PROJECTION AND INFLECTION:
A STUDY OF PERSIAN PHRASE STRUCTURE

by

Jila Ghomeshi

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Graduate Department of Linguistics
University of Toronto

© Copyright by Jila Ghomeshi 1996
This thesis is based on the following pivotal assumptions about morphology and syntax: (a) that monosemy is desirable in grammar; (b) only inflectional and not derivational affixes can correspond to syntactic projections; (c) projection is driven by the features borne by lexical and functional categories; and (d) base-generated $X^0$-adjunction is possible in the syntax.

Starting with the distinction between inflectional and derivational affixes, the thesis is organized around the inflectional affixes of Persian and the range of constructions they occur in. First, in looking at the Ezafe vowel -e, it is argued that the distribution of this vowel is best accounted for if nouns do not project structure in Persian. This explains why no phrasal material occurs below the DP (Determiner Phrase). It is proposed that modifiers to the head noun occur in an $X^0$-adjointed structure and that the Ezafe vowel is inserted at PF to identify potentially non-projecting heads as belonging to a single constituent.

Turning to agreement in Persian, it is shown that the differences in the distribution of the pronominal enclitics and the agreement suffixes follow from morphological properties of these affixes. For example, the fact that the enclitics can only cooccur with null, definite, arguments while the subject agreement affixes can appear with any noun phrase is accounted for by subcategorization and case requirements of the affixes themselves. This analysis also explains the fact that the enclitics must occur outside subject agreement, even though they refer to the direct object. Further, the proposal that the pronominal enclitics identify pro in an argument position is shown to explain binding facts in Persian.

Among the other inflectional morphemes considered in the thesis is the marker -râ which is shown to case-mark presupposed DPs in VP-adjointed position. It is argued that these DPs are construed as VP-level topics and a licensing condition is given to account for their occurrence. In general each of the language-particular constructions in Persian are shown to
be accounted for by the morphological properties of the lexical items and inflectional morphemes involved.
Acknowledgements

I thought this would be the easiest and most pleasurable task involved in completing a dissertation. It is not. It's easy to think of people to thank, but it is hard to come up with the words to thank them. Difficulties aside, it gives me much pleasure to acknowledge many of the people who have had a positive influence on my life both academically and personally during the time that I have become a linguist.

First, I would like to thank the two women who have had the most influence on my work, and on my life, during my graduate years. Elizabeth Cowper supervised my MA thesis and Diane Massam my PhD thesis. They have both generously given of their time and resources as my work has developed. It is hard to determine when and where the supervisory role of one stopped and the other began. Liz kept pushing me to believe in my ideas. Diane challenged me to think differently. I thank them for their rigour, for their integrity, for their willingness to let their students pursue their own interests and ideas, and for the positive example they set as women who teach and pursue research and at the same time have families and non-linguistic interests.

My interest in linguistics started during my undergraduate years at York University. I would like to thank my first linguistics professors Sheila Embleton, Rob Fink, Ruth King, Ian Smith and Susan Ehrlich who all (in different ways) taught me much about linguistics and encouraged me to pursue graduate work.

I would like to thank my fellow guinea pig (and sole classmate) in the rebuilding of U of T's PhD program, Carrie Dyck. I learned so much by simply observing her over the years. Tom Wilson, in addition to being a friend, was a valuable source of advice during the (more difficult) early years. Mirco Ghini's razor-sharp mind and wonderful laugh and Päivi Koskinen's steady flow of e-mail stories are treasured aspects of my PhD years.

I am grateful for having had the opportunity to work on several SSHRC-funded syntax projects at the University of Toronto. The project meetings were an incredible forum for testing new ideas and for carrying out debates. My fellow students on these projects were as much my classmates as anyone else. I would like to thank Barb Brunson, Sarah Cummins, Elaine Gold, Regina Moorcroft, Zhang Ning and Hitay Yukseler for suffering through hours and hours of Persian data and for piquing my interest in many other phenomena as they pursued their own research.

I would like to thank Michael Rochemont for showing confidence in my abilities as a linguist at a most critical time. The struggle to try and convince him of my ideas has undoubtedly made my work stronger. He is also responsible for three of the best months of my linguistic life, the months I spent at UBC. There I benefitted greatly from talking with Strang Burton, Guy Carden, Henry Davis (thanks for the lecture on the Binding Theory), Hamida Demirdache and especially Rose-Marie Déchaine.
I would like to thank the other members of my thesis committee: Yves Roberge, who has enriched my view of morphology, Elan Dresher, who has enriched my view of linguistics, Anne-Marie Brousseau, and Maria Subtelny. In addition, it was a pleasure to finally meet Simin Karimi who served as the external examiner.

For many useful discussions and/or encouragement I would like to thank Ed Burstynsky, Andrew Carnie, Alana Johns, Keren Rice, Betsy Ritter, Carson Schütze, Ron Smyth, Michael Szamosi, and Arnold Zwicky. Additionally, though I had not met either during the writing of my dissertation I would like to acknowledge Simin Karimi and Vida Samiiian. They were a constant source of inspiration to me by virtue of being Persian women who have written valuable dissertations on the Persian language.

I gratefully acknowledge SSHRC doctoral fellowships 452-90-2534 and 753-91-0149, as well as the funding I received from being employed on SSHRC grants to Elizabeth Cowper, Diane Massam, and Michael Rochemont. The work carried out in completing this thesis was made possible by this support.

Finally for the personal side, I would like to thank my dear friend Mary Daniel for first telling me to take a linguistics course and for many discussions we have had about linguistics (and other things) since then. Thanks to my brother Jean for, among other things, making me laugh. Thanks to Laurent Guardo for support and love this past year. Last but not least, thanks to my parents Farhang and Azar Ghomeshi. They are ultimately responsible for providing me with the desire and the means to indulge my intellectual curiosity to my heart's content. This thesis is dedicated to them.
# Table of Contents

## Chapter 1  
*Introduction* ............................................................................................................................ 1

1.0 Introduction .............................................................................................................................. 1  
1.1 Word Order in Modern Persian .............................................................................................. 3  
1.2 Theoretical Assumptions ........................................................................................................... 5  
   1.2.1 Monosemy ........................................................................................................................ 5  
   1.2.2 Inflection vs. Derivation .................................................................................................. 6  
   1.2.3 Inflection and Syntax ....................................................................................................... 8  
   1.2.4 Feature-Driven Projection ............................................................................................... 9  
   1.2.5 Constraints on Syntactic Configuration .......................................................................... 11  
   1.2.6 Referentiality, (In)definiteness, and Specificity ............................................................... 14  
1.3 Inflection vs. Derivation in Persian ......................................................................................... 17  
   1.3.1 The Stress System in Persian and Stress-affecting Affixes ............................................... 17  
      1.3.1.1 Nouns and Adjectives ............................................................................................. 18  
      1.3.1.2 Verbs ....................................................................................................................... 19  
   1.3.2 Non Stress-affecting Affixes ............................................................................................ 21  
   1.3.3 Section Summary ............................................................................................................ 22  
1.4 Outline of the Thesis ................................................................................................................. 24  

## Chapter 2  
*The Ezafe Construction and Non-Projecting Nouns* ................................................................ 26

2.0 Introduction .............................................................................................................................. 26  
2.1 Descriptions of the Ezafe Construction .................................................................................... 29  
2.2 Samiian's Analysis of the Ezafe Construction .......................................................................... 35  
2.3 Properties of the Ezafe Domain Within the Noun Phrase ....................................................... 38  
   2.3.1 Attributive Noun Phrases ............................................................................................... 39  
   2.3.2 Adjective Phrases .......................................................................................................... 41  
   2.3.3 Prepositional Phrases .................................................................................................... 44  
   2.3.4 The Possessive Noun Phrase .......................................................................................... 47  
2.4 The Filter vs. X0-Adjunction Approach to Attributive Modifiers .......................................... 48  
2.5 Ezafe Across Categories ......................................................................................................... 52  
   2.5.1 Two Kinds of Prepositions ............................................................................................. 54  
   2.5.2 Is There a Link Between Ezafe and Case? .................................................................... 61  
   2.5.3 Section Summary ............................................................................................................ 63  
2.6 The Projection Properties of Lexical Categories in Persian .................................................. 63  
   2.6.1 The Projection Properties of Nominal vs. "true" Prepositions ....................................... 64  
   2.6.2 The Projection Properties of Adjectives ......................................................................... 69  
   2.6.3 The Function of Ezafe Insertion ..................................................................................... 74  
2.7 The Difference between Ezafe-linked Constituents and Compounds .................................... 74  
2.8 Conclusion and Further Implications ..................................................................................... 77  

## Chapter 3  
*Possessors and the Structure of DP* .......................................................................................... 83

3.0 Introduction .............................................................................................................................. 83  
3.1 Samiian's Analysis of DP\textsubscript{poss} under N' ............................................................ 85  
   3.1.1 An argument that DP\textsubscript{poss} is separate from N and its modifiers ..................... 85  
   3.1.2 An argument that DP\textsubscript{poss} is a constituent with N and its modifiers ............... 90
3.2 Further Properties of DPposs .................................................................92
  3.2.1 Only one Syntactic Position ............................................................92
  3.2.2 The Semantics of DPposs ...............................................................94
3.3 Evidence that DPposs involves projection to DP .................................95
  3.3.1 Semantic Evidence that DPposs forms a DP with the head noun ....95
  3.3.2 Distributional Evidence that DPposs forms a DP with the head noun ....97
3.4 Structural Position of DPposs .................................................................100
  3.4.1 DPposs in [SPEC, NP] .................................................................101
  3.4.2 DPposs in [SPEC, DP] .................................................................107
  3.4.3 Summary .....................................................................................110
3.5 What Projects a DP in Persian? ..............................................................111
  3.5.1 Proper names and Pronouns .........................................................112
  3.5.2 The Indefinite Enclitic .................................................................113
  3.5.3 The Null Definite Determiner .......................................................115
  3.5.4 The Pronominal Enclitics .............................................................116
  3.5.5 Plural Marking on Nouns .............................................................118
  3.5.6 The Colloquial Marker -e ............................................................125
3.6 A Reformulation of the Ezafe Insertion Rule .........................................129
3.7 Conclusion ...........................................................................................133

Appendix to Chapter 3 Pre-head Material in the Noun Phrase ...................133
I Demonstratives .....................................................................................134
II Numerals and Classifiers ....................................................................135
III The Indefinite Article ..........................................................................140
IV Quantifiers ........................................................................................143
V Superlatives and Ordinals ....................................................................145
VI Summary ..........................................................................................146

Chapter 4 Two Types of Agreement .......................................................148
4.0 Introduction .......................................................................................148
4.1 Subject Agreement ............................................................................150
4.2 The Pronominal Enclitics .................................................................154
  4.2.1 Where They Occur ......................................................................155
    4.2.1.1 As Possessors ......................................................................155
    4.2.1.2 As Objects of Prepositions .................................................156
    4.2.1.3 As Objects of Verbs ............................................................157
  4.2.2 Analysis .......................................................................................159
    4.2.2.1 Pronominal Enclitics as Arguments ....................................159
    4.2.2.2 The Pronominal Enclitics and AgrP ...................................160
    4.2.2.3 Pronominal Enclitics as LICensers ....................................165
4.3 A Comparison of Subject Agreement and the Pronominal Enclitics ....169
  4.3.1 Subcategorization and Constraints on Spell Out .........................169
  4.3.2 Affixes vs. Clitics .........................................................................175
4.4 The Binding of Pronominals and Anaphors .......................................177
  4.4.1 The Binding Properties of Pronouns vs. pro ..............................181
    4.4.1.1 Binding and pro Subjects ...............................................181
    4.4.1.2 Binding and Possessor pro ...............................................185
    4.4.1.3 Binding and Objects ........................................................188
  4.4.2 Reinhart and Reuland and the Binding Theory .........................190
4.4.2.1 R&R's Condition B and Possessors ........................................192
4.4.2.2 R&R's Condition A and the Persian Anaphor xod ............195
4.4.3 Summary and Questions for Further Research ..................202
4.5 Conclusion and Residual Problems ......................................204
4.5.1 The Pronominal Enclitics and Indirect Objects .................205

Chapter 5  *Râ, VP-Level Topics, and Complex Predicates* ..........208

5.0 Introduction .................................................................................208
5.1 The Syntactic Category of -râ ....................................................211
5.2 The Semantic Contribution of -râ ..............................................212
5.2.1 Definiteness and Specificity ....................................................212
5.2.2 -râ as a Presupposition (Hearer-Old) Marker .......................220
5.3 The Structural Position of DP+râ .............................................223
5.3.1 Karimi's Analysis of the Syntax of -râ .................................224
5.3.2 DP+râ as the Direct Object of a Verb ....................................226
5.3.3 DP+râ with Intransitive Verbs ..................................................231
5.3.4 DP+râ and VP-level Topicalization ......................................234
5.3.5 Section Summary .................................................................245
5.4 Clause-Level Topicalization .....................................................246
5.5 The Position of Subjects and Predication Phrase .................251
5.6 Transitive Compound Verbs ......................................................254
5.6.1 On the Phrasality of the Non-Verbal Element ......................258
5.6.1.1 Complements to the Non-Verbal Element .........................258
5.6.1.2 A Unified Treatment of Reordering Operations .............262
5.6.2 VP-internal Predication Phrase ..............................................266
5.6.3 Theta Assignment ...............................................................272
5.8 The Impersonal Construction ..................................................276
5.9 Conclusion .................................................................................285

Chapter 6  *Summary of Thesis* ......................................................287

References .....................................................................................292
Glossary

Abbreviations used in examples

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>EZ</td>
<td>the Ezaf e vowel</td>
</tr>
<tr>
<td>DP&lt;sub&gt;poss&lt;/sub&gt;</td>
<td>the possessive noun phrase</td>
</tr>
<tr>
<td>1</td>
<td>first person</td>
</tr>
<tr>
<td>2</td>
<td>second person</td>
</tr>
<tr>
<td>3</td>
<td>third person</td>
</tr>
<tr>
<td>sg</td>
<td>singular</td>
</tr>
<tr>
<td>pl</td>
<td>plural</td>
</tr>
<tr>
<td>comp</td>
<td>comparative</td>
</tr>
<tr>
<td>sup</td>
<td>superlative</td>
</tr>
<tr>
<td>ord</td>
<td>ordinal</td>
</tr>
<tr>
<td>part</td>
<td>participle</td>
</tr>
<tr>
<td>subj</td>
<td>subjunctive</td>
</tr>
<tr>
<td>imp</td>
<td>imperative</td>
</tr>
<tr>
<td>def</td>
<td>definite</td>
</tr>
<tr>
<td>indef</td>
<td>indefinite enclitic</td>
</tr>
<tr>
<td>cont</td>
<td>continuous prefix</td>
</tr>
<tr>
<td>neg</td>
<td>negative prefix</td>
</tr>
</tbody>
</table>

• The subject agreement affixes are distinguished from the pronominal enclitics in the glosses by an 'S' following the person and number specification.

E.g.  
âmad-i 'You came.'  
came+2sgS

ketâb-at 'your book'  
book+2sg

• As is discussed in Chapter 5, -râ is a marker appearing on definite/presupposed direct objects and VP-level topics. Thus it is inaccurate to gloss it either as a definiteness marker (since it does not appear on definite subjects) or as an accusative case marker (since there can be more than one noun phrase marked with -râ in a given sentence). In all examples it will simply be glossed as -râ.

• Where examples are cited from other people's work the format is the following:

   [NAME DATE:PAGE #.EXAMPLE#]

Thus, if example (38) taken from Samiian (1983) found on page 39 of that work is cited, it appears as:

(x) (a) EXAMPLE
Gloss  
Translation  
[Samiian 1983:39.38]
1.0 Introduction

This thesis examines the phrase structure of Modern Persian. The general goal is to account for the structure of simple clauses in Persian using a small number of principles that are as simple and general as possible. The discussion of phrase structure is centred around the inflectional affixes of Persian under the assumption that inflectional morphology plays a role in syntax. We will see that many of the syntactic constructions that appear to be unique to Persian can be accounted for by the lexical properties of the inflectional morphemes or features involved.

Persian is one of the Iranian languages which form the Indo-Iranian branch of the Indo-European family. This group consists of two subgroups: Indic and Iranian. The Indic subgroup includes Sanskrit and many languages widely spoken in India and Pakistan today. According to Lehmann (1962), Old Iranian, which existed up to 300 B.C., was divided into two dialects: Avestan and Old Persian. Middle Iranian dates from 300 B.C. to A.D. 900, and one of its dialects, Pahlavi or Middle Persian, was the language of the Persian Empire from approximately A.D. 300 to 900. Windfuhr (1987) states that Classical Persian begins around the thirteenth century and continues to the beginning of the nineteenth century when contemporary standard Persian emerges. According to Windfuhr, the three major dialects of Modern Persian are Farsi, spoken in Iran by over 30 million people; Dari, spoken in Afghanistan by 5 million people; and Tadzhik, spoken in the Tadzhik republic of the former Soviet Union by 2.2 million speakers. Other Iranian languages still in use are Balochi spoken in West Pakistan; Pashtu or Afgan, the language of Afghanistan; Kurdish, spoken in western Iran and Turkey; and Ossetic spoken in the northern Caucasus.
The dialect of Persian under discussion here is spoken in and around Tehran and is usually referred to as the Tehrani or Standard dialect. It is spoken by educated speakers, and used in the mass media. It is this dialect that will be referred to by the terms 'Persian' or 'Modern Persian' hereafter. Within this dialect there is a difference between the speech used in formal situations (when conferring respect on an addressee for example, or on the television and radio) and that used colloquially among close friends and family. The former is the most studied and written about in the literature on Persian, however, the colloquial (spoken) variety is considered in this work.

Using the colloquial variety as the object of study raises several issues which, unfortunately, take us far beyond the scope of this dissertation. For example, there is much speaker-to-speaker variation in the spoken language and given that many people who reside in and around Tehran actually come from other regions of Iran, it is unclear how to distinguish between individual variation and dialectal variation. There is also variation among speakers depending on how familiar they are with the formal variety of the language. This results from the fact that many constructions that can be used colloquially are judged to be (prescriptively) unacceptable within the formal variety.

Another problem in dealing with colloquial speech is that it is harder to draw a line between truly ungrammatical utterances and utterances that are ill-formed on their own but are allowable given a sufficiently rich context. Further, it is not fully clear how we are to account for phenomena of the latter type within syntactic and semantic theories. Having to constrain the scope of this dissertation, I will proceed assuming that there is a standard Persian dialect and that there is a clear distinction between grammatical and ungrammatical

---

1The dissertations on Persian written within the generative framework in the last twenty years or so are an exception to this claim. The majority of them deal with the colloquial rather than the formal dialect.

2It should be noted that some examples from the formal dialect also appear. Where relevant this is pointed out in the text.
utterances even though some of the judgements made on the basis of these assumptions may be somewhat arbitrary.

As my source of data, I will be drawing on previous work on Persian, and on data from native speakers. I will also make reference to the grammar of Modern Persian by Lazard (1957), though I will be using the English translation published in 1992. This is probably the most thorough work and certainly is the grammar most cited by people who have worked on Persian.

The transliteration system used throughout the thesis is that of Lazard. The only differences are that sh, zh, and ch are used for [ʃ], [ʒ], and [ç] respectively. Where data from other sources is used the transliteration and glosses are modified to be consistent with the system being used here.

1.1 Word Order in Modern Persian

From the point of view of language typology Persian presents a very interesting case. If we start with Greenberg's (1963) work on language universals we see that Persian flouts one of the most basic tendencies, namely that the order of a verb and its complement and that of an adposition and its complement are the same. Thus, Persian as a Type III (i.e. SOV) language is predicted to have postpositions (Universal 4). Greenberg notes that this is "overwhelmingly" true for SOV languages but in a footnote mentions four exceptions of which Persian is one (the others are Iraqw, a southern Cushitic language; Khamti, a Thai language; and Amharic).

Greenberg also states that the position of the genitive correlates highly with whether a language has prepositions or postpositions. Thus his Universal 2 states that in languages with prepositions, the genitive almost always follows the governing noun. This universal is borne out in Persian, thus the locus of the problem Persian presents for the universals is in the position of the verb, and not the preposition, with respect to its complement. Comrie
(1981), who also makes the observation that Persian is exceptional with respect to language typology, points out that Persian is an OV language that in almost all other respects has the properties of VO languages.

More recent work (Dryer (1988, 1992)) continues to suggest that Persian goes against strong tendencies in language typology. For example, Dryer (1988) points out, with reference to a broader sample of languages than that of Greenberg, that there is no evidence for any universal relationship between the order of an adjective and noun and the order of an object and verb. He notes, however, that adjectives modifying nouns are often single words (cf. Chapter 2) and suggests that the position of branching categories such as genitives and relative clauses may be more important. He proposes the Branching Direction Theory, which is further refined in his later work (Dryer, 1992).

Branching Direction Theory

Verb patterners are nonphrasal categories or phrasal categories that are not fully recursive, and object patterners are fully recursive phrasal categories in the major constituent tree. That is, a pair of elements X and Y will employ the order XY significantly more often among VO language than among OV languages if and only if X is not a fully recursive phrasal category in the major constituent tree and Y is a fully recursive phrasal category in the major constituent tree. (Dryer, 1992:114)

Informally, this theory says that word order correlations reflect a tendency for phrasal categories to precede nonphrasal categories in OV languages and to follow them in VO languages. As Dryer puts it "languages tend towards one of two ideals: right-branching languages, in which phrasal categories FOLLOW nonphrasal categories, and left-branching

---

3This is the version of the Binding Direction Theory that Dryer settles on after several revisions, however, it should be noted that he gives a slightly different alternative version and remains agnostic as to which one is ultimately the better version. For our purposes, the subtleties are irrelevant.
languages, in which phrasal categories PRECEDE nonphrasal categories" (p. 109)\(^4\) Thus there will be a tendency in OV languages for genitives and relative clauses to precede the noun. Again, this is not true in Persian, where such constituents always follow the noun.

It is tempting then, given the above discussion, to posit an underlying SVO order for Persian so that it will then fit nicely into the predicted patterns. Unfortunately, there is no evidence whatsoever for this underlying order. On the contrary, positive evidence can be found for SOV order (Karimi, 1989, 1994).

While ultimately it would be desirable to explain the unusual constituent order of Persian on the basis of some other property of the language, for the purpose of the present work it remains a stipulation.

1.2 Theoretical Assumptions

For the most part I adopt the theory of Government and Binding (GB) (Chomsky (1981, 1986)). More recent theoretical innovations will be mentioned as they become relevant. However, a number of specific assumptions about morphology and syntax are crucial to the analysis developed in this thesis. These assumptions are outlined below.

1.2.1 Monosemy

The first assumption that I make has been recently pursued in work by Cowper (1989, 1995), and Johns (1992)\(^5\) (cf. also Nida (1948)):

\(^4\)Note that Dryer's Branching Direction Theory is specifically formulated to rule out predictions about elements such as adjectives, demonstratives and intensifiers, for example, which he says are nonphrasal elements that combine with other nonphrasal elements. This then accounts for the fact that these categories do not exhibit any correlations with OV vs. VO order.

\(^5\)Cowper (1995) adopts Johns' principle but also gives a stronger version as follows:
**One Form/One Meaning Principle**

Where morphemes are identical or similar in phonological properties, in the unmarked case, they are identical or similar in all lexical properties.

(Johns, 1992)

Under this assumption, if two morphemes, or sets of morphemes in the case of inflectional paradigms, are identical in form, it follows that they also have the same morphological, syntactic and semantic properties. This is not to say that homophony never exists but that it is marked.

1.2.2 Inflection vs. Derivation

I assume that inflectional and derivational morphology take place at different levels of the grammar. Derivational morphology is lexical, meaning that the internal structure of words created by derivational processes is opaque in the syntax. Inflectional morphology, on the other hand, is transparent to the syntax. I assume that there are at least two kinds of inflectional morphemes. Following Cummins and Roberge (1993, 1994) I assume that there is a pre-syntactic, but post-lexical interface level at which items drawn from the lexicon can acquire inflectional features. Cummins and Roberge refer to this interface level as the LSI

---

**Strong Monosemey Principle**

The conceptual structure of a lexical entry may contain no disjunctions and no optional elements. If the conceptual structures of two uses of a lexical item cannot be unified via underspecification, then they must be treated as distinct lexical entries. For our purposes Johns' principle will suffice.

How monosemy applies to paradigms is an interesting issue as pointed out by Yves Roberge (personal communication). For example, if two forms within two different paradigms are identical phonologically should they be treated as the same? Similarly, if one member of a paradigm has a different phonological form in two different contexts, is this sufficient to posit two different paradigms? Both of these situations arise in Persian. I take the view here that monosemy holds of paradigms as a whole, not within them. However, this area merits further research.
(Lexicon-Syntax Interface). These features then have to be checked in an appropriate syntactic configuration. They are phonetically realized at PF in the spell-out component.

In addition to inflectional morphemes that are the spell-out of inflectional features attached at the LSI, I also assume that there are inflectional morphemes which can head functional projections. I will show that, in Persian, the only morphemes heading functional projections are phrasal affixes (cf. Zwicky (1987), Yoon (1994)). Phrasal affixes combine not with words but with phrases, attaching to the periphery of the constituent for which they are subcategorized. Again I assume that these affixes come to be attached to the edge of the appropriate phrase at PF in the spell-out component.7

The view just outlined entails, of course, that a distinction can be made between inflectional and derivational morphemes. This distinction cannot be made on the basis of meaning alone. While there is a finite set of inflectional categories, not all of these categories are instantiated morphologically in every language.8 Thus the task of determining the inflectional morphemes in a given language is an empirical one. The criteria that I consider to distinguish inflection from derivation are standard ones such as stress assignment, organization into paradigms, inner vs. outer position with respect to a stem, etc.

Taking this view in its strongest form then, it is possible that morphemes marking a traditionally inflectional category (such as tense) may not be distinguishable in any empirical way from the derivational morphology. In such a case I will assume that the morphemes are derivational, despite their meaning.

7In the literature we find a third type of inflectional morpheme, namely one that heads a functional projection but attaches to heads. Usually this attachment is thought to take place via head movement (cf. Baker (1988)). Implicit in this view is the idea that the actual morpheme is present in the syntax. It is not clear whether this possibility is necessary, given the existence of the LSI.

8Note that to say that an inflectional morpheme is not instantiated morphologically in a given language is not to say that the language lacks that category. For example, a language may not have a morphological distinction among nouns between singular and plural, however, this does not mean that the language does not have a number distinction. Furthermore, as mentioned immediately below, I assume that what is normally considered an inflectional category may be expressed via derivational morphology.
1.2.3 Inflection and Syntax

Recently, in addition to the projections of the lexical categories verbs (V), nouns (N), adjectives (Adj), and prepositions (P), syntacticians have proposed that categories such as Tense, Agreement, Negation, Determiner and Number may also project. In a sense, projection of non-lexical categories has been present all along. For example, the phrase structure grammars of the 1960s and 1970s always included the category S, which corresponded to no lexical head, unlike all the other categories (leaving aside S', later CP). Non-lexical projection became more explicit when S was reanalyzed as IP (Inflection Phrase) in Chomsky (1986), from which followed the splitting of IP (Pollock (1989)), and a parallel development in the structure of the noun phrase. Thus we now find structures such as (1) and (2). (1) is taken from Chomsky (1993) and (2) is based on work by Abney (1987) and Ritter (1991) among others.

(1) AGRSP
   /   \     /   \
  AGRS'  AGRS TP

(2) DP
   /   \     /   \
  D'    NumP D    Num
     /   \     /   \
    Num' Num NP
       /   \    /   \
      N'    N VP
         /   \  /   \
        N  V' V

T' AGROP
     /   \    /   \
    T    AGRO' VP
       /   \  /   \
      ARO   NP
One question raised by structures such as these is: what constrains the number of functional (i.e, non-lexical) projections? One obvious constraint is that the functional projections correspond to inflectional categories. Thus there is no 'creating abstract nouns from adjectives and nouns' phrase, even though the language may have a derivational morpheme which does just this. But even supposing that functional projections are limited to inflectional categories, there are at least two possibilities. It might be that every inflectional category corresponds to a functional projection, regardless of which categories are instantiated in a given language. That is, the inventory of functional projections is determined by Universal Grammar and does not vary from language to language. Or, perhaps the set of functional projections varies from language to language, and depends on the inflectional morphology of the language.

This second view is explored here. I assume that only those inflectional categories instantiated by inflectional morphology are syntactically realized. Thus, if an inflectional category is instantiated by means of a derivational process it will not be syntactically represented. I do not assume, however, that all inflectional morphemes in a language must correspond to functional projections. Thus, as we will see, a subset of the inflectional morphemes in Persian, specifically the two types of agreement morphemes, do not head functional projections. Further, I do not assume that the functional projections in a language are headed only by inflectional categories instantiated in the language. That is, there may be a functional projection that does not correspond to a lexical or inflectional morpheme. However, in such a case there must be empirical evidence for the projection.

1.2.4 Feature-Driven Projection

In addition to inflectional features such as tense and agreement that may correspond to functional projections in the syntax, lexical categories are also thought to be composed of
features. Thus, for example Chomsky (1970)\textsuperscript{9} classified lexical categories on the basis of the features \([+/\text{-}N] \text{ and } [+/\text{-}V]\), while Jackendoff (1977) used \([+/\text{-}\text{subject}] \text{ and } [+/\text{-}\text{object}]\).\textsuperscript{10}

These features are thought to compose the terminal\textsuperscript{11} nodes in syntax, and in recent work by Halle and Marantz (1993) on Distributed Morphology it is suggested that only features are present in the syntax and that Vocabulary Insertion takes place after S-Structure.

\textsuperscript{9}There seems to be some confusion in the literature as to where Chomsky actually proposed the features \([+/\text{-}N, +/\text{-}V]\). His 1970 work, "Remarks on Nominalization," is usually cited in this regard. However, while there is much discussion of syntactic features, he does not actually propose that there are features that can cross-classify the lexical categories in this work (see the following quotation).

In fact, there is, to my knowledge, no convincing argument for a category including just verbs and adjectives (or, to take another traditional view, nouns and adjectives), although it is not excluded that some such subdivision may be correct. It is quite possible that the categories noun, verb, adjective are the reflection of a deeper feature structure, each being a combination of features of a more abstract sort. In this way, the various relations among these categories might be expressible. For the moment, however, this is hardly clear enough even to be a speculation. [Chomsky, 1970, p. 199]

Radford (1981) cites Chomsky (1972) which contains another version of "Remarks on Nominalization". Specifically Radford refers the reader to pp. 48-61 although even here the actual feature combinations are not given. Dechaine (1993) cites Chomsky's 1974 Amherst Lectures, however, unfortunately these are not readily available. Having noted the confusion regarding where the proposal is made, I will follow convention and cite Chomsky (1970).

\textsuperscript{10}For an alternate view on features borne by lexical categories see Dechaine (1993).

\textsuperscript{11}I am using 'terminal' node to mean a node that does not dominate any further nodes. Since I assume that every phrase must have a head, terminal nodes will always be \(X^0\)s. However, not every terminal node will be a head. The head is the term that gives its label to a constituent (cf. Carnie (1995)). For example, consider the diagram below:

\[
\begin{array}{c}
XP \\
\downarrow \\
ZP \\
\downarrow \\
Z^0 \\
\downarrow \\
X^0 \\
\downarrow \\
X_1^0 \\
\downarrow \\
Y^0
\end{array}
\]

In this diagram the terminal nodes are \(X_1^0, Y^0\) and \(Z^0\). \(X_1^0\) and \(Z^0\) are also heads. I will also (perhaps confusingly) use "head" to refer to the vocabulary item that is inserted under an \(X^0\) in the syntax.
Whether or not vocabulary items are present under terminal nodes in the syntax, I assume that it is the terminal nodes that project structure. That is, the node N, composed of the features [+N, -V] projects to NP. Two kinds of features constitute a terminal node in the syntax: features that are instantiated by inflectional morphology and features that determine the lexical categories in a language.\textsuperscript{12} Of course, it is an empirical question whether such features form terminal nodes or whether all terminal nodes project structure. This is discussed further in the next section, and in Chapter 2 we will see that the feature [+N] does not project in Persian.

1.2.5 Constraints on Syntactic Configuration

Under classic GB theory, lexical categories projected to maximal categories, each with a complement and a specifier position. The template for the structure of a phrase is shown in (3).

\textbf{(3)}
\begin{center}
\begin{tikzpicture}
  \node (X) {X'};
  \node (XP) at (0,0) {XP};
  \node (Specifier) at (-1,-1) {	extit{Specifier}};
  \node (X) at (0,-1) {X};
  \node (Complement) at (1,-1) {	extit{Complement}};
  \draw (XP) -- (Specifier);
  \draw (XP) -- (X);
  \draw (X) -- (Complement);
\end{tikzpicture}
\end{center}

The universality of this tripartite structure was questioned in early work on functional categories (Fukui \& Speas (1986), Fukui (1986) and Speas (1986)). More recently the issue of projection has been thoroughly reexamined (Speas (1990), Chomsky (1994)) and the idea

\textsuperscript{12}Note that if a language lacks a certain part of speech (e.g. adjectives) the number of features borne by the lexical categories in that language may differ in number and content. That is, I am assuming the features [+/-N] and [+/-V] for a language that has four distinct lexical categories. Note also that it is possible that a language may have more than four lexical categories.
that a head always projects to a maximal category complete with a specifier and a complement position has been called into question.

In this context I will propose modifications to the classic view that are essentially compatible in spirit with Speas (1990) and Chomsky (1994). I will propose that projection in the syntax is not always obligatory, and that it is possible for heads to adjoin to one another. Further I will show that the phrasality of a constituent in the syntax depends not only on whether a head projects, but on whether that head itself is selected by a projecting element. First, I will outline the assumptions I make about possible and impossible syntactic configurations.

Throughout the history of X-bar theory, at least two types of syntactic combination have been available. Ignoring specifiers for the moment, the theory has allowed head-complement structures such as the one shown in (4a), as well as XP-adjunction structures like the one shown in (4b). I make the further assumption that base-generated X0-adjunction structures such as that shown in (4c) are also possible.

Possible Structures (order irrelevant)

\[
(4) \quad \begin{array}{ccc}
(a) & XP & (b) & XP & (c) & X^0 \\
 & X^0 & YP & XP & YP & Y^0
\end{array}
\]

X0-adjunction has always been possible as a derived structure, i.e. as the result of head-movement. Cowper (1992) argues that base-generated X0-adjunction should therefore be possible given structure preservation. X0-adjointed structures are also argued for in Fabb (1984, in an earlier version of GB), Gold (1994) and Ghomeshi and Massam (1994).

There are several other logically possible combinations for phrases and heads that I assume are not possible syntactic structures. These are shown in (5) below. In (5a-c) we see structures in which a phrasal category is dominated by a head. I assume that this is not a possible syntactic configuration, whether the phrase is in complement position, head
position, or both. In (5d) we see that two adjoined heads cannot be dominated by a phrasal node. That is, if two heads are adjoined to one another they must be dominated by another head. In (5e) we see the reverse situation where a head has been adjoined to a phrase. Again, I assume that this is not a possible syntactic configuration.

Impossible Structures (order irrelevant)

(5) (a) * X\(^0\) YP
     \[ X\(^0\) \quad Y\(^0\) \]

(b) * X\(^0\) Y\(^0\)
     \[ X\(^0\) \quad Y\(^0\) \]

(c) * X\(^0\) YP
     \[ X\(^0\) \quad Y\(^0\) \]

(d) * XP
     \[ X\(^0\) \quad Y\(^0\) \]

(e) * XP
     \[ X\(^0\) \quad Y\(^0\) \]

These assumptions, taken together, have certain consequences, one of which will play a role in this work. Let us take a concrete example, and suppose that nouns and adjectives do not project to maximal categories. A noun with a modifying adjective will therefore not be dominated by an NP node in the syntax but rather by N\(^0\) (on the assumption that there is some mechanism that determines headedness). This is shown in (6a). But now suppose that a noun is the complement of a determiner, and that determiners project. The node D is therefore always dominated by DP. This should yield the structure shown in (6b), however, this is precisely one of the structures that was ruled out \textit{a priori}. In fact, the only possible result is the structure in (6c).

(6) (a) N\(^0\)
     \[ N\(^0\) \quad \text{Adj}^0 \]

(b) * DP
     \[ D\(^0\) \quad N\(^0\) \]

(c) DP
     \[ D\(^0\) \quad \text{NP} \]

To summarize then, I suggest that while certain X\(^0\) categories may be inherently non-projecting in syntax, they may still appear as XPs, provided they are selected by a projecting
head. However, the fact that they are inherently non-projecting means that they will never appear with filled specifier and complement positions.

As mentioned above, most of Chapter 2 will be dedicated to showing that nouns do not project in Persian. Thus we will see instances of $X^0$-adjunction within the noun phrase, but we will also see that a noun can be dominated by an NP node in accordance with the assumptions just outlined.

1.2.6 Referentiality, (In)definiteness, and Specificity

In this section I clarify my use of the terms 'referential,' 'definite,' and 'specific,' since there seems to be some terminological inconsistency in this area. Let us first consider the distinction between referential and non-referential noun phrases.13

Consider the following characterizations of referentiality. Givón writes that referentiality is "the speaker's intent to 'refer to' or 'mean' a nominal expression to have non-empty references - i.e. to 'exist' - within a particular universe of discourse (i.e. not necessarily within the real world)" (Givón 1978:293-4). According to Rapoport (1987), Akmajian (1970) uses 'referential' to describe a noun phrase that has a specific referent in the universe of discourse. Rapoport uses the term to indicate that a noun phrase denotes an individual in the world. What is common to all these views is that non-referential noun phrases do not introduce a discourse referent. Rather they are used as a label, referring to type not token.

In English non-referential nouns occur within compounds. Some examples are given in (7) below. Generic nouns are also non-referential. Some examples are given in (8). I will use the terms 'generic' and 'non-referential' interchangeably in this work.

(7) (a) towel rack

13Some of the discussion in this section is based on Matthewson (1995) which gives a general and very clear overview of these terms as they pertain to Salish.
(b) *truck* driver
(c) *food* shopping

(8) (a) *Beavers* are industrious.
(b) I like *apples*.

It was assumed until recently that referentiality was a property of maximal projections (Sproat (1984)). That is, within compounds, for example, the non-referentiality of nouns was a consequence of the fact that they were simply X0s (Di Sciullo and Williams (1987), Williams (1989)). However, with the advent of the DP hypothesis (Abney (1987)) and Higginbotham's (1985) view of variable binding there is another possibility. Higginbotham proposes that common nouns have an open place allowing them to denote a type. Recall that this is the definition of non-referentiality assumed here. These open places, or variable positions, can be bound by a determiner. I propose that this binding is what gives the noun a referential reading. It follows then that a DP will always be referential since it will be headed by an element that can bind the variable position of its nominal complement. NPs on the other hand will not be referential if they do not occur within a DP. This is what allows them to act as predicates. Stowell (1989) also makes this point, stating that whenever a common noun phrase functions referentially it must have a determiner. NPs are not referential on their own.14

14There are counterexamples in English to the claim that DPs are always referential. For example, consider the following examples. The sentences in (i) are taken from Stowell (1989:251). The examples in (ii) also show that generics can appear with the indefinite or the definite determiner.

(i) (a) Bob called Stan *(a) fool.*
    (b) This book will make John *(the) most famous person I know.*

(ii) (a) John is a *student.*
    (b) *The beaver* is an industrious animal.
Turning now to the distinction between definite and indefinite noun phrases, I assume that definites are \textit{familiar} in the discourse, i.e. known to the speaker and presupposed to be known to the hearer, while indefinites are \textit{novel} to the discourse, i.e. not presupposed to be known to the hearer (see Kamp (1981), Heim (1982) for a theory of discourse based on these definitions).

Finally let us consider the difference between specific and nonspecific noun phrases. Definite noun phrases are always specific, while non-specific noun phrase are always indefinite. So there is only one context in which specificity can actually be distinguished from (in)definiteness and that is in the specific use of indefinites. As Karimi (1989) notes, the specific use of an indefinite can be distinguished from the nonspecific use by pronominalizing with \textit{it} or \textit{one}:

\begin{quote}
(9) Mary was looking for a sloop, and
She found one. (Nonspecific)
She found it. (Specific)
\end{quote}

Karimi (1989, 1990) distinguishes among definites, indefinite specifics, and indefinite nonspecifics as follows: The referent of a definite noun phrase is known to the speaker and presupposed to be known to the hearer also. The referent of an indefinite specific noun phrase is not presupposed to be known to the hearer, i.e. the hearer is not supposed to be able to identify the referent, although the speaker has a particular referent in mind. The referent of an indefinite nonspecific noun phrase signals a yet unfamiliar referent, i.e. the speaker does not have a particular referent in mind.

Stowell discusses the presence of determiners in predicative noun phrases in English. He suggests that determiners are present for semantic reasons to enable the noun to form a certain kind of predicate. However, he does not deal with the difference between definite and indefinite determiners in non-referential noun phrases, nor with generics in non-predicative contexts. I leave these matters aside here as the problem of determiners within non-referential or generic noun phrases does not arise in Persian.
The resulting four way distinction is summarized below:

\[(10)\]

<table>
<thead>
<tr>
<th>Definite (Specific)</th>
<th>Indefinite Specific</th>
<th>Indefinite Non-specific</th>
<th>Non-referential/Generic (Non-specific)</th>
</tr>
</thead>
<tbody>
<tr>
<td>•particular referent •known to speaker •presupposed to be known to hearer</td>
<td>•particular referent •known to speaker •not presupposed to be known to hearer</td>
<td>•some referent but unfamiliar</td>
<td>•no referent</td>
</tr>
</tbody>
</table>

The distinction between referential and non-referential noun phrases is relevant to the discussion in Chapters 2 and 3 regarding the structure of the noun phrase, specifically the interpretation of NP vs. DP. The issue of definiteness vs. specificity plays a large role in the discussion of the morpheme \(-\text{râ}\) in Chapter 5.

1.3 Inflection vs. Derivation in Persian

Since a substantial part of this thesis is concerned with determining the syntactic structure of the Persian clause, including the functional projections, and given the assumptions outlined above, I now turn to the criteria for distinguishing inflectional from derivational morphology in Persian. Stress assignment seems to be a straightforward way to distinguish inflectional from derivational affixes. In this section I outline the stress system of Persian.

1.3.1 The Stress System in Persian and Stress-affecting Affixes

In Persian stress generally occurs on the final syllable of a word. Exceptions to this rule, according to Lazard (1992), are words used in isolation, or "words that are placed at the
beginning of a sentence or a sentence constituent and mark articulations of the discourse" (pp. 38-9). Within the lexical categories stress can be affected by affixation. Since prepositions do not bear affixes only nouns, adjectives and verbs will be discussed in the following two sections. The discussion and accompanying examples are based on Lazard (1992:39-41).

1.3.1.1 Nouns and Adjectives

Stress falls on the final syllable of the stem. Affixes that affect stress placement are the plural suffix on nouns, the comparative and superlative suffixes on adjectives, and the suffixes of nominal derivation. In each case stress appears on the final syllable of the suffix. Some examples are given below. The vowel on which the stress is realized appears in boldface. The example in (11d) shows that when there is more than one suffix, stress falls on the final syllable of the last suffix.

(11) (a) ketāb  (b) ketāb-hâ  (c) ketāb-che  (d) ketāb-che-hâ
     'book'       'books'       'little book, booklet'       'booklets'

(12) (a) kuchek  (b) kuchek-tar  (c) kuchek-tarin  (d) kuchek-i
     'small'       'smaller'       'smallest'       'smallness'

In compound nouns the main stress falls on the final syllable of the second element with the final syllable of the first element carrying a secondary stress. In very common compounds the secondary stress disappears.

(13) (a) fārsi zabān  Persian tongue  'Persian speaker'
     (b) tâze våred  new enter  'newcomer'
In order to discuss stress on Persian verbs it is necessary to have a basic understanding of the verbal morphology. Each verb in Persian has a present stem and past stem from which all tense/aspect forms are derived. These are referred to as Stem I and Stem II respectively by Lazard (1992). The past stem is also sometimes referred to as the short infinitive. Some examples of the past and present stems of regular verbs are given below. In these cases the past stem consists of the present stem plus the stress-bearing suffix -id.

<table>
<thead>
<tr>
<th>Past Stem</th>
<th>Present Stem</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) xâbid-</td>
<td>xâb-</td>
</tr>
<tr>
<td>(b) tarsid-</td>
<td>tars-</td>
</tr>
<tr>
<td>(c) xarid-</td>
<td>xar-</td>
</tr>
<tr>
<td>(d) xarâshid-</td>
<td>xarâsh-</td>
</tr>
</tbody>
</table>

'sleep'
'fear'
'buy'
'scratch'

While many Persian verbs are regular, many others, including the most common verbs, have irregular past and present stems. In Ghomeshi (1992) it is argued that the relation between the past and present stems of irregular verbs is not rule-governed and the stems must simply be memorized. Examples of irregular verbs are given in (15) below.

<table>
<thead>
<tr>
<th>Past Stem</th>
<th>Present Stem</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) xând-</td>
<td>xân-</td>
</tr>
<tr>
<td>(b) oftâd-</td>
<td>oft-</td>
</tr>
<tr>
<td>(c) bâxt-</td>
<td>bâz-</td>
</tr>
<tr>
<td>(d) nevesht-</td>
<td>nevis-</td>
</tr>
<tr>
<td>(e) kard-</td>
<td>kon</td>
</tr>
<tr>
<td>(f) did-</td>
<td>bin-</td>
</tr>
</tbody>
</table>

'read'
'fall'
'lose'
'write'
'make/do'
'see'

15It is interesting to note that many of the present stems of regular verbs can be used as nouns. This may reflect the fact that the suffix -id was at one time a productive affix forming verbs from nouns. However, now new verbs are created from nouns by means of compounding (see Chapter 5, section 5.6).
As stated above, all tenses are formed from one of the two stems of the verb. The present, the simple subjunctive and the imperative are formed with the present stem. These are shown in (16). The preterite and the imperfect use the past stem. These are shown in (17). The compound tenses are formed from the past participle, which consists of the past stem plus the stressed suffix -e. The compound tenses are shown in (18). The one exceptional compound tense is the future, shown in (19). The future consists of an inflected form of the verb 'to want' followed by the past stem.\footnote{In the colloquial language the future is indicated by using the present. The future given in (19) is very formal.} Labels and examples are taken from Lazard (1992).

(16) present
   simple subjunctive \textit{mi}+stem I+agr \textit{mi-xar-am} 'I buy'
   imperative \textit{be}+stem I+agr \textit{be-xar-am} 'that I buy'

(17) preterite
   imperfect \textit{mi}+stem II+agr \textit{mi-xarid-am} 'I was buying'

(18) perfect
   compound imperfect \textit{mi}+past part+agr \textit{mi-xaride-am} 'I have been buying'
   pluperfect past part preterite of 'be' \textit{xe} \textit{riede} bud-am 'I had bought'
   double compound past past part perfect of 'be' \textit{xe} \textit{riede} bude-am 'I had bought'
   compound subjunctive past part subj of 'be' \textit{xe} \textit{riede} bâsh-am 'that I bought'

(19) future
   \textit{xâh-}am \textit{xarid} xarid 'I shall buy'

There is also a long infinitive or verbal noun formed by suffixing stressed -\textit{an} to the past stem (e.g. \textit{xaridan} 'to buy'). Lazard states that this long infinitive has all the properties of a noun.
Negation is indicated by prefixing *na-* (*ne-* before *mi-*) to the verb. This prefix is in complementary distribution with the subjunctive/imperative prefix. Otherwise it appears on the first element. Examples are shown below:

(20) present  ne-mi-xar-am  'I don't buy'
simple subjunctive  na-xar-am  'that I not buy'
imperative  na-xar  'Don't buy!'
preterite  na-xarid-am  'I didn't buy'
perfect  na-xaride-am  'I haven't bought'
future  na-xâh-am xarid  'I shall not buy'

The tense morphology in Persian suggests that tense should be treated as derivational. The continuous prefix *mi-*, the imperative/subjunctive prefix *be-*, and the negative prefix *na-*, all attract stress. Otherwise stress appears on the last syllable of the stem or the participle. That is, all verbal morphology except the agreement morphology is stress-affecting.17

1.3.2 Non Stress-affecting Affixes

Lazard (1992) states that the following "particles" or suffixes are unstressed and pronounced with the preceding word.

(21) (i) The conjunction -*o* 'and'18
(ii) The particle *ke* in some of its uses (this particle introduces various kinds of subordinate clauses)

---

17The formal future is quite mysterious in this regard. In this form the last syllable of the auxiliary, which is usually an agreement morpheme, bears the main stress. This is the only context in which the agreement morphemes appear with stress. Note however, that this is also the only form in which the 'auxiliary' appears before the main verb. Unfortunately an analysis of this construction is beyond the scope of the present work and is left for future research.

18The conjunction also has a non-clitic form *va*. 
In accordance with the assumptions outlined in this chapter, these are the inflectional morphemes in Modern Persian. The first four morphemes will not be considered in this work. The syntactic and semantic properties of the last six morphemes will be the subject of the following chapters. It will be shown that these morphemes either head functional projections in the syntax (in the case of (v)-(vii)) or are spelled-out at PF in certain configurations ((viii)-(x)).

1.3.3 Section Summary

We have seen that most affixes in Persian affect stress. These are the derivational affixes. We have also seen that a small group of affixes do not affect stress placement and these have been identified as the inflectional affixes. Consider the following two examples, which demonstrate two of the uses of the suffix -i. In (22) it appears that the addition of this suffix to a noun yields an indefinite reading. In (23) addition of this suffix to a noun yields a restrictive reading for the relative clause that follows, modifying the noun.

(22) (a) ketâb book (generic) / 'the book'
     (b) (yek) ketâb-i 'a book'
These two apparently different meanings of the suffix -i have generally led to the conclusion that there are two different suffixes with the same phonological shape. Thus, in the list from Lazard that was given in (21) above, this suffix appeared twice as (v) and (vi). However, given the assumption of monosemy made above (see section 1.2.1), it would be desirable to unify the two uses of this morpheme. In fact, this appears to be possible if the meaning of -i involves selection or partitivity. Assuming that non-referential nouns refer to sets, and taking -i to mean 'one of a class', which is often the reading given in grammars, the presence of -i will yield a referential reading for a noun by specifying that an actual member of the set is being referred to.19 When attached to a noun modified by a relative clause, the fact that -i implies a partitive reading results in the head noun receiving a set interpretation and a restrictive reading of the modifying clause. While the details have to be worked out, it seems desirable and possible to posit a single morpheme -i in these two cases.20

Consider now the data in (24) (again stress is indicated by boldface):

---

19 In the absence of the suffix -i a bare noun is construed as non-referential in direct object position but definite in subject position. This fact is discussed in Chapter 3 where a null definite determiner is posited (see sections 3.4 and 3.5.3).

20 Of course, this morpheme also indicates that the referent of the noun phrase to which it attaches is new to the discourse (see Chapter 3, section 3.5.2 and Chapter 5, section 5.2.2), which is why it is referred to as the 'indefinite' enclitic.
(24) (a) bozorg 'big, great' bozorg-i 'bigness, greatness'
qermez 'red' qermez-i 'redness'

(b) mard 'man' mard-i 'manliness'
kolfat 'housekeeper' kolfat-i 'housekeeping'

We seem to have a third use of the suffix -i, here making abstract nouns out of adjectives and nouns. Note here that reducing the meaning of this morpheme to the characterization of the suffix shown in (22) and (23) seems almost impossible. Moreover there is one big difference between this suffix and the one discussed above. In (24), -i attracts the stress on the word. Most nouns and adjectives are stressed on the final syllable, and nouns derived by -i are stressed on -i. This phonological difference entails a morphological difference: this morpheme -i is derivational rather than inflectional. We can therefore consider this -i a different affix altogether from the one in (22) and (23).21

1.4 Outline of the Thesis

The rest of the thesis is organized around the inflectional affixes of Persian. In Chapter 2 the Ezafe construction is considered. This construction involves an unstressed vowel (-e) that links nouns to their modifiers and possessors. To account for the distribution of this vowel, a structure for the Persian noun phrase is developed involving non-projecting heads and X0-adjunction. We see that nouns and their modifiers do not project phrasal structure and that Ezafe insertion serves to identify a string of X0s as constituting a single constituent.

In Chapter 3 the structure of the noun phrase above the Ezafe domain is considered. This involves a discussion of possessors and the morpheme -i. It is shown that both possessors and -i form a DP above the Ezafe domain. It is also argued that there is a definite counterpart

21Note that I am not assuming here that the same affix can be both inflectional and derivational. For a different view see Cowper (1989, 1991, 1995).
to -\( i \) which is phonetically null. Both -\( i \) and this null definite marker are D-heads projecting to DP. Possessors appear in the specifier of, and receive case from, this null determiner.

In Chapter 4 the agreement morphemes and the pronominal enclitics are discussed and contrasted. It is shown that both sets of morphemes are the spell-out of F-features on a governing head. The difference between them is that the agreement morphemes are the spell-out of F-features that subcategorize a verb and must assign case to a noun phrase that they govern. The pronominal enclitics are the spell-out of F-features that can attach to any head. They absorb case and thus can be checked only by pro or a noun phrase that can receive case independently. The discussion in this chapter also covers the binding of pronominals and anaphors in Persian.

In Chapter 5 the final inflectional morpheme, -\( r\ddot{a} \), is discussed. This morpheme seems to appear only on definite direct objects. However, we see that this is not actually an accurate characterization of this morpheme (cf. Karimi (1989, 1990)). We see that -\( r\ddot{a} \) case-marks presupposed DPs adjoined to VP. It is shown that this position can be viewed as a clause-internal topic position. The second half of the chapter presents an analysis of verbal compounds. The analysis encompasses both transitive verbal compounds and intransitive verbal compounds, traditionally known as impersonal constructions. It is argued that both constructions involve a VP-internal Predication Phrase (Bowers (1993)).
Chapter 2

The Ezafe Construction and Non-Projecting Nouns

2.0 Introduction

In this chapter the Ezafe construction is examined. This construction is one of the most striking features of the noun phrase in Persian. Ezafe literally means 'addition' and, according to Karimi and Brame (1986), is named after the Arabic *idafa(t)*. It refers to an unstressed vowel -e (ye after vowels other than i)\(^1\), linking elements of certain phrases together. The origin of the Ezafe vowel is hard to trace. Windfuhr (1979) refers to "the traditional insight that e originally is the predicative 'is'" and cites Lumsden (1810) in this regard. Moyne (1971) states that Ezafe has developed historically from a non-enclitic relative pronoun in Old Persian. Moyne and Carden (1974) claim that the Ezafe construction is historically a reduced relative clause.

Before proceeding, let us consider some examples of the Ezafe construction. The Ezafe vowel appears between a noun and its (apparent) complement, between a noun and a possessor, and between a noun and a modifier. However, it does not appear on bare nouns, as shown in (2) where a bare noun occurs in subject position. (In these and the following examples in this section the Ezafe vowel is italicized.)

(1) (a) xordan-e āb
    drinking+EZ water
    'the drinking of water'

---

\(^{1}\) Samiian (1983, fn. 3, p. 77) gives some evidence for positing an underlying -e with a glide insertion rule rather than -ye with a glide deletion rule based on the fact that a rule of glide insertion is needed elsewhere.
(b) xordan-e jiān
drinking+EZ Jian
'Jian's drinking'

(c) āb-e xordan
water+EZ drinking
'drinking water'

(2) ketāb ru-ye miz-e
book on+EZ table+3sgS
'The book is on the table.'

It appears between an adjective and its complement (3a), between a preposition and its complement (3b), and more generally between any two items that have a connection to one another (3c-d).

(3) (a) montazer-e jiān
waiting+EZ Jian
'waiting for Jian'

(b) posht-e manzel
behind+EZ house
'behind the house'

(c) shahr-e kermān
city+EZ Kerman
'the city Kerman'

(d) āqā-ye qomeshi
sir+EZ Ghomeshi
'Mr. Ghomeshi'
It can iterate within a phrase linking several elements together, as shown in the following phrase taken from a story in the grammar by Thackston (1983:197):

\( \text{(4) } \ldots \text{ru-ye ânten-e rádyo-e posht-e bâm-e xâne-ye moqâbel on+EZ antenna+EZ radio+EZ back+EZ roof+EZ house+EZ opposite} \) '...on the radio antenna of the roof of the opposite house'

To summarize then, the Ezafe vowel appears between elements within noun phrases, adjective phrases and most prepositional phrases. The only phrasal category in which Ezafe is not found is the verb phrase.

In this chapter an analysis of the Ezafe construction is presented. The analysis reveals an interesting property of nominal elements in Persian, namely that they do not project to phrasal categories. The Ezafe vowel is analyzed as a linker affixed to \( X^0 \)s at PF to identify elements forming a nominal constituent.

The chapter is organized as follows. The immediately following section gives a brief overview of the way the Ezafe construction has traditionally been described in the literature. In section 2.2 one of the more recent and detailed analyses of the Ezafe construction, that of Samiian (1983) is presented. In contrast to traditional approaches, Samiian abstracts away from the different semantic relationships that the Ezafe construction can encode in order to provide a unified syntactic treatment. That approach is taken here as well.

In section 2.3 the properties of the elements linked via the Ezafe vowel to a head noun are presented. For convenience I refer to the elements that are linked by means of Ezafe as occurring in the "Ezafe domain". We see that the Ezafe vowel never attaches to a phrasal element. Thus elements within the Ezafe domain cannot take complements. This brings Persian in line with other languages (English, French) in which certain nominal modifiers cannot be phrasal. In section 2.4 two accounts of this fact are presented. The first involves a filter ruling out phrasal elements within the Ezafe domain. This approach is the one taken by Samiian, as well as a number of others for English and Romance (Emonds (1976), Williams
(1982), Giorgi and Longobardi (1991)). The second approach involves constraining the projection of nominal elements. We see that in Persian the second approach has many advantages.

In section 2.5 the distribution of Ezafe in the three categories (nouns, adjectives and prepositions) is examined. It is proposed that the trigger for Ezafe insertion is the feature [\(+N\)]. Evidence is presented that the prepositions that occur with Ezafe should be considered [\(+N\)]. The claim that Ezafe is present for case reasons, as has been suggested by several linguists (Karimi (1989, 1990), Hashemipour (1989), Samiian (1994)), is also discussed and shown to be implausible.

Since the analysis of the Ezafe construction developed in this chapter involves the idea of non-projecting heads, the question of projection is discussed in section 2.6. In section 2.7 the difference between items occurring in the Ezafe construction and true compounds is discussed. In section 2.8 some concluding remarks are made and some further implications of the ideas developed in the chapter are discussed.

2.1 Descriptions of the Ezafe Construction

Palmer (1971) provides a detailed study of the Ezafe construction within the framework of Case Grammar (cf. Fillmore (1968)). Palmer first summarizes other descriptions of the Ezafe construction, in particular those of Phillott (1919) and Lazard (1957). All of these descriptions classify the Ezafe construction in terms of the semantic relation expressed by the construction. Some of the classifications of Ezafe that Palmer gives, based on Phillott's work, are shown below. Phillott in turn bases his discussion on the Persian grammarians' traditional classes of the Ezafe.²

²I have taken some of the clearest examples. The examples have been modified slightly to make them consistent with the transliteration and glossing system used in this thesis.
(i) *The Qualifying Ezafe*: The second member of the Ezafe construction is an adjective that modifies the first member.

```
mard-e  xub
man+EZ  good
'the good man'  [Palmer 1971:11.36]
```

(ii) *The Agential Ezafe*: The first member expresses the idea "performer of the action," and the second member specifies the object of the action.

```
forushande-ye ketâb
seller+EZ  book
'seller of books'  [Palmer 1971:13.47]
```

(iii) *The Possessive or Property Ezafe*

```
shâh-e  irân
Shah+EZ  Iran
'the Shah of Iran'  [Palmer 1971:14.53]
```

(iv) *The Particularizing or Specifying Ezafe*

```
dokân-e  qassabi
shop+EZ  butcher
'butcher shop'  [Palmer 1971:15.60]
```

(v) *The Descriptive Ezafe*

```
sâ-at-e  talâ
watch+EZ  gold
'gold watch'  [Palmer 1971:16.65]
```
In total, Palmer gives 18 types of Ezafe construction, based on Phillott, a number of these having subtypes also. For example under the *Obligatory Ezafe* we find constructions such as the following:

(vi) *Proper Name + E + Profession*

mohammad-xân-e tâjer
Mohammad+Khan+EZ merchant
'Mohammad, the merchant' [Palmer 1971:20.92]

(vii) *The Words "Country", "City", "River", etc. + E + Name*

shahr-e kermân
city+EZ Kerman
' the city Kerman' [Palmer 1971:21.97]

rud-e nil
river+EZ Nile
' the river Nile' [Palmer 1971:21.98]

(viii) *"Book" + Title*

ketâb-e golestân
book+EZ Golestan

Palmer also summarizes Lazard (1957), with respect to the Ezafe construction. Lazard's work is one of the most comprehensive grammars of Persian. Since the English translation (1992) is being used in this thesis, I will take Lazard's description of the various Ezafe constructions from this more recent work.
Lazard states that the sole purpose of the Ezafe vowel is to mark "nominal
determination". He goes on to state that the Ezafe vowel "...indicates nothing as to the nature
of the semantic relation which unites the modifier and the head noun. This relation emerges
only from the meaning of the terms and from the context." (p. 68) He then goes on to
classify the different relations expressed through the Ezafe construction into five types. All
examples below are taken from Lazard (1992:68-70)

(i) **Qualification by an adjective (or a participle, or a noun used as a qualifier)**

áb-e
garm
water+EZ warm
'hot water'

ayyâm-e
gozashte
days+EZ past
'past days'

sa'di-ye
shirâzi
Sa'di+EZ Shiraz
'Sa'di of Shiraz' (lit., Sa'di the Shirazian)

(ii) **Situation: the modifier is an adverb or an adverbial phrase**

javân-hâ-ye
emruz
young+pl+EZ today
'the young people of today'

sâl-e
qabîl
year+EZ before
'the year before'
(iii) **Qualification by a noun (or an infinitive), which indicates a character or a property of the thing mentioned**

åb-e cheshme
water+EZ well

'well-water' Origin

tâj-e zar
crown+EZ gold

'gold crown' Material

åb-e xordan
water+EZ drinking

'drinking water' Appropriateness

åb-e zendegi
water+EZ life

'Water of [Eternal] Life' The Product

(iv) **Appurtenance: the determinant is a noun (or an infinitive)**

a) possession

xâne-ye hasan
house+EZ Hasan

'Hasan's house'

b) other relationships (origin, aim, partitive relation, locative relation, etc.) giving definite value to the head noun

dard-e jarâhat
ache+EZ wound

'the aching of the wound'
What we can observe from the brief summary above is that traditionally the Ezafe construction has been classified in terms of the relationships expressed between the joined elements. Palmer takes this further to develop a transformational account of the various types of construction from underlying paraphrases within Case Grammar. Specifically, he derives all instances of the Ezafe from underlying relative clauses involving the verbs dâshtan 'to have' and budan 'to be'.

A question to be considered is whether the various semantic relationships expressed through the Ezafe construction should be encoded syntactically. The alternative is to have a uniform syntactic configuration for all occurrences of the Ezafe vowel, encoding a number of different semantic relationships. Radford (1981), citing Chomsky (1977), refers to this approach as the Autonomous Syntax thesis, a view that is pursued here. This approach is also

---

3Even in more recent grammars like Thackston (1983) there are at least two kinds of Ezafe construction. Thackston calls one the descriptive adjectival, which he says serves to link a descriptive adjective to a noun. The second type he calls the possessive and he writes "The ezâfe construction, while called possessive, is used extensively in Persian and ranges into meanings far beyond the English possessive. It is used, in fact, to indicate any type of syntactical relationship between two nouns." (p. 39)
the one taken by Samiian (1983, 1994) and by Karimi and Brame (1986). In all of these works the authors seek to provide a unified syntactic analysis of the various types of Ezafe constructions. Their work will be summarized in the sections that follow.

2.2 Samiian's Analysis of the Ezafe Construction

Samiian (1983)\(^4\) notes that the Ezafe vowel occurs between a noun and its adjectival, prepositional, and possessive\(^5\) modifiers. It also occurs between an adjective and its noun phrase complement, and between most prepositions and their noun phrase complements. In all of these cases the Ezafe is obligatory. Examples of the Ezafe construction within the noun phrase are given below:

(5) Noun + Adjectival modifier
(a) \(\text{otâq-e kuchik}\)
 room+EZ small
 'small room' [Samiian 1983:39.34]

(b) \(\text{otâq-e zir-e shirvuni}\)
 room+EZ under+EZ roof
 'the room under the roof' [Samiian 1983:39.35]

---

\(^4\)This section is based on Samiian's (1983) thesis. Her more recent (1994) paper does not differ substantially from the analysis found in her thesis except with respect to prepositions. Minor differences are noted where relevant. The more recent work is discussed only in section 2.5 which includes a discussion of prepositions.

\(^5\)It should be noted that Samiian refers to the possessive noun phrase as the 'genitive noun phrase'. I am using the former term so as not to imply a link between this and genitive case.
Noun+Possessor

(c) otâq-e ali
room+EZ Ali

'Ali’s room' [Samiian 1983:39.36]

Samiian shows that all three modifiers can occur (see (6)) but only in a fixed order (see (7)).

(6) otâq-e kuchik-e zir-e shirvuni-e ali
room+EZ small+EZ under+EZ roof+EZ Ali

'Ali's small room under the roof' [Samiian 1983:39.38]

(7) *otâq-e zir-e shirvuni-e kuchik-e ali
*otâq-e zir-e shirvuni-e ali-e kuchik
*otâq-e ali-e kuchik-e zir-e shirvuni
*otâq-e ali-e zir-e shirvuni-e kuchik

Samiian gives (8) as the structure of the Ezafe domain, where the first NP following the head is the attributive noun phrase and the last NP in the sequence is the possessive noun phrase.6

(8)

Samiian claims that the Ezafe vowel is transformationally inserted. She writes: "Although phonologically the Ezafe seems to be attached to the preceding element, it is syntactically motivated by the relationship between the head noun and the phrasal modifier

6Samiian does not explain why the constituents under N' must appear in this order. This remains a problem under the account developed here (see footnote 16).
and therefore it is triggered by the occurrence of the latter." (p. 39)\textsuperscript{7}. She gives the following rule to account for the presence of the Ezafe vowel:

\begin{equation}
\text{Ezafe Insertion Rule}
\end{equation}

\[ X^{\text{max}} \rightarrow e + 1 \quad \text{When } X^{\text{max}} \text{ is immediately dominated by } Y', X \neq Y \neq V \]

[Samiian 1983:67.161]

This rule inserts the Ezafe vowel before every non-verbal phrasal category that occurs below the X' level. In the structure in (8) above, then, the vowel -e will occur before each of the phrasal constituents under N'.

The idea of a transformationally inserted morpheme with little or no semantic content is odd given current views on syntax and morphology. In fact, even to consider the Ezafe vowel as a morpheme is problematic. Samiian presents arguments against the Ezafe vowel being analyzed as a preposition (1983, section 2.3.1). Further, if the Ezafe vowel is viewed as a morpheme heading any sort of projection, it is highly unusual given that it iterates throughout the noun phrase and appears in adjective and prepositional phrases also. Nor can this vowel be thought to attach presyntactically at the LSI since it does not instantiate any known inflectional feature. Therefore, the view that the Ezafe vowel is inserted in a certain syntactic configuration is maintained here. The precise rule is given in Chapter 3 after a discussion of the possessive construction. The insertion rule is more akin to phonological phrasing rules rather than the spell-out of a morpheme.

In addition to the Ezafe Insertion Rule, Samiian (1983) gives the filter shown in (10) below. This filter accounts for the constraints on elements that appear within the Ezafe domain. (These constraints are outlined in detail in the following section.)

\textsuperscript{7}Samiian gives evidence in a footnote (1983, fn. 10, p. 78) from the distribution of pauses to support her claim that phonologically -e is attached to the preceding element.
The filter shown above prevents elements dominated by N' from having complements. Samiian notes (p. 60) that the filter above does not provide an explanatory account of the facts, although she points out that no other account has an explanation of the facts either. Further, the possessive noun phrase is an exception to the distributional restrictions Samiian observes, hence it appears in the structural description of the filter. Again she states that she has no account of why the possessive noun phrase is exceptional in this regard but she states that "[n]either do any alternative accounts of Ezafe provide an explanatory solution to the recursion restriction mentioned above." (p. 60) The next section presents the evidence for such a filter in Persian. Specifically, we see that all modifiers within the Ezafe domain occur only as X0's.

2.3 Properties of the Ezafe Domain Within the Noun Phrase

The structure that Samiian gives for the Ezafe domain (within noun phrases) is repeated in (11)

(11) \[ N', N \ldots X \ Y \ ] N_{\text{max}} \]

a. if \( Y \neq \emptyset \) for \( X = N, A, P^8 \)

b. if \( X = P \), then \( Y = S \)  

[Samiian 1983:58.132]

As mentioned above, Samiian shows that each of the constituents following the head noun and preceding the possessive noun phrase has a restriction on it. In this section each of these

---

8I have given Samiian's filter exactly as she states it. Note however, that since in (b) there is a special statement about prepositions, (a) should probably read "for \( X = N, A \)".
constituents will be considered in turn, with the goal of seeking a generalization about these restrictions.

2.3.1 Attributive Noun Phrases

The attributive noun phrase is the closest modifier of the head noun within the noun phrase. Examples are given in (12) below:

(12) (a) angoshtar-e firuze
    ring+EZ turquoise
    'turquoise ring'  \[Samiian 1983:45.63\]

(b) kif-e charm
    bag+EZ leather
    'leather bag'  \[Samiian 1983:45.65\]

(c) mâshin-e raxtshui
    machine+EZ washing
    'washing machine'  \[Samiian 1983:45.66\]

(d) sabzi-e ash
    herbs+EZ soup
    'vegetables for the soup'  \[Samiian 1983:45.67\]

Attributive noun phrases can cooccur with adjective phrases, but must precede them, as shown in (13). They can also cooccur with possessive noun phrases, as shown in (14).

---

9For some reason Samiian glosses \textit{kif} in this particular example as 'purse', but in the rest of her examples glosses it as 'bag'. I have therefore changed her gloss and translation for consistency. In fact, \textit{kif} can mean either 'bag' or 'purse'.

Samiian shows that while the possessive noun can be modified, the attributive noun cannot. For example in (15) we see that the possessive noun, but not the attributive noun, can occur with a demonstrative.

(15) (a) kif-e in mard
    bag+EZ this man
    'this man's bag' [Samiian 1983:53.103]

(b) *kif-e in charm
    bag+EZ this leather
    *'this leather's bag' (i.e. a bag of this leather) [Samiian 1983:53.104]

The attributive noun phrase and the possessive noun phrase are also distinguished by the fact that the former occurs closest to the head noun while the latter is the farthest from the head, within the Ezafē domain. An example involving an attributive noun, an adjective, and a possessive noun phrase is shown in (16).

(16) kif-e charm-e bozorg-e dust-am
    bag+EZ leather+EZ big+EZ friend+1sg
    'my friend's big leather bag'
Now, as Samiian mentions, and as can be seen from the examples above, attributive noun phrases surface only as bare nouns. However, Samiian rejects the idea that they are syntactically N⁰s rather than NPs for three reasons. First, she notes that if they were N⁰s, the notion of the 'head of Xⁿ' would have to be redefined in terms of the leftmost (in Persian) Xⁿ⁻¹. Second, she notes that allowing non-maximal projections as complements to heads violates the X-bar principles she is assuming (cf. Jackendoff (1977)). Finally, she notes that allowing attributive noun phrases to surface as just N⁰s would cause problems for her formulation of the Ezafe Insertion Rule.

Note that Samiian's three reasons for not allowing attributive nouns to surface as X⁰s are theoretical rather than empirical. Of these three problems, one is eliminated by the version of X-bar theory assumed here, and a second by a reformulation of the Ezafe Insertion Rule. The problem of headedness remains and is discussed in section 2.8.

2.3.2 Adjective Phrases

We have already seen that adjectives can modify nouns in the noun phrase and are linked to the noun by the Ezafe vowel. Samiian shows that, like attributive nouns, these modifiers too involve certain restrictions. Using the following examples, Samiian presents adjectives linked to head nouns via Ezafe (17a)-(19a). She then shows that these adjectives cannot take noun phrase, prepositional phrase or sentential complements in the Ezafe construction (17b)-(19b). However, such adjective phrases can occur as appositive modifiers. This is shown in (17c)-(19c). She concludes from this that the adjective phrase must be head-final within the Ezafe domain.
Adjective with a noun phrase complement

(17) (a) mard-e negarân-i vâred shod-Ø
man+EZ worried+indef enter became+3sgS
'A worried man entered.'

(b) *mard-e negarân-e bache-ha-sh-i vâred shod-Ø
man+EZ worried+EZ child+pl+3sg+indef enter became+3sgS
'A man worried about his children entered.'

(c) mard-i, negarân-e bache-ha-sh, vâred shod-Ø
man+indef worried+EZ child+pl+3sg enter became+3sgS
'A man, worried about his children, entered.' [Samiian 1983:42.47a-c]

Adjective with a prepositional phrase complement

(18) (a) mardom-e xashmgin-e tehrân bepâxâst-and
people+EZ angry+EZ Tehran rose-up+3plS
'The angry people of Tehran rose up.'

(b) *mardom-e xashmgin az ertejâ-ye tehrân bepâxâst-and
people+EZ angry at reactionaries+EZ Tehran rose-up+3plS
'The people of Tehran angry at the reactionary forces rose up.'

(c) mardom-e tehrân, xashmgin az ertejâ, bepâxâst-and
people+EZ Tehran, angry at reactionaries rose-up+3plS
'The people of Tehran, angry at the reactionary forces, rose up.'

[Samiian 1983:42.48a-c]
As we saw in section 2.2, Samiian's account of this involves a filter preventing sisters of N^0 dominated by N' from having complements. Crucially then, for Samiian these modifiers are base-generated as phrasal constituents, even though, like the attributive nouns, they surface as bare heads. In the next section we see some apparent evidence against the generalization that modifiers linked via Ezafe are simply bare heads. However, a further examination of the data reveals that the counterevidence is only apparent, and that the generalization is correct.
2.3.3 Prepositional Phrases

So far we have seen that attributive nouns and adjectives do not seem to be phrasal within the Ezafe domain but instead occur as bare heads modifying the head noun. Prepositional phrases can also be linked via the Ezafe to the nouns they modify. We therefore expect these modifiers also to be non-phrasal. The facts here are a little more complex, however. While it is true that there are restrictions on prepositional phrases in the Ezafe domain, at first glance the restriction does not seem to have to do with an XP vs. X⁰ distinction.

Samiian shows that adverbial prepositional phrases of time, location and manner can follow attributive modifiers.

(20)  
(a) bahs-e bā ajale-ye hasan jâleb bud-Ø  
discussion+EZ with haste+EZ Hasan interesting was+3sgS  
'Hasan's hasty discussion was interesting.' [Samiian 1983:55.117]

(b) molâqât-e qabl az zohr-e mâ jâleb bud-Ø  
meeting+EZ before-noon+EZ us interesting was+3sgS  
'our meeting before noon was interesting' [Samiian 1983:55.118]

(c) xâne-ye yeylâqi-ye kenâr-e daryâ-ye hasan  
house+EZ summer+EZ next-to+Ez beach+EZ Hasan  
'Hasan's summer house on the beach' [Samiian 1983:56.119]

However, there are restrictions on the types of prepositional phrases that can occur. While appositive prepositions can have sentential complements (see (21c)), they cannot when they are modifiers (see (21b)).
These facts lead Samiian to the particular formulation of her filter, given in section 2.2 and repeated below. The filter states that a constituent between the head noun and the possessive noun phrase cannot take a complement. However, in the case of prepositions, the restriction is that the complement cannot be sentential.

(10) \* N[N [... X Y ] N_{\text{max}} ]
    a. if \( Y \neq \emptyset \) for \( X = N, A, P \)
    b. if \( X = P \), then \( Y = S \)  

The filter above predicts that a fully phrasal NP should be perfectly grammatical when serving as a complement to a preposition within the Ezafe domain. However, this is not so. If we go back to one of the early examples, we see that the noun in the modifying PP cannot appear with a possessor, while it can if the PP is appositive.
Furthermore, the noun phrase in the PP cannot be modified by a demonstrative article.

(23)  *otâq-e kuchik-e zir-e in shirvuni-e ali
room+EZ small+EZ under+EZ this roof+EZ Ali
'Ali's small room under this roof'

The examples in (22) and (23) show a further restriction on the complements of prepositions within the Ezafe domain, namely that they cannot be phrasal. Now, since we have seen that attributive noun and adjectival modifiers cannot be phrasal, the behaviour of prepositions can be unified with that of these other two categories. Assume that when a preposition takes a non-phrasal complement, the preposition does not project to a phrasal category. This is illustrated below. In (24a) and (24b) we see that a preposition occurring with a noun phrase or a sentence as its complement will project to PP. In (24c) we see, however, that if a bare noun occurs as the complement to a preposition the resulting category is not phrasal.

(24)  (a) PP (b) PP (c) P
        P     DP     P     S     P     N

This and other properties of prepositions in Persian will be discussed further in section 2.5.
2.3.4 The Possessive Noun Phrase

The rightmost element linked by the Ezafe vowel within the noun phrase is the possessive noun phrase. As we have seen, all other modifiers within the Ezafe domain must be X\(^0\)s. In contrast, the possessive noun phrase is fully phrasal and can occur with a variety of modifiers. A few of Samiian's examples are given below. (Samiian refers to the possessive noun phrase as the genitive noun phrase.)

(25) **Restrictive Relative Clause Modifying the Genitive Noun Phrase**
(a) ketâb-e un mard-i ke diruz âmad-Ø
    book+EZ that man+indef that yesterday came+3sgS
    'The book of the man who came yesterday.' [Samiian 1983:59.133]

**Non-Restrictive Relative Clause Modifying the Genitive Noun Phrase**
(b) ketâb-e pedar-am ke dar iran zendegi mi-kon-e
    book+EZ father+1sg that in Iran live cont+do+3sgS
    'The book of my father who lives in Iran' [Samiian 1983:59.134]

**Recursive Genitive Noun Phrase Complement to the Genitive Noun Phrase**
(c) ketâb-e barâdar-e dust-e xâhar-am
    book+EZ brother+EZ friend+EZ sister+1sg
    'The book of the brother of the friend of my sister' [Samiian 1983:59.135]
    (i.e. 'My sister's friend's brother's book')

Samiian notes that this "freedom of expansion for the genitive noun phrase stands in sharp contrast to the restrictions pointed out earlier on the expansion of the attributive modifiers and the prepositional phrases of time and location." (p. 60) This is why the genitive, or possessive, noun phrase appears in the structural description of her filter.
For the moment let us simply note that the possessive noun phrase can be fully phrasal and return to the other modifiers within the Ezafe domain. The possessive noun phrase will be dealt with in Chapter 3. In the next section two approaches to the non-phrasality of modifiers within the noun phrase are considered.

2.4 The Filter vs. X^0-Adjunction Approach to Attributive Modifiers

The facts that we have seen in the preceding sections are not limited to Persian. For example, English has attributive nouns that must surface as bare nouns. These [N N] constructions are generally thought to be compounds, but note that not all such constructions have compound stress, (e.g. stone wall).

Also in English, prenominal adjectives cannot take complements even though they may do so in predicate position. Thus we get the following contrasts:\textsuperscript{10}

\begin{align*}
(26) \quad & \text{(a)} \quad \text{This man is proud.} \\ & \text{(b)} \quad \text{a proud man}
\end{align*}

\begin{align*}
(27) \quad & \text{(a)} \quad \text{This man is proud of his son.} \\ & \text{(b)} \quad *\text{a proud of his son man}
\end{align*}

These facts have most recently been discussed by Sadler and Arnold (1994), who analyze prenominal adjectives in English as X^0s. Their discussion of alternative approaches to the problem posed by prenominal adjectives is detailed and comprehensive. A subset of the

\textsuperscript{10}Note the following contrast in appositive position to complete the paradigm of cases:

\begin{align*}
& \text{(i)} \quad \text{a man, proud of his son, ...} \\
& \text{(ii)} \quad *\text{a man, proud, ...}
\end{align*}

This too was noted by Emonds (1976), who stated that adjective phrases without complement PP's or S's never appear post-nominally (p. 171), although he was stating it in terms of a rule of postposing from prenominal to postnominal position. The contrast between adjectives in predicate and appositive position is discussed in section 2.6.2.
literature they cover will be discussed here, but the reader is referred to their paper for analyses using other approaches such as Categorial Grammar.

One of the earliest accounts of prenominal attributive modifiers (APs and also participles) within the generative framework was given by Emonds (1976) who proposed that such modifiers must be head-final. He assumed that this followed from a more general recursion restriction, given in (28).\footnote{Note also that this is very similar to Dryer's (1992) Branching Direction Theory, Chapter 1, section 1.2, in which this kind of recursion restriction is correlated with VO/OV word order.}

\begin{align*}
(28) & \textbf{Surface Recursion Restriction:} \\
& \text{Given a surface configuration of the form } [H_1'' \ldots A \ldots H_i \ldots], \text{ if the base rules permit right sisters } H_k'' \text{ to } H_i, \text{ then } A \neq X S Y, A \neq X PP Y, \text{ where } PP \text{ dominates a lexical preposition, and } A \neq W A Z, \text{ where } W \text{ and } Z \neq \emptyset. \text{ In such cases we say that } A \text{ does not exhibit free recursion.} \quad [\text{Emonds 1976:19.21}]
\end{align*}

Williams (1982), making the same observation, proposed the Head Final Filter, barring post-head material in prenominal modifiers. He used this filter to argue that verbal passives must involve a post-verbal trace because they can't occur prenominally while adjectival passives can.

Most recently, Giorgi and Longobardi (1991), citing the observations made by Emonds and Williams, have proposed their own version of a filter given in (29) to account for similar facts in Romance languages. They assume that this filter applies at S-structure.

\begin{align*}
(29) & \textbf{Consistency Principle} \\
& \text{An XP immediately expanding a lexical category on the non-recursive side is directionally consistent in every projection.} \quad [\text{Giorgi & Longobardi 1991:98.111}]
\end{align*}

For G&L 'recursive side' means the side of the head where complements occur. Thus their principle says that if a phrase (XP) occurs on the opposite side of a complement (say to the
left of a head, where the head takes its complements on the right) then this category can only be expanded to the left. The reason that they explicitly mention 'lexical category' in their formulation of the filter is to account for the fact that functional heads such as I and C do not seem to be subject to the same restriction. This was also noted by Emonds, who pointed out that subjects are not limited in their recursion properties.12

Let us look at Giorgi and Longobardi's filter to see how it might apply to Persian. First, note that their filter is formulated to refer to the side of a lexical category opposite to the side on which complements occur (the 'non-recursive side'). Recall that in Persian only the verb takes its complements to the left. Thus, for all other categories the non-recursive side is on the left.13 This means that a filter such as the one above should not hold of the right side of non-verbal projections in Persian.

A second problem is that, according to Giorgi and Longobardi, the filter does not hold in languages that are cross-categorically inconsistent in terms of which side is the recursive one. They cite German in this regard, where Ns, Cs and most Ps take their complements to the right but Vs take their complements to the left. This prediction is simply false for Persian, which, like German, appears to be cross-categorically inconsistent. Thus (31a), the Persian equivalent to their German example, is ungrammatical. The examples in (31b-c) show that the adjective can take a complement in predicate position and can modify a noun when it does not take a complement.14

(30)  ein seinen FreundetreuerMann
     a to his friends faithful man  [Giorgi and Longobardi 1991:98.112a]

---

12Interestingly, he used this as an argument against subjects being VP-internal.
13Of course, this depends on whether nouns take complements at all in Persian. Under the analysis developed here, nouns do not take complements or specifiers.
14With respect to (31c) note that 'faithful' can also be translated in Persian as bâ vafâ lit. 'with fidelity' showing an example of a prepositional compound, see section 2.5.3.
(31)  (a) *mard-e  be dust-ā-sh      vafādār-i  
    man+EZ to friend+pl+3sg faithful+indef  
    *'a faithful to his friends man'

    (b) in  mard bedust-ā-sh      vafādār ast  
        this man to friend+pl+3sg faithful is  
        'This man is faithful to his friends.'

    (c) mard-e  vafādār-i     mi-shnās-am  
        man+EZ faithful+indef cont+know+1sgS  
        'I know a faithful man.'

A third problem is the behaviour of prepositional phrases. For Giorgi and Longobardi, 
Romance PPs are a problem in that they are head initial and can appear within APs. This 
forces them to say that the recursive side for APs in Romance (and in fact, in German too) is 
the left one. However, this is somewhat unsatisfactory in that the generalization that 
Romance languages are uniformly recursive on the right is lost.15 More generally, head-
initial modifier PPs are a problem for any of the filters proposed above. In English, of 
course, such modifiers are not possible, but in Persian, where they are possible, some 
exceptional statement about them has to be made, hence Samiian's formulation of the filter 
given in (10), section 2.2.

There are alternatives to the filter approach. Sadler and Arnold mention transformational 
accounts as well as Abney's (1987) proposal that adjectives are heads of projections 
dominating NP, however they point out that both approaches have problems. Another 
alternative is that modifiers such as the prenominal adjectives in English are heads (X0s) 
simply adjoined to other heads, rather than projecting their own structure. This has been 
proposed by Travis (1988) for prenominal adjectives in English (cf. also Bessler (1992)). It

15They also mention a second problem with assuming that the recursive side of APs is on the 
left, see Giorgi and Longobardi 1991, p. 233, fn. 46.
is also proposed by Sadler and Arnold (1994) who refer to the resulting A-N constructions as 'small' constructions.

Such a solution is extremely attractive for Persian in that it will unite the properties of adjectives and attributive nouns, and will also explain why prepositional phrases can occur as modifiers as long as their complements are non-phrasal. Let us say then that all elements preceding the possessive noun phrase within the Ezafe domain are base-generated in an X\textsuperscript{0}-adjoined structure (see section 2.6 for more discussion on this point). The consequences of this view are explored throughout the rest of this and the following chapter, but note that the first consequence is that a filter is no longer necessary.

2.5 Ezafe Across Categories

So far we have concentrated on the occurrence of Ezafe within the noun phrase. However, this is not the only place that the Ezafe vowel appears. Thus, part of the task in reformulating the Ezafe Insertion Rule of Samiian (1983) is to determine what the trigger for this rule is. While the exact reformulation of this rule is not given until the possessive noun phrase has been discussed (see Chapter 3), in this section we examine what sorts of categories appear with Ezafe.

Let us summarize, so far, where the Ezafe vowel appears. It appears on any X\textsuperscript{0} modifier within the noun phrase up to the possessive noun phrase. From the discussion up to now, we can give the following structure for the noun phrase, where the subscripts indicate headedness (to be discussed in section 2.8).\textsuperscript{16}

\textsuperscript{16}As mentioned in footnote 6, one issue that is not dealt with here is the order of elements within the noun phrase. While this issue is an important one, I believe it is a question of semantics and not something that requires a syntactic account.
The Ezafe vowel also appears between prepositions and adjectives and their complements, as shown below:

Samiiian (1983) points out that with a syntactic feature system like that of Chomsky (1970) or Jackendoff (1977) these three categories do not form a natural class. For example, using Chomsky's system, while nouns are [+N, -V] and adjectives are [+N, +V], the common feature being [+N]; prepositions are [-N, -V] so that there is no feature common to all three categories. It is for this reason that she formulates her Ezafe Insertion Rule as applying to all X, X≠V. However, there are several other possibilities. For example, it may be that in Persian the lexical categories don't divide as predicted. Thus, adjectives may pattern with nouns and prepositions and perhaps should be considered [-V], or prepositions might pattern with nouns and adjectives and might be specified [+N]. Another possibility is that these features are not relevant to Ezafe Insertion at all.

Two proposals have been made in the literature on this point. Karimi and Brame (1986) argue that the trigger for Ezafe Insertion is the class of 'nominal' elements (they do not employ a formal system of features) and they include prepositions in this class. Samiiian (1994) on the other hand suggests that the trigger for Ezafe Insertion is the lack of a case-
assigning property. In the next two sections these two works are considered, along with their major claims. In section 2.5.1 the status of Persian prepositions under the two views is discussed. In section 2.5.2 the linking of Ezafe Insertion to case is discussed. Ultimately we will see that the trigger for Ezafe Insertion is best characterized as the feature [+N], the view implicit in both Samiian (1983) and Karimi and Brame (1986).

2.5.1 Two Kinds of Prepositions

Prepositions in Persian do not form a homogeneous class. Thus while most prepositions take their complements via the Ezafe, there is a small group that do not, and another small group that optionally take their complements with the Ezafe. Samiian (1983) proposed a phonological account for the distribution of the Ezafe vowel in prepositional phrases. However, this account had a number of problems, and since Samiian herself has more recently proposed an alternative analysis I will not outline the phonological approach here, but refer the reader to Ch. II, section 3.2 and Ch. VI, section 3.1 in her thesis.

Let us consider the properties of the three classes of prepositions in Persian. First, a very small, closed class of prepositions never take their complements via Ezafe. These are given below. Following Samiian (1994) let us refer to these as P1 prepositions.

(34) *Prepositions that must not take Ezafe (P1)*

<table>
<thead>
<tr>
<th>Preposition</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>dar</td>
<td>'in/at'</td>
</tr>
<tr>
<td>tâ</td>
<td>'to/until'</td>
</tr>
<tr>
<td>be</td>
<td>'to'</td>
</tr>
</tbody>
</table>

Samiian (1994) puts *bar* 'on/onto' here, however, according to my consultants this appears to be a mixed preposition. She also includes *barāye* 'for' and *bedune* 'without' in this class noting that these originated as P+EZ sequences, but are now properly included in this list, presumably because they have become unanalyzable. While it is true that they differ from other P+EZ sequences in that they never appear alone, I will still include them with the latter prepositions since I am only interested in the trigger for Ezafe Insertion here and not other properties.
Second, a fairly open set of prepositions obligatorily take Ezafe before their complements. Samiian (1983) gives a partial list of these prepositions. She notes that this class contains the majority of prepositions in Persian. Karimi and Brame (1986) also give a list of prepositions that must take the Ezafe morpheme before their complements. The two lists are conflated below. What should be clear is that this is a fairly open class. Following Samiian (1994) these will be referred to as P2 prepositions.

\[(35)\] Prepositions that must take Ezafe (P2)

- bedun-e ‘without'  birun-e ‘outside'
- barâ-ye ‘for'  dâxel-e ‘inside'
- bâvojud-e 'in spite of'  xârej-e 'outside'
- zir-e ‘under’  miun-e ‘among'
- kenâr-e 'next to'  vasat-e ‘in between/the middle of'
- posht-e 'behind'  atrâf-e 'around'
- pain-e 'below'  in/un taraf-e ‘this/that side of'
- nazdik-e 'close to/near'  sar-e ‘at'
- mesl-e 'like'  beyn-e ‘between'
- dor-e 'around'  pish-e ‘beside'
- alâraqm-e 'in spite of'  aleyh-e ‘against' (in an argument)
- nazd-e 'with/near' (formal)  lah-e ‘for' (in an argument)

A third, small class of prepositions optionally takes Ezafe. This class is given in (36) below. The prepositions preceded by a question mark are those for which there is some disagreement as to whether they belong here or in the class requiring Ezafe. For example, Samiian (1983) does not include kenâr ‘beside' and barâ ‘for' in the list of 'mixed' prepositions, while Karimi and Brame do; thus they appear in the list above also. I will leave
aside the question of which prepositions properly belong in this list, since it does not bear on the analysis.

(36) **Prepositions that can optionally take Ezafe**

<table>
<thead>
<tr>
<th>Preposition</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>tu (ye)</td>
<td>'inside'</td>
</tr>
<tr>
<td>ru (ye)</td>
<td>'on'</td>
</tr>
<tr>
<td>jelo (ye)</td>
<td>'in front of'</td>
</tr>
<tr>
<td>pahlu (ye)</td>
<td>'next to'</td>
</tr>
<tr>
<td>?bâlâ (ye)</td>
<td>'on top of'</td>
</tr>
<tr>
<td>?kenâr (e)</td>
<td>'beside'</td>
</tr>
<tr>
<td>?barâ (ye)</td>
<td>'for'</td>
</tr>
</tbody>
</table>

Samiian (1983, 1994) and Karimi and Brame (1986) agree that the first class of prepositions, which do not take their complements via Ezafe, are the 'true' prepositions in Persian. However, they differ slightly on how to treat the second class of prepositions. Let us look at the properties of P2s more closely.

Samiian (1983, 1994) notes that P2s, unlike P1s, seem nominal. For example P1s are obligatorily followed by a noun phrase complement, while P2s can occur alone. Also, P2s but not P1s can occur with demonstratives, with plural marking and in case positions. Some examples are given below.

**P2s can occur alone**

(37) (a) birun-e manzel-o tamiz kard-im  
outside+EZ house+râ clean did+1plS  
'We cleaned the outside of the house.'

(b) birun raft-im  
outside went+1plS  
'We went outside.'
P1s cannot occur alone

(38) (a) bâ simâ raft-im
    with Sima went+1plS
    'We went with Sima.'

(b) *bâ raft-im

P2s can occur with demonstratives and plural marking

(39) (a) in zir-e miz
    this under+EZ table
    'this under(neath of) the table'

(b) in zir-hâ-ye miz
    this under+pl+EZ table
    'these underneaths of the table'

P1s cannot occur with demonstratives and plural marking

(40) (a) *in bâ hasan
    this with Hasan

(b) *bâ-hâ
    with+pl

P2s can occur in subject (nominative) and direct object (accusative) positions

(41) (a) zir-e miz kasif-e
    under+EZ table dirty+3sgS
    'Under the table is dirty.'

---

18Samiian (1994) gives a third instance of P2s occurring in a case position, namely when they occur in the Ezafe domain. Note that in this work Ezafe is not being considered a case assigner (see section 2.6).
Given that P2s appear to behave like nouns, the question is whether they actually are. Samiian (1994) argues against this, giving four arguments to which we now turn.

Two of Samiian's arguments have to do with restrictions on P2s that she seems to believe do not hold of true nouns. She notes (a) that P2s cannot take the full range of NP "specifiers", i.e. while they can take demonstratives, they can't cooccur with numerals and quantifiers; and (b) that P2s can't be followed by a relative clause. The question is whether these two properties are necessary conditions for membership in the category Noun, given that there are items that are generally accepted as nouns that cannot occur in these contexts either. For example, an abstract noun such as bozorgi 'bigness', can't be quantified or appear with a relative clause. Perhaps even more relevantly, spatial nouns in English can't occur in such contexts either. Thus we can't say three middles or the middle that the flowers are on.

Samiian's two other arguments that P2s are not nouns are actually used by Karimi and Brame to show that P2s are nominal. Samiian states that P2s can't take adjectival modifiers. However, Karimi and Brame state that in certain contexts adjectival modification is acceptable. They give the following example:19

---

19The speakers I consulted found this to be grammatical only in a sufficiently rich context.
She also states that P2s pattern more like P1s than like nouns in that with some P2s Ezafe is optional (see (36)) above). Recall that while a P1 does not appear with the Ezafe vowel when taking a complement, a noun can never be linked to a modifier without Ezafe except in a compound. This is precisely the analysis Karimi and Brame give of P2+N constituents without Ezafe, namely that they are compounds. Let us consider their analysis.

Let us call the prepositions that optionally take their complements via the Ezafe mixed prepositions, following Karimi and Brame. They state that many of the arguments showing the nominal nature of the prepositions that take their complements via Ezafe, apply to these mixed prepositions also. However, the question is: what is the status of these prepositions when they do not appear with the Ezafe vowel? Karimi and Brame note that nouns in the Ezafe construction with other nouns will sometimes lose the Ezafe vowel, becoming compounds. The example they give is âb-e jo which literally means 'water of barley'. This phrase has now been lexicalized as a compound noun as âbjo 'beer'. They then provide three arguments for treating this class of prepositions followed by nouns without Ezafe as compounds also. First, unlike the true prepositions, the compound cases allow demonstratives.

Second, on analogy with English compounds that do not allow inflectional elements on the first member of the compound, Karimi and Brame show that when these words occur without Ezafe they cannot be inflected.
Third they show that while true prepositional phrases cannot function as direct objects, and are thus incompatible with the direct object marker -rā, these compounds do occur with this marker:

(46) (a) *be madrese-ro
to school+rā

(b) *bā doxtar-o
with girl+rā

(47) (a) ru miz-o
on table+rā

(b) tu shahr-o
in city+rā

To summarize, there are three classes of prepositions. Those that obligatorily appear with the Ezafe vowel (the P2s) appear to be nominal. Karimi and Brame show that adjectival modification and compounding are both possible with P2s lending further support to the hypothesis that they are nouns, and refuting two of Samiian's arguments against P2s being nouns. Samiian's two other arguments, namely the cooccurrence restrictions between P2s and numerals, quantifiers and relative clauses, do not necessarily rule out nominal status for P2s. As was pointed out above, these restrictions hold of other nouns also and it may
ultimately be the case that such nouns form a subclass with a clearly definable characterization. Certainly subclasses within the category Noun are required for distinctions such as animate vs. inanimate or count vs. mass. However, I suggest that there is still something that identifies these subclasses as belonging to a superset distinct from adjectives, verbs and prepositions. Let us return to the notion that for the purposes of the syntax a noun is inserted under a node bearing the features [+N, -V] (for which N is an abbreviation in the trees given here). P2s belong this category, whatever other features they may have distinguishing them from other nouns. This means that the trigger for Ezafe Insertion can now be given as the feature [+N].

2.5.2 Is There a Link Between Ezafe and Case?

Samiian's (1994) solution to the puzzle posed by P2s depends on dividing lexical items into two classes depending on whether or not they assign case. Let us first review her analysis and then more generally assess the claim that the Ezafe morpheme is linked to case.

Samiian adopts a 'Neutralization Hypothesis' that she says is akin to one proposed by van Riemsdijk (1983) for adjectives in German. Van Riemsdijk suggests that adjectives in German are neutralized with respect to their [+N] feature. They bear only the feature [+V] making them nondistinct from verbs which are [+V, -N]. Samiian proposes that P2s are neutralized with respect to their [-N] feature. The paradigm she gives for lexical categories in Persian is shown below.

\[(48) \begin{array}{ll}
N: & [-V, +N] \\
A: & [+V, +N] \\
V: & [+V, -N] \\
P1: & [-V, -N] \\
P2: & [-V] \\
\end{array} \quad \text{[Samiian 1994:38 .46]}\]

She states that case-assigning categories are [-N] and that the Ezafe vowel only appears on categories that do not assign case. Note that under this view the trigger for Ezafe Insertion
cannot be characterized by features. If we say Ezafe insertion is triggered by [+N] it will not appear on P2s, and if we say it is triggered by [-V] it will not appear on adjectives. However, an advantage to this system for Samiian is that she can maintain a distinction between P2s and nouns (see the preceding section).

If Ezafe Insertion is not triggered by the presence or absence of a feature but rather by a property of lexical items, what motivates this rule? Samiian (1994) states that Ezafe is a dummy case assigner, similar to English of. Thus it occurs within phrases containing non-case-assigning heads, enabling these heads to case-mark their complements. In fact, Samiian is not alone in believing that Ezafe is present for case reasons. For example, Hashemipour (1989) claims that Ezafe is a structural case marker attached to nouns, adjectives, and some prepositional heads. Karimi (1989, 1990) states that Ezafe structurally relates a head to its governed phrases. More specifically, in noun phrases she says that it transfers the case of the head noun to its complements.

One problem with the view that Ezafe is a case marker/transferer is the definition of 'complement' for the noun phrase in Persian. The attributive modifiers that we saw in section 2.3 are not semantically complements to the head noun. If the notion of 'complement' is purely syntactic, it is unclear how it should be defined given that more than one such 'complement' can occur in a noun phrase (i.e. multiple attributive modifiers are possible). In fact, the best characterization seems to be that a 'complement' in a noun phrase is an element that appears with the Ezafe vowel, but of course this definition is circular.

Another question that arises if Ezafe is viewed as a case marker/transferer is why there are case requirements within the NP in Persian. That is, while it is assumed that noun phrases require case, and furthermore that possessive noun phrases within noun phrases also require case, it is not necessarily clear that other types of modifiers within noun phrases, such as adjectives and attributive nouns, need case. Samiian (1994) states that there are languages that case-mark attributive adjective phrases, namely Sanskrit and Latin. Thus she suggests that in Persian anything that does not assign case, needs case in order to be licensed.
However, this does not hold across the board. For example, intensifiers (see section 2.8) occurring with adjectives are not marked with Ezafe. Nor are adverbs within the sentence. Yet it is unlikely that these items are case-assigners.

If we view the Ezafe domain as an X^0-adjointed structure, the waters are muddied even further, since usually only maximal projections need to be case-marked.\(^{20}\) The view that all attributive modifiers within the Ezafe domain are heads, then, casts doubt on the hypothesis that Ezafe is present for case reasons.

### 2.5.3 Section Summary

In this section we have looked at possible triggers for Ezafe Insertion. We have seen that considering the Ezafe vowel to be case-related leads to undesirable stipulations (e.g. having to redefine the notion of ‘complement’, and stating that nominal, adjectival and prepositional modifiers require case). In contrast to this view, considering the trigger for Ezafe Insertion to be the feature [\(+\text{N}\)] unites all the environments in which the Ezafe vowel occurs. The idea that most prepositions in Persian bear this feature is independently supported by their nominal properties elsewhere. There remains a question as to why Ezafe Insertion occurs in Persian. This is discussed in section 2.6.3.

### 2.6 The Projection Properties of Lexical Categories in Persian

We now turn to the question of projection that arises if the Ezafe domain within the noun phrase consists of X^0-adjointed heads. That is, given that X^0-adjunction is possible, what determines whether a terminal node will project to a maximal category in the syntax or

---

\(^{20}\)There are proposals in the literature that certain X^0s absorb case, for example -en (Baker, Johnson and Roberts (1989)), incorporated nouns (Baker (1988)), and clitics (Roberge (1990)). However, X^0s in general do not seem to require case.
appear simply as an X0? I propose that the lexical categories themselves can be classified as projecting or non-projecting. Specifically, in Persian, I propose that nouns never project to a maximal category on their own. This explains why they never appear with true phrasal complements. It also explains why they can be used as modifiers both within the noun phrase (i.e. attributive nouns) and as predicates within the clause (see Chapter 5). In accordance with the claims outlined in Chapter 1, section 1.2.5, nouns may appear as NPs if selected by a projecting category. However, since noun heads do not themselves project, no phrasal constituent can appear within the NP. The only possibility for further elements occurring within an NP is if they are X0-adjoined to the head.

Unlike nouns, Persian verbs always project structure (cf. Chapter 5). However, the classification of categories into purely projecting or non-projecting breaks down when we consider prepositions and adjectives. While an understanding of the projection properties of these categories requires further research, some initial observations are made in the next two sections.

2.6.1 The Projection Properties of Nominal vs. "true" Prepositions

Recall from section 2.3.3 that prepositions taking non-phrasal complements do not form a phrasal projection. Recall also that in section 2.5 it was determined that there are two classes of prepositions in Persian. Let us now consider the projection properties of both sets of prepositions.

Nominal prepositions (P2s) will not be phrasal when they involve a non-phrasal complement, since nouns also do not project to a phrasal category. In fact we can view P2s with their non-phrasal complements as being in the same configuration as nouns modified by attributive nouns.
One apparent difference between nominal prepositions and nouns is that nouns can never have *phrasal* complements, while nominal prepositions seem to be able to.

There are two possible explanations for this difference, however, neither is fully satisfactory. One solution is to maintain that the nominal prepositions are exactly like nouns in every way. While it is true that nouns do not take phrasal complements, they do occur with phrasal possessors. A nominal preposition appearing with a phrasal complement could be analyzed as having the same syntactic structure as a noun occurring with a possessor. This means that the locative in (50) above is really best rendered as *that table's top*, which is not an implausible reading. The structure developed for possessed noun phrases in Chapter 3 is given below.

```
(51)  
       DP 
      /   
     D'   DP_{poss} 
   /     
NP     D 
  |     | 
 N     Ø_{def} 
```
The problem with adopting the structure in (51) for prepositional modifiers is that syntactically they become indistinct from the noun phrases that function as the arguments of verbs. Since noun phrase arguments, which are DPs, require case and a theta-role, we can ask if the same is true of these prepositional DPs. If prepositional DPs do require case and a theta-role where do they get them from? If prepositional DPs do not require case and a theta-role, how are they distinguished from the DPs that do?

An alternative is to consider these nominal prepositions not to be true nouns but rather a separate category that may optionally project. This is shown below, where \( P_{nom} \) is a label for 'nominal preposition'.

\[
\begin{align*}
(52) \quad \text{(a)} & \quad P_{nom} \\
& \quad P_{nom} \quad N \\
\text{(b)} & \quad PP_{nom} \\
& \quad P_{nom} \quad DP
\end{align*}
\]

Recall that category labels are simply abbreviations for the feature matrices that constitute terminal nodes in the syntax. Thus we must determine the feature matrix for which the category \( P_{nom} \) is an abbreviation. Prepositions are composed of the features \([-N, -V]\). However, since the category \( P_{nom} \) triggers Ezafe Insertion, and since the trigger for Ezafe Insertion is the feature \([+N]\), the category \( P_{nom} \) must have the features \([+N, -V]\) making it indistinguishable from a noun. Under this alternative then, the features \([+/-N]\) and \([+/-V]\) are not sufficient to give the five-way division that is required.

A further problem with considering \( P_{nom} \) to be a separate lexical category is with respect to projection. We have seen that nouns may only appear with head-adjoined modifiers, while (as we will see) adjectives and verbs appear only with phrasal complements. In (52) we see a category that can appear in both configurations. I leave these issues about the true nature of P2s for further research. For convenience I will adopt the structures in (52) throughout the rest of the thesis.
Let us now turn to the other set of prepositions in Persian. Recall that true prepositions, P1s in Samiian's terms, do not take their complements in an Ezafe construction. Samiian makes an interesting claim about these prepositions in her (1994) paper, namely that PPs headed by P1s cannot occur in the Ezafe domain. This follows from her analysis. Since P1s are case-assigners, they do not require case themselves, and thus cannot appear in a case position (she cites the Case Resistance Principle of Stowell (1981) in this regard). Note, however, that if Ezafe is not connected to case assignment we require a different explanation for the absence of P1s in the Ezafe Domain.

The analysis of the Ezafe domain as an $X^0$-adjunction structure suggests that perhaps P1s cannot ever form non-phrasal constituents with their complements. In other words, if we assume that P1s are obligatorily projecting categories, their ungrammaticality within the Ezafe domain will follow. We also predict that they will never take bare N complements, where the noun is construed as non-referential. The structure for P1s is given below:

\[
(53) \quad \text{PP} \quad \begin{array}{c}
\text{P1} \\
\text{DP}
\end{array}
\]

It turns out that PPs headed by P1s are not entirely impossible within the Ezafe domain. Interestingly there are counterexamples to this claim in Samiian's (1983) thesis. Two examples that were given above in section 2.3.3 are repeated below, with the relevant prepositional modifiers in bold face.

(54) (a) bahs-e bā ajale-ye hasan jâleb bud-Ø
discussion+EZ with haste+EZ Hasan interesting was+3sgS
'Hasan's hasty discussion was interesting.' [Samiian 1983:55.117]

(b) molâqât-e qablaz zohr-e mā jâleb bud-Ø
meeting+EZ before noon+EZ us interesting was+3sgS
'our meeting before noon was interesting' [Samiian 1983:55.118]
These prepositional modifiers are problematic both for Samiian's (1994) claim that case-assigners can't occur in case positions; and for the claim here that P1s obligatorily project to PP, while elements linked by the Ezafe morpheme are non-maximal. However, these constituents do not look like prototypical prepositional phrases. The item in (54b) is actually lexicalized just like its counterpart (ba'd az zohr 'afternoon'). The item in (54a) is best translated as an adjective ('hasty' rather than 'with haste'). I suggest then that the general claim is true, i.e. that constituents headed by P1s cannot occur within the Ezafe domain, and that the counter-examples that exist are actually compounds.

Examples of compounds formed with P1 prepositions are easy to find. Thackston (1983:137) shows examples of compounds involving bi 'without':

(55) (a) bi kâr without-work 'unemployed, idle'
     (b) bi hosele without-patience 'impatient'

He states that bâ 'with' is also a productive element for these sorts of compounds but gives no examples. However, the following examples appear in Lazard (1992:288)

(56) (a) bâ-vafâ with-fidelity 'faithful'
     (b) bi-hayâ without-shame/modesty 'impudent'
     (c) be-jâ to-place 'opportune, convenient'
     (d) bar-qarâr on-stability/establishment 'established, assured, in order'

Compounding takes place in the lexicon (see section 2.7). Thus these compounds are inserted under a single terminal node with their internal structure opaque for the purposes of the syntax.

To summarize, there are two types of prepositional elements in Persian. Most of them are [+N], i.e. have nominal properties, and they can occur with either phrasal or non-phrasal
complements. Only when they take a phrasal complement do they project to a maximal category. The second type of preposition, the so-called 'true' prepositions, always project to a phrasal category (PP) and take phrasal complements. Members of both types can also occur in compound structures in which case the entire compound is inserted under a single terminal node.

2.6.2 The Projection Properties of Adjectives

Let us now consider the contexts in which adjectives can appear. Recall that within the Ezafe domain, bare adjectives can appear as modifiers where they are clearly non-phrasal. This suggests that they are not obligatorily projecting categories. However, the same adjectives can appear with phrasal complements. In these cases the resulting constituent cannot appear in the Ezafe domain and thus seems to be phrasal. Examples are given below. In the (a) cases we see a noun modified by a bare adjective. In the (b) and (c) examples we see this adjective occurring with a complement in predicate and appositive positions respectively. In the (d) examples we see that the adjective cannot take its complement within the Ezafe domain.

(57) (a) mard-e negarân-i
man+EZ worried+indef
'a worried man'

(b) in mard negarân-e bachche-hâ-ast
this man worried+EZ child+pl+3sgS
'This man is worried about the children.'

(c) in mard, negarân-e bachche-hâ, asabâni shod
this man worried+EZ child+pl angry became+3sgS
'This man, worried about the children, got angry.'
(d) *mard-e negarân-e bachche-hâ asabârâ shod
   man+EZ worried+EZ child+pl angry became+3sgS
   *'The worried about the children man got angry.'

(58) (a) mard-e mashqul-i
   man+EZ busy+indef
   'a busy man'

(b) in mard mashqul-e âshpazi-ast
   this man busy+EZ cooking+3sgS
   'This man is busy (with) cooking.'

(c) in mard, mashqul-e âshpazi, man-o na-did-Ø
   this man busy+EZ cooking me+râ neg+saw+3sgS
   'This man, busy (with) cooking, didn't see me.'

(d) *mard-e mashqul-e âshpazi man-o na-did-Ø
   man+EZ busy+EZ cooking me+râ neg+saw+3sgS
   *'The busy (with) cooking man didn't see me.'

The structures corresponding to the observations that have just been made are given below. In (59a) we see that an adjective with a DP complement projects to AdjP. In (59b) we see that an adjective alone will not project. Adjectives with the structure in (59b) can modify nouns within the Ezafé domain.

(59) (a) AdjP
     \[\underline{\text{Adj}^0} \quad \text{DP}\]
(b) AdjP
     \[\underline{\text{Adj}^0}\]
The fact that an adjective alone does not project to AdjP explains a fact that was briefly noted in section 2.4. This concerns the difference between predicate position and appositive position. We saw above that in both positions an adjective can take a complement. However, in predicate position the adjective can also appear 'bare', while such an option is not possible for an adjective in appositive position. This is true of the English translations also.21

(60) (a) in mard negarân-e
   this man worried+3sgS
   'This man is worried.'

   (b) *in mard, negarân, asabânî shod-Ø
       this man worried angry became+3sgS
       *'This man, worried, got angry.'

(61) (a) in mard mashqul-e
   this man busy+3sgS
   'This man is busy.'

   (c) *in mard, mashqul, man-o na-did-Ø
       this man busy me+râ neg+saw+3sgS
       *'This man, busy, didn't see me.'

21Sadler and Arnold (1994) briefly discuss this fact in English. They suggest that perhaps A0s that occur postnominally in English should not be ruled out by the grammar (since speakers find it difficult to give grammaticality judgements) but that a Gricean approach is more appropriate. Specifically, they say that since [A0 N0] constructions are simpler that [N AP] constructions, "the latter should only be used if the 'extra' meaning of the N-AP construction is somehow significant." (Sadler and Arnold 1994:224, fn. 41) As an alternative I suggest that the syntactic explanation we are about to see for the facts in Persian can also hold in English.
I suggest that even when an adjective appears alone in predicate position it is dominated by an AdjP since it is the complement to the copula. Recall that non-projecting elements can appear as maximal categories if selected by a projecting element (cf. section 1.2.5). The resulting structure is given in (62a). In contrast, assuming that appositives are simply adjoined to the constituent that they modify, there is no item that selects the adjective forcing it to be phrasal. Thus the (b) sentences are ruled out because an X0 cannot adjoin to a maximal category. The illicit structure is shown in (62b).

(62) (a) VP (b) * DP
    V' DP Adj 0
    AdjP V
    Adj 0

There is one property of adjectives which may ultimately be explained when a full analysis of the projection properties of lexical categories is undertaken. Adjectives behave like verbs (see Chapter 5) in that they must take phrasal complements. For example, recall that nouns and nominal prepositions can appear with non-referential, X0, modifiers or complements as shown in (63a) and (63b) respectively.

(63) (a) kif-e pul
      bag+EZ money
      'money bag'

      (b) kenâr-e daryâ
          side+EZ sea
          'seaside'

This is not possible with adjectives as the examples in (64) show.
(64) (a) negarân-e bachche-hâ
    worried+EZ  child+pl
    'worried about the children'

    (b) *negarân-e bachche
        worried+EZ  child
        *'child worried'

While in this case the ungrammaticality of (64b) may be attributable either to syntax (adjectives cannot take $X^0$ complements) or to semantics (adjectives cannot take non-referential complements), there are cases where the referential status of the complement is not as easy to determine yet the complement must still be phrasal. This is shown below, where the phrasality of the complement to the adjective accounts for the ungrammaticality of (65b).

(65) (a) mashqul-e âshpazi bud-am
    busy+EZ  cooking was+1sgS
    'I was busy with cooking.'

    (b) *mard-e mashqul-e âshpazi
        man+EZ busy+EZ  cooking
        *'busy with cooking man'

Descriptively then, adjectives cannot appear in the following configuration.

(66) * Adj$^0$
    Adj$^0$
       N$^0$
Since this is also true of verbs, as we will see in Chapter 5, this fact may be a property of a feature that adjectives and verbs share. I leave this open for future research.

To summarize, adjectives can surface as non-projected Adj0s. They may also take complements, but their complements must be phrasal, in which case the dominating node is AdjP. Bare adjectives can occur as attributive modifiers or predicates. Adjectives with complements can appear in appositive position or as predicates.

2.6.3 The Function of Ezafe Insertion

In this section we have seen that nouns, adjectives and nominal prepositions are all potentially (and in the case of nouns, obligatorily) non-projecting. Note that these are precisely the three categories on which the Ezafe vowel appears. The truly projecting lexical categories, namely the verbs and 'true' prepositions, never occur with the Ezafe vowel. Given that (projected) phrase structure serves to identify constituent structure, a possible function for Ezafe begins to emerge. Specifically, the function of this vowel is to link non-projecting heads to the elements with which they form a constituent. That is, given a string of X0s, the Ezafe vowel identifies them as a constituent. We will return to the Ezafe vowel and the exact formulation of the Ezafe Insertion Rule in Chapter 3, section 3.6.

2.7 The Difference between Ezafe-linked Constituents and Compounds

Given that elements linked with the Ezafe vowel are X0s, we now turn to the question of whether there is a difference between two X0s adjoined to one another and a compound. Consider the following data:

(67) (a) mādār-e bozorg
    mother+EZ big
    'big mother'
In each of the examples above, we have the same lexical items in (a) and (b), however in the latter case the unit receives only one stress and crucially the Ezafe vowel does not appear. Note also that the semantic relationship between the two items is less transparent in the (b) cases. Clearly then these two modes of combining X0s are distinct from one another. I suggest that the (b) cases illustrate true compounding.

Karimi and Brame (1986) also note a difference between compounds and the Ezafe construction. They give the following contrastive pair to illustrate this difference.
They state that in the (a) case, āb portoqāl has developed into a compound since orange juice is very common. Presumably then, at one point this construction did involve the Ezafe vowel. In (b) we see an example of a non-compound. Karimi and Brame state that apple juice is not as common thus the construction has not been lexicalized yet. Lazard (1992) makes a similar point stating that "[i]n colloquial language, the ezâfe frequently is dropped in rapid speech when its omission does not make the utterance ambiguous or obscure ... This happens particularly in frequent expressions or expressions of a frequent type, notably in groups which have taken or which tend to take the value of lexical phrase ..." (Lazard (1992:67).22

Of course, compounds do not develop only from the Ezafe construction. In addition to the left-headed compounds shown above, there are also many right-headed compounds in Persian. Some examples taken from Thackston (1983:137) are given below:

(71)  mâh-ru  moon-face  'beautiful'
sang-del  stone-heart  'hardhearted'
xar-gush  donkey-ear  'rabbit'
xosh-lebâs  good-clothes  'well-dressed'
shekaste-del  broken-heart  'broken-hearted'

I propose that compounds differ from the Ezafe construction in that compounds are formed pre-syntactically and inserted into the syntax under a single terminal node. Compounds can be lexicalized juxtapositions of two distinct items, or, to the extent that there may be a productive process of compounding, formed in the lexicon. That is, compounding is a derivational process in Persian morphology. Consequently the internal structure of compounds is opaque to the syntax. On the other hand, the Ezafe construction involves base-

---

22Note that the fact that the Ezafe vowel can be dropped when no ambiguity results further supports the idea that its presence is simply to identify constituenthood.
generated head-adjoined structures that are syntactically transparent. The two different structures are given below:

\[
(72) \quad \begin{array}{ll}
(a) \text{ Two nouns in an Ezafe construction } & (b) \text{ Two nouns in a compound }\\
\begin{array}{c}
N^0_i \\
N^0_i \quad \overrightarrow{N^0} \\
N^0 \\
[\text{N N}]
\end{array}
\end{array}
\]

2.8 Conclusion and Further Implications

In this chapter we have considered the Ezafe construction within noun phrases, adjective phrases, and nominal prepositional phrases. We have seen that base-generated X^0-adjunction accounts for the constraints on elements that occur within the Ezafe domain far more elegantly than a filter-based approach. The presence of the Ezafe vowel is accounted for by a rule inserting it at PF on X^0s bearing the feature [+N] that are followed by another item. In this way Ezafe serves to identify constituent-hood. The idea of base-generated X^0-adjunction structures has led to a classification of lexical categories as either projecting or non-projecting. In this section we briefly consider headedness, and some issues surrounding this notion that are raised by the possibility of base-generated X^0-adjunction.

There are two properties implicit in the notion of the head of a phrase in classic GB theory. First, the head is the item that gives its label to a constituent. Second, the head and the constituent of which it is the head differ by one bar-level. For example, Jackendoff (1977:30) defines a head as the following: "The head of a phrase of category X^n can be defined as the X^{n-1} that it dominates." Given that here, not all lexical items project, i.e. not all X^n must be dominated by an X^{n+1}, we must redefine the notion of head simply in terms of which constituent gives its label to the dominating node.

Let us informally characterize headedness for Persian in the following way. Let the term 'left-dominant' refer to a structure in which the lefthand member gives its label to the
dominating node. Similarly the term 'right-dominant' will refer to a structure in which the righthand member gives its label to the dominating node. We can then say that in Persian the category [-N, +V] (i.e. the category 'verb') is right-dominant. All other lexical categories are left-dominant.

This characterization of headedness has interesting implications for the theory. It predicts that complements, specifiers and adjuncts should all occur on the same side. While this is true for most categories in Persian it is clearly too strong for other languages (e.g. English where complements and specifiers occur on opposite sides of the head). However, right or left dominance can easily be relativized to categories of certain bar-levels (e.g. V0 is right-dominant, V' is left-dominant). Since the whole notion of intermediate bar levels is being questioned (Speas (1990), Chomsky (1994)) I leave this open for further research.

To account for X0-adjoined structures, a characterization of headedness is not enough. The semantic properties of lexical items also play a role in the composition of X0s. For example, consider the following noun phrases:

(73) (a) xâne
    house
    'house'

(b) xâne-ye bozorg
    house+EZ big
    ‘big house’

(c) xâne-ye bozorg-e kenâr-e daryâ
    house+EZ big+EZ beside+EZ sea
    'big seaside house'

---

23The one counterexample to the general claim about headedness in Persian that has been made here is the DP headed by the null definite determiner that will be discussed in the following chapter. In this projection the head (D) occurs on the right, but D' occurs on the left of the possessor, which is in [SPEC, DP]. The distinction between specifier and adjunct is being blurred, however (see Speas (1990) and Kayne (1994) for example). In fact, if the possessor is an adjunct to DP, adjunction within the noun phrase will uniformly take place to the right. The fact that the possessor must be the closest DP adjunct can be explained by the fact that it requires case. I leave an exploration of adjunct vs. specifier for future research.
(d) *xâne-ye bozorg-e kenâr
      house+EZ big+EZ beside
*big beside house'

In (73c) we have to assume that kenâr 'beside' and daryâ 'sea' combine together and as a unit modify xâne-ye bozorg 'big house'. This is so because kenâr 'beside' alone cannot modify a noun as shown in (73d). This suggests that the structure of (73c) is the following:

(74) N_i
    / \
   /   \
N_i  P_{nom}
    / \A   \
   /   \N

In the above structure the subscripts indicate the headedness of the phrase. Note that under the analysis of the Ezafè domain developed here, wherein none of these elements are phrasal, nothing structural forces the above syntactic configuration. I believe this particular configuration follows from the semantics of nominal prepositions. That is, stated informally, nominal prepositions being inherently relational must take a complement (either phrasal or non-phrasal) before they can modify a noun.

Intensifiers, which can cooccur with adjectives in the Ezafè domain are also interesting from the point of view of headedness and semantic composition. These items precede the adjectives they modify and do not appear with the Ezafè vowel. Some examples are given below with the intensifier and the adjective it modifies given in square brackets.

(75) (a) xâne-ye [besyâr bozorg]-i
     house+EZ very  big+indef
     'a very big house'

     (b) ketâb-e [xeyli jâleb]-e jiân
         book+EZ very  interesting+EZ Jian
         'Jian's very interesting book'
The lack of Ezafe can be accounted for if we assume that the trigger for Ezafe Insertion is \([+N]\) and that intensifiers do not bear this feature. In terms of their structure, intensifiers clearly take adjectives as their complements, however, they cannot project to a phrasal category if we are to maintain the claim that all elements within the Ezafe domain are X\(^0\)s. Thus, we must posit one of the following structures:

\[
\begin{align*}
(76) \quad & (a) \quad \text{Int}^0 \quad \text{Adj}^0 \\
& \quad \text{Int}^0 \quad \text{Adj}^0
\end{align*}
\]

Again we have to appeal to the semantic properties of intensifiers to ensure that the noun phrases in which they occur have the structure in (77a) rather than (77b).

\[
\begin{align*}
(77) \quad & (a) \quad \text{N}_i \quad \text{Int}^0 \\
& \quad \text{N}_i \quad \text{Int}^0 \\
& \quad \text{N}_i \quad \text{Adj}^0
\end{align*}
\]

There is one last observation to be made here that initially seems to bear on the Ezafe construction, but I believe, ultimately has nothing to do with it. First, let us consider the data.

As we have seen, the Ezafe construction links a noun and any number of modifiers to one another. An example is given below with a noun modified by two adjectives.

\[
\begin{align*}
(78) \quad & \text{xune-ye bozorg-e qadimi} \\
& \quad \text{house+EZ big+EZ old} \\
& \quad \text{'a big old house'}
\end{align*}
\]

\[\text{24The structure in (76b) presents a problem for the generalization that all non-verbal lexical categories are left-dominant in Persian.}\]

\[\text{25Thanks to Elizabeth Cowper for useful discussion on this point.}\]
According the analysis developed in this chapter the noun and both adjectives in (78) are X0s, satisfying one of the conditions on the insertion of the Ezafe vowel. We have also seen that the Ezafe vowel appears between an Adj0 and a following DP. Now consider the following examples.

(79) (a) in xune bozorg-e
        this house big+3sg
        'This house is big.'

        (b) in xune bozorg-o qadimi-e
            this house big+and old+3sgS
            'This house is big and old.'

        (c) *in xune bozorg-e qadimi-e
            this house big+EZ old+3sgS
            *'This house is big old.'

We see in (79c) that two adjectives cannot be joined with Ezafe when they occur in predicate position. Yet they seem to meet the structural description for Ezafe Insertion. However, note that the same facts hold in English. We can say a big old house and This house is big but not *This house is big old.

I propose that this can be accounted for by a constraint that adjectives cannot modify adjectives. Thus, two adjectives modifying a noun are permitted since each occurs with an N0 sister as shown in (80a) below. Two conjoined adjectives are also permitted within a noun phrase as shown in (80b). However, an adjective cannot modify another adjective as shown in (80c).
In predicate position where the adjective occurs as sister to a verb and thus is dominated by AdjP (see section 2.6.2), a bare adjective can appear (see (81a) and two conjoined adjectives are also grammatical (see (81b)). However, two adjectives linked by Ezafe are ruled out (see (81c)) since this would involve the illicit structure shown above in (80c).

This concludes our discussion of the Ezafe construction and the portion of the noun phrase that involves Ezafe. The following chapter deals with the structure of the noun phrase outside the Ezafe domain in order to determine the position of the possessive noun phrase. At the end of that chapter a precise formulation of the Ezafe Insertion Rule is given.
3.0 Introduction

Chapter 2 dealt with the Ezafe construction in Persian. We saw that within noun phrases, the Ezafe vowel occurs on all modifiers of the head noun, and that these modifiers must be X₀s. We saw that any number of X₀ elements can be adjoined to one another, subject to semantic ordering constraints.

Possessors are also linked by the Ezafe construction as the following example shows.

(1) ketâb-e bozorg-e jiân
    book+EZ big+EZ Jian
    'Jian's big book'

Possessors contrast with all other modifiers in the noun phrase in that they can be fully phrasal. While they trigger Ezafe Insertion, they also mark the end of the Ezafe domain. That is, neither a possessor nor any following material will be followed by the Ezafe vowel. Assuming that possessors are always DPs, the following structure emerges for the Ezafe domain:

(2) \[
\begin{array}{c}
\text{XP}_1 \\
\text{N}_1^0 \\
\text{DP}_{\text{poss}} \\
\text{P}_{\text{nom}} \\
\text{A} \\
\text{N}_0 \\
\text{N}_0
\end{array}
\]
As before, subscripts indicate headedness. X^0 elements are adjoined to the head. The possessor is shown as a DP also adjoined to the head but in this case the dominating node is XP.\footnote{The abbreviation DP_{poss} will be used interchangeably with 'possessor' throughout the rest of the chapter.} If Ezafe Insertion is triggered by X^0s, the phrasal nature of the possessor predicts that Ezafe insertion will stop here. However, this preliminary structure raises a number of questions. Is the possessor base-generated to the right of the noun phrase or has it moved from a position inside the noun phrase? Is the dominating node NP or some other category? These and other questions are discussed in this chapter.

The chapter is organized as follows. In section 3.1 Samiian's (1983) analysis of the position of possessors is discussed. The section includes some of her arguments for constituenthood that are relevant to the following sections. In section 3.2 further properties of possessors are presented. In section 3.3 it is argued that the constituent including the head noun, its modifiers, and the possessor is a DP. It is shown, for example, that a possessor is in complementary distribution with the indefinite enclitic -i which projects a DP.

In section 3.4 the structural position of the possessor is discussed. Various hypotheses are compared and contrasted, including some based on analyses of similar constructions in other languages. Of these hypotheses, we see that the base-generation of the possessor in the right specifier of DP best accounts for the facts. In section 3.5 we take a more general look at noun phrases in Persian to see what else projects to DP. The set of items includes proper names, pronouns, the indefinite enclitic, and the null definite determiner. In this section two quite puzzling morphemes, the plural marker -hâ and the colloquial marker -e, are also discussed and shown to have quite a simple and natural account under the analysis so far developed.

Section 3.6 presents a reformulation of the Ezafe Insertion Rule of Samiian (1983, 1994) in keeping with the structure of the noun phrase that has been developed. Section 3.7 contains some concluding remarks. Section 3.8 is an appendix in which a description is
given of other material that can occur in the noun phrase, which has not been dealt with in this and the preceding chapters. This appendix is included in order to provide a more complete catalogue of nominal material in Persian. Note that there is nothing in the appendix that bears on the analysis being developed for the Ezafé domain and the position of possessors.

3.1 Samiian's Analysis of DP\text{poss} under N'

As we saw in Chapter 2, section 2.2, Samiian (1983) accounts for the presence of the Ezafé vowel with an Ezafé Insertion Rule. She assumes that all elements linked by Ezafé are generated under N'. The Ezafé Insertion Rule inserts the vowel -e before maximal projections dominated by N'. While Samiian generally shows a flat structure for elements dominated by N', in Chapter 3 she presents arguments for the structure in (3).

(3) \[ \text{N'} \]
\[ \text{N PP NP} \]
\[ \text{N NP AP} \]

[Samian 1983:96.16]

In the following two sections we review Samiian's arguments that (a) the possessor is at a different level from the head noun and its modifiers; and (b) the head noun, its modifiers, and the possessor form a single constituent. We also see that the PP in (3) patterns with the modifiers of the noun rather than the possessor.

3.1.1 An argument that DP\text{poss} is separate from N and its modifiers

Samiian describes an anaphoric process in Persian that seems to treat possessors as separate constituents from the elements preceding them. This process involves the word
mâl-e which she translates as 'the property of'. mâl-e can be followed by a DP<sub>poss</sub>, shown in (4). However, it cannot be followed by an adjective or attributive noun, shown in (5) and (6) respectively.

(4) ketâb-e inglisi-e ali-râ mi-xâh-i yâ mâl-e hasan-râ
book+EZ English+EZ Ali+râ cont+want+2sgS or the one+EZ Hasan+râ
'Do you want Ali's English book or Hasan's.' [Samiiian 1983:89.18]

(5) *ketâb-e inglisi-e ali-râ mi-xâh-i yâ mâl-e
book+EZ English+EZ Ali+râ cont+want+2sgS or the one+EZ
farânse-ye hasan-râ
French+EZ Hasan+râ
*‘Do you want Ali's English book or Hasan's French.' [Samiiian 1983:89.19]

(6) *dastkesh-e charm-e hasan-râ mi-xâh-i yâ mâl-e
glove+EZ leather+EZ Hasan+râ cont+want+2sgS or one+EZ
pashm-e ali-râ
wool+EZ Ali+râ
*‘Do you want Hasan's leather gloves or Ali's woolen.' [Samiiian 1983:89.20]

Samiiian concludes from this that the attributive noun phrase and adjective phrase must form a constituent with the head noun excluding the possessor. She does not mention the prepositional phrase, but from the structure shown in (3) above, she apparently believes that it patterns with the possessor. In fact it does not. Prepositional modifiers do not pattern with possessors, but behave like attributive adjectives and nouns with respect to mâl-e. Consider the following data:2

2Note that the second sentence is slightly better than the first. This may be attributable to the fact that in the second sentence the prepositional modifier is a compound and therefore...
(7)  (a)  *xâne-yê kenâr-e daryâ-e ali-râ did-i yâ mâl-e house+EZ next-to+EZ beach+EZ Ali+râ see+2sgS or the one+EZ kenâr-e rutxâne-yê hasan-râ next-to+EZ river+EZ Hasan+râ.  
*'Did you see Ali's seaside house or Hasan's riverside.'

(b)  ?molâqât-e qabl azzohr-e mâ-râ dust dasht-i yâ mâl-e meeting+EZ before-noon+EZ us+râ liked+2sgS or the one+EZ ba'daz zohr-e ânhâ-râ afternoon+EZ them+râ 
*'Did you like our before noon meeting or their afternoon.'

If we include the prepositional modifier with the attributive noun modifier and the adjectival modifier in Samiian's structure in (3) above, we have a structure almost identical to that shown in (2). It is interesting to note that Samiian has N rather than N' as the dominating node of the sister to the possessor. She rejects N' because of the predictions such a structure would make about recursion and also brings in data from the definiteness marker -e and the indefinite marker -i to support the idea that this node is simply an N0. (These markers will be discussed in 3.5.) Since all constituents beneath this node are non-phrasal we have a structure such as (2), repeated below. mâl-e substitutes for the highest N0.4

4It is not clear what rules out mâl- substituting for a lower N0. As suggested by Diane Massam (personal communication) this may be attributable to the A-over-A Principle (Chomsky (1964) although it was Ross (1967) who named it as such).
There are two facts that an eventual analysis of *mâl-e* must account for. First, we have seen that *mâl-e* is a form that must appear with a possessor. While in most contexts a pronominal possessor is interchangeable with a pronominal enclitic, (see (8) below), a pronominal enclitic is not possible with *mâl-e* (see (9b)). In Chapter 4 we see that the pronominal enclitics serve to identify a pro possessor. This suggests that *mâl-e* must appear with a phonetically realized possessor.

(8) (a) ketâb-e man
book+EZ me
'my book'

(b) ketâb-am
book+1sg
'my book'

(9) (a) mâl-e man
property+EZ me
'mine'

(b) *mâl-am
A second point of interest is that mâl-e surfaces in precisely the places where a null head is possible in English.

(10) (a) This book is John's [ø]

(b) in ketāb mâl-e jiān-e
   this book property+EZ Jian+3sgS
   'This book is Jian's.'

(11) (a) John's book is more interesting than Bill's [ø]

(b) ketāb-e jiān az mâl-e hasan jāleb-tar-e
   book+EZ Jian from property+EZ Hasan interesting+comp+3sgS
   'Jian's book is more interesting than Hasan's.'

(12) (a) I like Bill's shirt better than John's [ø]

(b) pirhan-e jiān-o az mâl-e hasan bishtar dust dār-am
   shirt+EZ Jian+rå from property+EZ Hasan better like+1sgS
   'I like Jian's shirt better than Hasan's.'

The fact that null heads are not possible in Persian might explain the existence of the form mâl. I put these two points aside here. The property of mâl-e that is relevant here is that it substitutes for all elements up to the possessor, suggesting that these elements form a single constituent.
3.1.2 An argument that $\text{DP}_{\text{poss}}$ is a constituent with $N$ and its modifiers

Samiian provides several arguments that the possessor forms a constituent with the head noun and its modifiers. She proposes that this constituent is dominated by $N'$ (as opposed to $N''$). I will focus on one of her arguments here and briefly mention the others.

Samiian's main argument is based on an anaphoric element similar to 'the one' in English. She shows that $\text{un yeki}$ 'that one' can substitute for the head noun and all following elements, including a possessor. If $\text{un yeki}$ is followed by a possessor the resulting sentence is ungrammatical. The relevant sentences are given below. In (13a) and (14a) we see that $\text{un yeki}$ cannot be followed by an adjective or a possessor.

(13) (a) *in ketâb-e inglisi-râ xund-am vali
this book+EZ english+râ read+1sgS but

$\text{un yeki-e farsi-râ}$ na-xund-am
that one+EZ Persian+râ neg+read+1sgS
'I read this English book but didn't read that Persian one.'

[Samiian 1983:92.24a]

(b) in ketâb-e inglisi-râ xund-am vali $\text{un yeki-râ}$ na-xund-am
this book+EZ english+râ read+1sgS but that one+râ neg+read+1sgS
'I read this English book but didn't read that one.'

[Samiian 1983:93.24b]

(14) (a) *in ketâb-e ali-râ xund-am vali
this book+EZ Ali+râ read+1sgS but

$\text{un yeki-e hasan-râ}$ na-xund-am
that one+EZ Hasan+râ neg+read+1sgS
'I read Ali's book but not Hasan's.'

[Samiian 1983:93.25a]
(b) in ketáb-e ali-râ xund-am vali un yeki-râ na-xund-am
this book+EZ Ali+râ read+1sgS but that one+râ neg+read+1sgS
'I read this book of Ali's but not that one.' [Samiian 1983:93.25b]

Samiian also shows that elements that follow DP$_{poss}$ (which she calls N'' complements) can follow un yeki without any difficulty.

(15) (a) bahs-e hasan-râ râje be in mozu shenid-am vali
discussion+EZ Hasan+râ about this issue heard+1sgS but

un yeki-râ râje be jang na-shnid-am
that one+râ about war neg+heard+1sgS
'I heard Hasan's discussion about this issue but didn't hear the one
about the war.' [Samiian 1983:93.26]

(b) bahs-e hasan râje be in mozu-râ shenid-am vali
discussion+EZHasan about this issue+râ heard+1sgS but

un yeki râje be jang-râ na-shnid-am
that one about war+râ neg+heard+1sgS
'I heard Hasan's discussion about this issue but didn't hear the one
about the war.'

In the examples above, the fact that the direct object marker -râ can occur before the prepositional constituent (15a) suggests that this constituent is actually adjoined to the noun phrase. The example in (15b) is given to show that -râ can also appear after the whole noun phrase.$^5$

Taking the structure proposed in (2) above, what we see is that un yeki can substitute for a phrasal category (XP$_1$, below) but not for anything below this node. (An explanation for

$^5$There is some preference for -râ to appear before the adjoined phrase in these cases. I will leave aside an exploration of why this is the case.
Phrasal categories outside the Ezafe domain appear to be adjoined to XP, judging from the fact that -râ can appear before or after YP.

(16)

\[
\begin{array}{c}
\text{XP}_2 \\
\text{XP}_1 \\
\text{YP} \\
\hline
\text{N}_1^0 \\
\text{DP}_{\text{poss}} \\
\hline
\text{N}_1^0 \\
\text{A}
\end{array}
\]

Samiian also applies other tests (taken from Jackendoff (1977)) to support her claim that the possessor is an N' complement, rather than an N'' complement. These tests include question formation, relativization and the scope of quantifiers. I will not discuss them here for two reasons. First, the arguments, especially those based on question formation and relativization, are very brief and somewhat incomplete. Second, it is not clear that the N' vs. N'' distinction is being tested, or even that this distinction is relevant if we assume that all modifiers between the head noun and the possessor are X0s.

What is important here is that un yeki cannot substitute for the noun phrase excluding DP_{poss}. An explanation for this fact is given in section 3.5.2. We now turn to consider some further properties of possessors in Persian.

3.2 Further Properties of DP_{poss}

3.2.1 Only one Syntactic Position

Only one possessive DP can appear in a Persian noun phrase. While this is true in other languages as well (compare the prenominal genitive in English), in Persian the possessor is the only phrasal constituent in the noun phrase. This explains why there is no Persian equivalent to the English the enemy's destruction of the city, since in that example both a
phrasal possessor (subject) and a phrasal object (complement) occur. The two paradigms below illustrate this.

(17) (a) taxrib-e shahr
destruction+EZ city
' the city's destruction'

(b) taxrib-e doshman-hâ
destruction+EZ enemy+pl
' the enemy's destruction' 6

(c) *taxrib-e shahr-e doshman
destruction+EZ city+EZ enemy
' the enemy's destruction of the city' 7 [Samiian 1983:92.41]

(18) (a) xordan-e sib
eating+EZ apple
' the eating of an apple'

(b) xordan-e jiân
eating+EZ Jian
' Jian's eating'

(c) *xordan-e sib-e jiân
eating+EZ apple+EZ Jian
' Jian's eating of an apple'
(Ok as 'the eating of Jian's apple')

6 This can also be construed as the 'destruction of the enemy' in which case it is ungrammatical since taxrib can't apply to animate objects.
7 This is grammatical if it is understood as 'the destruction of the enemy's city'. Note (as pointed out by Diane Massam) that a grammatical reading is also predicted if shahr 'city' can be interpreted as an attributive (i.e. non-referential) noun, if there are different types of destructions, of which city-destruction was one sub-type. This has been confirmed by speakers I have consulted with.
This fact about Persian would be quite mysterious if we were to assume that nouns project structure and that the Ezafe vowel is a case assigner - both commonly held assumptions in the literature on Persian (cf. Chapter 2, section 2.5.2). On the other hand, the facts are as expected if there is only one case position in which a phrasal argument of a noun can occur. Crucially, note that there is no structure equivalent to the of-phrase, found in English. If phrases of this type are generated in NP, their absence in Persian follows from the fact that nouns in Persian do not project structure.

3.2.2 The Semantics of DP_{poss}

As the possessor position is the only phrasal position within the Persian noun phrase, it should come as no surprise that this position can freely contain a DP bearing any sort of relationship to the head noun (e.g. theme, goal, agent, etc.). Thus as we saw in the last section the possessor of a deverbal noun such as xordan 'eating', can be either the agent or the patient of xordan. Samiian (1994) cites Philott (1919) as stating "that there is no proper genitive in Persian and that the possessive is expressed by the occurrence of Ezafe between two nouns" (Samiian 1994, fn. 2, p. 18). Further examples are given below showing possessors in the Ezafe construction that cannot appear in the prenominal genitive position in English.

(19) (a) dâñesh-e jabr
knowledge+EZ algebra
'knowledge of algebra' (*'algebra's knowledge')
3.3 Evidence that DP\textsubscript{poss} involves projection to DP

The structure for the Ezafe domain was given in (2) and is repeated below:

\[(2) \quad \text{XP} \quad \begin{array}{c} N_1^0 \text{DP}_{\text{poss}} \\ N_1^0 \text{P}_{\text{nom}} \\ N_1^0 \text{A} \\ N_1^0 \text{N}^0 \end{array} \]

If (2) is a base-generated structure, assuming X-bar principles, the dominating node (XP) should be NP with DP\textsubscript{poss} in its (right-hand) specifier. In this section, however, we see that this projection is a DP rather than an NP.

3.3.1 Semantic Evidence that DP\textsubscript{poss} forms a DP with the head noun

The first thing to note about noun phrases that include a possessor is that they are obligatorily marked with -\textit{râ} in direct object position. Normally, an object noun can occur
with -râ, with the indefinite enclitic, or be bare (generic). However, a noun phrase with a possessor can only occur with -râ. This is shown below:

(20)  
(a)  ketâb-o  xund-am  
book+râ  read+1sgS  
'I read the book.'

(b)  ketâb-i  xund-am  
book+indef  read+1sgS  
'I read a book.'

(c)  ketâb  xund-am  
book  read+1sgS  
'I read books.'

(21)  
(a)  ketâb-e  jiân-o  xund-am  
book+EZ  Jian+râ  read+1sgS  
'I read Jian's book.'

(b)  *ketâb-e  jiân-i  xund-am  
book+EZ  Jian+indef  read+1sgS

(c)  *ketâb-e  jiân  xund-am  
book+EZ  Jian  read+1sgS

If we take -râ as a marker of definite or presupposed information (see Chapter 5, section 5.2.2), then noun phrases that include a possessor are obligatorily construed as definite/presupposed. This suggests that they are DP s.

While the above argument is primarily semantic, there is distributional evidence that possessors form a DP with the nouns that they occur with to which we now turn.
3.3.2 Distributional Evidence that DP_{poss} forms a DP with the head noun

In this section we see that possessors are in complementary distribution with the indefinite enclitic -i. First, let us briefly consider the properties of this enclitic.

The indefinite enclitic can occur on a subject, the object of a verb, or the object of a preposition. It has been characterized as meaning 'a certain, a particular' thing, or 'one of a class' (see Thackston (1983), for example). Unlike the indefinite _a/an_ in English, it cooccurs with plural marking. This is shown below.

\[
\begin{align*}
(22) & \quad \text{mard-i âmad-Ø} \\
& \quad \text{man+indef came+3sgS} \\
& \quad \text{'}A man came.' \\
(23) & \quad \text{mard-i did-am} \\
& \quad \text{man+indef saw+1sgS} \\
& \quad \text{'}I saw a man'} \\
(24) & \quad \text{ketâb-o be pesar-i dâd-am} \\
& \quad \text{book+râ to boy+indef gave+1sgS} \\
& \quad \text{'}I gave the book to a boy.'} \\
(25) & \quad \text{ketâb-hâ-i xarid-am} \\
& \quad \text{book+pl+indef bought+1sgS} \\
& \quad \text{'}I bought some (certain) books.'}
\]

The enclitic -i also occurs on nouns modified by restrictive relative clauses, giving the contrast in (26).

\[
\begin{align*}
(26) & \quad \text{(a) ahmad, ke diruz âmad-Ø, injâ-st} \\
& \quad \text{Ahmad that yesterday came+3sgS here+3sgS} \\
& \quad \text{'}Ahmad, who came yesterday, is here.'}
\]
Ahmad+indef+that yesterday came+3sgS here+3sgS
'The Ahmad who came yesterday is here.'
(as opposed to the one who came today, e.g.) [Thackston 1983:82]

In traditional grammars two morphemes -i are posited to account for the two uses shown above. However, as was discussed in Chapter 1, section 1.3.3, there is a semantic characterization of this morpheme that unites its two uses. In fact this is a desirable goal given the assumption of Monosemy (see section 1.2.1).8

If -i is an indefiniteness marker that marks its noun phrase as new to the discourse (cf. Chapter 5, section 5.2.2), it would be reasonable to treat it as a determiner. A potential problem for this view is that the suffix -i is usually in complementary distribution with numerals, suggesting that it may be of the same category as they are.

(27) (a) se-tâ ketâb xarid-am
three+unit book bought+1sgS
'I bought three books.'

(b) se-tâ ketâb-râ xarid-am
three+unit book+râ bought+1sgS
'I bought the three books.'

(c) *se-tâ ketâb-i xarid-am
three+unit book+indef bought+1sgS

There are three reasons for not considering -i to be of the same category as numerals. First, numerals cannot ordinarily occur with plural marking, while -i can. Second, numerals

8Samiian (1983) also states that she sees no reason for positing two morphemes -i, one for the indefinite usage and one for the relative clause usage.
can freely cooccur with possessors, while -i cannot. Third, when-i marks a noun followed by a restrictive relative clause it can cooccur with numerals. Examples of -i cooccurring with a numeral and with a demonstrative are given below.9

(28) (a) se-tâ ketâb-i ke diruz xarid-am injâ-and 
three+unit book+indef that yesterday bought+1sgS here+3plS
'The three books that I bought yesterday are here.'

(b) un ketâb-i ke diruz xarid-am injâ-st 
that book+indef that yesterday bought+1sgS here+3sgS
'That book that I bought yesterday is here.'

We have seen that -i is in complementary distribution with DP\textsubscript{poss}. Crucially however, even in the context of a restrictive relative clause, when -i can cooccur with numerals and demonstratives, it cannot cooccur with a possessor.

(29) *ketâb-e jiân-i ke diruz xarid-am injâ-st 
book+EZ Jian+indef that yesterday bought+1sgS here+3sgS

Let us assume then that -i is a determiner, a D that projects a DP. We have seen that it is in complementary distribution with possessors. Further, we have seen that possessors impose a definite/specific reading on the possessed noun phrase. Therefore, it appears that possessed noun phrases are DPs.

---

9These two examples may actually involve different structures. Recall from Chapter 1, section 1.3.3 that -i gives a set reading of the noun it attaches to when this noun is followed by a relative clause. In (28a) this means that there is a set consisting of sets of three books. It is among this set of sets of three books that a unique referent is picked out (i.e. the set of three books that I bought yesterday). In (28b), however, it is possible that the demonstrative un 'that' takes the noun and the relative clause as its complement, i.e. [that [book that I bought yesterday]].
3.4 Structural Position of \( \text{DP}_{\text{poss}} \)

Let us consider again the structure in (2):

(2)

\[
\begin{array}{c}
\text{XP}_i \\
\quad \text{N}_i^0 \\
& \quad \text{DP}_{\text{poss}} \\
\quad \text{N}_i^0 \\
& \quad \text{P}_{\text{nom}} \\
\quad \text{N}_i^0 \\
& \quad \text{A} \\
\quad \text{N}_i^0 \\
& \quad \text{N}^0
\end{array}
\]

If we assume, based on the preceding section, that \( \text{XP}_i \) in (2) is DP, then by X-bar principles there must be a D in the structure. This D-head will in turn have as its sister an NP, since an XP cannot dominate two heads (see Chapter 1, section 1.2.5).

Since this D is null, it is not immediately obvious whether it is to the left or the right of its sister (NP). As we will see in the following sections, one potential analysis of the position of \( \text{DP}_{\text{poss}} \) forces us to assume that it is on the left. However, in the interests of symmetry, we might argue that since this null head is in complementary distribution with (and thus of the same category as) the indefinite enclitic \(-i\), and since \(-i\) is analyzed as a D-head occurring on the right, the null D-head involved in licensing \( \text{DP}_{\text{poss}} \) should also occur on the right.

Given both an NP and a DP projection there are now two places that \( \text{DP}_{\text{poss}} \) could originate: [SPEC, NP] and [SPEC, DP]. In the analysis of Persian noun phrases developed thus far, a DP should not occur within NP, however, let us investigate this possibility further to see whether it yields a favourable result.
3.4.1 DP_{poss} in [SPEC, NP]

There are four logical possibilities for the structure of the noun phrase with DP_{poss} in [SPEC, NP]. These are shown below:

(30) (a) \[
\begin{array}{c}
\text{NP} \\
\text{DP}_{\text{poss}} \\
\text{D} \\
\text{N}
\end{array}
\]

(b) \[
\begin{array}{c}
\text{NP} \\
\text{DP}_{\text{poss}} \\
\text{D} \\
\text{N}
\end{array}
\]

(c) \[
\begin{array}{c}
\text{DP} \\
\text{D} \\
\text{NP} \\
\text{DP}_{\text{poss}} \\
\text{N}
\end{array}
\]

(d) \[
\begin{array}{c}
\text{DP} \\
\text{D} \\
\text{NP} \\
\text{N} \\
\text{DP}_{\text{poss}}
\end{array}
\]

The structures in (30a) and (30b) have the empty D-head on the right, while those in (30c) and (30d) have the empty D-head on the left. In (30a) and (30c) the specifier containing DP_{poss} is on the left, while in (30b) and (30d) it is on the right. While (30b) and (30d) give the right surface order, these structures are less desirable from a theoretical point of view.\(^{10}\)

Looking at (30a) and (30c) first then, the question is whether the right surface order can be derived. In fact for (30c) there is a way to derive the correct surface word order using an analysis currently assumed for noun phrases in Hebrew.

Let us consider the analysis of the Hebrew Construct State proposed in Ritter (1991, 1992). There are several similarities between this construction and the Ezafe construction in Persian, notably the fact that the possessor follows the head noun and that the presence of a possessor gives a definite reading for the head noun. Ritter's structure is given below:

\(^{10}\)Kayne (1994) argues that specifiers can never appear on the right.
The essence of Ritter's analysis is the following: possessors are generated in [SPEC, NP] on the left of the head. However, the whole NP is embedded within a DP projection headed by a null determiner element. This element assigns case to DP\textsubscript{poss} in [SPEC, NP]. The head noun gets its definiteness feature through spec-head agreement with DP\textsubscript{poss}, then moves up to D\textsubscript{gen} to identify that functional head. This head movement accounts for the surface word order.

Clearly this analysis has great appeal, given that it explains how DP\textsubscript{poss} gets case, how the head of the DP comes to be construed as definite, and how the surface word order is derived. Specifically, in both Hebrew and Persian it is possible to maintain that specifiers are on the left while deriving the order possessed noun - possessor. Note, however, that it is crucial to this analysis that the D-head occur on the left also. That is, in a structure such as (30a) where the D-head occurs on the right, movement of the head noun to D will not derive the desired order of the possessed noun followed by the possessor.

\[\text{(31)}\]

\[
\begin{array}{c}
\text{DP} \\
\downarrow \\
\text{D'} \\
\downarrow \\
\text{D}\textsubscript{gen} \\
\downarrow \\
\text{NP} \\
\downarrow \\
\text{DP}\textsubscript{poss} \\
\downarrow \\
\text{N'} \\
\downarrow \\
\text{N} \\
\downarrow \\
\text{XP}
\end{array}
\]

\[\text{Her analysis is actually more complex once Num(ber) Phrase is introduced between DP and NP. However, the additional details are irrelevant for the point being made here. NumP is the projection corresponding to singular vs. plural marking. It is not clear, therefore, whether this projection is the relevant one for overt numerals. Consequently overt numerals will be represented as heading the projection NmrP in the text below.}\]

\[\text{Carstens (1991) drawing on work by Ritter makes a similar proposal for Swahili, in which the surface word order with the possessed noun preceding the possessor is derived by head movement to a higher functional projection.}\]
Carrying this type of analysis over to Persian raises several problems. For example, note that the structure proposed for Hebrew involves both a specifier (subject) and a complement (object) position within the NP. However, as we saw in section 3.2.1 above, nouns cannot have both a subject and an object in Persian. Only one phrasal argument is allowed within the noun phrase. Thus while in Hebrew Dan's eating of the apple is grammatical, in Persian we can only get the eating of Dan, or the eating of the apple.

Another problem in using the analysis of Hebrew outlined above for Persian, is that the wrong order of constituents is predicted once numerals are introduced. There is evidence from semantic scope that numerals occur below the DP projection. Specifically, where a noun phrase involves both a numeral and a possessor, the possessor has scope over the numeral. This is evident if we compare the following two phrases:

(32) (a) se-tâ ketâb-e hasan
    three+unit book+EZ Hasan
    'Hasan's three books'

    (b) se-tâ az ketâb-hâ-ye hasan
    three+unit of book+pl+EZ Hasan
    'three of Hasan's books'

In (32a) Hasan appears to have scope over three books, whereas in (32b) Hasan's books is embedded within a prepositional phrase and the numeral now has scope over it. Consequently, in (32a) three books has a definite reading (the phrase suggests that Hasan has only three books), whereas in (32b), the partitive reading, three is construed as indefinite. This suggests that the possessor is higher than the numeral in the non-partitive construction.

If we assume that numerals are generated below the DP projection (represented as Nmr below) then we predict that a noun phrase with both a numeral and a possessor will have the
order: Noun - Numeral - Possessor (illustrated in the diagram in (33)). This prediction is not borne out in Persian as the example in (34) shows.

(33) DP
   D'
   D_gen NmrP
     Nmr NP
       DP_pos N'
         d N

(34) (a) se-tâ ketâb-e jiân
      three+unit book+EZ Jian
     'Jian's three books'

   (b) *ketâb-e se-tâ jiân
      book+EZ three+unit Jian

Perhaps the most compelling evidence, however, for the structural relationship between the possessor and the possessed noun comes from the reflexive/emphatic element xod, the binding of which can be used as a test for c-command. The reflexive and emphatic uses of this morpheme have generally been treated separately (see Moyne (1971), Moyne and Carden (1974) and Hashemipour (1989)). An analysis of xod is given in Chapter 4. For our purposes now, it is important to note that a unified analysis can be provided for this morpheme in both its reflexive and emphatic uses.

In colloquial speech reflexives/emphatics always appear with the possessive enclitics attached. However, it is also possible to get xod on its own, meaning 'self'. The use of xod
on its own occurs in more formal or literary discourse. As we would expect, this element seems to have to be bound.

(35) (a) jiânì xodì-râ did-Ø
    Jian self+râ saw+3sgS
    'Jian saw himself.'

(b) *xodì jianì-râ did-Ø
    self Jian+râ saw+3sgS

Let us consider cases where xod is a possessor in direct object position. In such cases the presence of xod indicates emphasis. As a possessor xod is coindexed with the subject, just as it was when it appeared alone in (35a). However, if there is another possessor in the noun phrase then xod will be coindexed with that possessor. This is shown below:

(36) (a) jiânì barâdar-e xodì-râ did-Ø
    Jian brother+EZ self+râ saw+3sgS
    'Jian saw his (own) brother.'

(b) jiânì barâdar-e xod*i*j-e hasanj-râ did-Ø
    Jian brother+EZ self+EZ Hasan+râ saw+3sgS
    'Jian saw Hasan's (own) brother.' (as opposed to Hasan's wife's brother)

(c) jiânì xod*i*j-e hasanj-râ did-Ø
    Jian self+EZ Hasan+râ saw+3sgS
    'Jian saw Hasan, himself.' (lit. 'Hasan’s self', as opposed to a picture of Hasan)

This suggests that Hasan in (36b) and (36c) e-commands the reflexive/emphatic element xod.

Let us return to the analysis of the noun phrase in which the head noun moves up to a higher functional head from within the NP. If we take a noun phrase such as the direct object
in (36c), we see that the required c-command relation does not appear in the derived configuration (shown in (37) below). That is, in a noun phrase like *xod-e hasan 'Hasan's self', the possessor cannot bind the possessed noun.

\[
\begin{align*}
&\text{(37)} \\
&\text{DP} \\
&\downarrow \\
&D' \\
&\downarrow \\
&D_{\text{gen}} \\
&\downarrow \\
&\text{NP} \\
&\downarrow \\
&\text{DP}_{\text{poss}} \\
&\downarrow \\
&\text{hasan} \\
&\downarrow \\
&xod
\end{align*}
\]

A possible solution would be to have the binding take place prior to movement, when the possessed noun is c-commanded by the possessor. However, if such coindexing does take place we then have a violation of the *i-within-i* Condition (Chomsky 1981:212.73) which states (informally) that a constituent cannot bear the same index as the constituent within which it is contained.

\[
\begin{align*}
&\text{(38)} \\
&\ast \\
&\text{NP}_i \\
&\downarrow \\
&\text{DP}_i \\
&\downarrow \\
&\text{N}' \\
&\downarrow \\
&\text{N}_i
\end{align*}
\]

If we return now to the four structures we saw in (30) above, we can see that the problem posed by the binding of the reflexive/emphatic element *xod* exists for all the structures. In all four cases coindexing of *xod* with a possessor will give a violation of the *i-within-i*
Condition. We must therefore consider the alternative, namely the option of generating $\text{DP}_{\text{poss}}$ in the specifier of a higher projection.

3.4.2 $\text{DP}_{\text{poss}}$ in [SPEC, DP]

There are four logical possibilities for the structure of the noun phrase with $\text{DP}_{\text{poss}}$ in [SPEC, DP]. These are shown below:

(39)  

(a) \[
\begin{array}{c}
\text{DP} \\
\text{DP}_{\text{poss}} \\
\text{D'} \\
\text{NP} \\
\text{D}
\end{array}
\]

(b) \[
\begin{array}{c}
\text{DP} \\
\text{D'} \\
\text{NP} \\
\text{D}
\end{array}
\]

(c) \[
\begin{array}{c}
\text{DP} \\
\text{DP}_{\text{poss}} \\
\text{D'} \\
\text{D} \\
\text{NP}
\end{array}
\]

(d) \[
\begin{array}{c}
\text{DP} \\
\text{D'} \\
\text{D} \\
\text{NP}
\end{array}
\]

The structures in (39a) and (39b) show the empty D on the right, while those in (39c) and (39d) show the empty D on the left. In (39a) and (39c) the specifier containing $\text{DP}_{\text{poss}}$ is on the left, while in (39b) and (39d) it is on the right. It is immediately obvious that there is no way (without further stipulation) to derive the correct word order if either (39a) or (39c) is adopted. That is, these structures show the possessor preceding the possessed noun while in Persian the possessor must follow the possessed noun. Let us then consider (39b) and (39d), leaving aside for now the question of the right or left headedness of DP.

In the previous section we saw that a possessor cannot originate within the specifier of NP. There is also theoretical motivation for assuming that the possessor in Persian is in the specifier of DP. The argument here is based on the distinction drawn by some linguists (Higginbotham (1983), Stowell (1989, 1991), Massam (1993)) between possessors that
receive a theta-role from the head noun and those that do not. Theta-marked possessors are assumed to be structurally lower than those not theta-marked by the possessed noun. Of course, the question of whether there are possessors that do not receive theta-roles is a matter of debate. Massam (1993) argues that the structural relation between the prenominal genitive and the head noun in English can encode three different relations. The first is a theta-relation (Hilary's destruction of the city), the second an adjunct-argument relation (Hilary's sister), and the third is a looser relation which she characterizes as one of modification (Hilary's book).

Following Massam (1993) I will assume there is (at least) a distinction between possessors that receive a theta-role from a head noun and those that do not. I will refer to these two types of possessors as thematically and non-thematically selected, respectively. The question then arises whether the difference between these two types of possessors is syntactically realized. Let us briefly consider the proposals that have been made that thematically selected possessors are structurally lower than non-thematically selected possessors.

In Higginbotham (1983), for example, we find that thematically selected arguments are generated under N'. The subject position (sister to N') is where a non-thematically selected argument can be generated. Higginbotham further proposes that only if a thematic role is not realized under N' is it available for the argument in subject position. This gives the contrast illustrated in (40). In (40a) Mary can't be the goal because that role is assigned to John, but in (40b) Mary can be the goal.

(40) (a) Mary's gift to John
(b) Mary's gift

Giorgi and Longobardi (1991) use binding facts in Italian to build on Higginbotham's idea and they propose a layered structure for NP. This structure is given below in (41). They
claim that \( N' \) is the domain of strictly subcategorized arguments, and that both agents and possessors are excluded from this domain. Possessors, by which they mean genitive elements "semantically connected with the head in a looser way than through a specific [theta]-role" (p. 117), are the highest argument within the NP.

(41) \[
\begin{array}{c}
N'' \\
\downarrow \\
N'' \quad \text{possessor} \\
\downarrow \\
N' \quad \text{external thematic argument} \\
\downarrow \\
N \quad \text{internal argument} \\
\end{array}
\]

[Giorgi & Longobardi 1991:117.5]

A similar idea is found in work by Stowell (1989, 1991) who employs the DP-hypothesis instead of a layered NP. Stowell argues that the subject (SPEC) of NP is where arguments that are theta-marked by the noun originate, while [SPEC, DP] is where the Possessor theta-role is assigned. He cites Abney (1987) as assuming that the Possessor theta-role is assigned to the SPEC of DP. Stowell justifies this by saying that the SPEC position of a lexical category can only be assigned a theta-role that is specified as part of the lexical entry of the head. Since almost any object can be possessed, however, he states that it is unlikely that the Possessor theta-role is part of the lexical entry of every object (cf. Massam (1993)).

To summarize, we have seen that there is only one possessor position in Persian. Given that this position can contain a noun phrase bearing any sort of thematic relation to the head noun, it must be structurally high within the noun phrase (Higginbotham (1983), Giorgi and Longobardi (1991)). Further, given that this position can also contain non-thematically selected noun phrases, it must be outside NP and in [SPEC, DP] (Stowell (1991)).

---

13The proposal that the one possessor position in Persian is outside NP raises the question of how theta-assignment takes place when a possessor does bear a thematic relation to the possessed noun. This question is left for future research.
3.4.3 Summary

We have seen two possible analyses of possessed noun phrases in Persian. On the one hand, we have seen that it is possible to base-generate $\text{DP}_{\text{poss}}$ in the specifier of NP, either on the left or on the right of $\text{N}'$, and if necessary derive the correct surface word order via head-movement. This analysis preserves the idea that possessors are the subjects of NP and further, gives a simple account of the Ezafe domain. Under this approach all elements occurring within NP are marked with Ezafe. However, the binding properties of the reflexive/emphatic element $\text{xod}$ casts doubt on the $[\text{SPEC}, \text{NP}]$ approach. Since $\text{xod}$ can be bound by a possessor, an $i$-within-$i$ violation results if the possessor is in the $[\text{SPEC}, \text{NP}]$ position.

On the other hand we have seen that it is possible to base-generate $\text{DP}_{\text{poss}}$ in the specifier of DP. This is theoretically more attractive in that it accounts for the fact that possessors in Persian are not necessarily thematically selected. It also accounts for the binding of the reflexive/emphatic element $\text{xod}$, and the scope properties of noun phrases containing numerals. It also allows us to maintain the generalization (from Chapter 2) that nouns in Persian do not project to NP and thus an NP node cannot dominate any phrasal material. However, this analysis requires that the specifier of DP be generated on the right.

Under both analyses an empty case-assigning D-head is posited, which bears definiteness features. The fact that this element is null means that there is no obvious way to determine whether it heads a head-initial or head-final projection. However, since it is in complementary distribution with the indefinite enclitic $-i$, and since $-i$ appears on the right, we will assume that the empty definite determiner also appears on the right.

The structure for a possessed noun phrase is given below.
Given the structure in (42) we can now account for the complementary distribution between possessors and the indefinite enclitic \(-i\). As shown above, the empty determiner bearing the feature [+def] is a case-assigner. Let us assume that \(-i\) is not a case-assigner. This means \(-i\) cannot license a DP in its specifier. Thus, the apparent complementary distribution between possessors and the morpheme \(-i\), is in fact attributable to the properties of the two D-heads in Persian: one that is definite and can assign case, the other that is indefinite and cannot assign case.\(^{14}\)

3.5 What Projects a DP in Persian?

So far we have seen that the presence of a possessor requires a DP projection above common nouns in Persian. In this section we consider all the other elements that project a DP. These include proper names and pronouns, the indefinite enclitic \(-i\), and the (empty) definite D-head. In addition we see two derivational morphemes occurring on nouns in Persian, the plural marker and the colloquial marker \(-e\), both of which require a DP projection.

---

\(^{14}\)I leave as an open question here whether the case-assigning property of a determiner head is connected to its specification for definiteness (e.g. it is the feature [+def] that assigns case), or whether the ability to assign case is completely independent of the specification for the feature [definite].
3.5.1 Proper names and Pronouns

There is only one set of independent pronouns in Persian. These pronouns can function as subjects, possessors (linked with Ezafe), and objects of verbs or prepositions. The set is given in (43), with colloquial pronunciations given in brackets.

(43) Independent Pronouns

<table>
<thead>
<tr>
<th></th>
<th><em>sg</em></th>
<th><em>pl</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>man</td>
<td>mà, māhā</td>
</tr>
<tr>
<td>2</td>
<td>to</td>
<td>shomā, shomāhā</td>
</tr>
<tr>
<td>3</td>
<td>u</td>
<td>ishān (ishun)</td>
</tr>
<tr>
<td></td>
<td>ân (un), in</td>
<td>ânhā (unhā, unā), inhā (inā)</td>
</tr>
</tbody>
</table>

Lazard (1992) mentions that there is a third person pronoun _vey_ , referring only to animate beings, but he notes that this is only used in very literary language. He states that the first and second person plural pronouns are now used for singular referents in polite usage and thus (colloquially) they are inflected with the plural marker to have plural reference. Similarly, _ishān_ can only be used colloquially for the third person singular polite form and thus _ânhā_ is always used for the plural. Finally, note that _ân_ and _in_ are also used as demonstratives, meaning 'that (one)' and 'this (one)', respectively.

Proper names and pronouns project a DP since they are referential and bear features for definiteness. Consequently they are obligatorily marked with _-râ_ in direct object position, since _-râ_ selects a DP complement (see Chapter 5, section 5.1). Similarly, proper names and pronouns cannot be followed by the indefinite enclitic _-i_, nor by a possessor, both of which select an NP complement (in the latter case via the empty definite determiner). Structurally then, proper names and pronouns will appear under D as shown below.
3.5.2 The Indefinite Enclitic

In section 3.3.2 above we saw that the indefinite enclitic -i (so-called by traditional grammarians), is in complementary distribution with DP\textsubscript{poss}. It was proposed there that -i is a determiner heading a DP projection. Thus it appears in the following configuration:

\begin{equation}
(45) \begin{array}{c}
\text{DP} \\
D' \\
NP \\
\text{Ezafe Domain} \\
\text{-i}
\end{array}
\end{equation}

Recall that the specifier of this projection cannot be filled because the morpheme -i is not a case-assigner.

The structure in (45) and the fact that nouns don't project explains another interesting construction in which the morpheme -i appears.\(^{15}\) Consider the following example:

\begin{equation}
(46) \text{ketâb-e bozorg-e jâleb-i xarid-am}
\end{equation}

book+EZ big+EZ interesting+indef bought+1sgS

'I bought a big interesting book.'

Here the indefinite enclitic appears at the end of a string consisting of a noun and two adjectives. Recall that nouns and modifying adjectives are in a base-generated head

\(^{15}\)Thanks to Guy Carden for useful discussion regarding this construction.
adjunction structure. Now, the adjectives in the above example can also appear conjoined as shown below:

(47) ketāb-e bozorg-o jâleb-i xarid-am  
    book+EZ big+and interesting+indef bought+1sgS
    'I bought a big and interesting book.'

Of interest here is that in formal Persian the indefinite enclitic can appear on the head noun.

(48) ketāb-i bozorg-o jâleb xarid-am  
    book+indef big+and interesting bought+1sgS
    'I bought a big and interesting book.'
    (perhaps better translated as 'I bought a book, big and interesting.')

At first glance, if we compare (47) and (48) it looks like the morpheme -i can either attach to the end of a noun 'phrase' or to its head. However, if this were true we would expect to find a counterpart to (46) with -i occurring on the head. Such a construction is not possible. When -i occurs on a head noun, modifying adjectives must be conjoined and cannot be linked by the Ezafe morpheme.

(49) *ketāb-i bozorg-e jâleb xarid-am  
    book+indef big+EZ interesting bought+1sgS
    'I bought a big and interesting book.'
    (cf. *'I bought a book, big interesting.')

Recall from Chapter 2, section 2.8 that adjectives cannot modify other adjectives. Thus the ungrammaticality of (49) can be accounted for if the adjectives linked with Ezafe are not modifying the noun directly. This suggests that when -i appears on a head noun followed by modifying adjectives, the adjectives are in appositive position.
Under our analysis here, the fact that \(-i\) appears to be an affix that only attaches to head-adjoined material follows from the fact that phrasal constituents are not possible below the NP node. Thus we can maintain that \(-i\) is a phrasal affix (see Chapter 1, section 1.2.2) selecting NP as its complement.

This analysis of \(-i\) also explains why un yeki 'that one' cannot substitute for the noun phrase excluding the possessor (see section 3.1.2). Recall that for Samiian (1983) this was evidence that the possessor forms a constituent with the head noun and its modifiers. If we consider this form more closely, we see that it consists of a demonstrative element un 'that' (which we will leave aside for now), the numeral yek 'one' and the indefinite enclitic \(-i\). The fact that this pro-form cannot be followed by a possessor is explained by the fact that the indefinite determiner \(-i\) is in complementary distribution with the null definite determiner and cannot assign case to its specifier.

3.5.3 The Null Definite Determiner

A null definite determiner has been posited for noun phrases involving possessors. However, bare nouns in Persian can be construed as definite or generic. In direct object position these two uses are distinguished by the presence or absence of the direct object marker \(-râ\) (see Chapter 5). In subject position, however, there is no overt marking to distinguish the two readings. For example, consider the sentences below:

(50)  ketâb ru-ye miz bud-Ø
      book on+EZ table was+3sgS
      'The book was on the table.'

(51)  ketâb ru-ye miz bud-Ø
      book on+EZ table was+3sgS
      'There were books on the table.'
Abstracting away from tense which plays a role in whether a subject receives a generic or definite reading, the relevant point here is that there is no overt marking to distinguish the two uses of a bare noun.

I propose that there is a difference between generic and definite nouns and that the lack of any surface distinction is due to the fact that the definite determiner in Persian is null. The structures for both are given below.

\[
\begin{array}{c|c}
\text{Generic} & \text{Definite} \\
\hline
\text{NP} & \text{DP} \\
\text{N} & \text{NP} \\
\text{ketâb} & \text{D} \\
\text{N} & \text{Ø}_{\text{def}} \\
\text{ketâb} & \\
\end{array}
\]

3.5.4 The Pronominal Enclitics

In Persian, pronominal possessors can be expressed in two ways. An independent pronoun can be linked to the possessed noun in the Ezafe construction. Following the analysis of noun phrase structure just developed, the pronoun will be generated in a right-hand specifier of DP, and this DP will be headed by a null definite head. However, there is an alternative construction in which the possessor appears affixed to the possessed noun. The two alternatives are shown below.

\[
\begin{array}{c|c}
\text{(a) ketâb-e man} & \text{(b) ketâb-am} \\
\text{book+EZ me} & \text{book+1sg} \\
\text{'my book'} & \text{‘my book’} \\
\end{array}
\]
When the head noun appears with modifiers the enclitic attaches to the end of the X⁰-adjoined string. This is shown in (54).

(54)  
(a) ketâb-e jâleb-e man
      book+EZ interesting+EZ me
      'my interesting book'

(b) ketâb-e jâleb-am
      book+EZ interesting+1sg
      'my interesting book'

The paradigm of the pronominal enclitics is given in (55) with colloquial pronunciations given in parentheses.

(55)  

\[ \begin{array}{c|cc|c}
   & sg & pl \\
   \text{I} & \text{-am} & \text{-emân (-emun)} \\
   \text{2} & \text{-at (-et)} & \text{-etân (-etun)} \\
   \text{3} & \text{-ash (-esh)} & \text{-eshân (-eshun)} \\
\end{array} \]

The pronominal enclitics are the main focus of Chapter 4. However, for our purposes here we note the similarities between these morphemes and full pronoun possessors. For example, just like all other possessors, the pronominal enclitics are in complementary distribution with the indefinite enclitic -i.

(56)  
(a) *ketâb-e jiân-i
      book+EZ Jian+indef

(b) *ketâb-e man-i
      book+EZ me+indef
The properties that pronominal enclitic possessors share with full noun phrase possessors and full pronoun possessors can be captured structurally in the following way. Let us assume that with pronominal enclitic possessors, a pro is generated in the [SPEC, DP] position - the usual position for pronominal and noun phrase possessors. The difference here is that person and number (F) features are affixed to the null definite head to identify the pro possessor and are spelled out at PF. The structure is given below.

So far what has been proposed is that there are two types of DP in Persian, one headed by the indefinite enclitic -i, and the other headed by a null definite determiner. The latter head, by virtue of being a case-assigner, is also involved in the licensing of possessors, which occur in the specifier of DP. These proposals, taken together, will begin to explain plural marking in Persian.
Plural marking in Persian is a derivational rather than an inflectional affix in that it attracts stress (see Chapter 1, sections 1.2.2 and 1.3.1). However, it is usually in complementary distribution with overt numerals, shown in (58), suggesting that it is of the same category as they are.

(58) (a) se-tâ ketâb xarid-am  
three+unit book bought+1sgS  
'I bought three books.'

(b) *se-tâ ketâb-hâ xarid-am  
three+unit book+pl bought+1sgS

Even if it did not behave phonologically as a derivational affix, it would be hard to analyze plural marking as inflectional. For example, if the plural morpheme headed a functional category in the syntax, we would expect it to behave as a phrasal, not a lexical, affix. As can be seen in (59) below, however, in noun phrases involving adjectival modifiers the plural marker appears on the head noun, while the indefinite enclitic, a true phrasal affix, appears at the end of the string of modifiers. Since the noun and its modifiers are all X0s it is not clear how an inflectional morpheme heading a syntactic projection could be realized on the head.

---

16It should be noted that the plural marker being discussed in this section is -hâ, however, this is not the only way to mark plurality on Persian nouns. There is also a suffix -ân, which according to Lazard (1992) is used in literary language to mark plurality on animate nouns and a few others. The morpheme -ân is realized in a number of different ways depending on the phonological properties of the stem to which it attaches (see Lazard 1992:59-60). Lazard also notes that in words of Arabic origin the plural may be indicated according to Arabic rules (cf. Lazard 1992:60-62). Sometimes these rules are also extended to words not of Arabic origin. This section deals only with the plural -hâ for two reasons. First, it is the default plural marker in that almost all nouns that take their plural in some other form can also take this suffix. Second, all the comments made about -hâ can easily be extended to all other plural forms.
The other possible analysis for inflectional affixes (see the agreement affixes and the pronominal enclitics discussed in Chapter 4) is that they are features on a head that are spelled out at PF. However, in Persian the agreement affixes and the pronominal enclitics serve to identify some other constituent. With plural marking there is not a plausible syntactic constituent that the plural morpheme might be identifying.

Given all this then, we will take plural marking to be derivational. It is therefore attached to nouns in the lexicon. In other words, nouns are inserted into the syntax already marked for plural. However, we still must explain the syntactic properties of the plural marker, such as its complementary distribution with overt numerals.

A closer look at the syntactic properties of plural marking yields even more interesting facts. First, singular nouns in direct object position can appear with the indefinite enclitic -i, with the definite/presupposed direct object marker -râ, with both markers together, or bare, in which case they are construed as non-referential. However, nouns with plural marking must appear with one or both of the suffixes -i and -râ.

(60) (a) ketâb-o xarid-am
    book+râ bought+1sgS
    'I bought the book.'
(b) ketâb-i xarid-am
    book+indef bought+1sgS
    'I bought a book.'
(c) ketâb-i-ro xarid-am
    book+indef+râ bought+1sgS
    'I bought a (certain) book.'
(d) ketâb xarid-am
    book bought+1sgS
    'I bought books.'

(61) (a) ketâb-hâ-ro xarid-am
    book+pl+râ bought+1sgS
    'I bought the books.'
(b) ketâb-hâ-i xarid-am
    book+pl+indef bought+1sgS
    'I bought some books.'
Second, a bare noun in a copular construction is construed predicatively, while a plural noun is construed as referential, giving an equative reading for the sentence.17

(62)  
(a) ânhâ dâñeshju-and  
they student+3plS  
'They are students.'
(b) ânhâ dâñeshju-hâ-and  
they student+pl+3plS  
'They are the students.'

Third, while plural marking often cannot cooccur with a numeral, as was shown in (57) above, there are times when it can. If the noun phrase is construed as definite, plural marking is no longer ungrammatical, though it is interpreted as redundant. This is shown below, where the marker -râ (see Chapter 5) forces a definite reading for the noun phrase in object position. Plural marking is shown as optional to indicate that it is no longer ungrammatical, but does not usually appear.

(63)  
se-tâ  ketâb(-hâ)-ro xarid-am  
three+unit book(+pl)+râ bought+1sgS  
'I bought the three books.'

Finally, in subject position where there is no overt definiteness marking, bare nouns can be construed as either definite or generic. Again, nouns inflected for plural must have a

---

17In the present tense there is no verb stem in a copular construction and the agreement affix (-and '3pl' in (62)) is attached directly to the predicate.
definite interpretation. In fact, for a noun phrase containing a numeral to be construed as
definite in subject position, plural marking is obligatory.  

(64) (a) ketâbru-ye miz bud-Ø  
book on+EZ table was+3sgS  
'The book was on the table.'

(b) ketâbru-ye miz bud-Ø  
book on+EZ table was+3sgS  
'There were books on the table.'

(c) ketâb-hâ ru-ye miz bud-and  
book+pl on+EZ table were+3plS  
'The books were on the table.'

(65) (a) se-tâ ketâbru-ye miz bud-Ø  
three+unit book on+EZ table was+3sgS  
'Three books were on the table.'

(b) se-tâ ketâb-hâ ru-ye miz bud-and  
three+unit book+pl on+EZ table were+3plS  
'The three books were on the table.'

---

18 Plural marking is obligatory in the absence of demonstrative determiners. However, the
demonstratives are more commonly used to indicate definiteness. Note also that there are
some cases where the plural marker may not be able to appear (e.g. do nafar 'two people' but
*do nafar-hâ). In this case, stress is also a diagnostic for distinguishing the indefinite from
the definite reading. In the indefinite case stress appears on the numeral (dó nafar) but in the
definite it appears on the noun (do nafár) (cf. Lazard 1992:42-3). This raises some very
interesting questions about phrasal stress, but unfortunately an exploration of these questions
is beyond the scope of this work.

19 Interestingly, though not directly relevant to the point at hand, in this example the verb
takes third person singular rather than third person plural agreement.
These facts taken together make the semantics of the plural maker rather baffling. For example, the behaviour of the plural marker with overt numerals, where it forces a definite reading, seems to indicate that it is [+definite]. However, in other contexts it appears on indefinite nouns. This is noted by Karimi (1989, p. 114, fn. 16)\(^{20}\) who gives the following examples, showing that with an overt numeral the presence of the plural marker forces the presence of -râ (marking definite/presupposed direct objects), while without an overt numeral the plural marker does not force a definite reading.

(66) (a) do-tâ ketâb-â-ro xarid-am
    two+unit book+pl+râ bought+1sgS
    'I bought the two books.'

(b) ketâb-â-ye xub-i xarid-am
    book+pl+EZ good+indef bought+1sgS
    'I bought good books.' \[Karimi 1989:114, fn. 16\]

All these facts can be accounted for if plural marking on a noun requires the presence of a DP dominating the noun phrase. In direct object position this means that a plural noun must appear with either -i or the null definite determiner, which in turn forces the presence of -râ (see Chapter 5). When a plural noun appears as the predicate in a copular construction or as the subject of a sentence the null definite determiner must be present. With plural noun phrases containing numerals (which presumably don't project to DP themselves) again the null definite determiner is required in order to license the plural marker, giving a definite reading for the phrase.

---

\(^{20}\)Karimi assumes that the plural morpheme marks specificity although she notes that others (e.g. Samiian (1983)) have considered it a definite marker. She does not present arguments for either view simply adopting the former. Presumably this is because she assumes -râ marks specificity rather than definiteness. This issue will be dealt with in Chapter 5, section 5.2.1.
This description accounts for the apparent default reading of plural nouns as definite despite the fact that the plural marker can also cooccur with the indefinite enclitic -i (see (66b)). It also accounts for the necessarily referential reading of plural noun phrases. Recall from Chapter 1, section 1.2.6, that DPs are always referential since the determiner binds the variable position of the noun. Thus, in Persian, only the singular common noun can be generic, or non-referential.

Let us consider now how to implement the generalization that plural nouns must occur within a DP. I propose that although the plural morpheme is derivational, it can only be licensed by being checked by a D-head. This entails movement of the noun to D. Since there is no evidence for this movement in the overt syntax, we can assume that it takes place at LF.

I end this section with the following passage, cited by Windfuhr (1994), on the connection between plural marking and referentiality on the one hand, and singular nouns and genericity on the other.

21 The idea that some derivational morphemes may also have to be checked in certain syntactic configurations raises a number of questions that, unfortunately must be left for further study. Even in this specific case however, there are questions as to how the requirement of the plural marker should be formalized. Elizabeth Cowper (personal communication) has suggested that the plural marker bears a feature, for example [adefinite], which must be checked by either a [+definite] or [-definite] head. However, as we see in the following section, plural nouns are also licensed if they are inserted directly under D, i.e. when they are treated as proper names. If we assume that the D-head under which proper names and pronouns are inserted does not bear features of its own, then it seems that it is not a feature of D that licenses plural marking but simply that the plural noun appears under a D node.

22 Chomsky (1993) states that LF movement is "cheaper" than overt movement and calls this principle Procrastinate.
An interesting and insightful discussion of what nowadays is called genericity was made by the Persian Ibrahim in his manual of 1841, written for the East India Company's students in Haileybury, England. It shows that Ibrahim was thoroughly knowledgeable of his native grammatical tradition. The bulk of his grammar is contained in eight dialogues between a Persian and an open-minded Persian-speaking Englishman in Iran. In one of the dialogues the Englishman uses the plural in the phrase: /vaxti ke asb-ha [horse+pl: JG] amade and, soma ma-ra xabar konid/ "When the horses are ready give me notice." Politely, the Persian explains at length that a /lafz-e mofrad/ "a single word" does not always denote a unit. "But don't we need two horses?" "Yes," replied the Persian, now returning to English for a lengthy discussion: "But we didn't talk about specific horses, nor had we talked earlier about certain horses to which that expression could have referred; we just wanted to ride out on a horse" (Fleischer 1875:117)

3.5.6 The Colloquial Marker -e

To complete the survey of elements that force a DP projection in noun phrases in Persian, the morpheme -e should be mentioned. This marker only occurs in spoken Persian, and according to Karimi (1989) is restricted to an informal speech style. It appears as a stressed -e word-finally and as stressed -a non-word-finally. It occurs on nouns in any position marking them for definiteness (Samiian (1983)) or specificity (Karimi (1989)). I will not deal with the precise semantics here, simply noting that previous mention in the discourse is apparently relevant to the occurrence of this morpheme. Examples are given below:

(67) (a) doxtar-e âmad-Ø
girl+def came+3sgS
'The girl came.'

(b) ketâb-o be doxtar-e dâd-am
book+râ to girl+def gave+1sgS
'I gave the book to the girl.'
As mentioned above, Samiian (1983) considers -e a definiteness marker. She notes two further interesting properties of this affix. First, it can follow not just bare nouns, but also nouns followed by an attributive adjective. In such a case the Ezafe morpheme does not appear between the noun and the modifying adjective, as shown in (68b) below.

(68) (a) ketâb-e bozorg-o xarid-am
book+EZ big+râ bought+1sgS
'I bought the big book.'

(b) ketâb bozorg-a-ro xarid-am
book big+def+râ bought+1sgS
'I bought the big book.' (where the big book has been under discussion)

The other interesting fact noted by Samiian is that the plural counterpart to the marker -e is the regular plural marker -hâ. Thus we get the following contrast corresponding to the pair of sentences given above.

(69) (a) ketâb-hâ-ye bozorg-o xarid-am
book+pl+EZ big+râ bought+1sgS
'I bought the big books.'

(b) ketâb bozorg-â-ro xarid-am
book big+pl+râ bought+1sgS
'I bought the big books.' (where the big books have been under discussion)
Samiian shows that this use of the plural marker is different from its regular use in that, not only does it occur after the attributive adjective here, but it cannot be followed by the indefinite marker \(-i\), whereas normally it can be. Note that the marker \(-e\) also cannot be followed by \(-i\).

(70) (a) ketâb-hâ-ye bozorg-i xarid-am
        book+pl+EZ big+indef bought+1sgS
        'I bought some big books.'

        (b) *ketâb bozorg-å-i xarid-am
            book  big+pl+indef bought+1sgS

(71) (a) ketâb bozorg-a-ro xarid-am
        book  big+def+râ bought+1sgS
        'I bought the big book.'

        (b) *ketâb bozorg-a-i xarid-am
            book  big+def+indef bought+1sgS

Samiian's analysis of \(-e\) involves a restructuring rule. Specifically, to prevent the occurrence of the Ezafe marker between the noun and the attributive adjective, she proposes a rule that converts the noun and following adjective into a complex noun. She states that the rule is triggered by the presence of the definite marker, \(-e\). Essentially then, this is a rule creating a compound structure.

If we draw on the spirit of Samiian's analysis, but also maintain the assumptions that have been held thus far, an analysis suggests itself. Note that both the definite suffix \(-e\), and (as we have seen) the plural marker are stressed. Recall that stressed affixes in Persian are derivational. Let us assume then, that \(-e\) is attached pre-syntactically forming lexically definite nouns. This essentially turns them into proper names, which occur in the syntax
under a D-node rather than an N-node. The derivational nature of this affix explains why it
does not occur at the edge of phrases. Only when a compound has been formed will the affix
be able to occur. Further, assuming that a noun and a following adjective which bears the \(-e\)
suffix form a compound explains the lack of Ezafe between them. Note that the formation of
compounds to refer to presupposed, previously mentioned entities is not unusual.\(^{23}\)

The analysis of \(-e\) as a marker turning common nouns into definite D's by supplying them
with a definiteness feature predicts that items already specified for definiteness will not occur
with this suffix. This prediction is borne out. Proper names and pronouns cannot occur with
\(-e\) as shown in (72).

\[(72)\hspace{1em} (a)\hspace{1em} *jiân-a-\text{ro} \hspace{0.5em} \text{did-am}
\hspace{2em} \text{Jian}+\text{def}+\text{râ saw}+1\text{sgS}
\hspace{2em} *'I saw the Jian.'
\]

\[(b) \hspace{1em} *\text{un-a-\text{ro}} \hspace{0.5em} \text{xarid-am}
\hspace{2em} \text{that}+\text{def}+\text{râ bought}+1\text{sgS}
\hspace{2em} *'I bought the that.'
\]

Nothing further needs to be said about the plural marker when it functions as the
counterpart to \(-e\). Recall from the previous section that the plural marker must by checked
by a D-head. This can be achieved by movement of the head noun to D at LF. However, the
plural marker can also be licensed if the noun on which it attaches is inserted under a D in
the syntax. In this case, the plural noun is interpreted as a proper name.

\(^{23}\)See Downing (1977) for a particularly interesting discussion of compounds in English.
3.6   A Reformulation of the Ezafe Insertion Rule

In Chapter 2, section 2.6.3 it was suggested that the function of the Ezafe vowel is to identify constituenthood. In this section the Ezafe Insertion Rule of Samiian (1983) is reformulated in light of the structure of the noun phrase that has been developed in this and the previous chapter. In doing so the function of Ezafe Insertion will become clearer.

Recall that in Samiian's formulation, the Ezafe Insertion Rule inserted -e before maximal projections dominated by X' (X ≠ V). Given that transformations of this kind are no longer thought to belong to syntax proper, this rule must be reformulated as a post-syntactic rule. I suggest that it takes place in the spell-out component at PF.

Let us consider the kinds of operations that take place in the spell-out component. Some assume that inflectional morphemes are spelled-out at PF (e.g. Bonet (1991), Anderson (1992), Chomsky (1993), Cummins and Roberge (1993, 1994) among others). There are also proposals that all vocabulary insertion takes place post-syntactically (e.g. Halle and Marantz (1993)). However, given that the Ezafe vowel does not correspond to any lexical or semantic feature and, in fact, seems to have no semantic content at all, it is questionable whether it should be viewed as a morpheme.

PF is also the level at which phrasing and intonational rules are applied. Included here are rules that concern syntactic constituents smaller than the sentence and which affect the phonological realization of these constituents. While in English the prototypical example of such a rule involves stress, phenomena such as liaison in French are presumably also derived at this level. I suggest that the Ezafe Insertion Rule should be viewed in this vein.24

We saw in Chapter 2 that the trigger for Ezafe Insertion is the feature [+N] borne by an X^0. However, not every X^0 bearing the feature [+N] is marked with the Ezafe vowel, but

24As pointed out by Elan Dresher (personal communication) phonological phrases are not always isomorphic with syntactic phrases. This makes the Ezafe Insertion Rule rather different from a prototypical rule of phonological phrasing. What is relevant here, however, is that the Ezafe Insertion rule belongs in the same component as these other rules.
only those followed by a modifier or complement. An initial description of the rule is given in (73).

(73) **Ezafe Insertion Rule** (Version 1)

Insert the vowel -e on an X\textsuperscript{0} that bears the feature [+N] and is followed by another constituent within the same projection.

This rule will insert the Ezafe vowel in the following three environments (where P\textsubscript{nom} stands for a nominal preposition). Note that the rule applies whether the following constituent is an X\textsuperscript{0} or an XP.

(74) (a) \[ N_{i0} \]

(b) \[ AP \]

(c) \[ P_{nomP} \]

Recall, that the Ezafe vowel also occurs on a noun followed by a possessor. As we have seen in this chapter the possessor is not within NP, but in the specifier of DP. This is shown below, with the Ezafe vowels indicated.\textsuperscript{25}

(75)

\[ \text{DP} \]

\[ \text{DP}_{\text{Poss}} \]

\[ \text{NP} \]

\[ \text{D-e} \]

\[ \text{Ø}_{\text{def}} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ P_{nom}^{0} \]

\[ A^{0} \]

\[ N^{0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \]

\[ N_{i0} \] \textsuperscript{25}

The reason why the Ezafe vowel appears on D in (75) and not on the last element in NP will be discussed below.
The Ezafe Insertion Rule has to be amended to account for the presence of the Ezafe vowel before the possessor. This can be done if we use the notion of 'extended projection' proposed by Grimshaw (1991). The key idea is that N and D on the one hand, and V and I on the other, share the same categorial features. N and V are distinguished from D and I in their lexical (as opposed to functional) status. Grimshaw provides a feature {F} to distinguish lexical from functional categories. Lexical categories are specified as {F0} while functional categories are specified as {F1}. She then gives the following definitions (where (76) corresponds to the traditional notion of head and projection):

(76) x is the perfect head of y, and y is a perfect projection of x iff:
   (a) y dominates x
   (b) y and x share all categorial features
   (c) all nodes intervening between x and y share all categorial features
   (d) the F value of y is the same as the F value of x  \[Grimshaw 1991:3.3\]

(77) x is the extended head of y, and y is an extended projection of x iff:
   (a) y dominates x
   (b) y and x share all categorial features
   (c) all nodes intervening between x and y share all categorial features
   (d) If x and y are not in the same perfect projection, the F value of y is higher than the F value of x  \[Grimshaw 1991:4.4\]

A revised version of the Ezafe Insertion Rule can now be given that accounts for the presence of the Ezafe vowel on the constituent preceding a possessor.

(78) Ezafe Insertion Rule (Version 2)
    Insert the vowel -e on an X₀ that bears the feature [+N] and is followed by another constituent within the same extended projection.
Finally, we have to account for the places where Ezafe does not appear. The Ezafe vowel does not appear on the last $X^0$ within NP if this NP is followed by a phrasal affix (e.g. -$i$, the indefinite enclitic, or the pronominal enclitics). Of course, it may be that the rules cliticizing (in the case of -$i$) or spelling out (in the case of the pronominal enclitics) these morphemes apply prior to the application of the Ezafe Insertion Rule. Nevertheless, until the interaction of rules and processes within the PF component are better understood, we must block the rule in these environments.

The Ezafe vowel also fails to appear on the last $X^0$ within the NP when this element is followed by the empty definite determiner but no possessor. This can be explained by assuming that the null definite determiner is affixal. Note that the Ezafe vowel does appear when the DP contains a possessor. In this case it appears on the D since this element is also [+N] under Grimshaw's (1991) claim that N and D share categorial features. The final version of the Ezafe Insertion Rule can now be formulated and is given in (79).

(79) **Ezafe Insertion Rule** (Final Version)

Insert the vowel -$e$ on an $X^0$ that bears the feature [+N] and is followed by another non-affixal constituent within the same extended projection.

To summarize, the Ezafe vowel is inserted at PF on a [+N] element that is followed by another independent constituent within the same extended projection. Given that all of the lexical categories on which the Ezafe vowel appears are potentially non-projecting, its presence identifies constituents that are not clearly given by the phrase structure. That is, under the view taken here, the non-projection of [+N] elements in Persian is connected to, and part of the motivation for, the rule of Ezafe Insertion.

---

26 There is an alternative to assuming that the null definite determiner is an affix. It may be that Ezafe Insertion is only triggered by phonetically realized material following the relevant [+N] item. Thus, Ezafe Insertion would not be triggered by the null definite determiner, but would be triggered by a following possessor since the possessor has phonetic content. Further research is required to determine which alternative is the preferred one.
3.7 Conclusion

In this chapter we have examined the noun phrase above the Ezafe domain. Focussing on possessors, we have seen that they occur in the right specifier of a DP projection headed by a null definite determiner. We have seen that this null determiner is in complementary distribution with the indefinite enclitic -i, which also heads a DP projection. The latter element lacks a case-assigning feature and is thus unable to license a possessor.

The existence of a null definite determiner explains why bare nouns in Persian have two possible readings. They can be construed as generic, or non-referential, in which case they occur without a DP projection. They can also be construed as definite when they are selected by the null definite determiner.

We have also seen how the analysis developed here explains the curious properties of the plural morpheme in Persian, the generalization being that the plural morpheme can only occur on a noun that is contained within a DP.

Finally, given the analysis developed in the last two chapters, the Ezafe Insertion Rule has been formulated as a rule that applies at PF in order to identify [+N] elements and following constituents as being contained within the same extended projection.

Appendix to Chapter 3

Pre-head Material in the Noun Phrase

This appendix contains information about the elements that can precede the head noun within the Persian noun phrase. These items do not occur with the Ezafe vowel and are not relevant to the analysis just developed. They include demonstratives, numerals, an indefinite article, classifiers, quantifiers and superlatives.
I Demonstratives

Demonstratives, if present, are always the first element in the noun phrase. When they occur on a direct object, the definite/presupposed direct object marker -râ is required. There are two demonstratives in Persian, in 'this' and ân 'that' (pronounced as un, colloquially). They occur with both singular and plural noun phrases and can also act as independent pronouns. They appear with plural marking only in the latter case. Examples are given below.

(1)  (a) in ketâb jâleb-e
      this book interesting+3sgS
      This book is interesting.'

      (b) un ketâb-o xarid-am
          that book+râ bought+1sgS
          'I bought that book.'

      (c) un ketâb-hâ-ro xund-am
          that book+pl+râ read+1sgS
          'I read those books.'

      (d) un-o dust dâr-am
          that+râ like+1sgS
          'I like that.'

Demonstratives can cooccur with all other elements of the noun phrase, including numerals, possessors, and even the indefinite enclitic -i under certain circumstances (see section 3.3.2). Examples are given below.
II Numerals and Classifiers

Numerals occur between demonstratives and the head noun. Recall from section 3.5.5 that in direct object position, nouns with numerals are construed as indefinite unless they bear the direct object marker -râ, in which case they are construed as definite.

Numerals obligatorily occur with a classifier. Examples of classifiers, taken from Samiian (1983) are given below:

(3)  (a) **classifier proper**

- tâ  
  'unit,' used with all count nouns

  nafar  
  'person'

  jeld  
  'unit' used for books

  pors  
  'unit' used for meals

(b) **measurement nouns**

  metr  
  'metre'

  kilo  
  'kilogram'

  litr  
  'litre'
(c) **group nouns**

(i) used for count nouns

<table>
<thead>
<tr>
<th>classifier</th>
<th>meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>daste</td>
<td>'bunch'</td>
</tr>
<tr>
<td>goruh</td>
<td>'group'</td>
</tr>
</tbody>
</table>

(ii) used for mass nouns

<table>
<thead>
<tr>
<th>classifier</th>
<th>meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>fenjun</td>
<td>'cupful'</td>
</tr>
<tr>
<td>qāshoq</td>
<td>'spoonful'</td>
</tr>
</tbody>
</table>

[Samiian 1983:145.17]

Samiian (1983:159) notes that there is in fact only a two way syntactic distinction among the classifiers listed above. The classifiers proper and measure nouns on the one hand are distinguished from the group nouns on the other. I will refer to classifiers proper and measure nouns as 'true classifiers', in contrast with the group nouns. There are three differences between true classifiers and group nouns. First, group nouns can be preceded by -tâ whereas true classifiers cannot.27

\[(4) \quad \text{se} \quad \text{jabe ketâb} \quad \text{three box book} \quad \text{'three boxes of books'} \quad \text{[Samiian 1983:162.83]} \]

\[(b) \quad \text{se-tâ} \quad \text{jabe ketâb} \quad \text{three+unit}^{28} \quad \text{box book} \quad \text{'three boxes of books'} \quad \text{[Samiian 1983:162.82]} \]

(5) (a) se kilo gusht

three kilo meat

'three kilos of meat'

---

27 This suggests that -tâ is actually in a third category and should not be included with the true classifiers. Note that it is affixal while the rest of the classifiers are independent words. Further, it is the only classifier which can appear with the plural marker (see section 3.5.5).

28 Samiian glosses -tâ here as "class.", even though she glosses it as 'unit' elsewhere, so I have changed the gloss for consistency.
(b) *se-tâ kilo gusht
three+unit kilo meat
'three kilos of meat'  
[Samiian 1983:163.90]

Second, Samiian states that group nouns have independent status as nouns whereas true
classifiers do not, since group nouns can occur without numerals and with demonstratives
instead. The examples she gives are repeated below.

(6) (a) in daste gol
this bunch flower
'this bunch of flowers'
[Samiian 1983:164.97]

(b) in livân âb
this glass water
'this glass of water'
[Samiian 1983:164.98]

(7) (a) *in-tâ ketâb
this+unit book
'this book'
[Samiian 1983:164.95]

(b) *in kilo gusht
this kilo meat
'this kilo of meat'
[Samiian 1983:164.96]

Third, Samiian shows that group nouns can be modified by an adjective when preceded
by -tâ. When they are not preceded by -tâ, however, they behave like true classifiers in not
being able to be modified by an adjective. Her examples are given below:

(8) (a) do-tâ daste-ye qashang gol
two+unit bunch+EZ beautiful flower
'two beautiful bunches of flowers'
[Samiian 1983:164.99]
(b) *do daste-ye qashang gol
   two bunch+EZ beautiful flower

   [Samiian 1983:164.100]

(c) *do kilo-ye sangin gusht
   two kilo+EZ heavy meat
   'two heavy kilos of meat'

   [Samiian 1983:164.101]

However, the speakers I have consulted find (8a) ungrammatical. They consistently change it to the following:②⁹

(9) do(-tâ) daste-ye qashang-e gol
    two+unit bunch+EZ beautiful+EZ flower
    'two beautiful bunches of flowers'

In this case the adjective too bears the Ezafe vowel which suggests that the group noun is now functioning as a true noun in an Ezafe construction. Even in this case, however, group nouns are still distinguished from true classifiers in that the latter can't function as the head of a noun phrase.

(10) *do(-tâ) metr-e pahn-e pârche
    two(+unit) metre+EZ wide+EZ material

We can note that Samiian's point still holds even without this third argument, namely that there is a difference between true classifiers and group nouns.

②²⁹What is puzzling about example (9) is that the suffix -tâ is not obligatory on the numeral, i.e. the sentence would still be grammatical without it. This is not true of other nouns which must be preceded by a numeral with -tâ.
Note that neither a true classifier and a noun, nor a group noun and its nominal complement, can stand alone as an argument of a verb.

(11) (a) *jeldketâb xarid-am
       unit book bought+1sgS
       *'I bought volume book.'

       (b) *daste gol xarid-am
           bunch flower bought+1sgS
           *'I bought bunch flowers.'

Finally, let us consider the possessor as it interacts with numerals, classifiers, and other elements within the noun phrase. It turns out that there is a restriction on when the possessor can appear. The possessor can follow a numeral + true classifier as shown in (12a) below, strings involving a numeral + group noun followed by a possessor are considerably harder for speakers to process. This is shown in (12b). A group noun + noun followed by a possessor is completely ungrammatical, as seen in (12c). This is surprising since, unlike true classifiers, group nouns can occur without numerals and can follow a demonstrative.

(12) (a) se jeld ketâb-e hasan
       three unit book+EZ Hasan
       'Hasan's three books'

       (b) ?se livân âb-e hasan
           three glass water+EZ Hasan
           'Hasan's three glasses of water'

       (c) *livân âb-e hasan
           glass water+EZ Hasan
           'Hasan's glass of water'
Note that the example below, with the group noun functioning as the head of the phrase (note that it now takes Ezafe) is perfectly grammatical.

(13)  livân-e  âb-e      hasan  
glass+EZ water+EZ Hasan  
'Hasan's glass of water'

The cooccurrence restrictions among the various elements discussed up to now are shown in the chart in (14).

(14)

<table>
<thead>
<tr>
<th></th>
<th>demonstrative</th>
<th>numeral</th>
<th>true classifier</th>
<th>group noun</th>
<th>noun</th>
</tr>
</thead>
<tbody>
<tr>
<td>√</td>
<td>demonstrative</td>
<td>numeral</td>
<td>true classifier</td>
<td>group noun</td>
<td>noun</td>
</tr>
<tr>
<td>√</td>
<td>demonstrative</td>
<td>numeral</td>
<td>true classifier</td>
<td>group noun</td>
<td>noun</td>
</tr>
<tr>
<td>*</td>
<td>demonstrative</td>
<td>numeral</td>
<td>true classifier</td>
<td>noun</td>
<td></td>
</tr>
<tr>
<td>√</td>
<td>numeral</td>
<td>true classifier</td>
<td>noun</td>
<td></td>
<td></td>
</tr>
<tr>
<td>*</td>
<td>numeral</td>
<td>group noun</td>
<td>noun</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

III The Indefinite Article

An element which I will call the indefinite article can also precede the head noun in the noun phrase. Recall that Persian has an indefinite enclitic -i that can attach to noun phrases. This enclitic often cooccurs with yek (ye in colloquial speech) which on its own is the numeral 'one'. In the colloquial language yek and -i together, or yek alone is preferred to -i alone. The three options are shown below.
I propose that this use of *yeκ* is different from its use as a numeral. First recall from the previous section that numerals must occur with classifiers, yet when *yeκ* occurs with -i as in (15b) there is no classifier. When *yeκ* is used as a true numeral, a classifier must occur, as shown in (16).

(16) ye dune ketâb xarid-am
    one unit book bought+1sgS
    'I bought one book.'

Second, recall that the suffix -i is ordinarily in complementary distribution with numerals (see (17a)). When *yeκ* occurs with a classifier this generalization holds (see (17b)). However, as we saw above, when no classifier in present -i and *yeκ* cooccur.

(17) (a) *se-tâ ketâb-i xarid-am
    three+unit book+indef bought+1sgS

(b) *ye dune ketâb-i xarid-am
    one unit book+indef bought+1sgS
Having distinguished *yek* from the numerals, I suggest that it also differs (albeit subtly) from *-i*. First, it does not have the same element of meaning (characterized as 'selection') that allows *-i* to occur on the heads of restrictive relative clauses (see section 3.3.2). Second, while *-i* can cooccur with plural marking, *yek* cannot, as shown in (18).

(18) (a) ketâb-hâ-i xarid-am
    book+pl+indef bought+1sgS
    'I bought some books.'

(b) *ye ketâb-hâ(-i) xarid-am
    one book+pl(+indef)bought+1sgS

Third, as we saw in section 3.3.2, *-i* cannot cooccur with a possessor, while *yek* can.30

(19) (a) *ketâb-e hasan-i
    book+EZ Hasan+indef

(b) ye ketâb-e hasan
    one book+EZ Hasan
    'a book of Hasan's'

(20) ye ketâb-e barâdar-am
    one book+EZ brother+1sg
    'a book of my brother's' [Samiian 1983:103.78b]

The conclusion to be drawn from this is that *yek* is developing into a prenominal singular indefinite determiner along the lines of English *a/an*.

30According to the Persian speakers I consulted these examples are not completely natural, thought they improve in context.
IV Quantifiers

Quantifiers also precede the head noun within the noun phrase. Quantifiers cannot occur with demonstratives or numerals. Samiian (1983) divides quantifiers into two groups. The first group (Group A) take their nominal complements either with Ezafe or in a partitive structure with the preposition az 'from, of'. These quantifiers appear to be able to take phrasal complements. Samiian's list of Group A quantifiers is given in (21), followed by some examples.

(21) Group A Quantifiers

<table>
<thead>
<tr>
<th>Quantifier</th>
<th>Complement</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>hame</td>
<td>(EZ)</td>
<td>'all'</td>
</tr>
<tr>
<td>tamâm</td>
<td>(EZ)</td>
<td>'all'</td>
</tr>
<tr>
<td>bishtar</td>
<td>(EZ)</td>
<td>'most'</td>
</tr>
<tr>
<td>bazi</td>
<td>(az)</td>
<td>'some'</td>
</tr>
<tr>
<td>besiâri</td>
<td>(az)</td>
<td>'many'</td>
</tr>
<tr>
<td>hich kodum</td>
<td>(az)</td>
<td>'none'</td>
</tr>
<tr>
<td>har kodum</td>
<td>(az)</td>
<td>'each one'</td>
</tr>
</tbody>
</table>

[Samian 1983:146]

(22) (a) hame-ye bachche-hâ

all+EZ child+pl

'all of the children' [Samian 1983:147.20a]

(b) baziaz bachche-hâ

someof child+pl

'some of the children' [Samian 1983:147.21a]

Group B quantifiers either take their complements in the partitive construction or simply precede the noun. In the latter case the noun cannot appear with plural marking, suggesting

31See Samiian (1983), Chapter IV for more on the properties of quantifiers.
that these quantifiers do not take DP but rather NP as their complement. Some Group B quantifiers are given in (23) followed by examples.

(23) Group B Quantifiers

<table>
<thead>
<tr>
<th>Example</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>xeyli (az)</td>
<td>'many/much'</td>
</tr>
<tr>
<td>kami (az)</td>
<td>'a little/a few'</td>
</tr>
<tr>
<td>qadri (az)</td>
<td>'some'</td>
</tr>
<tr>
<td>ye zarre/xurde (az)</td>
<td>'a little'</td>
</tr>
<tr>
<td>har no/gune (az)</td>
<td>'every kind'</td>
</tr>
<tr>
<td>har che/qadr (az)</td>
<td>'whatever'</td>
</tr>
<tr>
<td>hame no/gune (az)</td>
<td>'all kinds'</td>
</tr>
<tr>
<td>che no (az)</td>
<td>'what kind'</td>
</tr>
<tr>
<td>che qadr (az)</td>
<td>'how much'</td>
</tr>
<tr>
<td>hich no (az)</td>
<td>'no kind'</td>
</tr>
<tr>
<td>hich chi (az)</td>
<td>'none'</td>
</tr>
<tr>
<td>in/un hame (az)</td>
<td>'all this/that much'</td>
</tr>
</tbody>
</table>

(Samian 1983:146)

(24) (a) xeyli az bachche-hâ many of child+pl 'many of the children'  
(Samian 1983:147.23)

(b) xeyli bachche many child 'many children'  
(Samian 1983:147.25)

(25) (a) kami az in qazâ little of this food 'a little of this food'  
(Samian 1983:147.24)

(b) kami qazâ little food 'a little food'  
(Samian 1983:147.26)
V Superlatives and Ordinals

Adjectives have both a comparative and superlative form. The comparative is formed by adding the suffix \(-tar\) to the adjective. The superlative is formed by adding \(-in\) to the comparative. While positive or absolute adjectives and comparatives follow the noun in the Ezafe construction, superlatives precede the noun and do not take Ezafe.

\[(26) \begin{align*}
(a) & \quad \text{kétáb-e bozorg} \\
& \quad \text{book+EZ big} \\
& \quad \text{'big book'} \\

(b) & \quad \text{kétáb-e bozorg-tar} \\
& \quad \text{book+EZ big+comp} \\
& \quad \text{'bigger book'} \\

(c) & \quad \text{bozorg-tar-in kétáb} \\
& \quad \text{big+comp+sup book} \\
& \quad \text{'biggest book'}
\end{align*}\]

This is also true of the ordinals. With the exception of the number one (\(yek\)) which has a suppletive form (\(avval\)), and the numbers two (\(do\)) and three (\(se\)), which take the suffix \(-vvom\), ordinals are formed from numerals by adding the suffix \(-om\). These ordinals follow the noun linked by the Ezafe vowel. However, the addition of the suffix \(-in\) to these ordinals yields another set of ordinals which precede the noun and do not take Ezafe.

---

32Traditionally the unmarked term in the three-way classification of adjectives into degrees is referred to as as the 'positive' or 'absolute' degree. This contrasts with the terms 'comparative' and 'superlative' (Crystal (1991)). Thanks to Elizabeth Cowper for bringing this to my attention.

33It is not at all clear how to characterize the difference in meaning between these two sets of ordinals. Lazard (1992) offers the following, not particularly helpful, explanation:
VI Summary

We have surveyed the elements that can precede the head noun in the noun phrase. None of these items (with the exception of some quantifiers) appear with the Ezafe vowel. All are single words, and none of them seem to occur with any kind of specifier. This raises the question of what sorts of projections these items head and how many such projections are possible above the NP.

An adequate analysis of these items must also account for their selectional properties. Questions such as the following must be answered: Why do numerals always occur with a classifier? Why can group nouns but not classifiers occur alone with a demonstrative? Why can neither true classifiers nor group nouns stand alone followed by a noun? What determines whether quantifiers take phrasal or non-phrasal complements? The interaction and relative position of these items vis-a-vis DPs headed by the null definite determiner and indefinite enclitic must also be explored.

There are also questions about the nature of these prenominal positions. For example, the prenominal equivalent (yek) to the indefinite enclitic (-i), acts more like the English indefinite article in that it cannot occur with plural nouns, but can occur with a possessor.

While the ordinals of the first series indicate simply numerical rank, the ordinals in -in indicate the unity which completes a series: it (sic) is used particularly in order to designate an anniversary. [Lazard 1992:101]
This may be a function of its structural position and not of its inherent features. We have also seen that items that can function as head nouns, and must therefore be specified as [+N], can occur prenominally as group nouns, in which case they do not trigger Ezafe insertion. This suggests that something about the prenominal position renders this feature opaque to the Ezafe Insertion Rule.

Having outlined some of the questions that remain about noun phrases in Persian, we now leave the noun phrase in order to focus on larger constituents. In the next two chapters we will consider the pronominal elements of Persian, and some issues in the structure of the simple sentence.
Chapter 4

Two Types of Agreement

4.0 Introduction

The focus of this chapter is the following two paradigms (the colloquial pronunciations are omitted here, but are included when each paradigm is further discussed in this chapter).

(1) **Subject Agreement**

<table>
<thead>
<tr>
<th></th>
<th>sg</th>
<th>pl</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-am</td>
<td>-im</td>
</tr>
<tr>
<td>2</td>
<td>-i</td>
<td>-id</td>
</tr>
<tr>
<td>3</td>
<td>Ø/-ad</td>
<td>-and</td>
</tr>
</tbody>
</table>

(2) **Pronominal Enclitics**

<table>
<thead>
<tr>
<th></th>
<th>sg</th>
<th>pl</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-am</td>
<td>-emân</td>
</tr>
<tr>
<td>2</td>
<td>-at</td>
<td>-etân</td>
</tr>
<tr>
<td>3</td>
<td>-ash</td>
<td>-eshân</td>
</tr>
</tbody>
</table>

These two paradigms encode person and number features. The morphemes in (1) indicate subject agreement. The morphemes in (2) are construed as possessors when attached to nouns, but can also be used as direct objects and objects of prepositions. Unlike full pronouns in Persian, which also encode person and number features, the morphemes in both (1) and (2) require a host to attach to. They do not affect stress placement, and are thus classified as inflectional rather than derivational morphemes (see Chapter 1).

The goal of this chapter is to provide a full account of the similarities and differences between the two sets of morphemes shown above. The chapter consists of two main parts. The first part presents an analysis of the structural configurations in which these two sets of
morphemes occur. It is shown that both sets of paradigms are the spell out of F-features that identify noun phrases in argument positions. It is in doing so that these two sets of morphemes are themselves licensed.¹ They differ in that the pronominal enclitics absorb case and therefore only cooccur with pro. The subject agreement affixes assign nominative case and therefore must cooccur with a subject whether it be pro or a full noun phrase. A further difference is that the subject agreement affixes subcategorize a verb, while the pronominal enclitics cannot occur directly on a verb stem.

Given that both sets of affixes can license pro, the second part of the chapter compares the binding properties of independent pronouns to those of pro showing that they are markedly different. An account of the binding facts is presented which also includes a discussion of the anaphor xod in Persian.

The chapter is organized as follows. Section 4.1 deals with subject agreement. In section 4.2 we turn to the pronominal enclitics, surveying the environments in which they occur, and then discussing a number of different approaches to account for their properties. In section 4.3 the subject agreement suffixes and the pronominal enclitics are compared with respect to their morphological properties, their spell-out at PF, and with respect to the distinction between affixation and cliticization.

The binding of pronominals and anaphors in Persian is discussed in section 4.4. This includes a description of the data, as well as a comparison of classic binding theory and the theory of Reinhart and Reuland (1993). It is shown that a combination of these two approaches is required in order to account for binding in Persian. In section 4.5 some concluding remarks are made and one outstanding problem is briefly discussed.

¹To avoid terminological confusion I will refer to the licensing of morphemes as 'checking'. That is, the agreement suffixes and the pronominal enclitics must be checked in order for a derivation to be successful.
4.1 Subject Agreement

The subject agreement suffixes are referred to by Lazard (1992) as *inflectional endings*, and as *personal endings* or *enclitic personal endings* in other grammars. They are referred to as agreement suffixes here to reflect more closely their function in the clause. The paradigm is repeated below, with the colloquial pronunciations given in parentheses.

(3) **Subject Agreement**

<table>
<thead>
<tr>
<th></th>
<th><strong>sg</strong></th>
<th><strong>pl</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-am</td>
<td>-im</td>
</tr>
<tr>
<td>2</td>
<td>-i</td>
<td>-id (-in)</td>
</tr>
<tr>
<td>3</td>
<td>Ø/-ad (-e)</td>
<td>-and (-an)</td>
</tr>
</tbody>
</table>

The third person singular forms differ depending on the tense of the verb. In the past tense the third person singular ending is null, while in the present tense it is realized as *-ad*.

Persian is a Null Subject language, in that overt pronouns in subject position are generally used only for emphasis. The person and number features of the subject are recoverable from the agreement marking on the verb. Examples of transitive and intransitive sentences are given below, with pronominal subjects enclosed in parentheses to indicate that they are not usually present.

(4) (a) (man) ketâb xarid-am
    I book bought+1sgS
    'I bought books.'

---

2Hashemipour (1989, p. 27-8) shows that Persian exhibits some other characteristics common to null subject languages, such as free subject-verb inversion, extraction out of a WH-island, null resumptive pronouns, lack of overt expletives, and an apparent lack of *that*-t effects.

3Lazard (1992) makes this observation also, but goes on to claim that in the colloquial language "the personal pronoun is often used in this position without any nuance of special insistence" (p. 136) I don't think this is true - at least for my consultants.
(b) bijan shomâ-rodid-Ø
Bijan you+râ saw+3sgS
'Bijan saw you.'

(5) (a) (ânhâ) xandid-an
they laughed+3plS
'They laughed.'

(b) (shomâ) xabid-in
you slept+2plS
'You slept.'

The agreement affixes are obligatory in all tenses and do not affect stress placement. Thus, with a regular verb, stress falls on the last syllable of the past stem. This is shown in (6) below (with stress indicated by boldface):

(6) xâbid-am 'I slept'
    xâbid-im 'We slept'
    xâbid-i 'You (sg) slept'
    xâbid-id 'You (pl) slept'
    xâbid 'He slept'
    xâbid-and 'They slept'

According to our assumptions (see Chapter 1), agreement is inflectional in Persian, while tense is derivational.

Let us assume that subjects are base-generated in the specifier of a functional projection above VP. I will call this projection IP, for now. The position of subjects and the label for the functional projection above VP are discussed in Chapter 5, section 5.5. The verb acquires F-features pre-syntactically in the Lexicon-Syntax Interface (LSI, cf. Cummins and Roberge (1993, 1994), see also Chapter 1, section 1.2.2).4 It is inserted into the syntax

4Apart from the theoretical position taken here that the only affixes that head projections in the syntax are phrasal affixes (Chapter 1, section 1.2.3), there are two problems with
bearing these features which must then be checked in order for the derivation to be grammatical.

Chomsky (1993) proposes that agreement features on a verb are checked by features (which he calls L-features) borne by the functional head to which the verb moves. This movement can take place in the overt syntax or at LF depending on whether the L-features are strong or weak, respectively. He states that the functional head must also bear a different set of features to check the noun phrase in its specifier. If the functional head is Agr it checks the "NP-features" of the argument in its specifier and the F-features on the verb. If the functional head is T(ense) it checks the Case of the subject and the tense on the verb. Chomsky states that once the mediating role of the functional projection has been performed, it disappears.

I propose to do away with the intermediate step altogether. That is, I propose that the F-features on the verb in Persian must be checked against a noun phrase bearing the same features. This noun phrase must be in a local relation to the inflected verb, i.e. governed by the verb (or in its minimal domain in terms of Chomsky (1993)). The question is why this noun phrase has to be the subject of the clause. That is, in a structure such as (7) below, why does the verb move to I₀, and why is it the subject and not the object that checks the agreement features?

assuming that the agreement morphemes head projections (AgrPs) in Persian. First, if the morphemes are actually present in the syntax, head movement of the verb is required for their affixation. However, other movement operations such as VP-fronting (see Chapter 5, section 5.6.1.2) treat the verb as being within VP. Crucially, the verb undergoes these movement operations with the agreement affixes attached. Thus, determining the order of these movement rules becomes problematic. Second, I see no reason why the agreement affixes should be considered to be present in the syntax, while the pronominal enclitics are not so considered. That is, I favour treating both sets of morphemes in the same way. However, if both sets of morphemes are actually present in the syntax problems arise regarding their relative order (see section 4.2.2.2).
Note that we have not yet dealt with the question of how the subject gets nominative case. I suggest that the agreement affixes are checked by the argument that gets nominative case because they are the nominative case-assigners. Thus the checking of a subject agreement affix occurs when the affix assigns nominative case to a noun phrase that agrees in person and number features with the affix. The fact that the subject agreement affixes must be checked forces verb-movement even if the subject is pro, which does not need case (see section 4.2.2.3). That is, verb movement occurs to satisfy the the requirements of the affix that the verb bears and not because the subject needs case. This is consistent with Chomsky's (1993) principle of Greed, whereby an element only moves to satisfy its own morphological properties and not because some other element needs its properties satisfied.

If we eliminate the idea that agreement affixes must be checked by features in the projection to which the verb moves, and assume that they can be checked by assigning case to a noun phrase sharing the same person and number features which occurs in a local relation to the inflected verb, note that nothing forces the movement shown in (7). That is, in principle, the morphological requirements of the inflected verb should be satisfied if it is the direct object that receives nominative case and agrees with the agreement suffix on the verb.
However, in this case the derivation will be ruled out because the subject does not receive case.\textsuperscript{5}

Chomsky (1993) states that under a checking theory there is no need for the Case Filter to be satisfied at S-Structure. That is, like other morphological features, case must be checked somewhere for a grammatical derivation, however, it need not be checked before LF. Note that in order for a derivation to 'converge', i.e. to be successful, it must converge at both PF and LF. Therefore, since there is no evidence that verb movement takes place in the syntax, let us assume that it takes place at LF (see Chomsky's (1993) principle \textit{Procrastinate} based on the idea that LF movement is "cheaper" than overt movement). If the F-features on the verb do not match the features of the argument in [SPEC, IP] the derivation is ruled out. Similarly, if a noun phrase does not receive case at LF the derivation is ruled out. The agreement features on the verb are spelled-out at PF.

IP in (7), as a projection in which agreement is checked, might be called AgrP, in keeping with proposals made by Pollock (1989) and Chomsky (1991, 1993) (but cf. Iatridou (1990)). Since the agreement morphemes are not generated as the head of this projection but are simply checked here, in some sense the label is irrelevant. However, as we will see in the following section there are arguments against the pronominal enclitics being checked in an agreement projection. To keep the treatment of these two sets of morphemes somewhat parallel, I will continue to call the highest projection by the more neutral term, IP.

4.2 The Pronominal Enclitics

The pronominal enclitic paradigm is repeated below, with the colloquial pronunciations given in parentheses.

\textsuperscript{5}Presumably the fact that it is possible to have a successful derivation if the verb moves, overrides \textquoteleft Greed\textquoteright.\textsuperscript{5}
In this section we consider how the pronominal enclitics are checked. The various environments in which the enclitics occur is catalogued in 4.2.1. In section 4.2.2 a unified account of the distribution and checking of the pronominal enclitics is presented.

4.2.1 Where They Occur

4.2.1.1 As Possessors

As we saw in Chapter 3, section 3.5.4, the pronominal enclitics can be used as possessors, attached to the end of the Ezafe domain. Some examples are given below:

(9) (a) ketāb-e bozorg-et
    book+EZ big+2sg
    'your big book'

    (b) manzel-eshun
    house+3pl
    'their house'

The structure of a noun phrase with a DP or pronoun possessor (see Chapter 3) is given in (10). Since the enclitics can be construed as possessors, the default assumption is that they also occur in this structure.
4.2.1.2 As Objects of Prepositions

According to Lazard (1992) the enclitics may follow a preposition only in colloquial language. According to Thackston (1983) "[i]n literary Persian the enclitics do not serve as complements of the "true" prepositions (be 'to', az 'from', bâ 'with', bar 'over', and dar 'in')" (p. 72). Examples in which the pronominal enclitics replace a full pronoun linked to a nominal preposition by Ezafe are shown in (11b) and (12b) below. Examples of the enclitics attached to "true" prepositions follow in (13b) and (14b).

(11) (a) darbâre-ye u
        concerning+EZ him

(b) darbâre-ash
    concerning+3sg

(12) (a) barâ-ye man
        for+EZ me

(b) barâ-m
    for+1sg

(13) (a) az u ye so'âl-i porsid-am
        from him one question+indef asked+1sgS
    'I asked him a question.'
(b) **az-esh** ye so'âl-i porsid-am
from+3sg one question+indef asked+1sgS
'I asked him a question.'

(14) (a) ketâb-o be to dâd-am
book+râ to you(sg) gave+1sgS
'I gave the book to you.'

(b) ketâb-o be-t dâd-am
book+râ to+2sg gave+1sgS
'I gave the book to you.'

While there is no difference in meaning between the sentences with full pronouns and those with enclitics as the object of a preposition, only the full pronouns can be used for emphasis.

The structure of a nominal preposition and its DP or pronoun complement is given in (15a) (see section 2.6.1 for a discussion of nominal prepositions). The structure of a true preposition and its DP or pronoun complement is given in (15b). Again the default assumption is that the enclitics also occur in these structures.

(15) (a) \[ P_{nomP} \]
\[ \begin{array}{c} P_{nom} \\
\text{DP} \end{array} \]
(b) \[ PP \]
\[ \begin{array}{c} P \\
\text{DP} \end{array} \]

4.2.1.3 As Objects of Verbs

So far we have seen that a pronominal enclitic can replace a full pronoun as a possessor and as the object of a preposition. Since a full pronoun can also serve as the object of a verb,
we expect to find that the enclitics can as well. This prediction is borne out, and as in the other two cases there is no difference in meaning.

In the examples that follow two points should be noted. First, the full pronouns are obligatorily followed by -râ, the marker that marks presupposition/definiteness on direct objects (see Chapter 5). However, if the direct object is expressed by a pronominal enclitic, -râ does not appear. Second, a pronominal enclitic must appear outside the subject agreement affix.

(16) (a) un-o xarid-am
        it+râ bought+1sgS
        'I bought that.'

(b) xarid-am-esh (*xarid-esh-am)
bought+1sgS+3sg
        'I bought it.'

(17) (a) to-ro did-am
        you+râ saw+1sgS
        'I saw you.'

(b) did-am-et (*did-et-am)
saw+1sgS+2sg
        'I saw you.'

Let us assume for now that the direct object of a verb is base-generated as the complement of that verb. The base-generated structure for a simple transitive sentence is given in (18). Once again, the default assumption is that the enclitics serving as direct objects also occur in this structure.
This section deals only with pronominal enclitics on simple verbs. Their behaviour with compound verbs is discussed in detail in Chapter 5, section 5.6.

4.2.2 Analysis

There are at least three possible analyses of the pronominal enclitics. Two approaches are outlined briefly in the following two sections and shown to present serious problems for the syntactic theory assumed here. The third approach, in which the pronominal enclitics are checked by agreeing in person and number features with a noun phrase that they govern, is discussed in section 4.2.2.3.

4.2.2.1 Pronominal Enclitics as Arguments

One fact that any analysis must account for is the complementary distribution of pronominal enclitics and full noun phrases. A standard account of this would generate the
enclitics themselves in the relevant argument position. Specifically then, the enclitics would appear as sister to a verb or preposition, or in the specifier position of DP to be construed as a possessor. The enclitics would then move to adjoin to the head that governs them. This type of analysis has been proposed for clitics in Romance (see Kayne (1975)).

Such an analysis encounters immediate theoretical problems. First, a phrasal constituent (the DP argument containing the enclitic) is moving to adjoin to a head. This type of movement should be ruled out, as pointed out by Cummins and Roberge (1993), who discuss clitics in Romance languages. Further, no generalization can be made about the position in which the enclitic originates. As we have seen above, in some cases the enclitic originates in a specifier position (e.g. as a possessor); in others in a complement position (e.g. as an object of a preposition or verb). I therefore reject this approach on the basis of the complexity of the movement operation required.

4.2.2.2 The Pronominal Enclitics and AgrP

If the pronominal enclitics themselves are not the argument of the item on which they occur, then the relevant argument could be pro. That is, the enclitics are present when pro is a possessor or the object of a preposition or verb.

Now let us consider a simple transitive sentence with pro as both the subject and the direct object.

(19) \[ pro_i \text{ proj xarid-am}_i-\text{esh}_j \]
\[ \text{bought+1sgS+3sg} \]
'I bought it.'

Suppose that the pronominal enclitics are exactly analagous to the subject agreement suffixes. The verb thus acquires two sets of F-features at the LSI. We have already seen that the subject agreement morphemes must be checked in a Spec-Head relationship with the
subject *pro* and that this checking takes place in a functional projection above VP.\footnote{This will be motivated in Chapter 5, section 5.5.} The pronominal enclitics must therefore also be in a Spec-Head relationship with the *pro* they identify. Since the direct object is base-generated as the complement, not the specifier, of the verb, there must be a higher projection in which this checking takes place. The verb will move to the head of this projection, and the direct object will move to the specifier. The structure of a transitive sentence with these two projections above VP is shown in (20).

\[
(20) \quad \text{Agr2P}\]
\[
\quad \text{Agr2}'
\]
\[
\quad \text{Agr1P} \quad \text{Agr2}
\]
\[
\quad \text{Agr1}'
\]
\[
\quad \text{VP} \quad \text{Agr1}
\]
\[
\quad \text{V}'
\]
\[
\quad \text{V} \quad \text{verb}+F_i+F_j
\]

In (20) the two projections in which the subject agreement affixes and the pronominal enclitics are checked are both labelled Agr, for convenience. In fact, it has been proposed that there are two Agr projections in the clause (Chomsky (1991, 1993) where the two Agr projections are Agr$_S$ and Agr$_O$, and Johns (1992) where they are Agr$_V$ and Agr$_N$). However, let us now consider the problems. Let us suppose that Agr1P is where the pronominal enclitics are checked against the direct object, and Agr2P is where the subject agreement affixes are checked. This is shown in (21) below.\footnote{I am not adopting the VP-internal subject hypothesis, for reasons that are outlined in Chapter 5, section 5.5. However, the discussion in this section also holds if the subject originates in [SPEC, VP] and then moves to the specifier of a higher projection.} This derivation makes the incorrect...
prediction that the pronominal enclitics should occur closer to the verb stem than the subject agreement suffixes.\(^9\)

(21)

\[
\begin{array}{c}
\text{Agr2P} \\
\text{proi} \\
\text{Agr1P} \\
\text{Agr1'} \\
\text{VP} \\
V' \\
\text{proj} \\
\text{verb}+F_j+F_i
\end{array}
\]

If, on the other hand, the order of morphemes on the verb stem is taken to indicate that Agr2P is the projection in which the pronominal enclitics are checked, we get the derivation shown in (22).

\(^9\)That is, of course, if we assume something along the lines of the Mirror Principle (Baker (1985, 1988)).
As shown in (22), the direct object must move to a position higher than the subject. Such a derivation is explicitly ruled out by Chomsky (1993). Specifically, the direct object cannot move past [SPEC, VP] to [SPEC, Agr2P] without violating the "Shortest Movement" Condition.

The AgrP approach is similar in spirit to the analysis proposed by Johns (1992) for Inuktitut. However, putting aside the fact that Persian does not exhibit the evidence found in Inuktitut supporting such a proposal, let us consider the consequences of the AgrP approach for the other positions in which the pronominal enclitics occur. If pronominal enclitics must be checked in an agreement projection, and if (by monosemy, see Chapter 1, section 1.2.1) all instances of the pronominal enclitics are to be treated the same way, then an agreement
projection must also be present when a pronominal enclitic occurs as a possessor or as the object of a preposition. The relevant derivations under the AgrP approach are shown below, where (23a) shows a pronominal enclitic as a possessor and (23b) shows a pronominal enclitic as the object of a preposition.

(23)  (a)

In each case above, head movement from D and P to Agr is required, and the pro must move to [SPEC, AgrP]. This movement is motivated only by the desire to maintain the generalization that agreement affixes are checked in an agreement projection. Another difficulty with this approach is that nothing prevents a pronominal enclitic from occurring
along with a noun phrase. Thus, further stipulations would be required to block this configuration.

The assumption that the pronominal enclitics are checked in an agreement projection leads to an otherwise unmotivated proliferation of these projections throughout the Persian clause. It also requires ad-hoc applications of syntactic head movement. We now turn to a third alternative that preserves the advantages of the AgrP approach while eliminating the undesired projections.

4.2.2.3 Pronominal Enclitics as Licensees

Given that the pronominal enclitics and the subject agreement suffixes are both sets of F-features, let us assume that they are checked in a similar way. We saw that the subject agreement affixes are checked by assigning case to a noun phrase that they govern and with which they share person and number features. Let us suppose that the pronominal enclitics are also checked by agreeing in person and number features with a noun phrase that they govern. However, let us further suppose that unlike subject agreement the pronominal enclitics not only lack a case-assigning feature, but absorb case if attached to a lexical item that can assign case. The idea that clitics absorb case has been proposed by Roberge (1990). Thus only a noun phrase that (a) does not require case or (b) can receive case in some other way will be able to check the pronominal enclitics. Let us consider these two cases in turn.

We have seen that the pronominal enclitics can occur in noun phrases identifying a possessor, on prepositions identifying the object of the preposition, and on verbs identifying the direct object. If the enclitics must be checked by a noun phrase that they govern, yet absorb the available case it follows that the only noun phrase that can appear must contain

11Note that this predicts that a noun phrase in the specifier or the complement position of a projection can check a pronominal enclitic occurring on the head of the projection. In this we diverge from Chomsky (1993) who allows checking only between a specifier and a head.
Note that I am assuming, following Roberge (1990) that pro does not need case. The relevant configurations for the pronominal enclitics as a possessor, and the objects of a preposition or verb are given below:

(24) (a) DP
    /\   /\  
   D'  DP_{poss}  proj_{i}  preposition+F_{i}  proj_{j}  
  /\   /\     /\     /\     /\  
 NP   D    N   O_{def}+F_{i}  
 N O

(c) IP
    /\   /\  
   IP'  I'   
  /\   /\     /\  
 DP   VP    I
    /\     /\     /\  
 proj_{i}  verb+F_{i}+F_{j}  proj_{j}  
  /\     /\     /\  
 V

Given that pro in Persian can be a possessor, the object of a preposition, and the object of a verb, in addition to occurring in subject position, we have to rule out its occurrence in these positions when a pronominal enclitic does not occur. Roberge (1990) proposes the following principle:

(25) **Recoverability Principle**
    An empty category must be licensed.       [Roberge 1990:8]

As an empty category then, pro is subject to this principle. Roberge argues that just like languages in which rich agreement licences pro, clitics can also be licensing elements for
pro. That is, pro is licensed if it is coindexed with and governed by a clitic. He notes that in most cases clitics have the same intrinsic features as agreement markers in the same language.

In a sense then, there is a mutual dependence between pro and the pronominal enclitics. On the one hand pro is contained in a noun phrase against which the enclitics can be checked for a successful derivation. On the other hand, the enclitics serve to license the occurrence of pro in an argument position. The necessity for a pro in the argument position that the pronominal enclitics govern will become clearer when we turn to binding in Persian (see section 4.4).

As stated above, the pronominal enclitics can also be checked by a full noun phrase if this noun phrase can receive case from some source other than the item to which the enclitics are attached. While such examples are rare, there is one case discussed by Lazard (1992) that illustrates this phenomenon. Lazard notes that a pronominal enclitic can appear on a verb to indicate the subject. He notes that this is frequent in colloquial language and "sometimes indicates a slight insistence on the identity of the subject." (p. 112) Some examples taken from Lazard (1992:112) are given in (26) and (27).

(26) mosyo C ... mi-ād-Ø-esh
mister C ... cont+come+3sgS+3sg
'Mr. C...is coming.'

(27) qadd-e kutā-i dāsht-Ø-esh
height+EZ short+indef had+3sgS+3sg
'He was of small stature.'

In these examples the subject can receive nominative case from the agreement on the verb. However, a pronominal enclitic can also appear. It is checked by the subject when the verb moves to I0 at LF. Note that the occurrence of both an agreement suffix and an enclitic that
are coindexed should only be possible with intransitive verbs, given the Reflexivity Condition (see section 4.4.3). This prediction is borne out by the examples Lazard provides.¹²

There is another argument in favour of the analysis proposed here, in which the pronominal enclitics appear on a head but are not syntactic heads themselves. Note that there is a problem if we consider the pronominal enclitics to actually be the head of DP.¹³ This predicts that just like pronouns, which are also D-heads (see Chapter 3, section 3.5.1), the F-features of the pronominal enclitics should percolate up, triggering agreement. This is illustrated in (28) below. In (28a) we see that a first person pronoun in subject position triggers first person agreement on the verb. However, a first person pronominal enclitic attached to a noun in subject position does not trigger first person agreement on the verb, as shown in (28b). In (28c) we see that a third person singular noun appearing with a first person singular pronominal enclitic triggers third person singular agreement on the verb.

(28)  
(a) man oftâd-am  
I    fell+1sgS  
'I fell.'  

(b) *ketâb-am oftâd-am  
book+1sg fell+1sgS  
'My book fell.'  

(c) ketâb-am oftâd-Ø  
book+1sg fell+3sgS  
'My book fell.'

¹²The double occurrence of agreement and a pronominal enclitic only occurs in the third person according to Lazard. I will leave aside an exploration of why this is the case, though it may be linked to the fact that third person agreement is null. The relevance of these examples to the analysis of the pronominal enclitics still holds.

¹³Thanks to Elizabeth Ritter for pointing this out to me.
Given that the features of a D-head percolate up to the DP to trigger agreement (as in the case of full pronouns), the fact that the pronominal enclitics do not trigger agreement means that they cannot be the head of D. This then supports the idea that the pronominal enclitics are the spell-out of F-features on the head of a projection, but are not the head of the projection themselves.

4.3 A Comparison of Subject Agreement and the Pronominal Enclitics
4.3.1 Subcategorization and Constraints on Spell Out

We have seen that both the subject agreement suffixes and the pronominal enclitics are represented syntactically as F-features on heads. We have also seen that they differ in one morphological property. The subject agreement suffixes assign case while the pronominal enclitics absorb case. However, if this were the only difference between them we would expect to find that both sets of morphemes could appear on the same lexical items. This is not so. The only lexical item on which both sets of morphemes can appear is the verb, and in this case they must appear in a particular order (pronominal enclitics following subject agreement).

The difference in the distribution of the subject agreement affixes and the pronominal enclitics can be characterized informally in the following way: the subject agreement suffixes are the spell out of F-features attached directly to a verb stem, while the pronominal enclitics are the spell out of F-features attached elsewhere. That is, the subject agreement suffixes must attach directly to verbs while the pronominal enclitics cannot do so. I propose to capture this generalization, at least in part, by using the idea of subcategorization. Let us say that the subject agreement suffixes subcategorize a verb. Therefore, at the LSI they will only attach to verb stems.
We are still left with the task of accounting for the distribution of the pronominal enclitics. The distribution of this set of morphemes is not easily captured by subcategorization, given that they can attach to any head except a verb stem. That is, the items to which the pronominal enclitics attach do not form a natural class. Let us suppose then that unlike the subject agreement suffixes, the pronominal enclitics do not have a subcategorization frame. Thus, any item that passes through the LSI can be inflected with these morphemes as long as they are checked by a noun phrase that they govern.

Let us consider again a case where both a subject agreement suffix and a pronominal enclitic appear on a verb. Such an example is given in (29) showing that the subject agreement suffix must precede the pronominal enclitic.

(29) xarid-am-esh (*xarid-esh-am)
    bought+1sgS+3sg
    'I bought it.'

It is not clear whether the ungrammaticality of *xarid-esh-am is due to the fact that the pronominal enclitic is attached directly to a verb stem, or whether the presence of this enclitic renders the verb stem 'opaque' for affixation of subject agreement. If we consider examples involving complex verb forms the source of the ungrammaticality becomes clearer.

In complex tenses, an auxiliary verb is usually used with a participial form of the main verb.14 In these cases the pronominal enclitics cannot occur on the participle but must follow

---

14In what follows I assume that the participial form of the verb (formed by adding the stressed suffix -e to the past verb stem) shares the same categorial features as a regular verb ([−N, +V]). However, this particular form can also occur as an adjective within noun phrases (see next footnote). In its adjectival use, the participle can take Êzafe (suggesting that it bears the feature [+N] and can also cooccur with the pronominal enclitics. It seems then that these two uses of the participle involve different feature specifications, yet, under the assumption of Monosemy, it is desirable to posit only one participial form. It may be possible to resolve this problem by developing a system whereby some of the feature specifications of lexical items come from the syntactic context in which they occur. I leave this as an area to be pursued in future research.
subject agreement on the auxiliary. It should be noted that even the grammatical examples sound odd, which can be attributed to the fact that in the formal register in which these tenses are used the enclitics are generally avoided.

\[(30)\] (a) sib-o xaride bud-am
apple+râ bought was+1sgS
'I had bought the apple.'

(b) xaride bud-am-esh
bought was+1sgS+3sg
'I had bought it.'

(c) *xaride-y-esh bud-am
bought+3sg was+1sgS

\[(31)\] (a) qazâ-ro suxte bud-am
dinner+râ burned was+1sgS
'I had burned the dinner.'

(b) suxte bud-am-esh
burned was+1sgS+3sg
'I had burned it.'

(c) *suxte-ash bud-am\(^{15}\)

\(^{15}\)When speakers are asked to assign the (c) sentences some sort of interpretation, it is interesting that they assign a nominal reading to the participle. That is, the sentence lends itself to being translated as 'I was its bought/shopping,' or 'I was its burnt.' A related point is that, while participles cannot occur with the pronominal enclitics attached when they occur as verbs, they can when they occur as adjectives. This is shown in the example below:

(i) qazâ-ye suxte
dinner+EZ burned
'the burnt dinner'

(ii) qazâ-ye suxt-ash
dinner+EZ burned+3sg
'his/her burnt dinner'
burned+3sg was+1sgS

In the future tense, formed with the verb xâstan 'to want', the auxiliary occurs before the bare verb stem. This future tense is highly formal, the usual way of indicating the future being to use the present tense. In the formal future there is a preference for the enclitics to occur on the auxiliary, although again, the enclitics are generally not used in the formal register.

(32) (a) ketâb-o xâh-am xarid
        book+râ want+1sgS buy
      'I will buy the book.'

      (b) ?xâh-am-esh xarid
          want+1sgS+3sg buy
      'I will buy it.'

      (c) ??xâh-am xarid-esh
          want+1sgS buy+3sg

The examples above show that regardless of where the subject agreement suffix appears within a complex verb form, a pronominal enclitic cannot appear directly on a verb stem. A pronominal enclitic can appear following a subject agreement suffix, however. This can be explained if we assume that affixes are also specified for categorial features. I propose that subject agreement affixes be specified as [+N]. The idea that agreement is nominal has been proposed before. For example Chomsky (1981) comparing AGR to PRO, suggests that it is

---

16See Cowper (1995), for example, who states that the participial affix-\textit{en} in English which is involved in the passive construction, the perfect construction, and adjectival formation, is specified for the feature [+N].
composed of the features [+N,-V]. See also Ritter (1995) who states that the syntactic category of agreement is the same as pronouns.

In terms of how to account for the distribution of the pronominal enclitics, as stated above, there is no generalization that can be made about the class of items on which the enclitics can appear. Thus, rather than having a subcategorization requirement that must be satisfied at the LSI, I suggest that the pronominal enclitics can be freely attached to any head but are subject to a PF filter that rules out their occurrence directly on a verb stem. This filter is given in (33).

(33) **The PF Filter on the Spell Out of Pronominal Enclitics**

*X+pronominal enclitic, where X bears the features [-N, +V]*

To summarize, the subject agreement suffixes and the pronominal enclitics differ in two ways. First, the subject agreement affixes assign case while the pronominal enclitics absorb case. Second, the subject agreement suffixes subcategorize verbs at the LSI while the pronominal enclitics can attach to any head at this level but are subject to the PF filter shown in (33).

A point of interest here is the observation that in the formal (written) language pronominal enclitics are not found on prepositions (Lazard (1992)), or at least not on the true prepositions (Thackston (1983)). Since formal and written language tends to be more conservative than colloquial language these observations may reflect an earlier stage of Modern Persian in which the pronominal enclitics were subcategorized for [+N] elements only.

There is another constraint on the spell-out of subject agreement and the pronominal enclitics, ruling out forms that are potentially ambiguous. Hashemipour (1989) points out two cases where ambiguity may arise, stating that in these cases the enclitics are not used. Both involve the first person singular enclitic form *am*. First, since the initial vowel of a
clitic is elided when it is preceded by another vowel, the first person singular enclitic form (-am) preceded by the second person singular agreement form (-i) will be identical to a verb marked simply for first plural agreement (-im). This is shown below.

\[(34) \quad (a) \quad \text{did-i-am} \rightarrow \text{did-i-m} \\
    \text{saw+2sgS+1sg} \\
    'You saw me.' \\

    (b) \quad \text{did-im} \\
    \text{saw+1plS} \\
    'We saw.' \]

Second, the fact that third singular agreement in the past tense is null, while the first singular enclitic and agreement forms are identical (-am) could also give rise to ambiguity, as shown below.

\[(35) \quad (a) \quad \text{did-Ø-am} \\
    \text{saw+3sgS+1sg} \\
    'He saw me.' \\

    (b) \quad \text{did-am} \\
    \text{saw+1sgS} \\
    'I saw.' \]

While in both (34) and (35) the (a) forms are not ungrammatical, as Hashemipour states, they are simply not used. As noted above, this suggests that there is a PF constraint against spelling out the pronominal enclitics in cases where an ambiguous form is created.

Finally, if the agreement affix and the enclitic bear identical non-third person and number features, they cannot cooccur on the verb.
This fact is not due to a spell-out constraint, however, but rather follows from the Reflexivity Condition that will be proposed in section 4.4.3.

4.3.2 Affixes vs. Clitics

In the treatment proposed here the subject agreement affixes and the pronominal enclitics both spell out F-features at PF. Both sets of morphemes must be checked by sharing person and number features with a noun phrase that they govern. Thus both sets of morphemes can be viewed as agreement. However, in previous work on Persian the former paradigm is treated as the manifestation of agreement, while the latter paradigm is treated as the manifestation of cliticization. The difference between agreement and cliticization is notoriously elusive (see Roberge (1990, section 4.2) and Spencer (1991, section 9.3) for some discussion of this distinction). In this section we review the differences between the two paradigms in Persian. However, whether these differences are sufficient to motivate a clear theoretical distinction between the two is a question that will be left unanswered pending further evidence from other languages.

As mentioned above, subject agreement is generally considered to be an instance of inflectional affixation while the pronominal enclitics are considered to be clitics (e.g. Samiian (1983), Karimi (1989), Hashemipour (1989), Browning and E. Karimi (1994)). Hashemipour gives three reasons for treating the pronominal enclitics as clitics rather than inflectional affixes. First she claims that their position is best described in terms of syntactic constituents rather than roots or stems. Second she says that unlike the agreement
morphemes, the enclitics are optional when attached to verbs. Third, she says they show a low degree of selection with respect to their host (citing Zwicky and Pullum (1983) in this regard).17

An additional point that Hashemipour does not make, but which could be used as further evidence that the pronominal enclitics are not inflectional affixes, is their order with respect to subject agreement. Since there is a strong cross-linguistic tendency for object agreement to occur 'inside' subject agreement, it can be argued that the pronominal enclitics in Persian do not behave like true object agreement markers and instead should be treated as clitics.

Let us consider Hashemipour's arguments in turn. First, note that under our analysis there is no difference in the structural position of the two paradigms. While we have seen that there are phrasal affixes in Persian (such as the indefinite enclitic -i, discussed in Chapter 3 and also the direct object marker -rā which will be discussed in Chapter 5) the morphemes in both of the paradigms under discussion here are realized on heads, not on phrases.

We have seen that the pronominal enclitics do indeed have a lower degree of selection with respect to their host. Specifically, we have seen that the subject agreement affixes must subcategorize a verb while the pronominal enclitics can attach to any head except directly to a verb stem. It is this difference between them that accounts for the relative order of subject agreement and enclitics on verbs.

Finally, it is not true that the enclitics are 'more optional' than subject agreement. We have seen that they are obligatory in order to license a null argument of a verb or a preposition or a null possessor. What distinguishes the enclitics from the agreement markers, however, is that they absorb case while the subject agreement markers assign case. Thus,

17She gives a fourth reason: "...we are in agreement with Heny 1986 that the placement of the clitic is a structural process, i.e. occurring at S-structure, resulting in the coindexation of the clitic with some structural position." (p. 33) The work she cites is an unpublished conference paper which I have not seen. While, this fourth argument is rather theory internal, it seems to be compatible in spirit with the analysis that has been developed here.
while the subject agreement markers can appear with any noun phrase, null or overt, definite or indefinite, the pronominal enclitics can only appear with a noun phrase that does not need case or can receive case independently. Most often then, the pronominal enclitics will appear with pro since pro does not need case.¹⁸

To summarize, the similarity between the subject agreement and pronominal enclitic paradigms is that they are both collections of F-features that must be checked against a noun phrase with the same person and number features. The difference between them is that one paradigm assigns case and has a subcategorization requirement. The other paradigm absorbs case and has no subcategorization requirement but is subject to certain filters at PF. There are many questions that are raised at this point (for example: Do all affixes either assign or absorb case or are there affixes that are neutral with respect to case? How loose can a subcategorization frame be for an affix?) Among these questions is the one that has run through this section: what is the difference between an agreement affix and a clitic? Whether subcategorization requirements and case properties are what, in general, distinguishes affixes from clitics is among the questions left for further research.

4.4 The Binding of Pronominals and Anaphors

Having provided a structural analysis of the pronominal enclitics, we now turn to their binding properties. Pronominal enclitics behave quite differently from independent pronouns¹⁹ with respect to binding. To account for their behaviour we will first look at

¹⁸Since the enclitics on verbs can only appear with pro as the direct object, or with a direct object marked with -râ (see Chapter 5), this explains their apparent sensitivity to the definiteness of the direct object in contrast with the subject agreement markers. Note that pro as a pronominal element is definite. As we will see in Chapter 5, DPs marked with -râ are also definite or presupposed. Thus we have a syntactic explanation for an otherwise mysterious semantic difference between the two paradigms.

¹⁹See Chapter 3, section 3.5.1 for a chart giving the one set of independent pronouns and a brief discussion of this set.
binding in Modern Persian more generally. We must also account for the properties of the reflexive/emphatic element *xod* 'self', which can occur alone or inflected with the pronominal enclitics.

The binding properties of all these elements can be accounted for in a relatively simple way provided a few basic assumptions are made. We will see that the pronominal enclitics are coindexed with and license *pro* in an argument position that they govern. Thus, the discussion here will provide support for the analysis of pronominal enclitics developed in the first half of this chapter.

Let us first outline the various pronominal elements under discussion. The following chart is given by Hashemipour (1989, p. 36) to show the entire set of third person pronominal elements that can occur in Modern Persian.

(37)

<table>
<thead>
<tr>
<th>Third Person Subject Pronouns</th>
<th>Third Person Object/Genitive Pronouns</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Singular</strong></td>
<td><strong>Plural</strong></td>
</tr>
<tr>
<td><em>pro</em></td>
<td><em>pro</em></td>
</tr>
<tr>
<td><em>u</em></td>
<td><em>ishân</em></td>
</tr>
<tr>
<td><em>un</em></td>
<td>_ unhâ_</td>
</tr>
<tr>
<td><em>in</em></td>
<td>_ inhâ_</td>
</tr>
<tr>
<td><em>taraf</em></td>
<td><em>taraf</em></td>
</tr>
<tr>
<td><em>xodesh</em></td>
<td><em>xodeshun</em></td>
</tr>
</tbody>
</table>

We can note that apart from the first line, the two halves of the chart are identical. That is, according to Hashemipour, the same forms are used in subject position and in object/genitive position, except for *pro* which only occurs in subject position and the pronominal enclitics which only occur in non-subject position. If, however, the pronominal enclitics are taken to license *pro* in non-subject position, then the apparent complementary distribution between *pro* and the enclitics disappears.
In the second line of Hashemipour's chart above, we see *u* and *ishân*, which she calls the formal Persian pronouns because they do not often appear in colloquial speech. The next two lines contain the pronouns *un/in* and their plural counterparts. These pronouns also function as deictic elements (*this/that*) and as demonstrative determiners. While Hashemipour considers them together, for my consultants *un* seems to pattern with *u* and *in* is used as the deictic pronoun. Next there is *taraf*, a pronoun that literally means 'side' and seems to be favoured when overt mention of someone's name would be considered impolite. Finally, the reflexive/emphatic elements are given in the last line of the chart. (Recall that *xod* can also appear alone as we saw in Chapter 3, section 3.4.1.)

In this section we compare and contrast the binding properties of *pro*, the independent pronouns, which in third person appear as *u/un*, and the reflexive/emphatic element *xod* in both its inflected and non-inflected forms. We do not consider the deictic pronouns *in* and *taraf*. Hashemipour (1989) states that *in* and *taraf* have as their binding domain a larger constituent than the sentence they appear in. The reader is referred to Hashemipour (1989), particularly Chapter 2, for discussion of these elements.

The version of Binding Theory that we start with is taken from Chomsky (1981) and involves the three conditions given in (38). This will be referred to as classic Binding Theory,

\[
\text{(38) Binding Theory} \\
\text{(A) An anaphor is bound in its governing category} \\
\text{(B) A pronominal is free in its governing category} \\
\text{(C) An R-expression is free} \quad \quad \quad \quad \quad \quad \quad \quad \quad \text{[Chomsky 1981:188.12]}
\]

Pronominals include both pronouns and *pro*. R-expressions include proper names and common nouns. As for anaphors, I assume for now that an anaphor is defined as an overt element that must be bound by a pronominal or an R-expression to receive its referential
content. I do not consider NP or wh-traces, nor wh-elements in general in the discussion that follows.

I assume the following definition of governing category:

(39) \(b\) is a governing category for \(a\) if and only if \(b\) is the minimal category containing \(a\), a governor of \(a\), and a SUBJECT accessible to \(a\). \[Chomsky 1981:211.70 (II)\]

Chomsky (1981:209) defines a SUBJECT as being (a) the subject of an infinitive, an NP, or a small clause or (b) AGR in INFL. The idea in the latter case is that AGR, when present, serves as a kind of antecedent to the NP subject.\(^{20}\) In fact Chomsky gives the following statement:

(40) \(\text{AGR is coindexed with the NP it governs}\) \[Chomsky 1981:211.70 (I)\]

Let us assume that in Persian it is the F-features on the verb that serve as the accessible subject in creating a governing category for the subject of a clause.\(^{21}\) This view has a very interesting consequence. If the presence of F-features that are coindexed with a noun phrase and appear on the head that governs that noun phrase count as an accessible subject, then the pronominal enclitics also count as accessible subjects. That is, if the governing category for \(pro\) in subject position is IP, given that agreement on the verb counts as an accessible subject,

\(^{20}\)Note that AGR cannot bind an anaphor in subject position. Chomsky simply stipulates (1981:212) that while AGR creates a governing category it cannot serve as a binder.

\(^{21}\)Since it was claimed earlier in the chapter that verb movement to I\(^0\) does not take place until LF, this entails that Binding Theory holds at LF also. On the other hand, if Binding Theory can be shown to hold prior to LF in Persian then the verb must move in the overt syntax. This issue merits further research, however, in what follows it is sufficient to assume that both Binding Theory and verb movement hold at the same level. That is, the analysis here is not affected by the choice of overt syntax vs. LF as the level at which these conditions hold.
then the governing category for \textit{pro} in non-subject position will be the maximal projection in which it occurs, since the pronominal enclitics must appear on the governing head to license \textit{pro}. As we will see, this result accounts nicely for the binding properties of \textit{pro}.

4.4.1 The Binding Properties of Pronouns vs. \textit{pro}

4.4.1.1 Binding and \textit{pro} Subjects

\textit{Hashemipour} (1989) states that \textit{pro} in Persian is subject to Binding Principle B which is to be expected as \textit{pro} is a pronominal element. Accordingly, \textit{pro} must be free in its governing category, but may refer to nominals outside its governing category. \textit{Hashemipour} claims, however, that in the latter case \textit{pro} can refer to only a c-commanding nominal outside of the clause that contains it. She gives the following examples:

(41) (a) moin\textsubscript{i} goft [CP ke [IP pro\textsubscript{i/j} ne-mi-âd-Ø]]
Moin\textsubscript{i} said [CP that [IP pro\textsubscript{i/j} neg+cont+come+3sgS]]
Moin\textsubscript{i} said [CP that [IP he\textsubscript{i/j} is not coming]]

(b) moin\textsubscript{i} goft [CP ke [IP nasser\textsubscript{j} fekr mi-kon-e]
Moin\textsubscript{i} said [CP that [IP Nasser\textsubscript{j} think cont+do+3sgS

[CP ke [IP pro\textsubscript{i/j/k} ne-mi-âd-Ø]]]]
[CP that [IP pro\textsubscript{i/j/k} neg+cont+come+3sgS]]]]
Moin\textsubscript{i} said [CP that [IP Nasser\textsubscript{j} thinks [CP that [IP he\textsubscript{i/j/k} is not coming]]

(c) shâgerd\textsubscript{i} az mo'allemporsid-Ø [CP agar [IP pro\textsubscript{i/j/k}
student\textsubscript{i} from teacher asked+3sgS[CP if [IP pro\textsubscript{i/j/k}

mi-tun-e emruz [boland pro\textsubscript{i/j/k} be-xun-e]]
cont+can+3sgS today [loud pro\textsubscript{i/j/k} subj+read+3sgS]]
The student\textsubscript{i} asked the teacher\textsubscript{j}[CP if [IP he\textsubscript{i/j/k} would read out loud today]]

[Hashemipour 1989:152.93-95]
Obviously Hashemipour's claim that *pro* can be coindexed only with a c-commanding nominal outside of the clause that contains it, entails that *pro* **cannot** be coindexed with a nominal that does not c-command it. She gives the following examples to illustrate this:

(42) (a) \(\text{moin}_i\) be \(\text{pesar-esh}_j\) tozih \(\text{dâd-Ø}\) ke chera \(\text{pro}_i*/j\,?/1\) mariz-\(e\)
Moin\(_i\) to son+3sg\(_j\) explain gave+3sg\(_S\) that why \(\text{pro}_i*/j\,?/1\) sick+3sg\(_S\)
Moin\(_i\) explained to his son\(_j\) why he\(_i*/j\,?/1\) is sick.

(b) \(\text{moin}_i\) berais-esh\(_j\) yâdâvari kard-\(Ø\) [ke diruz
Moin\(_i\) to boss+3sg\(_j\) remind did+3sg\(_S\) [that yesterday

\(\text{pro}_i*/j\,?/k\) kâr-\(o\) tamom kard-\(Ø\)]
\(\text{pro}_i*/j\,?/k\) work+\(râ\) finish did+3sg\(_S\)]

Moin\(_i\) reminded his boss\(_j\) that he\(_i*/j\,?/k\) had finished the work yesterday.

[Hashemipour 1989:43.38a-b]

Let us consider these examples in some detail. In fact, there are more possibilities for reference than the subscripting above shows. It seems that Hashemipour intends the subscript \(j\) to belong to 'son' in (42a) and 'boss' in (42b). In fact, just as in the English translation it is possible for *pro* to be coindexed with these two noun phrases in Persian. There are even further possibilities if we take into account the referent of -esh 'his' in the two sentences. The possessor can be coreferential with the matrix subject, in which case, as indicated by the subscripting, the *pro* embedded subject can be coreferential with both the matrix subject and the possessor. However, the possessor can also refer to a third person. For example, supposed we have been discussing John's illness at some length. In such a case, sentence (42a) can be used to indicate that Moin explained to John's son that John is sick. Again, this is true of the English example also.
While it is true that *pro* can be coindexed with a either a c-commanding or non-c-commanding nominal outside of its governing category, this is not a binding relation. In fact, *pro* behaves exactly like a pronoun should. As long as it is free within its governing category, it can be coindexed with any other nominal, under the right discourse conditions. Further, I suggest that the apparent preference for coindexing an embedded *pro* subject with the matrix subject is a strategy used for 'out-of-the-blue' contexts.

With respect to embedded subjects there is another issue that should be mentioned here. Recall that overt pronouns are normally used in subject position in Persian only for emphasis. Hashemipour (1989) claims that this null pronoun strategy is "more prominent" for first and second person. That is, she states that the use of an overt third person pronoun does not necessarily signal emphasis, as do the first and second person pronouns. However, the examples she gives all involve embedded sentences. The use of an overt pronominal subject in an embedded sentence signals a change in reference. An example is given below:22

(43) (a) moini goft-Ø ke *pro* i/?j mi-âd-Ø
Moini said+3sgS that *pro* i/?j cont+come+3sgS
'Moini said that he*i/j is coming.'

(b) moini goft-Ø ke *un*i/j mi-âd-Ø
Moini said+3sgS that he*i/j cont+come+3sgS
'Moini said that he*i/j is coming.' [Hashemipour 1989:30.17a-b]

Now, with first and second person pronouns, a change in reference is not possible between a matrix and embedded subject and, unsurprisingly, an overt pronominal subject is ruled out. This is shown below:

---

22 I have included the '?' Hashemipour puts on the second index in (43a). However, as already mentioned, while it is true that in an out-of-the-blue context there is a tendency to favour coreference with the matrix subject in such an example, the second reading is readily available if the context is set up to favour it.
(44) (a)  (mani) goft-am   ke  \textit{pro}_{i/*j} \textit{mi-âm}  \\
I_{i} \text{ said+1sgS that } \textit{pro}_{i/*j} \text{ cont+come+1sgS}  \\
'I said that I'm coming.'

(b)  *(mani) goft-am   ke  man  \textit{mi-âm}  \\
I \text{ said+1sgS that } \text{ I } \text{ cont+come+1sgS}  \\
'I said that I'm coming.'

(45) (a)  (toi) goft-i     ke  \textit{pro}_{i/*j} \textit{mi-ây}  \\
You_{i} \text{ said+2sgS that } \textit{pro}_{i/*j} \text{ cont+come+2sgS}  \\
'You said that you're coming.'

(b)  *(to)  goft-i     ke   to   \textit{mi-ây}  \\
You \text{ said+2sgS that } \text{ you } \text{ cont+come+2sgS}  \\
'You said that you're coming.'

I propose that the availability of a \textit{pro} subject in an embedded clause renders marked the use of an overt pronoun. This may be derivable from the Avoid Pronoun principle (Chomsky 1981:65.5). This accounts for the ungrammaticality of (44b) and (45b) and also explains why (43b) is only grammatical if the pronoun is not coreferential with the matrix subject. Further, the fact that disjoint reference is not possible in first and second person between an embedded \textit{pro} subject licensed by the same person and number features as a matrix subject is not a syntactic fact but rather a 'real world' fact. Given this then, I see no reason to treat third person as differently from first and second person.

\footnote{Chomsky formulates this principle as "Avoid Pronoun" thus creating perhaps the only principle that is identical to its title. He states that "[the Avoid Pronoun principle] might be regarded as a subcase of a conversational principle of not saying more than is required, or might be related to a principle of deletion-up-to-recoverability, but there is some reason to believe that it functions as a principle of grammar." (1981:65)}
4.4.1.2 Binding and Possessor *pro*

The following examples illustrate the difference between pronouns and the pronominal enclitics as possessors. In the first two examples we see that a full pronoun possessor of an object is not possible if the possessor has the same person and number features as the subject. A pronominal enclitic is, however, possible as a possessor in these cases. In the third example we see that if both the subject and a full pronoun possessor are third person, the sentence is grammatical only if they are not coindexed.

(46) (a) jiân ketâb-e man-o xund-Ø
    Jian book+EZ me+râ read+3sgS
    'Jian read my book.'

    (b) *(man) ketâb-e man-o xund-am
        (I)    book+EZ me+râ read+1sgS
        'I read my book.'

    (c) (man) ketâb-am-o xund-am
        (I)    book+1sg+râ read+1sgS
        'I read my book.'

(47) (a) jiân ketâb-e shomâ-ro xund-Ø
    Jian book+EZ you+râ read+3sgS
    'Jian read your book.'

    (b) *(shomâ) ketâb-e shomâ-ro xund-id
        (you)    book+EZ you+râ read+2plS
        'You read your book.'

    (c) (shomâ) ketâb-etun-o xund-id
        (you)    book+2pl+râ read+2plS
        'You read your book.'
The fact that independent pronoun possessors cannot be coindexed with a subject in the same clause follows from Condition B of the Binding Theory. That is, since a pronoun cannot be bound within its governing category, and since the governing category for these possessors is the whole sentence, it follows that they cannot be coindexed with the subject. The mystery here is why the corresponding pronominal enclitics do not yield a Condition B violation. This mystery is explained if the governing category for the pro possessor is the direct object DP itself. That is, with pro as possessor the pronominal enclitics appear as the spell-out of F-features on the D-head and can serve as accessible subjects. On the other hand with a full pronoun as possessor, the head of DP is simply the null definite determiner (see Chapter 3, section 3.4).

In Persian, even when there is no c-command relation between a pronoun (including pro) and a proper name, and both are thus free within their governing categories, the pronoun must follow the proper name. This is shown in (49).

\[(49)\] (a) *mâdar-e uî ali-ro dust dâr-e
mother+EZ him Ali+râ likes+3sgS
'Hisî mother likes Ali.'

(b) *mâdar-eshî ali-ro dust dâr-e
mother+3sg Ali+râ likes+3sgS
'Hisî mother likes Ali.'
If, however, Ali has been under discussion, (49b) is more acceptable, consistent with the discourse anaphoric nature of \textit{pro}. The same facts hold of the following examples, where again there is no c-command relation between the pronominal possessor and the proper name in the prepositional phrase.

(50) (a) \textit{jiâni ketâb-e u*i/*j/*k- ro az alij xâst-Ø}  
\textit{Jian book+EZ him+râ from Ali wanted+3sgS}  
'Jiani wanted his*i/*j/*k book from Alij.'

(b) \textit{jiâni ketâb-eshi/'j/*k-o az alij xâst-Ø}  
\textit{Jiân book+3sg+râ from Ali wanted+3sgS}  
'Jiani wanted hisi/'j/*k book from Alij.'

Again, note that in (50a) the pronoun possessor cannot be coindexed with the subject, while in (50b) the \textit{pro} possessor can be, following from the fact that in (50b) the DP is the governing category for \textit{pro}. In both cases coindexing of the possessor with Ali is highly marked. This is parallel with the sentences in (49), in which the pronominal precedes the proper name.

Finally note that if the prepositional phrase is scrambled to the left of the direct object in the sentences in (50), coreference is possible between the proper name and the possessor, shown in (51). This is to be expected if the relevant constraint refers simply to the order of mention of coreferential elements.

(51) (a) \textit{jiâni az alij ketâb-e u*i/*j/*k- ro xâst-Ø}  
\textit{Jian from Ali book+EZ him+râ wanted+3sgS}  
'Jiani wanted his*i/*j/*k book from Alij.'

(b) \textit{jiâni az alij ketâb-eshi/'j/*k-o xâst-Ø}  
\textit{Jiân from Ali book+3sg+râ wanted+3sgS}  
'Jiani wanted hisi/'j/*k book from Alij.'
4.4.1.3 Binding and Objects

Let us now compare sentences with pronouns as the object of a verb or preposition with sentences in which the pronominal enclitics identify the same positions. Some examples are given below. In the (a) sentences a full pronoun is used as the object of a preposition (52-53) and as the direct object of a verb (54-55). In the (b) sentences the pronoun is replaced by a pronominal enclitic.

(52) (a) jiânì be u*i/j telefon zad-Ø
    Jian to him telephone hit+3sgS
    'Jianî telephoned (to) him*i/j.'

    (b) jiânì be-sh*i/j telefon zad-Ø
    Jian to+3sg telephone hit+3sgS
    'Jianî telephoned (to) him*i/j.'

(53) (a) jiânì az u*i/j ye aks-i gereft-Ø
    Jiân from him one photo+indef got+3sgS
    'Jianî took a photo of him*i/j.'

    (b) jiânî az-esh*i/j ye aks-i gereft-Ø
    Jiân from+3sg one photo+indef got+3sgS
    'Jianî took a photo of him*i/j.'

(54) (a) jiânî u*i/j-ro did-Ø
    Jian him+râ saw+3sgS
    'Jianî saw him*i/j.'

    (b) jiânî did-Ø-esh*i/j
    Jian saw+3sgS+3sg
    'Jianî saw him*i/j.'
The ungrammaticality of the (a) sentences above, with full pronouns, is not surprising. It follows from Condition B of the Binding Theory since in each case the pronoun is coindexed with a c-commanding noun phrase within its governing category. However, the ungrammaticality of the (b) sentences is unexpected given the analysis outlined thus far. If the pronominal enclitics serve as accessible subjects for the pro object of the preposition or verb, then the VP or PP respectively should be the governing category for that pro. Outside of its governing category, pro should be free to be coindexed with another noun phrase. Why is this not possible in the sentences given above?

An answer to this question lies in the fact that in the cases above, the pro is an argument of the verb (in the first two cases this is mediated through the preposition\textsuperscript{24}). Recall that pro as possessor, unlike pro as a direct object, could be coindexed with a c-commanding noun phrase. In that case, however, it was the possessum, not the pro possessor that was an argument of the verb. It appears, then that we must constrain the coindexation of pro with another argument of its own predicate even if Binding Theory violations are not created by this coindexation. Just such a constraint is proposed by Reinhart and Reuland (1993). In

\textsuperscript{24}I simply assume this here but see Reinhart and Reuland (1993, section 4) for a discussion of syntactic and semantic predicates. They show that semantic predicates are relevant for the binding of pronominals and that a noun phrase can be a semantic argument of a predicate without receiving a theta-role from that predicate in the syntax. This means that even if a noun phrase is not clearly an argument of a predicate in the syntax (if it is contained within an adjunct PP, say), it may be an argument of the predicate at the relevant semantic level and thus still be subject to Binding Theory.
fact, they propose to restate binding theory in terms of constraints on predicates rather than on the forms that can serve as arguments of predicates. We will therefore look at Reinhart and Reuland's theory of binding in general since it seems that at least part of it is required to account for pro in Persian. We will see however, that we cannot do away with classic binding theory and will still need to state conditions on anaphors and pronouns in Persian.

4.4.2 Reinhart and Reuland and the Binding Theory

Reinhart and Reuland (R&R, hereafter) give the following definitions (p. 670), followed by their version of Conditions A and B of the Binding Theory (p. 670-1). Crucially, they intend these definitions to replace the classic ones, which are conditions on arguments rather than on predicates. This entails that notions such as c-command and governing category are also no longer required for the purposes of binding theory.

(56) Definitions
   a. a predicate is reflexive iff two of its arguments are coindexed
   b. a predicate (formed of P) is reflexive-marked iff either P is lexically reflexive or one of P's arguments is a SELF anaphor

(57) Condition A
   A reflexive-marked predicate is reflexive.

(58) Condition B
   A reflexive predicate is reflexive-marked.

---

25I am giving their abbreviated definitions here. The formal definitions, given on their p. 662, are the same as above except that they are relativized to an index, i.e. they are stated in terms of i-reflexivization and i-reflexive-marking. I am following R&R in using the abbreviated versions throughout the text.
By lexically reflexive predicates R&R mean predicates whose heads are marked as reflexive in the lexicon, with or without overt morphological marking. Persian does not have this type of lexical marking. What R&R mean by SELF anaphor is discussed in section 4.4.2.2. For our purposes here, we will simply interpret SELF anaphor as anaphor.

Let us return now to the problem raised in the preceding section, namely the ungrammaticality of sentences such as (54b), repeated below:

(54) (a) jiâni u*_{i/j}-ro did-Ø
     Jian him+râ saw+3sgS
     'Jiani saw him*_{i/j}.'

(b) jiâni did-Ø-esh*_{i/j}
     Jian saw+3sgS+3sg
     'Jiani saw him*_{i/j}.'

Both of these examples are ruled out by R&R's Condition B. There is a reflexive predicate in that the two arguments of the verb see are coindexed, however, there is no reflexive marking. Note that if an anaphor is used in the above examples, providing reflexive marking, the sentences are grammatical.

(59) (a) jiâni xod_{i}-râ did-Ø
     Jian self+râ saw+3sgS
     'Jiani saw himself_{i}.'

(b) jiâni xod-esh-ø_{i} did-Ø
     Jian self+3sg+râ saw+3sg
     'Jiani saw himself_{i}.'

The difference between the two anaphors in (59a) and (59b) above is discussed further in section 4.4.2.2 below.
There are a number of problems that arise if R&R's theory is adopted in the place of classic binding theory. Part of the problem may be that R&R do not deal with possessors or possessive anaphors. In the next two sections we see the problems for R&R's theory posed by the elements we have been considering in Persian. Specifically, it is argued that we still need the notion of c-command and governing category in addition to one of R&R's conditions on predicates.

4.4.2.1 R&R's Condition B and Possessors

One problem for R&R has to do with pronouns as possessors. Recall that full pronoun possessors cannot be coindexed with a c-commanding noun phrase in the same clause, as shown in (60a). (60b) shows that an anaphor can occur as possessor, in which case it must be coindexed with the subject.

(60) (a)  jiânî ketâb-e  u*i/j-râ xund-Ô
         Jian book+EZ  him+râ read+3sgS
         'Jiani read hisj book.'

(b) jiânî ketâb-e  xod*i*j-râ xund-Ô
         Jian book+EZ  self+râ   read+3sgS
         'Jiani read hisi (own) book.'

The problem is that to explain the complementary distribution of the pronoun and the anaphor (with respect to coindexation with the subject) by using R&R's revised binding theory, the possessor must be considered to be an argument of the predicate, in this case the verb 'read'. However, the possessor is not a direct argument of the verb at any level (syntactic or semantic) and thus that anaphor is not serving a reflexivizing function in this case.
R&R state that all aspects of local anaphora other than the reflexivizing aspect do not fall under binding theory but rather under chain theory. That is, for R&R Condition B is not the only constraint in the grammar that rules out the occurrence of pronouns in certain environments. Also at work is the Chain Condition. Loosely speaking, the Chain Condition states that an A-chain contains exactly one link that is case-marked and +R, where R stands for the property of referential independence which is shared by R-expressions and pronouns. According to R&R, this means that only one link of an A-chain projects a syntactic argument and they state further that this link must be at the head of the chain. Thus, if for example a chain is created between two pronouns or between a proper name and a pronoun, it will be ruled out by the Chain Condition since more than one link of the chain will be +R.

Returning to our problem above, exemplified in (60), note that if coindexing of the subject with the pronoun possessor of the direct object creates an A-chain, then we have an account of the ungrammaticality of this sentence. That is, such coindexation yields a Chain Condition violation since the pronoun is +R and is at the tail of the chain. Further, we can account for why the anaphor xod is acceptable here, since anaphors are -R.26

A problem arises if we consider the third type of possessor possible in Persian, namely pro. Consider the following example:

(61) jiânî ketâb-eshi/j-o xund-Ø
    Jian book+3sg+râ  read+3sgS
    'Jian read his book.'

As was noted in section 4.4.1.2, if a pronominal enclitic is used instead of a full pronoun as a possessor, coreference with the subject is possible (though not obligatory). This is because

26Of course, both the pronoun and the anaphor as possessors are getting case-marked (cf. Chapter 3, section 3.4.3). However, I will leave this problem aside for now since a more serious problem arises if we want to use the Chain Condition to account for the possessor facts.
enclitics serve to identify *pro* possessors, which in turn have as their governing category the DP in which they occur. Thus they are free to be coindexed with another noun phrase in the sentence. However, in spite of this difference between *pro* and full pronoun possessors, the two should behave alike with respect to case and specification for the feature +/-R. That is, the creation of a chain between the subject and a *pro* possessor should also be ruled out by the Chain Condition.

Of course, coindexation between the *pro* possessor and the subject is possible if the two do not form a chain. However, there is no way to prevent the creation of a chain between the subject and a *pro* possessor, if a chain exists between the subject and the pronoun possessor. Given that A-chains can be formed as long as no barriers intervene, and given that barriers involve projections that are not L-marked, it is impossible to have a DP in direct object position forming a barrier in one case (where the possessor is expressed by an enclitic) and not in another case (where the possessor is a pronoun). That is, whether or not a projection is a barrier does not depend on its internal constituents but rather on its relationship with its sister. In both cases the DP is sister to the same verb.

The best way to account for the difference between the pronoun and enclitic possessors seems to be the analysis outlined in section 4.4.1. That is, if we maintain that pronouns must be free in their governing category, then in (60a) where the governing category for the pronoun possessor is the whole sentence, the pronoun cannot be coindexed with the subject. By the same reasoning, if an anaphor must be bound by a c-commanding antecedent within its governing category (see the next section for more on this) then we can account for (60b). Finally, if we allow agreement heads to count as accessible subjects, then the governing category for the *pro* possessor is the DP and it can be freely coindexed with any noun phrase outside DP, as seen in (61).
4.4.2.2 R&R's Condition A and the Persian Anaphor *xod*

In order to examine R&R's version of Condition A, we must first determine what sort of anaphor exists in Persian. R&R state that there are two kinds of anaphors: long-distance anaphors and local anaphors. Further they state that there is a correlation between the type of anaphor and its morphological complexity: long-distance anaphors are universally simple while complex anaphors are universally local. They refer to the long-distance anaphors as SE anaphors (for simplex expression) and the local anaphors as SELF anaphors. In a footnote they state that they are not going to consider possessive anaphors.

They suggest the following structures for SE and SELF anaphors. First, they adopt the view that pronouns are in determiner position yet project as full NPs. (Recall, under the view taken here, pronouns are heads of DP.) They give the same structure for pronouns and SE anaphors shown in (62a) and (62b) respectively. In (62c) we see the structure for SELF anaphors, where the SELF is an N and combines with either a pronominal or SE determiner.

\[
(62) \quad \begin{align*}
(a) & \quad \text{NP} \\
& \quad \text{Pron} \quad \text{N'} \\
& \quad \text{e} \\
(b) & \quad \text{NP} \\
& \quad \text{SE} \quad \text{N'} \\
& \quad \text{e} \\
(c) & \quad \text{NP} \\
& \quad \text{Pron/SE} \quad \text{N'} \\
& \quad \text{self}
\end{align*}
\]

The structures given above can easily be translated into the structure we have adopted for Persian noun phrases. The pronouns and SE anaphors head DP projections. In (62c) this DP is generated as a possessor in the specifier of a DP, whose head takes NP as its complement.

Now, in Persian there are two types of anaphors also. The element *xod* can occur on its own, although, as pointed out by Moyne (1971) it is used only in formal or literary styles. More commonly, especially in the colloquial language, *xod* is inflected with the pronominal enclitics. This is shown below. I will refer to the elements in (63b) as instances of inflected
xod. Note also that as possessive anaphors, the elements in (63b) are, strictly speaking, not included in the scope of anaphors discussed by R&R.

(63) (a) xod 'self'

(b) xodam 'myself' xodemun 'ourselves'
    xodet 'yourself' xodetun 'yourselves'
    xodesh 'him/herself' xodeshun 'themselves'

Given that xod is a nominal element, and given the analysis developed for the pronominal enclitics, the structure for inflected xod is predicted to be the following.

(64) DP
    D' DP
    NP D pro_i
    | N Ø_{def+Fi}
    | xod

If we return to R&R's categorization of anaphors, Persian seems to have a third type. Inflected xod corresponds to R&R's SELF anaphors, but the morphologically simplex xod, which is generated under N, does not meet the structural criterion for being an SE anaphor. (Under our assumptions both pronouns and SE anaphors head DPs.) In fact, it seems to have a third set of properties altogether. Let us consider this more closely.

Long-distance anaphors are so named because their antecedents need not be local. Consider the following example from Dutch, where either an SE anaphor or a pronoun can occur.
(65) Jan zag [jou achter zich/hem staan].
Jan saw [you behind SE/him stand]
'Jan saw you stand behind SE/him.' [Reinhart & Reuland 1993:661.9]

Now consider the corresponding sentence in Persian.27

(66) (a) jiân to-râ did-Ø ke posht-e u istâd-i
Jian you+râ saw+3sgS that behind+EZ him stood+2sgS
'Jian saw that you stood behind him.'

(b) jiân to-râ did-Ø ke posht-e xod istâd-i
Jian you+râ saw+3sgS that behind+EZ self stood+2sgS
'Jian saw that you stood behind yourself.'

Note that xod cannot be used to refer to Jian. The presence of xod in the embedded clause signals reflexivization (and the resulting odd reading).

Hashemipour (1989) also points out that xod cannot be considered a long-distance anaphor, although the examples she gives involve inflected xod. These examples are given below.

(67) (a) moin goft-Ø [ke xod-esh ne-mi-âd-Ø]
Moin said+3sgS that EMP+3sg neg+cont+come+3sgS
'Moin said that he is not coming.'

(b) moin fekr mi-kon-e [ke xod-esh-o na-did-am]
Moin think cont+do+3sgS that EMP+3sg+râ neg+saw+1sgS
'Moin thinks that I did not see him.' [Hashemipour 1989:84.126]

27This example also involves a 'raising-to-object' type structure, however, this is irrelevant for the point at hand.
She points out that unlike long-distance anaphors in languages like Icelandic and Italian, obligatory coreference does not obtain between the anaphors in the above sentences and the matrix subjects. In this sense they act like the pronouns in the English translations. This corresponds to what R&R call the logophoric28 use of anaphors, which will be discussed shortly. Of note here is the fact that the sentences above are ungrammatical if uninflected xod is used.

(68) (a) *moin; goft-Ø [ke xod; ne-mi-åd-Ø]
    Moin said+3sgS that EMP neg+cont+come+3sgS
    'Moin said that he is not coming.'

(b) *moin; fekr mi-kon-e [ke xod;râ na-did-am]
    Moin think cont+do+3sgS that EMP+râ neg+saw+1sgS
    'Moin thinks that I did not see him.'
    (but fine as 'Moin thinks that I did not see myself')

As indicated, the sentence in (68b) is fine if xod is construed as coreferential with the embedded subject. This points to the clause bound nature of uninflected xod. That is, unlike both SE and SELF anaphors, xod does not seem to have a logophoric usage.

The clause bound nature of xod and the difference between it and the inflected xod forms cause problems for R&R's version of Condition A of the binding theory. Recall that rather than having a condition on the anaphor itself (i.e. that it must be bound in its governing category), their Condition A states that a reflexively marked predicate must be reflexive. For uninflected xod this means that if it occurs in a sentence it must be coindexed with another argument. However, consider the following contrast:

28R&R state that they "... refer to the grammatical occurrence of SELF anaphors in nonreflexive contexts as logophoric" (p. 672). However, they note that they are using the term in a wider sense than originally intended (cf. Sells (1987) for a discussion of the uses of this term).
(69) (a) jiâni xod_{i/*j-}râ did-Ø
     Jian  self+râ    saw+3sgS
     'Jian saw himself_{i/*j-}.'

(b) jiâni xod-esh_{i/j-}o did-Ø
     Jian  self+3sg+râ saw+3sgS
     'Jian saw himself_{i/j-}.'

(c) jiâni xod-e u*{i/j-}râ did-Ø
     Jian  self+EZ  him+râ saw+3sgS
     'Jian saw HIM*{i/j-}.' (lit 'Jian saw self-of-him.')</div>

Let us first look at the difference between (69a) and (69b). We see that in (69a) xod can only refer to the subject, Jian. This is consistent with the idea that it obligatorily reflexivizes the predicate. However, in (69b) we see that inflected xod can optionally be used to refer to someone other than Jian. This would occur in a context where, for example, someone asked if Jian had seen a picture of Bill and the reply would be 'No, Jian saw HIMSELF' (where HIMSELF = Bill). The problem for R&R's Condition A is that if xod is a SELF anaphor that obligatorily reflexivizes the predicate then it should do so in (69b) also. If it can optionally reflexivize a predicate then we cannot explain the obligatory coreference in (69a).

R&R discuss examples of the logophoric use of SELF anaphors in argument positions under the heading of Focus Anaphors. They note that this use of discourse anaphors has been labeled emphatic. This is true in Persian also where it is often claimed that there are two elements xod, one a reflexive marker, i.e. an anaphor, and the other an emphatic element (cf. Moyne (1971), Moyne and Carden (1974), Hashemipour (1989)) The use of anaphors as focus elements in argument positions presents a problem for R&R's Condition A. However, they state that if focus elements undergo movement at LF, then they no longer occupy argument positions hence Condition A will not apply.
Returning to the contrast between (69a) and (69b) the question is why only inflected *xod* can be used as a focus element, if this is the way it can escape being bound by the subject, Jian. Furthermore, if we look at (69c) we see that *xod* followed by a full pronoun in direct object position cannot be coindexed with the subject at all. Thus we need to say that *xod* itself cannot be used as focus, inflected *xod* can optionally be used as focus, and *xod* followed by a full pronoun possessor **must** be used as focus.

There is an alternative analysis, however, that does not require any new conditions, nor even the idea of focus movement at LF. Looking at the difference between (69b) and (69c) note that this is exactly the contrast we saw in section 4.4.1.2 between enclitic and pronoun possessors. The former may be optionally coindexed with the subject, while the latter can never be. It seems then that these two sentences do not exemplify any properties of *xod* but rather the fact noted above that pronouns must be free in their governing category, and that the enclitics identify a *pro* possessor and act as an accessible subject for it.

Let us say then, that *xod* is an element that must be bound by the closest c-commanding antecedent within the same governing category. In other words, *xod* is an anaphor under classic binding theory. In (69a) then, *xod* is c-commanded by the subject and must be coindexed with it. However, in (69b) and (69c) there is a closer c-commanding antecedent, namely *pro* and the pronoun, respectively. The structures for the direct objects in (69b) and (69c) are given below.

(70) (a)  
```
        DP
          \   /
           D' DP
             |   |
              NP D pro
                 |   |
                  N O_{def}^F\_i
                     |   |
                      xodi
```

(70) (b)  
```
        DP
          \   /
           D' DP
             |   |
              NP D Pronoun\_i
                 |   |
                  N O_{def}
                     |   |
                      xodi
```
The properties of the sentences in (69b) and (69c) now follow from the differences between pro and full pronoun possessors.

The analysis of xod developed here explains another interesting fact noted by Moyne (1971). In Persian the subject can be reduplicated, or emphasized (cf. also Moyne and Carden (1974). While the exact analysis of this phenomenon will not be outlined here, the fact is that either xod or inflected xod can be used for this purpose. Of interest is the fact that with inflected xod a pronominal subject need not be expressed, while with uninflected xod an overt subject is required. Moyne's two paradigms are shown below:

(71) (a) man xod âmadam 'I came myself.'
(b) to xod âmadi 'You (sg.) came yourself.'
(c) u xod âmad 'He came himself.'
(d) mà xod âmadim 'We came ourselves.'
(e) shomâ xod âmadid 'You came yourselves.'
(f) ishân xod âmadand 'They came themselves.'
(g) hushang xod âmad 'Hushang came himself.'
(h) hushang va zhâla xod âmadand 'Hushang and Zhala came themselves.'

(72) (a) (man) xodam âmadam 'I came myself.'
(b) (to) xodat âmadi 'You (sg.) came yourself.'
(c) (u) xodash âmad 'He came himself.'
(d) (mâ) xodemân âmadim 'We came ourselves.'
(e) (shomâ) xodetân âmadid 'You came yourselves.'
(f) (ishân) xodeshân âmadand 'They came themselves.'
(g) hushang xodash âmad 'Hushang came himself.'
(h) hushang va zhâla xodeshân âmadand 'Hushang and Zhala came themselves.'

[Moyne 1971:145.13-14]

These facts are consistent with the idea that xod needs to be bound by a c-commanding antecedent, while inflected xod need not be since its antecedent is the pro identified by the inflection.
Note that in claiming that there is only one *xod* which can either mark reflexivity or, if properly bound, can be used as an emphatic element, the question of when or why *xod* is used for emphasis belongs to discourse rather than to syntax. R&R make this point also:

In the contexts where the syntax allows both a pronoun and a SELF anaphor to be coindexed with a given antecedent, the choice between them is motivated by discourse considerations, as is often the case when there is more than one syntactic option to express the same proposition. For this reason, the use of an anaphor in such contexts may appear more marked than in the reflexivity environments, where the anaphor is the only grammatical option. There is no reason to assume that this type of discourse consideration is encoded in the syntax. (R&R 1993:672)

### 4.4.3 Summary and Questions for Further Research

In this section we have looked at the conditions on pronominal elements and anaphors in Persian in light of two versions of binding theory. One version, classic binding theory, states the conditions on the nominal elements themselves. However, we have seen that under this version we cannot explain why *pro* cannot serve as an argument of a predicate when it is coindexed with another argument of the same predicate. On the other hand, R&R's version of binding theory, with the binding conditions stated over predicates, cannot account for the contrast between full pronoun and *pro* possessors. Further, we have seen that this distinction is relevant to the behaviour of the anaphor *xod*.

The generalization that can be made then is the following: Classic binding theory holds in Persian with the additional constraint that when two arguments of a predicate are coindexed, this must be marked. In other words, classic binding theory, and R&R's Condition B, hold in Persian. The anaphor *xod* is the only element in Persian that can be coindexed with something bearing a different theta-role with respect to the same predicate (a
coargument in R&R's terms). However, R&R's Condition A does not hold. That is, the presence of $xod$ in a sentence does not mean that it has to be coindexed with a coargument. If $xod$ is bound within its own DP it will not necessarily reflexivize a predicate. Thus while it can be coindexed with a coargument, it need not be coindexed. I will rename R&R's Condition B as the Reflexivity Condition so as not to confuse it with the classic Condition B.

(73) (a) **Reflexivity Condition** (R&R's Condition B)
A reflexive predicate must be reflexive-marked.

(b) **Reflexivity Condition relativized to Persian**
If two arguments of a predicate are coindexed, one must be the anaphor $xod$.

One remaining question pertains to the semantic contribution of inflected $xod$ with intransitive predicates. As mentioned throughout this section, traditionally it has been claimed that there are two elements $xod$, a reflexive and emphatic. While the analysis proposed here unifies these two syntactically, there is still the question of whether the non-coindexed use of the anaphor, i.e. its logophoric use, is only for focus or emphasis.

Consider the following examples:

(74) (a) ketāb oftād-Ø
     book  fell+3sgS
     'The book fell.'

(b) (ketāb) xodesh oftād-Ø
     book itself fell+3sgS
     'The book fell by itself.'
(a) cherâq xâmush shod-Ø
    light  closed  became+3sgS
    'The light turned off.'

(b) (cherâq) xodesh xâmush mi-she
    light  itself  closed  cont+become+3sgS
    'The light turns off by itself.'

The semantic contribution of the anaphor in subject position does not seem to be solely one of emphasis. In both cases the (b) sentences have an implied sense that there is no external cause. Perhaps then, there is a reflexivizing function being performed here, but at a more abstract level. I leave this as a question for further research on the semantics of reflexivity.29

4.5 Conclusion and Residual Problems

In this chapter a structural account of subject agreement and the pronominal enclitics of Modern Persian has been presented. It has been argued that both sets of morphemes are the realization of F-features. These features are acquired by a head pre-syntactically at the Lexicon-Syntax Interface. The subject agreement suffixes are subcategorized for verbs only. To be checked they must assign nominative case to a noun phrase that they govern and with which they agree in person and number features. The pronominal enclitics, on the other hand, can attach to any head. They are subject to a PF filter that rules out their direct attachment to verb stems, however. They absorb the case of the head to which they attach (if there is one), thus they most often occur with pro. They must be checked by agreeing in person and number features with a noun phrase that they govern.

29For a discussion of these and other issues pertaining to reflexive predicates in French, see Cummins (1995).
The second half of the chapter has looked at the binding of pronominal elements and anaphors in Persian. We have seen that the binding properties of the pronominal enclitics can be accounted for if, in fact, pro is the potential bindee. By virtue of being accessible subjects, the pronominal enclitics allow pro to be coindexed outside of the projection within which it occurs, subject to the Reflexivity Condition.

Finally, our look at binding in Persian has also involved an account of the anaphor xod. We have seen that both the reflexive and emphatic uses of this element can be accounted for if the analysis developed for the pronominal enclitics holds of this item also, when it appears inflected. Thus, the binding properties of pronominals and anaphors in Persian provides support for the analysis developed in the first half of the chapter.

In the next chapter we consider properties of the clause in Persian, including a look at the direct object marker -râ, and also certain constructions in which the pronominal enclitics participate. However, we end this chapter with a brief discussion of one residual problem.

4.5.1 The Pronominal Enclitics and Indirect Objects

In section 4.4.1.3 we saw that the pronominal enclitics can replace the direct object of a verb. There are some cases where the enclitics appear to be able to replace indirect objects also. These cases present a problem in that, rather than appearing on the preposition, the enclitics appear directly on the verb.

This process appears to be quite restricted. For example, if we compare the following two cases, we see that the enclitic can replace the indirect object with the verb dâdan 'to give', shown in (76c), but not with the verb gozâshtan 'to put', shown in (77c), even though in this case too the indirect object is a subcategorized argument of the verb.

(76) (a) ketâb-o be jiân dâd-am
    book+râ to Jian gave+1sgS
    'I gave the book to Jian.'
It seems that the pronominal enclitics can replace indirect objects that are 'dative' objects, in the sense that in English these are the indirect objects that appear in double-object constructions. The link between the pronominal enclitics and dative arguments is further reinforced by the following example:

As pointed out by Simin Karimi (personal communication) the indirect objects that can be replaced by pronominal enclitics can also appear post-verbally without the preposition. This is shown below where (i) should be compared to (76a) and (ii) should be compared to (77a).

(i) ketâb-o dâd-am jiân
   book+râ gave+1sgS Jian
   'I gave the book to Jian.'
(a) be u  goft-am
    to him told+1sgS
    'I told him.'

(b) ??u-râ  goft-am
    him+râ told+1sgS
    'I told him.'

(c) ?goft-am-esh
    told+1sgS+3sg
    'I told him.'

The grammaticality of (78b), though somewhat marginal now, reflects the history of -râ as a dative marker (see Windfuhr (1979, p. 51 and references therein). Lazard (1992) also notes that -râ may mark indirect objects although "[t]his usage was much more frequent in preclassical and classical language; today indirect objects are marked usually by the directional preposition be 'to'." (p. 191) Note that in a construction such as (78b) the governing head on which an enclitic must appear if the indirect object is pro is the verb since there is no preposition. Thus it is possible that the instances where the pronominal enclitics can appear on the verb marking an indirect object, are precisely the cases where the indirect object was not in a prepositional phrase.

(ii) *ketâb-o gozâsht-am miz
    book+râ put+1sgS  table
    'I put the book on the table.'
5.0 Introduction

This first half of this chapter is devoted to the syntax and semantics of -râ and related issues in the structure of the clause in Persian. Râ has been the subject of much discussion and debate in the literature on Persian. Most recently it has been the focus of work by Karimi (1989, 1990).

According to Karimi (1989, 1990) and Bubenik (1992), the source of -râ is the Middle Persian relational postposition rây (<Old Persian rádi) meaning 'for, on behalf of, on account of'. Karimi (1990:179-180) states that in Middle Persian rây also "appears as an illustration of purpose, reference, beneficiary, or indirect object". Crucially, however, it does not appear with direct objects in Middle Persian or in what she terms early Modern Persian texts (see Karimi (1989, 1990) for more discussion and examples). This distribution persists in some dialects of Persian (e.g. Jahromi).

The behaviour of -râ raises two questions. First, what is its meaning? While it often seems to act like a definiteness marker (see (1a)), it can cooccur with the indefinite marker -i (see (1b)).

(1) (a) ketâb-o xarid-am
       book+râ bought+1sgS
       'I bought the book.'

---

1Karimi (1990) is a published version of the major claims in Karimi's (1989) thesis. For convenience I will cite the later work, except where the two diverge.
2In (1) we also see the phonological realization of this morpheme. In colloquial spoken Persian -râ appears as -ro following a vowel and as -o following a consonant.
(b) ketāb-i-ro xarid-am
book+indef+râ bought+1sgS
'I bought a certain/particular book.'

Browne (1970) observed for both Persian and Turkish that the cooccurrence of definite and indefinite markers is unexpected. This led him to propose that the feature [+specific], and not [+definite], is borne by -râ and Turkish -i. Karimi follows in this line, arguing that -râ is a specificity marker. It has also been argued that -râ is a topic marker (cf. Karimi (1989, 1990) for references).

The second question raised by -râ is whether it marks accusative case or direct object status. On the one hand, it never appears on subjects, even if they are definite/specific (see (2a)). On the other hand, as pointed out by Karimi, -râ can appear on adverbial noun phrases (see (2b)) and can sometimes appear twice in a single clause (see (2c)). This should not be possible if -râ marks accusative case on direct objects.

(2) (a) ketāb oftâd-Ø
book fell+3sgS
'The book fell'

(b) hafte-ye âyanda-ro esterâhat mi-kon-am
week+EZ coming+râ relax cont+do+1sgS
'As for next week, I will relax.' (I'll relax the whole week')

Karimi 1990:143.12

(c) mâshin-o dar-esh-o bast-am
car+râ door+3sg+râ closed+1sgS
'As for the car, I closed its door.'

Karimi 1990:143.13

The chapter is organized as follows. In section 5.1 the syntactic category of -râ is briefly discussed. It is argued that -râ is best treated as a phrasal affix heading a Kase Phrase. In
section 5.2 the semantic contribution of -râ is discussed with reference to Karimi's claim that -râ marks noun phrases for specificity. It is shown that the notion of presupposition better characterizes the semantic contribution of -râ. Section 5.3 discusses the syntax of noun phrases marked with -râ, arguing that DPs marked with -râ occur higher than other direct objects. It is proposed that presupposed DPs cannot receive case from the verb. These DPs must move to adjoin to VP where they receive case from -râ. It is also shown that VP-adjoined DPs may not have originated as direct objects and that -râ is available to case-mark any DP adjoined to VP. The Thematic Licensing Condition is proposed and is shown to license such DPs which are construed as VP-level topics. Section 5.4 deals with a second process of topicalization, which is referred to as clause-external topicalization. In section 5.5 the position of subjects is discussed in light of the analysis of -râ.

In the remainder of the chapter complex predicates in Persian are discussed. Section 5.6 deals with the structure of sentences containing transitive compound verbs, which typically involve a light verb and a non-verbal element. An analysis is presented in which the light verbs take a Predication Phrase (Bowers (1993)) as their complement. The phrasality of the non-verbal element and the fact that the pronominal enclitics can occur on this constituent are shown to follow from the proposed structure. In section 5.7 the analysis of transitive compound verbs is extended to the impersonal construction in Persian. It is shown that this construction is the intransitive counterpart to the transitive compounds. The properties of this construction are shown to follow from the fact that it involves only a VP. Section 5.8 contains some concluding remarks
5.1 The Syntactic Category of \(-râ\)

What category does \(-râ\) belong to? If it marks case\(^3\), as will be shown in section 5.3, then it cannot be a determiner. This is further confirmed by the fact that it can occur with the indefinite enclitic \(-i\) which is a determiner. Since I assume that an NP cannot be dominated by two DPs, I propose that \(-râ\) is a phrasal affix heading a KP (for Kase Phrase, cf. Travis and Lamontagne (1992)). The fact that \(-râ\) is a phrasal affix is easily seen in examples such as (3) below where the DP marked with \(-râ\) has a complex possessor. \(Râ\) appears not on the head noun, but at the end of the noun phrase.

(3) \[kolâh-e un mard-e qad-boland]-o na-pasandid-am\]
\[hat+EZ that man+EZ height+tall+râ neg+like+1sgS\]
'I didn't like the hat of that tall man.'

As a K-head, \(-râ\) selects a DP complement. It can therefore appear with either the indefinite enclitic \(-i\), or the null definite determiner (see Chapter 3, section 3.5.3). This is shown below.

(4) \[(a)\]
\[
\begin{array}{c}
\text{KP} \\
\text{DP} \\
\text{NP} \\
\text{D} \\
\text{K} \\
\end{array}
\]
\[-râ\]

\[(b)\]
\[
\begin{array}{c}
\text{KP} \\
\text{DP} \\
\text{NP} \\
\text{D'} \\
\text{D} \\
\end{array}
\]
\[Ø_{def}\]

\(^3\)I distinguish case marking from case assignment. A case marker appears on a noun phrase to license that noun phrase for the purposes of the Case Filter. A case assigner assigns case to a noun phrase in a structurally defined syntactic position. Note that a determiner can be a case-assigner. For example, the null definite determiner in Persian assigns case to a noun phrase in its specifier (see Chapter 3).
5.2 The Semantic Contribution of -rā

As pointed out by Browne (1970) and Karimi (1990), the cooccurrence of -rā and the indefinite determiner -i casts doubt on the hypothesis that -rā is a marker of definiteness. On the other hand, if Karimi is right in claiming that a language can mark either definiteness or specificity but not both, then -rā cannot mark specificity, given that Persian has a null definite determiner (see Chapter 3). Rather, the meaning of -rā must be compatible with both definiteness and indefiniteness.

Section 5.2.1 surveys Karimi’s arguments for considering -rā to be a specificity marker. We will see that most of these arguments are compatible with the view that -rā selects either an indefinite or a definite noun phrase as its complement. The remaining arguments show that -rā makes a semantic contribution to its noun phrase, but it is suggested that what it contributes is presupposition not specificity. If -rā is taken to mark noun phrases referring to entities presupposed in the discourse, all of its uses can be united. This view is discussed in section 5.2.2.

Note that in this section the data shows that -rā must select a DP. However, an analysis of the fact that definite DPs must appear with -rā, while indefinite DPs may optionally appear with this morpheme, is postponed until section 5.3.

5.2.1 Definiteness and Specificity

Recall from Chapter 1, section 1.2.6 that noun phrases can be divided into four types: definite, indefinite specific, indefinite non-specific, and generic. Karimi (1989) makes this four-way distinction and gives the following characterizations of the first three types, drawing on work by Heim (1982). A definite noun phrase has a referent known to the speaker, which is also presupposed to be known to the hearer. By definition definite noun
phrases are also specific. The referent of an indefinite specific noun phrase, while known to
the speaker, is not presupposed to be known to the hearer. That is, the hearer is not supposed
to be able to identify the referent, although the speaker has a particular referent in mind. The
referent of an indefinite non-specific noun phrase, according to Heim, signals a yet
unfamiliar referent. The speaker does not have a particular referent in mind. Karimi uses the
following test to illustrate the difference between indefinite specific and indefinite non-
specific use of a noun phrase. (The relevant indefinite noun phrase in (5) is italicized.)

(5) Mary was looking for a sloop, and
    she found one. (Nonspecific)
    she found it. (Specific)

The fourth category is the generic or non-referential noun phrase, which Karimi does not
formally define. Recall from Chapter 1 that I use the terms 'generic' and 'non-referential'
interchangeably to refer to any noun phrase that does not denote an entity or token, but rather
is used as a label for a type.

Karimi claims that -râ marks specificity in Persian. She writes that "[t]he fact that every
language has either a definite or a specific marker, but not both ...suggests that universal
grammar (UG) has a single category of specific/definite (=presumed known) whose
interpretation can differ from language to language, but which may play a role in every
language." (1990, p. 146) She defines specific reference as the denotation of a particular
entity out of a set of entities. The difference between definite and indefinite specific noun
phrases is that the former are presumed to be known to the hearer whereas the latter are not.

There is actually some evidence that she does not distinguish indefinite non-specific noun
phrases from generic noun phrases. That is, she seems to use 'specific vs. non-specific' the
same way as she uses 'referential vs. non-referential'. For us, the difference between an
indefinite non-specific noun phrase and a generic one is referentiality.
Karimi (1990) provides ten arguments that -râ is a specificity marker to which we now turn. In each case I will state the main points of Karimi’s argument and then go on to show that the data are also compatible with the view that -râ selects a DP. Specifically, we will see that seven of Karimi’s arguments are compatible with the view that -râ selects a DP headed by either a definite or indefinite marker. The remaining three arguments involve cases that can be captured as easily by the notion of presupposition as by the notion of specificity. Presupposition will be discussed in the immediately following section.

Argument 1:  -râ is obligatory when the direct object is a proper noun

Argument 2:  -râ is obligatory when the direct object is a pronoun

(6) (a) jiân-o did-am
       Jian+râ saw+1sgS
       'I saw Jian.'

(b) *jiân did-am
       Jiansaw+1sgS

(7) (a) shomâ-ro did-am
       you+râ saw+1sgS
       'I saw you.'

(b) *shomâ did-am
       you saw+1sgS

Karimi states that since proper names and pronouns are specific, they require the presence of -râ. Under the present analysis (see Chapter 3, section 3.5.1) proper names and pronouns are generated under D and project to DP. Note also that both proper names and pronouns are
definite as well as specific. These generalizations therefore do not argue against -râ as a marker of definiteness or as a selector of DPs which may be definite.

Argument 3:  *-râ is obligatory when the noun phrase is headed by a demonstrative  
  
determiner  

(8)  (a) in ketâb-o xund-am  
      this book+râ read+1sgS  
      'I read this book.'  

      (b) *in ketâb xund-am  
          this book read+1sgS

As with arguments 1 and 2, argument 3 involves noun phrases that are definite as well as specific. If we assume that demonstratives, like proper names and pronouns, project to DP, the presence of -râ is expected.

The next two arguments have to do with coindexing between the noun phrase bearing -râ and a pronominal element. They are given below.

Argument 4:  *An NP+râ may be coreferential with a clitic pronoun in the same  
  simple clause whereas the same noun phrase without -râ may not be  

(9)  (a) ketâb-o xund-am-esh  
      book+râ read+1sgS+3sg  
      'As for the book, I read it.'  

      (b) *ketâb xund-am-esh  
          book read+1sgS+3sg
Argument 5: The pronoun un ‘it’ substitutes for a phrase followed by -râ whereas yeki ‘one’ substitutes for a generic noun phrase

(10) râmin pirhan xarid-Ø man ham yeki/*un-o xarid-am
Ramin shirt bought+3sgS I also one/*that+râ bought+1sgS
'Ramin bought a shirt, I bought one too.' [Karimi 1990:148.25]

(11) man mi-xâst-am pirhan-e sabz-o be-xar-am
I cont+wanted+1sgS shirt+EZ green+râ subj+buy+1sgS

ammâ râmin un-o/*yeki zud-tar xarid-Ø
but Ramin that+râ/*one soon+comp bought+3sgS
'I wanted to buy the green shirt, but Ramin bought it first.' [Karimi 1990:148.26]

If, as Karimi claims, pronominal elements are specific, then any coindexed element should also be specific. Again, note that pronominals are definite as well as specific. The fact that a noun phrase bearing -râ can be coindexed with these elements does not show that -râ must mark specificity rather than definiteness.

Karimi presents two arguments that depend on the analysis of two morphemes: the plural marker and the marker -a/-e that occurs in the colloquial spoken language. Karimi assumes without argument that both of these markers indicate specificity. In a footnote (fn. 15, p. 114) she notes that Samiian (1983) has considered the plural suffix to be a definite marker. Presumably, for Karimi these markers must indicate specificity because languages mark either definiteness or specificity but not both.

Argument 6: The plural marker is not obligatory unless a specific reading is intended. The presence of this suffix forces the presence of -râ
(12)  (a)  ketâb-hâ-RO xund-am  
      book+pl+râ read+1sgS  
      'I read the books.'

      (b)  *ketâb-hâ  xund-am  
           book+pl  read+1sgS

Recall from Chapter 3, section 3.5.5, that the plural marker must be licensed by the head of a DP projection above the noun on which it occurs. Given that Persian has both a definite and indefinite determiner, the analysis from Chapter 3 predicts that the plural marker should be able to appear with the indefinite determiner. This prediction is borne out as the following example shows.

(13)  ketâb-hâ-Ô xund-am  
      book+pl+indef  read+1sgS  
      'I read some books.'

Karimi's statement, on the other hand, predicts that the cooccurrence of the plural marker and the indefinite enclitic should not be possible. The facts support the idea that the plural marker must be licensed by a D-head, whether it be definite or indefinite, and that -râ selects the DP as its complement.

Argument 7:  The suffix -a (-e word-finally), used only in spoken language, marks noun phrases for specificity. The presence of this marker forces the presence of -râ

5 Note that the way Karimi's argument is phrased, the plural marker can cooccur with the indefinite determiner if an indefinite specific reading is intended. However, it is not the case that the resulting noun phrase must be marked with -râ. If -râ only marks definite specifics then this amounts to saying that -râ is a definiteness marker.
The morpheme -a/-e was analyzed in section 3.5.6 as turning common nouns into definite nouns that then head DPs themselves. This analysis makes nouns marked with -a/-e identical to proper names. As with all of the preceding arguments then, this argument is consistent with -râ taking a DP as its complement.

Karimi's next argument has to do with the interrogative chi 'what', which may optionally appear with -râ.

Argument 8: The interrogative element chi 'what' may or may not cooccur with -râ. When it does, it receives a specific reading.

(15) (a) emruz ketâb xarid-am
today book bought+1sgS
'I bought books today.'

(b) chi-(*ro) xarid-i
what bought+2sgS
'What did you buy?' [Karimi 1990:149.30]

(16) (a) ketâb-i-ro ke be to gofte bud-am xarid-am
book+indef+râ that to you told was+1sgS bought+1sgS
'I bought the book that I had told you (about).'

(14) (a) ketâb-a-ro xund-am
book+def+râ read+1sgS
'I read that book.'

(b) *ketâb-e xund-am
book+def read+1sgS
According to my consultants, *chi-ro* can only be felicitously used in an echo question. For example if someone says 'I went shopping today' the hearer cannot say use (16b) but must use (15b) to ask 'what did you buy?'. Sentence (16b) is perhaps better characterized as meaning 'What is it that you bought?' or 'You bought the what?'. It seems then that previous mention in the discourse of some entity is relevant to the use of *chi-ro*. This falls under the characterization of presupposition that is discussed in the next section.

Karimi's last two arguments are given below. In both cases, *-râ* marks a noun phrase that is presumed known to the hearer. In fact, this is the characterization of presupposition that is proposed in the next section. Under Karimi's characterization of specificity, it is not necessary for a specific noun phrase to be presumed known to the hearer. Thus, it seems that the notion of presupposition better captures the meaning of these noun phrases.

Argument 9:  
*-râ* is obligatory when the noun is modified by words indicating a "certain" or "particular" reading

\[(17)\]

(17) (a)  
\[kâr-e \quad \text{mored-e nazar-o} \quad \text{peydâ kard-am}\]  
work+EZ case+EZ sight+râ found did+1sgS  
'I found the job (I had) in mind.'

(b)  
\[\ast kâr-e \quad \text{mored-e nazar} \quad \text{peydâ kard-am}\]  
work+EZ case+EZ sight found did+1sgS  
[Karimi 1990:148.28]
Argument 10: *Any relative clause modifying an NP adds to its descriptive content.*

*Thus the NP is set up to favor a specific reading, forcing the presence of -râ*

(18) (a) ketâb-i-ro ke diruz darbâr-ash harf mi-zad-im xarid-am
    book+indef+râ that yesterday about+3sg talk cont+hit+1plS bought+1sgS
    'I bought the book we were talking about yesterday.' [Karimi 1990:147.21]

    (b) *ketâb-i ke diruz darbâr-ash harf mi-zad-im xarid-am
        book+indef that yesterday about+3sg talk cont+hit+1plS bought+1sgS
        [Karimi 1990:147.22]

We have seen that the first seven of Karimi's ten arguments that -râ is a specificity marker are as easily accounted for by the proposal that -râ takes a DP as its complement. In most of the cases discussed above, this DP is definite, however, it need not be. We have also seen that there are three cases where -râ seems to involve presupposition, i.e. where the noun phrase marked with -râ is presumed to be known to the hearer. We now turn to a proposal that accounts for all of these facts.

5.2.2 -râ as a Presupposition (Hearer-Old) Marker

Let us consider again the four referential types that can appear in direct object position.

(19) (a) ketâb-o xarid-am
    book+râ bought+1sgS
    'I bought the book.'

    (b) (yek) ketâb-i xarid-am
        (one) book+indef bought+1sgS
        'I bought a book.'
I claim that -rå selects a DP as its complement. In Chapter 3 it was argued that there are two determiners in Persian, the indefinite enclitic -i and a null definite determiner. The appearance of -rå on a bare noun entails the presence of the null determiner, giving a definite reading for the noun phrase. The structure for the direct object in (19a), then, will be the following.

\[
\begin{array}{c}
\text{(20)} \\
\begin{array}{c}
\text{KP} \\
\text{DP} \\
\text{NP} \\
\text{N}
\end{array}
\end{array}
\]

\[
\begin{array}{c}
\text{K} \\
\text{D} \\
\text{Ø}_{\text{def}}
\end{array}
\]

\[
ketāb
\]

We have yet to account for the semantic contribution of -rå when it attaches to an indefinite noun phrase as shown in (19c). Note that the meaning of (19c) is not clear at all, although the sentence improves markedly if the noun phrase is followed by a relative clause.

\[
\begin{array}{c}
\text{(21) (a) ketāb-i-ro ke diruz darbār-ashharf mi-zad-im xarid-am} \\
\text{book+indef+rā that yesterday about+3sg talk cont+hit+1plS bought+1sgS} \\
\text{I bought the book we were talking about yesterday.'[Karimi 1990:147.21]}
\end{array}
\]
Recall that the indefinite enclitic was characterized as introducing a new referent into the discourse (cf. Chapter 3, section 3.3.2). Following Prince (1992), I distinguish three types of old/new information. On the one hand, information (i.e. an entity or referent) "may be old/new with respect to (the speaker's beliefs about) the hearer's beliefs" (Prince 1992:301). On the other hand an entity or referent may be old/new with respect to the discourse model being constructed between a speaker and hearer. Thus a noun phrase may refer to an entity that is (a) Hearer-new and Discourse-new, (b) Hearer-old and Discourse-old, or (c) Hearer-old and Discourse-new. (The one possibility that is ruled out is Hearer-new and Discourse-old.)

These are precisely the distinctions we need in order to account for the three ways in which object noun phrases can refer in Persian. A noun phrase marked with -i is Hearer-new and Discourse-new. A definite noun phrase is Hearer-old and Discourse-old. A noun phrase marked with both -i and -râ introduces a referent that is Discourse-new but Hearer-old. This is why these noun phrases are more felicitous when modified by a relative clause. The relative clause lets the hearer know that the referent is familiar.

This distinction also provides us with a way to characterize the suffixes -i, -râ, and the null definite determiner such that the meanings of the direct objects in (19a)-(19c) above can be compositionally derived. Let us say that -i and the null definite determiner mark noun phrases as new or old with respect to the discourse. That is, the morpheme -i marks Discourse-new noun phrases while the null definite determiner marks Discourse-old noun phrases (which by definition are also Hearer-old). Meanwhile, -râ marks Hearer-old noun phrase.

---

6It also has a partitive sense or a sense of selection which accounts for its presence on nouns followed by restrictive relative clauses (cf. Chapter 1, section 1.3.3).
phrases. Thus, when -râ occurs along with the null definite determiner its semantic content is redundant. When -râ cooccurs with the indefinite enclitic, the noun phrase is Hearer-old but Discourse-new. If Hearer-old information is what we mean by presupposition then we can say that -râ marks noun phrases as presupposed.

5.3 The Structural Position of DP+râ

In this section we consider where a noun phrase marked with -râ occurs within the clause. Since -râ does not appear on presupposed subjects, it seems plausible that it marks some kind of case in addition to marking presupposed information. If it does, then we must determine what kind of case this is and where it is assigned.

Karimi (1990) proposes that -râ marks oblique case. However, she shows that DPs marked with -râ appear to be outside the government domain of a head. She suggests that such DPs are marked as oblique by being coindexed with a clitic or a trace in an oblique argument position. Karimi's analysis of -râ is presented in section 5.3.1.

In sections 5.3.2 to 5.3.4 an alternative hypothesis is pursued. Under this view, the process of case transmission is not required in Persian. Instead -râ is analyzed as an element which case-marks any DP adjoined to VP. The constraints on DPs that can appear adjoined to VP are then considered and it is shown that the adjunct to VP position is a sort of VP-level topic position.

There are several advantages to taking this view. First, case transmission is eliminated in favour of theta transmission, a principle independently motivated for dislocated topics. Second, non-nominative cases no longer need to be referred to as a class. Third, the appearance of -râ on DPs is uniform. Râ appears only on DPs adjoined to VP. This explains why -râ does not appear on clause-external topics (see section 5.4). Finally, the proposed
connection between presupposed direct objects and VP-level topics captures a long-standing intuition that direct objects, like subjects, have special status.\(^7\)

5.3.1 Karimi's Analysis of the Syntax of \(-râ\)

The syntax of \(-râ\), like its meaning, is not immediately obvious. The fact that it most often occurs on direct objects of verbs (examples are repeated below) suggests that it is an accusative case marker.

(22) (a) sib-o xord-am apple+râate+1sgS 'I ate the apple.'

(b) ketâb-o xund-am book+râ read+1sgS 'I read the book.'

Karimi gives several arguments against viewing \(-râ\) as an accusative case marker. For example, \(-râ\) appears on noun phrases that are not theta-marked by the verb. There are two distinct instances of this. First, \(-râ\) can appear on adverbial noun phrases as shown in (23) below. Second, \(-râ\) can appear on noun phrases that occur in what Karimi calls clitic binder constructions as shown in (24) below. Clitic binder constructions will be discussed further in section 5.3.4, but for now they can be informally characterized as 'extracted'\(^8\) noun phrases coindexed with a clitic in the phrase in which they seem to have originated.

---

\(^7\)Bowers (1993), for example, calls direct objects secondary subjects.
\(^8\)I use this term since it is descriptively useful, however, I do not intend to suggest that there is syntactic movement involved in these constructions. As we will see in section 5.3.4, the 'extracted' DPs are base-generated as VP-adjuncts.
Clitic binder constructions can also involve direct objects. In such cases, both the direct object and the extracted noun phrase are marked with -râ, as shown in (25). Karimi argues that this double occurrence of -râ is a further reason for not treating it as an accusative case marker.

(25) mâshin-o dar-esh-o bast-am
    car+râ  door+3sg+râ closed+1sgS
    'As for the car, I closed its door.'  

Karimi claims that DPs marked with -râ are specific oblique noun phrases. We have seen why she considers them to be specific. By oblique case she means a case other than nominative. A further condition on -râ is that it must be outside the minimal governing domain of a lexical head. To resolve the problem of how a DP+râ is marked for oblique case outside the governing domain of a head, Karimi invokes the idea of case inheritance or transmission.\(^9\) Informally, this means that a DP is marked with -râ if it is coindexed with something bearing oblique case. Thus, when a DP is coindexed with a possessor, with the object of a preposition, or with a trace in direct object position, all non-nominative case positions, the DP is marked with -râ. She states the distribution of -râ as follows:\(^{10}\)

\(^9\)This notion comes from Safir (1985).
\(^{10}\)Presumably in (26) a is the relevant oblique-case-assigning head.
Specific-Oblique Principle

An NP must be marked by \( r\hat{a} \) in the following configuration:

\[
[\ldots \text{CP} \ldots \text{NP} \ldots [b \ldots \ldots ] \ldots ] \quad \text{where} \quad a = \text{the head of } b
\]

\[
e + \text{SPEC} \quad r
\]

\[
e - \text{NOM} \quad v
\]

[Karimi 1990:163.68]

Since the DP+\( r\hat{a} \) must be outside the government domain of a case assigning head, Karimi says it occurs either in adjunct position (adjoined to VP or IP), or in an A'-position ([SPEC, CP]). She states that discourse considerations (e.g. topicalization or focus) determine whether a noun phrase appears in a nonargument position.

To summarize, there are several facts that must be accounted for in an analysis of \(-r\hat{a}\): (a) its appearance on presupposed direct objects; (b) its appearance on certain adverbial noun phrases; (c) the phenomenon of clitic binder constructions; (d) the fact that two DPs marked with \(-r\hat{a}\) can appear in the same sentence; and (e) the fact that \(-r\hat{a}\) never appears on subjects. In the next three sections the first four facts are discussed. Subjects are taken up in section 5.5.

5.3.2 DP+\( r\hat{a} \) as the Direct Object of a Verb

A comparison of direct objects marked with \(-r\hat{a}\) with indefinite direct objects reveals that indefinite objects are closer to the verb. The presence of an indirect object brings out the difference as the following data from Browning & E. Karimi\(^{11}\) (1994) show:

\(^{11}\) There are two linguists, Simin Karimi and Ezat Karimi, both of whom are being cited here. To distinguish them I will identify the former as simply Karimi and the latter as E. Karimi.
(27) (a) hasan-râ be ali moa'refi kard-am
Hasan+râ to Ali introduced+1sgS
'I introduced Hasan to Ali.'

(b) be ali yek âdam-e xub moa'refi kard-am
to Ali one person+EZ good introduced+1sgS
'I introduced a good person to Ali.'

[Browning & E. Karimi 1994.17]

Browning and E. Karimi note that alternative word orders are possible in both cases, but native speakers judge such alternatives to involve scrambling.\(^{12}\) Thus, it appears that the unmarked position of a DP+râ is different from that of an indefinite DP, with respect to prepositional phrases.

DP+râ arguments and indefinite arguments also differ with respect to a reordering rule that Karimi (1989, 1994) calls XP-postposing. This rule allows arguments to follow the verb. Consider the sentences in (28):

(28) (a) sepide emruz se-tâ otâq-â-ro jâru kard-Ø
Sepide today three+unit room+pl+râ broom did+3sgS
'Sepide swept the three rooms today.'

(b) sepide emruz se-tâ otâq jâru kard-Ø
Sepide today three+unit room broom did+3sgS
'Sepide swept three rooms today.'

\(^{12}\)Actually, Karimi (1989, p. 200, fn. 11) states that the DP+râ may precede or follow the indirect object without any change in interpretation. However, this is not true of the indefinite object vis-a-vis the indirect object.
In (28a) the direct object *three rooms* is marked with *-rā* and is therefore definite and presupposed.¹³ It can be postposed around the verb as shown in (28c). In (28b), however, the direct object is indefinite. We see in (28d) that it cannot appear after the verb.

Why do DPs marked with *-rā* appear to be structurally higher and have more mobility within the clause than indefinite noun phrases? Although the answer to this question is likely to be structural, the fact that DPs marked with *-rā* are presupposed may also be relevant. In fact, the higher position of definite/specific/presupposed direct objects compared to that of indefinite direct objects has received a good deal of attention in recent work. Diesing (1992) states that presupposed direct objects appear to be structurally higher than true indefinites, citing Turkish as an example. Mohammad and Karimi (1992) cite Mahajan (1990) for Hindi, and Koopman and Sportiche (1991) for Dutch, as languages in which direct objects behave differently depending on their specificity/definiteness. Moorcroft (1995) discusses this phenomenon in Icelandic and German.

Interestingly, the appearance of definite, specific, or presupposed direct objects in a higher position than indefinite objects has been linked to case. For example, Browning and E. Karimi (1994), looking at Persian, suggest that the DP*-rā* is in a VP-external position. Their view is that noun phrases need both case assignment and case licensing. Case assignment takes place between a verb and its complement. Case can be licensed either by a

¹³Recall from Chapter 3, section 3.5.5, that the plural marker is licensed by the DP dominating the noun phrase.
verb or an inflectional head, however, the case licensed by a verb is incompatible with a specific interpretation. Therefore, a specific noun phrase must move to a VP-external position to be case licensed.

Similar proposals are found in the work of De Hoop (1992) and Moorcroft (1995). De Hoop proposes that there are two kinds of case that direct objects can bear, strong case and weak case. Strong case is associated with a strong interpretation for the noun phrase while weak case is associated with a weak interpretation. This accounts for not only the definite/specific vs. indefinite/nonspecific readings but also cases where the same items (e.g. quantifiers) can receive both a strong and a weak interpretation. De Hoop also proposes that only direct objects with strong case can scramble. This is because weak case requires adjacency. Recall that in Persian only the DP+râ can scramble, as we saw in (28) above.

Moorcroft (1995) develops a theory of case that involves three components. One of these components is called s-case and is crucially involved in licensing a presupposed argument. She proposes that a presupposed argument receives s-case in a structural position which c-commands the maximal projection of the category which assigns the other components of case to that argument. In this way she explains the fact that presupposed objects in Icelandic and German appear in a higher position than non-presupposed objects.

Let us assume then that there are two case positions in Persian for direct objects. One, adjacent to the verb, is where indefinite objects appear. However, in the spirit of the work cited above, let us assume that the case assigned by the verb to its noun phrase complement is incompatible with a presupposed reading for that noun phrase. Thus, such a noun phrase must adjoin to VP where it receives an oblique case and is marked with -râ. Recall that Karimi (1990) also proposes that the DP+râ occurs in adjunct position, but unlike her, I assume that DP+râ is adjoined only to VP. I assume that all other positions in which

---

14Note that this movement is implicit in Karimi's (1989, 1990) work also.
DP+râ is found are the result of scrambling operations (see section 5.6.1.2 for discussion).

The two positions for direct objects in Persian are shown in the trees below:

(29)

\[
\begin{array}{c}
\text{VP} \\
\text{VP} \\
\text{V'} \\
\text{DP_{indef}} \\
\end{array}
\]

\text{case/theta assignment}

(30)

\[
\begin{array}{c}
\text{VP} \\
\text{KP} \\
[\text{DP_{def}}_i] \quad -râ \\
\rightarrow t_i \\
\text{VP} \\
\text{V'} \\
\end{array}
\]

\text{oblique case}

\text{theta assignment}

To conclude, we have seen that presupposed direct objects occur in a different syntactic position from that of indefinite direct objects. Indefinite objects in complement position can receive case and a theta role from the verb. However, presupposed direct objects must adjoin to VP since the case assigned by the verb is incompatible with their semantic properties.\(^{15}\)

The DP+râ is therefore coindexed with an empty category (here a trace) which receives a

\(^{15}\)It should be pointed out that the base-generation of direct objects in two different positions has been proposed by Mohammad and Karimi (1992) for Persian. They have indefinite objects base-generated as sister to the verb and the DP+râ in [SPEC, VP]. I do not have DP+râ in [SPEC, VP] because in certain constructions more than one DP+râ is possible within a clause. While adjuncts can iterate, specifiers cannot. Mohammad and Karimi say that -râ is an inherent case marker, which means that it can only be assigned to a theta-marked noun phrase. As we will see in the following sections, under the present analysis the DP+râ does not directly receive a thematic role, and the notion of inherent case is therefore not applicable here.
theta role. This coindexing plays a role in clitic binder constructions as we will see in section 5.3.4.

5.3.3 DP+râ with Intransitive Verbs

Karimi (1990) shows that certain adverbials may appear with -râ. This is the basis of one of her arguments against -râ being an accusative case marker. Examples are given below. While Karimi translates these adverbials as topics, they also delimit the event described by the verb (see Ghomeshi and Massam (1994)). Note that the presence of -râ is optional. Without -râ these noun phrases are construed not as delimiting but as simply adverbial.\(^{16}\)

(31) (a) shab-e pish-o aslan na-xâbid-am
night+EZ last+râ at all neg+slept+1sgS
'As for last night, I didn't sleep at all.' [Karimi 1990:143.11]
('I didn't sleep the whole of last night')

(b) shab-e pish aslan na-xâbid-am
night+EZ last at all neg+slept+1sgS
'I didn't sleep at all last night.'

(32) (a) hafte-ye âyanda-ro esterâhat mi-kon-am
week+EZ coming+râ relax cont+do+1sgS
'As for next week, I will relax.' [Karimi 1990:143.12]
('I'll relax the whole week')

\(^{16}\)Sentence (31b) is better with the adverb dishab meaning 'last night'. shab-e pish is perhaps better translated as the 'preceding night'.
Karimi gives similar examples from written Persian, noting in a footnote that -râ can be absent but in such cases the interpretation of the adverbial element will be purely temporal.

Recall that for Karimi, a noun phrase is marked with -râ if it is specific and does not bear the feature [+NOM]. Furthermore the noun phrase must be outside the governing category of a head. Presumably, though Karimi does not state this explicitly, the adverbials in (31) and (32) are adjuncts and are thus outside the governing category of the verb. However, it is not clear that they need, or get, any sort of case. This leads her to propose that certain intransitive verbs, such as those in (31) and (32) above, are oblique case assigners.

There are some problems with Karimi's approach. First, she gives no criteria for distinguishing verbs that assign oblique case from those she calls 'purely intransitive' verbs. Second, she does not explain what happens when the adverbial is not marked with -râ. That is, since these adverbials can always also appear without -râ, do these verbs assign oblique case only optionally? Finally, the claim that these verbs are oblique case assigners does not explain the delimiting role played by adverbials marked with -râ.

Suppose that the appearance of -râ on adverbials indicates that these adverbials are actually direct objects (as opposed to indicating that -râ is not a direct object marker). Lazard (1992:191) makes essentially this point:

Certain expressions of time are marked with the postposition râ; the idea expressed by this construction is that one has occupied such an instant or space of time by doing such and such a thing. The expression of time is not then a simple adverbial phrase, but in a certain way the object of the idea expressed by the rest of the sentence...
Karimi translates the adverbial noun phrases marked with -rā as topicalized, while I have claimed that -rā marks presupposed DPs. However, recall that -rā takes a DP as its complement and that the null determiner is [+definite]. The semantics of these adverbials can be partially attributed to the presence of this null definite determiner. A similar contrast is found in English, in the difference between *I slept last night* and *I slept the night*. As support for the presence of a DP in these -rā-adverbials note that, according to Karimi, only adverbials that can be interpreted as nouns can precede -rā.

The same question arises in English as to the status of adverbials that appear to "measure out" (Tenny (1987)) or delimit the event described by the verb in sentences such as *I slept the whole night*, or *I ran a mile*. We can also include in this class cognate objects appearing in sentences such as *I danced a short dance, He smiled a wicked smile*, etc.

Unfortunately, a detailed discussion of all the properties of these adverbials is beyond the scope of this thesis. However, I suggest that they should be considered as direct objects, just like the noun phrases considered in the previous section. This view is supported by the fact that they have a number of prototypical direct object properties (Ghomeshi and Massam (1994)). Further, there is a precedent in the literature for this view (Massam (1990) for cognate objects). The question of which adverbials can function as direct objects and what, if any, thematic role they receive I leave for future research.

So far we have seen that DPs marked with -rā are base-generated as sisters to the verb. However, the case assigned by the verb is incompatible with a presupposed reading for the DP. This constituent must therefore move to adjoin to VP in order to receive case. As a result, the DP+rā is coindexed with a trace in direct object position. In the next section I turn to what Karimi calls clitic binder constructions, and show that these constructions can be easily accounted for if the DP+rā can be coindexed with another empty category, namely *pro*. In such cases, DP+rā functions as a VP-level topic. The fact that -rā marks VP-level topics may also play a role in the semantic interpretation of adverbial noun phrases discussed above, which Karimi translates as topics.
5.3.4 DP+rā and VP-level Topicalization

Let us now turn to the clitic binder constructions. These constructions allow two occurrences of -rā in a clause, leading Karimi to claim that -rā cannot be an accusative case marker. Consider the example below in (33):

(33) (a) mâshin-o dar-esh-o bast-am
car+rā door+3sg+rā closed+1sgS
'As for the car, I closed its door.' [Karimi 1990:143.13]

(b) dar-e mâshin-o bast-am
door+EZ car+rā closed+1sgS
'I closed the door of the car.'

Karimi states that the construction in (33a), illustrating apparent extraction out of an Ezafe construction, has been discussed in the Persian literature under the heading fakke-ezafe (Ezafe disjunction). She proposes to treat examples such as (34), where the DP+rā is coindexed with the object of a preposition, as instances of the same phenomenon. She refers to these constructions in general as clitic binders.

(34) sepida-ro beh-esh goft-am
Sepide+rā to+3sg told+1sgS
'As for Sepide, I told her.' [Karimi 1990:161.65]

These constructions, all of which involve the apparent extraction of a DP, exhibit a number of interesting properties. First, if this DP is extracted from a direct object, there are two occurrences of -rā as in (33) above. Note that clitic binders involving indefinite direct objects are not possible. This follows from the fact that only a DP can be extracted and the
only DP within the noun phrase is the possessor. Only definite DPs can have possessors (see Chapter 3) and they will therefore be marked with -râ.

Karimi points out that -râ cannot appear within what I have been calling the Ezafe domain. The relevant example is (35c) below.

(35) (a) dar-e otâq-e sepida-ro bast-am  
        door+EZ room+EZ Sepide+râ closed+1sgS  
        'I closed the door of Sepide's room.'

       (b) sepida-ro dar-e otâq-esh-o bast-am  
           Sepide+râ door+EZ room+3sg+râ closed+1sgS  
           'As for Sepide, I closed the door of her room.'

       (c) *dar-e otâq-o sepida-ro bast-am  
           door+EZ room+râ Sepide+râ closed+1sgS  
           [Karimi 1990:172.93]

For Karimi (35c) is ungrammatical because the head noun, dar 'door', governs otâq 'room' and -râ cannot appear on any noun phrase within the government domain of a lexical head. For us, -râ is a phrasal affix that takes a DP as its complement. Since, as shown in Chapter 2, the Ezafe domain does not include phrasal material, -râ cannot appear within this domain.

Looking at (35a) it is interesting to note what can be extracted and what cannot. Given a direct object such as the room's door, the DP room can be extracted. This is shown in (36b).

(36) (a) dar-e otâq-o bast-am  
        door+EZ room+râ closed+1sgS  
        'I closed the door of the room.'

       (b) otâq-o dar-esh-o bast-am  
           room+râ door+3sg+râ closed+1sgS  
           'As for the room, I closed its door.'
If we now consider a direct object such as *Sepide's room's door* (see (37a)), we see that *Sepide* can be extracted (see (37b)), as can *Sepide's room* (see (37c)). However, *room* alone cannot (see (37d)).

(37) (a) dar-e otâq-e sepida-ro bast-am
  door+EZ room+EZ Sepide+rà closed+1sgS
  'I closed the door of Sepide's room.'

(b) sepida-ro dar-e otâq-esh-o bast-am
  Sepide+rà door+EZ room+3sg+rà closed+1sgS
  'As for Sepide, I closed the door of her room.'

(c) otâq-e sepida-ro dar-esh-o bast-am
  room+EZ Sepide+rà door+3sg+rà closed+1sgS
  'As for Sepide's room, I closed its door.'

(d) *otâq-o dar-esh-e sepida-ro bast-am
  room+rà door+3sg+EZ Sepide+rà closed+1sgS
  *'As for the room, I closed its door of Sepide.'

This can be straightforwardly accounted for. If only DPs can be extracted and if the noun phrase in (37a) has the structure shown in (38) below, then we expect that DP$_3$ can be extracted corresponding to (37b) and that DP$_2$ can be extracted corresponding to (37c). However, (37d) is ungrammatical since *room* does not head its own DP.
Returning to the properties of the clitic binder construction in general, we have already seen that a DP can be extracted from a prepositional phrase (see (34)). In this case the extracted DP is also marked with -râ. Karimi also gives examples in which an extracted DP is coindexed with more than one possessor. In (39) below the extracted DP Hushang is coindexed both with the possessor in the direct object and with the possessor in the prepositional phrase.

(39) hushang-o man dâstân-esh-o be mâdar-esh goft-am
    hushang+râ I story+3sg+râ to mother+3sg told+1sgS
    'As for Hushang, I told his story to his mother.' [Karimi 1990:144.17]

Let us now consider extraction from subjects. In constrast to DPs extracted from direct objects and prepositional phrases, a DP extracted from a subject is not marked with -râ. In (40a) below we see a subject containing a possessor. Recall that subjects are not marked with -râ, as shown in (40b). In (41a) we see the corresponding clitic binder construction. The ungrammaticality of (41b) shows that a DP extracted from a subject cannot be marked with -râ.

(40) (a) barâdar-e sepide âmad-Ø
    brother+EZ Sepide came+3sgS
    'Sepide's brother came.'
(b) *barâdar-e sepeida-rö  âmâd-Ø
      brother+EZ Sepide-râ  came+3sgS

      [Karimi 1990:169.85]

(41)  (a)  sepide  barâdar-esh  âmâd-Ø
      Sepide  brother+3sg  came+3sgS
      'As for Sepide, her brother came.'

      (b) *sepide-ro  barâdar-esh  âmâd-Ø
      Sepide+râ  brother+3sg  came+3sgS

      [Karimi 1990:169.86]

The absence of -râ on a possessor extracted from a subject is accounted for if -râ only marks DPs adjoined to VP. However, the grammaticality of (41a) must still be explained. This issue is taken up in section 5.4.

Let us now consider Karimi's analysis of clitic binder constructions such as (33a) above, repeated here:

(33)  (a)  mâshin-o  dar-esh-o  bast-am
      car+râ  door+3sg+râ  closed+1sgS
      'As for the car, I closed its door.'

      [Karimi 1990:143.13]

She claims that the specific direct object ( 'its door') inherits [-NOM] case from its trace in the verb phrase. The noun phrase 'car', being coreferential with -esh, inherits specificity and [-NOM] case from the clitic pronoun. Further, since neither of these noun phrases is in the governing category of a lexical head, -râ shows up on both, marking them for specificity and oblique case.

Let us now pursue an account of these constructions that does without case inheritance. So far we have proposed that -râ can case-mark any adjunct to VP. The adjuncts we have considered have been coindexed with a trace within VP. Note, however, that in each of the
clitic binder constructions the DP+râ is also coindexed with something. Recall from Chapter 4 that the pronominal enclitics appear on a head in order to identify a pro argument of that head. Thus, in the cases above, the DP+râ is coindexed with the pro possessor or the pro object of a preposition. This is illustrated below in (42) and (43) respectively.

(42) (a) mâšin-o dar-esh-o bast-am
car+râ door+3sg+râ closed+1sgS
'As for the car, I closed its door.'

(b)

In (42b) above the direct object position (sister to V) contains a trace coindexed with the lowest DP+râ. This DP in turn contains a pro possessor coindexed with the higher DP+râ.

(43) (a) sepida-ro beh-esh goft-am
Sepide+râto+3sg told+1sgS
'As for Sepide, I told her.'

(b)
In (43b) the prepositional phrase is base-generated as an adjunct to VP. Its object is *pro*, with which the DP+*râ* is coindexed.

We are now in a position to state the conditions under which a DP+*râ* is licensed. Note that the DP+*râ* as direct object forms a chain with the trace in VP. The only difference between this and the cases where it is coindexed with *pro* is that in the latter cases the chain does not arise from movement. However, in all other respects *pro* functions like a trace in that it does not receive case (the relevant case being absorbed by the clitic) but does receive a theta-role. Note also that all the cases we have seen involve a DP+*râ* coindexed with an empty category that it c-commands. That is, it is not possible to coindex a DP+*râ* with a *pro* in subject position. These facts can be captured by the principle given in (44).

(44)

<table>
<thead>
<tr>
<th>Thematic Licensing Condition (Version 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A KP headed by -<em>râ</em> must be coindexed with and c-command an empty category in a theta position</td>
</tr>
</tbody>
</table>

I propose that this is a type of topicalization, similar to left dislocation in English. As we will see, a similar process can be observed at the clause level. I will refer to DPs marked with -*râ*, particularly those coindexed with a *pro*, as VP-level topics hereafter.

---

17Examples involving three DPs marked with -*râ* are rare, however, the fact that such examples can be constructed (see (i)) provides further evidence against a movement analysis. Note that the 'unextracted' version of the sentence is ungrammatical.

(i) lubiýâ-ro poxt-ash-o sard-esh-o dust dâr-am beans+râ cooked+3sg+râ cold+3sg+râ like+1sgS
   'As for beans, when they're cooked, I like them cold.'

(ii) *sard-e poxte-ye lubiýâ-ro dust dâr-am cold+EZ cooked+EZ beans+râ like+1sgS

18I assume that in a noun phrase that contains an NP, DP and KP projection, the projections all share the same index.
The view that -râ is available as a case-assigner for any thematically licensed DP accounts for another construction in which DP+râ can occur. This is the case where both a clitic and DP+râ appear, serving as the direct object. Lazard (1992:189-190) has the following to say about these constructions:

In colloquial language a direct object (followed by râ) is frequently taken up again by a personal suffix affixed to the verb...The verb may be separated from the object by other words or may follow it immediately...This construction is especially used when the object, being topicalized, is placed at the beginning of the sentence...

An example is given below:

(45) (a) sib-o xord-am-esh
    apple+râ ate+1sgS+3sg
    'As for the apple, I ate it.'

(45) (b) [Diagram of sentence structure]

As in sentences with DP+râ possessors and objects of prepositions, the DP+râ in (45) is coindexed with pro. In this case, however, pro is a direct object. This example reinforces the topic-like nature of DPs marked with -râ.
The Thematic Licensing Condition, as formulated above, states that DPs marked with -\(r\)a can only be licensed by empty categories.\(^{19}\) However, it is not clear that this stipulation is required, since all other cases can be ruled out by the Binding Theory.

Let us consider the following (ungrammatical) sentences in which a VP-level topic is coindexed with a proper name or a pronoun.

(46)  *jián-o be jián ye ketâb-i dâd-am
      Jian+râ to Jian one book+indef gave+1sgS
      *'As for Jian, I gave a book to Jian.'

(47)  *otâq-o dar-e un-o bast-am
      room+râ door+EZ it+râ closed+1sgS
      *'As for the room, I closed the door of it.'

(48)  *sepida-ro be u ye ketâb-i dâd-am
      Sepide+râ to her one book+indef gave+1sgS
      'As for Sepide, I gave a book to her.'

All the examples are ruled out by binding conditions B and C. As shown in Chapter 4, \(pro\) is the only pronominal element that can be coindexed with a c-commanding antecedent within the same clause. This is because the constituent within which it occurs always contains an accessible subject (the pronominal enclitic that serves to identify it). That constituent thus serves as the governing category for \(pro\), and \(pro\) can be bound by an antecedent outside the constituent within which it occurs but within the same clause. Since traces are neither pronominal expressions nor R-expressions they also escape Conditions B and C of the Binding Theory. Thus the Thematic Licensing Condition can be stated as though any noun

\(^{19}\)The question of why traces and clitics can occur in clitic binder constructions but pronouns and proper names cannot is a problem for Karimi's theory. That is, she doesn't explicitly state why only clitics and traces can transmit case.
phrase can license a VP-level topic, with the ungrammatical examples being ruled out by the Binding Theory.\textsuperscript{20}

A more general version of the Thematic Licensing Condition is given in (49).

(49) \textbf{Thematic Licensing Condition (Version 2)}
A KP headed by -rā must be coindexed with and c-command a theta-marked nominal argument

The Thematic Licensing Condition predicts that both a delimiting adverbial and a VP-level topic should be possible in the same clause. This prediction is borne out as shown in (50a). In (50b) and (50c) the adverbial and the object of the preposition are shown in their non-rā forms.

(50) (a) bachcha-ro faqat in ye sā'at-o bâ-sh bâzi kon
child+rā only this one hour+rā with+3sg play do(imp)
'As for the child, only this one hour play with him.' [Karimi 1990:165.73]

\textsuperscript{20}There is one possible problem with this view. The Binding Theory might be taken to predict that anaphors, which \textbf{can} be bound within the clause, should also license VP-level topics. This is not so, as the examples below show.

(i) *jiān-o be xod ye ketâb-i dâd-am
   Jian+rā to self one book+indef gave+1sgS
   '*As for Jian, I gave a book to (him)self.'

(ii) *jiān-o barâdar-e xod-o did-am
    Jian+rā brother+EZ self+rā saw+1sgS
    'As for Jian, I saw his own brother.'

However, Condition A of the classic Binding Theory requires anaphors to be bound by a c-commanding \textbf{argument}. If VP-level topics are in an A\textsuperscript{'}-position, they will be unable to bind an anaphor.
(b) bachcha-ro (barâ-ye) ye så'at bâ-sh bâzî kon
    child+râ (for+EZ) one hour with+3sg play do(imp)
    'As for the child, play with him (for) an hour.'

(c) ye såat bâ bachche bâzî kon
    one hour with child play do(imp)
    'Play with the child for an hour.'

Note, however, that it is not possible to have two delimiting noun phrases marked with -râ. That is, an event can only be delimited once (Tenny (1987)). On the assumption made here that delimiting arguments are generated in direct object position, a sentence such as (51) is ruled out because it is not possible to generate two sisters to V. Thus, while one of the DPs marked with -râ will be licensed by being coindexed with its trace, the other will not be thematically licensed.

(51) *ketâb-o faqat in ye så'at-o be-xun
    book+râ only this one hour+râ subj+read
    *'As for this book, only this one hour read.'

Finally, let us consider the interaction between the Thematic Licensing Condition and the Reflexivity Condition, given in Chapter 4 and repeated below.

(52)

<table>
<thead>
<tr>
<th>Reflexivity Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>A reflexive predicate must be reflexive-marked.</td>
</tr>
</tbody>
</table>

Recall that a reflexive predicate has two coindexed coarguments. In other words, if two noun phrases that receive theta-roles from a single predicate are coindexed, the Reflexivity Condition requires one of them to be an anaphor. The difference between such a case and the cases we are considering here is that one of the coindexed arguments (namely the DP+râ)
does not receive a theta-role at all. In (53a) below there are two coarguments Jian and the clitic -esh and they cannot be coindexed. In (53b) we see that if one of these is an anaphor the coindexing is possible. (54) has the same preposition but a different verb. Jian does not receive a theta-role from the verb and thus can be coindexed with, and licensed by, the clitic. In fact, in this case Jian and the clitic must be coindexed because otherwise Jian is not thematically licensed.

(53) (a)  jiān₁-o  b-esh*i/j moa'refî kard-am
        Jian+râ to+3sg introduced+1sgS
        'I introduced Jian to him/her.'

        (b)  jiān₁-o  xe xodeshi mo'a'refî kard-am
        Jian+râ to self+3sg introduced+1sgS
        'I introduced Jian to himself.'

(54)  jiān₁-o  b-eshi/*j ye ketâb-i dâd-am
        Jian+râ to+3sg one book+indef gave+1sgS
        'As for Jian, I gave a book to him.'

5.3.5 Section Summary

In this section we have considered two possible analyses of -râ. Under Karimi’s (1990) analysis, -râ marks an oblique case that is transferred or inherited from a coindexed trace or clitic. The view argued for here is that -râ is a freely available case-assigner for DPs adjoined to VP. To constrain the occurrence of such DPs the Thematic Licensing Condition has been proposed, requiring the DP+râ to be coindexed with a noun phrase in an argument position.

---

21Note that, as in English, this sentence is pragmatically odd but syntactically well-formed.
These two views are very similar. One relies on a notion of case inheritance, the other on a notion of theta-role inheritance. However, there are several advantages to the second view. For example, the licensing of topic-like constituents is now characterized in semantic rather than in syntactic terms. As we see in the next section, the Thematic Licensing Condition is independently motivated, in that it also accounts for the licensing of clause-external topics. Without Karimi's syntactic condition on 

\[-râ\], there is no need to refer to all non-nominative cases as a class, the notion of case transmission is no longer required, and we can dispense with the idea that some intransitive verbs are oblique case assigners. Further, the more constrained view of the syntactic distribution of DPs marked with \[-râ\] explains the absence of \[-râ\] on clause-external topics.

Since DP+\(râ\) only occurs adjoined to VP, and given that \(-râ\) most often marks direct objects, we must consider whether the DP+\(râ\) is adjoined to VP, or whether it occupies the specifier of a functional projection such as AgroP (cf. Chomsky (1993) for example).

One phenomenon favouring the adjunction view is the double occurrence of DP+\(râ\). Consider example (50a), repeated below:

(50) (a) bachcha-ro \(\text{faqat in ye sâ'at-o bâ-sh bâzi kon}\)
child+râ only this one hour+râ with+3sg play do(imp)

'As for the child, only this one hour play with him.' [Karimi 1990:165.73]

If DP+\(râ\) occupied [SPEC, AgroP], an example like this would require two AgroP projections, an otherwise unmotivated structural complication. I take this as an argument against placing DP+\(râ\) in the specifier of a particular functional projection.

5.4 Clause-Level Topicalization

We have seen that DPs marked with \(-râ\) are viewed as VP-level topics, semantically licensed by being coindexed with a noun phrase in an argument position. I will now show
that a similar phenomenon exists at the sentence level. The crucial difference, however, is that clause-level topics are not overtly marked for case.

Let us first consider what Thackston calls the Resumptive Sentence Construction.

A common type of sentence in Persian is the resumptive (or topic-comment) sentence, wherein a topic which is not the subject of the main verb is introduced as subject of the sentence; the predicate then affords some information about that subject. Since the sentence-subject is not the subject of the main verb, the sentence-subject must be referred to pronominally in its proper position in the predicate. Both the independent and the enclitic pronouns are used for this purpose. [Thackston 1983:92]

(55) ân zan-i-ke darbâre-ash sohbat mi-kard-im:
that woman+indef+rel about+3sg talk cont+did+1plS
az xâne-ash radd mi-shod-am
from house+3sg pass cont+became+1sgS
'That woman that we were talking about, I passed by her house'22
('I passed by the house of that woman we were talking about.') [Thackston 1983:92]

(56) ân qâliche: fekr ne-mi-kon-id qeymat-esh kam-iziâd-ast?
that carpet thought neg+cont+do+2plS price+3sg little much+3sgS
'That carpet, don't you think its price is a bit too much?'
('Don't you think the price of that carpet is a bit too much?') [Thackston 1983:92]

In the examples above, there are obvious similarities to the VP-level topic constructions that were discussed in section 5.3.4. Here clause-initial DP is coindexed with a pronominal enclitic which, in the examples above, is a possessor. However, there are two important

22I have changed the translation of this and the following example to more accurately reflect the topic-comment structure of the sentences. Thackston's translation is given in parentheses.
differences between the two constructions. First, note that the topicalized DP here is not marked with \(-râ\). Second, a full pronoun can occur in place of the pronominal enclitic within the main clause, whereas with VP-level topics, a coindexed full pronoun was impossible. Sentence (55) is repeated as (57a) below, with the topicalized constituent and the enclitic it is coindexed with in bold-face. The same sentence with a full pronoun instead of the enclitic is shown in (57b). In (57c) we see the non-topicalized version of the sentence.

(57) (a) \(ân \ ) zan-i-ke \[ \text{darbâre-ash sohabat mi-kard-im:} \] 
that woman+indef+rel about+3sg talk cont+did+1plS

az xâne-ash radd mi-shod-am
from house+3sg pass cont+became+1sgS
'That woman that we were talking about, I passed by her house'

(b) \(ân \ ) zan-i-ke \[ \text{darbâre-ash sohabat mi-kard-im:} \] 
that woman+indef+rel about+3sg talk cont+did+1plS

az xâne-ye u radd mi-shod-am
from house+EZ her pass cont+became+1sgS
'That woman that we were talking about, I passed by her house'

(c) az xâne-ye \(ân \ ) zan-i-ke \[ \text{darbâre-ash sohabat} \] 
from house+EZ that woman+indef+rel about+3sg talk

\[ \text{mi-kard-im } \] radd mi-shod-am
cont+did+1plS pass cont+became+1sgS
'I passed by the house of that woman we were talking about.'

I propose that these topics are licensed by the Thematic Licensing Condition. In other words they are coindexed with a noun phrase that gets a theta-role. In this sense, they are exactly like the VP-level topics. However, they receive case by being in the specifier of a
Topic Projection (TopP). This projection has been posited by Brunson (1992) to account for base-generated topics in English. The structure of a Persian sentence involving one of these topics given in (58).²³

(58)

This structure accounts for the lack of overt case marking on the topic DP, which receives case from the null Top head. It also explains why pronouns can be coindexed with the topic, since the governing category for the full pronouns is IP, and the topic DP lies outside this domain.

Given that DPs can appear in [SPEC, TopP], we must now reformulate the Thematic Licensing Condition so as to include these DPs as well. Let us refer to both DPs adjoined to VP where they are marked with -rā and DPs in [SPEC, TopP] as being DP topics. The revised Thematic Licensing Condition is given in (59).

(59)

<table>
<thead>
<tr>
<th>Thematic Licensing Condition (Final Version)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A DP topic must be coindexed with and c-command a theta-marked nominal argument</td>
</tr>
</tbody>
</table>

²³Brunson's TopP, proposed for English, is head-medial. I assume that the Persian TopP is head-final, like other functional categories in the language.
Let us now return to the one clitic binder construction not accounted for in section 5.3.4, namely clitic binders that involve subjects. Recall that unlike all other "extracted" DPs that are coindexed with an enclitic, a DP extracted from a subject is not marked with -rā.

(60) (a) barādar-e sepide āmad-Ø
    brother+EZ Sepide came+3sgS
    'Sepide's brother came.'

(b) sepide barādar-esh āmad-Ø
    Sepide brother+3sg came+3sgS
    'As for Sepide, her brother came.'

Recall that Karimi's analysis of clitic binder constructions involves case inheritance from the enclitic. To account for cases like (60b) above, Karimi has to further specify that the case inherited by the extracted noun phrase cannot be nominative. Under her account this explains why the extracted noun phrase in (60b) is not marked with -rā.

Given the fact that a Topic Phrase is needed in order to explain the topicalized constructions above, a straightforward analysis of the clitic binder construction out of subject position is available. That is, the "extracted" DP can be base-generated in the specifier of the TopP and receive case from the Top head.24

---

24We have seen that clause-level topics, in contrast with VP-level topics, can be coindexed with a full pronoun. The explanation for this is that the clause-level topic is base-generated outside of IP, which is the governing category for the pronoun. However, this predicts that clause-level topics should be able to be coindexed with full pronouns occurring within subjects also. This is not so, as shown in the following example.

(i) (a) sepide barādar-esh āmad-Ø
    Sepide brother+3sg came+3sgS
    'As for Sepide, her brother came.'

(b) *sepide barādar-e u āmad-Ø
    Sepide brother+EZ her came+3sgS
    'As for Sepide, her brother came.'
Interesting questions remain with respect to the choice between VP-level and clause-level topicalization. For example, the latter is favoured when the topicalized noun phrase is heavy, containing a relative clause, for example. However, these issues do not bear on the analysis being developed here.

5.5 The Position of Subjects and Predication Phrase

In our discussion of DPs marked with -râ we have not considered the position of subjects. It is to this issue that we now turn. Consider the following structure for a transitive sentence with a definite direct object, assuming that subjects are base-generated in [SPEC, VP] (Koopman and Sportiche (1985, 1991), Fukui and Speas (1986), Kitagawa (1986), Kuroda (1988), and many others).

Since the F-features that are checked under I should serve as an accessible subject, the governing category for the pronoun possessor in (ib) above should be IP. The 'heavyness' of the topicalized constituent seems to play a role in whether or not a pronoun can appear, however. Thus, perhaps the presence or absence of a pronoun in this case is not to be accounted for by Binding Theory but rather discourse principles and 'Avoid Pronoun'.
As indicated by the subscripts, it is assumed that the subject moves from \([\text{SPEC}, \text{VP}]\) to \([\text{SPEC}, \text{IP}]\), while the direct object moves from complement position within the VP to adjoin to VP. However, we must consider why the derivation proceeds in this fashion. That is, what prevents the subject from adjoining to VP and the direct object from moving to \([\text{SPEC}, \text{IP}]\)?

In Chomsky (1993) this issue is discussed although with respect to a slightly different structure. In his work it is assumed that the object moves to the specifier of an intermediate projection called \(\text{Agr}_\text{OP}\), while IP is referred to as \(\text{Agr}_\text{SP}\). However, the same question arises as to why the subject and direct object must move in crossing rather than nested paths. Chomsky provides a theory which achieves the desired result. However, there is a further problem not accounted for by the analysis in Chomsky (1993).

Let us consider sentences involving intransitive verbs. Under Chomsky's system a language may choose which Agr projection is "active". The single noun phrase in an intransitive clause then moves to the specifier of that Agr projection. If the active Agr projection is \(\text{Agr}_S\) then that noun phrase will pattern with subjects of transitive clauses. If the active Agr projection is \(\text{Agr}_O\) then the noun phrase will pattern with objects of transitive clauses. This corresponds to Nominative-Accusative and Ergative-Absolutive languages, respectively.

In Persian we have suggested that the DP+\(r\dot{a}\) is not in the specifier of a functional projection but rather is simply adjoined to VP. Thus even if there is one "active" Agr projection in an intransitive clause, nothing rules out the subject raising to adjoin to VP and getting case-marked by \(-r\dot{a}\). Note that the Thematic Licensing Condition would be met in this configuration also.

I suggest that the solution to this problem is that subjects are base-generated above VP. So far I have used IP as the label for the projection above VP. However, for reasons that will become clearer in the next section, I follow Bowers (1993) in assuming that subjects are base-generated in the specifier of \(\text{Pr}()\text{edication} P()\text{hrase}\). This is illustrated below:
Bowers (1993) states that subjects are base-generated in [SPEC, PrP] but move to the specifier of a higher projection to get nominative case. He takes this view because it is important for his analysis of certain constructions that [SPEC, PrP] not be a case position. I modify his claim slightly here. While it is true that [SPEC, PrP] is not a case position in that the head Pr is not a case-assigner, I propose that it is within this projection that the assignment of nominative case takes place. Recall that in Chapter 4, section 4.1, it was argued that the subject agreement affixes assign nominative case. They do so when the verb moves to the Pr head. If the verb does not move to the Pr head there will be no case available for the specifier position. Note that this view is similar to the one found in Chomsky (1993) where case is assigned to the specifier of AgrOP or AgrSP but only by virtue of the verb or tense (respectively) having moved to the head of the projection.

The fact that the subject agreement affixes assign nominative case and are checked by the noun phrase to which they assign this case provides a further argument that subjects are not base-generated within VP. Given that the subject agreement affixes are already on the verb when it is inserted under V in the syntax (having been attached at the LSI), if the subject was base-generated in [SPEC, VP], nominative case assignment and checking could take place within VP. However, there is clear evidence from the default word order (where subjects
appear before DP+rā), as well as from VP-joined adverbs (see next section), that the subject is outside VP in the syntax.

5.6 Transitive Compound Verbs

Persian verbal compounds, or complex predicates, have been discussed by many linguists, (Moyne (1970), Samiian (1983), Heny and Samiian (1991), Mohammad and Karimi (1992), Ghomeshi and Massam (1994) (in which the aforementioned two papers are summarized) and Vahedi (1995)). Let us consider their properties.

Most verbs in Persian are compound verbs. They consist of a non-verbal element that may be a noun, adjective, adverb or participle. The verb is usually a 'light' verb, i.e. one with little inherent semantic content.²⁵ Samiian (1983) gives two arguments that compounding is the means of verbal innovation in Persian. First, she shows cases where simplex verbs have been replaced by compounds. Some examples are shown below, taken from Samiian (1983.258) with the light verbs kardan 'to make/do' or dâdan 'to give', and where the non-verbal element is an adjectival or nominal form of another verb.

²⁵This term originates with Jespersen (1954). See also Grimshaw and Mester (1988) for an analysis of complex predicates involving light verbs in Japanese that differs from the one presented below.
The second argument Samiian gives for the productivity of compounding comes from borrowing. She states that foreign words are generally treated as nominals and verbal compounds are built again using a small set of light verbs. Some examples are given in (64).

(64)  
<table>
<thead>
<tr>
<th>Compounding Type</th>
<th>Example</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal Compound</td>
<td>tayp kardan</td>
<td>'to type'</td>
</tr>
<tr>
<td>Nominal Compound</td>
<td>fotokopi kardan</td>
<td>'to photocopy'</td>
</tr>
<tr>
<td>Nominal Compound</td>
<td>telefon kardan</td>
<td>'to call'</td>
</tr>
<tr>
<td>Nominal Compound</td>
<td>shut kardan</td>
<td>'to shoot'</td>
</tr>
<tr>
<td>Nominal Compound</td>
<td>eski kardan</td>
<td>'to ski'</td>
</tr>
</tbody>
</table>

As stated above, verbal compounds can involve adjectives, adverbs, particles and participles as well as nouns. The following examples are taken from Ghomeshi and Massam (1994).

(65)  
(a) **Adjective+Verb**

<table>
<thead>
<tr>
<th>Example</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>tang kardan</td>
<td>tight+to do/make 'to tighten'</td>
</tr>
<tr>
<td>bidâr kardan</td>
<td>awake+to do/make 'to wake up' (tr.)</td>
</tr>
<tr>
<td>bidâr shodan</td>
<td>awake+to become 'to wake up' (intr.)</td>
</tr>
<tr>
<td>tamiz kardan</td>
<td>clean+to do/make 'to clean'</td>
</tr>
</tbody>
</table>

(b) **Adverb/Preposition/Preverb+Verb**

<table>
<thead>
<tr>
<th>Example</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>pish raftan</td>
<td>forward, in front+to go 'to progress'</td>
</tr>
<tr>
<td>bâlâ keshidan</td>
<td>up, top, above+to pull 'to hoist, elevate'</td>
</tr>
<tr>
<td>pas dâdan</td>
<td>behind, back+to give 'to give back'</td>
</tr>
<tr>
<td>bar dâshtan</td>
<td>up, on+to have 'to pick up'</td>
</tr>
</tbody>
</table>

(c) **Participle+Verb**

<table>
<thead>
<tr>
<th>Example</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>parâkande kardan</td>
<td>scattered+to do/make 'to scatter'</td>
</tr>
<tr>
<td>parâkande shodan</td>
<td>scattered+to become 'to be scattered'</td>
</tr>
<tr>
<td>âsude kardan</td>
<td>tranquil, soothed+to do/make 'to give peace of mind'</td>
</tr>
</tbody>
</table>
Following Ghomeshi and Massam (1994), and in contrast to most other work cited above, I assume that all instances of compound verbs have the same syntactic structure. That is, compounds containing nouns are not treated differently from those with adjectives, adverbs and participles.

In Ghomeshi and Massam (1994) it is argued that verbal compounds are $X^0$-adjoined structures, as shown below.

\[
\begin{array}{c}
\text{\textit{V}}^0 \\
X^0 \\
\text{\textit{V}}^0
\end{array}
\]

This proposal accounts for several properties of verbal compounds. First, they form a single intonational unit, with the entire compound receiving only one stress. Thus, while stress usually occurs on the final syllable of the verb stem, in compounds it occurs on the final syllable of the non-verbal element. This is shown below with stress indicated in boldface. In (67a) the noun does not form a compound with the verb and the verb receives stress. In (67b) the same noun is used in a compound structure, and the verb does not receive any stress.

\[(67)\]

\[(a) \quad \text{hamum-o shost-am} \]

\[\text{bath+râ washed+1sgS} \]

'I washed the bath(tub).'

\[(b) \quad \text{bachcha-rohamum kard-am} \]

\[\text{child+râ bath did+1sgS} \]

'I bathed the child.'

This analysis also accounts for the fact that nothing can intervene between the two parts of a compound, while this is not true of an XP - V sequence. Consider the following examples in
which the adverb *diruz* 'yesterday' appears between the verb and its direct object ((68), (69a)) but not within a compound structure (69b).

(68)  
\[
\text{hamum-o diruz shost-am} \\
\text{bath+râ yesterday washed+1sgS} \\
\text{'I washed the bath(tub) yesterday.'}
\]

(69) (a)  
\[
\text{bachcha-ro diruz hamum kard-am} \\
\text{child+râ yesterday bath did+1sgS} \\
\text{'I bathed the child yesterday.'}
\]

(b)  
\[
*\text{bachcha-ro hamum diruz kard-am} \\
\text{child+râ bath yesterday did+1sgS}
\]

Finally the X⁰-adjunction structure accounts for the fact that while verbal compounds function in some respects as a single lexical item, they nevertheless are syntactically transparent. Recall that with simple transitive verbs, pronominal enclitics can appear on the verb, licensing a *pro* direct object. The pronominal enclitics can also appear with transitive verbal compounds, however, in this case they appear on the non-verbal element.

(70) (a)  
\[
\text{bachcha-ro hamum kard-am} \\
\text{child+râ bath did+1sgS} \\
\text{'I bathed the child.'}
\]

(b)  
\[
\text{hamum-esh kard-am} \\
\text{bath+3sg made+1sgS} \\
\text{'I bathed him/her.'}
\]

This section presents an alternative analysis to the one proposed by Ghomeshi and Massam, which accounts for some properties of verbal compounds not previously explained.
Specifically, we will see evidence that verbal compounds involve phrasal constituents. The revised analysis also accounts for the predication relation between the direct object and the non-verbal element of the compound in sentences such as (71a) and captures the apparent identity between this and the one between the subject and a predicate in a copular construction, such as (71b).

\[(71)\]
\[(a)\] bijan-o bidâr kard-am  
\hspace{1cm} Bijan+râ awake did+1sgS  
'I woke Bijan up.' (lit. 'I made Bijan awake.')

\[(b)\] bijan bidâr bud-Ø  
\hspace{1cm} Bijan awake was+3sgS  
'Bijan was awake.'

The compounds discussed below are semantically transparent. There are many other compounds whose meaning is no longer transparent (e.g. dust dâshtan 'to like (something)', lit. friend+to have), or where the non-verbal element no longer appears elsewhere as an independent lexical item (e.g. pas dâdan 'to return, give back (something), lit. back?+to give). It is assumed that such examples are syntactically identical to the semantically transparent compounds regardless of what accounts for the differences in the way their meaning is derived.

5.6.1 On the Phrasality of the Non-Verbal Element

5.6.1.1 Complements to the Non-Verbal Element

This section argues that the non-verbal elements of verbal compounds are phrasal. This evidence has to do with their 'separability' (in Thackston's (1983) terms). Consider, for example, the compound vâred shodan 'to enter'. This consists of the verb shodan 'to become'
plus a non-verbal element. While the construction in (72a) is possible, (72b) is the more usual way to express the same proposition.

(72) (a) be in shahr vâred shod-am
to this city enter became+1sgS
'I entered this city.'

(b) vâred-e in shahr shod-am
enter+EZ this city became+1sgS
'I entered this city.'

Note that while the non-verbal constituent in (72b) is an Ezafe construction, it cannot simply involve X0-adjunction. The complement to the adjective is referential and must therefore be a DP. Thus the structure for (72b) is the following:

(73)

V0

| AP | DP |

Samiian (1983) gives a number of examples of the same phenomenon. In (74b) we have a noun modified by an adjective, in (75b) an adjective taking a complement and in (76b) a prepositional element with a complement.

(74) (a) ali hasan-râ davat kard-Ø
Ali Hasan+râ invitation did+3sgS
'Ali invited Hasan.' [Samiian 1983:259.117]
(b) ali hasan-râ davat-e rasmi kard-Ø
Ali Hasan+râ invitation+EZ formal did+3sgS
'Ali formally invited Hasan.' 27

[Samian 1983:259.118]

(75) (a) ali hasan-râ negarân kard-Ø
Ali Hasan+râ worried did+3sgS
'Ali made Hasan worried.'

[Samian 1983:260.121]

(b) ali hasan-râ negarân-e bachche-hâ kard-Ø
Ali Hasan+râ worried+EZ child+pl did+3sgS
'Ali made Hasan worried about the children.'

[Samian 1983:260.122]

(76) (a) ali ketâb-râ pas dâd-Ø
Ali book+râ back gave+3sgS
'Ali returned the book.'

[Samian 1983:260.125]

(b) ali ketâb-râ pas-e hasan dâd-Ø 28
Ali book+râ back+EZ Hasan gave+3sgS
'Ali gave the book back to Hasan.'

[Samian 1983:260.126]

---

27 As Samian points out in a footnote (fn. 17, p.268) it is also possible to use an adverbial here. The adverbial can occur either before or after the Hasan suggesting that it is adjoined to VP.

(i) (a) ali rasman hasan-râ davat kard-Ø
Ali formally Hasan+râ invitation did+3sgS
'Ali invited Hasan formally.'

(b) ali hasan-râ rasman davat kard-Ø
Ali Hasan+râ formally invitation did+3sgS
'Ali invited Hasan formally.'

28 This example is not so acceptable for my consultants, suggesting that there is some variation among speakers.
These constructions are truly mysterious if compounds always involve X^0-adjunction. However, if the non-verbal constituents of verbal compounds are phrases, then they can contain phrasal material.

The analysis of nouns developed in this thesis, predicts that they should contrast with adjectival, prepositional and participial elements in not being able to take a DP complement within a verbal compound. Recall from Chapter 3 that when a DP is adjoined to a noun the only way it can be licensed is as a possessor, receiving case from the null definite determiner. Consequently the noun phrase is referential and definite and cannot remain within the VP since the case assigned by the verb is incompatible with a definite interpretation for a noun phrase.\(^29\) This prediction is borne out. (77a) below illustrates a compound containing a noun. Unlike other compounds however, (see (72b), (75b)) the non-verbal element cannot occur with a complement linked with Ezafe. Note that the same noun can occur with a DP possessor outside of a verbal compound.

(77) (a)  otâq-râ jâru kard-am  
       room+râ  broom did+1sgS  
       'I swept the room.'

(b)  *jâru-e otâq kard-am  
      broom+EZ room  did+1sgS

(c)  jâru-e otâq-râ na-did-am 
      broom+EZ room  neg+saw+1sgS  
      'I didn't see the broom for the room.' (lit. 'the room's broom')

\(^{29}\)Another way to state this is that definite noun phrases cannot act as verbal modifiers. Since a DP possessor always entails that the noun phrase within which it occurs in definite, a nominal modifier containing a DP will not be able to occur.
5.6.1.2 A Unified Treatment of Reordering Operations

While it is assumed here that the Persian clause has an underlying order, there is fairly free order among constituents at the surface. I will refer to the operations by which these alternate orders are derived as reordering rules. Viewing the non-verbal elements of verbal compounds as phrasal has a favourable consequence for reordering rules in Persian. Specifically, we can maintain that all reordering operations target phrasal constituents and move them leftwards.

Let us first consider the rule of verb preposing discussed by Karimi (1989, 1994). She shows that the verb can appear at the beginning of the clause. An example is given below showing that the subject can intervene between the preposed verb and the direct object.

(78) xarid-am (man) un ketâb-o
bought+1sgS I that book+râ
'I BOUGHT that book.'

This rule can be seen as an instance of a more general operation that moves phrases leftwards. (79a) below illustrates a sentence with unmarked word order. However, as we see in (79b-d) various other orders are possible.

(79) (a) sepide ketâb-o barâ-ye simâ xarid-Ø
Sepide book+râ for+EZ Sima bought+3sgS
'Sepide bought the book for Sima.'

(b) sepide barâ-ye simâ ketâb-o xarid-Ø
Sepide for+EZ Sima book+râ bought+3sgS
'Sepide bought the book for Sima.'
The one difference between verb-preposing as shown in (78) and the reordering rule illustrated in (79) is that verb-preposing targets an X₀ rather than a phrase. This presents an obstacle to unifying the two phenomena.

Let us now consider what happens with compound verbs. We see in the examples below that the whole compound must prepose. That is, compounds seem to be inseparable for the purposes of verb preposing.

(80)  (a)  járu kard-am (man) un otâq-o
      broom did+1sgS I that room+râ
      'I SWEPT that room.'

      (b)  *kard-am (man) un otâq-o járu
          did+1sgS I that room+râ broom

This inseparability is baffling if verb-preposing targets V₀s, and the non-verbal elements in a compound are generated as phrasal complements. This was one of the reasons that the X₀

adjunction structure was proposed in Ghomeshi and Massam (1994). If, however, definite direct objects and PPs are adjoined to VP, and if the subject is base-generated outside VP as was suggested in section 5.5 then, there is nothing in the VP except the verb and compounding elements (if there are any). Therefore, we can say that the movement rule
illustrated above is actually VP preposing. In other words, all preposing operations in Persian can now be said to uniformly target phrases.

This also explains another fact noted by Karimi (1989, 1994). She proposes a rule of XP-postposing in Persian that can only target noun phrases marked with -râ. This was mentioned in section 5.3.2 and an example is repeated below:

(81) (a) sepide emruz hayât-o tamiz kard-Ø
   Sepide today court-yard+râ clean did+3sgS
   'Sepide cleaned the court-yard today.'

   (b) sepide emruz hayât tamiz kard-Ø
       Sepide today court-yard clean did+3sgS
       'Sepide cleaned court-yards today.'

   (c) sepide emruz tamiz kard-Ø hayât-o
       Sepide today clean did+3sgS court-yard+râ
       'Sepide CLEANED the court-yard today.'

   (d) *sepide emruz tamiz kard-Ø hayât
       Sepide today clean did+3sgScourt-yard

   [Karimi 1989:213.10]

Note that instead of assuming that the DP+râ postposes around the verb, one could treat these cases as further instances of XP-preposing, though not to sentence-initial position. The ungrammaticality of (81d) is then explained, since the noun phrase hayât is in VP and the VP cannot move around the adverb without also taking this element. Another point in favour of this second analysis is that it allows all reordering phenomena in Persian to be analyzed as leftward movement (cf. Kayne (1994)).

Finally, consider the following example in which the non-verbal element of a compound has been moved from its base-generated position in (82b).
Further examples are given below showing a non-verbal element of a compound appearing before the VP-adjoined adverb aslan 'at all'. In (85) an example taken from Vahedi (1995) shows an adjectival complement appearing sentence-initially.

(83)  
\begin{verbatim}
gush  aslan ne-mi-kon-e  
\end{verbatim}
\begin{verbatim}
  ear  at all neg+cont+do+3sgS  
\end{verbatim}
'S/he doesn't listen at all.'

(84)  
\begin{verbatim}
jāru  aslan ne-mi-kon-e  
\end{verbatim}
\begin{verbatim}
broom  at all neg+cont+do+3sgS  
\end{verbatim}
'S/he doesn't sweep at all.'

(85)  
\begin{verbatim}
garm-esh  tu madrase mi-kon-am  
\end{verbatim}
\begin{verbatim}
  warm+3sg in school  cont+do+1sgS  
\end{verbatim}
'I will heat it up at school.'

The movement of the non-verbal element of a compound is rare. However, it is possible. If reordering rules in Persian always involves phrases and always move these phrases to the left, the examples above serve as further evidence that non-verbal elements in compounds are phrases.
We now turn to consider the structure of sentences containing compound verbs. Specifically, we must determine the structure of VP in sentences such as the following:

(86) (a) miz-o tamiz kard-am
    table+râ clean did+1sgS
    'I cleaned the table.'

    (b) bijan-o bidâr kard-am
    Bijan+râawake did+1sgS
    'I woke Bijan up.'

In such examples, it is clear that the non-verbal element forms a compound with the verb, both semantically, and in terms of stress placement (the entire compound receives only one stress). However, as mentioned in the introduction, there is also a relation between the direct object and the non-verbal element. I suggest that this relation is one of predication. That is, the relation between the direct object and the adjective in the above sentences is exactly the same as the relation in the copular sentences below.

(87) (a) miz tamiz-e
    table clean+3sgS
    'The table is clean.'

    (b) bijan bidâr-e
    Bijan awake+3sgS
    'Bijan is awake.'

A possible analysis for the types of constructions in (86), given that there is a relationship of predication between the direct object and the non-verbal element, is to assume that they
form a clause-like constituent, namely a small clause (cf. Chomsky (1981), Stowell (1981)).
Under such a view the structure for the VP in a sentence such as the one in (86a) (prior to
movement of the direct object to adjoin to VP) would be the following:

(88) (a) miz-o tamiz kard-am
       table+râ clean did+1sgS
       'I cleaned the table.'

(b)

\[
\begin{array}{c}
\text{VP} \\
\text{AP} \\
\text{V} \\
\text{DP} \\
\text{AP} \\
\text{kard-am} \\
\text{NP} \\
\text{D} \\
\text{A} \\
\text{Ødef} \\
\text{tamiz} \\
\text{miz}
\end{array}
\]

In the structure in (88b), the direct object receives a theta-role from the AP (in this case tamiz
'clean') but also receives a theta-role from the verb as its direct object. That is, the direct
object is an argument of both the non-verbal element and the main verb.

There is a problem with the view that transitive compounds involve small clauses, having
to do with the pronominal enclitics. Recall that Persian allows pro to appear as a direct
object, licensed by a pronominal enclitic on the verb. As noted in the introduction to this
section, the pronominal enclitics may also appear with verbal compounds, however, they
attach to the non-verbal element. Generally pronominal enclitics license pro arguments by
appearing on the governing head (cf. Chapter 4, section 4.2.2.3).

It is interesting to note that in an example such as the following, where the non-verbal
element is a noun, the pronominal enclitic is not interpreted as a possessor.
The question is, in a small clause configuration, what is the governing head? It might be the head of the predicate in the small clause. This is shown below.

(89) \[
\begin{array}{c}
\text{AP} \\
\text{DP} \quad \text{AP} \\
\text{D} \quad \text{A} \\
\end{array}
\]

\[\text{pro}_i \quad \text{adjective}+\text{Fi}\]

This hypothesis makes an incorrect prediction about predicates that contain a head with its complement, however. Consider the examples below, where (90) is taken from Samiian (1983:269.fn.20) and (91) shows three possible ways to express the same proposition:31

(i) (a) jâru-sh kard-am
    broom+3sg did+1sgS
    'I swept it.'

If the pronominal enclitic were to be construed as a possessor the presence of the null definite determiner would be required. Since definite DPs cannot receive case directly from the verb, this would mean the DP containing the possessor would have to move to adjoin to VP to get case. This is shown in (ii) below. The sentence in (iia) is ungrammatical because if the noun and the enclitic form a DP, there is no case marking. Further, if the noun is construed as a modifier (e.g. broom buying), it also absorbs the internal argument position of the verb (see section 5.6.3), thus leaving no internal argument position for the enclitic to absorb. In contrast, the sentence in (iib) where the the DP is construed as the direct object is well-formed.

(ii) (a) *jâru-sh xarid-am
    broom+3sg bought+1sgS

(b) jâru-sh-o xarid-am
    broom+3sg+râ bought+1sgS
    'I bought his/her broom.'

31As Samiian points there is sometimes a preference for the pronominal enclitic to be interpreted as a possessor of the complement rather than as a direct object of the verbal compound. She gives the following example:

(i) ??ali negarân-e bacheh-hâ-sh kard-Ø
    ali worried+EZ child+pl+3sg did+3sgS
    'Ali made him worried about the children.'
There is a tendency to construe the predicate as 'worried about his children' in which case the sentence is ungrammatical because it lacks a direct object. The possessor interpretation for the enclitic is reinforced if the constituent to which it attaches is definite. Thus compare the following example to (91b) in the text.

(ii)  ??ali âmâde-ye un  safar-am kard-Ø  
      ali ready+EZ that trip+1sg did+3sgS  
      'Ali made me ready for that trip./Ali made ?? ready for that trip of mine.'

If the complement is indefinite, there is no determiner to licence the possessor, and the object interpretation is more readily available. This can be seen as a parsing problem and not one of syntax proper.
These examples illustrate verbal compounds in which the non-verbal element takes a complement. In these cases the enclitic is attached to the end of the adjective+complement constituent. Thus we appear to have arrived at a paradox. If enclitics attach to governing heads then they should not be able to attach to the end of the whole adjective phrase. If, on the other hand, the adjective phrase is functioning as a head (cf. Carnie (1995) for a discussion of phrases that act like heads in Irish), then we need to explain why an adjective phrase cannot appear within the Ezafe domain. Recall from Chapter 2 that the reason given for why adjectives cannot take complements within the Ezafe domain was that this domain includes only heads, not whole phrases.

I propose that the small clause constructions in (90)-(91) have a null head to which the enclitic attaches. I follow Bowers (1993) in assuming that main clause and small clause predication both involve a functional projection that he calls Predication Phrase (PrP). Bowers (1993:595) gives the following basic properties for this projection:

(92) (a) the canonical D-Structure position for external arguments is [SPEC, Pr]
     (b) Pr0 F-selects the maximal projection YP of any lexical category Y
     (c) either PrP is F-selected by I0, or it can be subcategorized as a complement by V
     (d) the semantic function of Pr is predication

Given that PrP can occur as the complement of a verb, and given that that complement of Pr0 can be any lexical category, this proposal yields the following underlying structure for a sentence such as (88a), repeated below. Note that the definite DP in [SPEC, PrP] will have to move to adjoin to VP in order to be case-marked.

(88) (a) miz-o tamiz kard-am
     table+râ clean did+1sgS
     'I cleaned the table.'
This structure provides a solution to the problem posed by the pronominal enclitics in complex verb constructions. If the subject of PrP is pro, a pronominal enclitic will be required on the governing head, namely on Pr\(^0\). I assume that the enclitic comes to be attached to the edge of the complement to the Pr-head by cliticization rules that apply at PF. The relevant structure is given below:\(^{32}\)

\[^{32}\text{Note that with transitive compound verbs, as with simple transitive verbs (see section 5.3.4), a clitic and a direct object marked with -rā can cooccur.}\]

(i) (a) chamedun-opor-esh kard-am
suitcase+rā full+3sg did+1sgS
'As for the suitcase, I filled it.'

(b) bachcha-rohamum-esh kard-am
child+rā bath+3sg did+1sgS
'As for the child, I bathed her/him.'

This follows directly from the analysis developed. The enclitic is licensed by governing and being coindexed with pro. The DP+rā is licensed by c-commanding and being coindexed with pro.
5.6.3 Theta Assignment

Consider sentence (88a), repeated below:

(88) (a) miz-o tamiz kard-am
    table+râ clean did+1sgS
    'I cleaned the table.'

In such a sentence the subject receives the external theta-role of the verb. (Recall that the subject is base-generated in the specifier of a higher PrP that takes the VP as its complement. The verb moves to the head (Pr⁰) of this projection and assigns its external theta-role to the specifier.) The verb kardan 'to do/make' is a two-place predicate. This is shown in (95). In (95b) we see that a sentence involving this verb is ungrammatical if it lacks a direct object. For comparison an intransitive (i.e. one-place) predicate is given in (96).

(95) (a) un kâr-o kard-am
    that work+râ did+1sgS
    'I did that work.'

    (b) *kard-am
        did+1sgS
        *'I did.'
In a sentence such as (88a) then, the PrP complement to the verb serves as the internal argument. The direct object, *miz* 'table' receives a theta-role from the predicate *tamiz* 'clean' in [SPEC, PrP]. This is shown in (97) below.

\[
\text{(97) internal theta-role of verb}
\]

\[
\text{theta-role of predicate}
\]

In Ghomeshi and Massam (1994) it is argued that all compound verbs have the same syntactic structure, regardless of the relationship between the non-verbal element and the verb. However, under the view developed here this is not the case. That is, nouns that absorb a theta-role of the verb, even if they form a compound with the verb, are syntactically different from those that are purely predicative. Let us look at this more closely.

Consider the verb *kardan* 'to make/do', which is a two-place predicate, as shown above. This verb can occur with a bare noun that absorbs its internal argument position. Together they form a compound. This is shown in (98a) below. In (98b) we see that the resulting compound cannot occur with a direct object. However, the nominal complement to the verb can be 'pulled out' so as to function as the direct object. This is shown in (98c), where we see that the noun is now marked with -râ, i.e. that it is contained within a DP.
The difference between the compound vs. non-compound version of these sentences lies in the nature of the complement to the verb. If the complement is a DP it is construed as referential. However, if it is simply a bare noun it is construed as non-referential and acts like a modifier of the verb. However, in both cases the verb's internal theta-role is assigned to the complement position. This is illustrated below.

As for the fact that verbal compounds bear only one stress, one of the reasons the X₀-adjunction analysis was proposed in Ghomeshi and Massam (1994), I propose that this follows from the phrasal phonology of Persian clauses. Lazard (1992) writes the following about phrasal stress:
In the sentence, personal verbal forms are not always stressed: their being stressed or not depends on how far they are syntactically and semantically autonomous...If the verbal form is the only element in its syntactic phrase, it is stressed...If the verb (as a rule placed at the end of the clause) is preceded by nominal elements which are more or less closely tied to it, the last nominal element carries a strong stress and the verb is usually unstressed...

[Lazard 1992:43]

While the phrasal phonology of Persian has yet to be studied in depth, I suggest that it is possible that all constituents within the VP can be considered an intonational unit bearing only one stress.

Returning to the occurrence of bare nouns that form compounds with a verb, let us consider a second case, where the noun does not absorb the internal argument of the verb. In the example given below, the same verb kardan 'to do/make' is used, but here note that a direct object is permitted.

(100) (a) járu kard-am
         broom did+1sgS
         'I swept.'

         (b) otâq-o járu kard-am
             room+râ broom did+1sgS
             'I swept the room.'

In this case the noun is now acting as a true predicate and will appear in the following configuration.
The generalization that can be made then is that noun phrases must either receive a theta-role, in which case they are arguments, or they must assign a theta-role in which case they are predicates.

Finally note that in structures such as (97) and (101), the DP in [SPEC, PrP] must receive case. If it is indefinite it can remain in its base-generated position and it will be case marked by the verb by Exceptional Case Marking. However, recall that the case assigned by the verb is incompatible with a definite/presupposed reading for the direct object. Thus, just as in simple transitive structures, direct objects in sentences involving complex predicates must move to adjoin to VP where they are case-marked by -rā.

5.8 The Impersonal Construction

The analysis developed in the preceding section for sentences involving transitive complex predicates can be easily extended to explain another construction in Persian, often referred to as the 'impersonal construction'. Similar constructions (referred to as Dative or Experiencer Subject constructions) are found in many South Asian languages (see Verma and Mohanan (1990)). Generally in such constructions the 'subject', which is an experiencer rather than an agent, appears in genitive or dative case. In Persian the 'subject' is indicated by a pronominal enclitic while the verb does not bear agreement. Often these contructions
have a regular (or direct) counterpart. Examples are shown below where the regular version is given in (a) and the experiencer subject version is given in (b).33

(102) (a) xâbid-am
slept+1sgS
'I slept.'

(b) xâb-am âmad
sleep+1sg came
'I was/felt sleepy.'

(103) (a) un-o dust dasht-Ø
that+râ friend had+3sgS
'S/he liked that.'

(b) az ân xosh-esh âmad
from that good+3sg came
'S/he liked that.'

33Note that in (104) the difference in meaning between the impersonal construction and the copular sentence is akin to the difference between constructions involving 'to have' vs. 'to be' in Romance languages. Thus the French sentences below exhibit the same contrast in meaning.

(x) (a) il est froid
he is cold
'He is cold.' (as a person)

(b) il a froid
he has cold
'He is (feels) warm.'

Furthermore, note the experiencer subject versions in Romance involve the verb 'to have', while the experiencer subject version in Persion involves the pronominal enclitics, which can be used to indicate possession. Unfortunately this remains simply an intriguing observation for now.
(104) (a) jiân sard-e  
Jian cold+3sgS  
'Jian is cold' (as a person, or to touch)

(b) jiân sard-esh-e  
Jian cold+3sg+3sgS  
'Jian is (feels) cold'

There are similarities between these verbs and the transitive compound verbs. For example, these constructions involve a non-verbal element and a light verb, although in these cases the verb is intransitive. Like the transitive compounds, these 'compounds' receive only one stress. Further examples are given below:

(105)  
bad ámadan (az)  
bad=to come(=from)  
'to dislike'

xâb ámadan  
sleep=to come  
'to get sleepy'

xâb bordan/raftan  
sleep=to take/go  
'to fall asleep'

sard shodan/budan  
cold=to become/to be  
'to become/be cold'

garm shodan/budan  
warm=to become/to be  
'to become/be warm'

xaste shodan/budan  
tired=to become/to be  
'to become/be tired'

gorosne shodan/budan  
hungry=to become/to be  
'to become/be hungry'

dard gereftan  
ache=to get  
'to ache'

Another similarity between these constructions and transitive compounds is that nothing can intervene between the non-verbal element and the verb. This is shown below with a prepositional phrase in (106) and a VP-adjoined adverb in (107).

(106) (a) az un ketâb xosh-am âmad  
from that book  good+1sg came  
'I liked that book.'
(b) *xosh-am az un ketâb âmad
good+1sg from that book came

(107) (a) diruz sard-esh bud
yesterday cold+3sg was
'Yesterday, s/he was cold.'

(b) *sard-esh diruz bud
cold+3sg yesterday was

This property distinguishes an impersonal construction from an intransitive sentence in which the subject contains a pro possessor identified by a pronominal enclitic. In all other respects these two constructions look like they should be syntactically identical.\(^{34}\)

(108) (a) barâdar-am âmad-Ø
brother+1sg came+3sgS
'My brother came.'

(b) barâdar-am diruz âmad-Ø
brother+1sg yesterday came+3sgS
'My brother came yesterday.'

(109) (a) xâb-am âmad
sleep+1sg came
'I felt sleepy.'

---

\(^{34}\)Another difference is that the non-verbal elements occurring in impersonal constructions are not nouns. Most must undergo a derivational process to be used as nouns.

E.g. gorosne 'hungry' gorosnegi 'hunger'
sard  'cold'  sardi  'coldness'

This suggests that the pronominal enclitic is not a possessor in these cases.
The lack of agreement on the verb, the stress property, and the lack of separability of the impersonal construction suggests that it only involves a VP projection. In the previous section it was suggested that any material that stays within the VP forms one intonational unit with the verb. We also saw that no adjuncts could appear between constituents within VP. Finally, recall that it was suggested in Chapter 4, section 4.1, that agreement assigns nominative case, and is checked, in a higher projection than VP.

I propose the following structure for impersonal constructions:

Crucially this VP is not selected by a PrP. Thus there is no position in which a 'true' subject can appear (recall that pro does not need case). This means that the verb stays within VP and does not move to check agreement features.

A note should be made here about subject agreement. While in the past tense, third person agreement is null and thus a past stem without agreement looks exactly the same as one with third person agreement; in the present tense third person agreement is not null and, in fact, does appear on the verb in these impersonal constructions.

Consider the past and present paradigms for two verbs that participate in the impersonal construction (colloquial pronounciations are given in parentheses).
Unlike the past stem which is used with an inflected form of the verb xâstan 'to want' to indicate the future, the present stem (without inflection) never appears alone. Further, as can be seen above, in the colloquial pronounciati ons, often the present stem is simply a consonant. Given that verbs are inserted fully inflected into the syntax (see Chapter 1, section 1.2.2), I propose that the third singular version of the present stem is the default form used when no agreement is used. That is, the fact that it appears in third singular form is simply a reflection of a morphological requirement that the present stem not appear bare. Thus, the agreement features in the present third singular do not have to be checked in order for the form to be licensed.

The parallel between impersonal constructions and transitive complex predicates is that in both cases the DP in [SPEC, PrP] receive a theta-role from the complement to Pr⁰. The Predication Phrase itself receives a theta-role from the intransitive verb. This is shown in (113).
The difference between impersonal constructions and transitive complex predicates is that in impersonal constructions there is no Exceptional Case Marking from the verb to [SPEC, PrP]. Note that the verbs involved in these constructions are intransitive and thus have no case to assign. Nominative case, which is assigned by the agreement affixes, is also unavailable for the argument in [SPEC, PrP] since there is a lack of agreement on the verb. Thus, the only noun phrase that can appear in [SPEC, PrP] is one that does not require case, namely pro, which in turn must be identified by the pronominal enclitics.

Finally, let us consider the structure of sentences involving an impersonal construction and an apparent subject. Some examples are given below.

(114) (a) \textbf{man az} \textbf{ân} \textbf{xosh-am} âmad
I \hspace{1em} \text{from that good+1sg came}
I liked that.'

(b) \textbf{shomâ xâb-etun} âmad
you \hspace{1em} \text{sleep+2pl came}
'You got sleepy.'

(c) \textbf{barâdar-e rezâ} az \textbf{tehrân xosh-esh} ne-mi-âyad
\hspace{1em} brother+EZ Reza from Tehran good+3sg neg+cont+comes
'reza's brother doesn't like Tehran.'
The sentences in (114) show that when the subject is overtly expressed it must appear at the beginning of the sentence and be construed as coreferential with the enclitic. The verb does not agree with an overt subject and still appears in third person singular form. The resulting construction resembles clause-level topicalization (see section 5.4) and in fact some grammars assume they are the same thing (see Thackston (1983) who refers to this as the Resumptive Construction).

Given that subjects in the specifier of a PrP above VP cooccur with agreement on the verb, I suggest that overt subjects that appear in the impersonal construction are actually generated in the specifier of Topic Phrase. Thus, such sentences take the form of clause-level topicalization, however, these particular constructions lack a PrP between the VP and the TopP. The relevant structure is shown below.

(115)

\[
\text{TopP} \\
/ \quad \text{\textit{overt subject}_i} \quad \text{Top'} \\
\text{VP} \quad \text{Top} \\
\quad \text{PrP} \\
\quad \text{DP} \quad \text{Pr'} \\
\quad \text{pro}_i \quad \text{XP} \quad \text{Pr} \\
\quad \text{Ø+F}_i
\]

One difference between the impersonal construction and the clause-level topicalization structures discussed in section 5.4 is that the possibility of having a full pronoun or the whole constituent in place of the enclitic is not available. This is shown below where in (116) an example of a topicalized structure is given along with its two other potential forms. In (117) we see that only the (a) example is grammatical.
(116) (a) ân zan-e az xâne-ash radd shod-am
that woman+def from house+3sg pass became+1sgS
'That woman, I passed by her house'

(b) ân zan-e az xâne-ye u radd shod-am
that woman+def from house+EZ her pass became+1sgS
'That woman, I passed by her house'

(c) az xâne-ye ân zan-e radd shod-am
from house+EZ that woman+def pass became+1sgS
'I passed by the house of that woman.'

(117) (a) barâdar-e man az tehrân xosh-esh ne-mi-âyad
brother+EZ me from Tehran good+3sg neg+cont+comes
'My brother doesn't like Tehran.'

(b) *barâdar-e man az tehrân xosh-e u ne-mi-âyad
brother+EZ me from Tehran good+EZ him neg+cont+comes

(c) *az tehrân xosh-e barâdar-e man ne-mi-âyad
from Tehran good+EZ brother+EZ me neg+cont+come

This follows from the fact that in (116) the topicalized constituent is coindexed with a possessor. Pronouns and full noun phrases can also appear as possessors. However, in (117) the topicalized constituent is coindexed with a pro that is not in a case position. Therefore, neither a pronoun nor a full noun phrase can appear in this position.

To summarize, the analysis of transitive complex predicates developed in section 5.6 has been extended here to explain the impersonal construction in Modern Persian. We have seen that if the impersonal construction involves only a VP, we can explain the compound-like nature of the construction. We have also seen that this construction does not involve subject
agreement. Given that subject agreement is responsible for nominative case assignment, and
given that pro does not need case, we have an explanation for why only pro can appear in
these constructions. Finally, we have seen that overt 'subjects' are generated in [SPEC, TopP]\(^{35}\) where they can receive case and therefore do not require the presence of agreement
on the verb.

One of the most interesting properties of impersonal constructions, namely that the
'subject' identified by the pronominal enclitic is construed as an experiencer, remains
unexplained here. However, future research on VP-internal vs. VP-external Predication
Phrases, as well as similar constructions in other languages, may provide an answer.

5.9 Conclusion

In this chapter we have looked at the properties of the morpheme -râ and in doing so
have also considered topics and subjects. We have seen that the morpheme -râ case-marks
presupposed noun phrases that are adjoined to VP. These noun phrases are of two types: (a)
direct objects, which being presupposed, cannot be directly case-marked by the verb and

\(^{35}\)Note that the coindexing between these arguments and the pro in [SPEC, PrP] satisfies the
Thematic Licensing Condition (see section 5.4). This raises a question of why a DP could
not adjoin to VP, receive case-marking from -râ, and satisfy the Thematic Licensing
Condition by being coindexed with pro. That is, nothing in the analysis developed thus far
rules out a sentence such as the following:

(i) \[\text{*} \text{jîn-o sard-esh } \text{bud} \]
\[\text{Jian+râ cold+3sg was} \]
\[\text{'Jian was (felt) cold.'} \]

Note that a DP+râ can appear if it is not coindexed with the pro in [SPEC, PrP].

(ii) \[\text{kêtâb-o az-esh } \text{xosh-am nay-âmäd} \]
\[\text{book+râ from+3sg good+1sg neg+came} \]
\[\text{'As for the book, I didn't like it.'} \]

I leave this question for future research.
must move to receive case; and (b) base-generated noun phrases which are case-marked by -rā but must be thematically licensed by being coindexed with a noun phrase that receives a theta-role. The notion that the case assigned by the verb cannot license presupposed direct objects has been shown to have precedent in the literature.

The Thematic Licensing Condition on noun phrases that appear in VP-adjoined position (where they do not receive a theta-role), accounts for the similarity between direct objects and what are referred to here as VP-level topics. Further, this condition also plays a role in licensing clause-level topics.

In light of the fact that noun phrases marked with -rā are adjoined to VP, and that -rā is an available case marker for any such adjunct it has been shown that subjects cannot originate in [SPEC, VP]. Under such a view, subjects would be able to adjoin to VP to receive case and would appear marked with -rā. This is an incorrect prediction. It has been proposed instead that subjects are base-generated in the specifier of Predication Phrase.

In the final two sections of the chapter we have examined compound verbs or complex predicates. It has been shown that these constructions involve a second Predication Phrase which is selected by the verb. In the cases where the verb is transitive, this hypothesis has been shown to account for the semantic relationship between the direct object and the non-verbal element of the compound. In cases where the verb is intransitive the presence of a Predication Phrase has been shown to account for the properties of the impersonal construction in Persian. Finally, we have seen that presence of this phrase allows for a consistent description of where the pronominal enclitics appear.
Chapter 6

Summary of Thesis

This thesis is based on some crucial (and not uncontroversial) assumptions about morphology and syntax. These assumptions are (a) that monosemy, i.e. uniting the various occurrences of morphemes that are identical in their phonological properties, is desirable in grammar; (b) that only inflectional affixes and not derivational affixes can correspond to syntactic projections; (c) that projection is driven by the features borne by lexical and functional categories; and (d) that base-generated X0-adjunction is possible in the syntax. These assumptions, taken together, are shown to lead to further theoretical claims in order to explain Persian clause structure.

Starting with the distinction between inflectional and derivational affixes, the thesis is organized around the inflectional affixes of Persian and the range of constructions that they participate in. Chapter 2 takes as its starting point the Ezafe vowel -e. This vowel is present in noun phrases, adjective phrases, and most prepositional phrases linking heads to following constituents. To explain the properties of the Ezafe domain it is argued that nouns do not project structure in Persian. That is, they appear only as N0s in the syntax unless they are selected by a projecting head. Even if so selected (where they will be dominated by an NP node) they can never take phrasal complements or modifiers which follows from their non-projecting property. In contrast to nouns it is claimed that adjectives and nominal prepositions in Persian are optionally projecting categories, while true prepositions and verbs obligatorily project.

Chapter 3 deals with determiners in Persian which are expressed by inflectional affixes. It is argued that in addition to the indefinite determiner -i, there is a null definite determiner. The presence of this null determiner captures the difference between generic and definite noun phrases which look identical in subject position. This determiner is also responsible for assigning case to a possessor. The fact that the indefinite determiner -i lacks this case-
assigning property explains why it is in complementary distribution with possessors. It is
argued that possessors are base-generated in \([\text{SPEC, DP}]\) in Persian which is consistent with
the fact that Persian NPs cannot contain phrasal constituents.

In light of the structure developed in Chapters 2 and 3 for the extended noun phrases in
Persian an Ezafe Insertion Rule is given that applies at PF. It is argued that this rule should
be seen as part of the phrasal phonology, similar to rules deriving intonational phrasing. The
presence of this vowel serves to identify constituent structure.

In Chapter 4 the subject agreement suffixes and the pronominal enclitics are discussed. It
is argued that the subject agreement affixes are responsible for nominative case assignment
in the Persian clause. It is also proposed that they subcategorize verbs only. In contrast the
pronominal enclitics can attach to any head, including verbs. The sole constraint on their
appearance is that they cannot attach directly to a verb stem. This is accounted for by a PF
filter. The pronominal enclitics also differ from the subject agreement suffixes in that they
absorb case if attached to a case-assigning head. Thus they can only cooccur with a noun
phrase that does not require case. Since \(\text{pro}\) does not require case, and since its presence can
be licensed by the pronominal enclitics, the analysis developed allows for \(\text{pro}\) to occur as a
possessor, or the object of a preposition or a verb, in addition to occurring as a subject. The
analysis developed for these two sets of morphemes accounts for two intriguing facts in
Persian. First, the fact that when both subject agreement and a pronominal enclitic occur on
a verb the enclitic must follow the agreement affix even though it refers to the direct object.
Second, the fact that the pronominal enclitics appear to be sensitive to definiteness while the
subject agreement suffixes are not.

The second half of Chapter 4 deals with the binding of pronominal elements in Persian.
It is shown that the fact that the pronominal enclitics cooccur with \(\text{pro}\) explains the
differences in binding between constructions involving the enclitics and constructions
involving independent pronouns. In looking at binding in Persian a revision is made to
classic Binding Theory to include the Condition B of Reinhart and Reuland (1993), renamed
the Reflexivity Condition here. The discussion of binding in Persian includes also an analysis of the anaphor *xod* 'self'. The differences between this form when it occurs alone and the inflected version, where *xod* bears a pronominal enclitic, are accounted for. In keeping with the assumption of monosemy a unified analysis of *xod* in both its reflexive and emphatic uses is developed.

In Chapter 5, we turn to the last inflectional morpheme considered in the thesis, namely the marker -râ. It is argued that -râ is a case-marker for presupposed DPs adjoined to VP. A VP-adjoined DP can either have moved from the base-generated direct object position, on the assumption that the case assigned by the verb is incompatible with a presupposed reading for the direct object; or be base-generated as a VP-adjunct. It is shown that such DPs are subject to the Thematic Licensing Condition, i.e. that they must be coindexed with and c-command a theta-marked noun phrase. (Where the DP has moved from direct object position this noun phrase will contain a trace.) The Thematic Licensing Condition is generalized to account for clause-level topicalization in Persian. It is proposed that DPs marked with -râ are VP-level topics. The structure of a Persian sentence involving a simple transitive verb is thus the following:

(1)

```
(1)          PrP
            /  \
           DP    Pr'
          /  \    /
         subject VP Pr
               /  \    /
             KP  VP K
            / \   / \  
           DP K DP V
          /   \   /   
presupposed indefinite direct object -râ direct object/verb
```

In the last half of Chapter 5, complex predicates in Persian are discussed. Complex predicates generally involve a light verb and a non-verbal element. First transitive predicates
are dealt with. Evidence is presented that the non-verbal element in transitive 'compounds' must be phrasal. To account for the phrasality of this element, the thematic relationship between it and the direct object, and to account for the fact that pronominal enclitics can occur on this element a structure is proposed in which the verb takes a Predication phrase as its complement. This structure is shown below.

Finally, the impersonal construction is discussed. In this construction the subject is expressed by a pronominal enclitic attached to a non-verbal element. It is construed as an experiencer. The verb is intransitive and bears no agreement. It is proposed that these constructions are the intransitive counterpart to the structure shown in (2) above. The lack of agreement on the verb means that, structurally, these constructions lack the higher Predication Phrase. It also means that there is no case available for the 'subject' and thus it always contains a pro. The structure for an impersonal construction is given in (3).
In each of the syntactic constructions considered in this thesis the inflectional and syntactic category features play a role in determining the phrase structure. In this way much of the uniqueness of Persian syntax ultimately derives from the morphological properties of the items involved.
References


Chomsky (1972) Studies on Semantics in Generative Grammar, Mouton.


Lumsden, Matthew (1810) *A grammar of the Persian language; comprising a portion of the elements of Arabic inflexion, together with some observations on the structure of either language considered with reference to the principles of general grammar*. 2 vols. Calcutta.


*Linguistic Inquiry* 20, pp. 365-424.

*Discourse Description, Diverse linguistic analyses of a fund-raising text*, William C.  
Mann and Sandra A. Thompson (eds), John Benjamins Publishing Company,  
Amsterdam, The Netherlands.

England.

Ph.D. dissertation, MIT.


Riemsdijk, Henk C. van (1983) "The Case of German adjectives," in *Linguistic Categories:  
Auxiliaries and Related Puzzles I*, F. Heny and B. Richards (eds), Reidel, Dordrecht.

Ritter, Elizabeth (1991) "Two Functional Categories in Modern Hebrew Noun Phrases," in  
*Syntax and Semantics 25, Perspectives on Phrase Structure: Heads and Licensing*,  


Language and Linguistic Theory* 13, pp. 405-443.

Roberge, Yves (1990) *The Syntactic Recoverability of Null Arguments*, McGill-Queen's  
University Press, Montreal.


Sadler, Louisa and Douglas J. Arnold (1994) "Prenominal adjectives and the phrasal/lexical  


Thackston, Jr., Wheeler M. (1983) *An Introduction to Persian*, Department of Near Eastern Languages and Civilizations, Harvard University, Cambridge, MA.


