh bh dh ɖh
 jh gh
 s h
 m n ŋ
 l r ɽ /

Kostić and Das (1972) supported the findings of Ferguson and Chowdhury (1960) by carrying out experimental acoustic investigations of the phonemes of Bengali.

Bengali is written in a special script of the Debanāgarī type, which is believed to be a descendent of the ancient Indian Brāhmī Lipi: 'Brahmi Alphabet' (cf., Banerji (1919), Chatterji (1926), Lambert (1953)). The script is syllabic, and is written from left to right. The letters hang from a line called Mātrā, and use diacritic signs in all four directions. The letters are angular. All the vowels, except 'ɨ': a, have distinct graphemes for post-consonantal syllable formation. The Bengali alphabet includes a large number of conjunct letters. Many of them do not look like conjuncts any more. There are some simple letters in Bengali orthography which are seldom or never used.

There is no universally accepted convention of Romanization for Bengali in English (cf., Chatterji (1926), Mukherji (1963)). Transliteration, in this work, has been performed in accordance with the following table.

(3) Transliteration.

a : অ	ā : আ	i : ই	ī : ঈ	u : উ
ū : ঊ	ri : ঝ	e : এ	o : ও	
oi : ঔ	au : ঝ			
k : ক	kh : খ	g : গ	gh : ঘ	ṅ : ঙ, ঞ
c : চ	ch : ছ	j : জ	jh : ঝ	ñ : ঞ
ṭ : ট	ṭh : ঠ	ḍ : ড	ḍh : ঢ	ṇ : ণ
t : ত	th : থ	d : দ	dh : ধ	n : ন
p : প	ph : ফ	b : ব	bh : ভ	m : ম
y : য, ষ	r : র	l : ল	v : ব	ś : শ
ṣ : ষ	s : স	h : হ	ṛ : ঠ	ṛh : ঠ
kṣ : ক্ষ	jñ : ঞ	~ : ঔ		

A.2 Sādhu and Calit Bengali.

The greater part of Bengali literature is in Sādhu Bhāṣā. Sādhu Bengali assumed a well-defined shape in the 19th century with the birth and regular cultivation of Bengali prose. It seems that Sādhu Bengali was never used in oral communication, but it was the vehicle of Bengali literature until a few decades ago. Sādhu Bengali had been used in writing, sometimes in oratory, and in literature (including conversations in novels and

full-page dialogues in historical drama). Calit Bāṅlā developed out of the colloquial speech of Calcutta (cf., Chatterji (1926), and was taken up by educated speakers of Bengali as the language of oral communication. Although there is ample evidence that Calit Bengali was taking over Sādhu Bengali, Pramatha caudhurī launched a movement in his Sabuj Patra (1914) against the use of Sādhu Bengali in literature. It was a literary, but not a social movement. He was supported by some, not all talented writers of the time. Gradually Calit Bengali has become the language of Bengali literature. Calit Bengali is not universally used. We are not aware of the situation in West Bengal, but in Bangladesh Calit Bengali is used in literature, university and college lectures (if the teacher has a good command of Calit), scholarly lectures and in formal conversation among educated people. Not many Bangladeshi's use Calit Bengali in all spheres of life.

Sādhu and Calit are basically two different styles: the former is strictly a written language with a huge number of Sanskritic lexical items and expressions; the latter is both written and spoken with colloquial expressions and lexical items. In the days of a clear-cut division between these two styles it was possible to generalize that Sādhu Bengali was Sanskriticized and high-sounding, Calit Bengali the converse. But now it is difficult to draw a clear-cut distinction between these two styles. There is enough literary evidence of colloquial Sādhu Bengali, and highly Sanskriticized Calit Bengali. The main difference between Sādhu and Calit Bengali, now, is not stylistic, but lexical (cf., Dimock (1960)). In Sādhu Bengali certain lexical items are used in their archaic form,

but in Calit they are used in a derived form. Calit Bengali is passing through a period of rapid phonological change, so its phonological rules are more complex; but the derived forms are shorter (therefore simpler) than those of Sādhu Bengali. Their main differences are summarized below:

- (4) a. The verb forms (V) are shorter in Calit as the forms of the AUX have become shorter due to phonological change. For example, the form of the AUX which is itechī: [+AUX, +PRES, +prog, +1] in Sādhu Bengali has become chī in Calit Bengali. If the verb root is kar: 'Do', the verb form in Sādhu will be karitechī: '(I am) doing', and in Calit karchī.
- b. Forms of certain pronouns and deictics have become shorter in Calit due to phonological change.
- c. There are some plural markers which are used in Sādhu Bengali, but are not or rarely used in Calit Bengali.
- d. The non-initial /h/ sound is either reduced or dropped in Calit Bengali.
- e. Sādhu Bengali generally prefers Sanskritic lexical items and expressions, but Calit Bengali prefers indigenous and derived (mainly from Sanskrit) lexical items.

Those lexical items which are central in differentiating Sādhu Bengali from Calit Bengali are given below (Sādhu forms are at the left and Calit forms at the right of the arrows).

(5) Pronouns.

tāhārā	==>	tārā: 'They (nonhonorific)'
tā̃hārā	==>	tā̃rā: 'They (honorific)'
tāhā	==>	tā: 'That (thing)'
yāhā	==>	yā: 'Which (Relative)'

yāhārā	==>	yārā: 'Who (Relative, non-honorific)
yāhārā	==>	yārā: 'Who (Relative, honorific)
keha	==>	keu: 'Anyone, Someone' etc.;
kāhā	==>	kā: 'Who (interrogative, Objective)
kāhārā	==>	kārā: 'Who (interrogative, plural)

Sādhu Bengali has two abstract pronouns: ihā: 'This (thing); near', and uhā: 'That (thing); far, in sight'. They are not used in Calit Bengali. The third member of this group of abstract pronouns is tāhā, which survives in Calit Bengali as tā.

(6) Deictics:

ei	==>	e(i): 'This'
oi	==>	o(i): 'That, in sight'
sei	==>	se(i): 'That, out of sight'

(7) Plural Markers:

guli	==>	gulo
diga	==>	der

(8) A Few Forms of the AUX:

i	==>	i: [+PRES,+sim,+1]
itechī	==>	chi: [+PRES,+prog,+1]
iyāchi	==>	echi: [+PRES,+perf,+1]
itecha	==>	cho: [+PRES,+prog,+2,-HON]
itechen	==>	chen: [+PRES,+prog,{+2,+3},+HON]
...		
iyā	==>	e: 'Serial Infinitive Marker'
ite	==>	te: 'Infinitive Marker'
ile	==>	le: 'Conditional Infinitive Marker'

(9) Negative Markers:

nāi ==> nei: 'Existential Negative Marker'

nāi ==> ni: 'Perfective Negative Marker'

In copulative negative sentences in the present tense and simple aspect the 'be' verb ha is realized in the surface in Sādhu Bengali, but not in Calit Bengali (cf., § 1.5).

The present distinction between Sādhu and Calit Bengali is mainly in the selection of the above mentioned lexical items, which are shorter in Calit. Except for these superficial differences, Calit as used in contemporary Bengali literature is not far away from Sādhu Bengali, in its liberal use of Sanskritic words. Calit, of course, borrows lexical items widely from other languages, a procedure which Sādhu Bengali restricts. We have mentioned above that there is literary evidence of colloquial Sādhu Bengali and non-colloquial Calit Bengali. To give examples: Sudhīndranāth Datta's (1901-1960) prose is considered as Calit because it selects colloquial verb forms, forms of pronouns etc.,; the prose of Rabīndranāth Thākūr (1861-1941) in Jībansmṛiti is considered as Sādhu because it selects Sādhu verb forms, forms of pronouns etc.,. But there is no denying that stylistically Sudhīndranāth's prose is Sādhu and Rabīndranāth's prose is colloquial.

In prose literature a clear choice is made between Sādhu and Calit, but poetry allows a mixing up of both styles. So Jībanānanda Dās's (1899-1954) famous line bhālabese dekhiyāchi meyamānuṣere: '(I) loved and saw womenfolk' contains Calit and Sādhu verb forms side by side: in pure Sādhu the line would be bhālabāsiyā dekhiyāchi meyamānuṣere, and in pure Calit it would be bhālabese dekhechi meyamānuṣere. In both cases the

force and beauty of the line is lost. Sādhu Bengali is almost obsolete now: it is restricted to conservative news papers, text-books for schools and colleges and in Government circulars.

A.3 Works on Bengali.

Although Bengali is one of the major languages in South Asia, and perhaps the most distinguished in literary works, few linguistic studies have been done on Bengali. The enthusiastic introduction to the bibliography compiled by Čižikova and Ferguson (1969) gives an impression that Bengali has been studied extensively. This is in fact not the case. Very little substantial work has been done on Bengali. The most neglected fields of study are in the syntax and semantics of Bengali, which have not been seriously investigated in any framework. Bengali studies are usually carried out in Bangladesh, West Bengal, Russia, Britain and America. The work done on Bengali is classifiable into the following groups: (a) Historical Studies, (b) Descriptive Studies, (c) Dialect Studies, (d) Pedagogical Grammars, (e) Sketches, (f) Lexicography and (g) Others. Regrettably we could not include a class of works dealing with the syntax and semantics of Bengali. We do not wish to imply that nothing has been written in this field, but the few short works which has been done in this field do not merit a heading of their own. There are some short works dealing with the syntax of Bengali, such as Sableski's (1965) 'Equational Clauses in Bengali', Ferguson's (1972) 'Verbs of 'Being' in Bengali, with a Note on Amharic' and some other papers dealing with Bengali sentence types, negation etc., (cf., Čižikova and Ferguson (1969), Dil (1969)).

A. Historical Studies.

These works deal with the history and development of Bengali with special reference to its morphology and phonology. The studies devoted to old Bengali are included in this group. Among the historical studies of Bengali, Chatterji's (1926) The Origin and Development of the Bengali Language is a monumental work, both in size and merit. It deals with the origin of the Bengali language and its morphological and phonological change through different periods. It deals with lexical items individually, and shows their changes through time. Other works that deal with the development of Bengali or with various aspects of old Bengali are Mazumder (1920), Shahidullah (1928, 1960, 1965), Pattanayak (1966) and Mukherji (1963) (see Čižikova and Ferguson (1969) and the bibliographahy in Grierson (1903)).

B. Descriptive Studies.

These works deal mainly with the phonetics, phonology and morphology of Bengali in various contemporary frameworks. In this field too, Chatterji (1928) is the pioneer. He used Daniel Jonsean framework in his A Bengali Phonetic Reader (1928) in analysing the phonemes of Bengali. This work has been recently replaced by Ferguson and Chowdhury (1960) with the publication of their paper 'The Phonemes of Bengali'. Ferguson and Chowdhury (1960) deal with sounds of Bengali, using intensively the recent developments in phonological theory. Hai (1960) deals with nasals and nasalization in Bengali in a Firthian Framework. Hai (1964) later published his Dhvani-Bijñān o Bāṅlā Dhvanitattva: 'Phonetics and Phonology of Bengali' in Bengali.

It is an important work simply because it is written in Bengali. Other noteworthy works on the phonetics, phonology and morphology of Bengali are Ferguson (1945), Hai (1958), Hai and Ball (1961), Din Muhammad (1961) and Kostić and Das (1972).

C. Dialect Studies.

These works deal with various dialects of Bengali with special reference to their phonetics and phonology in relation to those of Calit Bengali. Among the dialect studies the most important one is Grierson (1903). It exemplifies some major dialects of Bengali without much linguistic description. Other works which can be mentioned are Chowdhury (1960), Shahidullah (ed., 1964), Hai (1964, 1965, 1966), Sen (1972) and Ućida (1972).

D. Pedagogical Grammars.

A work of this sort is traditionally known as Byākarana: 'Grammar'. These grew mainly out of the necessity to teach students 'how to read, write and speak Bengali grammatically'. These works deal mainly with Sādhu Bengali. Early works of this type are Halhed (1778) and Ray (1833) (cf., Qayyum (1974) for a review of the grammars of Carey, Halhed and Haughton). These grammars are written by people who are familiar with Sanskrit and English grammatical traditions, and so Bengali looks sometimes like Sanskrit and sometimes like English. This is the type of work which led Rabīndranāth Thākur to comment that 'We read Sanskrit grammars in the guise of Bengali grammars, which are flavoured with a little bit of Bengali.' A typical and representative work of this type is Chatterji's (1939) Bhāṣā-Prakāś Bāṅgālā Byākarana, which deals with morphology,

phonology, cases and other aspects of Bengali in more than five hundred pages. However it deals with syntax in only a few pages (427-42). As in any other pedagogical grammars, the prescriptive nature of Chatterji (1939) is apparent. It seldom explains the language. One feature common to all such works is that they have almost identical topics to deal with, and they deal with those topics almost identically, in almost identical language with similar types of examples and interpretations. Although these grammars are intended to teach students how to 'read, write and speak' Bengali grammatically, most of them are self-contradictory even in the first page, for they deal with Sādhu Bengali, which is never spoken. Some other grammar books of this type are Sen (1950) and Ghos (1956) (cf., Čižikova and Ferguson (1969)).

E. Sketches.

This type of work is written for foreigners who are interested to learn Bengali as quickly as possible. In this group are Page (1934), Chaudhury (1963), Hudson (1965), Ray et al (1966) and Islam (1970). The authors of this type of work use various methods for teaching Bengali to would be speakers. These usually deal with phonetics, sentence types, words and expressions used in every-day life and Bengali alphabet. These usually include a sketchy 'reference grammar' of Bengali. Some of them are quite pleasant books. Some, nevertheless, use hocus-pocus techniques in dealing with morphology and syntax (cf., Ray et al (1966)). This may not be much help to anyone trying to learn Bengali.

F. Lexicography.

This type of work includes functional mono-lingual and bilingual dictionaries. But there are some noteworthy works, such as, Dās¹ (1937), which is a valuable etymological lexicon, and Shahidullah (ed., 1964), which is a lexicon of the dialects of Bangladesh. Among the mono-lingual dictionaries Bisvās (1955) and Basu (1962) are helpful.

G. Others.

This group includes those works which deal with various aspects of Bengali, and works which deal with work done on Bengali. Such works are Chatterji (1962), which deals with the relationship between the written and colloquial Bengali; Khondkar (1971), which deals with the Portuguese contribution to Bengali lexicography, and Qayyum (1973), which reviews the Bengali grammars of Carey, Halhed and Haughton.

1 Dās, Jñānendramohan (1937): Bāṅgālā Bhāṣār Abhidhān.

2 Vols. Calcutta.

ABBREVIATIONS

The following abbreviations have been used frequently in this work. Although explained in the text, they are given here for ready reference. Those abbreviations, used seldom, should be easily checkable in the relevant sections.

ABS: Abstract	DEF: Definite
ADJ: Adjective	DEIC: Deictic
ADV: Adverb	DEM: Demonstrative
AGT: Agent	DO: Direct Object
ANI: Animate	EMOT: Emotive
AUX: Auxiliary	ESS: Essive
BIB: Bibhakti	FACT: Factive
C: Consonant	FUT: Future
CAM: Case Marker	GEN: Genitive
CL: Classifier	HAB P: Habitual Past
COM: Common	HON: Honorific
COMP: Complementizer	HUM: Human
CONJCO: Coordinate Conjunction	IMP: Imperative
CONJSUB: Subordinate Conjunction	INF: Infinitive
CONS: Constituent	INT: Interrogative
COREF: Coreferential	INS: Instrumental
CS: Complex symbol	IO: Indirect Object
D: Determiner	LOC: Locative
DAT: Dative	M: Modality
DD: Direct Discourse	N: Noun

NEG: Negative	SUBJ: Subject
NEUT: Neutral	SUBM: Subject Marker
NP: Noun Phrase	TEMP: Temporal
NP ^a : Antecedent Noun Phrase	V: Verb Form, Vowel
NP ^p : Pronominalizable or Pronominalized Noun Phrase	VB: Verb Root
NP-Rel: Relativizable or Relativized Noun Phrase	1: First Person
	2: Second Person
	3: Third Person
OBJ: Object	
ORD: Ordinal	
PART: Partitive	Expressions like (1a, b)
PAS: Past	and (1, 2a) have been used
PEJ: Pejorative	in this work. They should be
PL: Plural	interpreted as follows:
PL SEG: Plural Segment	(1a, b) = (1a) and (1b)
perf: Perfective	(1, 2a) = (1a) and (2a)
PRES: Present	
PRO: Pronoun	
prog: Progressive	
Q: Question	
QUANT: Quantifier	
REFL: Reflexive	
REL: Relative	
REP: Repetitive	
S: Sentence	
SC: Structure Change	
SI: Structure Index	
sim: Simple	
SPEC: Specifier	

BIBLIOGRAPHY

The bibliography is a list of items either referred to in the body of this work, or directly helpful to the writing of this dissertation. I have not had direct access to those items indicated by *, but they have been included for information purposes.

Abbreviations:

- FL: Foundations of Language. International Journal of Language and Philosophy. Dordrecht, Holland.
- IJAL: International Journal of American Linguistics. Baltimore
- IL: Indian Linguistics. Journal of the Linguistic Society of India. Poona.
- JL: Journal of Linguistics. The Journal of the Linguistics Association of Great Britain. London and New York.
- Lg: Language. Journal of the Linguistic Society of America. Baltimore.
- LI: Linguistic Inquiry. The MIT Press, Cambridge, Massachusetts.

A. Theoretical:

Akmajian, Adrian and Chisato Kitagawa (1976)

'Deep Structure Binding of Pronouns and Anaphoric Bleeding'. Lg, Vol 52, No 1, 61-77.

Anderson, John M. (1971)

The Grammar of Case. Cambridge University Press, London.

Bach, Emmon (1968)

'Nouns and Noun Phrases'. In Bach and Harms (1968).

_____ and R. T. Harms (eds., 1968)

Universals in Linguistic Theory. Holt, Rinehart and
Winston Ltd. New York.

_____ (1970)

'Problominalization'. LI, Vol 1, No 1, 121-22.

_____ (1971)

'Syntax since Aspects'. In O'Harrien (1971).

Bellert, Irena (1966)

'On Certain Syntactical Properties of the English
Connectives "and" and "but". In Transformational and
Discourse Analysis Papers, Zellig Harris, University
of Pennsylvania, Philadelphia.

Bierwisch, Manfred and Karl Eric Heidolph (eds., 1970)

Progress in Linguistics. Mouton, The Hague.

Botha, Rudolf P. (1968)

The Function of the Lexicon in Transformational
Generative Grammar. Mouton, The Hague.

Brown, Edward Keith (1972)

Towards a Case Grammar of Twi. Unpublished PhD
dissertation, University of Edinburgh.

Browne, Wayles E. (1970)

'Noun Phrase Determiners in Relatives and Questions:
Evidence from Macedonian'. LI, Vol 1, No 2, 267-70.

Chafe, Wallace L. (1970)

Meaning and the Structure of Language. Chicago.

Chomsky, Noam A. (1965)

Aspects of the Theory of Syntax. The MIT Press.

_____ (1970a)

'Remarks on Nominalization'. In Jacobs and Rosenbaum (1970); also in Chomsky (1972).

_____ (1970b)

'Deep Structure, Surface Structure, and Semantic Interpretation'. In Steinberg and Jakobovits (1971); also in Chomsky (1972).

_____ (1972)

Studies on Semantics in Generative Grammar. Mouton, The Hague.

Cole, Peter (1974)

'Indefiniteness and Anaphoricity'. Lg, Vol 50, 665-74.

Curme, George O. (1931)

Syntax: A Grammar of the English Language. Vol III.

D. C. Heath and Company, Boston.

Cushing, Steven (1972)

'The Semantics of Sentence Pronominalization'. FL, Vol 9, 162-168.

Dougherty, Ray C. (1969)

'An Interpretive Theory of Pronominal Reference'. FL, Vol 5, 488-519.

_____ (1970)

'A Grammar of Coordinate Conjunction, Part I'. Lg, Vol 46, 850-98.

_____ (1971)

'A Grammar of Coordinate Conjunction, Part II'. Lg, Vol 47, 298-339.

_____ (1974)

'The Syntax and Semantics of Each Other Constructions'.

FL, Vol 12, 1-47.

Essien, Okon E. A. (1974)

Pronominalization in Efik. Unpublished PhD dissertation,

University of Edinburgh.

Fiengo, Robert and Howard Lasnik (1972)

'The Logical Structure of Reciprocal Sentences in

English'. FL, Vol 9, 447-68.

Fillmore, Charles J. (1966a)

'Deictic Categories in the Semantics of "Come"'. FL,

Vol 2, 219-27.

_____ (1966b)

'Toward a Modern Theory of Case'. In Reibel and

Schane (1969).

_____ (1968a)

'The Case for Case'. In Bach and Harms (1968).

_____ (1968b)

'Lexical Entries for Verbs'. FL, Vol 4, 373-93.

_____ (1969)

'Types of Lexical Information'. In Kiefer (1969);

also in Steinberg and Jakobovits (1971).

_____ (1970)

'The Grammar of Hitting and Breaking'. In Jacobs and

Rosenbaum (1970).

_____ (1971)

'Some Problems for Case Grammar'. In O'Brien (1971)

_____ and D. T. Langendoen (eds., 1971)

Studies in Linguistic Semantics. Holt, Rinehart and

Winston, New York.

Fodor, Jerry A. (1970)

'Three Reasons for Not Deriving "Kill" from "Cause to Die"'. LI, Vol 1, No 4, 429-38.

_____ and Jerrold J. Katz (eds., 1964)

The Structure of Language: Readings in the Philosophy of Language. Prentice-Hall Inc., New Jersey.

Garvin, Paul L. (ed., 1970)

Method and Theory in Linguistics. Mouton, The Hague.

Gleitman, Lila R. (1965)

'Coordinating Conjunctions in English'. In Reibel and Schane (1969).

Greenberg, Joseph H. (ed., 1963a)

Universals of Language. The MIT Press, Cambridge, Mass., (1963b)

'Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements'. In Greenberg (1963a).

_____ (1966)

Language Universals. Mouton, The Hague.

Grinder, John (1971)

'Chains of Coreference'. LI, Vol 2, No 2, 183-202.

_____ and Paul M. Postal (1971)

'Missing Antecedents'. LI, Vol 2, No 3, 269-312.

Gruber, Jeffrey S. (1965)

Studies in Lexical Relations. Indiana University Linguistics Club.

_____ (1967)

Functions of the Lexicon in Formal Descriptive Grammars. Santa Monica, California.

Hagège, Claude (1976)

'Relative Clause, Centre-Embedding, and Comprehensibility'.

LI, Vol 7, No 1, 198-201.

Halliday, M. A. K. (1970)

'Language Structure and Language Function'. In Lyons (1970a).

Hinds, John and Nobuo Okada (1975)

'Pronominalization Across Coordinate Structures'.

LI, Vol 6, No 2, 330-35.

Huddleston, Rodney D. (1970)

'Some Remarks on Case Grammar'. LI, Vol 1, No 4, 501-11.

Hudson, R. A. (1976)

'Lexical Insertion in a Transformational Grammar'.

FL, Vol 14, 87-107.

Jackendoff, Ray C. (1972)

Semantic Interpretation in Generative Grammar. The MIT Press, Cambridge, Mass.

Jacobs, Roderick A. and Peter S. Rosenbaum (1968)

English Transformational Grammar. Ginn and Company Ltd.

_____(eds., 1970)

Readings in English Transformational Grammar. Ginn and Company, Waltham, Mass.

Jespersen, Otto (1927)

A Modern English Grammar. Part III, Syntax, Second Vol. Heidelberg.

Katz, Jerrold J. (1967)

'Recent Issues in Semantic Theory'. FL, Vol 3, 124-94.

_____(1971)

'Generative Semantics in^s Interpretive Semantics'.

LI, Vol 2, No 3, 313-31.

_____ and Jerry A. Fodor (1963)

'The Structure of a Semantic Theory'. In Fodor and Katz (1964).

_____ and Paul M. Postal (1964)

An Integrated Theory of Linguistic Descriptions.

The MIT Press, Cambridge, Mass.

Karttunen, Lauri (1971)

'Definite Descriptions with Crossing Coreference'.

FL, Vol 7, 157-82.

Keenan, Edward L. and B. Comrie (1972)

'Noun Phrase Accessibility and Universal Grammar'.

Unpublished Paper, Cambridge University.

Kiefer, Ference (ed., 1969)

Studies in Syntax and Semantics. FL Supplementary Series.

_____ (1970)

'On the Problem of Word Order'. In Bierwisch and

Heidolph (1970).

Kiparsky, Paul and Carol Kiparsky (1971)

'Fact'. In Steinberg and Jakobovits (1971).

Klima, Edward S. (1964a)

'Relatedness Between Grammatical Systems'.

In Reible and Schane (1969).

_____ (1964b)

'Negation in English'. In Fodor and Katz (1964).

Koutsoudas, Andreas (1968)

'On the Wh-words in English'. JL, Vol 4, No 2, 267-73.

_____ (1971)

'Gapping, Conjunction Reduction, and Coordination

Deletion'. FL, Vol 7, 337-86.

Kuno, Susumu (1971)

'The Position of Locatives in Existential Sentences'.

LI, Vol 3, No 3, 333-78.

_____ (1972)

'Pronominalization, Reflexivization, and Direct

Discourse'. LI, Vol 3, No 2, 161-95.

_____ (1974)

'The Position of Relative Clauses and Conjunctions'.

LI, Vol 5, No 1, 117-36.

Kuroda, S. -Y. (1968).

'English Relativization and Certain Related Problems'.

In Reibel and Schane (1969).

_____ (1969)

'Remarks on Selectional Restrictions and Presuppositions'.

In Kiefer (1969).

Lakoff, George (1968)

'Pronouns and Reference'. Indiana University Linguistic
Club.

_____ (1970a)

Irregularity in Syntax. Holt, Rinehart and Winston.

_____ (1970b)

'Pronominalization, Negation, and the Analysis of Adverbs'.

In Jacobs and Rosenbaum (1970).

_____ (1971a)

'Generative Semantics'. In Steinberg and Jakobovits (1971).

_____ (1971b)

'Presupposition and Relative Well-Formedness'.

In Steinberg and Jakobovits (1971).

- _____ and Stanley Peters (1966)
 'Phrasal Conjunction and Symmetrical Predicates'.
 In Reibel and Schane (1969).
- Langacker, Ronald W. (1969)
 'On Pronominalization and the Chain of Command'.
 In Reibel and Schane (1969).
- Lees, Robert B. (1970)
 'On Very Deep Grammatical Structure'. In Jacobs and
 Rosenbaum (1970).
- _____ and Edward S. Klima (1963)
 'Rules For English Pronominalization'. In Reibel and
 Schane (1969).
- Lyons, John (1966)
 'Towards a "Notional" Theory of the "Parts of Speech"'.
 JL, Vol 2, 209-36.
- _____ (1968)
Introduction to Theoretical Linguistics. Cambridge
 University Press.
- _____ (ed., 1970a)
New Horizons in Linguistics. Penguin Books.
- _____ (1970b)
 'Generative Syntax'. In Lyons (1970a).
- _____ (1970c)
Chomsky. Fontana Modern Masters Series.
- Masica, Colin (1972)
 'Relative Clause in South Asia'. In Peranteau et al (1972).
- McCawley, James D. (1968a)
 'Concerning the Base Component of a Transformational
 Grammar'. FL, Vol 4, 243-69.

_____ (1968b)

'The Role of Semantics in a Grammar'. In Bach and Harms (1968).

_____ (1970)

'Where Do Noun Phrases Come From'. In Jacobs and Rosenbaum (1970); also in Steinberg and Jakobovits (1971).

_____ (1971)

'Prelexical Syntax'. In O'Brien (1971).

Moyne, J. A. (1971)

'Reflexive and Emphatic'. In Lg, Vol 47, 141-63.

O'Brien, Richard J. (ed., 1971)

Report of the Twenty-Second Annual Round Table Meeting on Linguistics and Language Studies. Georgetown University, Washington D. C.

Perlmutter, David (1971)

Deep and Surface Structure Constraints in Syntax. Holt, Rinehart and Winston, New York.

_____ and John Robert Ross (1970)

'Relative Clauses with Split Antecedents'. LI, Vol 1, No 3, 350.

Peranteau, P. M., J. N. Levi and G. C. Phres (eds., 1972)

The Chicago Which Hunt. Chicago Linguistic Society.

Postal, Paul M. (1967)

'Crazy Notes on Restrictive Relative Clauses and Other Matters'. Unpublished Paper.

_____ (1970a)

'On So-Called "Pronouns" in English'. In Jacobs and Rosenbaum (1970); also in Reibel and Schane (1969).

- _____ (1970b)
 'The Method of Universal Grammar'. In Garvin (1970).
- _____ (1970c)
 'On Coreferential Complement Subject Deletion'. LI,
 Vol 1, No 4, 439-500.
- _____ (1971a)
Cross-Over Phenomena. Holt, Rinehart and Winston, New York.
- _____ (1971b)
 'On the Surface Verb "Remind"'. In Fillmore and
 Langendoen (1971).
- _____ (1972)
 'A Global Constraint on Pronominalization'. LI, Vol 3,
 No 1, 35-59.
- _____ (1974)
On Raising. The MIT Press, Cambridge, Massachusetts.
- Reibel, David A. and Sanford A. Schane (eds., 1969)
Modern Studies in English: Readings in Transformational
 Grammar. Prentice-Hall Inc., New Jersey.
- Rosenbaum, Peter S. (1967)
The Grammar of English Predicate Complement Constructions.
 The MIT Press, Cambridge, Mass.
- Ross, John Robert (1966)
 'A Proposed Rule of Tree-Pruning'. In Reibel and
 Schane (1969).
- _____ (1967a)
Constraints on Variables in Syntax. Mimeographed.
- _____ (1967b)
 'On the Cyclic Nature of English Pronominalization'.
 In Reibel and Schane (1969).

_____ (1970)

'Gapping and the Order of Constituents'. In Bierwisch and Heidolph (1970).

Schacter, Paul (1973)

'Focus and Relativization'. Lg, Vol 49, No 1, 19-46.

Schane, Sanford A. (1966)

A Schema for Sentence Coordination. Mitre Corp., Bedford.

Shibatani, Masayoshi (1973)

'Semantics of Japanese Causativization'. Fl, Vol 9, 327-73.

Smith, Carlota S. (1964)

'Determiners and Relative Clauses in a Generative Grammar'. In Reibel and Schane (1969).

_____ (1969)

'Ambiguous Sentences with "And"'. In Reibel and Schane (1969).

Starosta, Stanley (1974)

'A Place for Case'. Work in Progress, No 7, Department of Linguistics, University of Edinburgh.

Steinberg, Danny D. and Leon A. Jakobovits (eds., 1971)

Semantics. Cambridge University Press.

Stockwell, Robert P., Paul Schacter and Barbara Hall Partee (1973)

The Major Syntactic Structures of English. Holt, Rinehart and Winston, Inc.

Thompson, Sandra Annear (1971)

'The Deep Structure of Relative Clauses'. In Fillmore and Langendoen (1971).

Vendler, Zeno (1967)

Linguistics in Philosophy. Cornell University Press, Ithaca.

Wasow, Thomas (1975)

'Anaphoric Pronouns and Bound Variables'. Lg, Vol 51,
No 2, 368-83.

Weinreich, Uriel (1966)

Explorations in Semantic Theory. Mouton, The Hague;
also in Current Trends in Linguistics, Vol 3, Edited by
T. A. Sebeok, and in Steinberg and Jakobovits (1971).

Wierzbicka, Anna (1967)

'Against "Conjunction Reduction"'. Unpublished Paper.

B. On Bengali:

Banerji, R. D. (1919)

The Origin of the Bengali Script. University of Calcutta.

Basu, Rājśekhara (1962)

Calantikā. Calcutta.

Bisvās, Śailendra (1955)

Saṅsad Bāṅgālā Abhidhān. Sāhitya Saṅsad, Calcutta.

Chatterji, Suhas (1962)

*A Study of the Relationship Between Literary and
Colloquial Bengali. Unpublished PhD dissertation,
University of Chicago.

Chatterji, Suniti Kumar (1926)

The Origin and Development of the Bengali Language.
Part I: Introduction and Phonology; Part II: Morphology.
University of Calcutta.

_____ (1928)

A Bengali Phonetic Reader. University of London Press.

_____ (1939)

Bhāsā-Prakāś Bāṅgālā Byākaraṇ. University of Calcutta.

_____ (1972)

Saral Bhāsā-Prakāś Bāṅlā Byākaraṇ. Bāk Sāhitya Private Ltd., Calcutta.

Chaudhury, Mufazzal Haider (1963)

Colloquial Bengali. Bengali Academy, Dacca.

Chowdhury, Munier (1960)

'The Language Problem in East Pakistan'. IJAL, Part III, Vol 26, No 3, 64-78.

Čižikova, Ksenija L. and Charles A. Ferguson (1969)

'Bibliographical Review of Bengali Studies.'
In Sebeok (1969).

Dil, Anwar S. (1969)

'Linguistic Studies in Pakistan'. In Sebeok (1969).

Dimock, Edward C. (1960)

'Literary and Colloquial Bengali in Modern Bengali Prose'. IJAL, Part III, Vol 26, No 3, 43-63.

Din Muhammad, Quazi (1961)

The Phonology of the Verbal Piece in Colloquial Bengali.
Unpublished PhD dissertation, School of Oriental and African Studies.

Ferguson, Charles A. (1945)

*Phonology and Morphology of Standard Colloquial Bengali.
Unpublished PhD dissertation, University of Pennsylvania.

_____ (1972)

'Verbs of "Being" in Bengali, with a Note on Amharic'.
In The Verb 'BE' and Its Synonyms, Part 5, FL Supplementary Series, Edited by John W. M. Verhaar. Dordrecht, Holland.

_____ and Munier Chowdhury (1960)

'The Phonemes of Bengali'. Lg, Vol 36, No 1, 22-59.

Ghoṣ, Śrī Jagadīscandra (1956)

Ādhunik Bāṅlā Byākaraṇ. Calcutta.

Grierson, G. A. (ed., 1903)

Linguistic Survey of India. Vol 5, Indo-Aryan Family,
Eastern Group, Specimens of the Bengali and Assamese
Languages, Part I. Calcutta.

Hai, Muhammad Abdul (1958)

'Aspiration in Standard Bengali'. IL, Turner Jubilee
Vol, Vol 1, 142-46.

_____ (1960)

A Phonetic and Phonological Study of Nasals and
Nasalization in Bengali. University of Dacca.

_____ (1964)

Dhvani-Bijñān o Bāṅlā Dhvanitattva. Bengali Academy,
Dacca.

_____ (1966)

'Dacca Dialect'. In Ray et al (1966).

_____ and W. J. Ball (1961)

The Sound Structures of English and Bengali. Dacca
University.

Halhed, Nathaniel Brassey (1778)

*Grammar of the Bengali Language.

Hudson, D. F. (1965)

Bengali: An Introductory Course for Students and
Beginners. The English Universities Press, London.

Islam, Rafiqul (1970)

An Introduction to Colloquial Bengali. Dacca.

Khondkar, M. A. Rahim (1971)

The Portuguese Contribution to Bengali Prose, Grammar and Lexicography. Unpublished PhD dissertation, School of Oriental and African Studies.

Kostić, Djordje and Rhea S. Das (1972)

A Short Outline of Bengali Phonetics. Calcutta.

Lambert, H. M. (1953)

Introduction to the Devanagari Script. Enlarged Vol. Oxford University Press.

Mazumder, Bijaychandra (1920)

The History of the Bengali Language. University of Calcutta.

Mukherji, Tarapada (1963)

The Old Bengali Language and Text. University of Calcutta.

Page, Sutton W. (1934)

An Introduction to Colloquial Bengali. Cambridge.

Pattanayak, Debi Prasanna (1966)

A Controlled Historical Reconstruction of Oriya, Assamese, Bengali and Hindi. Mouton, The Hague.

Qayyum, M. Abdul (1974)

A Critical Study of the Bengali Grammars of Carey, Halhed and Haughton. Unpublished PhD dissertation, School of Oriental and African Studies.

Ray, Punya Sloka, M. A. Hai and Lila Ray (1966)

Bengali Language Handbook. Washington D. C.

Rāy, Rāmmohan (1833)

*Gaurīya Byākaraṇ.

Sableski, Julia A. (1965)

'Equational Clauses in Bengali'. Lg, Vol 41, No 3, 439-46.

Sebeok, Thomas A. (ed., 1969)

Current Trends in Linguistics. Vol 5, Linguistics in
South Asia. Mouton, The Hague.

Sen, Nilmadhav (1972)

'Some Dialects of Bangladesh - An Outline'. IL, Vol 33,
143-52.

Sen, Sukumar (1950)

*Bāṅlā Bhāṣā Byākaraṇ. Calcutta.

_____ (1958)

'The Position of the Negative Particle in Bengali'. IL,
Turner Jubilee Vol, Vol 1, 100-02.

Shahidullah, Muhammad (1928)

*Les Chants Mystiques de Kanna et de Saraha. Paris.

_____ (1960)

Buddhist Mystic Songs. Karachi University.

_____ (ed., 1964)

Pūrba Pākistānī Añcalik Bhāṣār Abhidhān. Bengali Academy,
Dacca.

_____ (1965)

Bāṅlā Bhāṣār Itibritta. Bengali Academy, Dacca.

Učida, Norihiko (1970)

*Der Bengali-Dialekt Von Chittagong: Grammatik, Texte,
Wörterbuch. Wiesbaden, Otto Harrassowitz.