SRI International

Technical Note 493 • July 1990

SYNTACTIC CONSTRAINTS ON ANAPHORIC BINDING

Prepared by:

MARY E. DALRYMPLE
Computer Scientist
Artificial Intelligence Center
Computing and Engineering Sciences Division
and
Center for the Study of Language and Information
Stanford University

A dissertation submitted to the Department of Linguistics and the Committee on Graduate Studies of Stanford University in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

333 Ravenswood Avenue • Menlo Park, CA 94025-3493 • (415) 326-6200 • FAX: (415) 326-5512 • Telex: 334486

To my sister Martha,
'the wisest woman in the world',
in gratitude for her love and support.

SYNTACTIC CONSTRAINTS ON ANAPHORIC BINDING

Mary Dalrymple, Ph.D.
Stanford University, 1990

An important goal in the study of anaphoric binding, as in most syntactic research, is the factoring apart of constraints that are universal, those that are language-specific, and those that are associated with individual lexical items. Some previous work on anaphoric binding has assumed that the binding conditions of reflexives, reciprocals, and pronominals should be thought of as invariant across languages, or as a property or parameter set for each language.

For languages with a relatively simple anaphoric system, one or the other of these positions might seem to be adequate. However, if a language has, for example, two or more reflexives, each obeying different binding constraints, it becomes clear that binding constraints must be associated with the particular lexical item for which they are applicable. The position taken in this work is that each anaphoric element is lexically associated with some set of binding constraints drawn from a universally-available inventory.

These universally-available binding constraints are characterizable in terms of one or more of a set of three grammatical concepts: subject, predicate, and tense. Constraints on the possible grammatical function of the antecedent are stated in terms of coreference with or disjointness from a subject. The domain in which the antecedent may or may not be found is constrained relative to either the syntactic predicate of which the anaphoric element is an argument, the minimal domain containing the anaphoric element and a subject, or the minimal tensed domain containing the anaphoric element. Combinations of constraints stated with reference to these three grammatical concepts form the full range of constraints which anaphoric elements can obey.

A precise formal encoding of the binding constraints enables a clear statement of their properties and a characterization of the interactions between them. The theory of Lexical-Functional Grammar provides a rich formal vocabulary for stating anaphoric binding constraints; in particular, *functional uncertainty* allows a statement

of the syntactic relation between an anaphor and the set of its possible or disallowed antecedents. Binding equations involving functional uncertainty are lexically associated with each anaphoric element, specifying the set of its possible antecedents or the set of elements with which it may not corefer.

Acknowledgements

I owe a number of people a debt of gratitude for helping me bring this dissertation to completion. Joan Bresnan was one of the foremost reasons I wanted to come to Stanford, and she has been an inspiring influence since my arrival here. She provided invaluable advice and guidance throughout the writing of this dissertation; no one could have asked for a better dissertation chair.

Annie Zaenen has been a great colleague, collaborator, and friend. She went above and beyond the call of duty in reading and commenting on the dissertation; because of her generous and thoughtful comments, it's a much better work than it would have been otherwise.

Ivan Sag kept me honest, always making sure I had good reasons for doing things the way I did. I was lucky to have had Ivan on my committee; his input has helped to make this a stronger work.

Paul Kiparsky brought a refreshing perspective to this work. Although he may not always have agreed with what I wanted to say, he was always willing to listen and provide criticism and encouragement.

My years as a student at Stanford have been a dream come true. Graduate school here has been every bit as wonderful as I thought it would be; nowhere else is so much groundbreaking linguistic work being conducted, and I'm proud to have been able to be a student here. I've had the opportunity to work with a number of people during the past few years: in particular, I've learned a lot about computation from Lauri Karttunen and Terry Winograd and a lot about semantics from Stanley Peters.

Working at SRI has been a tremendously interesting experience — almost like going to graduate school twice. I've learned as much from my colleagues at SRI as

from my classes at Stanford. I'm especially grateful to Phil Cohen for hiring me in the first place and for being a stimulating influence. If not for him, my experience at Stanford wouldn't have been as rich.

Fernando Pereira has been an inspiration and a privilege to work with. He has had more influence on my thinking than he is probably aware of; I'm fortunate to have had the chance to work with him.

Stuart Shieber has been a wonderful colleague and a friend through thick and thin. He deserves a special vote of thanks for reading this dissertation as carefully and critically as my committee members and providing insightful criticism and comments. I'm grateful to him for his help and support.

SRI wouldn't be the same without my swimming partners Martha Pollack and Sharon Oviatt, or without the after-swim lunches we've enjoyed together.

I was fortunate to have been given office space at Xerox-PARC during my third year of graduate school; PARC was a very exciting environment to work in. Kris Halvorsen has been more than generous with his time and facilities. Ron Kaplan was a great teacher and a pleasure to talk to; I've learned a lot from Ron. John Maxwell was always ready to patiently answer my questions, no matter how trivial.

The 1983 Linguistic Institute in Los Angeles was a tremendous influence in making me aware of the riches that the field of linguistics has to offer. I'm particularly grateful to Gerald Gazdar and Geoff Pullum for making the GPSG class so memorably good.

The University of Texas was the place where I first learned what linguistics is — and what a revelation that was! My first syntax teachers there were Lee Baker, Robert Bley-Vroman, Sue Schmerling, and Tony Woodbury; their classes were always interesting and thought-provoking. I had my first exposure to LFG at Texas, in a class taught by Mohanan when he was there — it seems to have taken. Herman Van Olphen taught me all the Hindi I know, and encouraged me to present my first paper at the South Asian Languages Analysis meeting there. I'll always appreciate the good times I had at Texas with Keith Walters, Yuen-Mei Yin, and Deirdre Mendez.

My family has been an invaluable source of support throughout my years in graduate school. It's always been nice knowing my sister was only a phone call away; she provides an oasis of sanity in the midst of a sometimes not-so-sane world. My parents'

faith in me was always reassuring; their love and much-needed encouragement has always been important in my life, and I will always be grateful.

I'd also like to acknowledge the support of SRI International, the agency that supported this work. Thanks especially to David Israel, Phil Cohen, and Bob Moore for providing me with the time to work on the thesis. Thanks are also due to Emma Pease for her help with the TeX macros used in this work and to Felix Ingrand and Paul Martin for their help with PostScript.

Contents

Chapter 1

Introduction

An important goal of syntactic research is the factoring of syntactic phenomena into those that are cross-linguistically invariant, those that vary from language to language, and those that are associated with individual lexical items. In the case of syntactic constraints on anaphoric binding, a number of widely divergent proposals have been made; in particular, it has been assumed in some work that the distribution patterns and antecedency conditions of reflexives, reciprocals, and pronominals are invariant across languages. Other analyses have assumed that these properties belong to particular languages; that is, that one may speak of domain or antecedent requirements on anaphoric elements as a property or parameter set for each language.

If a language has, for example, only one reflexive, it might seem to be an adequate approach to characterize the 'domain of reflexivization' as a property of universal grammar, or of a particular language. However, if a language has two or more reflexives, each with a different domain, it becomes clear that the domain of reflexivization must be a property of the particular lexical item for which it is applicable. Similarly, if a language has two or more reflexives, each with different requirements on the syntactic role of its antecedent, these antecedent requirements must be taken as lexically specified for each anaphoric element.

Seen in this light, the task of determining universals of anaphoric binding is seen not as defining what conditions on anaphoric binding are associated with a particular language, but as determining the universally-available set of possible binding

constraints which each anaphoric element can obey. Examining languages with multiple anaphoric elements provides evidence for such a universally-available inventory. In Chapter 2, two such languages are examined in detail: Norwegian, a Germanic language, and Marathi, an Indo-Aryan language.

Marathi has two reflexive elements: *aapaṇ* and *swataah*. The short-distance reflexive *swataah* has a distribution that is somewhat similar to the English reflexive *himself*, though there are differences between the two; for example, the antecedent of *swataah* must be a subject. The long-distance reflexive *aapaṇ* must be bound to the logical subject. It may not corefer with a coargument, but it must have an antecedent in the sentence in which it appears. The two Marathi reflexives differ from each other, then, both with regard to the domain in which they must be bound and with regard to the elements which can be their antecedents.

Norwegian has several anaphoric elements. One, *seg*, has a distribution similar to that of Marathi *aapaṇ*; it may not have a coargument as its antecedent, but it must have an antecedent in the same sentence. It is unlike *aapaṇ* in two respects: first, its antecedent must be a subject, not a logical subject; and second, its antecedent must appear in the same minimal finite domain. Another anaphoric element, *seg selv*, must have a subject coargument as its antecedent. Another, *ham selv*, must have a nonsubject antecedent; the antecedent must appear in the minimal domain containing a subject. These three elements, then, exemplify three different binding domains and two different conditions on the possible grammatical function of the antecedent. Other conditions govern the distribution of the reciprocal and the possessive pronoun and reflexive in Norwegian. It is particularly clear in the case of Norwegian that binding constraints must be associated with each anaphoric element rather than specified for Norwegian as a whole.

Patterns of anaphoric binding in both of these languages illustrate one of the basic claims of this work: constraints on anaphoric binding are lexically associated with each anaphoric element. In fact, generalizations have been noted that deal specifically with the lexical form of the anaphoric element: elements of a particular morphological form are usually or always associated with particular sets of anaphoric binding constraints. In Chapter 3, I discuss some proposals that do not state binding

constraints lexically; these are shown not to be viable.

The complete inventory of binding constraints can be characterized in terms of three grammatical concepts:

- subject,
- tense, and
- predicate.

It is interesting to note that these concepts all denote some syntactically or semantically 'complete' entity. In a complete, consistent f-structure, a PRED denotes a syntactically saturated argument structure; presence of a SUBJ entails a predication involving some property and the subject; and presence of TENSE indicates an event that has been spatiotemporally anchored. These 'complete' entities are the relevant domain for binding conditions.

Constraints as to the grammatical function of the antecedent are always stated in terms of coreference with or disjointness from a subject; these constraints are referred to as *antecedent* constraints. The domain in which the antecedent of an anaphor must or must not be found is always constrained relative to either the syntactic predicate of which the anaphoric element is an argument, the minimal domain with a subject containing the anaphor, or the minimal tensed domain containing the anaphor. These constraints are referred to as *domain* constraints. Combinations of constraints stated with reference to these concepts form the universally-available set of binding constraints which anaphoric elements can obey.

Data from Marathi and Norwegian illustrate that the possible range of anaphoric elements is much more varied than is reflected by a simple opposition of elements that must be bound (reflexives and reciprocals) and those that must be free (pronominals). There are two kinds of binding constraints: *positive* constraints state which elements an anaphoric element may be coreferent with, and *negative* constraints state which elements an anaphoric element may not be coreferent with. Generally, the term 'reflexive' refers to an element that must be bound; a reflexive must have an antecedent

within some syntactically-definable domain. English *himself* is such an element, obeying only positive constraints. The term 'pronominal' refers to an element that must be free; a pronominal is noncoreferent with elements in some syntactically-definable domain. English *him* is an example, obeying only negative constraints.

Some Norwegian and Marathi anaphoric elements are, by this criterion, both reflexives and pronominals: they must have an antecedent within a wide domain but must be noncoreferent from elements within a narrow domain. Furthermore, the kinds of positive and negative constraints obeyed by these elements are the same as those obeyed by elements that are subject to only a single kind of constraint. These facts support the relevance of a typology of *constraints* rather than a typology of anaphoric elements. Anaphoric elements may be subject to positive constraints, negative constraints, both, or neither; a theory which correlates this difference with a difference in types of anaphoric elements does not account for the fact that the same inventory of binding constraints is relevant for each of the various elements.

Certain interactions between the various binding requirements are also observable. In particular, domain constraints can be shown not to be independent of antecedent constraints. A single antecedent must satisfy both kinds of constraints; the domain and antecedent constraints may not be satisfied separately. The two kinds of requirements are actually two aspects of a single positive or negative requirement.

There are also certain universal conditions on possible kinds of binding constraints. For example, binding constraints are always local, in that constraints never apply only to non-local elements. There are anaphors that require their antecedents to appear in a higher clause, as data presented in Chapter 2 illustrate; constraints of this nature can always be reduced to a combination of a requirement of disjointness from elements in a local domain and a requirement of coreference with elements in a larger domain. There do not seem to be any anaphors which *require* an antecedent in a non-local domain but also *permit* an antecedent in a local domain, for example.

Chapter 4 provides a formal characterization of the binding constraints illustrated in Chapter 2. Stating binding constraints formally has several advantages: it provides for a precise statement of the constraints, and it allows a rigorous prediction of the interactions between them. Often, linguistic analyses are stated in such an informal

manner that it is not possible to determine the exact nature of binding constraints or how they should interact. Lexical-Functional Grammar (LFG) provides a rich formal vocabulary for stating anaphoric binding constraints formally and precisely.

In LFG, aspects of an utterance are represented by various levels of linguistic representation, known as *projections.* For example, a semantic projection might encode semantic relationships, a discourse projection might encode discourse information, a phonological projection might encode phonological information, and so on. The levels of representation that encode surface syntactic relations are the *constituent structure* and the *functional structure.*

The constituent structure or *c-structure* is represented as a tree; it encodes syntactic category information and dominance and precedence relations. The functional structure or *f-structure* is represented as a directed graph. It encodes syntactic argument structure and functional relations such as 'subject-of'. The anaphoric binding constraints discussed in this work are, for the most part, statable in terms of f-structure relations.

These constraints can be stated by use of the technique of *functional uncertainty.* Functional uncertainty is a method for stating relations between two positions in a graph structure such as the f-structure. In the case of anaphoric binding constraints, the positions to be related are the f-structure representations of the pronoun and its antecedent. In Chapter 4, the use of functional uncertainty to state anaphoric binding constraints is illustrated.

Functional uncertainty is used in expressing *binding equations*, equations delineating possible relations between the f-structure of a pronoun and that of its antecedent. That is, the lexical entry for each anaphoric element may contain a positive or a negative binding equation involving functional uncertainty. These equations provide negative and positive constraints by stating permissible relations between the f-structure of an pronoun and the f-structure of the elements with which it may or may not corefer.

By stating binding conditions in terms of binding equations, the kinds of interactions between binding requirements that were noted above are captured. Domain and antecedent requirements are not independent of each other; the use of binding

equations makes it possible to encode these two kinds of constraints as two aspects of a single requirement. Theories in which the two constraints are treated separately do not predict interactions of this kind.

Conditions on the possible forms of these equations can be used to limit the types of anaphoric binding conditions that are predicted. These limitations can be stated as restrictions on possible types of binding equations. For example, the fact that binding constraints are always local can be stated as a condition on possible domain specifications of binding equations.

Other universal conditions on anaphoric binding, explored in Chapter 5, have to do with interactions among positive and negative binding requirements. These can be stated as generalizations holding of the relation between the positive and the negative binding equations. Pronouns may be associated with both positive and negative equations, but the domain in which a negative constraint holds is never larger than the domain in which a positive constraint holds.

The set of elements that an anaphoric element may be required to corefer with or be disjoint from is also constrained by other factors. For instance, the antecedent of an anaphor must be *superior* to the anaphor, in a sense discussed in Chapter 5. Superiority is defined in terms of configurational relations statable on the f-structure representation, as well as in terms of a thematic argument hierarchy. Binding equations are stated in such a way as to refer only to elements that are superior to the anaphor.

In sum, this work provides a statement of the possible anaphoric binding constraints that are universally available. Universal constraints on anaphoric binding amount to restrictions on the type of constraints that each anaphoric element can obey and on the possible cooccurrence of these constraints; that is, the analysis is stated in terms of a typology of constraints rather than a typology of anaphoric forms. The precise formal statement of these constraints enables a clear characterization of the nature of the constraints as well as of the interactions between them.

Chapter 2

An Inventory of Binding Requirements

Examining constraints on anaphoric binding in several languages provides evidence for a range of conditions on domains in which anaphoric elements are free and bound, and for conditions on permissible and impermissible syntactic roles for antecedents of various anaphoric elements. In the following, we will first examine anaphoric binding conditions on two anaphoric elements in Marathi. Next, we will consider data from Norwegian; Norwegian, too, has multiple anaphoric elements.

These languages provide evidence for a certain inventory of binding constraints, definable in terms of three grammatical concepts: PRED, SUBJ, and TENSE. Combinations of requirements statable in terms of these concepts form the complete set of constraints which Marathi and Norwegian anaphors obey.

Data from Marathi and Norwegian also show that the range of anaphoric elements is much more varied than can be captured by a simple distinction involving reflexives, reciprocals, and pronominals. The term 'reflexive' is often used to refer to an element that is required to find an antecedent within some domain such as the minimal sentence in which it appears. The term 'pronominal' generally refers to an element that is required to be noncoreferent with other elements within some domain. As we will see, some anaphoric elements exhibit both of these characteristics; there are anaphors which are required to be disjoint from elements in a local domain but coreferent with an elements in a higher domain. Thus, a simple two-way distinction between 'pronominals' and 'reflexives' is not sufficient to characterize the range of anaphoric elements that is found. In the following, I will often refer to elements that

must stand in a certain structural relation to their antecedent as 'reflexives' and to elements that must be noncoreferent with some set of elements as 'pronominals'. It should be kept in mind, though, that this terminology does not reflect the complex nature of the binding constraints that anaphoric elements can obey.

2.1 Lexically Specified Anaphoric Domains: A Case Study

Theories in which some particular domain for anaphoric binding is specified as a language universal or as a property of a particular language do not allow for cases in which a language has more than one reflexive, each with a different domain. Marathi, an Indo-Aryan language spoken in west-central India, is a language with two reflexives, a 'long-distance' and a 'short-distance' one. The long-distance reflexive is required to be disjoint in reference from (i.e., noncoreferent with) arguments within a local domain but bound to (i.e., coreferent with) a higher argument in the same sentence. Its antecedent must be a 'logical' subject. The short-distance reflexive is required to be bound within a local domain, though not the same domain within which the long-distance reflexive is required to be free, and its antecedent must be a 'surface' subject.

Thus, to characterize the 'domain of reflexivization' in Marathi, reference is needed to three domains: the domain in which the long-distance reflexive must be free (disjoint from arguments that stand in a certain structural relation to it), the domain in which it is bound (required to corefer with an argument of a certain type), and the domain in which the short-distance reflexive is bound. Additionally, the two reflexives require different kinds of antecedents: one requires a 'logical' subject, while the other requires a 'surface' subject. I claim that specifying these constraints as lexical properties of each reflexive provides not only the simplest way of representing the facts about anaphoric binding in general, but the only way of specifying the facts in a language such as Marathi.

In the following, I will first note some general syntactic properties of Marathi; I will then examine the distribution and antecedency conditions of the two Marathi reflexives.

2.1.1 Typology

Marathi, like most Indo-Aryan languages, is verb-final; unmarked word order is SOV, though scrambling is also possible. As we will see, subjects in Marathi can be marked with dative or ergative case; they can also appear without overt casemarking.

As is the case in many other Indo-Aryan languages, verb agreement in Marathi is a rather complex phenomenon. The verb agrees with the subject, unless it is marked with an overt casemarker. Otherwise it agrees with the object, unless it too is marked with a casemarker. If both subject and object are marked with case, the verb takes default agreement.

Intransitive verbs and transitive verbs in aspects other than perfect agree with the subject:

(2.1) Mary swataah-hovun uṭhali
Mary of her own accord got up-FemSg
'Mary got up of her own accord.'

Subjects of transitive verbs are marked with ergative case in the perfect aspect. In this situation, the verb agrees with the unmarked object:[1]

(2.2) John ni sapharchãd khaalle
John ERG apple ate-NeutSg
'John ate an apple.'

When the verb is transitive and perfect, ergative marking appears on the subject; if accusative marking appears on the object, the verb agrees neither with the subject nor with the object but takes default (third person neuter singular) agreement:

(2.3) masterãã ni mulaa laa maarle
teacher ERG boy ACC hit-NeutSg
'The teacher hit the boy.'

[1]ERG=ergative; DAT=dative; ACC=accusative; GEN=genitive; LOC=locative.

Marathi also has subjects marked with dative case, referred to as 'dative subjects'.[2] The verb does not agree with the dative subject:

(2.4) John laa Jane ci aaṭhwaṇ aali
John DAT Jane GEN memory came.
'John remembered Jane.'
('To John, a memory of Jane came.')

The reasons for considering these ERG- and DAT-marked arguments to be subjects will be discussed below; here we note simply that agreement is not a reliable test of subjecthood, since the verb sometimes agrees with nonsubjects.

2.1.2 Subordination in Marathi

As mentioned above, unmarked word order in Marathi is SOV. In structures not involving subordination, the tensed verb always appears at the end:

(2.5) John jhoplaa
John slept
'John slept.'

(2.6) Jane ni John ce pustak phekun dile
Jane ERG John GEN book throw give
'Jane threw John's book.'

In cases involving tenseless VP subordination — gerunds (example (2.7)) and VP complements (example (2.8)) — the main verb appears at the end as well (the subordinate VP is set off in square brackets):

(2.7) Jane laa [John ne swataaci pustake phekun dilyaace] kaḷḷe
Jane DAT John ERG self-GEN books throw give learned
'Jane learned about John throwing away self's books.'

[2]In Marathi (as in many Indo-Aryan languages) dative and accusative marking are homophonous. The two are distinguished on the basis of distribution: ACC marking appears only on animate or definite direct objects, whereas DAT marking appears regardless of the characteristics of the DAT-marked argument. That is, the presence or absence of ACC marking depends on properties of the NP as well as on properties of the verb, while the presence of DAT marking depends only on the verb. In the following, the marker *-laa* is glossed as either DAT or ACC on this basis.

(2.8) Jane ni Bill laa [pustak phekaaylaa] saangitle
Jane ERG Bill DAT book throw told
'Jane told Bill to throw the book.'

There are two kinds of tensed complement clauses in Marathi. With some main verbs, only one or the other type may be used; with others, either is possible. In the first type, the subordinate clause follows the main verb, as in the following example:

(2.9) Jane laa vaaṭtaa ki [John guptaher aahe]
Jane DAT believes that John spy is
'Jane believes that John is a spy.'

This contrasts with the situation in which a VP complement is involved, as in (2.8); in that case, the VP complement precedes the main verb.

In the second type, the subordinate clause appears in sentence-initial position:

(2.10) [John gharii gelaa] asa Bill laa waaṭta
John home went so Bill DAT thinks
'Bill thinks that John went home.'
('John went home — so Bill thinks.')

Here, the tensed sentential complement precedes the main clause. The distinction between simple sentences and subordinated clauses — both tensed S and tenseless VP subordinate clauses — will be important later, in our discussion of the distribution of the two Marathi reflexives.

2.1.3 Subjecthood in Marathi

The antecedent of one Marathi reflexive, short-distance *swataah*, must be a 'surface' subject. I will call this the Subject Binding Condition.[3] The antecedent of the other, long-distance *aapaṇ*, is the 'underlying' or 'logical' subject;[4] the logical subject is the argument that bears the grammatical relation 'subject' in an active sentence, or

[3]Faltz (1985) calls this condition the Subject Antecedence Condition, citing a number of other examples of reflexives which obey this condition.

[4]I use the terms 'underlying' and 'surface' subject merely for convenience, without postulating a transformational analysis of passivization.

the argument that has been 'demoted' and appears as an oblique phrase in a passive sentence. Antecedent requirements for the two Marathi reflexives will be discussed in greater detail below; here, some independent tests for surface and logical subjecthood in Marathi are examined.

The verb *paahije* 'must' provides a test for surface subjecthood. The subject of *paahije* is marked with DAT, and it must be coreferent with the implicit subordinate clause subject:

(2.11) tyaalaa gele paahije
he-DAT go must
'He must go.'

(2.12) tyaalaa bet aaṭhwaalaa paahije
he-DAT plan remember must
'He must remember the plan.'

All types of subjects — intransitive subjects, transitive subjects (sometimes marked with ERG), and dative subjects — can participate in this construction.

Joshi (1989) discusses adjective formation, which constitutes another test for surface subjecthood. The following sentence is in the active voice:

(2.13) te gurujii mulaa-naa maartaat
that teacher children-ACC hits
'That teacher (usually) hits children.'

The corresponding adjective formation involves adding the suffix *-RAA*; the result is an adjectival clause predicated of the surface subject of the verb:

(2.14) maarṇaare mule
beat-RAA children
'children who beat'
*'children who are beaten'

The situation is the same when the verb is passive. Passivization in Marathi involves the use of the simple past form of the main verb together with the auxiliary verb *jaaṇe*

'go'; both verb forms bear agreement morphology. For clarity, the passive auxiliary is glossed not as 'go' but as 'PASS'. Adjective formation from a passivized verb complex involves predication on the surface subject:

(2.15) mulaa-naa tyaa gurujiin-kaḍuun maarle jaate
children-ACC that teacher-by hit PASS
'Children are (usually) beaten by that teacher.'

(2.16) maarlii jaaṇaarii mule
hit PASS-RAA children
*'children who beat'
'children who are beaten'

Joshi (1989) shows that there are also tests which pick out the logical subject in Marathi. One of these tests is a construction in which two clauses with identical subjects are adjoined; the verb of one of the clauses is participial:

(2.17) mi tikḍe jaauun tyaalaa bheṭṭo
I there going him-ACC meet
'(Upon) $\emptyset_{i,*j}$ going there, I_i meet him_j.'

(2.18) diwas bhar kheḷuun John laa bhuuk laaglii
day all playing John DAT hunger stuck
'(Upon) $\emptyset_i$ playing all day, $John_i$ became hungry.'

(2.19) gharii jaauun John-nii Bill-laa shaalet paathawle
home-LOC having gone John-ERG Bill-ACC school-LOC sent
'(Upon) $\emptyset_{i,*j}$ going home, $John_i$ sent $Bill_j$ to school.'

In these examples, involving sentences in the active voice, the unrealized subject of the subordinate clause must be coreferent with the subject of the main clause. Where the main clause verb is in the passive voice, though, coreference is with the logical subject: (ACC marking is obligatorily retained on the subject of passivized transitive verbs):

(2.20) [gharii jaauun] John-kaḍuun Bill-laa maarle gele
home-LOC having gone by John Bill-ACC hit was
'(Upon) $\emptyset_{*i,j}$ going home, $Bill_i$ was beaten by $John_j$.'

Application of these tests confirms that the (surface) subject is indeed the antecedent for the short-distance reflexive *swataah*, and that the antecedent for long-distance *aapaṇ* is the logical subject.

2.1.4 Reflexives in Marathi

There are two reflexives in Marathi: the short-distance reflexive *swataah* and the long-distance reflexive *aapaṇ*.[5,6] The antecedent of both reflexives must be the surface subject if the sentence is in the active voice:

(2.21) John nii Bill laa swataabaddal maahiti dili
John ERG Bill DAT about self information gave
'John$_i$ gave Bill$_j$ information about self$_{i,*j}$.'

(2.22) Jane ne John laa kaḷavle ki aapaṇ turangaat aahot
Jane ERG John ACC informed that self prison-LOC was
'Jane$_i$ informed John$_j$ that self$_{i,*j}$ was in prison.'

The situation is different with regard to passive sentences, however. Joshi (1986) shows that the long-distance reflexive *aapaṇ* has the logical subject as its antecedent, while the short-distance reflexive *swataah* has the surface subject as its antecedent. In the following sentence, the antecedent of *aapaṇ* is *Bill*, the logical subject/agent of *hit*:

(2.23) John laa Bill kaḍuun aaplyaa gharii maarle gele
John ACC Bill by self-GEN house-LOC hit PASS
'John was hit by Bill$_i$ at self's$_i$ house.'

As noted in Section 2.1.3, there are a number of grammatical processes in Marathi which refer to the surface subject, while others refer to the logical subject. The question of how to define 'logical subject' is a very interesting one; here, however,

[5]The form *aapaṇ* has other pronominal uses: it is used as a first person singular or inclusive plural pronoun and as a second person honorific pronoun. In the following, I will discuss only the reflexive use of *aapaṇ*.

[6]Some data in this section are taken from Dalrymple (1984), Joshi (1989), Wali (1976), and Wali (1979). The rest are due to Smita Joshi, Kashi Wali, and my informant Chitra Lele.

the constraints that we are primarily interested are surface syntactic constraints on anaphoric binding, and these can be stated with reference to the f-structure.

To define 'logical subjecthood', reference is needed to a level of representation encoding the predicate-argument structure of verbs — that is, to a level where the active/passive distinction is not encoded and where the argument that is the 'logical subject' can be characterized nondisjunctively. The level of f-structure does not suffice, since there is no uniform characterization of the logical subject at f-structure; the logical subject is the subject of an active sentence, but in a passive sentence, the logical subject either appears as an oblique argument of the verb or does not appear at all. Rather, the logical subject is definable as the argument of a predicate that is highest on a thematic hierarchy (Kiparsky 1989, Bresnan and Kanerva 1989). In the following, I will not provide a formal characterization of reflexive antecedency in terms of the notion of logical subjecthood; however, I will return to the question of the role played by a thematic hierarchy in determining anaphoric binding possibilities in Chapter 5, Section 5.6.3.

In order to determine further constraints on the domain of each of these reflexives, let us consider simple sentences: sentences with only one tensed verb which do not have tenseless VP complements. In some positions in sentences like these, *swataah* is obligatory; if the reflexive is in direct object position, for example, only *swataah* and not the long-distance reflexive *aapaṇ* (or its accusative case counterpart *aaplyaalaa*) can appear:

(2.24) Jane ne swataahlaa bockaarle
Jane ERG self-ACC scratched
*aaplyaalaa
self-ACC
'Jane$_i$ scratched herself$_i$.'

(2.25) Jane ne swataahlaa baḍavle
Jane ERG self-ACC beat
*aaplyaalaa
self-ACC
'Jane$_i$ beat herself$_i$.'

As we will see below, the inability of *aapaṇ* to appear here is not due to the absence of an accusatively casemarked version of *aapaṇ* — that is, to a gap in the paradigm. In other contexts, accusative *aapaṇ* does appear: for example, in subordinate clause object position. Examples (2.41) and (2.47) contain felicitous occurrences of accusative *aapaṇ*.

In other simple sentences, too, the short-distance reflexive must be used:[7]

(2.26) Jane swataahši baḍbaḍte
Jane with self mutters
*aaplyaaši
with self
'Jane$_i$ mutters to/with herself$_i$.'

(2.27) Jane laa swataacaa raag yeto
Jane DAT self-GEN anger comes
*aaplaa
self-GEN
'Jane$_i$ gets angry at herself$_i$.'
('To Jane$_i$, self's$_i$ anger comes.')

(2.28) Jane swataacaa dweš karte
Jane self-GEN hate does
*aaplaa
self-GEN
'Jane$_i$ hates herself$_i$.'
('Jane$_i$ does self's$_i$ hate.')

Again, it would not be possible to explain these facts by positing a gap in the lexical paradigm for *aapaṇ*, since the forms in the starred sentences can appear elsewhere.

In other contexts, however, either reflexive is possible, although in some of these cases *swataah* is 'preferred' (Wali 1976, p. 48):

[7]Example (2.27) contains a dative subject which is the antecedent of the reflexive; recall the discussion above concerning subjects with non-nominative case marking.

(2.29) Jane ne swataahkartaa saaḍi gheṭ li
Jane ERG for self sari bought
aaplyaakartaa
for self
'Jane$_i$ bought a sari for herself$_i$.'

(2.30) Jane ne John laa swataahbaddal maahiti dili
Jane ERG John DAT about self information gave
aaplyaabaddal
about self
'Jane$_i$ gave John information about herself$_i$.'

(2.31) Jane ne John laa swataahvišayi kaahic saangitle naahi
Jane ERG John DAT about self anything told not
aaplyaavišayi
about self
'Jane$_i$ didn't tell John anything about herself$_i$.'

These facts indicate that the short-distance reflexive may appear as a *coargument* (an argument of the same predicate) with its antecedent, while the long-distance reflexive may not. I will refer to this requirement as the Coargument Disjointness Condition: *aapaṇ* must be disjoint from — that is, noncoreferent with — its coarguments.

It might be noted that the structure of the following two examples appears to be similar, yet the possibilities for reflexivization are different:

(2.32) Jane laa swataacaa raag yeto
Jane DAT self-GEN anger came
*aaplaa
self-GEN
'Jane$_i$ gets angry at herself$_i$.'
('To Jane$_i$, self's$_i$ anger comes.')

(2.33) Jane ni aaplye pustak phekun dile
Jane ERG self-GEN book threw give
'Jane$_i$ threw self's$_i$ book.'

In both cases the reflexive is marked with genitive case and appears to be modifying a direct object: 'self's anger' in (2.32) and 'self's book' in (2.33). However, in (2.32), the long-distance reflexive *aapaṇ* is disallowed, whereas in (2.33) it is permitted.

In fact, reanalysis has occurred in the examples in (2.32); that is, the sequence *raag yeto* 'anger comes' acts as a verb with a genitively-marked object. Thus the reflexive is an argument of the verb in (2.32) — it is a genitively-marked object — whereas it is not in (2.33). The bracketing for these examples is as indicated:

(2.34) *[Jane laa] [aaplaa] [raag yeto]
Jane DAT self-GEN anger came
'Jane$_i$ gets angry at herself$_i$.'
('To Jane$_i$, self's$_i$ anger comes.')

(2.35) [Jane ni] [aaplye pustak] [phekun dile]
Jane ERG self-GEN book threw
'Jane$_i$ threw self's$_i$ book.'

In (2.34), then, *aaplaa* is a coargument with the subject *Jane*, whereas in (2.35) it is not; *aaplye pustak* forms its own argument domain, so the reflexive *aapaṇ* is not a coargument with the subject of the main verb, *Jane*.

Independent evidence for reanalysis is provided by the following distinction:

(2.36) Jane laa John caa kharokharac raag yeto
Jane DAT John GEN indeed anger came
'Jane got angry at John, indeed.'
('To Jane, John's, indeed, anger came.')

(2.37) *Jane ni John ce kharokharac pustak phekun dile
Jane ERG John GEN indeed book threw
'Jane threw John's, indeed, book.'

It is possible to interpose the adverb *kharokharac* 'indeed' between the genitive object 'John' and the verbal sequence 'anger came' in (2.36). This is evidence that 'John' and 'anger' do not form a phrasal unit. On the other hand, it is not possible to insert an adverb between 'John's' and 'book' in (2.37), since 'John's book' is a noun phrase, the object of 'throw'.

2.1.5 Subordinate Tenseless VP's

Now we turn to the possibilities for reflexivization in clauses containing a tenseless VP. Here there seems to be a dialect split: some speakers allow the short-distance reflexive *swataah* anywhere inside a tensed clause, while others do not allow *swataah* in subordinate nontensed VP's.

For all speakers, example (2.38) is acceptable on the reading that the antecedent of the short-distance reflexive *swataah* is *Jane*. Some speakers also accept *John* as an antecedent (the symbol '%' indicates that the indexing is acceptable only for some speakers):

(2.38) John ne Jane laa swataahlaa maraaylaa saangitle
John ERG Jane DAT self-ACC hit told
'John$_i$ told Jane$_j$ to hit self$_{\%i,j}$.'

In example (2.38), the reflexive corefers with an object argument, but this is not a counterexample to the generalization that the antecedent of *swataah* must be a subject. This is because the object *Jane* of the verb *saangitle* controls the subject position of the infinitive *maraaylaa*. It is the subject position of the infinitive which is the antecedent of the short-distance reflexive for all speakers. Some speakers also allow the matrix subject to antecede the short-distance reflexive.

All speakers find the above example grammatical with the matrix clause subject as the antecedent of the reflexive if the long-distance reflexive is used instead of the short-distance reflexive:

(2.39) John ne Jane laa aaplyaalaa maraaylaa saangitle
John ERG Jane DAT self-ACC hit told
'John$_i$ told Jane to hit self$_i$.'

In example (2.40), the reflexives appear inside a gerundive NP. Again, the long-distance reflexive is acceptable for all speakers if the antecedent is the matrix subject. Some speakers also find the short-distance reflexive acceptable with the indexing as indicated:

(2.40) a. Jane laa John ne aapli pustake phekun dilyaace kaḷḷe
Jane$_i$ DAT John ERG self-GEN$_i$ books throwing learned
'Jane$_i$ learned about John throwing away self's$_i$ books.'

b. Jane laa John ne swataaci pustake phekun dilyaace kaḷḷe
Jane DAT John ERG self-GEN books throwing learned
'Jane$_i$ learned about John throwing away self's$_{\%i}$ books.'

In example (2.41), the reflexive appears inside a tenseless adverbial clause. All speakers find the long-distance reflexive acceptable in this environment; for some speakers, the short-distance reflexive is also acceptable.

(2.41) a. Jane John ne aaplyaalaa maarlyaavar rusun basali
Jane John ERG self-ACC hitting angrily remained
'Jane$_i$ remained angry upon John hitting self$_i$.'

b. Jane John ne swataahlaa maarlyaavar rusun basali
Jane John ERG self-ACC hitting angrily remained
'Jane$_i$ remained angry upon John hitting self$_{\%i}$.'

In sum, for some speakers, the short-distance reflexive must appear in the minimal domain containing the reflexive and a subjective function — SUBJ or POSS. In the terminology of LFG, a syntactic predicate and its arguments form a *nucleus*. A *complete* nucleus is a nucleus containing a subjective function. For these speakers, then, the short-distance reflexive *swataah* must find its antecedent in the minimal nucleus which contains it and a subjective function: the *minimal complete nucleus*. In the following, I will refer to this as the Minimal Complete Nucleus Binding Condition: *swataah* must be bound within the minimal complete nucleus.

For other speakers, the short-distance reflexive may find its antecedent anywhere inside a tensed domain. It is not restricted to positions inside the minimal complete nucleus. Even when it appears inside a nucleus with a SUBJ, it may take a higher subject as its antecedent. I will refer to this as the Minimal Finite Domain Binding Condition: in this dialect, *swataah* must be bound within the minimal finite domain. Even in this dialect, *swataah* is not completely unrestricted as to where it can appear in relation to its antecedent, as we shall see in the next section.

The long-distance reflexive, on the other hand, is restricted in that it may not appear as a coargument with its antecedent: it obeys the Coargument Disjointness Condition. Although it must be bound within the same sentence, the domain in which it must be bound is unrestricted; it can even be separated from its antecedent by a tensed boundary.

2.1.6 Tensed Subordinate Clauses

Let us now consider sentences with tensed subordinate clauses. Where the subject of the tensed subordinate clause is coreferent with the subject of the matrix clause, either the long-distance reflexive or the pronominal may be used. The following examples involve subordinate clauses which follow the main verb:

(2.42) Jane laa vaaṭtaa ki ti guptaher aahot
Jane DAT believes that she spy is
aapaṇ
self
'Jane$_i$ believes that she$_i$/self$_i$ is a spy'.

The pronominal *ti* may refer to *Jane* here; it may also have an extrasentential referent. The reflexive subject must corefer with the higher subject.

In the following example, the subordinate clause precedes the main clause:

(2.43) aapaṇ naapaas jhaalo hi baatmi Jane-ne naakaarli
self fail become this news Jane-ERG denied
'Jane$_i$ denied the news that self$_i$ failed.'

Here, too, the long-distance reflexive may appear where coreference with the matrix subject is intended.

The facts are more complex with regard to the appearance of the reflexive *swataah* in tensed subordinate clause subject position. For some (but not all) speakers, there are cases where *swataah* appears in or internal to a subordinate tensed clause subject position, as in the following:

(2.44) %Jane laa waaṭte ki swataa saglyaat sundar aaho
Jane DAT thought that self most beautiful was
'Jane$_i$ thought that self$_i$ was the most beautiful.'

(2.45) %Jane mhaṇaali ki swataaci parikshaa sampli
Jane said that self-GEN test finished
'Jane$_i$ said that self's$_i$ test was over.'

The division between speakers who do and do not accept sentences of this kind is not the same as the division between speakers who do and do not accept *swataah* in subordinate VP's. Some speakers accept both of the examples above, some accept neither, and some accept only one or the other. I will have more to say about binding of subjects and into subject positions in Chapter 5, Section 5.7.

In non-subject position of a tensed subordinate clause, the short-distance reflexive *swataah* cannot corefer with the matrix subject,[8] while the long-distance reflexive may:

(2.46) Sue ne Tom laa vinanti keli ki tyaane aaplaa kavitaa vaacu nayet
Sue ERG Tom DAT request did that he-ERG self-GEN poems read shouldn't
*swataacaa
self-GEN

'Sue$_i$ requested of Tom that he not read self's$_i$ poems.'

(2.47) Tom mhanat hota ki Sue ni aaplyaalaa maarle
Tom said that Sue ERG self-ACC hit
*swataahlaa
self-ACC

'Tom$_i$ said that Sue hit self$_i$.'

Here the short-distance reflexive corefers with the subordinate subject but not with the higher subject. Wali (1976) calls this the Intervening Subject Condition: she states that the short-distance reflexive *swataah* cannot corefer with a higher subject if an intervening subject exists.

[8]There may be dialects which differ with regard to this fact, particularly where scrambling is involved; see Wali (1976, p. 84 ff.). I will not discuss this here.

In sentences like the ones above, where coreference with a higher subject is intended, it is possible to use either the long-distance reflexive or the pronominal:

(2.48) Sue ne Tom laa vinanti keli ki tyaane aaplaa kavitaa vaacu nayet
Sue ERG Tom DAT request did that he-ERG self-GEN poems read shouldn't
ticyaa
her-GEN

'Sue$_i$ requested of Tom that he not read self's$_i$/her$_i$ poems.'

In general, then, either the pronominal or the long-distance reflexive *aapaṇ* can be used in subordinate clause positions where coreference with a higher subject is intended. The short-distance reflexive is not acceptable in tensed subordinate clause nonsubject position; for only some speakers, it may appear as the subject of a tensed subordinate clause. The long-distance reflexive must be bound within the sentence in which it appears, but it may appear at an indefinite distance from its antecedent. I will call this the Root S Binding Condition: *aapaṇ* must be bound within the root S in which it appears.

2.1.7 Subordinate clause subjecthood and *aapaṇ*

It might be thought that a possible analysis of complex constructions in which the long-distance reflexive appears is that the intervening subject somehow 'loses' its subjecthood; in this case, there would be only one subject: the one that is the antecedent of the reflexive. This may not be maintained.

As I have noted, among the properties peculiar to subjects in Marathi is participation in a process of subject coreference in which the subject of the matrix clause and the subject of an adverbial clause must be coreferent.

Now, the following sentence is acceptable on the coindexing indicated:

(2.49) John mhanat hota ki post-office saat jaauun Jane ni aaplyaalaa maarle
John said that post-office in going Jane ERG self-ACC hit
'John$_i$ said that (upon) going to the post-office, Jane hit self$_i$.'

Jane participates in the subject coreference construction here: it is Jane that goes to the post-office and it is Jane that hits John. Further, the higher clause subject

John is an acceptable antecedent for the long-distance reflexive. Therefore both have subject properties; both the higher clause subject *John* and the subordinate clause subject *Jane* are SUBJ's.

A more dramatic example follows:

(2.50) Jane laa vaaṭtaa ki John ni aaplyaa kaathi ni swataahlaa khaajavle
Jane DAT believes that John ERG self-GEN stick INST self-ACC scratched

'Jane$_i$ believes that John$_j$ scratched self$_j$ with self's$_i$ stick.'

Here the short-distance reflexive has the lower subject as its antecedent, while the long-distance reflexive has the higher subject. Therefore both *Jane* and *John* have subject properties in this example.

2.1.8 Logophoricity and *aapaṇ*

Clements (1975) presents an analysis of the Niger-Congo language Ewe in which he discusses *logophoric* pronouns — pronouns which take a logophoric antecedent. Logophoric antecedents are defined as those antecedents which lie in an indirect discourse relationship to the anaphor: the source of information or emotional expression, for example.

Clements provides these examples [pp. 157, 158]:

(2.51) Kofi gblɔ be yè-se Kɔku wò-nɔ yè dzu-m
Kofi said that LOG-hear Kɔku PRO-be LOG insult-A
'Kofi$_i$ said that he-LOG$_i$ heard Koku insulting him-LOG$_i$.

(2.52) me-se tso Kofi gbɔ be yè-xɔ nunana
PRO-hear from Kofi side that LOG-receive gift
'I heard from Kofi$_i$ that he-LOG$_i$ had received a gift.

In these examples, the antecedent of the logophor *yè* is the subject of the verb 'speak', *Kofi*.

Clements also provides an example of an extended discourse in which the antecedent of a logophoric pronoun is in a previous sentence [p. 171]:

(2.53) [The three of them$_i$ resolved that they$_i$ would take the moon out of the water.]

ne yèwoɖii tɔa me ko a
When LOG$_i$ had taken it out of the water ...

The distribution of logophoric pronouns is constrained with respect not to surface syntactic structure, but to a level of representation encoding discourse relations and notions such as 'point-of-view'. Kameyama (1985) provides an account of the Japanese anaphor *zibun* within the LFG framework; the account she presents makes use of the feature +LOGOPHORIC, first introduced by Bresnan et al. (1985). The constraints that are the focus of this work are those that are statable at a level of representation encoding surface syntactic relations, the f-structure of LFG; as such, I do not provide a formal account of logophoricity here.

If it can be shown that *aapaṇ* is a pure logophor, however, it would be inappropriate to constrain its distribution syntactically, with reference to the f-structure; its distribution would be determined with reference only to a discourse structure. In fact, Joshi (1989) states that 'when bound outside the clause, the reflexive *aapaṇ* shows some logophoric properties which are not fully studied or understood as yet'.

However, it does not seem that logophoric antecedents (in a strict sense) are acceptable as antecedents for *aapaṇ*. Example (2.54) is unacceptable if the antecedent is *John*:[9]

[9]The following sentence may appear to be a counterexample to the claim that logophoricity is not relevant to determining antecedency of reflexives:

(a) John laa Bill kaḍuun saangitla gela ki Jane ni aaplyaalaa maarla
John ACC Bill by tell PASS that Jane ERG self-ACC hit
'John was told by Bill$_i$ that Jane hit self$_i$.'

(As noted above, accusative marking is retained on subjects of passive sentences.) Here it appears that the logophoric antecedent/source *Bill* and not the subject may be the antecedent of the reflexive, contrary to the claims made above.

However, as we have seen, the antecedent of *aapaṇ* is always the logical subject:

(b) John laa Bill kaḍuun aaplyaalaa gharii maarla gela
John ACC Bill by self-GEN house-LOC hit PASS
'John was hit by Bill$_i$ at self's$_i$ house.'

This, then, is the explanation for the possibility of an apparently logophoric antecedent of a reflexive in the first example.

(2.54) Jane laa John kaḍuun kaḷḷe ki aapaṇ gharii jaaṇaar aahot
Jane DAT John by heard that self house-LOC is going
'Jane$_i$ heard from John$_j$ that self$_{i,*j}$ was going home.'

Although *John* is the source of expression in this example, the antecedent of *aapaṇ* must be *Jane* and not *John*.

'Jane' is not a possible antecedent of *aapaṇ* in example (2.55)b, although the use of the pronominal is permitted, as example (2.54)a shows:

(2.55) a. [Jane cyaa nawyaa yashaa-mule] [[ti paraabhut jhaale aahe] hi
Jane GEN new success-due-to she defeated happened is this
baatmi pasru shakli naahi]
news spread could not

b. *[Jane cyaa nawyaa yashaa-mule] [[aapaṇ paraabhut jhaalo aahot] hi
Jane GEN new success-due-to self defeated happened is this
baatmi pasru shakli naahi]
news spread could not
'Due to Jane's new success, the news that she/*self had failed could not spread.'

Here, too, the point of view expressed in the main clause is Jane's; nevertheless, *Jane* is not an appropriate antecedent for *aapaṇ*.

The following examples, in which the intended antecedent of *aapaṇ* is in a previous sentence, are also ungrammatical:

(2.56) Jane dukhi hoti. *aaplyaalaa jaataa aale naahi.
Jane sad was Self-DAT go could not
'Jane was sad. *Self could not go.'

(2.57) John laa waaiit waaṭle. *aapli aipat dhuḷi laa miḷaala hoti.
John DAT bad felt self-GEN reputation dust DAT joined was
'John felt bad. *Self's reputation was in shambles.'

If *aapaṇ* were a pure logophor, it ought to be able to appear in a different sentence from its antecedent, provided that the sentence it appears in could be construed as

standing in the appropriate discourse relation to the logophoric antecedent. This is not possible, however.

Maling (1984) and Sigurdsson (1989), building on work by Thráinsson (1976), present an analysis of Icelandic long-distance reflexives according to which a reflexive may find a logophoric antecedent in a higher clause. In Icelandic, the antecedent of the long-distance reflexive must have an antecedent that is both a subject and a logophoric antecedent.[10] Data from Icelandic show, then, that anaphors may be simultaneously subject to syntactic and semantic or pragmatic conditions on their distribution.

An analysis of a similar nature may also be appropriate for *aapaṇ*; the antecedent of *aapaṇ* must be the logical subject, as we have seen, but it may also be subject to discourse constraints on its distribution. It seems clear, though, that long-distance reflexivization in Marathi is syntactically constrained and does not depend entirely on logophoricity.

2.1.9 Distribution of pronominals

The Marathi pronominal *to* may corefer with the subject when it appears in non-nuclear position with relation to the subject; recall that this was just the situation where either the long-distance or the short-distance reflexive could appear in simple sentences. As in the case of the examples given above, the reflexives must have the subject as their antecedent; the pronominal may corefer either with the subject, the object, or an extrasentential element. In the following examples, the pronominal and reflexives are in adjunct position:

[10]This generalization holds whenever the antecedent for the reflexive f-commands it; however, there are also cases in which the antecedent of the reflexive appears in a previous sentence, where no f-command relation holds. Sigurdsson (1989, p. 11) presents this example:

(a) [Mary was always so nasty. When Olaf$_i$ came she would certainly tell self-LOG$_i$ to leave. She was always so nice to Olaf$_i$ ...]
Jà hùn segði sèr áreiðanlega að fara
yes she told-subjunctive self-LOG certainly to leave
'Yes, she would certainly tell self-LOG$_i$ to leave.'

(2.58) Jane ne swataahkartaa saaḍi gheṭ li
Jane ERG for self sari bought
aaplyaakartaa
for self
ticyaakartaa
for her
'Jane$_i$ bought a sari for herself$_i$/her$_i$.'

(2.59) Jane ne John laa swataahbaddal maahiti dili
Jane ERG John DAT about self information gave
aaplyaabaddal
about self
ticyaabaddal
about her
'Jane$_i$ gave John information about herself$_i$/her$_i$.'

(2.60) Jane ne John laa swataahvišayi kaahic saangitle naahi
Jane ERG John DAT about self anything told not
aaplyaavišayi
about self
ticyavišayi
about her
'Jane$_i$ didn't tell John anything about herself$_i$/her$_i$.'

If the pronominal is in the same nucleus as the antecedent, however, coreference is not possible:

(2.61) Jane ne swataahlaa bockaarle
Jane ERG self-ACC scratched
*aaplyaalaa
self-ACC
*tilaa
her-ACC
'Jane$_i$ scratched herself$_i$/*her$_i$.'

Unlike the long-distance reflexive, the pronominal need not have an antecedent in the same sentence:

(2.62) Mary dukhi hoti. tilaa jaataa aale naahi.
Mary saw was she-DAT go could not
*aaplyaalaa
she-DAT
'Mary$_i$ was sad. She$_i$/*Self$_i$ could not go.'

Thus the pronominal has just the same distribution as the long-distance reflexive when it is coreferent with a subject in the same sentence. There are two important differences between the two: the pronominal need not have a subject as its antecedent, and it need not have an antecedent in the same sentence.

2.1.10 Anaphoric Binding in Marathi

It is evident that it is not possible to describe the facts of Marathi with reference to a single 'domain of reflexivization'; the two reflexives have different domains. In fact, the data we have seen illustrate four domain conditions. In one dialect, *swataah* is bound in the minimal complete nucleus; in the other, it is bound in the minimal finite domain. The domain in which *aapaṇ* and the pronominal *to* are free is the coargument domain; *aapaṇ* is bound in the sentence in which it appears.

Additionally, each reflexive obeys a different condition on the function of its antecedent: *swataah* must be bound to the surface subject, while *aapaṇ* must be bound to the logical subject. Antecedent conditions for the Marathi reflexives must also be specified separately for each reflexive.

The properties of the Marathi reflexives are summarized in the following table:

	Bound to	Disjoint from
swataah (more restricted dialect)	Surface subject in minimal complete nucleus	
swataah (less restricted dialect)	Surface subject in minimal finite domain	
aapaṇ	'Logical' subject in root S	Argument in coargument domain

In sum, each anaphoric element in Marathi obeys a different set of anaphoric binding conditions; these conditions must be stated lexically, not with reference to Marathi as a whole. The conditions examined so far form a partial inventory of universally-available binding conditions. In the next section, the inventory will be expanded when data from Norwegian is considered.

2.2 Additional Binding Requirements: Norwegian

We have seen that several anaphoric binding conditions are attested by data from Marathi. In Norwegian, additional binding conditions are found. In the following, I will present a brief characterization of Norwegian anaphoric elements and the binding conditions they obey. In doing this, I rely heavily on data from Hellan (1988).[11]

Hellan's analysis is presented within the framework of Government and Binding; however, it is substantially different from many of the other GB analyses of anaphoric binding to be discussed in Chapter 3. His analysis associates binding constraints with individual anaphoric elements and so, like the one presented here, is lexical. Further, Hellan rejects many standard GB assumptions in his work; he states that the binding principles presented in Chomsky (1981) (to be discussed in Chapter 3, Section 3.1.1) 'cover only a very limited subpart of what constitutes a possible anaphoric system' (Hellan 1988, p. xi).

One of Hellan's proposals deals with what he calls the 'Principle of Independent Targeting'. This is a very general principle, according to which an anaphor must be bound by an antecedent that 'is the target of another basic grammatical relation' [p. 81], where the concept of 'basic grammatical relation' includes being the subject of predication or being the argument of some predicate. Such a claim follows without stipulation in a system such as the one proposed here: acceptable antecedents for an anaphor are f-commanding elements that bear a governable grammatical function. Stating binding relations in terms of f-structure relations, as will be done in Chapter 4, presupposes generalizations such as this one, since the grammatical function borne

[11] I would also like to thank Lars Hellan for providing some of the Norwegian examples in this section and in the following chapters.

by the antecedent is explicit at the f-structure level.

Similarities between Hellan's analysis and the analysis presented here will be discussed at greater length in Chapter 5. In particular, what Hellan calls the 'Command Principle' is paralleled by what I call the Superiority Condition; the antecedent of an anaphor must be superior to the anaphor in a structural sense and, often, in terms of a thematic hierarchy.

Norwegian has a larger inventory of anaphoric elements than does either English or Marathi. *Seg*, *sin*, *seg selv*, and *ham selv* are elements that must be bound within some syntactically definable domain; they might be thought of as the reflexive elements in Norwegian. *Hans*, the possessive pronominal, must be disjoint from subjects in a local domain. Norwegian also has a reciprocal element, *hverandre*, whose distribution is quite similar to English *each other*.

2.2.1 The Coargument Binding Condition: *seg selv*

As Hellan (1988) shows, the anaphor *seg selv* obeys the Subject Binding Condition; like Marathi *swataah*, its antecedent must be a surface subject.

(2.63) Jon fortalte meg om seg selv
Jon told me about self
'Jon$_i$ told me about self$_i$.'

(2.64) *vi fortalte Jon om seg selv
we told Jon about self
'We told Jon$_i$ about self$_i$.'

The domain condition holding of *seg selv* is one we have not yet encountered, however. We have seen that it is possible for an anaphor to respect a requirement of disjointness from higher coarguments; the Marathi reflexive *aapaṇ* is such an anaphor. The corresponding positive requirement may also hold: an anaphor may require a coargument as a binder.

Hellan (1988) shows that the Norwegian anaphor *seg selv* must be an argument of the same predicate as its antecedent: it respects the Coargument Binding Condition. In addition, it respects the Subject Binding Condition: its antecedent must be a

subject. Taking the two requirements together, its antecedent must be the subject of the same predicate it is an argument of. This is in contrast to the Norwegian reciprocal *hverandre*, which is subject to the less strict Minimal Complete Nucleus Binding Condition; it need not be a coargument with its antecedent, though its antecedent must appear in the minimal domain containing *hverandre* and a subject.

Both *seg selv* and *hverandre* can appear as coarguments with their antecedent (examples from Hellan (1988, p. 67 ff)):

(2.65) a. Jon fortalte meg om seg selv
Jon told me about self
'Jon$_i$ told me about self$_i$.'

b. de fortalte meg om hverandre
they told me about each other
'They$_i$ told me about each other$_i$.'

(2.66) a. Jon's angrep på seg selv var en overraskelse
Jon's attack on self was a surprise
'Jon's$_i$ attack on self$_i$ was a surprise.'

b. deres angrep på hverandre var en overraskelse
their attack on each other was a surprise
'Their$_i$ attack on each other$_i$ was a surprise.'

However, *seg selv* has a more limited distribution than *hverandre*, as these sentences illustrate:

(2.67) a. *hun kastet meg fra seg selv
she threw me from herself
'She$_i$ threw me away from herself$_i$.'

b. de kastet meg til og fra hverandre
they threw me to and from each other
'They$_i$ threw me to and from each other$_i$.'

I assume that '(to and) from' forms a nucleus, but not a complete nucleus, and that that *seg selv* must appear as a coargument with its antecedent, while *hverandre* must appear in the minimal *complete* nucleus containing it and its antecedent.

Hellan provides another sentence that makes much the same point:

(2.68) a. *Jon leste noen omtaler av seg selv
Jon read some reports about self
'Jon$_i$ read some reports about self$_i$.'

b. Jon og Marit leste noen omtaler av hverandre
Jon and Marit read some reports about each other
'[Jon and Marit]$_i$ read some reports about each other$_i$.'

The NP in which the anaphors appear does not have a POSS function (recall that POSS is defined as a prenominal genitive); thus it forms a nucleus, but not a complete one. *Seg selv* is not permitted here, since it is not a coargument with its antecedent; however, *hverandre* is permitted, since it appears in the minimal complete nucleus with its antecedent.

2.2.2 The Subject Disjointness Constraint: *ham selv*

The anaphor *ham selv* provides evidence for another binding constraint: the Subject Disjointness Constraint. This is the negative counterpart of the Subject Binding Constraint. An anaphor which obeys the Subject Binding Constraint, such as Marathi *swataah* or Norwegian *seg* or *seg selv*, must be coreferent with a subject; an anaphor which obeys the Subject Disjointness Constraint must be *disjoint* in reference from subjects.

The following examples show that *ham selv* must be bound to a non-subject:

(2.69) vi fortalte Jon om ham selv
we told Jon about self
'We told John$_i$ about self$_i$.'

(2.70) *Jon snakker om ham selv
Jon talks about self
'Jon$_i$ talks about self$_i$.'

As the following examples show, the distribution of *ham selv* is wider than that of *seg selv*, in that it need not be bound to a coargument:

(2.71) jeg ga Jon en bok om ham selv
I gave Jon a book about self
'I gave Jon$_i$ a book about self$_i$.'

(2.72) jeg introduserte Jon for en venn av ham selv
I introduced Jon to a friend of self
'I introduced Jon$_i$ to a friend of self$_i$.'

However, *ham selv* may not take an antecedent which is outside the minimal domain containing it and a subject. The following example is ungrammatical for this reason:[12]

(2.73) *jeg lovet Jon å snakke om ham selv
I promised Jon to talk about self
'I promised Jon$_i$ to talk about self$_i$.'

The subject of *lovet* 'promise' controls the subject of *talk*; *Jon* appears outside the minimal domain containing the anaphor and a subject, the infinitival clause. Thus *ham selv* obeys the Minimal Complete Nucleus Binding Condition: it must be bound in the minimal domain containing it and a subject.

Taken together, the requirements on *ham selv* are that it must be bound within the minimal domain containing a subject, but it must be disjoint from that subject. It must be bound to a nonsubject within that domain.

2.2.3 Long distance binding: *seg*

The anaphor *seg* appears in various constructions in Norwegian. There are cases in which it appears to be bound locally, as in the following example:

(2.74) Jon skammer seg
Jon shames self
'Jon is ashamed.'

[12]Example (2.73) is grammatical if *ham selv* is taken not to be a reflexive, but the pronoun *ham* followed by the emphatic marker *selv*. Like Hellan, I disregard such readings in the following.

However, Hellan (1988, Chapter 3) shows that in cases where *seg* appears to be bound locally, it is not a semantic argument of the verb. In example (2.74), for example, *seg* is not replaceable by any other NP; the sentence is parallel to the English *John perjured himself.* This is one example of the *non-argument* use of *seg.*

Other examples involve detransitivization, as proposed by Bresnan et al. (1985) as well as Hellan:

(2.75) døren åpnet seg langsomt
door opened self slowly
'The door opened slowly.'

In example (2.75), *seg* does not have argument status, but serves only as a marker that detransitivization has occurred. Examples of this sort have also been discussed by Sells et al. (1987) for Finnish, German, Dutch, and Serbo-Croatian.

Other types of examples occur in which *seg* appears to be bound locally; none involve the use of *seg* as an anaphoric element.

Seg in its anaphoric use is subject to the following constraints:

- It must be bound to a subject.
- It must be disjoint from its coarguments.
- It must be bound in the minimal tensed domain.

Like the anaphor *seg selv*, the antecedent of *seg* must be a subject:

(2.76) Jon hørte oss snakke om seg
Jon heard us talk about self
'Jon$_i$ heard us talk about self$_i$.'

(2.77) *jeg lovet Jon å snakke pent om seg
I promised Jon to speak nicely about self
'I promised Jon$_i$ to speak nicely about self$_i$.'

Example (2.76) is acceptable because the antecedent of *seg*, *Jon*, is a subject. Example (2.77) is not acceptable, however, because the binder of *seg*, *Jon*, is not a subject; the

subject of *lovet* 'promised' controls the subject position of *snakke* 'speak', and *Jon* bears only the grammatical function of object in this sentence.

Seg, like Marathi *aapaṇ*, must also be disjoint from its coarguments. The following examples, where *seg* is coindexed with a coargument, are ungrammatical:

(2.78) *Jon snakket om seg
Jon talked about self
'Jon$_i$ talked about self$_i$.'

(2.79) *Jon foraktet seg
Jon despised self
'Jon$_i$ despised self$_i$.'

The domain in which Marathi *aapaṇ*, must be bound is the root S: it is subject to the Root S Binding Condition. The binding domain of *seg* is more restricted, however. It must find a binder within the minimal finite domain in which it appears:

(2.80) *Jon var ikke klar over at vi hadde snakket om seg
Jon was not aware that we had talked about self
'Jon$_i$ was not aware that we had talked about self$_i$.'

Example (2.80) is unacceptable because the binder of *seg*, *Jon*, does not appear in the minimal finite domain in which *seg* appears.

In sum, binding conditions on Norwegian *seg* resemble those of Marathi *aapaṇ* to an extent. Their potential binders are different: *seg* is bound to the surface subject, while *aapaṇ* is bound to the logical subject.[13] However, both are required to be disjoint from coarguments, and both must be bound in a wider domain: *seg* within the minimal finite domain, and *aapaṇ* within the root S.

2.2.4 Possessive forms: *sin* and *hans*

The possessive anaphoric forms in Norwegian are the possessive reflexive *sin* and the possessive pronominal *hans*. *Sin* must have a subject as its antecedent, while *hans* requires disjointness from the immediately higher subject:

[13]See Chapter 5, Section 5.6.3, however, for discussion of reflexive binding in nominals in Norwegian; within nominals, a notion like 'logical subject' seems to play a role even for *seg*.

(2.81) Jon beundrer sin mor
Jon admires self's mother
*hans
his
'Jon$_i$ admires self's$_i$/his$_i$ mother.'

If the antecedent is a higher subject, but not immediately so, it is possible to use either *sin* or *hans*, although *sin* (or the neuter form *sitt*) is 'slightly preferred' (Hellan 1988, p. 132):

(2.82) Jon gjorde oss glad i huset sitt
Jon made us fond of house self's
?hans
his
'Jon$_i$ made us fond of his$_i$/self$_i$ house.'

Hans thus provides another example of a pronoun which obeys the Subject Disjointness Condition, the negative counterpart of the Subject Binding Condition.

The possessive reflexive *sin* obeys the Subject Binding Condition. A non-subject cannot be the antecedent of *sin*, as these examples show:

(2.83) Jon ble arrestert i sin kjøkkenhave
Jon was arrested in self's kitchen garden
'Jon$_i$ was arrested in his$_i$ kitchen garden.'

(2.84) *vi arresterte Jon i sin kjøkkenhave
we arrested Jon in self's kitchen garden
'We arrested Jon$_i$ in his$_i$ kitchen garden.'

Hellan (1988, p. 73) states that the domain in which *sin* must find an antecedent is the minimal finite domain, as it is for *seg*. Binding conditions on *sin* seem to be identical to those for *seg*: it must be bound in the minimal finite domain to a subject. The Coargument Disjointness Condition is not relevant for *sin*, but this is as expected; I will discuss this further in Chapter 5, Section 5.7.

Non-subjects are acceptable antecedents for *hans*:

(2.85) vi fant Jon under sengen hans
we found Jon under bed his
'We found Jon$_i$ under his$_i$ bed.'

(2.86) vi jaget Jon tilbake til huset hans
we chased Jon back to house his
'We chased John$_i$ back to his$_i$ house.'

Further constraints on reflexive possessives will be discussed in greater detail below, in 5, Section 5.6.3); there, it will be shown that a relation of *superiority* must hold between the antecedent of the anaphor and the noun phrase of which the anaphor is a possessor.

2.2.5 Anaphoric Binding in Norwegian

We have examined five anaphoric elements in Norwegian; these elements exemplify six different binding constraints. *Seg selv* obeys the Subject Antecedent Condition and the Coargument Binding Condition; its antecedent must be an argument of the same predicate. *Ham selv* obeys the Subject Disjointness Condition and the Minimal Complete Nucleus Binding Condition: its antecedent must appear in the minimal domain with a subject, and it must be disjoint from that subject. *Seg* obeys the Subject Binding Condition, the Coargument Disjointness Condition, and the Minimal Finite Domain Binding Condition; it resembles the Marathi long-distance reflexive *aapaṇ* in that it may not corefer with a coargument; its antecedent must be a subject of a higher clause, but it may not find its antecedent across a tensed boundary. Additionally, the possessive reflexive must be coreferent with a superior subject, while the possessive pronominal must be disjoint from the immediately superior subject. As in Marathi, then, binding conditions for Norwegian reflexives must be stated lexically, not with reference to Norwegian as a whole.

The binding conditions obeyed by the Norwegian anaphoric elements discussed above are summarized in the following table:

	Bound to:	Disjoint from:
seg selv	Subject in coargument domain	
ham selv	Argument in minimal complete nucleus	Subject in minimal complete nucleus
seg	Subject in minimal finite domain	Argument in coargument domain
ham		Subject in minimal complete nucleus
sin	Subject in minimal finite domain	

2.3 An Inventory of Binding Requirements

We have seen the need for the following requirements:

Subject Binding/Disjointness Condition: bound to or disjoint from a subject

Coargument Binding/Disjointness Condition: bound to or disjoint from an argument of the same syntactic predicate

Minimal Complete Nucleus Binding Condition: bound to an argument within the minimal domain containing a subjective function, SUBJ or POSS

Minimal Finite Domain Binding Condition: bound to an argument within the minimal tensed domain

Root S Binding Condition: bound to an argument within the S in which the anaphor appears

All of these binding requirements, except for the Root S Binding Condition, can be defined in terms of either SUBJ/POSS, PRED, or TENSE; the Root S Binding Condition is not actually a domain constraint, but the absence of any constraint on the

domain in which an antecedent must be found. Several conditions make reference to SUBJ/POSS: the Subject Binding and Disjointness Conditions require coreference with or disjointness from a subject, and the Minimal Complete Nucleus Binding Condition is defined as the minimal domain containing a subjective function. The Coargument Binding and Disjointness Conditions are defined in terms of a nucleus: a PRED and its arguments. The Minimal Finite Domain Binding Condition is defined in terms of the presence of TENSE. From this small set of syntactic concepts, then, the complete array of binding requirements can be defined.

Although there are seven syntactically definable conditions that can hold of anaphoric elements, there are not 128 ($=2^7$) possible sets of anaphoric binding requirements which anaphoric elements can obey, since it is impossible for some combinations to cooccur. For instance, the Coargument Binding Condition and the Coargument Disjointness Condition cannot both hold of the same anaphor; if they did, the anaphor would not be able to take an antecedent at all. Further, if a smaller domain is specified, then specifying a larger domain adds no new information: for instance, an anaphor bound in the minimal complete nucleus is necessarily bound in the minimal finite domain, so adding a Minimal Finite Domain constraint is redundant.

Other requirements also narrow down the set of possibilities. As we will discuss in Chapter 5, certain combinations of positive and negative requirements are universally unavailable. Both positive and negative constraints hold only in the positively-specified domain, so the domain of the negative constraints cannot be wider than the domain of the positive constraints.

Additionally, as we will discuss in Chapter 5, Section 5.3, there do not seem to be any long-distance anaphors that do not have a specified antecedent condition. This further disallows certain of these combinations.

We will discuss interactions between the constraints that are attested by Marathi and Norwegian in Chapter 5; there, it is shown that the number of possible combinations is smaller than it might appear.

2.4 Conclusion

The analysis presented here makes the claim that there is a certain universally-available, syntactically relevant set of binding constraints which anaphoric elements may obey. These constraints are definable in terms of the presence of a SUBJ, of a PRED, or of TENSE. They break down along two dimensions: domain conditions stand in opposition to antecedent conditions, and positive constraints stand in opposition to negative constraints.

Domain conditions specify that a pronoun must be bound or free in a particular, syntactically definable domain; the domain is universally characterizable in terms of the presence of a SUBJ, a PRED, or TENSE. Some anaphors must be bound, but the domain in which they are bound is unconstrained: this is the Root S Binding Condition. Antecedent conditions specify that a pronoun must be bound to or free from a SUBJ; some anaphors do not obey an antecedent condition and are bound to or disjoint from not only subjects, but arguments with any grammatical function.

Further, both domain and antecedent conditions can be positively specified: a pronoun may require an antecedent within a particular domain or of a particular syntactic type. Negative (disjointness) conditions also are found: a pronoun may require that no coreferent element appear within a certain domain, or that no coreferent element bear a particular grammatical function. Anaphoric elements may obey only positive conditions, only negative conditions, or a combination of the two kinds of conditions.

Chapter 3 will present a review of some previous work on anaphoric binding. The Norwegian and Marathi data will prove to be problematic for many of these analyses. Chapter 4 will show how the anaphoric binding constraints presented here can be expressed within the theory of LFG, using the technique of *functional uncertainty.*

Chapter 3

Previous Approaches to Anaphoric Binding

Approaches to anaphoric binding that take binding constraints to be universal or to be specified on a language-by-language basis encounter particular difficulties in describing binding constraints in languages with more than one anaphoric element, where different binding domains and antecedent conditions are associated with each anaphoric element. In the following, I will describe some of these approaches, showing how data from languages such as Norwegian and Marathi prove these analyses to be inadequate.

In contrast to these approaches, Bresnan et al. (1985) state binding constraints as features lexically associated with each anaphoric element. An approach like this one can correctly characterize binding conditions in languages with more than one anaphoric element, since binding conditions are taken to be properties of individual anaphoric elements.

Like Bresnan et al., I will state syntactic constraints on anaphoric binding as lexically-specified properties of individual anaphoric elements. However, rather than using features to specify these constraints, I will use the technique of *functional uncertainty*; equations involving functional uncertainty are associated with the lexical entry for each anaphoric element. This will be discussed in Chapter 4.

The constraints I am mainly concerned with in this work are those that are statable on a level of representation that carries information about surface syntactic relations, the *f-structure* of LFG. For that reason, I will not provide an analysis of phenomena requiring access to other levels of representation, such as *logophoricity*. A full analysis

of logophoricity of the sort that will be presented here would require access to a level of representation encoding discourse participants and discourse relations. I will, however, discuss the proposal of Pollard and Sag (1989); they propose that syntactic considerations fully determine when discourse constraints influence anaphoric binding in English. Their position will be shown to be too strong; discourse influences on anaphoric binding in English are found, but their applicability is not determined by surface syntactic conditions.

3.1 Government-Binding Proposals

3.1.1 Government-Binding Theory: Chomsky 1981

Chomsky (1981, p. 188) presents the following conditions, by now generally familiar, on binding in Government-Binding Theory (henceforth GB):

A. An anaphor is bound in its governing category.

B. A pronominal is free in its governing category.

C. An R-expression is free.

Some terminology: An R-expression (or 'referring expression') is a noun phrase that is not anaphoric; for example, a name. 'Bound' means coindexed with a c-commanding element. 'Free' means not bound.[1]

'Governing category' is defined in the following way (Chomsky 1981, p. 211):

> β is a governing category for α if and only if β is the minimal category containing α, a governor for α, and a SUBJECT accessible to α.

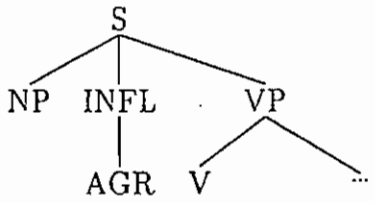
γ governs δ if γ assigns case to δ; often, the governor of α is the predicate that subcategorizes for α.[2] The governing category for α contains α and its governor; for subjects of tensed clauses, the governor is taken to be INFL, the inflectional element.

The basic tree structure for a sentence that is assumed in Chomsky (1981) is:

[1] I will not discuss conditions on empty categories such as trace or PRO.

[2] See Chomsky (1981, p. 162) for a more accurate characterization of the concept of government.

(3.1)



A SUBJECT accessible to γ is, again very roughly, the c-commanding subject of a nontensed S or the 'subject'/possessive element of an NP; for a tensed S, the SUBJECT is the set of agreement features of the tensed verb, referred to as AGR.[3] The concept of accessibility comes most importantly into play in determining the governing category for anaphoric elements which appear in subject position or in positions internal to the subject; this will be discussed in greater detail below, in Sections 3.1.2.1 and 3.1.2.2.

In many cases, the governing category for α is, then, the minimal category containing α, the category that subcategorizes for α, and either a subject or AGR. Often, the governing category is the minimal NP or S in which α appears.

According to Condition A of these definitions, the following sentence is grammatical:

(3.2) John$_i$ saw himself$_i$ (in the mirror).

The governing category for *himself* is the entire S, since that is the minimal category containing *himself*, its governor *saw*, and an accessible SUBJECT, the agreement features AGR of *saw*. *Himself* is required to be bound in that domain: that is, it must be coindexed with a c-commanding element which appears in the S; *John* is such an element.

This sentence is ungrammatical, though:

(3.3) *John$_i$ said that Mary saw himself$_i$ (in the mirror).

The governing category for *himself* is the subordinate S *Mary saw himself*, since this is the category containing *himself*, its governor *saw*, and an accessible SUBJECT, AGR. *Himself* is not bound in this domain, and the sentence is ungrammatical.

We turn now to Condition B. This sentence is ungrammatical:

[3]See Chomsky (1981, p. 211) for a more accurate characterization of the concept of accessibility.

(3.4) *John$_i$ saw him$_i$.

The governing category for *him* is the S, for reasons outlined above. *Him* is bound by *John* in this domain; pronominals are required to be free in their governing category, though, and so the sentence is ungrammatical.

This sentence is grammatical:

(3.5) John$_i$ said that Mary$_j$ saw him$_i$ (in the mirror).

The governing category for *him* is the subordinate S; *him* is free in this domain, and the sentence is grammatical.

Finally, this sentence is ungrammatical:

(3.6) John$_i$ saw John$_i$ (in the mirror).

Both instances of the name *John* are r-expressions. According to Condition C, r-expressions must be free, not just in their governing category but in the entire sentence. The second instance of *John* is bound by the first in this example, however, and the sentence is ungrammatical.

For present purposes, the salient features of this account — features which have since been shown to be undesirable by researchers working within the GB framework as well as other frameworks — are the following:

- Reflexives and reciprocals are taken to be in complementary distribution with pronominals. The governing category for any particular position is the same, whether it is filled by a pronominal, a reflexive, or a reciprocal. However, if a reflexive or reciprocal fills that position, it requires a binder within the governing category, while a pronominal requires that a binder not be present.

- Reflexives and reciprocals are anaphors, constrained by Principle A. They are taken to have exactly the same distribution in every language.

- Conditions A and B specify only domain requirements on anaphoric elements; that is, they specify the category within which an anaphor and its antecedent

must appear, or within which a pronominal and a coreferent element may not appear. Requirements as to the role of the antecedent (for example, that it must be a subject) are not specified.

- Binding conditions are taken to be universal. Reflexives and reciprocals in all languages are taken to obey Condition A, and pronominals in all languages are taken to obey Condition B.

Clear counterexamples to some of these claims were presented in Chapter 2. In particular, we have seen that neither of the two Marathi reflexives is in complementary distribution with the Marathi pronoun; data from both Marathi and Norwegian make it clear that antecedent conditions as well as domain conditions must be addressed in the analysis of anaphora; and many of the anaphoric elements in both Norwegian and Marathi obey neither Principle A nor Principle B. Some of the work to be discussed in this chapter provides further counterexamples. Below, I will discuss some efforts toward modifying and extending these conditions within GB.

3.1.2 Anaphoric Distribution: Noncomplementarity and Asymmetry

A problem for the binding theory as proposed in Chomsky (1981) is encountered when the relative distribution of pronominals, reflexives, and reciprocals violates one or another of the binding conditions. Two sorts of violations are attested.

First, there are many positions in which either a pronominal or an anaphor can appear; contrary to the statement of binding conditions in Chomsky (1981), pronominals and anaphors are not in complete complementary distribution. Huang (1983) discusses this fact and proposes a solution. A solution of a similar sort is proposed in Chomsky (1986). Both of these solutions are incomplete, since they treat only a small number of cases of noncomplementarity.

Second, although reflexives and reciprocals are both classified as anaphors, there are languages (including English) in which the distribution of the reflexive differs from the distribution of the reciprocal. Lebeaux (1983) discusses these phenomena for English and proposes an explanation.

3.1.2.1 Noncomplementary Distribution: Huang 1983

Huang (1983) notes that examples of the following sort present a problem for the binding theory as presented in Chomsky (1981):

(3.7) a. They$_i$ saw each other's$_i$ pictures.

b. They$_i$ saw their$_i$ pictures.

(3.8) a. They$_i$ expected that pictures of each other$_i$ would be on sale.

b. They$_i$ expected that pictures of them$_i$ would be on sale.

All of these examples are grammatical. The problem is that *each other* is an anaphor and *their* and *them* are pronominals; the binding conditions outlined above do not allow them in the same positions.

Chomsky (1981) proposes two different definitions of 'governing category', a preliminary version and the version which was presented above, in Section 3.1.1. As Huang points out, according to the definition of governing category presented above, only the sentences with *each other* should be grammatical. In example (3.7)a, the NP *each other's pictures* is not a governing category for *each other*, since the smallest category containing an accessible SUBJECT is the root S (the subject position of the NP *each other's pictures* is filled by *each other*, and a position is not accessible to itself). In example (3.8)a, the NP *pictures of each other* is not a governing category, since it does not contain a SUBJECT. The subordinate clause is also not a governing category, since AGR is not accessible to positions internal to a tensed clause subject. The governing category for *each other* in both cases is the root S, and *each other* is bound in this category. For this definition of governing category, the (b) sentences are problematic.

On the other hand, according to the preliminary definition of governing category presented in Chomsky (1981, p. 188), only the sentences with *them/their* should be grammatical. In his preliminary version, the governing category for an element is the smallest NP or S containing it and its governor. Given this definition, the governing category for *each other* is the NP *each other's pictures* in (3.7)a and the NP *pictures*

of each other in (3.8)a, and the reciprocal is not bound in these categories. But the pronominal is allowed in these positions under this definition, since it is free in the NP. For this definition of governing category, the (b) sentences are predicted to be grammatical, but the grammaticality of the (a) sentences is unexplained.

Huang's proposal is that the definition of 'governing category' should be different for anaphors and pronominals. His revised definition of 'governing category' follows (Huang 1983, p. 557, italics added):

> α is a governing category for β if and only if α is the minimal category containing β, a governor of β, and a SUBJECT that, *if* β *an anaphor*, is accessible to β.

The revised definition includes the stipulation that the notion of accessibility for the SUBJECT is only relevant for anaphors, but not for pronominals. This predicts noncomplementary distribution of pronominals and anaphors in certain cases.

For the example *They$_i$ expected that pictures of each other$_i$ would be on sale*, the governing category for *each other* is the entire sentence. This is because there is no SUBJECT in the lower clause that is accessible to *each other* — although the AGR of the lower clause is a SUBJECT, it is not accessible. *Each other* is thus required to be bound in the entire sentence; it need not be bound in the subordinate sentence. *They* is, for this reason, an acceptable binder.

Huang's definition of governing category also allows for the grammaticality of *They$_i$ expected that pictures of them$_i$ would be on sale.* The subordinate clause is the governing category for *them*, since it contains AGR, a SUBJECT. The fact that AGR is not an accessible subject is not relevant, since the governing category for a pronominal must contain a SUBJECT but not necessarily an accessible one. *Them* is free in the subordinate clause and thus satisfies Condition B.

Similar comments apply for the examples in (3.7), since noun phrases also contain SUBJECTs. For the example *They$_i$ saw each other's$_i$ pictures*, the governing category for *each other* is the entire sentence, since the SUBJECT of *each other's pictures* is not accessible to *each other.*[4] Since *each other* is bound in this category, the sentence

[4]See Huang (1983, p. 558) for discussion of this example.

is acceptable. On the other hand, for *their* in *They$_i$ saw their$_i$ pictures*, the governing category is the NP *their pictures*; the pronominal is thus acceptable, since it is free in that NP.

Huang's revision of the definition of governing category predicts noncomplementary distribution of pronominals and anaphors just in subject positions or in positions contained inside subjects. But there are cases of free variation between pronominals and anaphors that are not handled by this revision; consider:

(3.9) a. Mary$_i$ wrapped a blanket around her$_i$.

b. Mary$_i$ wrapped a blanket around herself$_i$.

Both of these sentences are grammatical, yet one contains a pronominal, *her*, and the other contains an anaphor, *herself*. On the assumption that no structural ambiguity is involved in this example, the governing category for both *her* and *herself* is the whole sentence. (3.9)b is predicted by Huang's theory to be grammatical, but (3.9)a is predicted to be ungrammatical.[5]

Further, data from Marathi and Norwegian provide many examples of noncomplementary distribution involving pronominals and anaphors. For example, in the following Marathi sentence, either the long-distance reflexive *aapaṇ* or the pronominal *ti* can appear:

(3.10) Jane ne swataahkartaa saaḍi gheṭ li
Jane ERG for self sari bought
aaplyaakartaa
for self
ticyaakartaa
for her
'Jane$_i$ bought a sari for herself$_i$/her$_i$.'

The solution proposed by Huang does not extend to these cases.

Another problem with Huang's analysis is that his data involve, to a large extent, 'picture nouns'. The sentences he cites include:

[5]The analysis of Bresnan et al. (1985), to be discussed in Section 3.4, allows for either the reflexive or the pronominal without positing a structural ambiguity.

(3.11) a. They$_i$ saw each other's$_i$ pictures.

b. They$_i$ saw pictures of each other$_i$.

c. They$_i$ expected that pictures of each other$_i$ would be on sale.

d. They$_i$ expected that it would be possible for pictures of each other$_i$ to be on sale.

However, picture noun anaphors have a notoriously wide distribution. As pointed out by Pollard and Sag (1989), they need not be c-commanded by their antecedent, and they may even appear in a different sentence from their antecedent:[6]

(3.12) a. The pictures of each other$_i$ with Ness made [Capone and Nitty]$_i$ somewhat nervous.

b. John's$_i$ intentionally misleading testimony was sufficient to ensure that there would be pictures of himself$_i$ all over the morning papers.

c. Sue$_i$ was sad. The picture of herself$_i$ in the school yearbook was very unflattering.

In these examples, as Pollard and Sag (1989) show, discourse factors influence the binding of *himself*. These facts will be discussed in more detail in Section 3.2. It would seem unwise to undertake a purely syntactic explanation of the distribution of picture noun reflexives, since their distribution often seems to be constrained by discourse factors.

3.1.2.2 Noncomplementary Distribution: Chomsky 1986

Chomsky (1986) notes the following case in which pronominals and anaphors are in noncomplementary distribution:

(3.13) a. The children$_i$ heard stories about each other$_i$.

[6]Some of these examples are from Pollard and Sag (1989).

b. The children$_i$ heard stories about them$_i$.

For this case, he posits a structural ambiguity. He assumes the noun phrase *stories about them* contains an unpronounced subject element; for this reason, the governing category for *them* is that noun phrase, and the pronominal is free in its governing category. The noun phrase *stories about each other* does not contain a subject, however; the governing category for *each other* is thus the whole sentence, and *each other* is bound in this category.

Evidence for this ambiguity is that the sentence *The children heard stories about them* has no reading paraphrasable as *The children$_i$ heard their$_i$ stories about them$_i$*. The unpronounced subject element of the noun phrase must be referentially disjoint from *the children*. This is what is predicted by the binding theory, since *them* is a pronominal and must be free in its governing category.

For another case, though, Chomsky does not posit a structural ambiguity, but revises the binding theory so either a pronominal or an anaphor can occur:

(3.14) a. The children$_i$ like each other's$_i$ friends.

b. The children$_i$ like their$_i$ friends.

Chomsky's analysis is in the spirit of Huang (1983), although a slightly different effect is produced; noncomplementary distribution of anaphors and pronominals is only predicted in the type of case presented in (3.14). He reformulates the binding theory in the following way (Chomsky 1986, p. 171, material in square brackets added):

> [An indexing] I is *BT-compatible* with [a pair consisting of an NP α and a domain β] (α,β) if:
>
> A. α is an anaphor and is bound in β under I
>
> B. α is a pronominal and is free in β under I
>
> C. α is an r-expression and is free in β under I

This is the basic definition for BT-compatibility. The following (re-)definition relies on the above one (material in square brackets added; the portion of the definition pertaining to r-expressions has been omitted):

> For some [anaphoric element] β such that [the condition in] (i) [holds], [an indexing] I is BT-compatible with [a pair consisting of a NP α and a domain β] (α,β):
>
> i. α is an anaphor or a pronominal and β is the least C[omplete] F[unctional] C[omplex] containing γ[, the governor of α,] for which there is an indexing J BT-compatible with (α,β)

Chomsky (1986, p. 169) states that:

> A governing category is a 'complete functional complex' (CFC) in the sense that all grammatical functions compatible with the head are realized in it — the complements necessarily ... and the subject, which is optional unless required to license a predicate, by definition.

Chomsky also states that a governing category can be only NP or S, since these are the only categories with subjects.

The intent of this reformulation is the following: an anaphor must be bound in the minimal NP or S in which it can possibly satisfy the binding conditions. That is, it must be bound in the minimal NP or S in which there is some indexing (not necessarily the actual one) under which the binding conditions are satisfied. A pronominal must be free in the minimal NP or S in which it can possibly satisfy the binding theory — in which there is some indexing under which binding conditions would be met.

For many examples, this reformulation gives the same results as the formulation in Chomsky (1981). For example, the governing category for ***himself*** in the following is the S:

(3.15) John saw himself.

This is because there is a possible indexing under which the binding theory can be satisfied: namely, where *himself* and *John* are coindexed. If they are coindexed, the reflexive is bound in its governing category, and the binding conditions are satisfied. Similarly, the governing category for *him* in the following example is the S:

(3.16) John saw him.

There is an indexing under which the binding theory can be satisfied: namely, where *John* and *him* have different indices. If this is the actual indexing, *him* is free in its governing category, and the binding conditions are again satisfied.

However, there is one case for which results obtain that are different from the 1981 binding conditions:

(3.17) a. The children$_i$ like each other's$_i$ friends.

b. The children$_i$ like their$_i$ friends.

There is no possibility for *each other* to satisfy the binding theory within the NP *each other's friends*, since there is no possible indexing whereby the binding theory could possibly be satisfied. For this reason, the governing category for *each other* is the entire sentence, and it is bound there.

There is, however, a possibility for *their* to satisfy the binding theory within the NP *their friends*, since there exists an indexing (namely, any indexing) under which the pronominal is free within the NP. In this case, then, the governing category for *their* is *their friends*; since *their* is free in that domain, the binding conditions are satisfied.

As is the case with Huang's analysis, other examples of noncomplementary distribution between pronominals and anaphors are attested; Chomsky's reformulation of the binding constraints does not address these cases. His reformulation is intended mainly to predict that the possessor position of an NP can be filled by either a pronominal or a reciprocal. Unconstrained, his reformulation seems to predict that both reflexives and reciprocals can appear as possessors of noun phrases, whereas in fact only reciprocals are permitted there. The fact that reciprocals and reflexives

cannot appear in subordinate clause subject position is also not predicted; Chomsky observes that sentences of the following type are unacceptable:

(3.18) a. *The children$_i$ thought that themselves$_i$ should go.

b. *The children$_i$ thought that each other$_i$ should go.

He gives two alternative explanations for the unacceptability of these examples. The first is that *themselves* and *each other* can satisfy the binding theory by being coindexed with AGR in the subordinate clause, and therefore their governing category is the subordinate clause and not the sentence as a whole. This explanation is weakened by the fact that AGR is not in general a proper antecedent for an anaphor, so the reason that coindexation with AGR is a possibility is somewhat mysterious.

Given the unsatisfactory nature of the first explanation, he advances a second explanation: Anaphors move to INFL at LF, leaving traces that must be properly governed. INFL does not itself constitute a proper governor, and so anaphors may not appear in tensed subject positions. Either stipulation seems to be more or less equivalent to a prohibition against reflexives or reciprocals in subject position in tensed clauses; this will be discussed in greater detail in the next section.

3.1.2.3 Reflexive and Reciprocal Differences: Lebeaux 1983

According to the theory of anaphoric binding presented in Chomsky (1981), both reflexives and reciprocals are anaphors, subject to Condition A. That is, reflexives and reciprocals are taken to have identical patterns of distribution.

Lebeaux (1983) notes that the distribution of reflexives and reciprocals is actually not identical. He presents the following set of data, noting a contrast in acceptability depending on whether a reflexive or a reciprocal is used (judgements are Lebeaux's):[7]

[7]Very similar examples to these have elsewhere been judged to be grammatical. For example, Pollard and Sag (1989) cite these examples as fully grammatical:

(a) John$_i$ wanted more than anything else for himself$_i$ to get the job.

(b) What John$_i$ would prefer is for himself$_i$ to get the job.

(3.19) a. [John and Mary]$_i$ brought some friends for each other$_i$ to meet.

b. ?? John$_i$ would like some books for himself$_i$ to read.

(3.20) a. It would please the boys$_i$ very much for each other$_i$ to win.

b. ?? It would please John$_i$ very much for himself$_i$ to win.

According to Lebeaux, the (b) examples are worse than the (a) examples; reciprocals seem to have a wider distribution than reflexives.

Lebeaux proposes to describe these facts in the following way:

(3.21) a. Reciprocals are subject to the binding theory.

b. Reflexives

1. are subject to the binding theory.
2. must be properly governed.

Chomsky (1981) provides several alternative definitions of proper government; one of them is the following:

(3.22) α properly governs β if and only of α governs β and α is lexical.

All of the definitions he gives entail that subject position of a tensed or nontensed clause is not properly governed. Lebeaux's description rests on this: subject positions are ruled unsuitable for reflexives (since reflexives must be properly governed) but suitable for reciprocals (since reciprocals need not be properly governed).

It seems to be quite common for anaphoric elements to vary as to whether they can appear in subject position of tensed and nontensed clauses. In some dialects of Marathi, a reflexive that in GB terms might be described as 'bound in its governing category' (though, as we have seen, this is not quite accurate) can appear in subject position of a tensed subordinate clause. This sentence, where the reflexive *swataah* appears in tensed subordinate clause subject position, is acceptable for some speakers, as indicated by the percent sign:

(3.23) %Jane laa waaṭṭe ki swataah sag|yaat sundar aaho
Jane DAT thought that self most beautiful was
'Jane$_i$ thought that self$_i$ was the most beautiful.'

In some languages, then, the same pattern of distribution that Lebeaux describes for *each other* — ability to appear in subject position of subordinate clauses — is also found for reflexives. In Lebeaux's terms, then, certain reflexives would also not be 'subject to the binding theory'.

In one sense, then, the difference between English *himself* and *each other*, then, should not be universally tied to a distinction between reflexives and reciprocals. Rather, the variation should be seen simply as a way in which binding conditions on anaphoric elements can vary; some anaphors may appear in subordinate clause subject position, and others may not. I will discuss reflexives which appear in subject position in Section 5.7. The difference Lebeaux describes can be taken as evidence that anaphoric binding conditions should be associated with particular lexical items rather than being stated universally. This is the position to be taken in this work.

However, Lebeaux proposes a deeper explanation for the data that he presents. Reflexives have a narrower distribution than reciprocals because reflexives are semantically vacuous at Logical Form; their meaning is totally determined by the meaning of their antecedents. Reciprocals, on the other hand, are not semantically vacuous; though they are like anaphors in that they must be syntactically bound, they also resemble quantifiers in that they have scope. In other words, reciprocals, unlike reflexives, are not simple anaphoric elements, since they have 'independent interpretive content' beyond the content of their antecedent.

For example, the contrast between example (3.24)a and (3.24)b can be explained by the fact that when semantic identity holds between the matrix and subordinate subject, as in (3.24)b, another form is available that carries the same meaning — namely, the form in (3.24)c:

(3.24) a. [John and Mary]$_i$ brought some friends for each other$_i$ to meet.

b. ?? John$_i$ would like some books for himself$_i$ to read.

c. John$_i$ would like some books to read.

The meaning expressed in (3.24)a is not expressible in another way, however.

Another semantic factor, unnoticed by Lebeaux, must also be taken into account in the analysis of the English reciprocal. There seems to be a wide range of dialect variation as to the interpretation of the English reciprocal. In particular, where the antecedent of a reciprocal is nonlocal, the reciprocal can be interpreted in some dialects of English in a manner similar to a pronoun.

Consider, for example, this sentence:

(3.25) [John and Bill]$_i$ said that Mary knew which of each other's$_i$ pictures was on sale.

This sentence is ungrammatical for some speakers of English. For others it is grammatical; for those speakers, three types of interpretations are attested, paraphrasable as:

(3.26) a. John said Susan knows which of Bill's pictures were on sale, and Bill said that Susan knows which of John's pictures were on sale.

b. Of John and Bill's pictures of each other (e.g. that they had painted of each other), John and Bill said Susan knows which were on sale.

c. Of John's pictures and Bill's pictures, John and Bill said Susan knows which ones were on sale.

I will refer to interpretation (3.26)a as the 'true reciprocal interpretation', since it involves a reciprocal relation between *John* and *Bill*.

The interpretation in (3.26)b is also a reciprocal interpretation; however, the reciprocal relation has been constructed locally, involving implicit arguments of the noun *picture*. I will refer to this as the 'local reciprocal interpretation'.

The interpretation in (3.26)c is more or less equivalent to the interpretation of the same sentence, where the reciprocal is replaced by a pronoun:

(3.27) [John and Bill]$_i$ said Susan knows which of their$_i$ pictures were on sale.

The difference is that the interpretation is necessarily distributive; a better paraphrase might be:

(3.28) [John and Bill]$_i$ said Susan knows which of the pictures of each of them$_i$ were on sale.

I will refer to this interpretation as the 'distributed pronoun interpretation'.

Notably, speakers for whom a nonreciprocal, distributed pronoun interpretation is available more readily allow the reciprocal to appear in environments where there is no local binder. Additionally, speakers for whom the local reciprocal interpretation is available permit the construction of a local reciprocal relation when such a relation is otherwise unavailable; these speakers, too, allow the reciprocal to appear in environments in which there is no local binder.

In brief, the distribution of the reciprocal is determined at least in part by the interpretations that speakers can give to it. Speakers who do not require a true reciprocal interpretation allow the reciprocal in a wider range of environments. It is evident, then, that interpretation of sentences involving the reciprocal should be taken into account when determining its binding domain, since these interpretations can affect binding possibilities.

3.1.3 Antecedent Conditions

Another problem for the theory presented in Chomsky (1981) is that antecedent conditions for anaphoric elements are not addressed. It is quite common for an anaphoric element to require a subject as its antecedent; this is the Subject Binding Condition. The Norwegian reflexive *ham selv* illustrates a requirement of disjointness from subjects, the Subject Disjointness Condition. Vikner (1985) proposes a factorization of binding conditions into those that specify domain requirements and those that specify antecedent requirements.

3.1.3.1 Domain and Antecedent Factorization: Vikner 1985

Vikner (1985) presents an analysis in which both possible binders and possible binding domains are taken into account. An anaphor can be a 'binder-pronominal', a 'binder-anaphor', or neither; it can also be a 'domain-pronominal', a 'domain-anaphor', or neither. The definitions of these categories follow (Vikner 1985, p. 19):

(3.29) a. Binder-pronominals:

A binder-pronominal is not bound by a subject inside its SUBJECT domain (i.e. inside the c-command domain of its lowest c-commanding SUBJECT).

b. Binder-anaphors:

A binder-anaphor is bound by a subject inside its AGR domain (i.e. inside the c-command domain of its lowest c-commanding AGR).

c. Domain-pronominals:

A domain-pronominal is not bound in its SUBJECT domain (i.e. in the c-command domain of its lowest c-commanding SUBJECT).

d. Domain-anaphors:

A domain-anaphor is bound in its SUBJECT domain (i.e. in the c-command domain of its lowest c-commanding accessible SUBJECT).

Anaphors subject to the binder parameter must be either bound to (in the case of binder-anaphors) or free from (in the case of binder-pronominals) a subject. Anaphors subject to the domain parameter must be either bound (in the case of domain-anaphors) or free (in the case of domain-pronominals) in the domain of the SUBJECT.

An element may be simultaneously a domain-pronominal and a binder-anaphor; this would mean that it must be bound by a subject in the minimal *tensed* S, but it must be free in the minimal (*non-tensed*) S. This kind of element would, then, appear in subordinate infinitival complements and would take a matrix argument as its antecedent; Vikner shows that Danish *sig* is such an element.

Vikner presents convincing argumentation that the four anaphoric elements in Danish exhaust the combinatoric possibilities of these features in the following way:

(3.30) a. *ham, hende*: binder-pronominal and domain-pronominal

b. *sig*: binder-anaphor and domain-pronominal

c. *ham selv, hende selv*: binder-pronominal and domain-anaphor

d. *sig selv*: binder-anaphor and domain-anaphor

He also shows that there are lexical items in Danish which lack specification in the 'binder' category or in the 'domain' category: *hans* is a binder-pronominal; *sin* is a binder-anaphor; *mig* is a domain-pronominal; and *mig selv* is a domain-anaphor.

Data to be discussed in Chapter 5 show that domain and antecedent constraints can interact; Vikner allows for this interaction by combining domain and antecedent requirements under a single heading. For example, the definition for 'binder-anaphor' refers simultaneously to the grammatical function of the binder and to the domain in which the binder must be found. However, Vikner's approach does not predict interactions between the negative requirements and the positive requirements; as we will see, interactions of this type are also found.

Vikner's analysis is, in other respects, close to the analysis that is presented here. His analysis is different from many presented within the GB framework in two ways: (1) it allows for specification not only of the domain requirements of an anaphoric element but also for specifications as to the type or role of its proper antecedent (that it must be or may not be a subject), and (2) it allows for different anaphors in the same language to have different binding requirements. For these reasons, his analysis can characterize binding constraints in Danish, which has more than one anaphoric element.

However, Vikner's analysis is not without its problems. For example, Vikner provides no way of describing an anaphor which may be bound outside a minimal tensed boundary, such as the long-distance reflexive in Marathi. The Marathi reflexive *aapaṇ* is not describable in Vikner's system.

The reflexive *aapaṇ* must be noncoreferent with its coarguments; its antecedent can be a subject in a higher clause, even when a tensed clause boundary separates *aapaṇ* from its antecedent:

(3.31) Sue ne John laa vinanti keli ki tyaane aaplaa kavitaa vaacu nayet
Sue ERG John DAT request did that he-ERG self-GEN poems read shouldn't

'Sue$_i$ requested of John that he not read self's$_i$ poems.'

As discussed in Chapter 2, Section 2.1.7, it is not possible for *aapaṇ* to take a discourse antecedent. It must be bound within the sentence in which it appears:

(3.32) Jane dukhi hoti. *aaplyaalaa jaataa aale naahi.
Jane sad was Self-DAT go could not
'Jane was sad. *Self could not go.'

In sum, Vikner's inventory of possible binding conditions does not seem to be complete, though his conditions are sufficient to describe anaphoric binding phenomena in Danish, a language with multiple anaphoric elements. The fact that Vikner discusses antecedent requirements as well as domain requirements is also a point in favor of his analysis; however, his analysis does not predict the full range of interactions among these requirements.

3.1.4 Domain Variation: Lexicalization and Parametrization

The binding conditions which Chomsky (1981) presents are taken to be universal: pronominals in every language are predicted to have the same distribution, and the same is true for reflexives and reciprocals. However, there are a number of cases of anaphoric elements that are subject to different binding conditions than those described above.

Long-distance reflexives — reflexives that are bound by an element outside their governing category — are an example. Huang (1982) proposes a solution to this problem for the Chinese reflexive *ziji*; the solution he presents will be shown to be not generally applicable. Sportiche (1986) proposes to allow for long-distance reflexivization by modifying the classification of anaphoric elements; the distribution of

anaphoric elements is described by a system which makes more fine-grained distinctions than the anaphor/pronominal choice.

Iatridou (1986) notes that the binding conditions of Chomsky (1981) do not allow for the distribution of the Greek anaphoric element *o idhios*; she makes a proposal to add a new condition, Condition D, to the binding theory. Her proposal will prove to be not generally applicable, though it is descriptively more adequate than the binding theory of Chomsky (1981) in that it allows for a language that has more than one reflexive, each with a different domain.

Other proposals have been made which involve parametrizing the binding conditions; under these approaches, some aspects of the binding conditions would vary from language to language or for individual anaphoric elements. In particular, Yang (1983) and Manzini and Wexler (1987) have proposed ways of parametrizing the binding theory. We will see that Yang's proposal is inadequate in that it assumes that parameters are set for a language as a whole rather than for a particular lexical item. Manzini and Wexler's proposal is different, however. Their proposal is similar to the one to be proposed in this work in that they assume that parameters are relevant not for languages as a whole but for particular lexical items.

3.1.4.1 Long-Distance Reflexivization: Huang 1982

One problem for an account like the one in Chomsky (1981) is the existence of long-distance reflexivization, where a reflexive may be bound by an element outside its governing category. Huang (1982) contains a brief discussion of long-distance reflexivization within the GB framework. The Chinese reflexive *ziji* can have an antecedent which is the subject of a subordinate tensed clause:

(3.33) (Huang's (47a), p. 331)

Zhangsan shuo ziji hui lai
Zhangsan say self will come
'Zhangsan$_i$ said that self$_i$ will come.'

The pronominal *ta* is also acceptable in this position:

(3.34) (Huang's (47b))

Zhangsan shuo ta hui lai
Zhangsan say he will come
'Zhangsan$_i$ said that he$_i$ will come.'

Ziji is a reflexive, so it should be bound in its governing category; the governing category for tensed subordinate clause subjects is the subordinate clause, and *ziji* can take an antecedent outside that category. Huang proposes to explain the fact that the Chinese reflexive may appear in subordinate clause subject position by saying that the verbal inflection INFL in Chinese does not contain AGR; in fact, there is no agreement morphology on tensed verbs in Chinese. This means that there is no SUBJECT accessible to the subject position in the lower clause (the subject is not accessible to itself), and thus the governing category for the subordinate clause subject is the matrix S, not the subordinate S. In this way, Huang predicts that a reflexive may appear in the subject position of a subordinate clause, even when it is bound by an antecedent in a higher clause.

For non-subject positions, the subject of the subordinate clause counts as the accessible SUBJECT; thus, the governing category for non-subject positions in a subordinate clause is the subordinate clause itself. The Chinese reflexive *ziji* may not appear in non-subject subordinate clause position unless it is bound within the subordinate clause.

There are two problems for an account like this one. First, this account correlates the possibility for long-distance reflexivization with the presence or absence of AGR in tensed clauses. This correlation does not seen to be legitimate, though. As we saw in Chapter 2, Marathi is a language with a long-distance reflexive; Marathi tensed verbs bear agreement morphology, however, and would therefore be expected to have AGR in tensed clauses.

Second, the analysis does not account for long-distance reflexives that can appear in subordinate clause nonsubject position. In any position other than subject position in the lower clause, the governing category for a reflexive is the lower clause itself, and Huang would predict that in non-subject position in a lower clause, reflexivization should not be possible with a higher-clause antecedent. However, Marathi and

Norwegian have reflexives that can appear in subordinate clause nonsubject position, even when their antecedent is the matrix clause subject.

3.1.4.2 Long-Distance Reflexivization: Sportiche 1986

Sportiche (1986) presents an analysis of the Japanese long-distance anaphoric element *zibun* according to which it has properties both of anaphors and of pronominals. He provides the following table:

	C-command required	C-command not required
Locality condition	Anaphors	*
Antilocality condition	Pronouns as variables	Referential pronouns

'Anaphors' are, as above, reflexives and reciprocals. 'Referential pronouns' are, roughly, pronouns that do not have a c-commanding antecedent. 'Pronouns as variables' are pronouns that are c-commanded by their antecedent.

On Sportiche's analysis, anaphors are subject to a locality condition: they must be bound in their governing category. Pronouns as variables and referential pronouns are subject to an antilocality condition; they must be free in their governing category. On a second dimension, both anaphors and pronouns as variables must be c-commanded by their antecedents; referential pronouns need not be.

The distinction between pronouns as variables, which must be c-commanded by their antecedent, and referential pronouns, which need not be, is often taken to be motivated by two factors. First, it is claimed that pronouns whose antecedent is a quantified NP must be c-commanded by that NP. Counterexamples to this claim have been presented by Reinhart (1983):

(3.35) Every boy's$_i$ mother thinks he$_i$ is a genius.

The quantified NP *every boy* does not c-command the pronominal *he*; nevertheless, the anaphoric relation is possible.

Second, facts concerning strict and sloppy readings under ellipsis are cited, and the claim is made that sloppy readings are only available when the antecedent c-commands the pronoun. However, Wescoat (1989) provides evidence that this generalization is incorrect: sloppy readings are available even when the antecedent does not c-command the pronoun, as in the following:

(3.36) The policeman who arrested John failed to read him his rights, and so did the one who arrested Bill.

A sloppy reading is available here, under which the policeman who arrested Bill failed to read Bill's rights to Bill. Dalrymple et al. (1989) provide an analysis of ellipsis in which the distinction between pronouns as variables and referential pronouns is unnecessary in obtaining strict and sloppy readings; it is not entirely clear, then, that the distinction should be maintained. Without a semantic basis to motivate the distinction, the category 'pronouns as variables' can be assumed to be simply shorthand for a syntactic requirement of c-command by an antecedent.

The GB analysis of English is that the locality and antilocality conditions distinguish between anaphors and pronominals. Sportiche's proposal is that *zibun* occupies the first column of the table rather than occupying a row; that is, *zibun* is both an anaphor and a bound-variable pronoun. It requires only a c-commanding antecedent, without constraining the antecedent to appear in a particular syntactic domain.

This analysis assumes that the distribution of *zibun* is syntactically constrained. Iida (1990) provides evidence that discourse factors determine the binding of *zibun* — that its distribution is not constrained by syntactic factors such as c-command. However, even if Sportiche's proposal faces difficulties in describing the full range of properties of *zibun*, it is interesting as an example of how the binding constraints presented in Chomsky (1981) can be modified to cover a broader range of data.

Sportiche's analysis is different from Chomsky's in that the two-way distinction between anaphors and pronominals is relaxed. That is, Sportiche's theory allows for at least three kinds of anaphoric elements rather than Chomsky's two. The English pronominal covers the 'pronouns as variables' and the 'referential pronouns' types,

while the English reflexive covers the 'anaphors' type; in contrast, Japanese *zibun* covers the 'anaphors' and 'pronouns as variables' type.

Sportiche's theory, like the analysis presented by Vikner (1985), seems to be a step toward allowing lexical specification of binding requirements. An anaphoric element can 'lexicalize' (in Sportiche's terminology) any of the three types he discusses; in fact, it can lexicalize any row or column of the table. (I assume Sportiche would not expect to encounter an anaphoric element that lexicalizes on the diagonal.) For Sportiche, although binding requirements like 'c-command required' and 'locality condition' are part of the universal inventory of possible binding conditions, it is possible to specify these requirements independently. Sportiche's system, then, tends toward the kind of analysis to be presented here, where binding requirements are taken to be lexically associated with each anaphoric element.

However, we will see that the inventory of possible binding domains is more varied than is allowed for in Sportiche's analysis. Like Vikner's analysis, discussed above in Section 3.1.3.1, the inventory of possible anaphoric elements allowed for in Sportiche's analysis is not rich enough. His analysis allows only for anaphors that are bound in their governing category (the 'anaphors' type) and those that are bound in the entire S (the 'pronouns as variables') type. Other binding domains are attested, as the data from Marathi and Norwegian illustrate.

3.1.4.3 Additional Binding Conditions: Iatridou 1986

Iatridou (1986) presents an analysis of the Greek anaphoric element *o idhios*. According to Iatridou's analysis, *o idhios* must be bound, but it may not be bound to an antecedent that is 'too close' in some sense. In the following sentence, the antecedent of *o idhios* must be *Yanis* and not *Maria*:

(3.37) O Yanis theli i Maria na voithisi ton idhio
John wants Mary helps himself
'John$_i$ wants Mary to help him$_i$.'

Example (3.38) is ungrammatical; the antecedent of *o idhios* may not be *Yanis*:

(3.38) *O Yanis aghapa ton idhio
John loves himself
'John$_i$ loves himself$_i$.'

There is another reflexive in Greek, *ton eafton tou*, which 'obeys binding condition A', according to Iatridou; if that reflexive is used in (3.38), the sentence is grammatical.

To account for the distribution of *o idhios*, Iatridou adds a new condition to the set of binding parameters:

(3.39) Condition D: bound in the whole sentence but free in the governing category

According to this condition, *o idhios* must be free in (roughly) the minimal S or NP in which it appears, but it must be bound by a higher antecedent.

Iatridou's analysis is confined to cases in which *o idhios* is syntactically bound. Evidence presented by Condoravdi (1989) seems to show that, contrary to Iatridou's claim, *o idhios* can be used even in cases in which it is not bound by its antecedent. Among these cases are those in which *o idhios* is used contrastively:

(3.40) O Yanis arnithike na mu kani ti hari
John-NOM refused-3SG subjunctive me-GEN do-3SG the favor
an ke o idhios poles fores mu ihe zitisi paromies
although self-NOM many times me-GEN had-3SG asked similar
'John$_i$ refused to do me this favor, although self$_i$/he$_i$ himself had asked me for similar favors many times.'

(3.41) O Yanis mas simulepse na figume amesos.
John-NOM us-ACC advised-3SG subjunctive leave-1PL immediately
O idhios tha efevge argotera.
self future leave-3SG later
'John$_i$ advised us to leave immediately. Self$_i$/he$_i$ (on the other hand) was to leave later.'

Iatridou discusses similar examples, claiming that examples such as these illustrate the 'emphatic' or the 'adjectival' use of *o idhios*; she glosses the 'adjectival' use as 'the same'. Her analysis, then, treats only the anaphoric, syntactically bound use of *o idhios*.

Iatridou's solution has the advantage of being able to characterize both reflexives in Greek; one of the reflexives obeys condition A, while the other obeys condition D. Iatridou's solution is also a move toward lexicalization of binding category specifications; on her analysis, an element can be either bound or free in the whole sentence, and it can be either bound or free in its governing category. Furthermore, different reflexives in the same language can be subject to different binding requirements.

Recall that Marathi resembles Greek in having two reflexive elements. One, *aapaṇ*, has a distribution similar to that of *o idhios*. It must be bound by an antecedent in the same sentence; it may not take a discourse antecedent. However, the following sentence, where *aapaṇ* is bound by an antecedent that is 'too close', is ungrammatical:

(3.42) *Jane ne aaplyaalaa bockaarle
Jane ERG self-ACC scratched
'Jane$_i$ scratched herself$_i$.'

A similar sentence, where the short-distance reflexive *swataah* is used in place of *aapaṇ*, is grammatical:

(3.43) Jane ne swataahlaa bockaarle
Jane ERG self-ACC scratched
'Jane$_i$ scratched herself$_i$.'

The distribution of *aapaṇ* is, then, strikingly similar to that of *o idhios*. The problem for Iatridou's analysis is that the solution she outlines will not work for Marathi without modification.

In one dialect of Marathi, the short-distance reflexive *swataah* has a distribution similar to that of the English reflexive *himself*; in GB terminology, *swataah* would be subject to Condition A of the binding theory. The distribution of *aapaṇ* is somewhat broader than that of *o idhios*, however; in particular, it is possible for *aapaṇ* to be bound in the same sentence in which its antecedent appears:[8]

(3.44) a. Jane ne aaplyaakartaa saaḍi gheṭ li
Jane ERG for self sari buy take
'Jane$_i$ bought a sari for herself$_i$.'

[8]Iatridou does not discuss the distribution of *o idhios* in sentences of this type, so it is not clear whether similar problems would arise for her analysis with regard to the Greek data.

b. Jane ne John laa aaplyaabaddal maahiti dili
Jane ERG John DAT about self information gave
'Jane$_i$ gave John information about herself$_i$.'

Assume that *aapaṇ* is subject to condition D, and is free in its governing category. Recall, though, that under the definitions of governing category given in Chomsky (1981) and Chomsky (1986), NP and S are the only possible categories which can be governing categories; further, only NP's that contain a SUBJECT — a possessive NP — count as a governing category. Notice that by these definitions, the NP 'information about self' does not count as a governing category for the reflexive *aapaṇ*; the S is the governing category for the reflexives in the examples above, and the long-distance reflexive *aapaṇ* should not be allowed in the examples in (3.44). Even if we define 'governing category' in such a way as to exclude the phrases *for herself* and *about herself* in the above sentences, it is less than clear how to characterize the domain in which the short-distance reflexive *swataah* must be bound. It is not the case that *swataah* is in complementary distribution with *aapaṇ*, since the following examples, identical to those given above except for the replacement of *aapaṇ* by *swataah*, are grammatical:

(3.45) a. Jane ne swataahkartaa saaḍi gheṭ li
Jane ERG for self sari bought
'Jane$_i$ bought a sari for herself$_i$.'

b. Jane ne John laa swataahbaddal maahiti dili
Jane ERG John DAT about self information gave
'Jane$_i$ gave John information about herself$_i$.'

Thus it is not possible to simultaneously describe the domain in which *swataah* must be bound and the domain in which *aapaṇ* must be free by making reference to the notion of governing category, since their distributions partially overlap.

The problem outlined here, that of overlap between domains in which anaphoric elements must be bound and domains in which pronominal elements must be free, was discussed above, in Section 3.1.2. It was mentioned there that certain cases of noncomplementarity of pronominals and anaphors were not covered by the solutions

presented by Huang and Chomsky and discussed there; among these are the cases described in this section. Neither Huang's nor Chomsky's revisions to the binding theory allow for such cases, since these revisions were intended only to explain cases of reflexives in or internal to subject positions.

Iatridou's analysis, like the analyses presented by Vikner and Sportiche, allows for different reflexives with different domains within the same language; in this, it describes the facts more successfully than the one presented in Chomsky (1981). However, the fact that her analysis uses the same notion, 'governing category', to describe both the binding domain and the disjointness domain for all anaphors makes it unable to correctly characterize the binding conditions for other, superficially similar anaphors such as Marathi *aapaṇ*.

Binding patterns of *aapaṇ* and *o idhios* also indicate that, as we have noted, a division of anaphoric elements into 'pronominals' and 'anaphors' is not an adequate characterization of the range of anaphoric elements that are found in natural language. Both *aapaṇ* and *o idhios* are like anaphors in that they must be bound by an element standing in a certain structural relation; they are also like pronominals in that they must be free from elements in a particular domain. The variety of attested anaphoric elements is greater than allowed for by a simple two-way distinction between 'bound' and 'free' elements.

3.1.4.4 Parametrized Binding: Yang 1983

Yang (1983) presents an analysis of anaphoric binding within the GB framework which takes into account cross-linguistic variation in the domain in which anaphors must be bound.

Yang's analysis involves the existence of *unmarked* and *marked* anaphors and pronominals. Unmarked anaphors are bound in the minimal domain containing a subject or AGR. An NP with a SUBJECT/possessor, a tenseless clause, or a tensed clause are all binding domains for unmarked reflexives. The domain of marked anaphors is wider, however: a marked anaphor is an anaphor whose domain is larger than that of an unmarked anaphor. For a marked anaphor, the binding domain is determined

only by reference to AGR, not by reference to a subject. A marked anaphor may be felicitously bound by an antecedent that appears in a higher clause, for example, as long as a boundary of a certain kind does not intervene.

Yang's Anaphor-Binding principle is as follows:

(3.46) (Yang's (80)): Anaphor Binding Principle

1. An anaphor is bound in the c-domain of its c-commanding minimal SUBJECT.
2. SUBJECT is parametrized.
   (a) SUBJECT = AGR *or subject* for unmarked binding (reciprocals, unmarked reflexives).
   (b) SUBJECT = AGR *only* for marked binding (marked reflexives).
3. AGR for marked binding is parametrized.
   (a) AGR = INFL of a finite clause (for Russian, etc.)
   (b) AGR = INFL of an indicative clause (for Icelandic, etc.)
   (c) AGR = COMP (for Dutch, etc.)
   (d) ...

Yang states that for unmarked binding the unmarked domain is the relevant one — the c-domain of the c-commanding minimal SUBJECT. The unmarked definition of SUBJECT is:

> A SUBJECT is AGR or the subject of an infinitive, a gerund, an NP, or a small clause.

The concept of 'c-domain' is defined in Manzini (1983) as follows:

> A is the c-domain of B if A is the minimal maximal category dominating B.

Thus if we take the subject to be SUBJECT (the possibility Yang presents for unmarked binding), the minimal maximal category containing the subject is the minimal S — since S is the smallest maximal projection that contains the subject. Similarly, if we take INFL to be SUBJECT, the minimal maximal category containing INFL is the minimal *tensed* S. If we take the Icelandic possibility of letting INFL of an indicative clause be SUBJECT, the minimal maximal category is the minimal *indicative* S.

An unmarked anaphor must, then, be bound in the minimal S or NP containing it (the minimal maximal category containing the anaphor and a subject or AGR). In Russian, a marked reflexive must be bound in the minimal *finite* S containing it. In Icelandic, a marked reflexive must be bound in the minimal *indicative* S containing it. In Dutch, a marked reflexive must be bound in the minimal clause containing a COMP. In this way, it is possible to specify different binding domains for different reflexives in different languages.

As Yang notes, this analysis differs from the standard GB analysis in several respects: for example, no mention is made of governor or governing category. Here, the relevant category within which an anaphor is bound is determined by the SUBJECT, which is defined in different ways for different reflexives and for different languages.

We will evaluate Yang's analysis with reference to Marathi. Recall that Marathi has a long-distance reflexive *aapaṇ* and a short-distance reflexive *swataah*. The long-distance reflexive *aapaṇ* must be free from coarguments but must be bound in the root S: it may not take a discourse antecedent.

The first question is what Yang's analysis of *aapaṇ* would be. Yang's analysis of reflexive anaphora in Korean, Japanese, and Kannada involves the claim that there is no AGR in these languages. If this is assumed, he says, there is no binding domain for anaphors, since the binding domain is defined in terms of AGR. This is why long-distance reflexivization is possible in these languages — the reflexive must be bound by a subject, but there is no domain requirement, so the reflexive can be bound by any higher subject. The only requirement is that it be bound.

The analysis is problematic for Marathi in the same way that it is for Kannada, however. In Korean and Japanese, the lack of AGR that Yang posits correlates with a lack of agreement morphology on the verb. However, in both Kannada and Marathi,

verbs do bear agreement morphology.[9] Yang asserts, though, that the presence of agreement morphology on the verb does not have any necessary connection to whether or not a language has AGR. His notion of AGR is thus relatively abstract, since it does not correlate with the presence of agreement morphology; let us take his notion of AGR merely as a shorthand for describing various binding possibilities. If we assume that there is no AGR in Marathi that is relevant for the binding theory, we can explain why *aapaṇ* can be bound to a higher subject even when it appears in a subordinate tensed clause.

But this assumption makes it impossible for Yang's account to describe the distribution of the short-distance reflexive *swataah*. As we have seen, *swataah* appears when its antecedent is in the same clause:

(3.47) Jane ne swataahlaa bockaarle
Jane ERG self-ACC scratched
'Jane$_i$ scratched herself$_i$.'

The situation in sentences with infinitival subordinate clauses is more interesting. Recall that there is a dialect of Marathi in which the following sentence is acceptable on either of the indicated indexings:

(3.48) John ne Jane laa swataahlaa maraaylaa saangitle
John ERG Jane DAT self-ACC hit told
'John$_i$ told Jane$_j$ to hit self$_{i,j}$.'

However, for all speakers, *swataah* may not appear in a tensed subordinate clause if its antecedent is in the matrix clause:

(3.49) *John mhanat hota ki Sue ni swataahlaa maarle
John said that Sue ERG self-ACC hit
'John$_i$ said that Sue hit self$_i$.'

[9]Recall that the verb does not always agree with the subject in Marathi; in some cases, the verb agrees with the object, and in some cases it shows neutral agreement. Presence of absence of agreement morphology does not correlate with varied reflexivization possibilities, though. It is not possible for the antecedent of a reflexive to be an object, even in cases where the verb agrees with the object.

For these speakers, *swataah* must be bound in the minimal finite clause in which it is contained.

In Yang's terms, there are two possibilities for *swataah*: it may be a marked reflexive or an unmarked one. If it is a marked reflexive, its distribution depends only on the presence of AGR, not of a subject. However, to account for the distribution of *aapaṇ*, we have assumed that there is no AGR in Marathi. If there is no AGR, *both* reflexives should be free to have long-distance antecedents; in fact, however, the short-distance reflexive *swataah* may not have antecedents in higher clauses.

The other possibility is that *swataah* is an unmarked reflexive. However, example (3.48) provides evidence against that possibility. Unmarked reflexives are bound in the minimal domain containing a subject. For example (3.48), that domain is the infinitival complement, and *John* does not appear in that domain. However, for at least some speakers, *John* may antecede *swataah* in this example.

The central problem for Yang's analysis seems to be that the domain of reflexivization is taken to be relevant for a language as a whole rather than for a particular lexical item. The analysis presented by Manzini and Wexler differs from Yang's in this respect; we will now turn to an examination of their analysis.

3.1.4.5 Parametrized Binding: Manzini and Wexler 1987

Manzini and Wexler (1987) present an approach to parametrization of binding constraints. For example, possible domains in which an anaphoric element must be bound or free include:

1. the minimal category containing a subject
2. the minimal category containing an Infl
3. the minimal category containing Tense
4. the root S.

An anaphor may also be marked as to its proper antecedent:

(3.50) A proper antecedent for [an anaphor] α is

1. a subject β; or
2. any element

An anaphor may require that its antecedent be a subject, or it may be bound by any element. A pronominal may also require disjointness from subjects or from every higher element.

According to Manzini and Wexler, these constraints are specified for individual lexical items, according to what they call the 'Lexical Parametrization Hypothesis' [p. 424]:

(3.51) Values of a parameter are associated not with particular grammars but with particular lexical items.

In spirit, their analysis is very similar to the analysis to be presented here, although there are substantial formal and substantive differences. In both analyses, binding constraints — among them, constraints on what Manzini and Wexler would refer to as the 'governing category' of an anaphoric element — are lexically associated with individual anaphoric elements. For this reason, the data presented in Chapter 2 would not be problematic for their analysis; for most other GB analyses, where anaphoric constraints are given universally or language-by-language, the data presented there are problematic.

There are particular respects in which Manzini and Wexler's analysis differs from the one to be presented here, however. Further, in several interesting ways it resembles the analysis of Pollard and Sag (1989). I will discuss their analysis at greater length in the next section.

3.2 Anaphoric Binding in English

On the analysis of Bresnan et al. (1985), to be discussed in Section 3.4, the English reflexive *himself/herself* must be bound in the minimal domain containing it and a

subject. Similarly, the Specified Subject Condition of Chomsky (1973) permits reflexivization in the minimal domain containing a subject. Here I refer to this condition as the Minimal Complete Nucleus Binding Condition. This analysis is supported by examples like the following:

(3.52) John$_i$ saw himself$_i$.

(3.53) *John$_i$ told Bill to look at himself$_i$.

(3.54) *John$_i$ said that Bill saw himself$_i$.

Example (3.52) is grammatical because the minimal domain containing *himself* and a subject also contains the antecedent of the reflexive, *John*. Examples (3.53) and (3.54) are ungrammatical because the minimal domain containing the reflexive and a subject is the subordinate clause *Bill saw himself*, and *himself* is not bound in that domain.

This analysis predicts that if the reflexive appears in a domain that does not contain a subject, the reflexive can be bound in a wider domain, since the domain containing a subject is not the minimal domain containing the reflexive. This prediction also appears to be borne out:

(3.55) John$_i$ wrapped a blanket [around himself$_i$].

The PP *around himself* does not contain a subject. The minimal finite domain containing the reflexive and a subject is the entire S; *himself* is bound in this domain, and the sentence is grammatical.

Similarly, the English reciprocal *each other* is bound in the minimal complete nucleus:

(3.56) [John and Bill]$_i$ saw each other$_i$.

(3.57) *[John and Bill]$_i$ told Mary to look at each other$_i$.

(3.58) *[John and Bill]$_i$ said that Mary saw each other$_i$.

(3.59) [John and Bill]$_i$ wrapped blankets [around each other$_i$].

Bresnan et al. (1985) analyze the English pronominal *him/her* as obeying the condition referred to above as the Coargument Disjointness Condition. *Him/her* may not be coreferent with a coargument:

(3.60) *John$_i$ saw him$_i$.

However, it may corefer with noncoarguments:

(3.61) [John and Bill]$_i$ told Mary to look at them$_i$.

(3.62) [John and Bill]$_i$ said that Mary saw them$_i$.

(3.63) [John and Bill]$_i$ wrapped blankets [around them$_i$].

In some respects, however, anaphoric binding in English is more complicated than the above examples indicate. For instance, Pollard and Sag (1989) provide evidence that antecedency of English reflexives and reciprocals is sometimes determined with respect not to syntactic constraints but to discourse constraints. We turn now to a discussion of their analysis.

3.2.1 Exemption from Binding Constraints: Pollard and Sag 1990

Pollard and Sag (1989) present an analysis of anaphoric binding in English in which, under syntactically determined conditions, an anaphor may be subject to discourse binding constraints. These conditions depend on the *obliqueness* of the anaphor.

Pollard and Sag's analysis is couched in the framework of Head-Driven Phrase Structure Grammar (Pollard and Sag 1987, Pollard and Sag 1990). According to HPSG, a lexical item that subcategorizes for complements, such as a verb, has a SUBCAT list consisting of the arguments that the item requires in order to become *saturated*; for an item to be saturated, it must have combined with all of its arguments.

The SUBCAT list is organized according to a hierarchy of obliqueness: the least oblique argument of a verb is its subject, the next is the object, and so on. In

treating subcategorization in this way, HPSG builds on the work of other researchers who have also noted a relation of obliqueness holding among arguments of a predicate. Ideas such as these are explicit in work by Keenan and Comrie (1977) as well as in work on categorial grammar (Dowty 1982); the ordering of final stratum terms in Relational Grammar (Perlmutter and Postal 1983, Perlmutter and Rosen 1984) also reflects such a hierarchy.

To take an example, the SUBCAT list of the verb *donate* would be:[10]

(3.64) SUBCAT = $\langle PP,\ NP_2,\ NP_1 \rangle$
more $\longrightarrow$ less oblique

Less oblique arguments appear farther to the right on the SUBCAT list. The PP which *donate* subcategorizes for is its most oblique argument. NP_2 is the object of *donate*. NP_1, the least oblique argument, is its subject.

The analysis of English anaphoric binding which Pollard and Sag (1989) present makes the following claims:

- A reflexive or reciprocal must be bound by a less oblique referential NP coargument, *if there is one.*
- A pronominal must be free from any less oblique referential NP coarguments.

Pollard and Sag's claim that it is only *less oblique* coarguments that are relevant for binding or disjointness constraints is an important one. We will return to a discussion of this claim in Chapter 5, Section 5.6.3; a distinction among arguments of a predicate along these lines is clearly necessary, and we will see how a similar result can be obtained within the LFG framework.

The other claim which Pollard and Sag make, and the one that is of immediate concern, is that reflexives and reciprocals are exempt from syntactic binding constraints when there is no less oblique referential NP coargument on the SUBCAT list. Their theory is similar in this respect to the one proposed by Manzini (1983) and Manzini and Wexler (1987).

[10]The members of the SUBCAT list, schematically represented here as the atoms PP, NP, and so on, are actually complex graphs containing phonological, syntactic, and semantic information.

Seen broadly, the similarity between Pollard and Sag's theory and the one presented by Manzini is the rejection of what Safir (1990) calls the Next Subject Up requirement. In most GB versions of the binding theory, as well as in the theory presented by Bresnan et al. (1985), the Next Subject Up requirement is relevant. The binding domain for an anaphor is the smallest domain satisfying some requirement: the smallest domain containing a subject, for example. If the minimal domain containing the anaphor does not satisfy the requirement, the next larger domain is checked; checking stops when the smallest domain fulfilling the requirement is found, and that domain becomes the relevant one for binding constraints.

In contrast, the theories of Manzini and of Pollard and Sag check only the smallest domain to see whether the requirement is met. If the requirement is not met in the smallest domain, larger domains are not checked; in that case, the anaphor is exempt from syntactic binding constraints.

The particulars of the theory of binding presented in Manzini (1983) are quite different from Pollard and Sag's, though the similarity between the two is plain. In Manzini's theory of binding, an anaphor is subject to two constraints: it must be bound in its governing category and in what Manzini calls its domain-governing category, if these categories exist.

Manzini's definition of 'governing category' is unlike Chomsky's: for Manzini, the governing category for a phrase α is the smallest category containing α, a governor for α, and a subject, provided that the subject of that smallest category is accessible. If the subject is not accessible, then α has no governing category. As for the domain-governing category, that may be roughly characterized as the governing category for the phrase properly containing the anaphor, with the identical notion of 'governing category' applying in both instances.

Manzini's intent in introducing the notion 'domain-governing category' seems to be to produce the Next Subject Up effect, but once only. That is, for Manzini, an anaphor is bound in the minimal category containing it and a subject, if that is a governing category; otherwise it is bound in the next larger domain containing a subject, if that is a governing category for the phrase containing it; if neither of those categories count as governing categories, the phrase is not subject to syntactic

constraints.

For Manzini, this sentence is grammatical:

(3.65) The boys$_i$ saw each other's$_i$ pictures.

The minimal domain containing *each other*, the governor of *each other*, and a subject is the NP *each other's pictures*. However, the subject of this NP is *each other*, and a subject is not accessible to itself. For this reason, *each other* has no governing category.

However, the phrase containing *each other*, *each other's pictures*, has a governing category, namely the entire S — and this counts as the domain-governing category for *each other*. *Each other* is bound in the S, and the sentence is grammatical.

This sentence is also grammatical, according to Manzini's theory:

(3.66) Each other's$_i$ pictures would please the boys$_i$.

The minimal domain containing *each other*, the governor of *each other*, and a subject is the NP *each other's pictures*. However, as in example (3.65), the subject of this NP is *each other*, and a subject is not accessible to itself. Again, then, *each other* has no governing category.

The category containing *each other* is *each other's pictures*. The minimal category containing *each other's pictures*, its governor, and a subject is the entire S. However, once again the subject of the S is not accessible to itself. Thus in (3.66) *each other* has no governing category or domain-governing category, and is thus not subject to syntactic constraints. It is, then, free to corefer with a non-c-commanding NP such as *the boys*.

On both Manzini's and Pollard and Sag's theories, then, there are some positions in which anaphors do not obey syntactic constraints on their distribution. Pollard and Sag's theory differs, though, in an important respect. On Manzini's theory, only anaphors appearing in subject position may be exempt from binding conditions. For Pollard and Sag, the positions exempt from the binding theory may have various syntactic realizations, though they share the characteristic of being the referential NP

that is least oblique on the SUBCAT list of the predicate. In particular, possessors of NP, subjects, and arguments of NP's that do not have possessors are all positions in which there is no less oblique referential NP on the SUBCAT list; anaphors appearing in any of these positions are exempt from syntactic binding constraints.

Additionally, Pollard and Sag's theory is more explicit than Manzini's in characterizing the nature of the constraints that hold of anaphors in syntactically 'exempt' positions: when anaphors appear in such positions, they must 'reflect the *point of view* that the sentence in question presents' (Pollard and Sag 1989, p. 19).

The categories which for Pollard and Sag are subsumed under the heading 'least oblique referential NP' can be divided into two subgroups: (1) possessors of NP and subjects (the *subjective* functions of LFG), and (2) arguments of NP's with no possessor, such as *himself* in the phrase *some pictures of himself*. In these categories, Pollard and Sag claim, only discourse factors are relevant in determining antecedency of reflexives. Anaphors that do not fall into one of these categories are those which are not the least oblique referential NP, and these are subject only to syntactic constraints.

In Chapter 5, Section 5.7, we will return to a discussion of the syntactic binding constraints that apply to anaphors in subject position. The position taken in this work is that in some languages anaphors in subject position do in fact obey syntactic binding constraints, not just discourse constraints.

Here we discuss Pollard and Sag's position that only anaphors that are least oblique can be subject to discourse binding in English — that syntactic considerations determine when discourse binding can occur. Their position seems to be too strong. Zribi-Herz (1989) provides a number of examples of anaphors which are not the least oblique referential NP on their SUBCAT list; nevertheless, discourse binding is possible.

Examples she cites, from naturally-occurring text, include:

(3.67) *Clara*$_i$ did not know whether to regret or to rejoice at their arrival; she$_i$ did not get on well with either of them (...) and yet, on the other hand their presence did not intensify the difficulty of an evening, but somehow dissipated and confused it, so that at least its burden did not rest upon *herself*$_i$ alone.

(from *Jerusalem the Golden*, Margaret Drabble.)

(3.68) But *Rupert*$_i$ was not unduly worried about Peter's$_j$ opinion of *himself*$_i$.

(from *A fairly honourable defeat*, Iris Murdoch.)

In these examples, the reflexive is not the least oblique NP on the SUBCAT list on which it appears. In example (3.67), *herself* is a syntactic argument of *rest*, and the NP *its burden* is a less oblique referential NP. In example (3.68), *himself* is an argument of *opinion*, and the possessive NP *Peter's* is a less oblique referential NP.

Zribi-Herz's position is that the possibility for discourse binding is always available, no matter what the syntactic realization of the anaphor. A reflexive may obey purely syntactic binding conditions, uninfluenced by discourse constraints; when it appears in violation of some syntactic binding constraints, its distribution is constrained by discourse considerations. In particular, all reflexives obey this condition:

(3.69) The English reflexive pronoun may not be separated from its antecedent by a domain-of-point-of-view boundary.

Within the domain of a single point of view, the reflexive is either bound to the 'subject of consciousness', perhaps in violation of syntactic constraints such as f-command or clause-boundedness, or to an antecedent which is eligible because it fulfills certain syntactic criteria. Zribi-Herz notes that the following example is ungrammatical if a reflexive is used (she provides additional examples as well):

(3.70) *He*$_i$ sat staring straight ahead of him with bright blue eyes that seemed a little screwed up, as if the glare of the East were still in them; and puckered at the corners as of the dust were still in them. Some thought had struck *him*$_i$/**himself*$_i$ that made what the others were saying of no interest to *him*$_i$/**himself*$_i$.

(from *The Years*, Virginia Woolfe)

According to Zribi-Herz, there is a domain-of-point-of-view boundary intervening between the 'subject of consciousness' *he* and the impermissible occurrences of reflexives; the narrator switches to an objective, reportative stance in telling the story. Thus

although *he* is a perfectly acceptable 'subject of consciousness' and should be an acceptable antecedent for a reflexive, it is prevented from anteceding the reflexives in this example because it is outside the domain-of-point-of-view.

On the other hand, it is not the case that the 'subject of consciousness' is the *only* possible antecedent for a reflexive; syntactic constraints also apply. For example, the following example provided by Zribi-Herz is ambiguous:

(3.71) But Rupert$_i$ was not unduly worried about Peter's$_j$ opinion of *himself*$_{i,j}$.

The reading available in the text in which the example occurred is the one where *himself* and *Rupert* are coindexed; the other reading is clearly also available, though.

This example makes the same point:

(3.72) Not till she had, with difficulty, succeeded in explaining to *him*$_i$ that she had done nothing to justify such results and that *his*$_i$ wife was equally incredulous of her innocence and suspected *himself*$_i$, the pastor, to be the cause of her distress, did *his*$_i$ face light up with understanding.

(from *Of mortal love*, William Gerhardie.)

Since the 'subject of consciousness' is *him*, it can antecede the reflexive *himself*. Note, though, that *herself* would also be grammatical in this example, with the meaning that the pastor's wife suspects herself to be the cause of distress and not the pastor. However, the pastor's wife is not the 'subject of consciousness' in this example; *herself* is possible because *his wife* is a syntactically suitable antecedent for the reflexive.

According to Zribi-Herz, then, anaphoric binding in English may be determined by purely syntactic factors; it may also be determined by discourse factors such as 'subject of consciousness'. Work on anaphors in other languages shows that anaphors can be simultaneously subject to both syntactic and discourse binding constraints. For example, Maling (1984) shows that the antecedent of the Icelandic anaphor *sig* is determined by both syntactic and discourse considerations: the antecedent of *sig* must be a subject as well as an acceptable logophoric antecedent.

Pollard and Sag's position, if weakened somewhat, may prove to be correct. When a reflexive or reciprocal appears in a position where there is no syntactically available

antecedent in the local syntactic nucleus in which it appears, discourse influences on binding may become stronger. It is not the case, though, that the possible influence of discourse factors is determined completely by syntactic considerations; in this, Pollard and Sag's position is too strong.

3.2.2 English Reflexives with Nonlocal Antecedents

Reflexives in constructions involving picture nouns are notorious in allowing non-local antecedents. Many of the examples upon which the reanalyses of the binding theory presented by Huang (1983) and Chomsky (1986) are based involve picture nouns; however, as noted above, it is often the case that picture noun reflexives are subject not to syntactic constraints but to discourse constraints. Revising the binding theory so that picture noun reflexives are syntactically permissible does not, therefore, seem to be the correct move. Some examples of nonlocal binding involving picture nouns are:[11]

(3.73) a. The picture of himself$_i$ in the museum bothered John$_i$.

b. John's$_i$ campaign requires that pictures of himself$_i$ be placed all over town.

c. John's$_i$ intentionally misleading testimony was sufficient to ensure that there would be pictures of himself$_i$ all over the morning papers.

It is also possible for the antecedent of the reflexive in such constructions to appear in a different sentence:

(3.74) a. John$_i$ was furious. The picture of himself$_i$ in the museum had been mutilated.

b. Mary$_i$ was extremely upset. That picture of herself$_i$ on the front page of the Times would circulate all over the world.

c. John$_i$ was going to get even with Mary. That picture of himself$_i$ in the paper would really annoy her, as would the other stunts he had planned.

[11]Some of these examples are taken from Pollard and Sag (1989).

As Pollard and Sag point out, these examples involve coreference between the reflexive and the individual whose 'point-of-view' is expressed in the sentence. For this reason, the following sentences are not felicitous:

(3.75) a. *The picture of himself in the museum bothered John's$_i$ father.

b. Without J's knowledge$_i$, his campaign managers made a crucial decision. ??Pictures of himself would be placed all over town.

However, the following sentence is infelicitous because it is not possible for a single sentence to express more than one 'point-of-view', as would be required by the use of two reflexives referring to different individuals:

(3.76) *John told Mary that the photo of himself with her in Rome proved that the photo of herself with him in Naples was a fake.

Examples provided by Pollard and Sag, as well as those provided by Zribi-Herz, amply show that there are cases of long-distance binding of reflexives in English in which the antecedent of the reflexive is determined by discourse factors. However, as we have seen, purely syntactic considerations do not determine when these discourse relations become relevant.

For cases where the antecedent of the reflexive is *not* the 'point-of-view' — that is, cases in which the relevant discourse configuration does not obtain — syntactic constraints determine possible antecedents for the English reflexive; in those cases, as the theories of Chomsky and of Bresnan et al. (1985) suggest, the reflexive must be bound in the minimal domain containing a subject. For cases such as those described by Pollard and Sag and by Zribi-Herz, some combination of syntactic constraints and discourse factors can determine reflexive antecedency; in these cases, the reflexive may find an antecedent in a wider domain.

3.3 Disjointness Constraints: Reinhart 1986

Since the work of Lasnik (1976), it is widely accepted that both coreference and disjointness conditions play a role in anaphoric binding; this is the position adopted in

this work. In Chapter 2, anaphoric elements were shown to obey two sorts of binding conditions: *positive* conditions on the elements with which an anaphoric element may corefer, and *negative* conditions on elements with which an anaphor may not corefer.

In some work, though, negative constraints are taken to be pragmatic in nature, not syntactic. For example, Dowty (1980) and Reinhart (1986) propose that negative constraints are a consequence of a pragmatic strategy of cooperativeness: a bound form must be used whenever possible. Reinhart gives the following as the speaker's and hearer's strategy:

(3.77) Reinhart (1986, p. 143):

> *Speaker's strategy*: When a syntactic structure you are using allows bound-anaphora interpretation, then use it if you intend your expressions to corefer, unless you have some reason to avoid bound-anaphora.
>
> *Hearer's strategy*: If the speaker avoids the bound anaphora options provided by the structure he is using, then, unless he has reasons to avoid bound-anaphora, he didn't intend his expressions to corefer.

On Reinhart's analysis, example (3.78)a is unacceptable on pragmatic grounds because example (3.78)b is allowed by the grammar:

(3.78) a. *He$_i$ saw him$_i$ in the mirror.

b. He$_i$ saw himself$_i$ in the mirror.

However, data from Norwegian provide evidence that this position cannot be correct — that negative as well as positive binding constraints exist.

Recall that the anaphor *ham selv* obeys the Subject Disjointness Condition; it cannot corefer with a subject. Example (3.79) is ungrammatical for this reason:

(3.79) *Jon fortalte seg selv om ham selv
Jon told self about self
'Jon$_i$ told self$_i$ about self$_i$.'

The ungrammaticality of this example is not due to *seg selv*, since binding conditions for *seg selv* are satisfied: it must be bound to a subject coargument, and *Jon* is such an element. Binding conditions for *ham selv* are violated, though. *Ham selv* must find an antecedent in the minimal complete nucleus in which it is contained; both *Jon* and *seg selv* satisfy this requirement. However, coreference obtains between *ham selv* and *Jon*, a subject, and this is disallowed by the Subject Disjointness Condition.

On an account such as Reinhart's, the ungrammaticality of example (3.79) is not explainable. This is because binding to the object position is possible for *ham selv*, as illustrated by example (3.80):

(3.80) vi fortalte Jon om ham selv
we told Jon about self
'We told Jon$_i$ about self$_i$.'

For Reinhart, binding to the object of *fortalte*, *seg selv*, ought to be a possibility in example (3.79); on a theory in which the only requirements for *ham selv* are positive, the ungrammaticality of example (3.79) remains mysterious. The position that *ham selv* is subject to negative constraints (it may not corefer with a subject) as well as positive constraints (it must corefer with an element in the minimal complete nucleus) provides an explanation for these facts.

It is possible to express the meaning that is intended in example (3.79), by use of the reflexive whose antecedent must be a subject in both instances, as in the following:

(3.81) Jon fortalte seg selv om seg selv
Jon told self about self
'Jon$_i$ told self$_i$ about self$_i$.'

The conclusion is, then, that negative as well as positive constraints play a role in the grammar of anaphora.

Lasnik (1989) also provides arguments for the necessity of disjointness conditions on anaphoric binding. He notes that a grammar without negative as well as positive conditions has no way of ruling out examples such as these:

(3.82) a. *John$_i$ and Mary like him$_i$.

b. *John told them that Mary should leave. [where John is in the group referred to by 'them']

Note that the reflexive cannot be used in these situations either; for the reflexive to occur felicitously, identity between the reflexive and its antecedent must obtain, and here there is only referential overlap:

(3.83) a. *John$_i$ and Mary like himself$_i$.

b. *John told themselves that Mary should leave.

Given a theory with no negative constraints on anaphoric binding, there is nothing to forbid the use of a pronoun in cases such as (3.82), where the reflexive cannot appear. Such a theory would incorrectly predict that sentences such as those in (3.82) are grammatical.

3.4 Lexical-Functional Grammar: Bresnan et al. 1985

In this section and in the remainder of this work, I assume familiarity with the basic concepts of the theory of LFG and also with the notation commonly used in expressing these concepts. See Bresnan (1982b) (especially Kaplan and Bresnan (1982)) and Levin et al. (1983) for an explanation of these concepts.

Bresnan et al. (1985) describe an approach to constraints on anaphoric elements in which binding requirements are determined by a small set of properties stated in terms of features; under their approach, both domain and antecedent constraints are accounted for. The analysis they present is lexical: each anaphoric element is lexically associated with some set of these features. Thus, for example, the fact that an anaphor must have a subject as its antecedent is stated with the feature +SUBJECT. If an anaphor may not have a subject as its antecedent, it is marked −SUBJECT. If an anaphor can have either a subject or a non-subject as its antecedent, it is unmarked with respect to the SUBJECT feature. Bresnan et al. (1985) provide an analysis of anaphoric relations in Norwegian in which nearly all the possible combinations of the

features SUBJECT and NUCLEAR are put to use. Here, I discuss the features SUBJECT and NUCLEAR as they appear in Bresnan et al. (1985).[12]

3.4.1 SUBJECT and NUCLEAR

The SUBJECT feature indicates whether a lexical item has a subject as its antecedent. Items that are +SUBJECT must be anteceded by a subject; those that are −SUBJECT may not be anteceded by a subject; and those that are unmarked for the feature have no subjecthood requirement on their antecedent. In Chapter 2, these conditions were referred to as the Subject Binding Condition and the Subject Disjointness Condition.

According to the analysis of Bresnan et al. (1985), cited in Sells (1985), the Norwegian reflexive *seg selv* is +SUBJECT:

(3.84) Jon fortalte meg om seg selv
Jon told me about self
'Jon$_i$ told me about self$_i$.'

(3.85) *vi fortalte Jon om seg selv
we told Jon about self
'We told Jon$_i$ about self$_i$.'

As we have noted, the antecedent of *seg selv* is the subject of the sentence; nonsubjects are not eligible antecedents for *seg selv.*

The anaphor *ham selv* is −SUBJECT; *ham selv* must be bound to a nonsubject:

(3.86) *Ola snakket om ham selv
Ola talked about himself
Ola$_i$ talked about himself$_i$.'

(3.87) vi fortalte Ola om ham selv
we told Ola about himself
'We told Ola$_i$ about himself$_i$.'

The characterization of the feature NUCLEAR given in Bresnan et al. (1985) and Bresnan (1987) is slightly more complicated. The −NUCLEAR requirement is stated

[12]For a more detailed exposition of the anaphoric feature system, see Sells (1985).

in terms of the notion of *nucleus*, a syntactic predicate and its arguments. A pronoun that is −NUCLEAR may not appear in the same syntactic nucleus as its antecedent; in other words, a −NUCLEAR pronoun obeys the Coargument Disjointness Condition.

An example of a −NUCLEAR pronoun is the English pronominal, which may not appear in the same nucleus with its antecedent:

(3.88) *Jane$_i$ defended her$_i$.

(3.89) Jane$_i$ hopes that Max will hire her$_i$.

(3.90) Jane$_i$ liked the story about her$_i$.

(3.91) Jane$_i$ wrapped the blanket around her$_i$.

In example (3.88), the pronominal and *Jane* are coarguments of 'defend', and *Jane* cannot antecede *her*. On the other hand, in examples (3.89)-(3.91), the pronominal and *Jane* are arguments of different predicates, and they may corefer.

The definition of +NUCLEAR is as follows (Sells 1985, p. 174):

> Pronominals that are +NUCLEAR must find an antecedent within the minimal nucleus containing the pronominal and a subjective function.

+NUCLEAR anaphors must find their antecedent inside the minimal complete nucleus — they obey the Minimal Complete Nucleus Binding Condition. Note, then, that there is an asymmetry with regard to the domain of +NUCLEAR and −NUCLEAR: the definition of +NUCLEAR refers to the minimal *complete* nucleus, while the definition of −NUCLEAR refers simply to the nucleus.

On the analysis of Bresnan et al. (1985), the English reflexive pronominal is an example of a +NUCLEAR anaphor. Consider the following sentence:

(3.92) John hit himself.

A simplified f-structure representation for this sentence is:

(3.93)
$$\left[\begin{array}{l}\text{PRED 'hit}\langle(\uparrow \text{SUBJ}), (\uparrow \text{OBJ})\rangle\text{'} \\ \text{SUBJ}\ \left[\text{PRED 'John'}\right] \\ \text{OBJ}\ \left[\text{PRED 'self'}\right]\end{array}\right] \longleftarrow (1)$$

The minimal complete nucleus containing the reflexive and a subjective function (in this case SUBJ) is the entire f-structure, labeled (1). *Himself* must find an antecedent in this domain; the only acceptable indexing for the sentence is where *John* and *himself* are coindexed.

In the following sentence, too, the antecedent of the reflexive is *John*:

(3.94) John wrapped the blanket around himself.

An f-structure representation of this sentence is:

(3.95)
$$\left[\begin{array}{l}\text{PRED 'wrap}\langle(\uparrow \text{SUBJ}), (\uparrow \text{OBJ})\rangle\text{'} \\ \text{SUBJ}\ \left[\text{PRED 'John'}\right] \\ \text{OBJ}\ \left[\text{PRED 'blanket'}\right] \\ \text{ADJ}\ \left[\begin{array}{l}\text{PRED 'around}\langle(\uparrow \text{OBJ})\rangle\text{'} \\ \text{OBJ 'self'}\end{array}\right]\end{array}\right] \begin{array}{l}\longleftarrow (1) \\ \longleftarrow (2)\end{array}$$

The minimal nucleus which contains the reflexive, labeled (2), is the one whose PRED is 'around $\langle(\uparrow$ OBJ$)\rangle$'; however, this nucleus is not a complete nucleus, as it does not contain a subjective function. The minimal nucleus containing the reflexive and a subjective function is the entire f-structure, labeled (1). Thus *John* is an acceptable antecedent for the reflexive.

Note that according to these definitions −NUCLEAR pronouns are not required to be in complementary distribution with +NUCLEAR anaphors. In particular, the following sentences are both acceptable:

(3.96) John$_i$ wrapped the blanket around him$_i$.

(3.97) John$_i$ wrapped the blanket around himself$_i$.

Examples such as these are problematic for theories that posit complete complementary distribution between the reflexive and the pronominal, such as the theory presented in Chomsky (1981). According to the above definitions of +NUCLEAR and −NUCLEAR, though, it is predicted that either the pronominal or the reflexive can appear as the object of *around* and can corefer with the matrix subject. As noted above, the +NUCLEAR reflexive finds its antecedent in the minimal *complete* nucleus; *John* appears in the minimal complete nucleus containing the reflexive and so can antecede it. The −NUCLEAR pronominal must be disjoint from its coarguments; *John* is not an argument of the same PRED as the pronominal and so can antecede it.

An approach of this kind is more successful than many of the GB analyses discussed above at specifying binding constraints in languages with more than one anaphoric element. The analysis to be presented in this work is similar to the analysis of Bresnan et al., in that binding constraints are lexically associated with each anaphoric element. Further, some of the constraints presented in this work (for example, those dealing with subjecthood requirements on the antecedent of the anaphor) are exact analogues of these constraints.

However, a feature approach also has drawbacks, ones which the present analysis does not share. Foremost is that given a feature approach to binding, it is not clear how the features should interact with each other, or even that they should interact at all. Each feature is independent of the others; any interaction between them would have to be stipulated, since no interaction is predicted.

In fact, binding requirements do interact, as we will see in the next section. We will also see that it is possible to model these interactions easily in the system to be presented in this work.

3.4.2 Feature Interactions: Two Positive Features

Anaphors are sometimes associated with multiple positive binding features. In this case, how do the features interact? Can the features be satisfied independently, by

two different antecedents? The answer is no; the features must be satisfied by the same element.

Consider the following ungrammatical sentence from Norwegian:

(3.98) *Martin ba oss snakke til ham om seg selv
Martin asked us to talk to him about himself
(+SUBJECT, +NUCLEAR)
'Martin$_i$ asked us to talk to him$_i$ about himself$_i$.'

Here is a skeletal f-structure for this sentence:

(3.99)
$$\begin{bmatrix} \text{PRED 'ask}\langle(\uparrow \text{SUBJ}), (\uparrow \text{OBJ}), (\uparrow \text{XCOMP})\rangle\text{'} \\ \text{SUBJ}\ \begin{bmatrix}\text{PRED 'Martin'}\end{bmatrix} \\ \text{OBJ}\ \begin{bmatrix}\text{PRED 'us'}\end{bmatrix} \\ \text{XCOMP}\ \begin{bmatrix} \text{PRED 'talk}\langle(\uparrow \text{SUBJ}), (\uparrow \text{OBL}_{to}), (\uparrow \text{OBL}_{about})\rangle \\ \text{SUBJ} \\ \text{OBL}_{to}\ \begin{bmatrix}\text{PRED 'him'}\end{bmatrix} \\ \text{OBL}_{about}\ \begin{bmatrix}\text{PRED 'himself'}\end{bmatrix} \end{bmatrix} \end{bmatrix}$$

As indicated in (3.98), the anaphor *seg selv* is [+SUBJECT,+NUCLEAR]. In this example, the +SUBJECT feature is satisfied, since the anaphor is coreferent with the matrix subject *Martin*. The +NUCLEAR feature is also satisfied, since the anaphor is coreferent with an argument within its minimal complete nucleus, *ham*. Nevertheless, the sentence is ungrammatical. The same element must simultaneously satisfy both the +SUBJECT and the +NUCLEAR requirement; the two features may not be satisfied independently. This, however, is not predicted by a feature account.

The ungrammaticality of the above example contrasts with the grammaticality of this one:

(3.100) Martin ba oss snakke til ham om ham selv
Martin asked us to talk to him about himself
(−SUBJECT, +NUCLEAR)
'Martin$_i$ asked us to talk to him$_i$ about himself$_i$.'

When the reflexive used is one that requires coreference with a nonsubject argument, the sentence is grammatical.

3.4.3 Feature Interactions: Dilemmas

Hellan (1988) shows that the Norwegian possessive pronominal *hans* is required to be disjoint from the immediately higher subject:

(3.101) (Hellan 1988, p. 61):

*Jon traff hans venner
Jon met his friends
'Jon$_i$ met his$_i$ friends.'

It may, though, be coreferent with a nonsubject:

(3.102) (Hellan 1988, p. 134):

vi fant Jon under sengen hans
we found Jon under bed his
'We found Jon$_i$ under his$_i$ bed.'

Or a subject that is not immediately higher:

(3.103) (Hellan 1988, p. 133):

Jon ba meg hjelpe seg og moren hans
Jon asked me to help him and mother his
'Jon$_i$ asked me to help him$_i$ and his$_i$ mother.'

(3.104) Jon gjorde oss glad i huset sitt
Jon made us fond of house self's
?hans
his
'Jon$_i$ made us fond of his$_i$/self's$_i$ house.'

In example (3.104), *sin* is 'slightly preferred' (Hellan 1988, p. 132); *hans* is not ungrammatical, though, as it is when coreference with the immediately higher subject is intended.

It is not clear how to represent a requirement like 'disjoint from immediately higher SUBJ' with a combination of the features SUBJECT and NUCLEAR. A −SUBJECT pronoun is required to be disjoint from *every* higher subject, while a −NUCLEAR pronoun is required to be disjoint from *every* coargument; both of these features are too strong to characterize the constraints that *hans* obeys. It is possible that some interaction could be defined between the features [−SUBJECT] and [+NUCLEAR] (or [−NUCLEAR]) to represent these facts; however, it is not straightforward to see how this can be done, or how it would interact with the interpretations of the features that are otherwise required.

In sum, the strength of the approach outlined by Bresnan et al. (1985) is that binding conditions are taken to be lexically specified for each anaphoric element. It is only by associating binding conditions with individual elements that facts concerning languages with multiple anaphoric elements, each with its own set of binding conditions, can be adequately described. The weakness of their approach is that an analysis of the interaction among the binding requirements is not provided; given a feature approach, no interactions among the requirements are predicted.

3.5 Conclusion

We have seen that binding requirements must be lexically associated with anaphoric elements; many proposals made within Government-Binding theory are unsuccessful because binding requirements are taken to apply universally or to be parametrized by languages as a whole. In fact, some of the GB proposals we examined owed their success to the lexical statement of binding constraints.

In Chapter 4, I will outline a theory of anaphoric binding which makes use of the LFG technique of *functional uncertainty* to relate anaphors and their antecedents; the same technique will be used to specify disjointness requirements for pronominals. In Chapter 5, I will explore interactions between these constraints.

Chapter 4

Anaphoric Binding, Projections, and Functional Uncertainty

The inventory of binding and disjointness conditions illustrated in the previous chapters demarcate certain permissible, structurally definable relations between an anaphor and its possible antecedents. This chapter provides a framework within which these constraints can be precisely formulated. A formalization of the binding requirements is useful in providing an explicit, testable statement of their properties and in predicting precisely how the requirements interact.

The f-structure representation of surface syntactic form is related to a semantic level of representation, called the semantic *projection*. The anaphor-antecedent relation represented thus far by coindexing will be given an interpretation in terms of a relation involving the f-structure representations of an anaphoric element and its antecedent and their corresponding semantic projections.

Syntactic binding constraints are statable by the use of *functional uncertainty*. Functional uncertainty provides a formal vocabulary for stating relations between f-structures; in the case at hand, the f-structures to be related are those of an anaphor and its permissible and impermissible antecedents. These relations can be stated in terms of *binding equations* involving functional uncertainty.

This chapter is devoted to a demonstration of how the binding constraints discussed in Chapter 2 can be modeled using these tools. In the following, I discuss only anaphoric elements whose antecedents are not quantified noun phrases; most of

the examples provided involve names. The same syntactic constraints that have been outlined above are at work when the antecedent is a quantified noun phrase; however, a treatment of quantification adds a great deal of complexity to the semantic representation, and I will not attempt to provide a theory of quantification here.

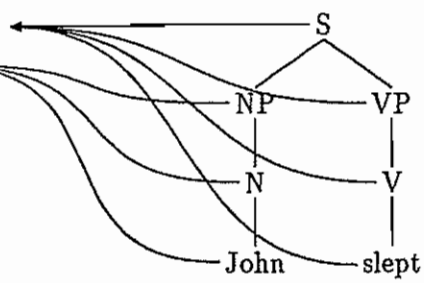
Discussion is also confined to the relation between a reflexive or pronominal and its antecedent; these cases are simplest, in that the relevant relation between the anaphor and its antecedent is identity at the semantic level. Reciprocals obey the same sorts of binding conditions as reflexives and pronominals, but their semantics is again much more complicated. In particular, the semantic relation between a reciprocal and its antecedent is not one of identity; rather, the reciprocal is a scope-bearing element, inducing a relation between a plural antecedent and a two-place relation. In the following, I treat only the simpler cases of identity of semantic relation.

4.1 Projections and Semantic Representation

4.1.1 Projections

In LFG, surface syntactic relations are represented in two ways: by a *c-structure* tree and by a directed graph, the *f-structure*. Each of these representations carries a different kind of information: the c-structure contains information about dominance, precedence and constituent relations, and the f-structure contains functional information. Kaplan and Bresnan (1982) note a correspondence between f-structures and c-structures of the following type:

(4.1)
$$\begin{bmatrix} \text{PRED 'sleep}\langle(\uparrow \text{ SUBJ})\rangle\text{'} \\ \text{TENSE PAST} \\ \text{SUBJ } \begin{bmatrix}\text{PRED 'John'}\end{bmatrix} \end{bmatrix}$$



The nodes in the tree labeled 'S', 'VP', and 'V' and the leaf node 'slept' are associated with the outermost f-structure; the nodes labeled 'NP' and 'N' and the leaf node 'John' are associated with the SUBJ f-structure.

In recent work, Kaplan (1987) and Halvorsen and Kaplan (1988) have introduced a new method for organizing levels of representation; these levels are called *projections*. Each projection represents some level of linguistic representation. For example, a semantic projection might encode semantic relationships, a discourse projection might encode discourse information, a phonological projection might encode phonological information, and so on. Projections are related to one another in a manner similar to the way the c-structure and the f-structure are related.

In other work within the LFG framework (Fenstad et al. 1987, Halvorsen 1988) it has been assumed that the semantic representation is represented as a directed graph, just as the f-structure is.[1] Here the simplifying assumption is made that the semantic representation consists of a formula of first-order logic. Since most of the examples we will be dealing with involve coreference or disjoint reference between an anaphor and a name, a representation like this one will suffice for present purposes. The representation would need to be more complex if examples involving quantification were also considered, however.

Here is a sample lexical entry for the verb *slept*:

(4.2) slept:

$$(\uparrow \text{ PRED}) = \text{'sleep}\langle(\uparrow \text{ SUBJ})\rangle\text{'}$$
$$(\uparrow \text{ TENSE}) = \text{PAST}$$
$$\uparrow_\sigma = \mathit{sleep}((\uparrow \text{SUBJ})_\sigma)$$

As in previous LFG work, the mapping from c-structure nodes to f-structures is represented by the up arrow $\uparrow$ and the down arrow $\downarrow$. According to these equations, the c-structure node that is the mother node of *slept* in the c-structure tree corresponds to an f-structure whose PRED is 'sleep $\langle(\uparrow$ SUBJ$)\rangle$' and whose value for TENSE is PAST.

[1] Pollard and Sag (1987) also represent semantic structure as a directed graph.

The new piece of notation introduced here is $\uparrow_\sigma$, which specifies a relation between projections.

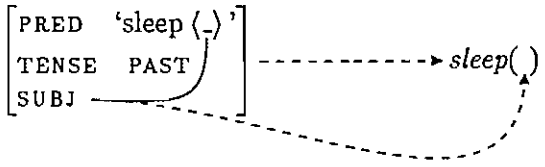
The subscript symbol σ is called a *projector*. Projectors map from pieces of one projection to pieces of another. In this case, the projector σ maps from f-structures to the semantic projection. For any f-structure f, f_σ is the corresponding semantic structure, reachable by the projector σ:

(4.3) $[\ \ldots\ ]_f \text{- - - - - - - - - - -} \blacktriangleright f_\sigma$

Here the projector σ is represented by a dotted line. Often, as in this example, the piece of structure which the projector maps to is associated with the same subscript label as the projector name. For example, the semantic structure reachable by the σ projector is referred to as the 'σ projection'.

The outermost f-structure, labeled (1), corresponds to $\uparrow$ in the above lexical entry; the following correspondences are induced:

(4.4) (1)



The equation ($\uparrow$ PRED) = 'sleep $\langle(\uparrow$ SUBJ$)\rangle$' induces identity between the first argument position of the PRED 'sleep' and the SUBJ of f-structure (1). The equation ($\uparrow$ TENSE) = PAST supplies the attribute TENSE with value PAST. The equation $\uparrow_\sigma = sleep((\uparrow \text{ SUBJ})_\sigma)$ induces a correspondence between the f-structure and the semantic form.

As above, the semantic projector σ is represented by a dotted line. $\uparrow_\sigma$ represents the semantic structure that corresponds to the f-structure $\uparrow$, and $(\uparrow \text{ SUBJ})_\sigma$ represents the semantic structure that corresponds to ($\uparrow$ SUBJ), the subject of *slept*. The

equation $\uparrow_\sigma = sleep((\uparrow \text{SUBJ})_\sigma)$ indicates that the semantic projection corresponding to $\uparrow$ (labeled (1) in example (4.4)) is the one-argument predicate *sleep*, where the argument of *sleep* is filled by the semantics of the subject of *sleep*. In general, the projector σ maps any f-structure (representable as $\uparrow$, $(\uparrow \text{SUBJ})$, etc.) to the corresponding piece of semantic structure.

4.1.2 Anaphoric Linkages

My intention in this chapter is to provide a means of modeling surface syntactic constraints on anaphoric binding — constraints that can be defined in terms of the f-structure and c-structure. These constraints involve the relation that has heretofore been represented by coindexing between two noun phrases.

The relation between noun phrases that is generally represented by coindexing is a very complex one; in fact, coindexing represents an array of possible semantic relations, depending both on the nature of the antecedent of the anaphor and on the type of the anaphor itself.

The cases of coreference/coindexing that are semantically the simplest are found where the antecedent of a pronominal or reflexive is a name or definite noun phrase. In these cases, coreference may be represented as identity of representation at the semantic level. That is, if a pronominal or reflexive and a name are coindexed, they have the same semantic representation.

However, where the antecedent is a quantified noun phrase, the situation is more complicated. In the case of quantified noun phrases, the anaphor and the antecedent should both be variables bound by the same quantifier. The relation between the anaphor and its antecedent is similar to the simpler case, where the antecedent of the anaphor is a name, in that identity between the two variables is what is wanted; however, a theory of quantifier scoping is also needed. Essentially, semantic identity between the anaphor and its antecedent should obtain at a level of representation at which quantifiers have not yet been scoped, one at which the quantified NP is represented simply as a variable.

A theory of quantifier scoping such as the one presented by Pereira (1989) would

allow for such an approach. On Pereira's theory, a quantified NP introduces a variable into the semantic form of a sentence, as well as a quantifier *assumption*; the relation between an anaphor and a quantified antecedent is that of identity between the semantic representation of the anaphor and the variable introduced by the quantifier. Scoping the quantifier involves *discharging* the quantifier assumption and binding the quantifier variable, including any occurrences of the variable corresponding to pronouns bound by the quantifier.

In the case of reciprocals, too, identity between the reciprocal and its antecedent is not what is wanted. Semantically, a reciprocal encodes a relation between a group and a two-place relation. For example, the semantics of a sentence like *John and Mary like each other* is a reciprocal relation involving the two-place relation

$$\lambda x \cdot \lambda y \cdot like(x, y)$$

and the group

$$\{john,\ mary\}$$

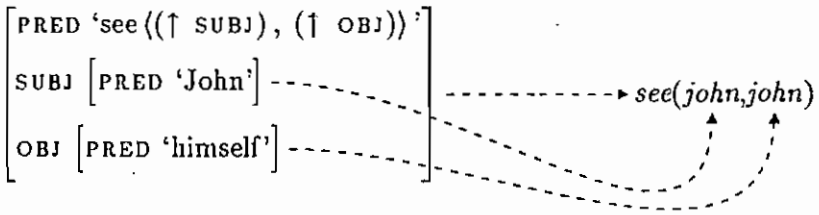
The reciprocal and its antecedent mark the abstracted argument positions of this two-place relation; the semantics of the reciprocal and its antecedent are not identical.

For ease of exposition, I will not discuss any but the simplest cases of the relation of coreference or coindexing in the following. The semantic relation between an anaphor and its antecedent which I will model is that of identity: the relation that holds between nonreciprocal anaphors and definite noun phrase antecedents. Extending the treatment of coindexing to include the more complex cases would pose no difficulty, however.

For a sentence such as $John_i$ *saw* $himself_i$, the semantic representation is:

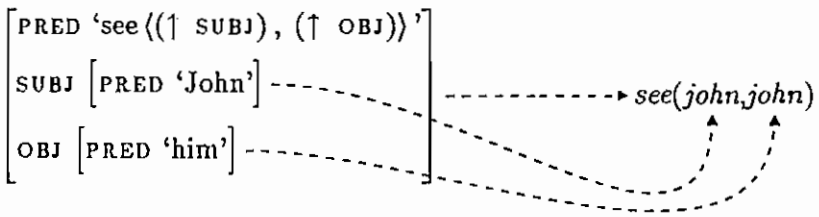
(4.5) *see(john, john)*

The representation of the reflexive *himself* is identical to that of its antecedent *John*. An f-structure for this sentence, together with the σ projector relations between the f-structure and the semantic form (again represented as dotted lines), is:

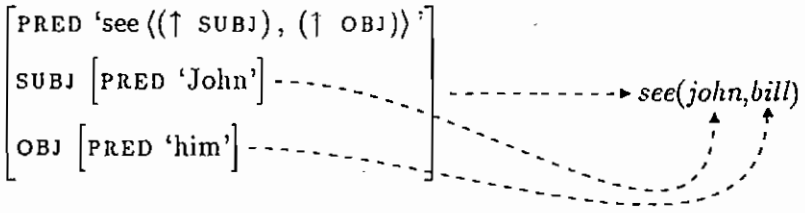
(4.6)


In the f-structure, the representations for the reflexive *himself* and its antecedent *John* are *not* identified; syntactically, they are autonomous and play different roles. Semantically, though, the anaphor has no intrinsic referent.

My assumption is that syntactic constraints on anaphoric binding perform the function of *checking* that coreferential relations are properly represented in the semantic structure. That is, I assume that referents for anaphoric elements are freely supplied. These referents are acceptable as long as they do not violate any syntactic constraints. In the example *John saw himself*, syntactic constraints on the relation between the anaphoric element *himself* and its antecedent require coreference between *John* and *himself*, since *John* is the only acceptable antecedent for the reflexive in this case. However, this is not the case with pronominals. In the sentence *John saw him*, *him* must be disjoint in reference with *John*; other than that, though, there are no syntactic constraints on the referent of *him*. Any suitable discourse entity will do as an antecedent for *him*. For *John saw him*, the following representation is disallowed by the Coargument Disjointness Condition:

(4.7)


This is not a permissible semantic representation for this sentence because the semantic representation for the pronominal *him* is identical with that of a coargument, and this is what the Coargument Disjointness Condition disallows. This representation is permitted, however:

(4.8)


Let us assume that *bill* is the semantic representation of some available discourse referent (or, alternatively, a parameter anchored to the individual Bill, in the sense of Gawron and Peters (1988)); I assume that discourse referents are freely available as interpretations of anaphors.

In this, I follow to some extent the work of Gawron and Peters (1988). In their theory, anaphors function as parameters that are anchored by a component they call the Circumstances to entities in the world. An approach of that nature would work well here. Syntactic constraints would require that when a binding relation obtains, a pronoun and its antecedent use the same parameter, supplied by the Circumstances; the parameter is anchored to some individual, the referent of both the antecedent and the anaphor. Disjointness constraints would require the use of distinct parameters by an anaphor and the set of elements with which it may not corefer.

There is a distinction between pronouns that can only be used anaphorically and those that may also be used deictically. For example, the English pronominal can be used deictically, to introduce a discourse referent (in the sense of Karttunen (1976)); it is possible to open a discourse by pointing to someone and saying:

(4.9) Who's he?

In some languages, pronouns cannot perform this function.[2] This does not seem to be a matter of the kinds of syntactic constraints a pronoun might be subject to; a pronoun may have no syntactic constraints whatever on the domain in which its antecedent must be found or the syntactic role its antecedent must bear, but yet be unable to perform the function of introducing a discourse entity. Rather, the

[2]This is true of the German pronouns *er*, *sie*, and *es*, which cannot be used deictically or to head restrictive relative clauses (Paul Kiparsky, p.c.).

distinction between purely anaphoric pronouns and those that may be used deictically is found in the relation between the semantics of the pronoun and the component of the grammar (here referred to as the Circumstances) which supplies the referent for the pronoun.

In sum, the coreference relations that have thus far been represented by coindexing are represented in the system to be developed here as identity at the semantic level. Anaphoric elements such as *himself* require that there be an element that is coreferent in this sense within a particular syntactic domain, standing in a particular structural relation within the f-structure to *himself*. Similarly, elements such as *him* require that there be no such element within a particular syntactic domain, standing in a particular structural relation.

The task now is to specify the set of elements which is relevant for binding each anaphoric element: the set of elements one of whose semantics might be identical with the semantics of the anaphor, or each of whose semantics must be different from that of the anaphor. We turn now to a specification of that set of elements.

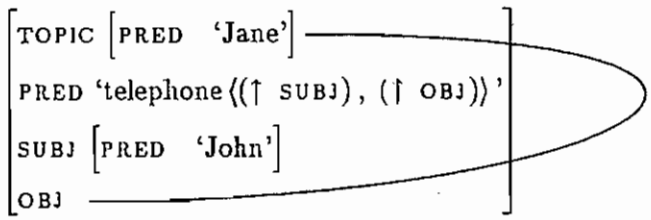
4.2 Functional Uncertainty: Modeling Binding Constraints

The technique of *functional uncertainty* can be used to model the conditions that have so far been proposed to describe possible binding relations between anaphoric elements.

Functional uncertainty involves the use of regular expressions as specifications of paths within feature structures, as illustrated by the LFG analysis of topicalization (Kaplan et al. 1987, Kaplan and Zaenen 1989).

Consider the topicalized sentence *Jane, John telephoned yesterday.* The TOPIC of the sentence is *Jane*; this NP also fills the role of the OBJ of *telephoned.* A skeletal f-structure for this sentence (ignoring the adverb *yesterday*) is:

(4.10)



Identification of the TOPIC and the OBJ might be expressed with the following equation at the top-level S:[3]

(4.11) (↑ TOPIC) = (↑ OBJ)

The following sentences are also grammatical:

(4.12) Jane, Bill claimed that John telephoned yesterday.

Jane, Bill claimed that Sue said that John telephoned yesterday.

Suppose we want to state a rule to the effect that the topic of the sentence should be identified with the OBJ, or with the COMP OBJ, or with the COMP COMP OBJ, and so on. In other words, the TOPIC is identified with the OBJ embedded inside any number (including zero) of COMPs. This can be stated with the following equation, associated with the rule expanding topicalized S:

(4.13) (↑ TOPIC) = (↑ COMP* OBJ)

The regular expression COMP* OBJ, involving the Kleene closure operator '*', stands for an infinite disjunction of paths within f-structures: paths involving zero or more COMPs followed by the attribute OBJ. This equation, then, identifies the TOPIC of the sentence with the value of any one of these paths.

An expression involving functional uncertainty is interpreted existentially, not universally. The equation is satisfied if there is *some* path picked out by the regular expression that makes the equation true.

[3]For the sake of clarity, the analysis of topicalization presented here is a simplified version of the one found in Kaplan et al. (1987). Constituents other than objects can, of course, be topicalized.

Formally, the interpretation of an expression not involving functional uncertainty is the following:

(4.14) $(f\ s) = v$ holds iff f is an f-structure, s is a symbol, and the pair $\langle s, v\rangle \in f$.

Extending this notation to equations in which the position occupied by s in the above example is occupied by a string of symbols rather than a single symbol, the following equivalences hold (Kaplan and Zaenen 1989, p. 6):

(4.15) $(f\ sx) \equiv ((f\ s)\ x)$, for a symbol s and a (possibly empty) string of symbols x.

$(f\epsilon) \equiv f$, where ϵ is the empty string.

These equivalences provide an interpretation for equations containing strings of symbols in the position following the f-structure specification f. The intent is the following. We are given an expression of the form $(f\ sx)$, where sx is a string whose first symbol is s and whose remainder is x. The f-structure reachable by the path sx through f is the same as the f-structure that is reachable from the f-structure fs by following the path x. The second clause of the definition states that by following the empty path through f, we reach f.

Now, the following holds of an expression involving functional uncertainty:

(4.16) $(f\ \alpha) = v$ holds iff for some x in the set of strings α, $(fx) = v$.

This definition gives an existential interpretation to an expression involving functional uncertainty: some solution must be found involving a string from the language described by the regular expression.

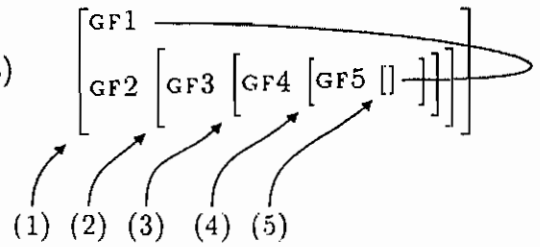
The type of functional uncertainty described above is often referred to as 'outside-in' functional uncertainty. The type of functional uncertainty which is of use in describing anaphoric binding constraints is instead 'inside-out' functional uncertainty. In general, outside-in functional uncertainty involves identification of a *less* deeply embedded f-structure with a *more* deeply embedded one. For inside-out functional uncertainty, the opposite is true.

Outside-in functional uncertainty (the type illustrated above, in the topicalization example) can be schematically represented in the following way:

(4.17) $(\uparrow \text{GF1}) = (\uparrow \text{GF}^* \text{GF5})$

where GF denotes the set of primitive grammatical function labels. Assume that the f-structure associated with $\uparrow$ is the outermost one, labeled (1):

(4.18)



The path picked out by the regular expression GF^* in this case is

GF2 GF3 GF4

and the result is that the less embedded GF1 is functionally identified with the more embedded GF5.

An alternative way of achieving the same result involves inside-out functional uncertainty. We might associate an expression involving functional uncertainty not with the f-structure labeled (1) in (4.18), but with the one labeled (5). Essentially, the idea would be that GF5 'looks outward' for a higher GF1 to identify itself with. If $\uparrow$ is now taken to be associated with (5), the following equation produces the same result as the one in (4.17):

(4.19) $((\text{GF}^* \text{GF5} \uparrow) \text{GF1}) = \uparrow$

Prefixing a regular expression to an f-structure identifier such as $\uparrow$ picks out a set of structures which contain the first structure. The expression $(\text{GF}^* \text{GF5} \uparrow)$ stands for any of the nodes labeled (1), (2), (3), or (4) in the f-structure in (4.18), since these are the f-structures containing a path specifiable by the regular expression $\text{GF}^* \text{GF5}$ which leads to the structure picked out by $\uparrow$ and labeled (5).

Since the expression $(\text{GF}^* \text{GF5} \uparrow)$ stands for an f-structure, we can specify a path within it, just as we do in the case of the expression $(\uparrow \text{SUBJ})$, where $\uparrow$ stands for an f-structure and SUBJ is a path of length one through it. Consider the f-structure displayed in (4.20). If $\uparrow$ is the f-structure labeled (5), the expression

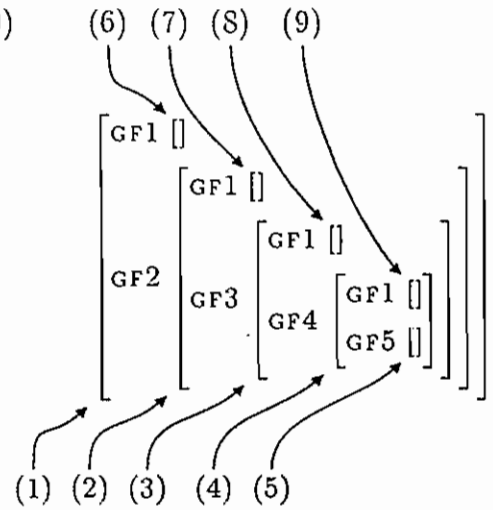
((GF* GF5 ↑) GF1)

picks out (among other f-structures) (6), the GF1 of f-structure (1). (1) contains (5), and through (1) there is a path to (5) characterizable by the expression

GF* GF5

Other f-structures picked out by the expression in (4.19) are the ones labeled (7), (8), and (9).

(4.20)

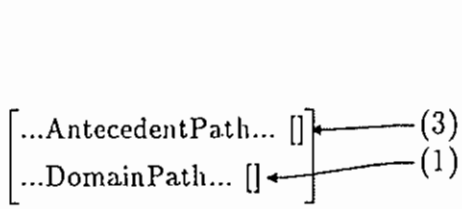


In the following, I will refer to the path prefixed to an f-structure identifier as the DomainPath, and the path through the containing f-structure as the AntecedentPath:

(4.21) ((DomainPath ↑) AntecedentPath)

Intuitively, the DomainPath is a path through some set of f-structures that ends at ↑. The AntecedentPath is a path through one of those f-structures. In the following, the f-structure ↑ is labeled (1), the f-structure (DomainPath ↑) is labeled (2), and the f-structure at ((DomainPath ↑) AntecedentPath) is labeled (3):

(4.22)



Constraining equations (Kaplan and Bresnan 1982, p. 207 ff.), including negative constraining equations, may also involve functional uncertainty, as in the following:

(4.23) $((\text{GF}^* \uparrow) \text{GF}) \neq \downarrow$

The interpretation of this equation is that $\downarrow$ is distinct from *every* f-structure reachable by $((\text{GF}^* \uparrow) \text{GF})$. The negation of an equation involving functional uncertainty has the effect of negating an existentially quantified expression: a negative constraint with functional uncertainty requires that there be *no* path picked out by the regular expression that makes the equation true. Returning to the observation that an expression involving functional uncertainty is equivalent to a (possibly infinite) disjunction of expressions involving paths picked out by the regular expression, negating the disjunction follows DeMorgan's Law in its equivalence to a (possibly infinite) conjunction of negative constraints.

4.3 Anaphoric Binding Equations

Inside-out functional uncertainty allows the statement of domain and antecedent conditions for anaphoric elements. These constraints are statable in terms of *binding equations*. Binding equations involve expressions of the following general form:

((DomainPath $\uparrow$) AntecedentPath)

This expression picks out the elements that are possible antecedents of the anaphor or the elements with which the anaphor may not cooccur.

Domain and antecedent constraints consist of restrictions on the two paths in this expression. Constraints on the DomainPath determine the domain of binding or

disjointness of the anaphoric element; constraints on the AntecedentPath determine the permissible or impermissible grammatical function of its antecedent.

Positive and negative constraints are expressible in terms of binding equations. Negative disjointness constraints are statable in terms of negative constraining equations involving functional uncertainty; positive binding constraints are statable in terms of positive constraining equations.

4.3.1 Binding Conditions

As we have seen, anaphors may require that their antecedents appear within a domain determined by the presence of a predicate, of a subject, or of tense. Anaphors may also require that their antecedent be a subject. I refer to constraints of this type as *positive* constraints; positive constraints are constraints on where an anaphor must appear in relation to its antecedent or what grammatical function its antecedent must bear. Constraints on impermissible antecedents for anaphors are referred to as *negative* constraints.

Anaphors with a positive domain requirement are associated with an equation of this form:

$$((\text{DomainPath} \uparrow)\ \text{AntecedentPath})_\sigma =_c \uparrow_\sigma$$

where the DomainPath may be required to meet certain conditions. The f-structure (DomainPath $\uparrow$) contains the anaphor by definition; it also contains the antecedent, reachable from the f-structure (DomainPath $\uparrow$) by following the path AntecedentPath. In other words, (DomainPath $\uparrow$) is the domain in which both the anaphor and its antecedent appear. *Domain* constraints consist of restrictions on DomainPath.

The DomainPath is also required to be nonempty. This has the effect that only f-structures that *properly* contain the anaphor are considered as anaphoric domains — that only f-commanding elements are eligible as antecedents for an anaphor — and this seems to be the right result.

As shown in Chapter 2, some anaphors require an subject antecedent. Recall that the AntecedentPath carries information about what grammatical function the antecedent bears; *antecedent* constraints consists of restrictions on the AntecedentPath.

4.3.2 Disjointness Conditions

We have noted that some anaphors are subject to disjointness conditions within some domain: the Coargument Disjointness Condition states that an anaphor may not corefer with a coargument, for example. Other anaphors are subject to disjointness conditions with respect to subjects: this is the Subject Disjointness Condition. I refer to constraints such as these as *negative* constraints. An anaphor with a negative constraint must be disjoint from elements in a particular domain or bearing a particular grammatical function.

In the simple model we are using, disjoint reference is interpreted as nonidentity at the semantic level. For most of the cases we are considering, where a singular name or definite noun phrase takes a pronoun or reflexive as its antecedent, this produces the right results; it should be noted, though, that in the case of a plural antecedent or a plural anaphor the wrong predictions are made. For plurals, not identity but referential overlap is important in applying disjointness constraints.

We have noted that a negative constraining equation with functional uncertainty has the effect of requiring that *no* solution be available that satisfies the equation. For disjointness conditions, this is the effect we want, as we now illustrate.

A sentence such as *John$_i$ saw him$_i$* is ungrammatical, since the pronominal *him* is coreferent with a coargument;[4] *him* is subject to the Coargument Disjointness Condition. This example does not, however, give us an answer to a question about the nature of this restriction: Does a negative condition necessarily impose disjointness, or just a requirement for an antecedent outside the negative domain? That is, does the Coargument Disjointness Condition consist simply of an instruction to look outside the nucleus in which the anaphor appears for an antecedent?

[4]We will see that this condition is actually too strong; *him* may not corefer with a coargument that is *superior*, in a sense to be defined in Chapter 5.

The answer is that negative conditions are requirements for disjointness within the relevant domain. Consider this sentence:

(4.24) *John$_i$ said that he$_i$ saw him$_i$.

The pronominal *him* is subject to the Coargument Disjointness Condition. Here there is a coreferent NP which is not a coargument; there is also a coreferent NP in the same nucleus. The sentence is ungrammatical. The Coargument Disjointness Condition, then, imposes a disjointness constraint on superior GF's in the same nucleus; it is not simply an instruction to look outside the nucleus for an antecedent.

Disjointness conditions will be of the following general form:

$$((\text{DomainPath} \uparrow) \text{ AntecedentPath})_\sigma \neq \uparrow_\sigma$$

An equation such as this one imposes a requirement of disjointness between an anaphoric element and *every* element picked out by the equation; as we noted in Section 4.2, the interpretation of a negative constraining equation is that no solution be found that satisfies the equation.

4.4 Domain Requirements

Domain requirements are constraints on values for DomainPath. That is, domain constraints consist of a specification of the f-structures in which the antecedent of a positively-specified anaphor can appear or of the f-structures which form a disjointness domain for a negatively-specified anaphor.

All domain requirements are statable in terms of binding equations involving expressions of the following general form:

(4.25) $((\text{DomainPath} \uparrow) \text{ AntecedentPath})_\sigma$

where DomainPath does not pass through an f-structure containing X.

The variable 'X' in the above equation will be replaced by a specific condition making reference to one of the three grammatical concepts in terms of which anaphoric binding

conditions can be stated: PRED, SUBJ, and TENSE. The Root S Binding Condition is the case in which the DomainPath is unconstrained and can pass through any sort of f-structure.

4.4.1 Coargument Condition

We have noted above that the Norwegian anaphor *seg selv* obeys the Coargument Binding Condition: its antecedent must be an argument of the same predicate. This condition is statable in this way:

(4.26) ((DomainPath $\uparrow$) AntecedentPath)$_\sigma$= $\uparrow_\sigma$

where DomainPath does not pass through the f-structure containing the PRED of which the anaphor is an argument.

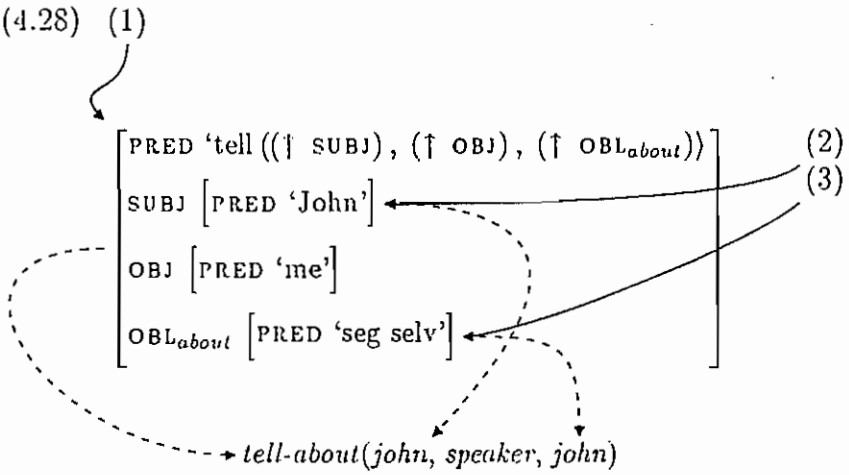
A formal statement of this requirement can be found in the Appendix. The antecedent of *seg selv* is the f-structure ((DomainPath $\uparrow$) GF). The semantic structure of *seg selv* is $\uparrow_\sigma$; this equation unifies $\uparrow_\sigma$ with the semantic structure of the antecedent. The antecedent must appear within the f-structure containing the PRED of which the anaphor is an argument (the *coargument domain* of the anaphor); usually the DomainPath is of length one, since the coargument domain f-structure is in the general case the immediately containing one.

As we have noted, sentences like the following are grammatical:

(4.27) Jon fortalte meg om seg selv
Jon told me about himself
'Jon told me about himself.'

The f-structure and semantic structure for this sentence are as follows:

(4.28)



Given the equation in (4.26):

- ↑ is labeled (3),
- DomainPath is OBL_{about}, a string of length 1,
- (DomainPath ↑), the coargument domain f-structure, is labeled (1),
- GF is SUBJ,
- ((DomainPath ↑) GF) is labeled (2), and
- the semantic structure of (2) is identical to that of (3).

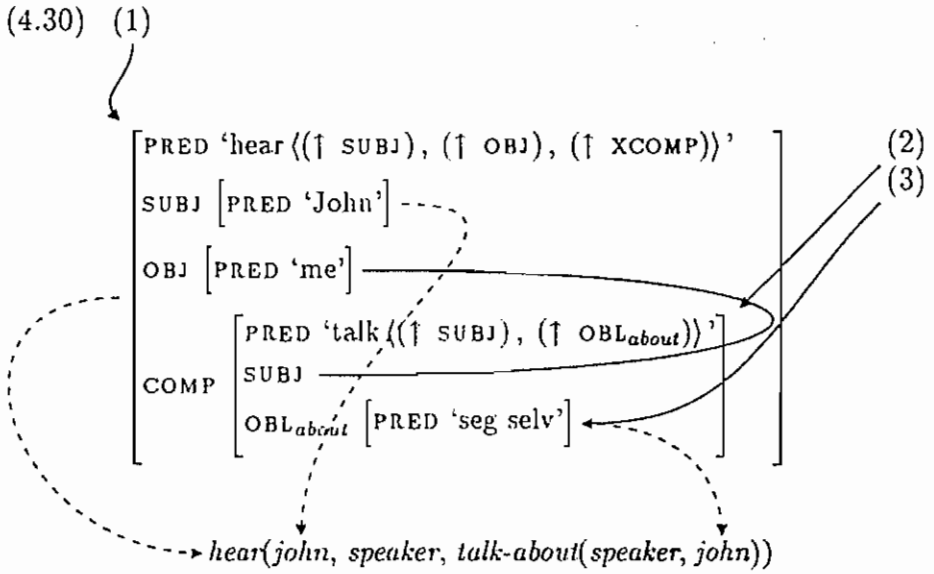
This configuration, then, satisfies the binding equation.

This sentence is ungrammatical, though, since the antecedent of the reflexive is not a coargument:

(4.29) *Jon hørte meg snakke om seg selv
Jon heard me talk about himself
'Jon heard me talk about himself.'

The f-structure and semantic structure for (4.29) are:

(4.30)



As above, (DomainPath ↑) is (1), the f-structure that contains the anaphor and its intended antecedent *John*. The f-structure corresponding to ↑ is (3). In this case, though, DomainPath is the string COMP OBL$_{about}$, since that is the path through (1) that leads to (3). The requirement on anaphors obeying the Coargument Binding Condition is that the anaphor and its antecedent appear in the coargument domain f-structure (DomainPath ↑), which contains the PRED of which the anaphor is an argument; in this case, that f-structure is labeled (2). The configuration is ruled out for this reason.

The English pronominal *him* is subject to the Coargument Disjointness Condition: it may not corefer with a coargument. The Coargument Disjointness Condition is just the negative counterpart of the Coargument Binding Condition:

(4.31) ((DomainPath ↑) GF)$_\sigma \neq \uparrow_\sigma$

where DomainPath does not pass through the f-structure containing the PRED of which the anaphor is an argument.

According to this condition, *him* must be disjoint from all arguments in the domain (DomainPath ↑), the coargument domain in which *him* appears.

This definition of the Coargument Condition also applies in some LFG analyses of prepositional phrases,[5] according to which a structure like the following is involved for a sentence like *John gave the book to Bill*, where *give* subcategorizes for the object of a prepositional phrase:

(4.32)
$$\begin{bmatrix} \text{PRED 'give}\langle(\uparrow \text{SUBJ}), (\uparrow \text{OBJ}), (\uparrow \text{OBL}_{goal}\ \text{OBJ})\rangle\text{'} \\ \text{SUBJ}\ \left[\text{PRED 'John'}\right] \\ \text{OBJ}\ \left[\text{PRED 'book'}\right] \\ \text{OBL}_{goal}\ \left[\text{OBJ}\ \left[\text{PRED 'Bill'}\right]\right] \end{bmatrix}$$

In an analysis like this one, the verb *give* subcategorizes for a SUBJ, an OBJ, and the OBJ of the preposition *to*. There is no PRED value associated with the preposition *to*; that is, the f-structure which is the value of the attribute OBL_{goal} does not contain a PRED.

The SUBJ and the OBL_{goal} must count as coarguments in this situation, to rule out the possibility of coreference in this sentence:

(4.33) *John$_i$ gave the book to him$_i$.

Given the definition provided above, the OBL_{goal} is a coargument of the SUBJ; in this case, the DomainPath which defines the disjointness domain for *him* may be of greater than length one, so long as it does not pass through the f-structure containing the PRED 'give' of which the anaphor is an argument.

4.4.2 Minimal Complete Nucleus Condition

As we have observed, in some dialects of Marathi the reflexive *swataah* obeys the Minimal Complete Nucleus Binding Condition: it must be bound in the minimal domain containing it and a subject. For anaphors obeying this condition, the DomainPath of

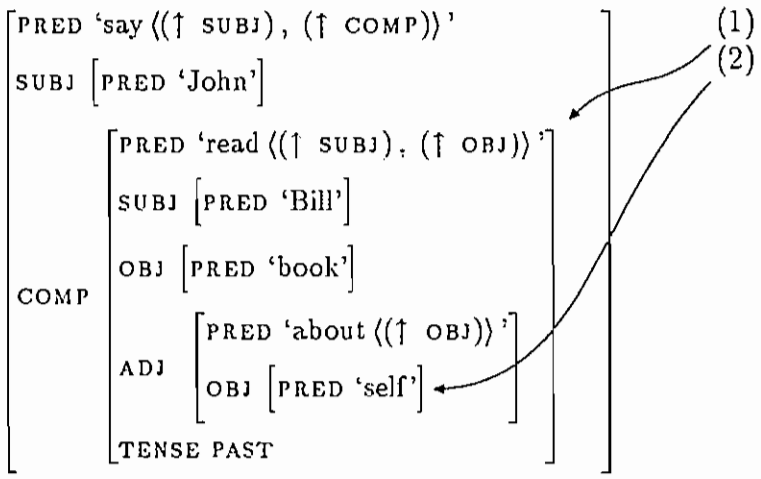
[5]See, for example, Kaplan and Bresnan (1982, pp. 196 ff.).

the binding equation may not pass through an f-structure which contains a subjective function.

Consider this sentence and its f-structure:

(4.34) John said that Bill read a book about himself.

(4.35)



The minimal f-structure containing the reflexive and a subjective function is the f-structure labeled (1). The reflexive, whose f-structure is labeled (2), and its antecedent must both be found within f-structure (1), with the result that *Bill* is a possible antecedent for *himself* in this sentence, while *John* is not. Recall that the DomainPath picks out the f-structure in which both the reflexive and its antecedent must be found; constraining the DomainPath produces the right results.

The Minimal Complete Nucleus Condition is statable as:

(4.36) $((\text{DomainPath}\ \uparrow)\ \text{GF})_\sigma = \uparrow_\sigma$

where DomainPath does not pass through an f-structure containing a subjective function.

A formal definition of this requirement appears in the Appendix.

4.4.3 Minimal Finite Domain Condition

In one dialect, the Marathi reflexive *swataah* obeys the Minimal Finite Domain Binding Condition: it must be bound in the minimal finite domain in which it appears.

The Minimal Finite Domain Condition is also statable as a constraint on the DomainPath, since, like the Minimal Complete Nucleus Condition, it determines the f-structure in which both the anaphor and its antecedent are found. For the Minimal Finite Domain Condition, the requirement is:

(4.37) ((DomainPath $\uparrow$) GF)$_\sigma$= $\uparrow_\sigma$

where DomainPath does not pass through an f-structure containing a TENSE attribute.

A formal definition of this requirement appears in the Appendix.

4.4.4 Root S Binding Condition

The Marathi reflexive *aapaṇ* obeys the Root S Binding Condition; its antecedent must appear in the same sentence, but there are no limits as to the possible distance between *aapaṇ* and its antecedent. This condition corresponds to a lack of constraints on the DomainPath.

(4.38) ((DomainPath $\uparrow$) GF)$_\sigma$= $\uparrow_\sigma$

The DomainPath is allowed to be of any length and to pass through any sort of f-structure.

We have noted that all of the anaphoric binding constraints discussed in the previous chapters can be stated with reference to three grammatical concepts: SUBJ, PRED, and TENSE. In this sense, the Root S Binding Condition is actually not a binding constraint at all: it amounts to the *lack* of a constraint on the domain in which the antecedent of an anaphor must be found.

4.5 Antecedent Requirements

Antecedent requirements are constraints on the grammatical function borne by the antecedent of the anaphor. In terms of the binding equations, antecedent requirements are constraints on AntecedentPath, since the AntecedentPath consists of the grammatical function of the antecedent.

4.5.1 Subjecthood Condition

The Marathi reflexive *swataah* must be bound to a subject: it obeys the Subject Binding Condition. The formulation of this constraint is straightforward: the AntecedentPath must be the singleton SUBJ.

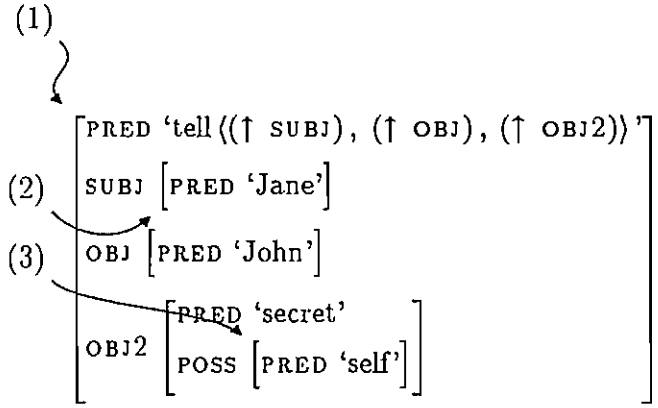
(4.39) $((\text{DomainPath} \uparrow)\ \text{SUBJ})_\sigma = \uparrow_\sigma$

As an example, consider this Marathi sentence:

(4.40) Jane ne John laa swataace gupit saangitle
Jane ERG John DAT self-GEN secret told
'Jane$_i$ told John$_j$ self's$_{i,*j}$ secret.'

Here is a simplified f-structure for this sentence:

(4.41)



The f-structure corresponding to the reflexive *swataah* is the value of the attribute POSS in the f-structure for the sentence, labeled (3). Assume that ↑ picks out this

f-structure. The f-structure represented by (DomainPath ↑) is the f-structure for which there is a path DomainPath (in this case, DomainPath is OBJ2 POSS) ending at the f-structure for *himself*. In the example at hand, (DomainPath ↑) is the f-structure for the entire sentence, marked (1). ((DomainPath ↑) SUBJ) picks out only one f-structure: (2), the value of the at the end of the path SUBJ inside f-structure (1).

Anaphors which obey the Subject Disjointness Condition must not corefer with f-commanding subjects. Recall the following Norwegian examples:

(4.42) *Ola snakket om ham selv
Ola talked about himself
'Ola$_i$ talked about himself$_i$.'

(4.43) vi fortalte Ola om ham selv
we told Ola about himself
'We told Ola$_i$ about himself$_i$.'

The constraint on anaphors obeying the Subject Disjointness Condition is the negative counterpart of the one above: the AntecedentPath is SUBJ, and the equation is a negative constraining equation.

4.5.2 Non-Subject Antecedency

It was noted above that the Root S Binding Condition consists of the lack of a constraint on the domain of an anaphor: the antecedent of an anaphor which obeys the Root S Binding Constraint may appear at an arbitrary distance from the anaphor.

Anaphors may also be unspecified as to the grammatical function of their antecedent. For example, the English anaphor *himself* may take an antecedent that is either a subject or a nonsubject:

(4.44) John$_i$ told Bill$_j$ about himself$_{i,j}$.

Anaphors that do not obey the Subject Binding or Disjointness Condition are associated with binding equations in which the grammatical function of the antecedent is unspecified:

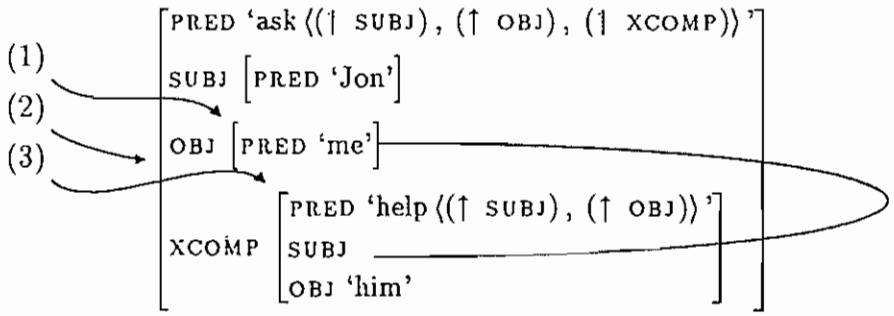
(4.45) ((DomainPath ↑) GF)$_\sigma$= ↑$_\sigma$

4.6 Anaphoric Binding and Functional Control

On the approach to anaphoric binding presented here, binding constraints are stated in terms of relations between positions in an f-structure. What, then, of cases in which an f-structure is the value of more than one attribute in a containing f-structure — that is, where structure-sharing is involved? Such a situation arises in cases of *functional control.*

In cases of functional control (or in other cases of identification of the value of two attributes in an f-structure) a single DomainPath might pick out two different f-structures. Consider this f-structure for the sentence 'John asked me to help him':

(4.46)



The node labeled (1) appears as the value of two different attributes in this f-structure: as the object of *ask* and as the subject of *help.* Assuming that ↑ represents f-structure (1), the expression given in example (4.47) picks out either the outermost f-structure, labeled (2), or the value of XCOMP, the one labeled (3).

(4.47) (GF ↑)

In other words, the structure reached by following the DomainPath GF is either the one in which (1) is an object or the one in which (1) is a subject.

4.6.1 Positive Binding Equations

Recall that a binding equation involving functional uncertainty holds if *some* solution to it can be found. In cases involving functional control and anaphors that obey

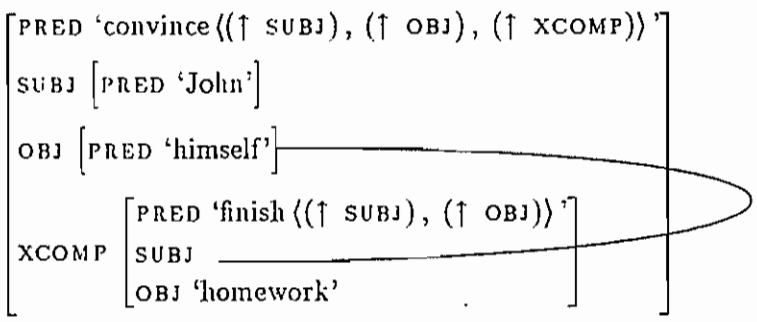
positive binding equations, then, the requirement is that an antecedent must bear the proper relation to some f-structure position in which the anaphor appears.

Consider, for example, example (4.48):

(4.48) John$_i$ convinced himself$_i$ to finish his homework.

A skeletal f-structure for this sentence is:

(4.49) (1)



Binding conditions for *himself* are satisfied here; the antecedent of *himself* is *John*. *Himself* must be bound in the minimal complete nucleus, the minimal domain containing a subject. That is, the antecedent of *himself* must be a superior element that lies in a domain characterizable by a DomainPath that does not pass through a f-structure with a SUBJ. In this case, there is such a path; the relation between *John* and *himself* is characterized by this equation (assume that ↑ is the f-structure for *himself*):

(4.50) $((\text{OBJ}\ \uparrow)\ \text{SUBJ})_\sigma = \uparrow_\sigma$

where (OBJ ↑) picks out the outermost f-structure, labeled (1).

It can also happen that the antecedent for an anaphor is involved in functional control, so the f-structure for the antecedent appears in more than one position in a containing f-structure. As long as there is some position occupied by the antecedent that satisfies the positive constraints, and as long as no negative constraints are

violated, anaphoric binding conditions are satisfied. In example (4.51), the antecedent for the anaphor appears as the object of *convince* and the subject of *hit*:

(4.51) John convinced Bill$_i$ to hit himself$_i$.

The f-structure for example (4.51) is:

(4.52)
$$\begin{bmatrix} \text{PRED } \text{'convince}\langle(\uparrow \text{ SUBJ}), (\uparrow \text{ OBJ}), (\uparrow \text{ XCOMP})\rangle\text{'} \\ \text{SUBJ } \left[\text{PRED 'John'}\right] \\ \text{OBJ } \left[\text{PRED 'Bill'}\right] \\ \text{XCOMP } \begin{bmatrix} \text{PRED } \text{'hit}\langle(\uparrow \text{ SUBJ}), (\uparrow \text{ OBJ})\rangle\text{'} \\ \text{SUBJ} \\ \text{OBJ } \left[\text{PRED 'himself'}\right] \end{bmatrix} \end{bmatrix}$$

Recall that *himself* obeys the Minimal Complete Nucleus Binding Requirement: it must be bound in the minimal domain with a subject. Its antecedent is *Bill*, the subject of *hit*. The binding equation that relates *Bill* and *himself* is:

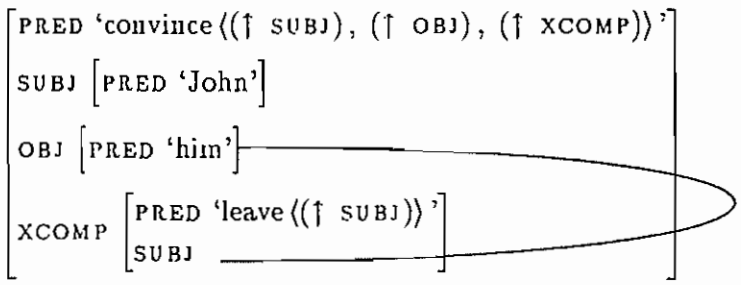
(4.53) $((\text{OBJ } \uparrow) \text{ SUBJ})_\sigma = \uparrow_\sigma$

The fact that the f-structure for *Bill* is also the value of a grammatical function outside the binding domain is not relevant; again, the requirement for *himself* is that *some* way to satisfy the binding equation must exist.

4.6.2 Negative Binding Equations

If an anaphor obeys negative constraints, the requirement is that *no* solution be found for a binding equation involving functional uncertainty. Consider the following sentence and its f-structure:

(4.54) *John$_i$ convinced him$_i$ to leave.



The pronominal *him* obeys the Coargument Disjointness Condition: it may not corefer with a coargument. Here, the constraint is violated. *Him* is an argument of both *convince* and *leave*; it is not coreferent with a coargument of *leave*, but it is coreferent with the subject of *convince*. This is not permitted, and the sentence is ungrammatical.

4.7 Conclusion

I have presented an inventory of binding constraints and described a way of formalizing these constraints. Functional uncertainty permits the description of possible syntactic relations between an anaphor and the set of its possible antecedents; the coreference relation between an anaphor and its antecedent is represented as identity of semantic structure, by use of the σ projector relating the semantic projection to the f-structure.

Given the formal vocabulary now at our disposal, we turn now to an examination of the systems of constraints on the antecedent-anaphor relation. For example, the antecedent must be *superior* to the anaphor in a sense which we will discuss in Chapter 5. Interactions among the various binding constraints will also be discussed.

Chapter 5

Anaphoric Binding Universals

In the preceding chapters, evidence was presented for a universally-available inventory of binding constraints definable in terms of the three grammatical concepts SUBJ, PRED, and TENSE. Chapter 4 showed how these constraints can be modeled equationally. Combinations of these primitive anaphoric binding requirements are involved in composing the full range of binding constraints obeyed by the anaphors discussed in Chapter 2.

In the following, it will be shown that domain and antecedent conditions are not separate requirements, but two aspects of a single, complex positive or negative constraint. The complex positive constraint picks out the set of possible antecedents of an anaphoric element, and the complex negative constraint picks out the set of elements with which an anaphoric element may not corefer.

Other interactions among binding requirements involve the relation between these complex positive and negative constraints. A pronoun may be subject to both positive and negative binding constraints; we will see that the domain in which a negative constraint holds is never wider than the domain in which the positive constraints hold. In this sense, the positive domain constraint is primary, in that all other constraints hold only within the positively-specified domain.

General syntactic constraints on anaphoric binding will also be defined; for example, a notion of *superiority* is relevant in determining permissible or impermissible antecedents for an anaphoric element. Only elements that are superior to an anaphor are considered when binding constraints are applied.

Additionally, conditions on the form of binding requirements will be shown to rule out some possible types of requirements. For example, the *Locality Condition* requires that binding constraints always refer to local elements, never exclusively to nonlocal elements.

5.1 Equational Binding Constraints

In Chapter 2, evidence was presented to show that anaphors may be lexically specified as to the domain in which they are bound or free and as to the permissible and impermissible grammatical functions of their antecedents. Some anaphors obey a combination of several of the binding requirements that have been shown to exist. In the following, we will see that the positive and negative binding requirements which each anaphor obeys are not independent of each other. Anaphors can obey a single complex positive requirement, a single complex negative requirement, or both.

5.1.1 Positive Constraints

Some anaphors obey both a positive domain constraint and a positive antecedent constraint. Recall, for example, that the short-distance Marathi reflexive *swataah* must be bound to a subject; in one dialect, it must be bound in the minimal nucleus containing a subject. The following examples are ungrammatical in that dialect:[1]

(5.1) a. *John mhaaṇaalaa ki aapaṇ aaplyaaši swataahbaddal bolle paahije
John said that we to self about self talk should
'John$_i$ said that we should talk to self$_i$ about self$_i$.'

b. *John nii aamhaalaa aaplyaaši swataahbaddal bolaay-ci vinanti keli
John ERG us-ACC to self about self talk-of request did
'John$_i$ requested us to talk to self$_i$ about self$_i$.'

In these examples, *swataah* is coreferent with a coargument, *aapaṇ*, satisfying the domain constraint; it is also coreferent with a subject, *John*, satisfying the antecedent

[1]The occurrence of *aapaṇ* as the subject of the subordinate clause in (5.1) is as a first person plural pronoun; recall that *aapaṇ* can also be used as a first person or second person honorific pronoun.

constraint. Each of its binding requirements is satisfied, though the requirements are satisfied by different elements. However, the sentences are ungrammatical. The domain and antecedent requirements cannot, then, be satisfied by different elements; a single element must satisfy both requirements.

Similar facts hold in Norwegian. The Norwegian anaphor *seg selv* obeys the Subject Binding Condition and the Coargument Binding Condition; its antecedent must be a subject, and it must be bound to a coargument. Example (5.2), in which the requirements are satisfied by two different elements, is ungrammatical:

(5.2) *Martin ba oss snakke til ham om seg selv
Martin asked us to talk to him about self
'Martin$_i$ asked us to talk to him$_j$ about himself$_i$.'

Seg selv is coreferent with a coargument, *ham*, although *ham* is not a subject. *Seg selv* is also coreferent with a subject, *Martin*, although *Martin* is not a coargument. *Martin* satisfies the Subject Binding Condition, and *ham* satisfies the Coargument Binding Condition. Nevertheless, the sentence is ungrammatical.[2]

The conclusion to be drawn is that the two binding requirements may not be satisfied by different elements; a single antecedent must satisfy both binding conditions. Example (5.3) is grammatical because both the Subject Binding Condition and the Coargument Binding Condition are satisfied by the antecedent of *seg selv*, *Martin*:

(5.3) Martin fortalte meg om seg selv
Martin told me about self
'Martin$_i$ told me about himself$_i$.'

In terms of the binding equations introduced in Chapter 4, the observation that all positive binding requirements must be satisfied by the same antecedent is statable as a constraint that an anaphor may be associated with at most one positive equation characterizing the set of its possible antecedents. Domain conditions constrain the DomainPath of the unique binding equation, and antecedent conditions constrain the AntecedentPath. The two requirements are not given separately, but appear as

[2]See Section 5.8, though, for further discussion of similar examples; some sentences that are structurally similar to this one are, mysteriously, grammatical.

constraints on two different portions of a single anaphoric binding equation. This positive equation, when it is present, picks out the f-structure for the antecedent of the anaphor.

It was noted in Chapter 2, Section 2.3, that certain combinations of binding requirements are not allowed. In particular, given that only one positive binding equation is associated with each anaphor, it is not possible for an anaphor to be associated with more than one positive domain constraint or more than one positive antecedent constraint; the DomainPath and AntecedentPath of the binding equation cannot be constrained in more than one way. An anaphor may obey at most one positive domain constraint and one positive antecedent constraint.

5.1.2 Negative Constraints

I also assume that an anaphor is associated with at most one negative equation characterizing the set of elements from which it must be disjoint. It is difficult to demonstrate that this assumption holds by appealing to positive evidence, since negative constraints have the effect of ruling out coreference between an anaphor and all elements picked out by the negative equation with which it is associated. One negative equation is sufficient to rule out a possible antecedent; if an antecedent for an anaphor were ruled out by two equations rather than one, the effect would be the same.

However, evidence for this restriction does exist, in that no anaphor seems to obey constraints that are only statable in terms of more than one negative equation. The substantive effect of this restriction, then, is to rule out anaphors which must be disjoint from all elements in a local domain and, additionally, all elements in a wider domain which bear a particular grammatical function. These requirements would be statable only in terms of two separate binding equations, and such anaphors are not attested.

The negative and positive constraints on anaphoric binding which were outlined in Chapter 2 are, then, statable in terms of complex binding equations. An anaphor may be associated with a positive equation, a negative equation, or both; the binding

conditions constrain either the AntecedentPath or the DomainPath of these equations.

5.2 The Locality Condition

It has been shown that constraints on domains of binding or disjointness are uniformly characterizable in terms of the presence of a SUBJ, of TENSE, or of the PRED of which the anaphor is an argument. The complex positive and negative requirements statable in terms of these concepts always refer to local elements, never exclusively to nonlocal ones.

In particular, evidence from Marathi and Norwegian shows that there are anaphors that require an antecedent in a higher clause; Marathi *aapaṇ* and Norwegian *seg* are examples of such anaphors. It may seem that binding requirements for these anaphors refer to a nonlocal domain — that is, the higher clause; however, as we have seen, these anaphors are actually subject to a complex combination of binding and disjointness requirements. They are required to be disjoint from elements in a local domain but coreferent with elements in a wider domain.

In general, the requirement for a nonlocal antecedent can always be reduced to a combination of a positive requirement (a requirement for an f-commanding antecedent of some type) and a negative requirement (a requirement for disjointness within some local domain). That is, there are no anaphors that *require* an antecedent in a higher clause but also *allow* an antecedent in a local domain. In this sense, constraints on anaphoric binding are local; binding constraints never refer exclusively to nonlocal elements. I will refer to this condition as the *Locality Condition.*

Notice that it would in principle be possible to define such a requirement in terms only of the concepts SUBJ, PRED, and TENSE; the Locality Condition does not follow from the fact that only these three concepts are allowed in statements of anaphoric binding constraints. For example, the Locality Condition rules out a constraint stating that an anaphor must have an antecedent outside of the coargument domain but inside the Minimal Tensed Domain.

Noyer (1989) presents a very interesting study of anaphoric binding constraints, observing a number of types of anaphors which are unattested; many of these types

are ruled out by the Locality Condition in combination with the requirement that anaphoric elements obey at most one positive and one negative requirement.

For example, Noyer notes that what he calls a 'double anaphor' is unattested; such an anaphor must have two antecedents, one local and one nonlocal. This type of anaphor would be associated with two binding equations, a situation that we have ruled out on other grounds; one of the equations would have a local domain, while the other would refer exclusively to a nonlocal domain. Equations that refer exclusively to a nonlocal domain are ruled out by the Locality Condition.

Also ruled out by this constraint is what Noyer refers to as an 'anti-double-anaphor', an element which is prohibited from having two binders, one local and one nonlocal. Again, such an anaphor would be associated with two disjointness equations, one referring exclusively to local elements and the other referring exclusively to nonlocal elements.

Each anaphoric element may, then, be associated with at most two complex constraints, one positive and one negative, stating the binding constraints which it obeys. The domain of binding or disjointness is stated in terms of maximality conditions, never minimality conditions. In terms of the equational representation of binding constraints, the Locality Condition amounts to a prohibition against constraints of the form 'DomainPath MUST go through an f-structure of a certain type' or of the form 'DomainPath MUST be of a certain minimal length'. Constraints on paths that pick out proper binding domains are always maximality constraints on the form of the DomainPath, never minimality constraints; a path in a binding equation never picks out a non-local domain.

In the case of negative equations, too, constraints on the DomainPath are always maximality constraints. There do not seem to be any anaphors which are required to be disjoint from elements in a higher clause but which may optionally corefer with elements in a local domain.

5.3 Binding Conditions and Morphological Form

On the analysis presented here, binding requirements are lexically associated with each anaphoric element. Further evidence for the lexical character of binding constraints comes from the observation that the kinds of binding constraints associated with an anaphoric element can depend on the morphological form of the element. This is true in two senses: anaphors that are polymorphemic tend to be associated with a different set of constraints from those that are monomorphemic; and the particular morphemes that compose a polymorphemic anaphoric element can each contribute to the complete set of binding properties of the anaphor.

5.3.1 Pronominal and Compound Anaphors

Faltz (1985) and, more recently, Kiparsky (1989) note a correspondence between the morphological form of an anaphor and the binding constraints which it obeys. Both domain and antecedent constraints are at least partially predictable on the basis of the form of the anaphor.

Kiparsky discusses four types of anaphoric elements: *pronominal*, monomorphemic anaphors, consisting of an affix, clitic, or pronominal form (such as Norwegian *seg*); *compound*, polymorphemic anaphors, which are lexical anaphors having full NP status (such as English *himself*); *adverbial* anaphors, which are often used emphatically; and *intransitivizing* anaphors, which induce a modification in the argument structure of a predicate. The anaphors whose binding constraints were discussed in the previous chapters are all either pronominal or compound anaphors.

Kiparsky (1989, p. 56) notes the following generalizations pertaining to pronominal and compound anaphors:

- Pronominal anaphors always require subject antecedents; compound anaphors usually permit nonsubject antecedents.
- Pronominal anaphors may allow long-distance binding; compound anaphors are always bound in a local domain.

These conditions can be stated in terms of generalizations over the type of binding requirements associated with anaphors of certain morphological forms.

In particular, pronominal anaphors obey the Subject Binding Condition, but they do not necessarily require an antecedent in any particular domain. In contrast, compound anaphors do not obey the Subject Binding Condition; nonsubject antecedents are permitted. However, the domain for compound anaphors is always relatively local.

In the same vein, Bresnan et al. (1985) note that the Minimal Finite Domain Parameter holds only for anaphors that obey the Subject Antecedent Condition — that there is a correlation between the possibility for (relatively) long-distance binding and the requirement that the antecedent be a subject. They posit a parameter which is set for the language as a whole: the Minimal Finite Domain parameter. If a language is positively specified for the Minimal Finite Domain parameter, any anaphor that respects the Subjecthood Condition is also subject to the Minimal Finite Domain parameter. In Bresnan et al.'s terminology, only +SUBJECT anaphors are subject to the minimal finite domain parameter; there is a correlation between the requirement for a SUBJ antecedent and the possibility for relatively long-distance binding.

The following possible requirements on anaphoric binding are attested:

- (Relatively) short-distance anaphors (those subject to the Minimal Complete Nucleus condition or the Coargument Condition) that must be bound to a SUBJ. (Norwegian *seg selv*; Marathi *swataah* [more restricted dialect])
- (Relatively) short-distance anaphors that are bound to any superior GF. (English *himself*)
- Long-distance anaphors (those subject to the Minimal Finite Domain condition, or those that have no domain restriction) that must be bound to a SUBJ (or a logical subject). (Marathi *aapaṇ*)

There do not seem to be any true long-distance anaphors that are not restricted as to the type of their antecedent.[3]

[3]The Greek anaphor *o idhios* seems to constitute a counterexample to this claim; as discussed

These facts fit with the tendencies noted by Kiparsky: either the antecedent or the domain is always constrained in an anaphoric binding requirement. If an anaphor does not obey the Subject Antecedent Condition and can take as its antecedent an element bearing any grammatical function, then the domain in which the anaphor must be bound is constrained. If the anaphor must have a subject as its antecedent, the domain in which it is bound need not be constrained — it can be bound in a wide domain.

Some pronouns have no restriction as to the type of their antecedent; Enç (1989) analyzes the Turkish pronoun *kendisi* in this way. It may appear without a sentence-internal antecedent (Enç 1989, p. 58):

(5.4) kendisi geldi
he/she came
'He/She came.'

---

in Chapter 3, Section 3.1.4.3, Iatridou (1986) analyzes *o idhios* as 'free in the governing category but bound in the whole S'. She shows that it may be bound to an antecedent with any grammatical function [p. 769]:

(a) O Yanis ipe ston Costa oti i Maria aghapa ton idhio
John said to Costa Comp Mary loves self
'John$_i$ told Costa$_j$ that Mary loves himself$_{i,j}$.

In this example, *o idhios* is bound to either a subject or an object. If *o idhios* must be bound, but to any higher antecedent, it constitutes a counterexample to the claim that all anaphors are restricted either as to their domain or as to the grammatical function of their antecedent.

However, recall that there are cases in which *o idhios* can appear with no syntactic binder (Iatridou 1986, p. 770):

(b) O idhios pighe sto scholio
self went to school
'Himself went to school.'

(c) (Do you want to talk to the doctor or to his nurse?)
Theloume na milisoume ston idhio
(We) want talk to self
'We want to talk to himself.'

Iatridou analyzes these cases as involving not the anaphoric use of *o idhios*, but as 'emphatic' or 'adjectival'. If these cases are analysed as uses of the same form as the apparently 'bound' cases such as (a), however, then *o idhios* does not require a binder at all, but is subject only to negative constraints, similar to the English pronoun *him*. Under such an analysis, *o idhios* would not constitute a counterexample to the claim.

It can also be bound by an NP with any grammatical function and within any domain:

(5.5) Ali Ayşe-ye kendisin-den bahsetti
Ali Ayşe-DAT he/she-ABL talked
'Ali$_i$ talked to Ayşe-ye$_j$ about him/her$_{i,j,k}$.

(5.6) Ali [Ayşe-nin kendisin-e kızmasın-a] şaştı
Ali Ayşe-GEN he/she-DAT be-angry-DAT was-surprised
'Ali$_i$ was surprised that Ayşe-ye$_j$ was angry at him/her$_{i,j,k}$.

The existence of pronouns of this sort makes it clear that the generalizations discussed above concern possible types of binding requirements, not anaphors that can take nonlocal antecedents. The Turkish pronoun *kendisi* can have an antecedent in a higher clause. However, the grammatical function of this antecedent is not constrained; *kendisi* is associated with no binding constraints whatsoever, and is free to take non-f-commanding, discourse antecedents. The generalization having to do with long-distance binding only applies to anaphors whose antecedent must stand in a certain structural relation to it; that is, it only applies to anaphors that obey a positive binding requirement.

In sum, the constraints that each type of anaphor obeys tend to be of particular forms, depending on the morphological form of the anaphor. Compound anaphors tend to obey domain and not antecedent constraints; pronominal anaphors tend to obey antecedent but not domain constraints.

5.3.2 Binding Conditions and Morphological Compositionality

Hellan (1988) notes that a kind of morphological compositionality holds to some extent of the anaphors in Norwegian. As we have seen, anaphors in Norwegian obey the following set of constraints:

seg: Subject Binding Condition, Coargument Disjointness Condition, Minimal Finite Domain Binding Condition

seg selv: Subject Binding Condition, Coargument Binding Condition

ham selv: Subject Disjointness Condition, Minimal Complete Nucleus Binding Condition

As Hellan notes, both *seg* and *seg selv* must be bound by a subject, while *ham selv* may not. In other words, anaphors containing the form *seg* are bound by subjects, while those not containing *seg* are disjoint from subjects.

On another dimension, both *seg selv* and *ham selv* are bound in a relatively local domain, whereas *seg* is bound in a wider domain. The presence of *selv* seems to induce a requirement for local binding.

In terms of the constraints that have been defined in the preceding chapters, anaphors containing the form *seg* obey the Subject Antecedent Condition: their antecedent is a subjective function. Anaphors containing the form *selv* are constrained to require a local domain. It should be noted, though, that the compositionality is not perfect, in that the presence of *selv* does not signal exactly the same domain for *seg selv* and *ham selv*; both elements are bound in a relatively narrow domain, but their domains are not the same.

The absence of *seg* induces a requirement for disjointness from a superior subject. This seems to be due to a blocking or complementarity effect that induces a requirement of subject disjointness in anaphoric elements not containing *seg*. Similarly, the absence of *selv* induces a requirement of disjointness from elements in a narrow domain: *seg* must be disjoint from its coarguments. The correlation between morphological form and binding requirements is not perfect; nevertheless, the tendencies pointed out by Hellan do hold.

5.4 Complex Binding Constraints

The complex binding requirements which anaphoric elements can obey are statable in terms of both domain requirements and antecedent requirements. Here we exemplify the possible combinations of binding conditions which make up the complex binding requirements that anaphoric elements may be subject to.

5.4.1 The Positive Requirement

We have seen that the English reflexive *himself* obeys the Minimal Complete Nucleus Binding Condition: it is bound in the minimal domain containing a subject. It is not associated with any negative conditions — it need not be disjoint from any class of elements. The complete set of binding conditions for *himself* may be stated as:

(5.7) $((\text{DomainPath} \uparrow)\ \text{GF})_\sigma =_c \uparrow_\sigma$

where DomainPath does not pass through an f-structure containing a subjective function.

The DomainPath is constrained not to pass through an f-structure with a subject, so *himself* must find its antecedent in the minimal complete nucleus. Since the AntecedentPath, GF, represents any member of the set of governable grammatical functions, the grammatical function of the antecedent is not constrained.

As we have seen, some anaphoric elements obey both a positive domain condition and a positive antecedent condition. Data presented in Chapter 2 show that the Norwegian anaphor *seg selv* obeys the Subject Binding Condition and the Coargument Binding Condition. The requirement for *seg selv* may be stated as:

(5.8) $((\text{DomainPath} \uparrow)\ \text{SUBJ})_\sigma =_c \uparrow_\sigma$

where DomainPath does not pass through the f-structure containing the PRED of which the anaphor is an argument.

Both the AntecedentPath and the DomainPath are constrained in this equation. The DomainPath must pick out the coargument binding domain, so only coarguments are appropriate binders. Further, the binder must be a subject, since AntecedentPath must be SUBJ. Taken together, the constraints entail that the binder of *seg selv* must be a subject coargument.

Marathi *swataah* also obeys two positive binding requirements; it is associated (in one dialect) with the Minimal Complete Nucleus Binding Condition and the Subject Binding Condition. Like *himself*, *swataah* obeys no negative conditions. Conditions on *swataah* in this dialect are given by a single positive equation:

(5.9) $((\text{DomainPath} \uparrow) \text{SUBJ})_\sigma =_c \uparrow_\sigma$

where DomainPath does not pass through an f-structure with a subjective function

The domain in which *swataah* and its antecedent are found is determined by constraints on the DomainPath; in this case, the DomainPath may not pass through an f-structure with a subjective function. Further, the antecedent of *swataah* must be a subject; the AntecedentPath of the positive equation is the singleton SUBJ.

In the less restrictive dialect, Marathi *swataah* must be bound within the Minimal Finite Domain to a subject; it obeys the Minimal Finite Domain Binding Condition and the Subject Binding Condition in this dialect. Conditions on *swataah* in this dialect are:

(5.10) $((\text{DomainPath} \uparrow) \text{SUBJ})_\sigma =_c \uparrow_\sigma$

where DomainPath does not pass through an f-structure containing a TENSE attribute.

Given the inventory of binding constraints that have been examined thus far, another possibility exists for a positive binding requirement: an anaphoric element may be required to be bound to a subject in the root S in which it appears. In other words, an anaphoric element may be required to obey the Root S Binding Requirement and the Subject Binding Requirement. According to the analysis of Yoruba pronominals presented by Pulleyblank (1986), the Yoruba anaphor *òun* is such an element. The conditions for *òun* as described by Pulleyblank are:

(5.11) $((\text{DomainPath} \uparrow) \text{SUBJ})_\sigma =_c \uparrow_\sigma$

The antecedent of *òun* must be a subject, but the domain in which it must be bound is unconstrained; the antecedent of *òun* may appear in any higher clause.

In sum, the following combinations of binding requirements are possible:

	Coargument	MCN	MFD	Root
SUBJ	Norw. *seg selv*	Mar. *swataah*$_1$	Mar. *swataah*$_2$	Yor. *òun*
GF	[predicted]	Eng. *himself*	[predicted]	*

The rows of the table indicate antecedent possibilities: SUBJ stands for the Subject Binding Condition, and GF indicates an anaphor that may have an antecedent with any grammatical function. The columns indicate domain possibilities: the Coargument Domain, the Minimal Complete Nucleus, the Minimal Finite Domain, and the Root S.

In Section 5.3, it was noted that anaphors associated with a positive binding requirement are always constrained with respect either to the domain in which the antecedent is found or with respect to the grammatical function of the antecedent; this constraint rules out the possibility of an anaphor that is unconstrained as to the domain in which its antecedent may be found and that can have an antecedent with any grammatical function. This is indicated by the asterisk in the fourth column of the second row.

All of the other possibilities are predicted to exist. As yet, I am not aware of attested examples of anaphors which can take an antecedent of any grammatical function but which must be bound within the coargument domain or the minimal finite domain; however, the theory presented here predicts that such anaphors can exist.

5.4.2 The Negative Requirement

Anaphors may also obey negative domain and antecedent requirements; they may be required to be disjoint from all elements in some domain, or only those with a certain grammatical function.

Norwegian *ham* is an anaphor that obeys only a negative domain constraint, the Coargument Disjointness Condition, according to which it must be disjoint from superior coarguments. The negative requirement for *ham* is statable in this way:

(5.12) $[((\text{DomainPath} \uparrow) \text{AntecedentPath})_\sigma \neq \uparrow_\sigma]$

where DomainPath does not pass through the f-structure containing the PRED of which the anaphor is an argument.

There is no positive equation associated with *ham*; that is, there is no requirement for *ham* to stand in a binding relation with an antecedent. The negative equation constrains *ham* to be disjoint in reference from its coarguments.

The Yoruba pronominal form *ó* as described by Pulleyblank (1986) obeys a negative antecedent constraint: it must be disjoint in reference from every higher subject. The requirement obeyed by *ó* is statable as:

(5.13) $[((\text{DomainPath} \uparrow)\ \text{SUBJ})_\sigma \neq \uparrow_\sigma]$

According to this requirement, Yoruba *ó* is disjoint from higher subjects; the domain in which the requirement holds is unconstrained, so the requirement applies to every higher subject.

Just as it is possible for both the AntecedentPath and the DomainPath to be constrained in the case of the positive equation, the AntecedentPath and the DomainPath of a negative equation can both be constrained. In this sense, both positive and negative binding equations reflect both domain and antecedent constraints.

Norwegian *hans* obeys a negative condition of this nature. Recall that *hans* must be disjoint from the immediately superior subject; the negative equation associated with *hans* is:

(5.14) $[((\text{DomainPath} \uparrow)\ \text{SUBJ})_\sigma \neq \uparrow_\sigma]$

where DomainPath does not pass through an f-structure containing a subjective function.

A negative equation in which both the AntecedentPath and the DomainPath are constrained rules out coreference between an anaphor and elements in the negatively-specified domain which bear a certain grammatical function.

The combinations of negative constraints forming the complex negative requirement are:

	Coargument	MCN	MFD	Root
SUBJ	[predicted]	Norw. *ham selv*	[predicted]	Yor. *ó*
GF	Eng. *him*	[predicted]	[predicted]	*

All of these combinations except for the starred possibility are predicted to exist; however, in several cases I do not as yet know of any anaphoric elements which obey the conditions indicated.

5.4.3 Negative and Positive Requirements

We have seen that some anaphors obey both negative and positive requirements; these anaphors are associated with both negative and positive binding equations. Norwegian *seg* is such an anaphor: it is bound to a subject in the minimal finite domain, but must be free from an argument in the coargument domain. *Seg* is, then, associated with the following two equations:

(5.15) $((\text{DomainPath1} \uparrow) \text{SUBJ})_\sigma =_c \uparrow_\sigma$

where DomainPath1 does not pass through an f-structure containing a TENSE attribute.

$[((\text{DomainPath2} \uparrow) \text{AntecedentPath})_\sigma \neq \uparrow_\sigma]$

where DomainPath2 does not pass through the f-structure containing the PRED of which the anaphor is an argument.

The first equation states the positive requirement obeyed by *seg*, according to which it must be bound to a subject in the minimal finite domain; the second equation states the negative requirement, according to which it must be noncoreferent with its coarguments.

5.5 Positive and Negative Constraints: Interactions

Some anaphors may obey both negative and positive binding requirements; in particular, anaphors may be associated with a negative antecedent requirement and a

positive domain requirement. We will see that the negative antecedent requirement is required to hold only in the positively-specified binding domain.

Consider, for example, an anaphor such as Norwegian *ham selv*, which obeys the Subject Disjointness Condition (it is noncoreferent with superior subjects) and the Minimal Complete Nucleus Binding Condition (it is coreferent with a NP in the minimal nucleus containing a subject). The combination of these two conditions might in principle produce one of the following three complex requirements which *ham selv* must obey:

1. Bound in the minimal complete nucleus to a nonsubject
2. Bound in the minimal complete nucleus to a nonsubject and disjoint from the *immediately* superior subject
3. Bound in the minimal complete nucleus to a nonsubject and disjoint from *all* superior subjects

We have already seen that the first alternative is not a possibility; the Subject Disjointness Requirement imposes a requirement of disjointness. It is not sufficient for the anaphor to find a nonsubject antecedent; it must also be disjoint from an appropriate set of subjects. This is the reason for the ungrammaticality of Example (5.16) (Hellan 1988, p. 192), previously discussed in Chapter 4, Section 4.3.2:

(5.16) *Jon fortalte seg selv om ham selv
Jon told self about self
'Jon$_i$ told himself$_i$ about himself$_i$.'

Although there is a nonsubject noun phrase in the sentence which is a possible antecedent for *ham selv*, the sentence is ungrammatical; this is because *ham selv* is also coreferent with a subject, *Jon*, and this is forbidden by the Subject Disjointness Condition. In contrast, Example (5.17) is grammatical:

(5.17) Jon fortalte seg selv om seg selv
Jon told self about self
'Jon$_i$ told himself$_i$ about himself$_i$.'

The grammaticality of this sentence can be attributed to the availability of an acceptable antecedent for both reflexives; *seg selv* must be bound to a subject coargument. No negative binding conditions are associated with *seg selv*, and so no conditions are violated.

Given that *ham selv* must be disjoint from superior subjects within some domain, consider the following grammatical Norwegian example:

(5.18) Martin ba oss snakke til ham om ham selv
Martin asked us to talk to him about himself
'Martin$_i$ asked us to talk to him$_i$ about himself$_i$.'

Although *ham selv* is coreferent with a subject, *Martin*, the subject is outside the binding domain of *ham selv*. The requirement of disjointness from subjects holds, then, only within the domain within which *ham selv* must be bound. Coreference of *ham selv* with a subject outside that domain does not constitute a binding violation.

The following relation holds between the negative and the positive binding requirements:

(5.19) The domain of the negative constraint is not wider than the domain of the positive constraint.

In other words, negative binding constraints hold only in the positive binding domain; the positive binding domain is the relevant domain for all constraints, both positive and negative. It is, of course, possible for the negative domain to be *narrower* than the positive domain; this is the situation with Marathi *aapaṇ* and Norwegian *seg*, which must be disjoint from coarguments but bound in a wider domain.

Interestingly, this generalization can be recognized only when the negative equation is restricted in its application to subjective functions. The constraint holds vacuously when the negative constraint applies to every element in a domain, because it is in principle impossible for there to be a situation where the negative constraint rules out coreference from *every* element in some domain that is larger than the binding domain. If a reflexive were required to be coreferent with some element in a smaller domain and disjoint from every element in a larger domain, it could never find an

antecedent at all. Imagine, for example, a reflexive subject to the Coargument Binding Condition (it must be bound to a coargument) and to a constraint that might be called the Minimal Finite Domain Disjointness Constraint, where it must be disjoint from every element within the minimal finite domain in which it appears (*including* its coarguments). In this case, the negative domain is larger than the positive binding domain, and the reflexive is not able to find any antecedent that satisfies the negative and positive constraints, since it is required to be bound to an element in the same domain in which it must be disjoint from every element.

Returning to example (5.2), repeated here, it is evident that there is a conceptual similarity with the constraint just outlined:

(5.20) *Martin ba oss snakke til ham om seg selv
Martin asked us to talk to him about himself
'Martin$_i$ asked us to talk to him$_i$ about himself$_i$.'

As we observed, this sentence is ungrammatical; the reason is that *seg selv* must be bound to an element which is both a subject and a coargument. In this example, the two binding requirements are satisfied by different elements. *Seg selv* is coreferent with an NP, *ham*, in its binding domain, but *ham* is not a subject. *Seg selv* is also coreferent with a subject, *Martin*, but *Martin* is outside the binding domain and thus does not count as a proper binder. This example illustrates that antecedent constraints (both positive and negative) hold only within the positive binding domain — that there is a sense in which the positive binding domain is prior, and other constraints are required to hold within it. The constraint given in (5.19) can, then, be made stronger:

(5.21) All binding constraints hold only within the positive binding domain.

5.6 Superiority Effects

Binding constraints require coreference with or disjointness only from elements that are *superior*, in a sense to be defined here. In other theories, a notion of superiority is sometimes stated in terms of c-command, a relation of structural superiority between

two positions in a tree. In GB analyses of anaphoric binding, for example, the antecedent is required to c-command the anaphor. Similarly, GB analyses require that a pronominal be disjoint in reference not from every element, but only from those that c-command it. Hellan (1988) discusses various conditions which he subsumes under the rubric 'command'; all of these conditions involve a notion of the superiority of one argument relative to another.

The senses of superiority that will be examined here are both structural and semantic in nature. The F-Command Condition and the Noncontainment Condition, described in the following sections, are definable in terms of structural relations between the f-structure of the anaphor and that of its possible antecedents; these two conditions pick out a set of elements that are structurally superior to the anaphor. Only these elements are required to be coreferent with or disjoint in reference from the anaphor.

In addition to these structural conditions, a notion of superiority among coarguments of a predicate plays a role in anaphoric binding. The ranking among coarguments is determined by reference both to syntactic features such as obliqueness and to semantic/thematic features.

Other conditions also play a role in determining eligibility for antecedency of anaphoric elements; these requirements. too, involve a relation of superiority between two arguments. For example, linear precedence conditions clearly play a role in anaphoric binding. Example (5.22)a is not grammatical, although it involves simply a linear reordering of the grammatical example (5.22)b:

(5.22) a. *Mary talked about himself$_i$ to John$_i$.

b. Mary talked to John$_i$ about himself$_i$.

See Kameyama (1989) for an LFG analysis of the distribution of the Japanese overt and zero pronominals; she makes crucial use of the relation of *f-precedence*, a property based on the c-structure property of linear order.

5.6.1 F-Command

Anaphors that obey a positive binding constraint are those that must be bound to a superior element of some type or within some domain. Similarly, anaphors that obey a negative binding constraint must be disjoint from each member of a set of superior elements.

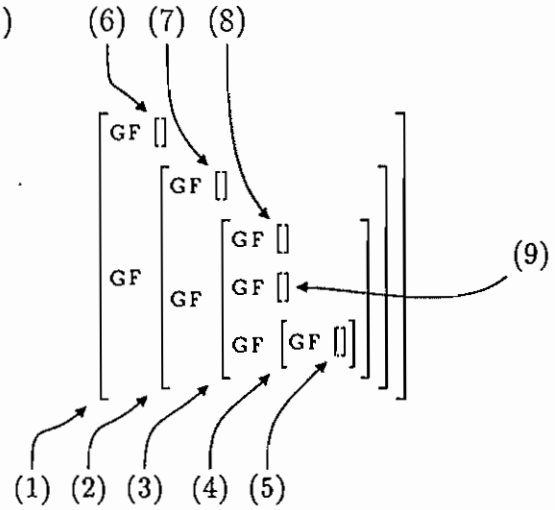
One sense of superiority which is relevant is the F-Command condition: an element superior to an f-structure f-commands it. Bresnan (1982a, p. 333 ff.) provides the following definition of f-command:

(5.23) For any occurrences of the functions α, β in an f-structure F, α *f-commands* β if and only if α does not contain β and every f-structure of F that contains α contains β.

Informally, the relation of f-command is analogous to the c-command relation, except that it is stated on f-structures instead of constituent structure trees.

To illustrate: In the f-structure shown in (5.24), the f-structures labeled (6), (7), and (8) f-command the f-structure labeled (9), since f-structures (1), (2), and (3) contain those structures and f-structure (8). However, f-structure (5) does not f-command (9), since (5) is contained in f-structure (4), but (9) is not.

(5.24)



The relation of f-command is particularly easy to state, given the formulation of inside-out functional uncertainty described above, since the notion of containment in an f-structure is the central notion in the statement of inside-out functional uncertainty. An expression like (GF^+ $\uparrow$) picks out all the f-structures that properly contain $\uparrow$, the notation 'GF^+' indicating a string of length one or greater. That is, the expression picks out the f-structures in which there is a path of length one or greater leading to $\uparrow$. These are the containing f-structures in the definition above. The f-structures that f-command $\uparrow$, then, are just those that appear at the end of a path of length one in those containing f-structures. The following expression picks out the f-structures that f-command the node represented by $\uparrow$:

(5.25) ((GF^+ $\uparrow$) GF)

The notion of f-command is a crucial notion in defining a set of possible binders for an anaphor. All positively-specified anaphors (anaphors which must be bound within a certain domain or to a certain binder) will be associated with an equation containing an expression of this general form; their antecedents must f-command them. Similarly, all negatively-specified anaphors will be associated with a negative constraining equation containing an expression of the same form; they will be required to be disjoint from f-commanding elements of a certain type.

Notice that the notion of f-command is already built into the statement of binding equations given in Chapter 4. The general form of a binding equation is

(5.26) ((DomainPath $\uparrow$) AntecedentPath)$_\sigma$=$_c$$\uparrow_\sigma$

As we noted, DomainPath is required to be nonempty. AntecedentPath consists of the grammatical function of the antecedent of the anaphor, a path of length one. To ensure that an anaphor is f-commanded by its antecedent, both negative and positive binding equations will be required to be of this general form.

5.6.2 The Noncontainment Condition

The definition of f-command given by Bresnan (1982a) has two clauses; requirement (2) is a condition on the form of binding equations, as we have just seen:

(5.27) For any occurrences of the functions α, β in an f-structure F, α *f-commands* β if and only if (1) α does not contain β and (2) every f-structure of F that contains α contains β.

According to requirement (1) of this definition, the f-command relation does not hold between an f-structure and one which contains it. I will refer to this as the Noncontainment Condition.

On the assumption that a relation of f-command holds between an anaphor and its antecedent, a condition forbidding f-structures that contain the anaphor from counting as f-commanding elements produces a result that is similar to the i-within-i condition in Government-Binding Theory (Chomsky 1981, p. 212):

(5.28) $*[_{\gamma} \ldots \delta \ldots]$

where δ and γ bear the same index

There are important differences between the i-within-i condition and the Noncontainment Condition as formulated here, however; in particular, the i-within-i condition is not a condition on possible antecedent-anaphor relations, but on the stronger relation of coreference. The i-within-i condition rules out coreference between a NP and a constituent it is contained in. Chomsky (1981) lists as ungrammatical several examples of violations of a condition of this kind; among them are:

(5.29) a. [the friends of their$_i$ parents]$_i$

b. [the owner of his$_i$ boat]$_i$

These examples seem to be truly uninterpretable. However, grammatical examples exist that seem to violate the i-within-i constraint:

(5.30) [a man with his$_i$ dog]$_i$

Additionally, Hellan (1988) cites examples such as the following, claiming that they are grammatical in Norwegian:

(5.31) a. [en beundrer av seg selv$_i$]$_i$ blir aldri helt ulykkelig
[an admirer of himself$_i$]$_i$ becomes never quite unhappy

b. jeg er på jekt etter [et bilde av seg selv$_i$]$_i$
I am hunting for [a picture of itself$_i$]$_i$

Hellan [p. 220-221] claims on the basis of these examples that the i-within-i condition is not a syntactic one, but a ban against referential circularity (Hellan also cites Williams (1982), Higginbotham (1983), and Haïk (1985) as making similar claims.). The argument structure of the nominal seems to play a large role in the determination of the acceptability of examples of this sort; examples in which the i-within-i violation comes about because of coreference between a possessor of an argument of a nominal and the nominal itself seem worst, as Chomsky's examples illustrate.[4] Examples involving reflexivization of the nominal argument seem strange, yet at least interpretable:

(5.32) ?? friends$_i$ of themselves$_i$

These differences suggest that an appeal to semantic argument structure rather than a syntactic condition condition such as the i-within-i condition may be a more promising approach. In this case, we would simply omit the Noncontainment Condition given below as a syntactic constraint on anaphoric binding.

The Noncontainment Condition may be stated informally in this way:[5]

[4]According to a test for the argument structure of nominals proposed by Fernando Pereira, the phrases 'of their parents' and 'of his boat' are arguments of the nominals 'friends' and 'owner'. Pereira observes that it is possible to negate an adjunct modifier of a nominal:

(a) the man not in a red jacket

However, negating an argument of a nominal is acceptable only on a contrastive reading:

(b) the population not of France

The following examples are acceptable only on a contrastive reading, showing that the prepositional phrases are indeed arguments of the nominals:

(c) the friends not of their parents

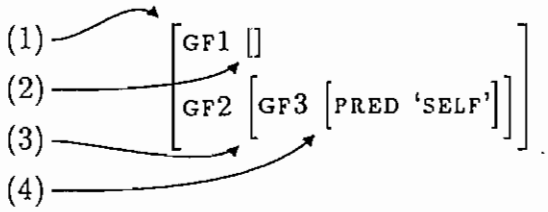
(d) the owner not of his boat

[5]A formal statement of the restriction may be found in the Appendix.

(5.33) The f-structure of the antecedent, ((DomainPath ↑) AntecedentPath), may not contain the anaphor.

To take a concrete example, consider this f-structure:

(5.34)



Assume that f-structure (4) is an anaphor subject to the Noncontainment Condition. Neither the outermost f-structure, labeled (1), nor the value of GF3, labeled (3), are eligible antecedents for (4), since they contain (4). Choosing the f-structure labeled (2) as an antecedent for (4) does not violate the Noncontainment Condition, though, since (2) does not contain (4).

According to this definition, the Noncontainment Condition holds only of anaphors that obey positive binding constraints — it applies only to reflexives and reciprocals and not to pronominals. This is because the constraint refers to the antecedent picked out by the positive requirement, so only anaphors associated with positive binding conditions will be subject to the Noncontainment Condition. It is less powerful than the i-within-i condition, since the i-within-i condition constrains possible coreference relations and applies to pronominals as well as reflexives and reciprocals; the i-within-i condition does not only constrain the relation between a reflexive or reciprocal and its antecedent, but constrains coreference between any two noun phrases.

The Noncontainment Condition as stated above will not rule out examples such as:

(5.35) a. [a man with his$_i$ dog]$_i$

b. *[friends of their$_i$ parents]$_i$

since *his* and *their* do not obey positive constraints. It will, however, rule out:

(5.36) *[a picture of itself$_i$]$_i$

since *itself* is associated with a positive binding constraint which picks out its antecedent.

5.6.3 Thematic Superiority

Pollard and Sag (1989) note contrasts such as the following:

(5.37) a. I sold the slave$_i$ himself$_i$.

b. *I sold himself$_i$ the slave$_i$.

(5.38) a. Mary explained Bill$_i$ to himself$_i$.

b. *Mary explained himself$_i$ to Bill$_i$.

(5.39) a. Mary talked to John$_i$ about himself$_i$.

b. *Mary talked about John$_i$ to himself$_i$.

Antecedency possibilities are not equal for arguments of a predicate; only certain arguments can antecede others.

Pollard and Sag (1989) take these examples as evidence that when the antecedent of an anaphor appears as an argument of the same predicate in which the anaphor appears, the antecedent must be higher on some hierarchy. In their theory, the relevant hierarchy is the obliqueness hierarchy; grammatical functions are defined in terms of an obliqueness ordering, as discussed in Chapter 3, Section 3.2.1.

Phenomena involving anaphors which obey negative binding constraints illustrate a similar point. For example, an anaphor such as English *he* is subject to the Coargument Disjointness Constraint: it may not corefer with a coargument. However, the following sentence is acceptable:

(5.40) He saw himself in the mirror.

Although *himself* and *he* are coarguments, this sentence is grammatical. The negative binding condition which *he* obeys is not violated, since a pronominal must be disjoint only from coarguments which are higher on some hierarchy.

Besides a structural notion of superiority such as f-command, then, a notion of superiority holding among arguments of the same predicate is also necessary. An anaphor associated with a positively-valued binding feature must find an antecedent that is superior to it in both senses; an anaphor associated with a negatively-valued feature must be disjoint from elements that are superior to it in both senses. Elements that are not superior to the anaphor in these senses are left out of consideration when binding constraints are applied.

Recent work by Bresnan (Bresnan and Moshi 1990, Bresnan and Kanerva 1989), Kiparsky (1989), and others has explored the thematic role structures of verbs. In both Bresnan's and Kiparsky's work, a hierarchy of thematic roles is posited; the 'logical subject' for each verb is taken to be the highest role in the hierarchy. Defining coargument superiority partly in terms of a thematic hierarchy seems a promising approach.

Jackendoff (1972) makes a proposal of this kind, citing data first presented by Postal (1971):

(5.41) a. I regard myself as pompous.

b. ?I strike myself as pompous.

(5.42) a. I smelled myself.

b. ?I smelled funny to myself.

The thematic structure of the sentences in (5.41) and (5.42) includes a Theme and a Goal; the Goal is higher on the thematic hierarchy than the Theme. The (a) sentences, where the Goal is the antecedent of the Theme, are grammatical; the (b) sentences, where the Theme is the antecedent of the Goal, are less good.[6]

Arguments against determining anaphoric antecedency conditions by reference to a thematic hierarchy often make reference to examples involving passivization, claiming that these examples constitute proof that a thematic hierarchy is not at work in determining possible antecedents for anaphors. The following example is ungrammatical:

[6]Wilkins (1988) also proposes that a thematic hierarchy is involved in reflexivization possibilities.

(5.43) *Himself was hit by John.

Assuming the agent role to be higher on a thematic hierarchy than the theme role, a purely thematic criterion of superiority would predict example (5.43) to be as grammatical as example (5.44):

(5.44) John hit himself.

An argument such as this shows only that the thematic hierarchy is not *by itself* sufficient to determine possible antecedents for reflexives. This argument does not preclude the possibility that a thematic hierarchy in combination with some set of syntactic criteria can determine the relevant notion of superiority.

Hellan (1988, Chapter 4) proposes that the relevant grammatical hierarchy at work in determining pronoun antecedency is both syntactic and thematic in nature. According to Hellan, the broad division that is relevant here is between direct and oblique functions: all direct functions (SUBJ, OBJ, and OBJ2) are superior to all oblique functions.

Within these two groups, antecedency is determined by the thematic hierarchy; arguments higher on the hierarchy are superior to lower arguments. Since the subject is in the general case higher on the thematic hierarchy than the object, the subject may antecede the object but not vice-versa.

Hellan provides a good deal of evidence in support for this analysis; evidence from other languages indicates that this is the right approach, and I will adopt this analysis here.[7]

The unacceptability of example (5.45)b is due to a violation of the thematic hierarchy constraint; the reflexive and its antecedent are both oblique arguments of the same predicate, but the reflexive is higher on the thematic hierarchy than its antecedent:

(5.45) a. Mary talked to John$_{goal}$ about himself$_{theme}$.

b. *Mary talked about John$_{theme}$ to himself$_{goal}$.

[7]See the Appendix for a formalization of this requirement.

Hellan (1988) also notes that the notion of superiority is not only relevant for coarguments but also for anaphors that are contained within coarguments. For Norwegian verbs subcategorizing for an OBJ and an OBJ2, two passives are possible:

(5.46) a. vi overlot Jon pengene
we gave Jon money
'We gave John the money.'

b. Jon ble overlatt pengene
Jon was given money
'Jon was given the money.'

c. pengene ble overlatt Jon
money was given Jon
'The money was given Jon.'

However, consider the following example (Hellan 1988, p. 160):

(5.47) barnet ble fratatt sine foreldre
child was taken self's parents
'The child$_i$ was deprived of self's$_i$ parents.'
*'The child$_i$ was taken away from self's$_i$ parents.'

Since either the OBJ or the OBJ2 of the verb 'deprive' can become the subject under passivization, the sentence would be expected to be ambiguous. The subject *barnet* could in principle be construed as either the malefactive/'underlying object' or as the theme/'underlying OBJ2' of *fratatt*. However, only one construal is possible: the one under which the malefactive argument is the SUBJ. Under this construal, the antecedent of the reflexive is the malefactive and is higher on the thematic hierarchy than the theme, the NP of which the reflexive is a possessor.

Hellan also shows that a thematic hierarchy is at work in determining the distribution of anaphora within nominals. As evidence, he presents example (5.48):

(5.48) Jons begravelse av sine naboer brakte tårer frem i manges øyne
John's burial of self's neighbors brought tears into many eyes
*by
'Jon's$_i$ burial of/*by self's$_i$ neighbors brought tears into many eyes.'

In general, the preposition *av* can mark an agent; here, though, the only possible interpretation is as indicated. For *Jon* to be an acceptable antecedent of the reflexive *sin*, *Jon* must outrank *sin naboer* on a thematic hierarchy. This is only possible if *av* is interpreted not as marking an agent but as marking a theme. Zaenen (1990) provides an analysis of these facts within the LFG framework; her analysis also makes use of a thematic hierarchy.

In sum, the kind of hierarchy that is relevant for the definition of superiority is both syntactic and semantic in nature. Direct arguments are superior to oblique arguments; within these two groupings, a thematic hierarchy determines superiority. Further, the relation of superiority is relevant not only for coarguments but for anaphors contained inside coarguments, as Hellan shows.

5.7 Anaphors in Subjective Position

Given the superiority constraints discussed in the previous section, anaphors in subjective positions are distinguished, since they appear in a position in which there is no superior element in the same nucleus. The antecedent of an anaphor in subjective position is necessarily nonlocal; if an anaphor in subjective position is to find an antecedent, that antecedent must appear in a higher domain. For this reason, subjective positions are privileged: if an anaphor can appear in only one position in a tensed subordinate clause, for example, that position is always subject position.

5.7.1 Subjective Position Superiority

As we have seen, there are some anaphors that can appear in subject position in subordinate clauses; for example, dialects of Marathi vary in this respect. In some dialects, the reflexive *swataah* can appear in subordinate tensed clause subject position; some (but not all) speakers judge the following sentences to be grammatical:

(5.49) %Jane laa waaṭṭe ki swataa saglyaat sundar aaho
Jane DAT thought that self most beautiful was
'Jane$_i$ thought that self$_i$ was the most beautiful.'

(5.50) %Jane mhaṇaali ki swataaci parikshaa sampli
Jane said that self-GEN test finished
'Jane$_i$ said that self's$_i$ test was over.'

Since different dialects of the same language can vary as to whether or not an anaphor can appear in these positions, the possibility for an anaphor to appear in subject position cannot always be tied to some intrinsic feature of the anaphor: its morphological form, for example. Ability to appear in subordinate subject position should be thought of as another way in which domain constraints can be specified.

Subordinate clause subject position is a privileged one for anaphora in general, and there are complex anaphoric phenomena that refer specifically to subordinate clause subject position. The phenomenon of *switch-reference* is defined by Haiman and Munro (1983) as 'an inflectional category of the verb, which indicates whether or not its subject is identical with the subject of some other verb.' In terms of the formal analysis presented here, switch-reference involves a binding equation associated with a verb indicating whether its subject is coreferent or disjoint in reference with the subject of another verb.

Obviation refers to a similar phenomenon, definable as 'the exclusion of certain possible controllers or antecedents for arguments' (Simpson and Bresnan 1983). Simpson and Bresnan analyze obviation in Warlpiri, showing that subjects of some Warlpiri subordinate clauses are obligatorily controlled by the subject, the object, or the oblique dative. The featural analysis they present is easily translatable into an equational analysis of the type presented here.

In these cases, it seems clear that syntactic binding constraints hold for anaphors that appear in subordinate subject position, and that claims such as those made by Pollard and Sag (1989) and Manzini (1983) that such positions are always 'exempt' from binding constraints seem to be incorrect.

5.7.2 The Defective Paradigm Explanation

Sentences such as the following are ungrammatical in English:

(5.51) *John$_i$ said that himself$_i$ did it.

It has been proposed that the explanation for this fact is the lack of a nominative case form of the reflexive; that is, the appeal is to a gap in the paradigm.

A proposal of this nature has also been made for Icelandic; there is no nominative form of the Icelandic reflexive *sig*, but where the subordinate clause subject is non-nominative, *sig* may appear (Maling 1984):

(5.52) a. hún sagði að sig vantaði peninga
she said that self-ACC lacked money
'She$_i$ said that self$_i$ wanted money.

b. hún sagði að sér þætti vænt um mig
she said that self-DAT was fond of me
'She$_i$ said that self$_i$ was fond of me.'

For Icelandic, appeal to a defective paradigm appears to be the correct approach. Anaphors can appear in subordinate clause subject position when the required morphological form of the anaphor exists. The generalization does not have to do with subject position but with the morphological paradigm of the anaphor.

Some evidence that this is the correct approach for English comes from the following examples, where a reflexive appears in subject position but is marked with non-nominative case[8]:

(5.53) a. John$_i$ wanted more than anything else for himself$_i$ to get the job.

b. What John$_i$ would prefer is for himself$_i$ to get the job.

However, a simple appeal to a gap in the paradigm is not a satisfying explanation. If a case form is missing from the morphological paradigm of a reflexive, it is invariably nominative case; there seem to be no languages in which, for example, there is a nominative but no accusative form of a reflexive. An account that appeals to an accidental gap in the paradigm fails to capture this generalization.

[8]These examples are taken from Pollard and Sag (1989). The grammaticality of examples of this form is not universally accepted; similar examples, discussed by Lebeaux (1983), were judged to have '??' marking.

A possible explanation for this generalization is the interrelatedness of nominative case and subject position (Kiparsky 1989). Since anaphors can vary as to whether or not they can appear in subject position, and subject position is often filled by noun phrases with nominative case, the privileged nature of subordinate subject position extends naturally to the privileged appearance of reflexives with nominative case. If a reflexive paradigm lacks nominative case, this is not purely an accidental gap, but reflects the fact that subject positions are treated specially by binding constraints.

5.7.3 Anaphors in Subjective Position: Binding Equations

Data presented in Chapter 3 show that some anaphors can appear in subordinate clause subject position; for example, we have seen that this is true of the Marathi reflexive *swataah* for some speakers. Huang (1982) also shows that this is true of the Chinese reflexive *ziji*, as discussed in Chapter 3, Section 3.1.4.1.

On the analysis presented by Huang (1982), the Chinese reflexive *ziji* is acceptable in subordinate clause subject position:

(5.54) (Huang's (47a), p. 331)

Zhangsan shuo ziji hui lai
Zhangsan say self will come
'Zhangsan$_i$ said that self$_i$ will come.'

In other positions in the subordinate clause, *ziji* is not permitted; *ziji* appears in non-subject subordinate clause position only if it is bound within the subordinate clause.

On Huang's analysis, *ziji* is bound in the minimal domain containing a subject; I will assume here that this is correct, and that *ziji* obeys the Minimal Complete Nucleus Binding Condition. Here is the binding equation which *ziji* obeys:

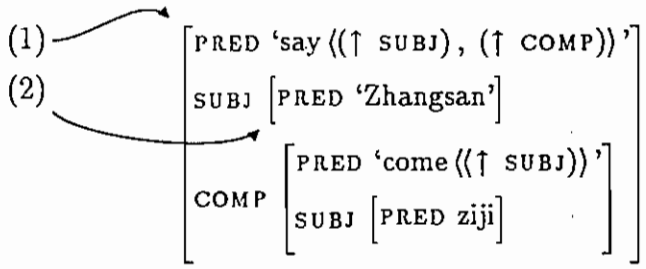
(5.55) $[((\text{DomainPath } (\text{SUBJ}) \uparrow) \text{ GF})_\sigma = \uparrow_\sigma]$

where DomainPath does not pass through an f-structure containing a subjective function.

This equation is similar to the one obeyed by other elements that are subject to the Minimal Complete Nucleus Binding Condition, such as English *himself*; the difference is the presence of the optional SUBJ at the end of the DomainPath. This equation allows *ziji* to appear in the minimal nucleus containing a subject other than itself; that is, if *ziji* is itself a subject, it may be bound in a wider domain — the higher domain containing a subject. All anaphoric elements that may appear in subordinate clause subject position only if they appear as subjects are associated with equations of this general form: the DomainPath contains an optional SUBJ attribute at its end.

Consider the following simplified f-structure for the sentence in example (5.54):

(5.56)



Given that *Zhangsan* is the antecedent of *ziji*, the instantiation of the binding equation given in (5.55) in this case is:

(5.57) $[((\text{COMP SUBJ} \uparrow)\ \text{SUBJ})_\sigma = \uparrow_\sigma]$

In this case:

- Domain Path is COMP and does not pass through a f-structure with a subjective function;
- (DomainPath SUBJ ↑), the outermost f-structure, is labeled (1);
- ((DomainPath SUBJ ↑) SUBJ), the antecedent of the reflexive, is labeled (2).

In sum, subordinate clause subject position is a privileged position for reflexives. This is as expected, since subjects are the most superior element in their local domain, and the antecedent of a reflexive appearing there is necessarily nonlocal. In terms of binding equations, binding constraints for anaphors which may appear in subordinate

clause subject position involve the appearance of an optional SUBJ at the end of the DomainPath.

5.8 Binding Asymmetries

In both Marathi and Norwegian, the applicability of disjointness conditions seems to depend in some cases on the form of other noun phrases in the same clause — whether or not the other noun phrases are pronouns, and if they are, what their surface realization is. In particular, we will see that there are some cases in which fully grammatical sentences involve binding configurations that would be incorrectly ruled out by the constraints that have been presented thus far. It seems that the conditions we have discussed in the foregoing apply only to certain elements in some cases, and that whether or not the conditions apply can depend on the morphological form of the element.

Because anaphoric binding conditions are stated in terms of binding requirements relating the f-structure representation of the anaphor with that of its antecedent, it would be an easy matter to provide for checks on morphosyntactic features of the antecedent, requiring different binding patterns for each type of antecedent. The form of such conditions is as yet unclear to me, however, and so I will not attempt to provide such refinements based on the small set of data presented here.

As we have seen, the Norwegian anaphor *seg* must be bound to a superior subject; it must also be disjoint from superior coarguments. The following sentence is ungrammatical because *seg* is coreferent with a subject coargument:

(5.58) *Jon snakket om seg
Jon talked about self
'Jon$_i$ talked about self$_i$.'

Seg is an example of an anaphor that is associated with a negative domain requirement (the Coargument Disjointness Condition) and a positive antecedent requirement (the Subject Binding Requirement).

The fact that *seg* is ungrammatical in the following example is, then, as expected; *seg* cannot appear in a position where it is coreferent with a coargument:

(5.59) Martin ba oss snakke til ham om ham selv
Martin asked us to talk to him about self
ham
him
*seg
self
*seg selv
self
'Martin$_i$ asked us to talk to him$_i$ about himself$_i$.'

The ungrammaticality of *seg* in this position is apparently due to the presence of *ham*; *ham* counts as a superior coargument in this case, and its presence disallows *seg*.

The fact that *seg selv* is ungrammatical here is also as expected. *Seg selv* requires a subject coargument as antecedent; here, the subject coargument is *oss* 'us', the understood subject of *snakke*, 'talk', and not the intended antecedent *Martin*.

The grammaticality of *ham selv* is also as expected. *Ham selv* requires a superior nonsubject antecedent in the same nucleus, and *ham* is such an element. Further, the Subject Disjointness Constraint obeyed by *ham selv* is not violated; *ham selv* is not coreferent with the immediately superior subject, in this case the understood subject of 'talk'.

However, the fact that *ham* is possible in example (5.59) is unexpected. *Ham* is required to be disjoint in reference from its coarguments, as illustrated by examples (5.60) and (5.61):

(5.60) ?*vi fortalte Jon om ham
we told Jon about him
'We told Jon$_i$ about him$_i$.'

(5.61) *vi gav Jon ham i julegave
we gave Jon him as Christmas present
'We gave Jon$_i$ him$_i$as a Christmas present.'

In example (5.59), though, it appears felicitously. Assuming that the first occurrence of *ham* is not a coargument of the second occurrence of *ham* would explain why *ham* is allowed here; however, in that case the ungrammaticality of *seg* in this position would

go unexplained, since the explanation for its ungrammaticality that was proposed above involved the claim that *seg* and *ham* were coarguments.

Puzzlingly, when the superior coargument is *seg* and not *ham*, grammaticality judgements are reversed:

(5.62) Martin ba oss snakke til seg om seg
Martin asked us to talk to himself about self
seg selv
self
*ham selv
himself
*ham
him
'Martin$_i$ asked us to talk to him$_i$ about himself$_i$.'

Here, *seg* and *seg selv* are unexpectedly grammatical, whereas they was disallowed in example (5.59). *Ham selv* and *ham* are allowed here, whereas they were disallowed in example (5.59).

The generalization that seems to hold in these cases is that an anaphor is sometimes allowed in positions where it would otherwise violate disjointness conditions; this happens where the argument that would bring about the violation is of a similar morphological form. Where the pronominal *ham* appears as a superior coargument, *ham* and *ham selv* are grammatical, and *seg* and *seg selv* are ungrammatical. Where the anaphor *seg* is a superior coargument, the reverse is true.

Similar facts hold in Marathi. Recall that both the Marathi pronominal and the reflexive *aapaṇ* obey the Coargument Disjointness Condition; they must be disjoint from coarguments. The ungrammaticality of the following sentence is, then, as expected:

(5.63) *John mhanaalaa ki Mary ni tyaalaa aaplyaa-baddal saangitle hote
John said that Mary ERG him-ACC about self told had
'John$_i$ said that Mary had told him$_i$ about self$_i$.

The presence of the pronominal *tyaalaa* violates the negative conditions that *aapaṇ* obeys, since *tyaalaa* is a coargument of *aapaṇ* and corefers with it. However, if a second

occurrence of *aapaṇ* is used in place of the pronoun, the sentence is unexpectedly grammatical:

(5.64) John mhanaalaa ki Mary ni aaplyaalaa aaplyaa-baddal saangitle hote
John said that Mary ERG self-ACC about self told had
'John$_i$ said that Mary had told self$_i$ about self$_i$.

Where coreference with *John* is intended in two positions in the subordinate clause, the reflexive must be used in both cases; a pronoun/reflexive combination cannot be used.

The ungrammaticality of example (5.63) does not seem to be due to the use of the two different anaphoric forms to refer to the same antecedent. In other cases, this is permitted, as in example (5.65):

(5.65) [[[Mary aaplyaa-baddal bolte] hi tyaalaa temwhaa maahiti hote] ase John
Mary about self talks that him-ACC then known was so John
malaa mhaaṇaalaa]
me-ACC said
'John$_i$ told me that he$_i$ knew that Mary had talked about self$_i$.'

It is only when the two anaphoric elements appear within the same clause that the constraint appears to hold.

Data such as these seem to indicate that the *form* of the antecedent as well as its referent play a role in anaphoric binding. Similar facts have also been observed by Koopman and Sportiche (1989) for the Kwa language Abe; they show that binding conditions for certain Abe pronouns depend on the form of their binder.

5.9 Conclusion

One of the primary advantages of the treatment of syntactic constraints on anaphoric binding that has been presented here is that it provides a framework for stating binding constraints in a precise manner. The framework makes explicit predictions as to which noun phrases will be considered when anaphoric binding constraints are applied. In this sense, anaphoric binding constraints are different from relations of

coreference. If two elements are coreferent, this may satisfy some positive constraint or violate some negative constraint; coreference between the two might also be irrelevant to any binding constraint. These cases can be easily distinguished in this framework.

Stating constraints in terms of binding equations makes it easy to express interactions between binding constraints; such interactions are even predicted in some cases. In the absence of such a framework, it is difficult even to frame the question of what interactions can occur; given an explicit statement of the binding constraints, such interactions are easy to observe.

Appendix: A Formalization of Binding Requirements

Following is a complete characterization of the syntactic binding constraints that were informally presented in Chapters 4 and 5.[9]

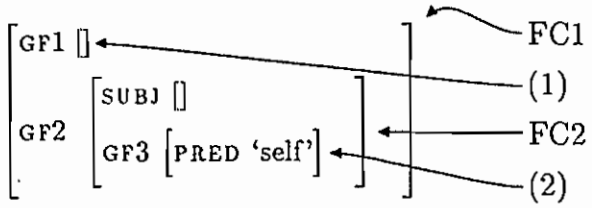
A.1 The Minimal Complete Nucleus Binding Condition

Anaphors obeying the Minimal Complete Nucleus Binding Condition must be bound in the minimal domain containing a subjective function, either SUBJ or POSS. To state the requirement in a slightly different way, if the minimal f-structure containing the anaphor and its antecedent is FC1, there must not be another f-structure FC2 that contains the anaphor and a subjective function but is properly contained within FC1.

Consider the f-structure in example (A.1). The f-structure labeled (1) is not a possible antecedent for the anaphor 'self', labeled (2), if 'self' obeys the Minimal Complete Nucleus Binding Condition:

[9]I would like to thank John Maxwell, Ron Kaplan, and Stuart Shieber for their assistance in formulating the constraints given in this section.

(A.1)



This is because the minimal complete nucleus containing the anaphor and a subjective function (here, SUBJ), labeled FC2, does not contain the antecedent of the anaphor. The outermost f-structure, labeled FC1, contains both the anaphor and its antecedent but is not the minimal complete nucleus containing the anaphor.

Here is the constraint on anaphors subject to the Minimal Complete Nucleus Binding Condition:

(A.2) Domain1 = $(\text{GF}^+ \uparrow)$

Domain2 = $(\text{GF}^+ \uparrow)$

$(\text{Domain1 GF})_\sigma =_c \uparrow_\sigma$

$\neg[(\text{Domain2 } \{\text{SUBJ} \mid \text{POSS}\}) \wedge$
$(\text{Domain1 GF}^+) =_c \text{Domain2}]$

In these equations, GF denotes a member of the set of primitive grammatical functions. Domain1 and Domain2 are variables ranging over f-structures, defined here as those f-structures through which the anaphor can be reached by following a path of length one or greater. In other words, Domain1 and Domain2 denote the f-structures that properly contain the anaphor.

Since Domain1 stands for an f-structure containing the anaphor, (Domain1 GF) picks out the antecedent of the anaphor — the value of GF for some grammatical function GF that f-commands the anaphor. In other words, this equation states that any f-commanding f-structure can represent the antecedent of the anaphor. The third line 'coindexes' the anaphor and its antecedent by requiring that the semantic representation of the anaphor and its antecedent (Domain1 GF) must be the same, in the manner discussed in Chapter 4. As it stands, though, this is too strong. Not all

f-commanding elements are proper antecedents for the anaphor, but only those lying within the minimal complete nucleus.

The fourth line states further constraints on possible antecedents for the anaphor. The negative constraint rules out the conjunction of the following two formulas:

(A.3) (Domain2 {SUBJ | POSS})

(Domain1 GF^+) $=_c$ Domain2

Recall that Domain2 is defined as ($GF^+ \uparrow$) — as any f-structure properly containing the anaphor. The expression (Domain2 {SUBJ|POSS}) is an existential statement; it picks out any f-structure containing the anaphor which has a subjective function SUBJ or POSS. Thus Domain2 is a complete nucleus containing the anaphor and a subjective function (though not necessarily the minimal one).

The second equation is satisfied if Domain2 is properly contained in Domain1, the minimal f-structure containing the anaphor and its antecedent. Domain2 is properly contained in Domain1 if by some non-null path GF^+ inside Domain1 it is possible to reach Domain2.

The conjunction of these two equations is constrained not to hold; there may not be a complete nucleus Domain2 properly contained in Domain1. This enforces the constraint that an anaphor which obeys this constraint finds its antecedent in the *minimal* complete nucleus.

A.2 The Minimal Finite Domain Condition

The definition of the Minimal Finite Domain condition is very similar to the definition of the Minimal Complete Nucleus condition:

(A.4) Domain1 = (GF^+ ↑)

Domain2 = (GF^+ ↑)

(Domain1 GF)$_\sigma$ =$_c$ ↑$_\sigma$

¬[(Domain2 TENSE) ∧
(Domain1 GF^+) =$_c$ Domain2]

As above, Domain1 is the f-structure containing the anaphor and its antecedent. The negative constraint rules out the possibility that there is a f-structure Domain2 that has a TENSE attribute and that is properly contained in Domain1, the binding domain for the anaphor.

A.3 The Coargument Binding/Disjointness Condition

Anaphors obeying the Coargument Binding Condition must be coreferent with an argument of the same predicate; anaphors obeying the Coargument Disjointness Condition must be disjoint in reference from an argument of the same predicate. The Coargument Binding and Disjointness Conditions pick out the domain containing the PRED of which the anaphor is an argument.

The Coherence Condition requires that a governable grammatical function be governed by the nearest PRED. For this reason, the PRED that subcategorizes for the anaphor will always be the most local one; an f-structure in which the anaphor appears as the argument of a non-local PRED would be incoherent and so would be ruled ill-formed.

The relevant binding domain for the Coargument Condition is, then, the minimal domain properly containing the anaphor and containing any PRED at all, since this PRED will always be the one that subcategorizes for the anaphor. In other words, the minimal f-structure properly containing the anaphor and containing a PRED will be the binding domain for anaphors subject to the Coargument Binding/Disjointness Condition.

The Coargument Binding Condition can be stated as:

(A.5) Domain1 = ($GF^+ \uparrow$)

Domain2 = ($GF^+ \uparrow$)

$(\text{Domain1 GF})_\sigma =_c \uparrow_\sigma$

$\neg[(\text{Domain2 PRED}) \wedge$
$(\text{Domain1 GF}^+) =_c \text{Domain2}]$

Domain1 and Domain2 are, as above, f-structures properly containing the anaphor. The negative constraint rules out the possibility that there is a f-structure Domain2 with a PRED that is properly contained in Domain1.

The Coargument Disjointness Condition is the negation of the above condition:

(A.6) $\neg[\text{Domain1} = (GF^+ \uparrow)$
$\text{Domain2} = (GF^+ \uparrow)$
$(\text{Domain1 GF})_\sigma =_c \uparrow_\sigma$
$\neg[(\text{Domain2 PRED}) \wedge$
$(\text{Domain1 GF}^+) =_c \text{Domain2}]]$

Domain1 and Domain2 denote f-structures containing the anaphor in the way described above; Domain1 is the domain in which the anaphor and its antecedent are both found. The constraint forbids a situation in which a pronoun obeying the Coargument Disjointness Condition is coreferent with an argument GF that is a coargument of the same PRED.

A.4 The Noncontainment Condition

The Noncontainment Condition prevents an anaphor from taking as its antecedent an f-structure in which it is contained. The following set of equations enforces this requirement:

(A.7) Domain = (GF^+ ↑)

AntecedentFunction = GF

(Domain AntecedentFunction)$_\sigma$ =$_c$ ↑$_\sigma$

¬[((Domain AntecedentFunction) GF^+) = ↑]

The first line of the equation defines the variable Domain as ranging over the set of f-structures containing the anaphor. The second line defines AntecedentFunction as ranging over the set of primitive grammatical functions. The expression (Domain AntecedentFunction) picks out the f-structures that f-command the anaphor — the set of its possible antecedents. The third line enforces semantic identity between the anaphor and its antecedent.

The fourth line of the equation enforces the noncontainment condition. It forbids a nonnull path GF^+ through the antecedent of the anaphor (Domain AntecedentFunction) by which the f-structure for the anaphor ↑ can be reached.

A.5 Thematic Superiority

The set of possible antecedents for an anaphor consists of the NP's that are *superior* to the anaphor. One relevant sense of superiority is thematic; here we define the notion of thematic superiority.

The first relevant division is between direct and oblique arguments. Direct arguments of a verb are always superior to oblique arguments:

(A.8) { SUBJ, OBJ, OBJ2 } > OBLIQUE

Within this broad division, a thematic hierarchy operates. The thematic hierarchy is:

(A.9) Agent > Goal > Instrument > Theme > Locative

An element A is thematically superior to a coargument B if

- A is a direct argument and B is an oblique argument; or
- A and B are both direct arguments and A is higher on the thematic hierarchy than B; or
- A and B are both oblique arguments and A is higher on the thematic hierarchy than B.

It was noted in Chapter 5, Section 5.6.3 that the thematic hierarchy is relevant not only for cases in which the anaphor and its antecedent are coarguments, but also in cases in which the anaphor is contained within a coargument of the antecedent. An example is:

(A.10) barnet ble fratatt sine foreldre
child was taken self's parents
'The child$_i$ was deprived of self's$_i$ parents.'
*'The child$_i$ was taken away from self's$_i$ parents.'

Hellan (1988) shows that the possibilities for construal of this sentence depend on the relation between the thematic roles assigned to the antecedent *barnet* 'the child' and the phrase *sine foreldre* 'self's parents', the noun phrase in which the reflexive *sin* appears as a possessor. In the following, I will assume that thematic superiority is relevant between the antecedent of an anaphor and the coargument in which the anaphor is contained.

The following set of equations enforces thematic superiority in the relevant range of cases:

(A.11) DIR1 = SUBJ ∨ OBJ ∨ OBJ2

DIR2 = SUBJ ∨ OBJ ∨ OBJ2

OBL1 = OBL_θ ∨ ADJ

OBL2 = OBL_θ ∨ ADJ

$((\text{DomainPath} \uparrow) \text{ AntecedentPath})_\sigma =_c \uparrow_\sigma$

$[[(\text{DomainPath} = \text{OBL1 GF}^*) \wedge (\text{AntecedentPath} = \text{DIR1})] \vee$
$[(\text{DomainPath} = \text{DIR1 GF}^*) \wedge (\text{AntecedentPath} = \text{DIR2}) \wedge \text{DIR2} >> \text{DIR1}] \vee$
$[(\text{DomainPath} = \text{OBL1 GF}^*) \wedge (\text{AntecedentPath} = \text{OBL2}) \wedge \text{OBL2} >> \text{OBL1}]]$

DIR1 and DIR2 are variables ranging over the grammatical functions SUBJ, OBJ, and OBJ2. OBL1 and OBL2 are variables ranging over the oblique functions and adjuncts. The antecedent of the anaphor is the element ((DomainPath ↑) AntecedentPath); the disjunction of equations is intended to properly constrain the DomainPath and AntecedentPath values.

The notation '>>' is intended to indicate thematic superiority of one argument relative to another; 'X >>Y' indicates that X is higher on the thematic hierarchy than Y. This notation should be thought of as standing for a relation between the thematic roles of two arguments; the level of f-structure is, of course, not the proper level at which to encode thematic relations, though the notation used might seem to indicate that thematic superiority must hold at f-structure. A more accurate notation to indicate the intended relation might be:

$\text{DIR2}_\theta >> \text{DIR1}_\theta$

indicating that the relation holds at the level of a thematic projection θ.

The first line of the disjunction enclosed in brackets:

(DomainPath = OBL1 GF*) ∧ (AntecedentPath = DIR1)

covers the case where the anaphor is an oblique argument and its antecedent is a

direct argument. Since all direct arguments are superior to all oblique arguments, the antecedent is necessarily superior to the anaphor in this case.

The second line of the disjunction:

$$(\text{DomainPath} = \textsc{dir1}\ \textsc{gf}^*) \wedge (\text{AntecedentPath} = \textsc{dir2}) \wedge \textsc{dir2} >> \textsc{dir1}$$

covers the case where the anaphor and its antecedent are both direct arguments. The grammatical function of the antecedent is required to be higher on the thematic hierarchy than the grammatical function of the anaphor, as indicated by the notation '>>'.

The third line of the disjunction:

$$(\text{DomainPath} = \textsc{obl1}\ \textsc{gf}^*) \wedge (\text{AntecedentPath} = \textsc{obl2}) \wedge \textsc{obl2} >> \textsc{obl1}$$

is the same as the second, except it is the case where the anaphor and its antecedent are both oblique. Again, the antecedent must be thematically superior to the anaphor.

References

Joan Bresnan and Jonni M. Kanerva. 1989. Locative inversion in Chichewa: a case study of factorization in grammar. *Linguistic Inquiry*, 20(1):1–50.

Joan Bresnan and Lioba Moshi. 1990. Object asymmetries in comparative Bantu syntax. *Linguistic Inquiry*, 21(2):147–186.

Joan Bresnan, Per-Kristian Halvorsen, and Joan Maling. 1985. Logophoricity and bound anaphors. MS, Stanford University.

Joan Bresnan. 1982a. Control and complementation. In Joan Bresnan, editor, *The Mental Representation of Grammatical Relations*, pages 282–390. MIT Press, Cambridge, Mass.

Joan Bresnan, editor. 1982b. *The Mental Representation of Grammatical Relations.* MIT Press, Cambridge, Mass.

Joan Bresnan. 1987. Class lectures, 1987 Linguistic Institute, Stanford University.

Noam Chomsky. 1973. Conditions on transformations. In Stephen Anderson and Paul Kiparsky, editors, *A Festschrift for Morris Halle.* Holt, Reinhart, and Winston, New York.

Noam Chomsky. 1981. *Lectures on Government and Binding.* Foris Publications, Dordrecht.

Noam Chomsky. 1986. *Knowledge of Language: Its Nature, Origin, and Use.* Praeger, New York.

George N. Clements. 1975. The logophoric pronoun in Ewe: Its role in discourse. *Journal of West African Languages*, 10(2):141–177.

Cleo Condoravdi. 1989. Class lectures, Stanford University.

Mary Dalrymple, Stuart M. Shieber, and Fernando C. N. Pereira. 1989. Ellipsis and higher-order matching. Paper presented at the 1989 Annual Meeting of the Linguistic Society of America.

Mary Dalrymple. 1984. Subject types in Nepali, Marathi, and Bengali. Master's thesis, University of Texas at Austin.

David Dowty. 1980. Comments on the paper by Bach and Partee. In Jody Kreiman and Almerindo E. Ojeda, editors, *Papers from the Parasession on Pronouns and Anaphora*, pages 29–40, University of Chicago. Chicago Linguistic Society.

David Dowty. 1982. Grammatical relations and Montague Grammar. In Pauline Jacobson and Geoffrey K. Pullum, editors, *The Nature of Syntactic Representation*, pages 79–130. Reidel, Dordrecht.

Mürvet Enç. 1989. Pronouns, licensing, and binding. *Natural Language and Linguistic Theory*, 7(1):51–92.

Leonard M. Faltz. 1985. *Reflexivization: A Study in Universal Syntax.* Garland, New York.

Jens Erik Fenstad, Per-Kristian Halvorsen, Tore Langholm, and Johan van Benthem. 1987. *Situations, Language and Logic.* Reidel, Dordrecht.

Mark Gawron and Stanley Peters. 1988. Anaphora and quantification in Situation Semantics. MS, Center for the Study of Language and Information, Stanford University.

Isabelle Haïk. 1985. *The Syntax of Operators.* Ph.D. thesis, MIT.

John Haiman and Pamela Munro. 1983. Introduction. In John Haiman and Pamela Munro, editors, *Switch Reference and Universal Grammar.* Benjamins, Amsterdam.

Per-Kristian Halvorsen and Ronald M. Kaplan. 1988. Projections and semantic description in Lexical-Functional Grammar. In *Proceedings of the International*

Conference on Fifth Generation Computer Systems, pages 1116–1122, Tokyo, Japan. Institute for New Generation Systems.

Per-Kristian Halvorsen. 1988. Situation Semantics and semantic interpretation in constraint-based grammars. Paper presented at the International Conference on Fifth-Generation Computer Systems.

Lars Hellan. 1988. *Anaphora in Norwegian and the Theory of Grammar.* Foris Publications, Dordrecht.

James Higginbotham. 1983. Logical form, binding, and nominals. *Linguistic Inquiry*, 14(3):395–420.

C.-T. James Huang. 1982. *Logical Relations in Chinese and the Theory of Grammar.* Ph.D. thesis, MIT.

C.-T. James Huang. 1983. A note on the binding theory. *Linguistic Inquiry*, 14(3):554–560.

Sabine Iatridou. 1986. An anaphor not bound in its governing category. *Linguistic Inquiry*, 17(4):766–772.

Masayo Iida. 1990. Untitled Stanford University Ph. D. thesis, in preparation.

Ray S. Jackendoff. 1972. *Semantic Interpretation in Generative Grammar.* MIT Press, Cambridge, Mass.

Smita Joshi. 1989. Logical subject in Marathi grammar and predicate argument structure. In *Proceedings of West Coast Conference on Formal Linguistics 8.*

Megumi Kameyama. 1985. *Zero Anaphora: The Case of Japanese.* Ph.D. thesis, Stanford University.

Megumi Kameyama. 1989. Functional precedence conditions on overt and zero pronominals. MS, Stanford University.

Ronald M. Kaplan and Joan Bresnan. 1982. Lexical-Functional Grammar: A formal system for grammatical representation. In Joan Bresnan, editor, *The Mental Representation of Grammatical Relations*, pages 173–281. MIT Press, Cambridge, Mass.

Ronald M. Kaplan and Annie Zaenen. 1989. Long-distance dependencies, constituent structure, and functional uncertainty. In Mark Baltin and Anthony Kroch, editors, *Alternative Conceptions of Phrase Structure*. Chicago University Press.

Ronald M. Kaplan, John Maxwell, and Annie Zaenen. 1987. Functional uncertainty. *CSLI Monthly Newsletter*. January.

Ronald M. Kaplan. 1987. Three seductions of computational psycholinguistics. In Peter Whitelock, Harold Somers, Paul Bennett, Rod Johnson, and Mary McGee Wood, editors, *Linguistic Theory and Computer Applications*, pages 149–181. Academic Press, London.

Lauri Karttunen. 1976. Discourse referents. In James D. McCawley, editor, *Syntax and Semantics 7: Notes from the Linguistic Underground*, pages 363–386. Academic Press, New York.

Edward L. Keenan and Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry*, 8(1):63–99.

Paul Kiparsky. 1989. Morphology and grammatical relations. MS, Center for the Study of Language and Information, Stanford University.

Hilda Koopman and Dominique Sportiche. 1989. Pronouns, logical variables, and logophoricity in Abe. *Linguistic Inquiry*, 20(4):555–588.

Howard Lasnik. 1976. Remarks on coreference. *Linguistic Analysis*, 2:1–22.

Howard Lasnik. 1989. On the necessity of binding conditions. In *Essays on Anaphora*. Kluwer, Dordrecht. To appear in R. Frieden (editor), *Principles and Parameters in Comparative Grammar*, MIT Press.

David Lebeaux. 1983. A distributional difference between reflexives and reciprocals. *Linguistic Inquiry*, 14(4):723–730.

Lori Levin, Malka Rappaport, and Annie Zaenen, editors. 1983. *Papers in Lexical-Functional Grammar*. Indiana University Linguistics Club, Bloomington.

Joan Maling. 1984. Non-clause-bounded reflexives in modern Icelandic. *Linguistics and Philosophy*, 7(3):211–241.

Maria Rita Manzini and Kenneth Wexler. 1987. Parameters, binding theory, and learnability. *Linguistic Inquiry*, 18(3):413–444.

Maria Rita Manzini. 1983. On control and control theory. *Linguistic Inquiry*, 14(3):421–447.

Rolf Noyer. 1989. What are the binding parameters? In search of the anti-pronoun. MS, MIT.

Fernando C. N. Pereira. 1989. A calculus for semantic composition and scoping. In *Proceedings of the 1989 Meeting*, pages 152–160. Association for Computational Linguistics.

David Perlmutter and Paul Postal. 1983. *Studies in Relational Grammar*, volume 1. University of Chicago Press, Chicago.

David Perlmutter and Carol Rosen. 1984. *Studies in Relational Grammar*, volume 2. University of Chicago Press, Chicago.

Carl Pollard and Ivan A. Sag. 1987. *Information-Based Syntax and Semantics*, volume 1. CSLI/University of Chicago Press, Stanford University. CSLI Lecture Notes, Number 13.

Carl Pollard and Ivan A. Sag. 1989. Anaphors in English and the scope of the binding theory. MS, Stanford University.

Carl Pollard and Ivan A. Sag. 1990. *Information-Based Syntax and Semantics: Topics in Binding and Control*, volume 2. CSLI/University of Chicago Press, Stanford University. To appear.

Paul M. Postal. 1971. *Crossover Phenomena.* Holt, Rinehart, and Winston, New York.

Douglas Pulleyblank. 1986. Clitics in Yoruba. In *Syntax and Semantics*, volume 19, pages 43–64. Academic Press, New York.

Tanya Reinhart. 1983. *Anaphora and Semantic Interpretation.* University of Chicago Press, Chicago.

Tanya Reinhart. 1986. Center and periphery in the grammar of anaphora. In Barbara Lust, editor, *Studies in the Acquisition of Anaphora*, volume 1. Reidel, Dordrecht.

Ken Safir. 1990. Implied non-coreference and the pattern of anaphora. MS, Rutgers University.

Peter Sells, Annie Zaenen, and Draga Zec. 1987. Reflexivization variation: Relations between syntax, semantics, and lexical structure. In Masayo Iida, Stephen Wechsler, and Draga Zec, editors, *Working Papers in Grammatical Theory and Discourse Structure*, pages 169–238. CSLI/University of Chicago Press, Stanford University. CSLI Lecture Notes, Number 11.

Peter Sells. 1985. *Lectures on Contemporary Syntactic Theories.* CSLI/University of Chicago Press, Stanford University. CSLI Lecture Notes, Number 3.

Halldór Sigurdsson. 1989. Moods and (long distance) reflexives in Icelandic. MS, University of Iceland.

Jane Simpson and Joan Bresnan. 1983. Control and obviation in Warlpiri. *Natural Language and Linguistic Theory*, 1(1):49–64.

Dominique Sportiche. 1986. Zibun. *Linguistic Inquiry*, 17(2):369–374.

Höskuldur Thráinsson. 1976. Reflexives and subjunctives in Icelandic. *Proceedings of NELS*, 6:225–239.

Sten Vikner. 1985. Parameters of binder and of binding category in Danish. In *Working Papers in Scandinavian Syntax 23*. University of Trondheim.

Kashi Wali. 1976. *Two Marathi Reflexives and their Implication for Causative Structure*. Ph.D. thesis, Syracuse University.

Kashi Wali. 1979. Two Marathi reflexives and the causative structure. *Studies in Language*, 3(3):405–438.

Michael Wescoat. 1989. Sloppy readings with embedded antecedents. MS, Stanford University.

Wendy Wilkins. 1988. Thematic structure and reflexivization. In Wendy Wilkins, editor, *Syntax and Semantics 21: Thematic Relations*, pages 191–214. Academic Press, San Diego.

Edwin Williams. 1982. The NP Cycle. *Linguistic Inquiry*, 13(2):277–296.

Dong-Whee Yang. 1983. The extended binding theory of anaphors. *Language Research*, 19(2):169–192.

Annie Zaenen. 1990. Modeling reflexivization in Norwegian nominal predicates. MS, Stanford University.

Anne Zribi-Herz. 1989. Anaphor binding and narrative point of view: English reflexive pronouns in sentence and discourse. *Language*, 65(4):695–727.