

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

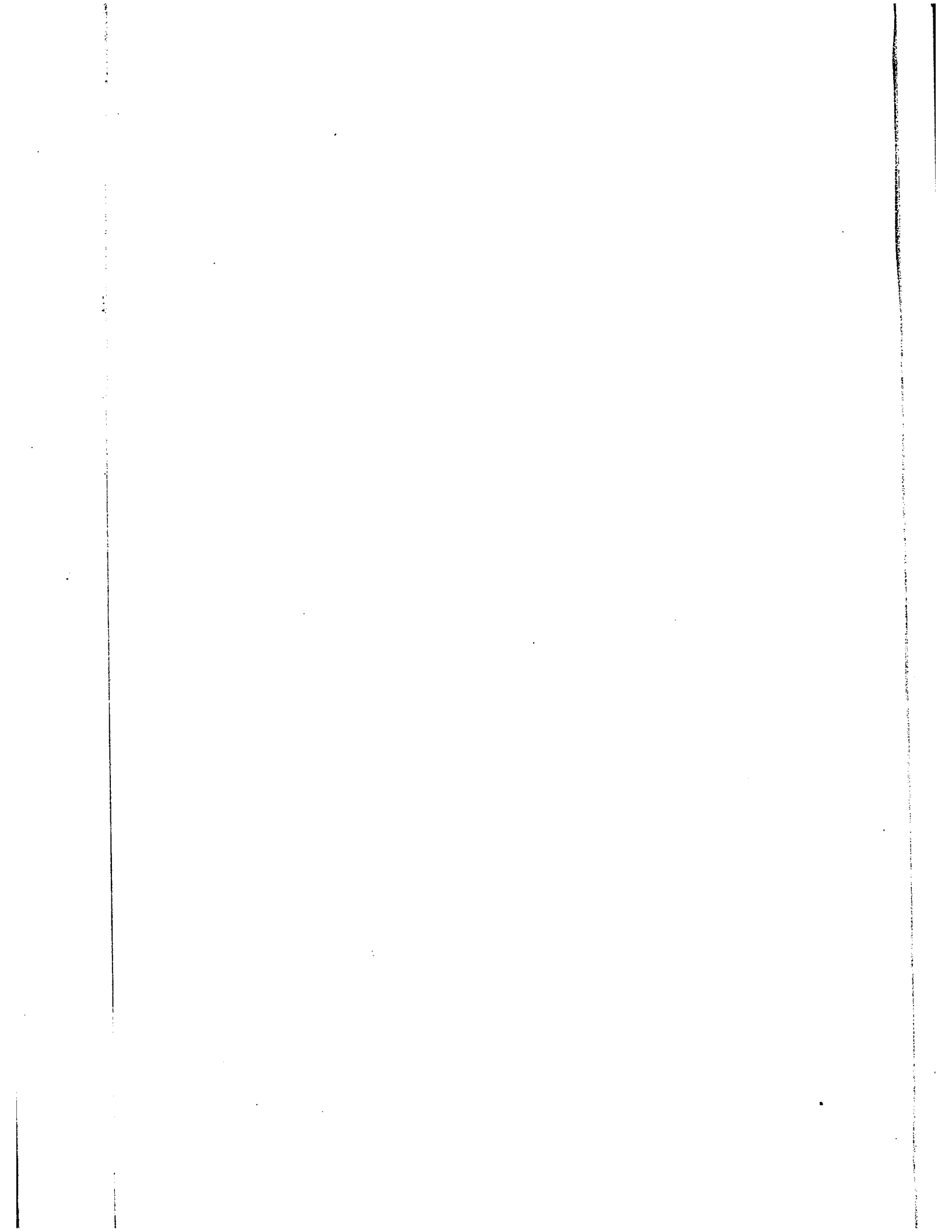
The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI[®]



UNIVERSITÉ D'OTTAWA



UNIVERSITY OF OTTAWA

ÉCOLE DES ÉTUDES SUPÉRIEURES
ET DE LA RECHERCHE

SCHOOL OF GRADUATE STUDIES
AND RESEARCH

VAHEDI LANGRUDI, Mohammad Mehdi
AUTEUR DE LA THÈSE-AUTHOR OF THESIS

M.A. (Linguistics)
GRADE-DEGREE

LINGUISTICS
FACULTÉ, ÉCOLE, DÉPARTEMENT-FACULTY, SCHOOL, DEPARTMENT

TITRE DE LA THÈSE-TITLE OF THE THESIS

CLAUSAL AND VERBAL STRUCTURE IN FARSI

P. Hirschbühler
DIRECTEUR DE LA THÈSE-THESIS SUPERVISOR

EXAMINATEURS DE LA THÈSE-THESIS EXAMINERS

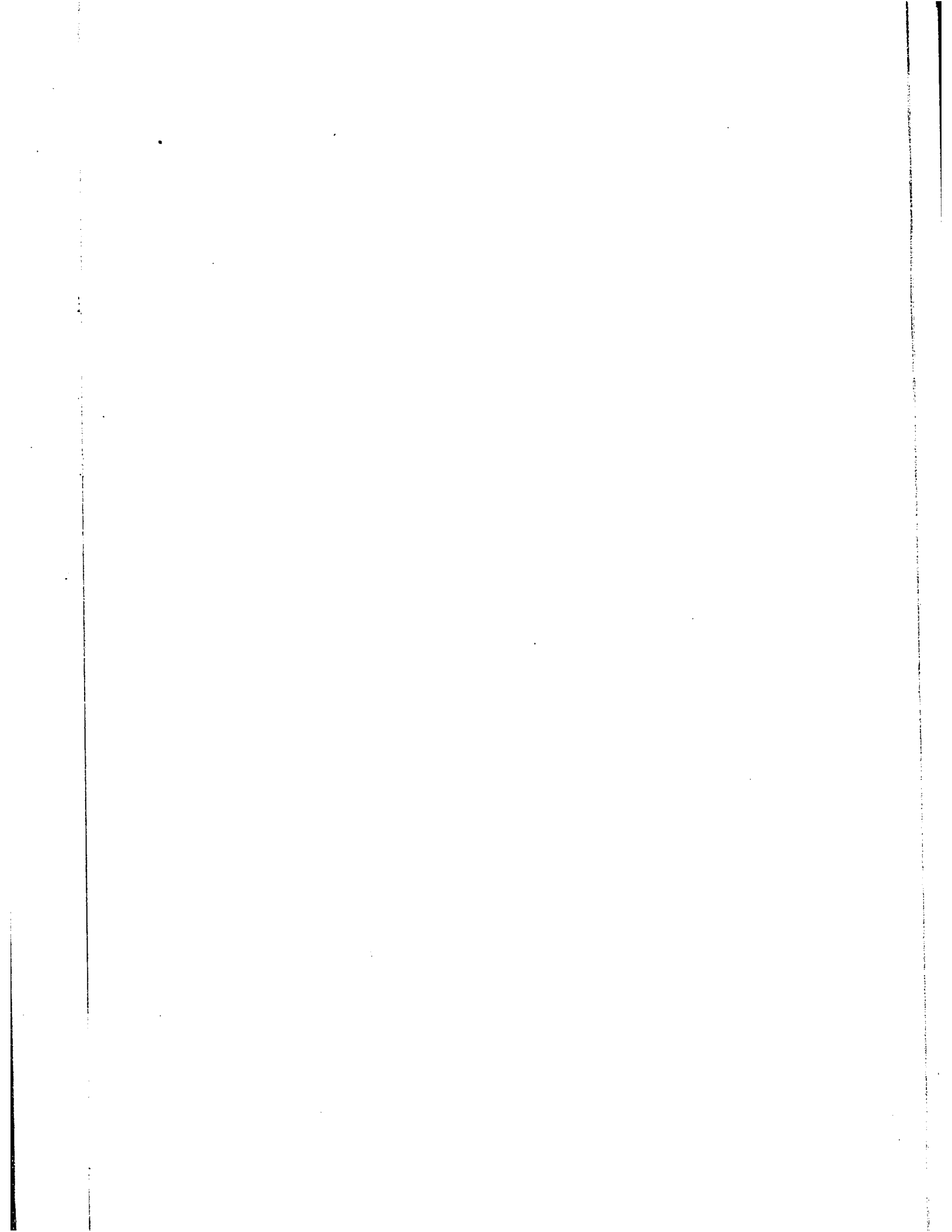
M. Authier

M.-L. Rivero

(LE DOYEN DE L'ÉCOLE DES ÉTUDES SUPÉRIEURES
ET DE LA RECHERCHE)

SIGNATURE

(HEAD OF THE SCHOOL OF GRADUATE STUDIES
AND RESEARCH)



Clausal and Verbal Structure in Farsi

Mohammad Mehdi Vahedi Langrudi



Thesis submitted to

the School of Graduate Studies and Research
in partial fulfillment of the requirements for the M.A.
degree in Linguistics

Université d'Ottawa / University of Ottawa

UMI Number: EC45139

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI[®]

UMI Microform EC45139
Copyright 2007 by ProQuest LLC
All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

ABSTRACT

This thesis intends to study very broadly and generally the verbal and clausal structure of the Farsi, the official language of the Islamic Republic of Iran. It tries to bring the clausal and verbal structure of Farsi language into the recent amendments of X-bar theory, i.e. the introduction of two different sets of categories, namely functional and lexical ones. We will use the principle and parameter model of generative grammar as our theoretical model of analysis.

The verbal structure of the language will be analyzed based on the proposal of Pollock (1989), Chomsky (1989), and Ouhalla (1990), where it is argued that Infl constitutes of two sets of functional projections, namely AgrP and T(ense)P. Further separate projections are proposed for Negation and Auxiliaries.

As far as X-bar theory is concerned, we will adopt a theory of X-bar which is least sophisticated and assumed to be more natural. We will adopt the theory of X-bar as proposed in Fukui and Speas (1986), Fukui (1986), and Speas (1990) among others which suggests that lexical items as well as functional ones project to a head, X^0 , and X^{max} , and any number of intermediate iterable bars. The other constraints on the structure of phrases and clauses are imposed by other modules and subtheories of grammar. We will show that the phrasal structure of Farsi does in fact match this system. Word order in main and subordinate clauses will be studied and shown that it is constrained with Specificity of the elements as discussed in Karimi (1989), as well as case theory and directionality of head placement. The Ezafe Construction will be studied in Chapter III. I will argue that Ezafe, which is a clitic-like suffix similar to English of and French de has three different functions in the language.

In chapter IV the structure of verb Phrases will be analyzed. We will see that the future auxiliary in Farsi is not a true auxiliary but a modal-like element which acts as a verb and subcategorizes for a subordinate main, lexical verb which causes the latter to incorporate into it at a higher level. Based on the function of negation in the language, we will also show that auxiliaries are base generated outside the projection of the verb. For the passives, we will argue that there are no real verbal passives in the language, rather passive constructions are adjectival formed in the lexicon. This is accounted for by the types of verb movements/raisings involved in the verbal structure, and the function and position of negation. Finally morphological causatives in Farsi will be brought in line with the "Incorporation" analysis of Baker (1988) and Li (1991).

ACKNOWLEDGEMENT

I should like to thank my thesis supervisor professor Paul Hirschbuhler who gave me enough confidence to get involved in this study. He has inspired and encouraged me in my keenness on theoretical linguistics all throughout my one-year period of study at the University of Ottawa. Had it not been his help, encouragement, and scientific preciseness, hardly half of this thesis would have been done by now. I would also like to express my gratitude to Professor M-L. Rivero for her excellent "Seminar on Syntax" which brought about almost all of the ideas and analyses in chapter IV of this thesis. Needless to add, of course, that all the errors remain my responsibility.

TABLE OF CONTENTS

	Page
CHAPTER I <u>THEORETICAL FRAMEWORK</u>	1
Introduction.....	1
Grammatical Relations in GB.....	2
Lexical Entry.....	3
The Semantic Portion of a Lexical Entry....	3
The Syntactic Portion of the Lexical Entry.	4
The Thematic Hierarchy.....	5
Dominance < Preference and the Phrase marker	5
Syntactic Projection and licensing.....	6
X-Bar Theory and the Projection of Heads...	7
The Aspects Model.....	7
Jackendoff's X-bar Schema.....	10
Stowell's Model of Phrase Structure.....	10
X-Bar Schema Revised.....	12
Complements and Adjuncts.....	13
Theta-Grids & Thematic Relations.....	14
Case Theory and Specificity.....	16
CHAPTER II <u>CLAUSAL STRUCTURE</u>	17
The Basic Word Order.....	17
VP-Internal & VP-External Adjuncts.....	26
Re-Ordering.....	31
XP-Postposing.....	34
Extraction to Main Clauses.....	38
Footnotes of Chapter II.....	41
CHAPTER III <u>PHRASAL STRUCTURE</u>	42
Sentential Complements.....	46
Objects/Modifiers of Nominals.....	47
Ezafe Construction.....	49
Attributive Adjectives.....	50
Attributive Noun Phrases (NP _a).....	51
Prepositional Modifiers (pp).....	53
Prepositional Modifiers Revised.....	56
Infinitives and Gerunds.....	64
Infinitives.....	64
Gerunds.....	71
Extending the Syntactic Feature.....	72
Footnotes of Chapter III.....	74

CHAPTER IV	<u>VERBAL STRUCTURE</u>	75
	Introduction.....	75
	Simple Tenses.....	77
	Future.....	82
	Causatives.....	86
	Passives.....	95
	Subjunctives & Imperatives.....	102
	Negation.....	104
	The Structure of IP.....	107
	Footnotes of Chapter IV.....	109
	Bibliography.....	111

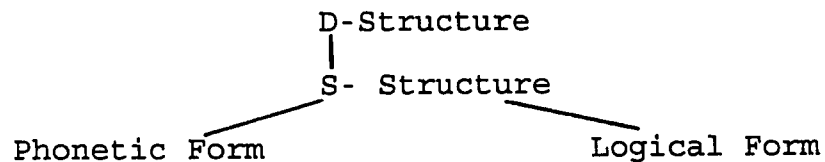
CHAPTER I

THEORETICAL FRAMEWORK

1- INTRODUCTION

In this chapter I will Present the theoretical framework upon which my analysis of the clausal and verbal structure of Farsi will be based. The general framework we adopt is Government and Binding (GB), in particular the theory of phrase structure and principles constraining the projection of sentence structure from lexical representations proposed by Speas (1990). It is a modular theory of grammar, regulating principles and systems which map between **Form** and **Meaning**, and further attempts to explicate Universal Grammar (UG). According to GB theory, mapping between form and meaning involves four levels of representation, i.e., (1) Deep Structure, (2) Surface Structure, (3) Logical Form, and (4) Phonetic Form, as indicated below:

(1) Levels of Representation



A structural representation at each of these levels may be indicated by a PHRASE MARKER in form of a tree diagram or bracketed structure which expresses relationships of hierarchical and/or linear constituency. These levels are related to each other by a general rule AFFECT ALPHA or (MOVE ALPHA), with α (=ALPHA) ranging over lexical and functional categories, and its nature varying cross-linguistically. We assume that D-structure is a pure projection of thematic information in the theta-grid of the lexical and/or functional items, (we will revise this point later on, in the course of exposing our theory). In other words D-structure is "a level of representation at which Grammatical Relations (GR) relevant to assignment of theta roles and only these have arguments bearing them" (Chomsky, 1981). This is accomplished when lexical entries are put together at D-structure. D-structure is formally subject to X-bar theory which regulates projection of lexical information from the lexicon. Semantically, D-structure must be a pure representation of theta structure of lexical items. D-structure is mapped into S-structure by application of the

Transformational rule, AFFECT ALPHA or MOVE ALPHA. S-structure is assumed to be an interpretive level of representation, simultaneously related to the other levels. It reflects the more superficial properties of sentences, the actual, linear ordering of the elements in the surface string, and their case marking properties. S-structure thus ties form and meaning. LF is the level of quantifier raising and Wh-raising in languages lacking wh-movement at S-structure. LF is linked to S-structure by Affect- α . LF also places elements of a clause in a position which represents their semantic scope and yields the semantic interpretation of the clause, though it is not assumed to be a semantic representation, rather a syntactic one; it is subject to syntactic rules and principles such as the ECP, and the Projection Principle. Phonetic Form (PF) is the superficial, phonetic realization of a clause which accounts for processes like simplifying and/or adjusting hierarchical structure, deleting null elements, stylistic movements, and contractions. (We will argue that deletion of prepositions of location and /or direction adverbials in Farsi occurs at this level.)

The Projection Principle is a fundamental principle of GB theory that regulates the projection of lexical information along with X-bar theory and constrains relations and mapping between the above levels. We will further explicate it in the next section. These levels of representation are further regulated, constrained and subject to other rules, principles and sub-systems of the grammar as listed below:

- (2) (i) Bounding Theory
- (ii) Government Theory
- (iii) Theta Theory
- (iv) Binding Theory
- (v) Case Theory
- (vi) Control Theory
- (vii) X-bar Theory
- (viii) Predicate Theory

In our study of clausal structure of Farsi we will study (ii), (iii), (v), and (viii).

1-1 GRAMMATICAL RELATIONS IN GB

Grammatical relations such as subject, object, are not primitive terms in the language, rather they are defined in structural terms. In other words grammatical rules and principles in a language i.e., English and Farsi, do not explicitly refer to GRs, rather they refer to structural positions such as [NP, S], [NP, VP], etc. It is not necessarily the case that all of the properties which are borne by a position such as [NP, S], subject, and [NP, VP], object in one language be equal to respective positions, [NP, S], and [NP, VP] in other languages. Speas (1990) following Travis (1984) assumes that rules of syntax may refer to syntactic entities like the following, rather than GRs,:

- (3)
- a. "external" or "most prominent" XP
 - b. parts of speech: V, N, A, P, I, C, and their projections
 - c. Arguments vs Non-arguments
 - d. Inherent case vs Structural case
 - e. Dominates, immediately dominates and the relations defined over these, such as government and C-command
 - f. +, - is assigned a theta role

1-2 LEXICAL ENTRY

Given the definition of the lexicon as a pure structural realization of the thematic properties of lexical items, we assume that the lexicon consists of a list of entries for all words, affixes, and stems each with its independent projection and subcategorization frames. Pollock (1989) argues in a similar fashion for the independent entries for functional categories, and attributes separate projections to them. We assume in this work that each lexical entry includes:

(4)

- (a) A phonetic Representation
- (b) Semantic Information
- (c) Syntactic Information
- (d) Morphological Information (i.e., stem, affix, and simple word)

1-2-1 THE SEMANTIC PORTION OF A LEXICAL ENTRY:

LEXICAL CONCEPTUAL STRUCTURE AND THE DEFINITION OF THE THEMATIC ROLES:

Thematic roles or theta roles (agent, patient, goal,...) are the semantic information of a lexical entry. Speas (1990), following Jackendoff (1987) and Levin and Rappaport (1986) assumes that grammatically relevant semantic information about a given lexical item is expressed in terms of a **LEXICAL CONCEPTUAL STRUCTURE** (LCS). An LCS expresses the particular situation named by a lexical item, with variables in it which represent the arguments of the lexical entry as below:

(5)

PUT:	X	causes	Y
		come to	
		be at	Z

The variables are projected into syntactic positions at D-S. The relation between LCS and the syntax is mediated through **predicate argument structure** (PAS). Any syntactic position associated with a variable in LCS is assigned a theta-role, hence in (5) the variable Z occurs in the context /be at-----/. Any such variable will always bear the location role. Thus theta roles are not primitives of grammar, either.

1-2-2 THE SYNTACTIC PORTION OF THE LEXICAL ENTRY

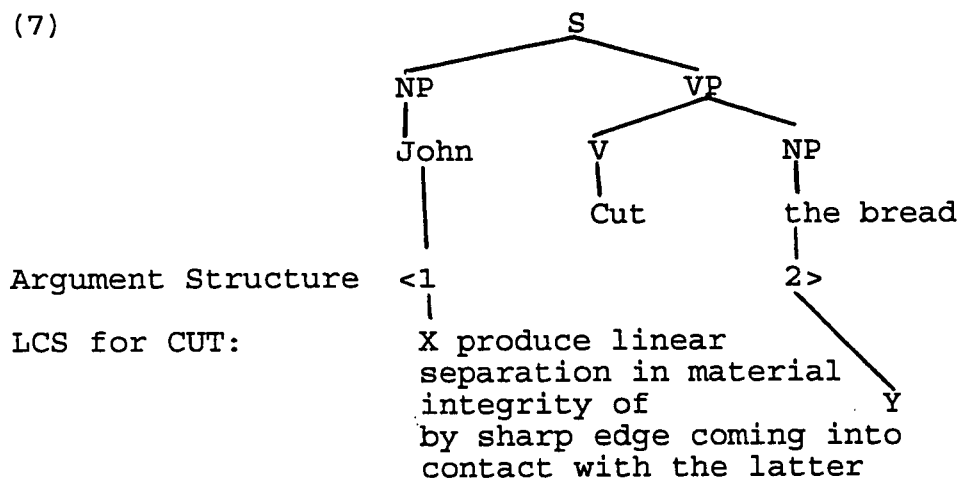
PREDICATE ARGUMENT STRUCTURE (PAS) = (THETA GRID):

PAS includes the syntactic information in a lexical entry. As mentioned above, Speas adopts Levin and Rappaport's (1986) and Hale and Keyser's (1987) view that PAS mediates the process of mapping between LCS and the syntactic representation. The argument structure indicates the number of (obligatory) arguments that a predicate requires which correspond to variables in LCS. The arguments are the participants minimally involved in the activity or state expressed by the predicate. Thus her PAS is similar to a theta-grid with variables rather than theta roles:

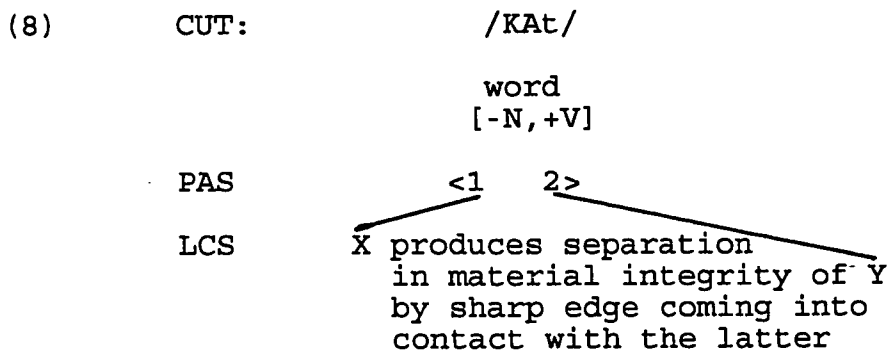
(6) CUT: <1,2>

She follows Higginbotham's (1985) view that all parts of speech may have an argument structure (theta grid). Contrary to the prevailing GB framework that assumes the "external" argument or the "most prominent" argument projects out of the maximal projection of its predicate (and indicates this by underlining), she proposes that the external argument does not project out of its predicate, rather it is projected as the highest daughter of its XP. We suggested that the variables in the LCS of a predicate are mapped into syntax through PAS. This is shown in (7) below, where the variables in LCS are linked to arguments in the PAS which are ordered according to the THEMATIC HIERARCHY, and are in turn linked to syntactic positions. Only under such a configuration "theta-roles" are assigned to syntactic positions:

(7)



The syntactic portion also indicates the part of speech of a lexical item in terms of syntactic features $[\alpha N, \beta V]$, where α and β range over + and -. As a result (8) counts as a partial lexical entry in our system:



1-3 THE THEMATIC HIERARCHY

We argued that variables in PAS are assigned a particular theta role only when they are linked to a variable in LCS. But how do we determine whether these variables project to subject, direct object, indirect object, ... position at D-structure? It is generally assumed in the literature that variables in PAS are arranged according to a THEMATIC HIERARCHY (TH), which regulates the association between theta roles and syntactic positions. We follow Speas who in turns adopts Ostler (1979) and Carrier-Duncan (1985) that Arguments (variables) in a theta grid are ordered according to a TH, so that the higher role of the TH in the PAS of a lexical item projects to a higher structural position in syntax:

Order a verb's theta roles according to the Thematic Hierarchy.
(Carrier-Duncan 1985,p.7)

The TH we adopt, following Speas, is as in (9) below:

- (9) AGENT < EXPERIENCER < THEME < GOAL < SOURCE < LOCATION
< MANNER / TIME

1-4 DOMINANCE< PRECEDENCE AND PHRASE MARKERS

We argued that underlying structures are projected from strings of lexical entries. That is to say that every string of words is associated with a PHRASE MARKER (PM) which expresses category information, constrained and regulated by X-bar theory, and two other separate linguistic relations i.e., **HIERARCHICAL** or **DOMINANCE**, and **LINEAR** or **PRECEDENCE** relations. They function as well-formedness conditions on every PM. X-bar theory differs from the last two relations in that the former "constraints the relationship between the PMs and strings of words whose syntactic organization the PM describes", while the latter defines the restrictions, relations, and conditions on the PM itself and not on the strings of words. These notions are important in that they express the configurations under which direction of theta/case marking, and the direction of head placement and other configurations for move α , relationship of strings in the PM as well as projection of lexical items take place. In other words they

refer to structural positions which are part of the innate faculty of the language acquisition and are relevant in the projection of categorial variables and may be considered as primitives of phrase structure.

2-SYNTACTIC PROJECTION AND LICENSING

This section deals with the projection of lexical information to D-structure and principles that regulate and constrain it. We will study the projection principle. We adopt the theory of the theta criterion as proposed by Higginbotham (1985), and X-bar theory. In the section on X-bar theory we will review the theory of phrase structure in the generative framework, starting from Chomsky (1965), and Chomsky (1970). We will study the essential framework of X-bar as proposed by Chomsky (1970) and developed by Jackendoff (1977), and the essential revision of its principles by Stowell (1981). We will then spell out in more detail the theoretical framework we will adopt in the subsequent sections.

The basic assumption in GB theory to account for the well-formedness of strings of words is the **LICENSING CONDITION** as suggested by Kim (1988) and Chomsky (1986):

(I) THE LICENSING CONDITION

Every element in a structure must be Licensed

Elements in D-structure, as mentioned before, must be a pure representation of the thematic relations of lexical items. To do so every structure must obey **UTAH** as proposed by Baker (1988):

(II) UTAH

Identical thematic relationships between items are represented by identical structural relationships between those items at the level of D-structure.

In other words all surface realizations of particular theta roles of a lexical item arise from identical structural relations at D-structure, (i.e., unaccusative and ergative predicates in contrast with their transitive, passive and raising equivalents). UTAH requires that arguments in a grid, which are arranged in conformity with TH, are mapped into syntax in such a way that prominence in TH directly corresponds to syntactic prominence.

The **Projection Principle** informally indicates that thematic and selectional properties of lexical items are projected and maintained at all syntactic levels (Chomsky (1981)). It also predicts that all daughters of V' (X') are theta marked by V (X). This, however, can not be a correct position, given the fact that in many languages non-theta marked phrases may and /or must intervene between the theta assigner and the theta receiver, as in Turkish, Farsi, and Navajo (Speas chapter II & III). We will argue for a theory of projection where all lexical categories project to a single bar level and allow for iteration of bar levels so that

all phrases (daughters of an X'-phrase) which are assigned a theta-role and are thematically related to the head predicate, are projected within a projection of the assigning head. Their surface ordering is determined by other principles of grammar. Theta roles are assigned directly to the internal arguments and compositionally to other arguments under sisterhood and government within the maximal projection (MP) of the assigning head, Fukui (1986).

Arguments at D-structure are also subject to the **Theta Criterion**, Chomsky (1986):

Each argument α appears in a chain containing a unique visible theta position P, and each theta position P is visible in a chain containing a unique argument α .

A position is visible if the chain that it is in contains a case-marked position. We will adopt the theory of theta criterion of Higginbotham (1985), who defines it in terms of "discharge" of positions in the theta grid, as below:

- (a) Every thematic position is discharged
- (b) If X discharges a thematic role in Y then it discharges only one

To summarize, the following general conditions constrain the mapping from lexicon to syntax:

- (I) The licensing condition
- (II) UTAH
- (III) Theta Criterion
- (IV) Projection Principle
- (V) X-bar theory

2-1 X-BAR THEORY AND THE PROJECTION OF HEADS

In this section I will study the theory of phrase structure in generative grammar as first proposed by Chomsky in **Aspects of the Theory of Syntax** (1965) and its subsequent development in **Remarks on Nominakization** (1970). I will then briefly examine the X-bar theory of Jackendoff and the later rejection of its basic assumptions by Stowell (1981). In the last section we will spell out the theory of X-bar schema as proposed by Fukui (1986), Fukui and Speas (1986), and Speas (1990).

2-1-1 THE ASPECTS' MODEL

Chomsky (1965) sets as the goal of his generative grammar the description of competence of a mature, native speaker of the language that enables him to understand and generate infinite number of sentences. He proposes that a generative grammar consists of three major components as: Syntactic, Phonological, and Semantic component. The syntactic component, which is the core of his investigation, consists of a lexicon, and a categorial component (or

phrase structure rules) which together constitute the base. The base allows generation of the Deep Structure (D-S) which determines the semantic interpretation of the sentences. A Transformational component applies to D-S to derive the surface structure. The categorial component of the base consists of a set of context free rewrite rules, similar to (10) below:

(10) A-----> Z / X---Y

where A is a single, non-terminal category symbol which expands into at least one non-terminal element. The categorial environment thus determines both the underlying linear and the hierarchical structure of the constituents of a phrase. The symbols appearing to the right of (10) are preterminal constituents. The derivation started with an initial structure from which all other related structures were derived by transformational rules. Sets of syntactic features characterize the lexical categories appearing in the pre-terminal string of phrase markers, which he names complex symbols. The final string is formed by lexical insertion rules which replace the complex symbols with the appropriate lexical formatives in the lexicon which match the information specified in the complex symbol. The set of complex symbols which contained specified distinctive syntactic/semantic features were produced by another subcomponent of the base, called strict subcategorization rules. The strict subcategorization rules for verbs further determined the categorial context in which it appeared together with the complex of its complements. The selectional rules, another subcomponent of the base, specified the distinctive, syntactic features of the context in which the lexical category, verb, appeared. Strict subcategorization rules and selectional rules together determined the properties of complex symbols of the preterminal strings. When lexical items were inserted into the phrase marker, the D-S was produced, from which the semantic interpretation of the sentence was derived. Now a set of powerful rules, called transformational rules were introduced to apply to D-S in order to determine the surface structure of words in a clause. An interesting but controversial property of words in this system, apart from expressive, powerful potency of phrase structure rules and transformational rules, was that derived nominals and adjectives were not independent lexical formatives of the lexicon, but were derived from the(ir) equivalent base sentences by transformational rules.

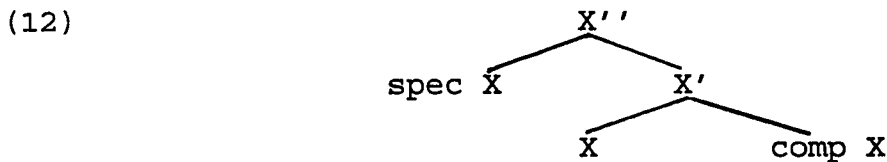
Stowell (1981) argues that the "rich expressive power of the transformational component, combined with the expressive power of the categorial rules of the base allowed for many possible formal accounts of phrase structure". Another problem with this system was that lexical similarities and relationships between the derived nominals, adjectives, and adverbs were attributed to a common DS and the respective surface forms were derived by ad hoc transformational rules which added lots of complexities, unnaturalness, and arbitrariness to the theory of grammar, apart from missing important cross-categorial parallelism between lexical

categories, and could not account for idiosyncratic properties. The derivation of semantic interpretation exclusively from the D-S led the generative semanticists to do away with Chomsky's view of the D-S and derive the surface structure straight from the semantic representation via transformations. Stowell argues on the basis of child language acquisition that phrase structure rules can not be taken responsible for the ordering of elements in D-S and proposes convincingly that phrase structure rules must be eliminated from the theory of grammar.

Chomsky (1970) attributes independent base values in the lexicon to nouns, adjectives, adverbs, and verbs, and later on to prepositions, (as proposed by Riemsdijk (1973) and Jackendoff (1973)). Their similarities in distribution were attributed to what Stowell calls "redundancy rules operating in the categorial component of the base". Chomsky (1970) also introduces the X-bar schema as in (11) below. He proposes a variable X to stand for the lexical category symbol adjectives (A), verbs (V), nouns (N), and later on prepositions (P). X stands for the head of a phrase i.e., A, V, N, and P. PSRs which introduced extensions of categories were replaced by X' and X'' introducing other constituents of the phrase. In this way he accounts for cross categorial generalizations and similarities among the lexical categories.

- (11) a. X''-----> (spec, X') X'
 b. X' -----> X...

where X'' is the specifier of the head, X' the complement structure. The rules for determining the structure of elements in X'' and X' for each category differed and were determined by constrained phrase structure rules, as in Jackendoff (1977). (12) below indicates the hierarchical organization of (11), adopted from Jackendoff (1977):



(12) indicates that every lexical item expands to a head and (may) consist(s) of a specifier at X'' level, and a complement on the X' level, each carrying the categorial features of the head. This implies that every phrase is endocentric. An abstract notion of (12) is assumed to constitute part of the language faculty and UG which controls and facilitates child language acquisition.

A third innovation of Remarks (1970) was the introduction of syntactic features +/-N and +/-V to characterize the major lexical categories which determine syntactic natural classes, similarities and dissimilarities of distribution and application of grammatical rule.

2-1-2 JACKENDOFF'S X-BAR SCHEMA

Jackendoff extends Chomsky's X-bar convention into "the uniform three-level hypothesis". He claims that all lexical categories constitute of three bar levels. He takes S as the maximal projection of verbs. His X-bar schema is presented in (13) below:

(13) X^n -----> ... X^{n-1} ...

He proposes that the value of n is the same for all categories, and even for categories that do not actually constitute of three bar levels, he assumes that they must have three levels in order to maintain the cross-categorial generalizations and claims this extension to be "harmless". He names the X''' a "Major Clausal Category". He classifies complements on semantic grounds into three distinct groups: (1) Functional arguments (arguments that are strictly subcategorized by a predicate); (2) restrictive modifiers; (3) non-restrictive modifiers. Now he correlates this 3-level semantic distinction to his 3-bar level of lexical categories, claiming that (1) above corresponds to X' , (2) corresponds to X'' , and (3) to X''' in English as well as cross linguistically. In case of more than one complement in each bar level, he attributes their ordering to some constrained PSRs. We will show that this is not true in the phrasal structure of Farsi, as the ordering are not due to X-bar levels in English either, as shown by Stowell (1981).

2-1-3 STOWELL'S MODEL OF PHRASE STRUCTURE

Stowell (1981, chap 2) provides convincing arguments against the theory of phrase structure as proposed by Chomsky and Jackendoff from the point of view of language acquisition. He correctly claims that the categorial rules of the base make no real predictions beyond the realm of X-bar theory (of Chomsky (1970)). The categorial identity and the mutual ordering of complements, adjuncts and specifiers within each bar level by categorial rules are descriptive and essentially ad hoc, and can not account for the facts of child language acquisition which exhibit specie-specific, principled, structured, innate and predetermined faculty in acquiring a language in a short time on the basis of a severely impoverished stimulus. Non-configurational languages which exhibit free word order, beside languages which stay in between configurationality and non-configurationality, would not receive a proper analysis given the categorial rules and would further complicate the process of language acquisition. Following Chomsky (1981) he argues that PSRs "recapitulate most of the information that is encoded independently in the strict subcategorization frame of individual entries". They have enormous descriptive but no explanatory power, and Stowell claims that they are redundant and must be eliminated from the theory of grammar. To conclude, by enumerating the pitfalls of the base component, he deprives it from

referring to categorial features and proposes a base component which lacks phrase structure rules. His base component now contains a rich lexicon with (1) strict subcategorization frames, which refer to the complements of predicates, (2) Theta-grid, which constitutes an unordered set of arguments of a predicate, and (3) Category neutral principles of X-bar theory, which determine the hierarchical structure of lexical items and function as a well-formedness condition on the Phrase marker. X-bar theory distinguished between head and non-head positions and also between specifiers and complements. In other words it referred to primitive structural positions valid for all categories. The order of complements and adjuncts within X-bars was determined by other independently required principles of language like case theory, theta role assignment and initial or final position of the head. X-bar theory turned out to play no role whatsoever in determining the fixed order of complements in such languages as English, and French, contrary to Jackendoff (1977), (see Stowell chap I, II, III). Stowell now concludes that the speakers of a language do not know the principles of phrase structure rules, rather they know the "innate" principles of X-bar theory. The basic principles that Stowell refers to are repeated here as (14):

- (14)
- a. every phrase is endocentric
 - b. specifiers appear at the X'' level, subcategorized complements within X'
 - c. the head always appears adjacent to one of the boundaries of X'
 - d. the head term is one bar level lower than the immediately dominating phrase node
 - e. only maximal projections may appear as non-head terms within a phrase

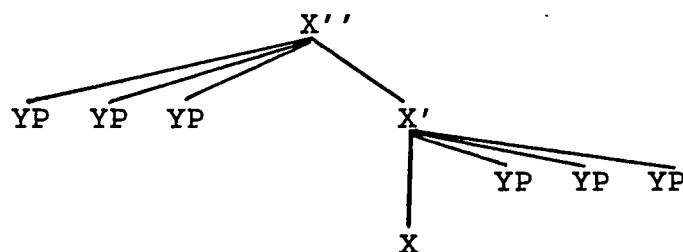
Speas (1990) argues that the X-bar conditions exemplified above are redundant and in fact empirically incorrect. She proposes a theory of X-bar which follows Koopman (1987):

- (15) Heads are dominated by maximal projections; maximal projections have heads

Reducing the redundancies of the X-bar schema in (14), we get the standard schema of X-bar theory of Chomsky (1986b, p 3)

- (16) a. a. X''-----> X' Y''*
- b. X' -----> X Y''*

b.



The formal schema restricts structures to only one node at each bar level. It does not allow iteration of intermediate bar levels (X') in particular, rather allows conjunction of elements at each bar. Binary branching of Kayne (1983), however, necessitates the iteration of intermediate bars. Speas, following Travis (1984) argues that principles of grammar do not specifically refer to X-one bar level, while they may refer to X^0 or X-max level. She argues that X-one bar or the intermediate levels have the status of a "sort of elsewhere case" to which no specific rules of grammar refers, and allows for the iteration of intermediate levels. There are no distinctions among intermediate bar levels from the point of view X-bar theory, and grammar in general. To conclude we would like to agree with Travis (1984) who says that:

although syntactic phrases may have internal structure which includes intermediate nodes, there is no justification for a component of grammar in which XP, X, and X' are treated as three types of syntactic categories. ...each phrase has a maximal projection and an X^0 , and any amount of structure in between the two... .

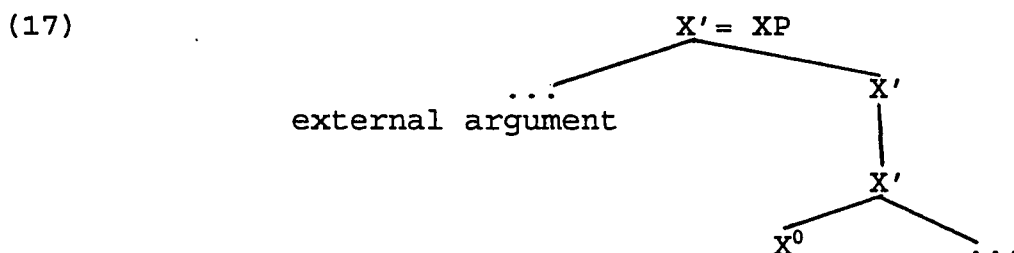
Our discussion so far implies that the number of complements and specifiers is not limited by the principles of phrase structure, X-bar. In the next section we will spell out the X-bar schema adopted in this thesis as proposed by Fukui and Speas (1986), Fukui (1986), Muysken(1983), Stuurman (1984) and Speas (1990).

2-1-4 X-bar Schema Revised

Speas (1990), following the other researchers adopts a theory of X-bar that does not differentiate between lexical and functional categories with regard to their mode of projection. All categories project to a single bar level and allow for iteration of that bar level. The number of specifiers and complements are not limited by principles of X-bar, rather by other sub-systems of grammar which are independently required by a modular theory of grammar. This is what she calls "PROJECT ALPHA":

A word of syntactic category X is dominated by an uninterrupted sequence of X nodes.

Presence and number of specifiers and complements are licensed by other modules and sub-theories of grammar, in particular case theory and the saturation principle, theta theory, (Fukui (1986) and Higginbotham (1985)) and direction of assignment of case and theta-role. Diagram (17) below represents the projection of lexical categories:



(17) indicates that all the arguments of a category project within the maximal projection of the head, no specific reference is made to intermediate levels, and all intermediate levels have identical hierarchical structure. Fukui (1986) proposes a theory of projection which differentiates lexical categories from functional ones. According to Fukui, lexical categories project to a single bar level, and allow for iteration of the bar level (similar to **PROJECT ALPHA** above) while functional categories may project to two levels, X'' , limited to a unique specifier and a single complement position if kase/case is assigned to their specifier positions; otherwise they will project to a single bar level. In the **PROJECT ALPHA** system there is no difference between lexical and functional category Projection, because bar levels are not assumed to be primitives of the system, and specifiers and complements are not distinguished from each other by the X-bar schema. The fact that functional categories are restricted to a unique specifier and do not allow for iteration (while LCs do) is not due to X-bar or phrase structure theoretic reasons, but is due to case and theta role assignment, and more precisely to a biuniqueness condition on kase assignment which requires spec head agreement (and adjacency) between a functional head and its specifier. As only one XP can be adjacent to the head and/or agree with a specifier (in English as well as Farsi), then there is only one specifier for functional heads. Thus in Speas' **PROJECT ALPHA** model which we adopt here, categories project to one-bar level, with the possibility of iteration, and any restriction on it must be attributed to other sub-theories of grammar.

2-2 COMPLEMENTS AND ADJUNCTS

One major consequence of a **PROJECT ALPHA** system which allows only one iterable bar level is the impossibility of XP over XP adjunction constructions at D-structure. Many linguists have argued that D-structure is a pure representation of the theta grid of a lexical item. This means that D-structure includes only heads and

arguments, and adjuncts are not present at D-structure, rather they are Chomsky-adjoined to the appropriate structure at the later stages of the derivation. Speas, however, distinguishes between two types of adjuncts:

(1) VP-internal adjuncts, which behave as though they must be present at D-structure. She calls them thematic adjuncts. They "behave as though they are governed by the verb and bear a thematic relation to the verb Speas (1990, P 52)", though they are not part of the theta grid of it.

(2) VP-external adjuncts, which are not present at D-structure, rather are added to it by Chomsky-adjunction at later stages of the derivation, by a generalized transformation, adjoin α .

She convincingly argues that locatives, benefactives, and instrumentals PPs fail to show anti-reconstruction effect in English while temporal and rational PPs do. She assigns the first group of adjuncts to type (1) and the second group of adjuncts to type (2). Rivero (1992) argues along the same line for two types of adverbs (=adjuncts) in Modern Greek: VP-internal adverbs, which may incorporate into the governing-verb, and VP-external adverbs, such as time and aspect adverbials, which fail to incorporate. For Farsi we will argue that time adverbials are VP-external adverbials, adjoined to VP/TP node. These observations support the **PROJECT ALPHA** system that claims that "a head projects indefinitely, allowing multiple VP internal adjuncts at D-structure...while VP-external adjuncts are added to the phrase marker by means of the transformation Adjoin α . D-structure contains all and only phrases which are licensed by theta theory. Phrases which are not arguments of the head (= which are not in its theta grid) are nonetheless licensed by theta theory if they bear some kind of thematic relation to the head, and thus may appear in the X^{\max} of the head as predicted by **PROJECT ALPHA**.

3-THETA-GRIDS & THEMATIC RELATIONS

We are adopting the theory of theta grids of Higginbotham (following Speas), according to which all lexical and functional items have a theta grid, and when they form a syntactic construction with each other, their grid positions must be discharged in order for the resulting constituent to be well-formed. In other words, for a given item in a syntactic structure to be licensed, it must enter into a "thematic relation" with its sister. "Modes of discharge" of grid positions indicate this thematic relation, which takes place under government and sisterhood. All heads have grids, whose positions must be discharged in order to license the non-head daughters. Discharge is informally defined as the "elimination of open thematic positions in lexical items and complex phrases", (Higginbotham (1985)). In fact "modes of discharge" of theta grids is a version of the theta

criterion which has been extended to all lexical items, as well as to functional categories. The relevant terms are defined below:

Grid G is saturated iff all positions in G are discharged.

Syntactic position X DISCHARGES a position P in grid G iff

X is a sister of Y which is labelled with G AND

X is a canonical structural realization of P

All grids appearing in the maximal projection of a lexical item must be discharged by entering into a grid relation with the theta-grid of the head. There are 4 methods of grid relations which allow the syntactic representation of a predicate to be composed into a constituent with the syntactic representation of its arguments and thus discharge (saturate) their grid positions, as follows:

(a) **THETA-MARKING**, exemplified by pairs consisting of a predicate and one of its arguments.

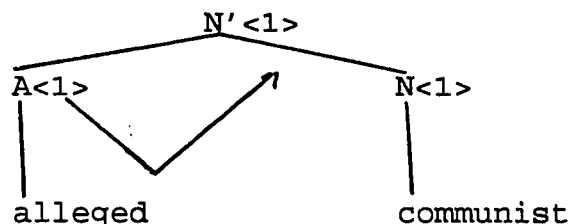
(b) **THETA-BINDING**, exemplified by determiners or measure words and their nominals, as in "every dog", interpreted as "for every x such that dog x'".

(c) **THETA-IDENTIFICATION**, exemplified in simple adjectival modification, as in "white wall", interpreted as "white(x) and wall(x)".

(d) **AUTONOMOUS THETA-MARKING**, where the value assigned to the open position in the theta marker is the attribute given by its sister constituent. (Higginbotham (1985))

The autonomous mode of discharge captures cases where the "properties denoted by the adjective (or adverb) are not absolute, but are relative to the choice of attribute". It concerns examples as fast, big, expensive, and alleged, etc., where the property of fastness, bigness, and amount of the price are relative to the object that they denote. A big butterfly is big for a battlerfly, but not for an elephant. An alleged communist is not necessarily a communist, but assumed to be a communist, as (18) below shows:

(18)



In (18) the two theta grids do not merge, rather the saturated position, as Speas defines it, "simply disappears", and it is the theta position of the modifier, N', which percolates up to the head N. This is contrary to the simple theta identification where the theta grid of the modifier, (that is the adjective or the adverb), is discharged by theta identification so that the two theta positions are merged into one unsaturated theta position, still required to be discharged. However, Speas attributes these semantic properties to the LCS rather than to operation in argument structure, as explained above.

An important case of relation between theta grids is when both sisters have saturated theta grids (their theta grids have no open position). This is the situation which distinguishes thematic (VP-internal) adjuncts from non-thematic (VP-external) ones, as defined below:

- (1) A phrase P is present at D-structure iff P is an argument or P has an open position in its theta grid.
- (2) A phrase P may be added by a Generalized Transformation only if P bears a saturated theta grid.

4-CASE THEORY AND SPECIFICITY

Case theory derives from the licensing condition which requires that every position must be licensed. Now as the projection from argument structure to D-structure is licensed by the theta criterion, the projection principle and other principles presented in the previous sections, the derivation from D-structure to S-structure must obey certain additional principles as well, the most important of which is case theory. Case theory imposes the case filter on the occurrence and distribution of any bare NP:

CASE FILTER: *NP if NP has phonetic content but not case.

Apart from the case filter which licenses the occurrence of NPs in a structure, the order and case marking properties of arguments in a structure may also be constrained and regulated by the nature of the arguments, i.e., whether they are definite or indefinite, specific or nonspecific, as discussed by Belletti (1988), Enc (1991) and Karimi (1989). We will see that the variable, and sometimes free order of constituents, in particular of the direct objects, can to a certain extent, be accounted for by the nature of the NPs, i.e., specificity/non-specificity of the arguments.

CHAPTER II

CLAUSAL STRUCTURE

1-THE BASIC WORD ORDER

In this chapter we study the basic word order in main and subordinate clauses in Farsi. The unmarked word order of elements in finite clauses and (Type-1) INF(initives) indicate that the language is a SOV one, as the following examples may indicate:

- (1) a. ali ketab-o bara-m kharid
ali book-DOM for-CL/me bought
Ali bought the book for me.

It is almost equally possible to exchange the position of direct object (DO) and the indirect object (IO) in (1) above as below:

- (1) b. ali bara-m ketab-o kharid
ali for-CL/me book-DOM bought

ra follows all specific DOs (Karimi (1989)). While DO and IO in (1) above seem to be free inside the VP, it is not so with regard to generic, nonspecific DO-NPs, as (2) may show:

- (2) ali bara-m ketab kharid
ali for-CL/me book bought
ali bought me (some) book.

The DO in (2) has an indefinite, generic reading. It must stay adjacent to the verb. Its movement inside the VP causes markedness, contrastive, and sometimes an ungrammatical reading, while preposing it to IP initial position causes a topic, and contrastive reading:

- (3) a. ali [_{VP} ketab bara-m kharid]
ali book for-CL/me bought
b. ketab ali bara-m kharid

While (3 a) may seem odd and marked, in (3 b) the DO-NP, ketab, is focused. As a study of word order in the language seems to be closely related to the phenomenon of specificity and the function and distribution of ra / ro / -o, I will present a short review of the phenomenon of specificity and its role in word order in the

language, as mainly proposed in Karimi (1989).

Karimi (1989) argues that ra, or its morphological variants ro /-o/, is a specific marker for NPs occurring in clauses, when it is not in the minimal government-projection of a N(oun), A(djective), and P(reposition). Thus ra is not a specificity marker for DO only, and other NPs may also be followed by it if they satisfy the structural description of its distribution, (e.i. 8, 9, 10, 11). ra must follow both definite and indefinite DO-NPs which are specific, as examples below indicate:

- (4) a. man diruz un mard-o did-am
I yesterday that man-DOM ps/saw-1S
I saw that man yesterday.
- b. man diruz ye mard-i ro did-am
I yesterday one man-Indf DOM ps/saw-1S
I saw a man yesterday.
- c. man amir*(-o) zad-am
I amir(DOM) ps/hit-1S
I hit Amir.

In (4 a) un mard "that man" is definite due to the determiner un "that". (4 b) is indefinite due to ye "one", and -i an indefinite suffix. The fact that ra / ro / -o must follow both the definite and the indefinite NPs indicate that it marks DOs for specificity, as proposed by Karimi (1989), who has in turn adopted it from Browne (1970). I would also like to add the following examples from Karimi (1989) chapter II, (her examples (9) and (10)) to further explicate the idea:

- (5) a. emruz bayad in doxtar-o be-bin-am
today must this girl-DOM Subj-prs/see-1S
I must see this girl today.
- b. emruz bayad ye doxtar-i-ro be-bin-am
today must a girl-IND-DOM Subj-prs/see-1S
I must see a girl today.

In both (4 b) and (5 b) the DOM must obligatorily follow the indefinite DO-NPs, otherwise the sentences would turn ungrammatical. We define specificity as "the selection of a Particular individual from a set of individuals", as Karimi suggests. Notice examples below where DO-NPs are not followed by ra. The (a) examples have an indefinite (non-specific) reading, and the (b) ones a generic (non-specific) one.

- (6) a. man bara ali ketab kharid-am
I for ali book ps/bought-1S
I bought a book for Ali.
- b. Man bara ali ye kitab-i kharid-am

- I for ali one book-IND ps/bought-1S
I bought one (some) book for Ali.
- (7) (Karimi's (5))
- a. diruz ru miz ye sib gozasht-am
yesterday on table one apple ps/put-1S
I put an apple on the table yesterday.
- b. diruz ru miz sib gozasht-am
yesterday on table apple ps/put-1S
I put an apple on the table yesterday.

As (6) and (7) clearly show ra /ro /-o does not follow DO-NPs when they are non-specific (generic or indefinite).

The following examples, adopted from Karimi (1989), however, indicate that ra is not only a specific DO-NP marker:

- (8) shab-e pish-o aslan na-khabid-am
night-EZ last-DOM at all Neg-ps/slept-1S
As for last night, I did not sleep at all.

(8) is also acceptable without -o, but the sentence (9) below is ungrammatical without it:

- (9) man zud shahr-o tark kard-am
I soon city- leave ps/made-1S
I left the city soon last night.

However (9) may be controversial as shahr "city" may also be considered as a DO-NP, subcategorized by the verb. (9) is ungrammatical without an NP which denotes a location. If it is not, then it indicates an adverb of location. It may be questioned with the Wh-word koja "where". In (10) below the verb tark kardan "to leave" selects a DO-NP which does not denote location, and must be obligatorily present:

- (10) u zan-esh-o tark kard
he wife-his-DOM leave made
He left his wife.

The DO may only be questioned with ki "who(m)", that may only replace an animate NP. However in both cases the NPs may not be used without ra, which denotes specificity of an NP. Now notice the example below where the repeated IO in topic position takes ra, but the real IO in its base position does not.

- (11) = (Karimi's 16)
man-o, be-h-em mi-khand-e
I-ra to- -me ASP-prs/laugh-3S
As for me, he/she laughs at me.

man "I" is in topic position and is coreferential with the IO be-h-

em "to me", which under no circumstances may be followed by ra. It is not in topic position but in its base one, and as such is governed by the preposition be, and receives oblique case from the preposition be. man in topic position inherits the oblique case from IO, but is not governed by the preposition¹. Let us now consider (12) below:

- (12) batche-ha-ro, be-h-eshun tazakor dad-am
 the child-PL-ra to- -them warning gave
 The children, I warned them.

tazakor dad-an "to warn" selects a PP-complement. batche-ha is in topic position, and is coreferential with the IO-PP, be-h-eshun which may not itself be followed by ra. The topic NP must be followed by ra, otherwise the sentence would be ungrammatical. It inherits the oblique case from its co-referent IO-NP. However, all NPs in topic position are not followed by ra, as (13) below shows:

- (13) ketab, pro be-h-esh bego ba khod-esh na-y-ar-e
 book, pro to-him tell with himself Neg-prs/bring-3S
 As for the book, you would better tell him not to bring
 with himself.

Thus we may agree with Karimi who considers ra as a sign of specific topic NP, and not topic in general. However, specific NPs which function as subject, IO, and predicate nominals may not be followed by ra either, as proposed by Karimi (1989):

- (14) a. ali (*ro) raft
 ali (ra) went
- b. pro ketab-o be ali(*-o) dad-am
 pro book-DOM to ali gave-1S
 I gave the book to ali.
- c. un mard [_{NP} baradar-e man]-e (=past)
 that man brother-EZ I is(= is)
 That man is my brother.
- d.* un mard [_{NP} baradar-e man]-o ?ast

In (14 a) ali a specific NP in subject position may not be used with ra. (14 b) shows that ra may not follow an NP which is governed by a P. (14 c-d) show that a predicate nominal which also receives NOM case by co-indexing with the subject may not be used with ra.

On the basis of these facts Karimi suggests the following condition (15) to account for the distribution of ra in farsi:

- (15) **SPECIFIC OBLIQUE CASE**

ra indicates oblique (#[+NOM]) case on an NP iff the latter is

- a. [+specific]
- b. not in the minimal government projection of α , $\alpha = N, A, \text{ or } P$.

As the subject position receives NOM case in (14 a), the sentence becomes ungrammatical. (14 b) with -o (ra) following the PP is bad, as it violates condition (15 b). (14 d) where the predicate nominal, baradar-e man is in an equative relation with the subject within the minimal government projection of S, is co-referent with the subject, thus it receives [+NOM] case, and does not satisfy (15). If we return to examples (8) and (9), we see that shab-e pish and shahr in (8) and (9) respectively are not in the NOM case positions. They are not furthermore, in the minimal government projection of α , but are specific; they are also in oblique case as we will see in chapter (3), thus must be followed by ra. Let us finish this subsection on specificity of NPs with the following example from Karimi's (1989) example (12) chapter II:

- (16) sepide, *(-ro) (pro) be-h-esh, goft-am
 sepide-ra to- -her told-I
 As for Sepide, I told her.

The clitic pronoun esh (see chapter IV for an analysis of clitics in Farsi) is governed and case marked by the preposition be "to", thus violates (15 b) and may not be followed by ra. But sepide is in an A-bar topic position and is co-indexed with esh, thus it inherits the [+oblique] case of esh. It thus satisfies both conditions (a) and (b) of (15) and must be followed by ra. Absence of it in (16) causes ungrammaticality¹.

Given these facts on the status and function of specificity, with thanks to Karimi's interesting analysis, I would like to agree with her (and disagree with the majority of the others working both in generative and structuralist framework) in proposing (17) as the unmarked word order of the major phrasal categories of a finite clause:

- (17) S IO DO V

I would like to emphasize that DO in (17) stands for nonspecific, generic DOs which obligatorily stay close to the verb and may not carry ra, (note that in an example like (3) the non-specific DO is in a contrastive topic position and complies with our discussions so far). But specific DO-NPs are freer in the VP. They must be followed by ra. Thus we adopt the view that ra is a sign of ACC case for specific (definite or indefinite) DO-NPs in Farsi (apart from other uses of ra with non-DO-NPs). Non-specific DOs do not carry a case marker. Nonspecific, generic DOs may not depart from their base position in (17), but specific DOs may. We will study certain conditions on the word order inside and outside the

projection of VP shortly in this section. The idea that specific DOs may carry some kind of case but IOs may not is not particular to Farsi. Enc (1990) argues for a similar situation in Turkish where specific DOs are followed by a specificity case marker but non-specific DOs are not, as (18) below shows:

- (18) a. İki kız-i taniyordum
two girl-ACC I-knew
I knew two girls.
- b. İki kız taniyordum
two girl I-knew
I knew two girls.

As Enc argues: " The difference in case marking correlates with a difference in the interpretation of these sentences". (18 a) is about two girls who are included in the set of children, already established in the context of discourse, hence specific, while (18 b) involves a first mention of the children, hence indefinite and non-specific. Belletti (1988) argues for a similar phenomena in Finnish. She claims that an object NP may bear case apart from ACC in Finnish which shows a rich case system. In example (19), below, as Belletti argues, the DO-NP may either bear ACC or Partitive case, depending on the reading of the DO. She further evaluates partitive case as an instance of inherent case which is assigned by a theta-role assigning head in a particular configuration.

- (19) a. Han pani kiriat poydolle
he put the books (ACC, PL) on the table
- b. Han pani kirjoja poydalle
he put (some)books (part, PL) on the table

As we notice the DO in (19 a) is definite, as reflected by the in the translation, but in (19 b) it is indefinite. She attributes this phenomenon to a "Definiteness Effect" (DE). She argues that although unaccusative verbs may not assign structural (ACC) case, they may still preserve the capacity to assign partitive (inherent) case. "The partitive case selects an indefinite meaning for the NP that carries it". She concludes that "DE typically appears in the object position of unaccusative verbs". The object of such verbs must be indefinite as (20) below shows:

- (20) a. A man is in the garden.
b. The man is in the garden.
c. There is a man in the garden.
d.* There is the man in the garden.

Enc (1990) argues that the semantic notion involved in Belletti's argument is "specificity", rather than "DE".

If these arguments are correct, then we assume that (17) is in fact the correct word order of major elements of the clause. Now

let us consider examples below:

- (21) a. ali diruz ketab-o bara-m ?avord
ali yesterday book-DOM for-me brought
Ali brought the book for me yesterday.
- b. ali diruz bara-m ketab-o ?avord

The order of DO and IO in (21 a-b) are exchanged. They are equally acceptable in the language, though I believe that the example (a) where the IO stays closer to the verb is more frequent. In fact in finite clauses and our type-1 INFs where the verb takes its arguments to the left, all arguments under the projection of VP may exchange position in most cases, apart from the non-specific-generic DO-NPs. Notice the following examples:

- (22) a. ali bara-t az ketab-khane ketab ?avord-e
ali for-you from library book ppr/has brought
Ali has bought book for you from the library.
- b. ali bara-t ketab az ketab-khane ?avord-e
ali for-you book from library ppr/has brought

The example (b) is marked, and hardly used unless it carries a contrastive or focused reading on the word ketab which is a non-specific generic DO, otherwise it might even be considered ungrammatical by some native speakers. But notice (23) where the DO-NP is specific and must carry the specificity marker ra:

- (23) a. man in ketab-o bara-t az ketab-khane ?avord-am
I this book-DOM for-you from library ppr/have brought
I have brought this book for you from the library.
- b. man bara-t in ketab-o az ketab-khane ?avord-am
I for-you this book-DOM from library ppr/have brought
I have brought this book for you from the library.
- c. man bara-t az ketab-khane in ketab-o ?avord-am

All the examples in (23) are equally acceptable, though I believe that (a) is more frequent than (b) and (b) is more frequent than (c). In other words when there is a specific DO-NP followed by ra, we seem to have (24) below as the dominant order, though other orders are also possible:

- (24) S DO IO(=PP) PP V

where the second PP stands for adverbials of location and direction. These latter especially in colloquial speech, tend to drop the head Preposition and appear post-verbally. Even when we seem to have a non-specific, indefinite DO, the order seem to be looser than with a non-specific but generic DO (i. e. , (22)) which

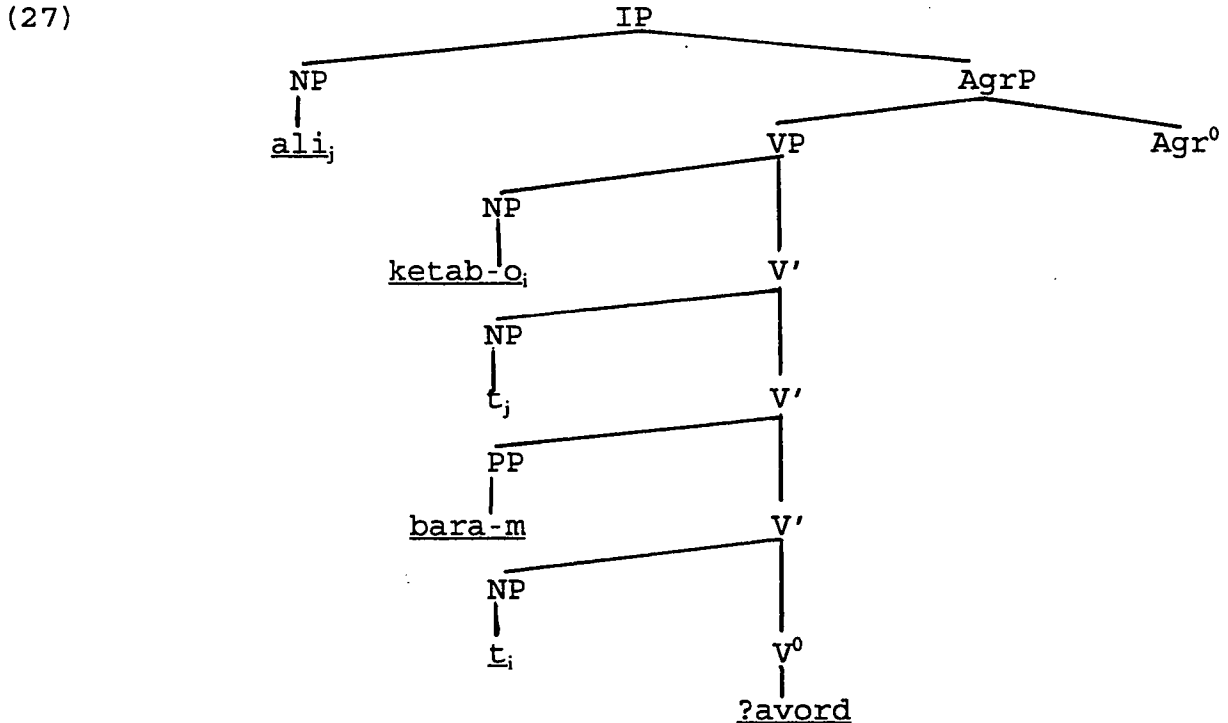
must stay close to the verb. Notice example (25) below:

- (25) a. ketab-ha-ye khubi bara-t kharid-e
 book-PL-EZ good for-you ppr/has bought-3S
 He has bought good books for you.
- b. bara-t ketab-ha-ye khubi kharid-e
 for-you book-PL-EZ good ppr/has bought-3S
 He has bought good books for you.

The underlined NPs in (25) are non-specific DOs as evidenced by the absence of ra. Both orders are equally good for me.

Given these facts I will maintain (17) as the unmarked word order in Farsi. In order to account for the non-adjacency between the verb and its specific (indefinite or definite) DO under the assumption that there is an adjacency requirement for case assignment (Stowell (1981)), I suggest (27) for the reordering of the specific DOs from their base position. It may be assumed to be an instantiation of the move α where the DO is Chomsky adjoined to a higher level of structure, still within the projection of VP.

- (26) ali ketab-o_i bara-m t_i ?avord.
 ali book-DOM for-me ps/brought
 Ali brought the book for me.

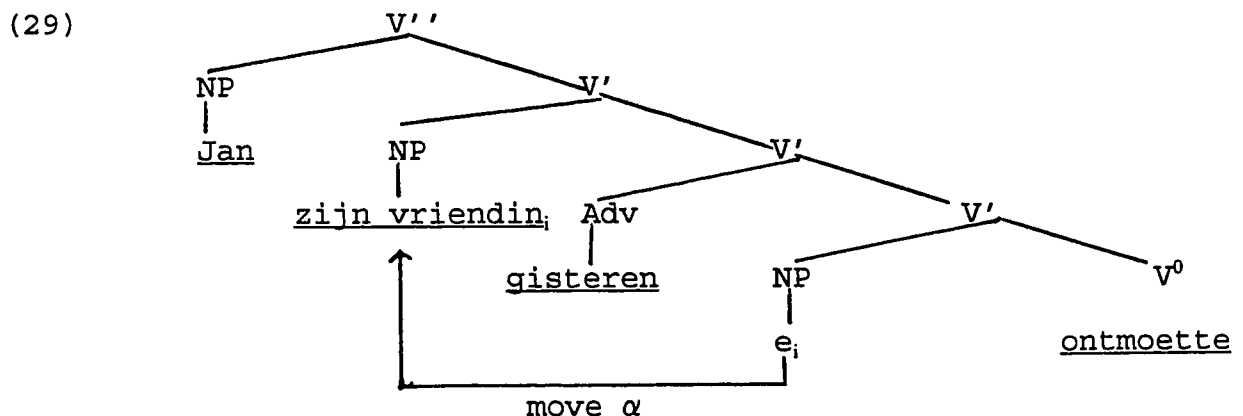


The specific DO, ketab-o, is base generated adjacent to the verb and is adjoined to a higher position within its projection. Remember that according to our PROJECT ALPHA the subject is also

base generated within the VP and is raised to [NP IP] position for case reasons. (This latter notion is however irrelevant for our present discussion). The adjoined position of the DO is still within the government domain of the head verb (see Fukui (1986)), but case is assigned to the trace adjacent to the verb and the DO-NP receives case from the chain $\langle \text{DO}, e \rangle$. The fact that the adjoined position of the DO is governed should not be surprising. Belletti & Rizzi (1981) and Rizzi (1982) convincingly argue that the subject in Italian, an SVO language, in the post-posed, right-adjoined position to the VP is governed by the verb, and hence extraction of the subject is possible only from the post-posed, adjoined position and not from the pre-verbal base position of the subject which is not governed by the verb. If they are right in their arguments, then we assume that our adjoined DO, which is still within the government domain/direction of the verb, is governed by the verb, if we define government as M-command within X^{max} , although the DO receives its case from the chain $\langle \text{DO}, t \rangle$, where the t is the trace adjacent to the verb. If we accept that our analysis of specificity and non-specificity of the DOs is correct for Farsi, (as I assume that we have to), since similar phenomena are attested in other languages as well, and observing the fact that the non-specific generic DO-NPs may not stay disjoint from the verb, our analysis of adjunction in (27) should not be regarded as an ad hoc rule for saving the adjacency requirement of case assignment. Hoekstra (1984) argues for a similar adjunction rule for Dutch as (28) below adopted from him indicates:

- (28) dat Jan zijn vriendin gisteren ontmoette
 that John his girl friend yesterday met
 That John met his girlfriend yesterday.

As Hoekstra claims (28) displays an apparent violation of the adjacency requirement. The DO-NP is separated from its case assigning head-verb by an adjunct modifier. It is also a counter example to his proposal that "modifiers always occur peripheral to complements". He proposes the following diagram, (29), for the structure of (28):



The higher V' node is created by Chomsky-adjunction of "his girl friend". Thus it was adjacent to its case assigner in D-structure. The empty category receives case, and thus the chain of <zijn vriendin, e> has a case marked position and is visible for theta role assignment. As we already indicated, the adjoined position is still within government domain of the verb, as Rizzi's (1982) adjoined subject to post-verbal position is. For the fact that adjunction takes place to the left of the verb, both in Farsi and Dutch and not to the right, we may claim that verbs in these languages govern to the left and not to the right.

Furthermore it is implausible and contrary to the X-bar schema adopted in this thesis and in general to assume two different types of projections; one for non-specific DOs adjacent to the verb, and the other apart from it. Thus we conclude our discussion of basic word order in Farsi by adopting (17) repeated as (30) for convenience:

(30) S IO DO V

In adopting (30) we are agreeing the view of word order as defended by Karimi (1989), and disagreeing with the majority who have proposed S + DO-ra + IO + V as the unmarked order; (see Karimi for a more detailed discussion of word order in Farsi).

2-VP-INTERNAL & VP-EXTERNAL ADJUNCTS

To talk of the position of certain elements within a clause in Farsi seems to me to be something ad hoc, as it sometimes looks that one may adopt a certain order as the base order and argue for it, against the other orders. In the study I made of the position of time adverbials both in colloquial speech and in written language, as school Farsi textbooks and Farsi papers, it looks that time adverbials are freer than other elements in a sentence. They normally appear at the periphery of the complements and subcategorized phrases. They mainly follow the subject and precede other elements in the clause. However their occurrence closer to the verb may lead to marked or odd sequences, particularly if the adverb of time is a long one.

(31) a. ma farda t_i mi-rim madrase_i
 we tomorrow ASP-prs/go/1S school
 We go to school tomorrow.

b. ma farda madrase mi-rim

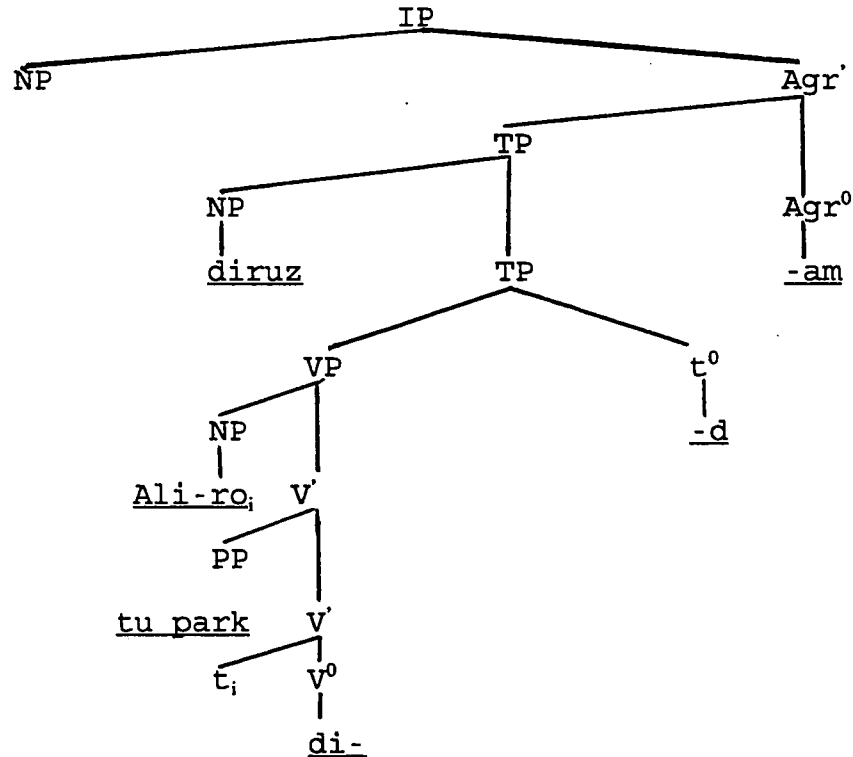
In (31 a) the base position of the underlined time-adverb is before the empty category. The adverb of location, madrase, has been postposed by move α . It is the most normal position for it in such sentences. (31 b) is also good and completely acceptable. However the sequence in (32) below where the location adverb precedes the one of time is odd and has a contrastive reading on madrase.

- (32) ma (be) madrase farda mi-rim.
 we to school tomorrow ASP-prs/go/1st/pl
 We go to school tomorrow.

I suggest that madrase is adjoined to TP in (32) to a position higher than farda, which is itself adjoined to TP in its base position. Now let us consider (33) below:

- (33) a. pro diruz ali-ro tu park did-am.
 pro yesterday ali-DOM in park ps/saw-1S
 I saw ali in the park yesterday.
- b. pro ali-ro diruz tu park did-am
- c. pro ali-ro tu park diruz did-am

(33 a-b) are quite normal, though I would consider the first variant better and more frequent. Example (33 c) is less normal than the other two. Assuming ali is adjoined to a higher position in the maximal projection of the verb as discussed above (rule 27), diruz is then at a position adjoined to TP in (33 a). The figure (33d) below, further indicates this point:



In the next two chapters I will argue that the verb phrase is always dominated by a tense projection (TP) whenever it takes its complements to the left under its maximal projection, and governs and case marks them. This is true both for finite clauses and type-

1 INFs in Farsi which take their complements under the maximal projection of the verb, V^{\max} , and in turn must be dominated by a TP. I would like to suggest that time adverbials in Farsi are VP external elements which are not part of the argument structure of the predicate, thus not present at D-structure, rather they are added in the course of derivation to S-structure. As such they are adjoined to the maximal projection of the TP which always dominates the VP. They do not constitute an argument and bear a saturated theta-grid, as discussed in chapter I. This is also compatible with the proposal of Hoekstra who suggests that adjuncts appear at the periphery of obligatory complements. Our examples (31-32-33) also indicate that adverbs of location and direction normally occupy a position adjacent to the verb but may be freely exchanged with the IOs(=PPs), or appear in adjoined post-verbal position. They constitute VP internal elements in our system even though they may not be complements of a predicate in its theta-grid, rather they have an open position in their theta-grids, and are considered as thematic adjuncts in our system. Now let us consider (34) below, where diruz "yesterday" does not constitute a complement of the verb, but be madrase "to school" does (see chapter I as well) :

- (34) a. ali diruz (be) madrase na-raft
 ali yesterday to school Neg-ps/go
 Ali did not go to school.
- b. ali (be) madrase diruz na-raft

Example (34 b) is quite marked, and odd, unless we attribute a contrastive reading to madrase in contrast with some other place. Even so, it must be accompanied with an appropriate intonation. If this analysis is correct, then I suggest that in (34 b) madrase (with the contrastive reading), is adjoined to TP, higher than diruz. Notice also example (35) below where the place and time adverbs are used with an INF. The complex phrase formed with the location adverbial and the INF is all right, but the one with time adverb is not.

- (35) a. madrase raft-an baraye bache-ha lazem bud
 school INF/to go for child-PL necessary was
 Going to school was necessary for the children.
- b. * diruz raft-an baraye bache-ha lazem bud
 yesterday INF/to go for child-PL necessary was
 Going yesterday was necessary for the children.
- c. raft-an(-e) (be) madrase / madrase raft-an
 INF/to go(-EZ) (to) school
 Going to school.
- d. !! raft-an(-e) diruz / diruz raft-an
 INF/to go(-EZ) yesterday

going yesterday.

- e. man un-o moghe-?e raft-an be madras / (or)
 I him-DOM the time-EZ INF/to go to school / or
- madrase raft-an did-am
 school INF/to go ps saw-1S
 I saw him while going to school.

- f. * man un-o moghe-?e raft-an-e diruz / (or)
 I him-DOM the time-EZ INF-to go-EZ yesterday / (or)
- diruz raft-an did-am
 yesterday INF/to go ps/saw-1S
 I saw him while going yesterday.

These examples show that while forming complex verbal (or infinitival) units with a location adverb and a verbal stem is quite normal, it is not so with regard to time adverbs and verb stems. The result with time adverbs seem to be unacceptable semantically and syntactically. Thus we may conclude that time adverbials may not be considered as VP-internal adverbials, but adverbials of location and direction must. They further suggest that be madrase "to school" does constitute a complement of the verb raftan "to go", and is a VP-internal adjunct as proposed in chapter I.

Given these observations, I suggest that adverbs of location and direction are VP-internal adjuncts, bearing an open theta grid and have some kind of thematic relation with the verb, while adverbs of time are VP-external adjuncts with saturated theta grids and no thematic relation with the verb. They are adjoined to the projection of TP which always dominates VP of finite and type-1 INFs in the language. These suggestions, I believe, to be well motivated, with the observations we made on the distribution and function of these adverbs in the language.

Apart from sentential adverbs which occupy a peripheral, adjoined position with regard to the finite clauses, and normally appear at the beginning of them as (36) below indicates:

- (36) a. bache-ha va.ghe.?an kheyli kar kardan(-d)
 child-PL really very work ps/made(-PL)
 The children have really worked a lot.
- b. va.ghe.?an bache-ha kheyli kar kardan(-d)

Adverbials of manner seem to be free in the VP. As I mentioned before, different possible word orders in Farsi make it possible to make opposite and contrasting claims with regard to the underlying position of elements, specially for adjuncts, in order to accommodate different purposes. As far as adverbials of manner are concerned, they may appear in different positions in the clause. For me their position in the VP initial position preceding or

following the specific DO-NPs is the most natural and normal one. They are more frequent in these positions both in formal and colloquial speech. The following examples are borrowed from Hashemipour (1989):

- (37) a. moin ketab-o tond khund
 moin book-DOM fast ps/read
 Moin read the book fast.
- b. moin tond ketab-o khund
 moin fast book-DOM ps/read
- c.* ketab-o tond khund moin
- d.* moin tond khund ketab-o

While (37 c-d) are reported as unacceptable by Hashemipour, they are not for me, though not frequent. Nor should they be unacceptable for Karimi (1989). In fact the core of Karimi's thesis is to provide arguments for the grammaticality of, and possibility of post-posing NPs such as those in (37 c-d), as we will study them in the next section. Notice also the following examples which are reported as ungrammatical by Hashemipour, but are all right for me with a contrastive or topic reading on the post-posed element; so are they for Karimi(1989):

- (38) a. * tond moin ketab-o khund.
- b. * tond ketab-o moin khund
- c. * ketab-o tond moin khund

The examples in (38) are correct for me both in colloquial and formal speech, but less common in writing where there no intonation accompanies the preposed elements. Before coming to any final conclusions, let us consider some examples of our own:

- (39) a. ali ba ajale khuna-ro tark kard
 ali hurriedly house-DOM leave ps/made
 Ali left home/the house hurriedly.
- b. ali khuna-ro ba ajale tark kard
 ali house-DOM hurriedly leave ps/made

both (39 a-b) are acceptable, and equally normal. However, given the fact that either of them is normal and acceptable, we may always adopt one order, and hypothesize on it. I would like to suggest, following Speas (1990) and Higginbotham (1985), that manner adverbs are VP-internal elements which enter into a theta-identification relation with the verb (VP) as the diagram (40) below shows (see also chapter I):

I will come home soon tomorrow.

- b. ali goft ke farda zud mi-y-ad khune
ali said that tomorrow soon ASP--prs/come-3S home
Ali promised that he would come home soon tomorrow.
- c. pedar ghol dad ke bara-m ye ketab be-khar-e
father promise gave that for-me a book Sbj-buy-3S
Father promised me to buy a book for me.
- d. bache-ha pedar-e-shun-o vadar kard-and
child-PL father-EZ-CL/their-DOM force ps/made-3PL
ke bara-shun ye saat be-khar-e
that for-CL/their a watch Sbj-buy-3S
The children made their father to buy them a watch.

As (42 d) clearly indicates the controlled embedded clause is finite and is equal to an infinitive in English. As these examples show the order of phrases in embedded and main clauses is the same, including specific, and non-specific generic DOs. Any phrase of a sentence may be topicalized by preposing and Chomsky-adjointing it to IP-initial position, (see chapter IV). However, the topicalized elements must preferably be accompanied by an appropriate intonation to avoid confusion.

- (43) a. ketab-o ali baraye man be madrase t; ?avord
book-DOM ali for I to school ps/brought
Ali brought the book for me to school.
- b. baraye man; ali t; ketab-o be madrase ?avord
Ali brought the book for me to school.
- c. (be)madrase ali ketab-o baraye man ?avord
Ali brought the book for me to school.
- d. baba bara-m ketab kharid
father for-CL/me book bought
Father bought me (some) book.
- e. ketab; baba bara-m t; kharid
- f. ketab bara pesar-at kharid-i ?
book for son-CL/your bought
Did you buy (some) book for your son?

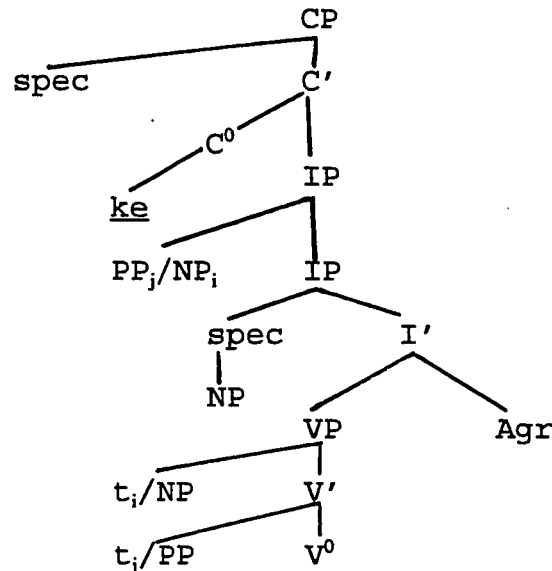
ketab "book" in (44 d) is a generic NP, it may only be moved to VP/IP initial position and/or stay away from the verb in order to be topicalized, focused, or express a discourse function, otherwise the sentence would be unacceptable. The other underlined phrases in (43) are all topicalized by preposing them to IP initial position.

The same is true for embedded clauses where the topicalized phrase may move to the initial position of the embedded or the main clause, as the following examples illustrate:

- (44) a. ali; ghol dad ke ketab_i-o baradar-esh_i bara-m t_j
 ali promise gave that book-DOM brother-CL/his for-CL/me
 be-ferest-e
 Sbj-prs/send-3S
 Ali promised that his brother would send the book for
 me.
- b. ali; ghol dad ke baraye man_i baradar-esh_i t_j
 ali promise gave that for I brother-CL/his
 ketab-o be-ferest-e
 book-DOM Sbj-prs/send-3S
 Ali promised that his brother would send the book for
me.
- c. ketab-o_i ali; ghol dad ke baradar-esh_i bara-m t_j
 be-ferest-e
- d. baraye-man_i ali; ghol dad ke baradar-esh_i t_j ketab-o
 be-ferest-e

The underlined elements in (44) above are topicalized. In (44 a-b) they have been moved/adjoined to IP initial position of the embedded clause, while in (44 c-d) they have been moved to the initial position of the main clause. In the first two examples the presence of an appropriate intonation with the initial elements is necessary. I assume that the same rule of preposing is responsible for topicalization in main and subordinate clauses, where an NP/PP is Chomsky-adjoined to the IP. The empty category left behind by preposing of the phrase is properly antecedent governed, hence the ECP is respected. However, in both cases of topicalization a proper intonation must be used with the preposed element. As Karimi (1989) also notices, topicalization does not seem to be a root phenomenon in Farsi. The rule we suggest for this process in Farsi is the one of Chomsky-adjunction which adjoins an XP to IP, in front of the clause, so that it is compatible with the featureless subordinator/complementizer ke "that" in the language. The fact that the XP may co-occur with ke for a discourse function, that is topicalization, contrastive topic, and question, proves that it does not occupy the COMP position, as (45) show:

(45)



Farsi displays unexpected re-orderings so that topicalization is also possible by re-ordering elements internal to IP together with an appropriate intonation, as we see in (46):

- (46) a. baba ketabⁱ bara-m tⁱ kharid-e
 father book for-CL/me bought-Prf/3S
 Father has bought (some) book for me.

ketab is a generic, non-specific NP which is topicalized by moving to a higher position adjoined to VP. For Karimi (1989) the above examples involve a VP-internal and/or S-internal topicalization which are formed by inversion of XPs. We have already discussed examples like (46 a) and those which involve specific NPs and claimed that they reflect V-internal adjunction as displayed in (27, section 1), or IP-adjunction as exemplified by (45) above. We conclude this section by agreeing with Karimi that topicalization in Farsi is not a root phenomenon, and claiming that the same rules of topicalization and re-orderings apply in both main and subordinate clauses.

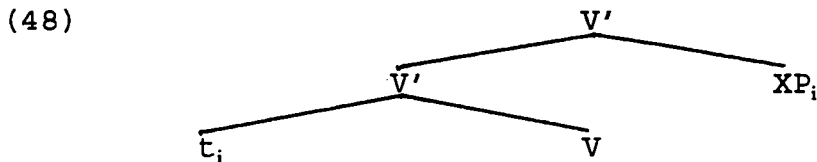
3-1 XP-POSTPOSING

The most important, and interesting phenomena of word order in Farsi is with regard to movement of phrasal arguments to post-verbal position. As we pointed out briefly in the first section of this chapter, it is quite normal and frequent to post-pose PPs of location/direction and IOs after the verb.

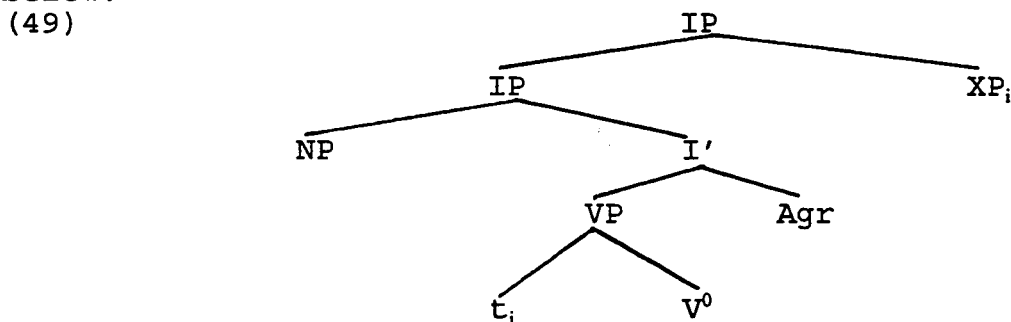
- (47) a. man t_i raft-am madrase_i
 I ps/went-1S school
 I went to school.

- b. pro t_i ketab-o ?avord-am bara- t_i
 book-DOM ps/brought-1S for-you
 I brought the book for you.
- c. in-o t_i khund-am bara dust-am;
 this-DOM ps/read-1S for friend-1S
 I read this for my friend.

As I already mentioned this process is highly frequent and used in both colloquial and formal style. This process is possible in both main and subordinate clauses. Our suggestion is that phrasal arguments are adjoined to IP, contrary to Karimi (1989) who neglects Infl, and suggests that the postverbal-XPs are adjoined to the verb. Karimi's proposal (1989, p 154) is that these constructions are instances of Chomsky-adjunction to V', as (48) below, corresponding to her (56) indicates:



She names this process a "local rule" which conforms to the "Structure-Preserving Constraint" of Emonds (1976, 1985): "That is, an XP is adjoined to a node in a position that is separated from its trace only by an X⁰ category (V⁰ in this case)". One point in order with (48) and the Structure-Preserving Constraint is that the post-posed phrases are actually adjoined to IP, and not to V', as they follow the Infl (Agr) in my and her examples. Another realization of the post-posed verbal-XPs is suggested in (49) below:



The only way to save her figure (48) of adjunction to VP is to argue that functional heads (Infl) do not count as barriers, as they really do not. Agr/IP adjoined positions count as A-bar positions. As a result t_i in (49) is properly antecedent governed by the post-posed XP. This is in fact what she properly argues for by adopting the **Structure Preserving Constraint** of Emonds (1976,1985) into her analysis. It is a general principle which restricts "the derived structures that transformations can

produce", (Emonds 1985: 139):

(50) = (134)

The Structure Preserving Constraint

In a major transformational movement, α can be phonologically realized adjoined to β only if the adjunction is structure preserving or if β is a root node.

A transformation is defined structure preserving if and only if " α and β are of the same category". In our structure (49) α , that is the preposed XP, and β , the adjoined position, are not for sure of the same category. Karimi (1989) describes ((48)=(her 55, P.154)) as a local rule which applies within the V'. In other words "an XP is adjoined to a node in a position that is separated from its trace only by an X⁰ category (V⁰ in this case)", Karimi (1989, P.154). While our figure (49) indicates that the adjunction is not to V' (=X⁰), rather to IP, the maximal projection of the Infl. Thus it seems to me that in (48) and (49) the adjunction is not a Structure Preserving one as her definition of Structure Preserving adopted from Emonds (1985) implies, rather adjunction is to a root node (=β).

An interesting proposal of Karimi's thesis which correlates the notion of specificity of NPs with the word order is that following Brame, Fukui & Speas she adopts a DP structure for all NPs in Farsi. She successfully argues that the Spec of DP is always occupied by specific lexical elements, like determiners, demonstratives, or with an abstract feature [specific] for definite and proper nouns in Farsi, English, and cross-linguistically. In case of non-specific DPs the spec-DP remains empty, leading to an empty node in the structure of any non-specific DP. This empty node as she argues is subject to ECP. Thus non-specific DO-NPs, subject DPs, adjectives, and predicate nominals which have an empty node in their spec-position may not appear in post verbal position (adjoined to IP) because the empty node of spec-DP remains ungoverned, thus violating the ECP. PPs (of location/direction and IOs) may freely appear adjoined to IP, since their governing head, that is the preposition, properly governs the empty node of the non-specific DPs under their projections. The following examples illustrate this interesting proposal. However, it is further interesting to note that the same/similar examples which are well-formed for us, (that is (51a)), according to her proposal and post-verbal IP-adjunctions in general, are bad and unacceptable for Hashemipour (1989),

(51) a. t_i umad un shaged-e,
 came that student-spc
 That student came.

b. * t_i umad shagerd,
 came student

The underlined DP in postverbal position in (51 b) is non-specific, while the one of (51 a) is specific. The traces of both DPs in spec-IP receive NOM case by spec-head agreement from the Agr. The chain containing the postposed DP and the preverbal subject-trace bears case, as a result the ungrammaticality of (51 b) can not be attributed to the lack of case, rather the empty node in spec-DP of the non-specific DP, shagerd, is not governed by the Infl/V, since these both govern to the left, thus the ECP is violated. As the spec-DP of the specific-DP in (51 a) is already filled and need not be governed, so it is well-formed. To conclude, I will present some more examples (adopted from Karimi) without any explanation, which contradict Hashemipour's contention that postposing of specific DPs is ungrammatical just as postposing of non-specific ones is.

(52) (Karimi 40 a, p.226)

- a. sepide mo?allem-e ma bud
Sepide teacher-EZ our was
Sepide was our teacher.
- b. sepide bud mo?llem-e ma
- c. * sepide bud mo?allem
Sepide was a teacher.

(53) (her 42, p. 227)

- a. sepide mariz shod-e
Sepide sick ppr/become
Sepide has become sick.
- b. sepide shod mariz

Even in case of IOs which may lack a preposition in quick colloquial speech, this rule applies:

(54) (her 44, p.228)

- a. ketab-o dad-am shagerd-e
book-DOM gave-1S student-spec
I gave the book to the student.
- b.* ketab-o dad-am shagerd

(55) a. man bara-t ketab kharid-am
I for-you book bought-1S
I bought book for you.

- b.* man bara-t kharid-am ketab

The underlined elements in the ungrammatical sentences are non-specific, and their empty nodes in the spec-DP remains ungoverned, violating ECP.

3-2 EXTRACTION TO MAIN CLAUSES

The last re-ordering phenomenon I will study here is preposing of XPs from embedded clauses to the main clauses, which we already briefly discussed in (44 c-d) section (3) of this chapter. The following examples illustrate this point:

- (56) a. (Hashemipour's 40, a.)
 ...gheremomken-e ke pedar-esh in ghaza-ro bedune
 impossible-is that father-his this food-DOM without

ab bo-khor-e
 water Sbj-eat-3S

It is impossible for his father to eat this food without water.

- b. (her 40, b.)
in ghaza-ro, ...gheremomken-e ke pedar-esh t_i bedune
 this food-DOM ...impossible-is that father-his without

ab bo-khor-e
 water Sbj-eat-3S

The subject of the main clause in (56) is a non-overt expletive equivalent to "it" in English, (lack of overt expletives is a property of pro-drop parameter in Farsi). gheremomken-e "is impossible" takes a CP complement to the left. The preposed XP, in ghaza-ro, does not count as the subject of the main predicate, and (56b) should not be considered as an instance of raising to subject, as we note from (57) below:

- (57) a. ...gheremomken-e ke bach-e-ha in ghaza-ro bedune
 impossible-is that child-pl this food-DOM without

ab bo-khor-e
 water Sbj-eat-3S

It is impossible for the children to eat this food without water.

- b. bach-e-ha, ...gheremomken-e ke t_i in ghaza-ro bedune
 -3S/is

ab bo-khor-e

The underlined subject of the embedded clause is plural in (57). In both (57 a-b) it agrees with the verb/Agr of the embedded clause, and not with gheremomken-e, the verb of the main one. This shows that bach-ha in (57 b) does not function as the subject of the main clause, rather it is adjoined to the IP of the main clause. It also indicates that there is no raising verb and/or adjective in Farsi,

assuming that gheremomken is the equivalent of the English raising adjectives.

Both subject and object of a subordinate clause may be preposed:

- (58) in ghaza-ro; bach-ha; ...gheremomken-e t; t; bedune ab
 this food-DOM child-PL ...impossible-is without water
- bo-hkor
 Sbj-eat-3rd/S

However to indicate a clear topicalized reading, the preposed phrases are to be accompanied by appropriate intonation. It is possible to embed (58) furthermore as (59) shows:

- (59) ali goft (ke) in ghaza ro; bach-ha; ...gheremomken-e (ke)
 ali said that
- t; t; bedune ab bo-khor-and

(59) confirms our suggestions that the preposed, (underlined elements) are in adjoined position. They may express a discourse function if accompanied by an appropriate intonation. Now, notice the following examples, taken from Hashemipour (1989):

- (60) a. (her 129, p 173)
 (man) omidvar-am [_{CP} ke hossein-o ali] be shome
 (I) hope-1S that Hossein-and Ali to you
- komak=kon-and
 help=make-3PL
- I hope that H & A may help you.
- b. hossein-o ali; (man) omidvar-am [ke t; be shoma
 komak=kon-and]

The embedded subject has been preposed to IP adjoined position of the main clause. As expected it may not appear after the subject of the main clause in (60). Both the subject and IO=PP of (60 a) may be preposed, as (61) below shows:

- (61) be shoma; H & A; (man) omidvar-am t; t; komak=kon-and

This ends our study of re-ordering in main and subordinate clauses, which I believe included the important instances of the almost free word order in main and subordinate clauses. Being an SOV language, all complements and adjuncts of the verb and clause fall under the canonical government projection of the verb and Agr, Kayne (1983), thus re-ordering of the elements is possible,

specifically in pre-verbal position, since the empty categories left behind by these re-orderings are always under the canonical government projection of the V^0/Agr^0 , hence the ECP is respected. That is why SOV languages exhibit rich possibilities of re-ordering and scrambling, Kayne (1983), Baker (class lectures). We classified the re-orderings in Farsi either as IP-initial or IP-final adjunction, based on their distributional, functional, semantic properties.

FOOTNOTES OF CHAPTER II

1- Karimi (1989) argues that in a sentence like (16), repeated as (i) below, the clitic esh is governed and case-marked by the preposition be "to" in its minimal government-projection. It, esh, is also co-referential with the NP Sepide in the A-bar (topic) position:

- (i) sepide-ro, be-h-esh goft-am
 Sepide-Ra, to - her told-I
 As for Sepide, I told her.

In order for the clitic pronoun esh to be appropriately A-bar bound by the antecedent, Sepide, in the A-bar topic position, she proposes the Extended Locality Principle (ELP) as below:

- (ii) =(her 73)

EXTENDED LOCALITY PRINCIPLE (ELP)

(Where X=A or A-bar)

An X-anaphor α can not be free in [β ... α ...] where β is the minimal government projection containing

- (i) α
 (ii) the governor of α
 (iii) and a SUBJECT accessible to α

Given the ELP in (ii), S' is the minimal government projection for esh. S' in (i) includes the governor of esh, that is the preposition be "to" ($=\alpha$), an accessible SUBJECT, that is AGR ($=Infl$). Since the NP sepide is also in the same government-projection, it can, therefore, be co-indexed with the clitic pronoun esh. As Sepide and esh are co-indexed, they may share and derive properties from each other only if one lacks the properties in question. Thus an NP can share (inherit) case or specificity from the other only if it does not have them itself, that is, "if it is in an A-bar topic position" as Karimi puts it herself. The NP Sepide is in an A-bar topic position and is co-indexed with esh by ELP. It lacks case property. It inherits case from esh which receives oblique case from the preposition be "to". Being a specific noun phrase, Sepide has to carry Ra by SPECIFIC OBLIQUE CASE (15), but esh does not. The same reasoning applies to (11 & 12) in the text.

The other analysis for the oblique case (#NOM) of the NPs followed by Ra in the topic or A-bar position might be to attribute their case to a sort of default case. Default, here has to be any case except NOM in Farsi, that is, +oblique (#NOM).

CHAPTER III

PHRASAL STRUCTURE

In the previous chapters we argued that the surface order of arguments within a clause, and phrase is not determined by principles of X-bar theory, rather case and theta theory, and direction of case and theta-role assignment, and specificity of NPs are relevant and determinant issues in the order of phrases within a clause/phrase in Farsi. From our discussions so far, we know that verbs in Farsi appear at the end of a clause and AUXs and Infl (=TP & Agr P) must obligatorily follow the verb. There are no real modals in the language, but there are some verbal items which function as defective modals, and appear to the left of the lexical verbs in surface structure. We will study them in the next chapter. Thus both VP and Infl (clause) are head final and must govern and case and theta mark their arguments to the left, while nouns, prepositions, and adjectives are head initial and govern their arguments to the right. In order to form acceptable syntactic structures, items which combine together must obey the theta criterion in terms of discharge of theta positions as explained in chapter I. In order for a position to discharge and enter into a theta relation with another position it must be visible. We adopt the visibility condition on theta role assignment as below:

- (1) A position is visible in a chain if the chain contains a case marked position.

Case theory imposes the case filter. Domain of case assignment is defined by Stowell (1981) as (2) below:

(2) CASE ASSIGNMENT UNDER GOVERNMENT

In the configuration [α β ...] or [... β α] α case marks β , where

- (I) α governs β and
 (II) α is adjacent to β , and
 (III) α is [-N]

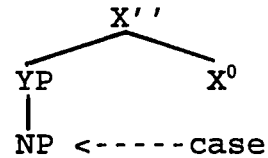
In Farsi Infl is not adjacent to the subject position. In fact they are at two opposite ends of the clause. Furthermore the DO is not always adjacent to the verb. Generic-nonspecific bare DO-NPs are obligatorily adjacent to the verb, other DOs may freely move within the VP, as well as out of the IP under certain conditions as explicated in chapter II. Rouveret and Vergnaud (1980) propose that case can be assigned under two configurations: (a) Government, (b) Agreement configuration, as (3) below:

(3) - a. Case assignment under government

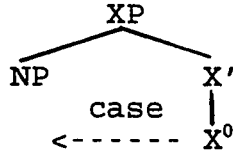
(i)



(ii)



b. Case assignment under agreement



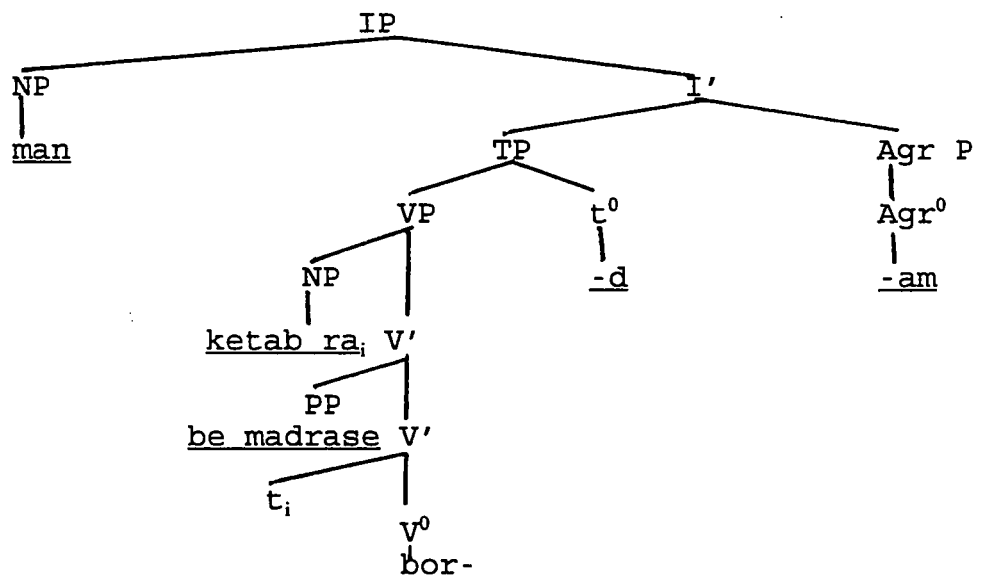
Assignment of ACC case in Farsi is under 3 (a). Given the unmarked word order for sentences containing DO-NPs, we adopted in chapter I, repeated here for convenience:

(4) S IO(=PP) PP DO V AUX Agr

The DOs, specific or non-specific, comply to the definition of case assignment of Stowell in (2) above, and the configuration (3 a), so that ACC case is assigned to them under adjacency, and government by the assigning head. At the same time the theoretical framework we suggested for theta role assignment in chapter I, that is **Theta Marking** by a predicate to one of its arguments, is also satisfied. Thus a sentence like (5) will have the tree diagram as (6):

(5) man ketab ra be madrase bord-am
 I the book DOM to school ps-took-3S
 I took the book to school.

(6)



However, it may either be realized morphologically as in Finnish, or in the interpretation as in Turkish and Farsi. Enc (1991) further suggests that the relevant semantic notion in examples (7) and (8) is not a "Definiteness Effect" as Belletti claims, but it is "Specificity". We follow Enc as the relevant semantic notions in Farsi too, as well as Turkish show a "Specificity" effect.

Nominative case assignment in Farsi does not respect adjacency as configuration (6) shows. I assume that NOM in Farsi is assigned both under spec-head agreement (3 b) and government¹ (3 a(ii)), see Sportiche (1990).

Prepositions immediately precede their arguments, and I assume that they assign oblique case. IOs are always preceded by prepositions. In very rare cases when it might not be present in colloquial speech or in verb postposed positions, the meaning of a preposition is always implied, and we might assume that it has dropped in the course of derivation to PF, as the following example adopted from Karimi (1989) indicates:

- (9) a. ketab-o dad-am shagerd-e
 the book DOM ps-3S student spec(ific)
 I gave the book to the student.

(9) is much better when we add a preposition be "to" before "student". In many cases the prepositions before VP-internal adjuncts (that is place and direction ones) drop. Again the meaning of a preposition is implied, and the sentences seem much better when the prepositions are added:

- (10) ketab-o bord-am madrase
 the book ps-1S school
 I took the book to school.

There seems to be two groups of prepositions in the language. We will study their properties in more detail in the next section.

There are many bare-NP location and time adverbials in the language as (11) indicates:

- (10) a. inja "here"
 b. anja "there"
 c. diruz "yesterday"
 d. anruz "that day"
 e. anvght "at that time"
 f. hafte-ye pish "last week"

Some of these may actually be used with a few of the prepositions, but most of the times they may be used alone, without any preceding preposition.

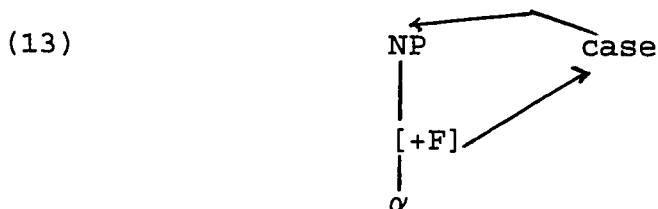
- (11) ali unja raft
 ali there ps-3S "Ali went there".

There are also many compound noun phrases which indicate time and location and may be used without a preceding preposition. These are mainly NPs and adverbials indicating time or location related to another NP by a dummy case assigner similar to of in English, called Ezafe, (EZ henceforth). It is a morpheme -e or its morphologically conditioned equivalent after the vowels, -ye which follows the head nouns. Notice the following examples:

- (12) a. shab-e pish
night-EZ before "last night"
- b. in taraf-e hayat
this side-EZ yard "this side of the court"

For their case marking I adopt Larson (1985) who convincingly argues that:

"...case-assignment in bare-NP adverbs occurs through a special feature, [+F], which is borne by these nouns. This feature is inherited by any NP having such an N as its head, and it assigns an oblique case to the NP it labels":



Thus he suggests that bare NP adverbs are inherently case-marked by a [+F] feature.

2-SENTENTIAL COMPLEMENTS

Finite clauses in Farsi can not appear in subject, direct object position, as well as after preposition. We attribute this property to CRC of Stowell (1981), (14) below, which requires that case may not be assigned to a category bearing a case assigning feature. Infl heads a finite clause and assigns NOM case, consequently a finite clause bearing a case assigning feature may not appear in a case position. As Farsi is verb final and structural and/or inherent case to DOs and/or bare DO-NPs and NOM case to the subject position are assigned leftward by the respective case assigners, we expect that they may not appear in any case positions to the left of the verb as (15) below indicates:

(14) CASE RESISTANCE CONDITION (CRC)

Case may not be assigned to a category bearing a case assigning feature

(15 a. ali goft [_{CP} ke pro ketab-e-sh-o bara-m mi-y-ar-e]
 ali told that pro book-EZ-his-DOM for-me brings
 Ali told that he would bring his book for me.

b.* ali [_{CP} ke pro ketab-e-sh-o bara-m mi-y-ar-e] goft

[-N] categories bear a case assigning feature, thus they may not be assigned case. We will study the structure and distribution of infinitives in Farsi in section (4). We will notice that infinitives in Farsi are NPs with hybrid functions both as a (1) tensed NP, and (2) as a non-tensed one. In function (1) they may appear only in [+NOM] case positions, but as (2) they appear in all case positions.

3-OBJECT/MODIFIERS OF NOMINALS

As we argued before, [+N] categories, adjectives and nouns, are not case assigners. The question is what happens to the subcategorized XPs (NPs and PPs) of derived nominals and adjectives in Farsi. How do these elements pass the case filter? Notice the following examples:

(16) a. doshman shahr ra nabud kard
 the enemy the city DOM destroy(N) made
 The enemy destroyed the city.

b. nabudi-y-e shahr
 destruction-EZ city
 the destruction of the city

In (16 a) nabud kard "destroyed" subcategorizes for the NP shahr "city". In (16 b) the NP derived from the compound verb has the same subcategorization frame. There is an EZAFE morpheme which is a dummy case marker in the language similar to of in English, de in French, and di in Italian. EZ is not a preposition. It is a bound affixal (clitic) morpheme which attaches to a head noun and adjective. I believe that [+N] categories do have case to assign, but unless an EZ is affixed to them they can not transfer the case to their NP-complements, to the right. Notice also the examples below:

(17) a. nabudi-ye nagahani-ye shahr
 destruction-EZ sudden-EZ city
 The sudden destruction of the city.

b. nabud kardan-e nagahani-ye shahr
 destroy to make-EZ sudden-EZ city
 to destroy the city suddenly

Nagahani "sudden(ly)" is an adjective/adverb in (17) above. It functions as an adjunct which has intervened between the head and the complement in (17), thus justifying our PROJECT ALPHA in

chapter I. The EZ following the adjunct is reiterated in order to transfer the case of the head to its NP-complement. Whatever number of adjuncts that intervene between the head and the complement, EZ must still reiterate. In English, where of precedes the complement, and is not attached to the head as in Farsi, it does not reiterate, rather it always precedes the complement and the intervening adjuncts do not carry of.

Now consider (18) where a head predicate subcategorizes for a PP:

- (18) a. amir [_{pp} az doshman] mi-tars-ad
 amir from the enemy fears
 Amir fears the enemy.
- b. tars [az doshman]
 fear from the enemy
 fear from the enemy
- c. tars-e [az doshman]
 fear-EZ from the enemy
 fear from the enemy
- d. tarsidan [az dashman]
 to fear from the enemy
 to fear from the enemy
- e. tarsidan-e [az doshman]
 to fear-EZ from the enemy

The nominal counterparts are possible both with and without EZ, according to my intuition, but many people and linguists may consider the ones with EZ ungrammatical. We may account for (18) by assuming that the PP complements have a case assigning feature, that is the P, and by CRC they resist any case from other assigners. They also show that EZ, itself, is not a real case assigner in these examples, but only transfers the case of the head to its complements.

A similar argument holds for the NP-complements of adjectival predicates, as (19) below indicates:

- (19) a. u be bache-ha alagemand ast
 he to children-pl interested is
 He is interested in children.
- b. alagemand be bache-ha
 interested to children-pl
 interested in children
- c. alagemand-e *(be) bache-ha
 interested-EZ *(to) children-pl
 interested in children

- (20) a. ali [az kar-e ziyad] khaste ast
 ali from work-EZ too much tired is
 Ali is tired of too much work.
- b. khaste *(az) [kar-e ziyad]
 tired *(from) work-EZ too much
 tired of too much work

alaghemand subcategorizes for a PP complement. In (19 a) it is used as a predicate but in (19 b-c) as a head. In both cases the sentences without the preposition be is ungrammatical, as indicated by (19 c). (19 b-c) also indicate that adjectives may or may not be followed by an EZ. As we claimed EZ is not a case assigner, thus it causes no problem for us by CRC. (19 C) with EZ is, however, controversial, and many including Samiian (1983) do not accept it. It is good with EZ for me and I consider it as a stylistic variant. Similarly in (20 a) the predicate adjective subcategorizes for a PP-complement, its adjectival phrase counterpart must select a PP as well. Now consider (21):

- (21) U [_{AP} nagaran-e [_{NP} bache-ha]] bud
 he worried-EZ children-pl was
 He was worried for children.

(21) shows that the adjective nagaran "worried" takes an NP-complement and an EZ has intervened in order to take care of the case of the NP. We conclude that adjectives may select a PP or NP complements to their right. In the latter case an EZ must follow the adjective, but the range of adjectives taking NP-complements is very limited, and most of their complements tend to be PPs.

3-1-EZAFE CONSTRUCTIONS

Ezafe is also a sign of real genitive case, as indicated by examples below:

- (22) a. ketab-e ali
 book-EZ ali "Ali's book"
- b. ketab-e u
 book-EZ he "his book"
- c. [_{NP} dust-e [baradar-e man]]
 friend-EZ brother-EZ I
 my brother's friend
- d. ketab-e-sh
 book-EZ-his (clitic)
 his book

(22) indicates a possession relation between the two nouns. Thus we

conclude that in these examples EZ functions as a genitive case assigner. (22 c) indicates two successive cases of genitive by EZ.

3-2 ATTRIBUTIVE ADJECTIVES

Ezafé also intervenes between a noun and its following attributive adjectives as indicated by the examples below:

- (23) a. ketab-e mofid
book-EZ useful "a useful book"
- b. ketab-e kheyli mofid
book-EZ very useful "a very useful book"
- c. ketab-e [kheyli mofid-o khub]
book-EZ [very useful-and good]
A very useful and good book.

In (23) EZ can not be clearly considered as a case assigner. I assume that EZ in (23) indicates the thematic and/or the modificational relation between the modified head noun and the following adjectives. Similar modificational relations between nouns and adjectives is also indicated by agreement or case relation between nouns and adjectives in French, German, Russian, and agreement between prenominal adjectives and nouns in Dutch, as presented in Stowell (1981), and Riemsdijk (1990).

This view is more justified when we consider superlative modificational adjectives that must obligatorily precede the head nouns, and do not take EZ in the pre-head position, as (24) below shows:

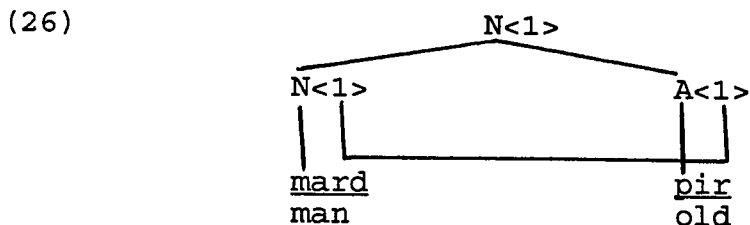
- (24) a. kuchik-tarin otagh
small-spr room
The smallest room.
- b. * kuchik-tarin-e otagh
small-spr-EZ room
The smallest room.
- c. * otagh-e kuchik-tarin
room-EZ small-spr
- d. otagh-e kuchik-tar
EZ small-cmp
a smaller room.

In (24 a) the superlative (spr) adjective modifies the head noun, but no EZ is allowed, as noticed by ungrammaticality of (24 b-c). But the comparative adjective which follows the head noun does require EZ. This observation, I assume indicates that EZ before modificational APs (NPs, and PPs) does not assign case. Pre-head position in non-verbal categories in Farsi is the specifier

position where, a closed set of elements appear and normally indicates a tighter and closer relation, similar to compounding, as a consequence no phonetically realized agreement relations is required. But the superlative adjective in (24 a) can not be counted as a specifier of the head noun, since it indicates the same modificational relation with the noun as the comparative and simple adjectival equivalents, as the following examples also may show:

- (25) a. pir mard
 old man "an old man"
- b. mard-e pir
 man-EZ old "a man who is old"
- c. javan-mard
 young man "a kind, and very helpful man"
- d. mard-e javan
 man-EZ young "a man who is young"

We tentatively assume that a noun enters into a kind of thematic relation with its modifiers, similar to theta identification as discussed by Higginbotham (1985). That is the modifiee is theta identified by the modifier. This relation is realized in the surface structure by EZ construction in post-head positions. In theta-identification the theta position in the grid of the modifier is identified with the grid position of the modifiee. In this type only the grid of the modifier is discharged, and the two theta positions are merged into one, but undischarged position, headed by the grid position of the modifiee which is still available to be discharged as (26) shows:



This type of relation is expressed at s-structure by an EZ in Farsi and some similar means in French, German, and Russian, etc.

3-3 ATTRIBUTIVE NOUN PHRASES (NP_s)

There is another kind of noun modifiers, apart from APs which occur in Ezafe construction as (27) below:

- (27) a. Kif-e charm
 purse-EZ leather "a leather purse"

- b. kif-e pul
 purse-EZ money "money purse"

the two modifiers of kif in (a) and (b) are incompatible as the example below:

- (27) c. * kif-e charm-e pul
 1 EZ 2 Ez 3
 d. * 1 EZ 3 EZ 2

Consider also the following examples:

- (28) a. saat-e tala
 watch-EZ gold "gold watch"
 b. roghan-e cheragh
 oil-EZ lamp "lamp oil"

These NP modifiers do appear with APs and must precede them, as:

- (29) a. kif-e charm-e bozorg(-e [ru-ye miz])-e ali
 purse-EZ leather-EZ big (EZ [on-EZ table])EZ ali
 Ali's big, leather purse which is on the table
 b. mo?allem-e khub-e riyazi
 teacher-EZ good-EZ mathematic
 The good math teacher

As Samiiian (1983) defines them, they are "limited semantically to the material of which objects are made (leather, plastic, brick,...) or the function performed by objects or individuals". These nouns are head final and can not be further expanded to take a modifier or specifier. They may make compounds with the head nouns which they modify, and as a result drop the EZ:

- (30) a. mo?allem-e riyazi
 teacher-EZ mathematics "the math teacher/the teacher of
 mathematics"
 b. mo?allem-riyazi "a math teacher"
 c. roghan-e cheragh
 oil-EZ lamp "the lamp oil/the oil for lamp"
 d. roghan-cheragh "lamp-oil"

Samiiian compares NP_i with compound NPs (NP_c) in Japanese and Farsi and shows their syntactic similarities. She assigns them identical positions in the X-bar schema, so that they are incompatible with each other, as (31) below indicates:

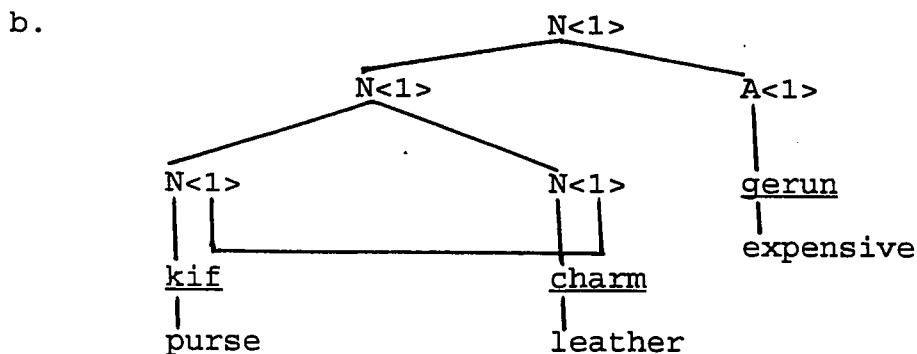
- (31) a. kif-e charm-e hassan (np_a)
 b. kif-e pul-e hassan (np_c)
 c. * kif-e charm-e pul-e hassan (np_c)

A similar points holds for English as well:

- (32) a. an old leather purse
 b. a leather old purse

From these points we conclude that NP_a function similarly to NP_c both in Farsi and Japanese. They together with APs modify the head nouns and the EZ preceding them is not a case assigner, but is added there by a transformation in the course of derivation to S-structure. We assume that NP_{a=c}, too, enter into a theta identification with the head nouns as (33) below:

- (33) a. kif-e charm-e gerun
 purse-EZ leather-EZ expensive
 an expensive leather purse



The top N<1> still has an open position to discharge. I assume that +N categories must be related to their complements and adjuncts to the right with some kind of obligatorily agreement or case relation in the language.

3-4 PREPOSITIONAL MODIFIERS (PP)

EZ must also appear between a head noun and a prepositional modifier as (34) below indicates:

- (34) a. [ketab-e [pp ru-ye miz]]
 book-EZ on-EZ table
 the book (which is) on the table
- b. [otagh-e [pp zir-e shirvani]]
 room-EZ under-EZ roof
 the room under the roof

The occurrence of second EZ inside the PP is irrelevant at the

moment. The ru-ye miz, and zir-e shirvani are claimed by Samiian (1983) to be PPs. We adopt her view at the moment, but we will return to it in the next section and argue that they are not in fact PPs, rather they are compound NPs that denote place and direction, but we will continue to refer to them as PPs. I suggest that the relation between the PPs and the head noun is similar to that of NP_a and APs. The PPs in (34) in fact modify the head nouns as APs and NP_a do, and the first EZ between the head noun and the PP is in fact a dummy case assigner as explicated for APs and NP_a. This should not count against our yet-to-come generalization that PPs do not appear in case position, as EZ in the above case does not really assign case, and more importantly we will argue that these PPs (according to Samiian (1983)) are not really PPs, rather compound NPs. Now let us consider (35) below:

- (35) [otagh-e [chubi-ye-] [kuchik-e] [_{pp} zir-e shirvani-ye]]
 room-EZ wooden-EZ small-EZ under-EZ roof-EZ
 1 2 3 4
- [ali]]]]
 ali
 5

ALI's small, wooden, under the roof room

The EZ in the PP (4) is irrelevant at the moment. NP (1) is the head noun, NP_a (2), AP (3), and PP (4) are modificational phrases. They are all followed by EZ. From the semantic interpretation I understand that they have a modificational relationship to the head noun, of the type of theta identification, and EZ is not a real case assigner, but a Dummy one. But the last NP (5) does receive a genitive case from the EZ as there is a possession interpretation in the relation between (1) and (5). Now let us study (36) below which contains a real PP:

- (36) a. [tars (*-e) [az doshman]]
 fear (EZ) from enemy
 fear of the enemy

Samiian (1983) consider the EZ before the PP in (36) ungrammatical. For me it is possible to have the EZ in (36) above as a stylistic variant, as was said before, while I still think that the phrase is better without EZ. It is mainly used in quick colloquial speech. We may consider its appearance as an optional stylistic variant. The relation between the head and the PP complement is not one of modification, while the relation between the PP (4) in the example (35) and its head is. Furthermore the obligatory appearance of EZ before the PP (4) in (35) indicates that it is not a real preposition bearing a case assigning feature. If it were a real PP, it would not appear after EZ, similar to (36). As we notice in (35) EZ reiterates after each of the modificational phrases, while it

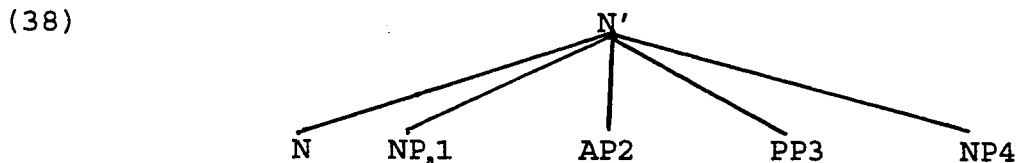
can not do so in (37) below, where an adjunct precedes the PP-complement in the example (36):

- (37) [tars (-e) [ziyad (*-e) [az doshman]]]
 fear (EZ) too much (EZ) from enemy
 too much fear from the enemy

EZ does not reiterate before the PP-complement as noticed by the ungrammaticality of (37), while it did in example (35) before the PP modifier. This clearly indicates that there must be some kind of difference between these two types of PPs. We will study them in more detail in the next section and show that the modifier PPs are not in fact PPs in Farsi. The first EZ in (37) indicates the modificational relation between the head noun, tars, and the following adjective, ziyad. Its appearance in (37) between the head and the PP-complement also indicates that contrary to Jackendoff (1977) complements do not stay at a lower X-bar level than adjuncts in Farsi. In fact all of the modificational phrases studied above (that is (2, 3, 4) in the example 35) must appear before NP/PP-complements of a head noun in Farsi, which is in agreement with the PROJECT ALPHA system we adopted in the first chapter. IF we look back at examples (17 a-b) in section (3) with a completely similar structure, but an NP complement instead of a PP one, we see that EZ must reiterate as noticed by the second occurrence of EZ before the NP-complement.

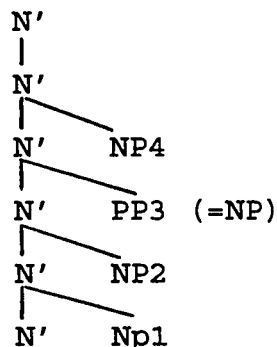
These observations about the EZ in Farsi indicate, I believe, that neither the dummy EZ before the complements nor the modificational EZ are real case assigners, but the genitive case EZ is. It also indicates that adjectives function similarly to nouns in Farsi, in that they may appear after EZ. In fact adjective are used as nouns in many instances and can be pluralized if used as nouns. I conclude that neither of+NP in English nor EZ+ NP(AP) in Farsi are underlying, D-structure elements. The elements after EZ and of are underlying ones and EZ and of are only added in the course of derivation to S-structure in order to license their surface-occurrence, possibly due to cross-linguistic property of nouns and adjectives in not being able to assign case to their complements, and enter into a modificational relation with their modifiers without such an element.

Samiian (1983) assigns the following tree diagram to NPs in Farsi where EZ construction also appears:



The subcategorized complements appear after NP4. Given our PROJECT ALPHA system we may draw the following tree diagram for the items that appear in the EZ construction of the NP:

(39)



The three functions of the EZ must be separated in the language. I suggest the following functions for EZ in the language, as (40) below indicates:

(40) The three functions of EZ:

(a) EZ as a dummy case assigner intervenes between derived nominals gerunds, infinitives and their NP complements

(b) Ez as a genitive case assigner intervenes between two NPs which have a possession relation with each other.

(c) EZ as a sign of agreement or thematic relation appears between a head noun and its NP_i, AP, and PP(=NP) modifiers.

3-4-1 PREPOSITIONAL MODIFIERS REVISED

Prepositions deserve more consideration as there seems to be two types of them in the language, as we notice from the examples below:

(41) a. ketab-e ru-ye miz
book-EZ on-EZ desk
The book on the table.

b. ketab ru-ye miz ast
book on-EZ desk is
The book is on the table.

(42) a. ketab ra az ali be-gir
book DOM from ali imp-prs
Take the book from Ali.

b. ketab ra az ru-ye miz bardar
book DOM from on-EZ desk imp-take
Take the book from on the table.

We assumed ru to be a preposition in the previous section, as also claimed so by Samiian (1983) and Hashemipour (1989). Samiian considers them as prepositions that take an EZ before the following

noun phrase (henceforth group-1 P) in contrast with prepositions as az, "from", that do not take an EZ (henceforth group-2 P). Hashemipour (1989) considers the former as prepositions that do not assign case and thus must take an EZ in order for case to be assigned to the NPs that follows them, in contrast with the latter that do assign oblique case to their NP-complements. The list of group-1 Ps is as (43) below, adopted completely from Samiian (1983):

- (43)
- | | | |
|----|----------------------|-----------------------|
| a. | bedun-e | "without" |
| | EZ | |
| b. | bara-e | "for" |
| c. | ba vojud-e | "in spite of" |
| d. | zir-e | "under" |
| e. | kenar-e | "next to" |
| f. | posht-e | "behind" |
| g. | pain-e | "below" |
| h. | nazdik-e | "close to" |
| i. | birun-e | "outside of" |
| j. | dakhel-e | "inside of" |
| k. | karej-e | "outside" |
| l. | miun-e | "in the middle of" |
| m. | vasat-e | "in between" |
| n. | atraf-e | "around" |
| o. | <u>in/up</u> taraf-e | "this / that side of" |
| p. | sar-e | "at" |

(44) below is adopted from Samiian:

- (44)
- | | | | |
|-----|------------|-------|--------------------|
| a. | zir-e | miz | |
| | under-EZ | desk | "under the desk" |
| b.* | zir | miz | |
| c. | birun-e | otagh | |
| | outside-EZ | room | "outside the room" |
| d.* | birun | otagh | |

In (43 a) bedun has no grammatical meaning. It is not a bound morpheme and is only used in the form bedun-e "without". In (43 b) bara is mostly used in colloquial speech. I will argue that both of them belong to group-2 Ps. In (43 c) ba in ba vojud-e is a group-2 preposition. vojud is a noun. All the other examples in (43 d-p)² are bare-NP adverbials of location, as my examples below show:

- (45)
- | | | | |
|----|--------|----------------------|-------------------|
| a. | boro | <u>pain</u> | |
| | imp-go | <u>down</u> | "go down" |
| b. | biya | <u>nazdik</u> (-tar) | |
| | come | <u>close(r)</u> | |
| c. | ma | <u>in taraf</u> | mi-rim |
| | we | this side | go |
| | | | "we go this side" |

Group-2 Ps do not take an EZ before their NP-complements, and are the true prepositions according to my analysis below, and can be characterized as [-N,-V], assigning the oblique case in Farsi. Their list follows below, adopted again from Samiian:

- (46) a. dar "in, at"
 b. ta "until"
 c. be "to"
 d. az "from"
 e. ba "with"
 f. bi "without"

- (47) a. dar in otagh
 in this room

- b. * dar-e in otagh
 EZ

(b) is all right if we interpret it as "the door of this room".

- c. az hasan
 from hasan

- d. * az-e hasan

Samiian also distinguishes a third type of preposition which optionally take an EZ before their following NPs, as (48) below:

- (48) a. tu (-ye) "inside"
 b. ru (-ye) "on"
 c. jelo (-ye) "in front of"
 d. bala (-ye) "on the top of"
 e. pahlu (-ye) "next to"

- (49) a. tu-ye otagh
 in-EZ room "in the room"

- b. tu otagh "in the room"

Similar to group-1 Ps all the Ps in (48) (excluding the EZ) are bare NP adverbials of location and direction, as (50) below indicates:

- (50) a. biya bala
 come up
 b. boro tu
 go in "go inside"

They do not make compound verbs as the English translations might imply. Samiian (1983) attributes the optionality of EZ in (48 & 49) to phonological motivations, as all of them end in a vowel. I agree with her who counts (48) as a subset of (43), that is group-1 Ps. I would like to add further that neither are true prepositions.

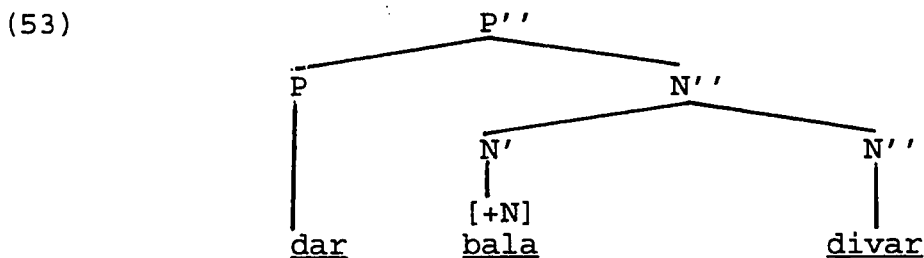
I will refer to them as group-1 Ps henceforth. I will argue that in fact they make a compound NP adverbial of direction or location and do not count as PPs, as Samiian herself inevitably characterizes them as [+P, +N]. However, her discussion in chapter II (PP 75-76) seems to me confusing as she mixes up reference to the three types of Ps. Samiian cleverly discovers the difference in distribution and function of the two types of prepositions, and refers to those which take an EZ as [+N] prepositions, but working in the descriptive framework of Jackendoff (1977), she is unable to give an explanatory account for EZ constructions and prepositions, as she herself admits. She correctly gives the following expansion rule for PPs in Farsi:

(51) P' -----> P- (N')

which should be incorrect if we take the following examples, (52) into account. They indicate that only group-2 Ps can take group-1 prepositional phrases as complements, and never the reverse, (examples adopted from Samiian (1983)):

- (52) a. [az [zir-e miz]] umad birun
 from under-EZ desk came out
 He came out from under the table.
- b. [ta [tu-ye otagh]] ovord-am-esh
 until in-EZ room brought-him/it
 I brought him up to the room.
- c. [bara-ye [tu-ye otagh]] kharidam-esh
 for-EZ inside-EZ room bought-it
 I bought it for inside the room.

(52 c) is the most interesting one, as bara-ye is among her group-1 Ps, but yet functions with group-2 Ps in taking a group-1 PPs as complement. The group-2 Ps as she correctly argues strictly subcategorize for an NP-complement, but in (52) above according to her analysis they have taken a PP-complement. First of all they can not be considered as compound Ps. Had the group-1 Ps in the example (52) been real prepositions, there would be a case clash, because case would be assigned to a phrase containing a case-assigning feature, thus violating CRC. She assigns the unnatural feature [+N, -P] to the group-1 Ps as (53) below, (her 35, P. 282):



The EZ is added by a transformation in the course of derivation. The above examples cast doubt on her assumption that bare NP adverb + EZ + NPs count as PPs. They also indicate that bara-ye and bedun-e belong to the second group of prepositions. As far as bara-ye is concerned, she considers it as a group-1 P which must take an EZ. She gives the following examples, ((54 a-b) (her 14 & 15 P. 273)) to prove her claim:

- (54) a. bara-ye hasan
for-EZ hasan "For Hasan"
- b.* bara hasan

She considers (54 b) ungrammatical, while it is not (see Karimi (1989) and Hashemipour (1989) for abundance of examples similar to (54 b)). It is completely normal, unmarked, and acceptable. Notice the following examples:

- (55) a. bara(-ye) to
for-EZ you "for you"
- b. bara-t
for-you (cl) "for you"
- c. bara-y-at
for- -you (cl)
- d. bara(-ye) u
for-EZ him "for him"
- e. bara-sh
for-him (cl) "for him"
- f. bara-y-ash
for- -him (cl) "for him"

All the above examples indicate that bara and bara-ye are two stylistic variants, and baraye does not consist of bara + ye. bedun-e is another preposition of the group-1 Ps that she has problem with. She notices that bedun, bara (and sar-e, according to me) are not adverbial prepositions (our bare-NP adverbials). There is no morpheme as bedun in Farsi. There is no reason to consider bedun-e as P+EZ, rather bedune is a monomorphemic preposition, belonging to group-2 Ps. Notice the examples below:

- (56) a. bedune tu
without you
- b. pro bedune dust-am ne-mir-am
pro without friend-my neg-prs-1S
I will not go without my friend.

Both bedune and baraye (=bara) may take animate NP-complements which denote the goal or source of action and function as IOs, while other group-1 Ps only denote location and direction. Thus I list these two prepositions among our group-2 Ps which assign oblique case. They do not show any nominal characteristics of group-1 Ps, as discussed above. Samiian (1983) notices that group-1 Ps also display certain syntactic/morphological characteristics of nominals. "They can be pluralized or take a determiner in some cases", Samiian (1983).

- (57) a. un [zir-a-ye sandali]
 that under-pl-EZ chair
 well under that chair
- b. un [tu-ha-ye jabe]
 that inside-pl-EZ box
 completely inside the box

As we notice the prepositional phrases are preceded by un, a determiner, and the prepositions, zir and tu, are pluralized which are assumed to be characteristic of nominals. In her group-1 Ps she wrongly considers un/in taraf-e "that/this side of" as a preposition, while in fact taraf "side" is a bare-NP and un/in "that /this" are determiners. Notice the following examples:

- (58) a. [DP in [NP taraf-a-ye shoma]] che khabar ?
 this side-pl-EZ you what news
 What is the news in your area?
- b. [PP az [DP in [NP taraf-e kheyabun]]] boro un taraf
 from this side-EZ street go that taraf
 Go to that side of the street from this side.

(58) clearly shows that the in/un taraf function as NP-adverbials of location which display some characteristics of nominals, that is pluralization, taking modifiers, and appearing in case marked positions, oblique case in (58 b). But Samiian based on these observations characterizes the group-1 Ps, that is P+ EZ+ (NP), as [+P, +N]; and group-2 Ps, that is P+ NP, as [+P, -N]. Even with these feature properties, she can not account for baraye and bedune, which function as group-2 Ps, but are listed in her group-1 Ps. We have argued that they are in fact group-2 Ps, and not group-1 Ps, and ye/e is not an EZ in them. There are cases of compound prepositions in the language, which indicate that examples like (50) can not be considered as compound prepositions, as (59) below:

- (59) a. ba?d az nahar
 after from lunch "after lunch time"
- b. pas az in
 after from this "after this"

- c. ru be divar
facing to wall "facing the wall"

It is interesting that (c) consists of a group-1 P + a group-2 P, but there is no EZ, and be divar has a meaning completely different from ru be divar. The following examples adapted from Samiiian (1983) further proves the view that they are compound prepositions.

- (60) a. hasan ghabl az ali amad
hasan before from ali came
hasan came before Ali.
- b.* Hasan ghabl ali amad
hasan before Ali came
- c.* hasan az ali amad

Notice the structure I would like to propose for these prepositions, and the fact that they may form compound nouns with their following NPs, as the NP_i did with their preceding heads:

- (61) ru = a bare NP adverb, "on"
miz = an NP, "desk"
ru-ye-miz = a compound NP adverb "on the desk"
az ru-ye-miz = "from on the desk"

- (62) a. ru-miz "on the table"
- b. ali tu khune-as
ali at home is "Ali is at home"

As we notice from (62 a) ru is a group-1 P, but is used without EZ. It has made a compound NP with the following NP. The same is true for (62 b), where tu khune "inside the house" is a compound adverbial of location. It is correct even without tu:

- (63) a. ali khune-as "Ali is at home"
- b. az tu-ye/tu khune biya birun
from inside-EZ/inside house come out
Come out of the house.
- c. * tu-ye/tu khune biya birun
- d. * khune biya birun

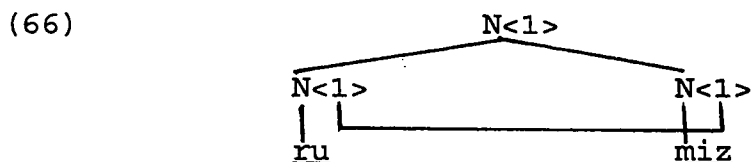
(63) shows that tu khune is an NP-adverb which needs a case assigner to pass case filter. To put it in more technical terms, birun amadan "to come out" subcategorizes for a PP-complement but tu khune does not count as one, rather az tu khune does, as we notice by the grammaticality of (b). Also notice these compounds:

- (64) a. zir-e zamin
under EZ ground "under the ground"
- b. zir-zamin "the basement"

In some dialects of north we still have compounds as (65) which are archaic in the standard Farsi:

- (65) a. bala-ye khune
over-EZ house "over the house"
- b. bala-khune "the house on the upside; second floor"
- c. Pain-e khune
under-EZ house "under the house"
- d. pain-khune "the house on the downside/lower ground"

These data again show that group-1 P+ EZ + NP is not a prepositional phrase. The other set of data we presented above, where group-1 Ps patterned with nominals, by being pluralized and taking determiners, provide further evidence that these really have nominal properties and count as NPs. We may now propose that group-1 Ps in Farsi are instances of bare-NP adverbials of location and direction which form compound adverbial-NPs with a following NP and an intervening EZ. EZ in these constructions is not a sign of genitive case. If so, then EZ is not an entity present in deep structure in these cases either, as we discussed before, rather it is there to license the occurrence of the NP following our bare NPs, thus forming a compound adverbial of location and direction which is in turn subject to case filter as we observed by their occurrence after the group-2 Ps. I assume that there is a thematic relation, theta identification, between the two NPs, similar to the one we explained for NP_a, attributive adjectives, and the prepositional modifiers in the previous section. They may make compound as NP_a did with their head nouns. Thus group-1 Ps are bare NPs that enter into a theta identification relation with their following NP, and the resulting prepositional phrase does not count as real prepositional phrase, rather it is a compound NP adverb of location:



The EZ is added in the course of derivation to S-structure, and the resulting NP still has an open theta position to discharge; that's why it may occur in case positions, specifically after group-2 Ps as we noted above; or they may inherently case mark themselves as

Larson (1985) proposed. This phenomenon is an excellent proof for Larson's theory of bare-NP adverbs, where they were case marked by a specific feature, [+F], which he assumes to be borne with these NPs. Farsi data shows that there may be compound NP-adverbs which may be inherently case marked.

As we noted before Samian characterized her group-1 Ps as [+P, +N], and group-2 Ps as [+P, -N]. Her Ezafe insertion rule which she originally proposed is as (67):

- (67) $X_{max} \rightarrow C+1$ when immediately dominated by Y' for all X
and $Y \neq V$

she later on improves it to $Y = [+N]$, in order to exclude the group-2 Ps from the rule (67), but keeping group-1 Ps in its domain, because the latter carries the feature [+N], but the former does not. Our analysis of prepositions showed that her division of Ps into two groups in Farsi is not correct and [+P, +N] encompasses an unnatural class which does not hold for Farsi. Her further conclusion, that non-verbal categories constitute a natural class from which the verb is excluded was based on the function of group-1 Ps. Our analysis indicates that the syntactic feature [+/-N, +/-V] does really hold for Farsi. However we will extend these features to account for the function of infinitives and gerunds in the next section, in the spirit of Stowell (1981, chap I, II, III).

4-INFINITIVES AND GERUNDS

4-1-INFINITIVES

In this section we study the structure and function of infinitives (INF) and gerunds in Farsi. I will show that they have both (derived) nominal and verbal behaviour. They may take their complements either to the left and function as verbs, or take them to the right and function as nouns. I will argue that there is a tense morpheme in INFs (as well as finite verbs) in the language. Gerunds and derived nominals on the other hand take their arguments to the right. They do not contain a tense morpheme. Notice the examples below:

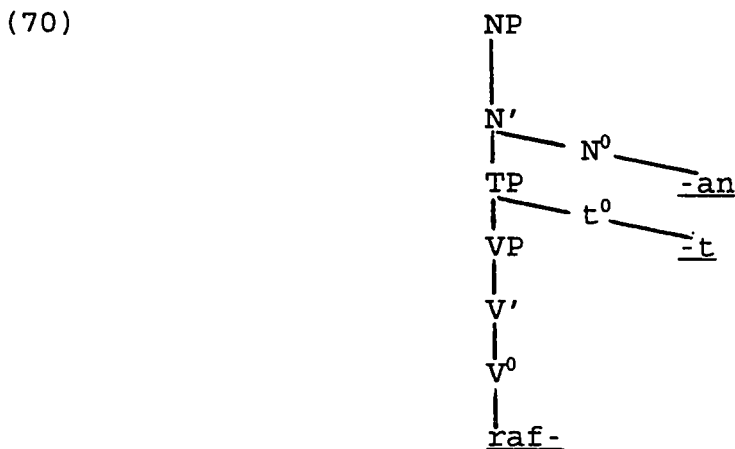
- (68) a. raft-an "to go"
b. khord-an "to eat"

-an is considered as INF marker, which is attached to the past stem of the verbs. I consider it as a functional category which projects to an NP^{max} . The initial part is the past stem, raft and khord. They are considered as the past stem (Ps) in the grammar books, but they are also 3rd person singular of the simple past tense, meaning he went and he ate respectively. To make the present stem, the last phoneme of the past stem, which is always -(i)d or -(i)t is dropped. There is a fairly large number of irregular present stems

(prs) as well. I would like to suggest that the -t and -d morphemes in the past stem of the verbs count as the tense morpheme (TM), and the present stem (prs) does not contain a phonetically realized tense morpheme; rather we postulate an abstract tense morpheme for present stem, (see chapter III for more detail). Accordingly, for (68) above we may propose (69):

- (69) a. raf-t = ps + TM
 b. khor-d = ps + TM
 c. ro(u)-... = prs + abstract TM
 d. khor-... = prs + abstract TM

Taking this for granted (we will come back to this in the next chapter), we return to the study of complement structure of INF and gerunds. I propose that INF in Farsi projects to the maximal projection VP which is immediately dominated by a TP, in turn dominated by the projection of the INF suffix, -an, as (70) below shows:



Thus we have assumed that TP and NP in (70) are independent lexical entries that head their own functional maximal projections. Given (70) the arguments of an INF may appear either on its left under the projection of VP/TP, or to its right under the projection of NP. In the former case, they count as the arguments of a finite clause and the verb assigns structural (and inherent) case to the right, and in the latter case they will count as complements of a derived nominal and appear in EZ construction. Our study also shows that, as Pollock (1989) proposes, Infl is made up of two sets of separate features, [+/- tense, +/- Agr], and each of them heads its own maximal projection. When the complements of an INF appear under VP in (70), the INF differs from finite clauses in the feature composition of its Infl (and COMP); that is INFs function similarly to finite clauses; apart from taking a subject since they lack the [Agr], in Farsi. They may be distinguished by the feature [-finite] and [+finite], if we adopt Pollock's terminology; or [+Agr] and [-Agr] to specifically refer to Farsi. Given these theoretical

(II) for the finite clauses. My contention is that similarly to what happens in finite clauses, the verb raises to t^0 , as the latter is a bound morpheme, and we may further assume that in its lexical entry t^0 selects a VP, (see IV for more details). The bracketed structure of (73 b) which also indicates the movement of the verb is shown in (74) below:

(74) [NP [TP [VP ketab ra baraye amir bor-] -d] -an]

We will characterize this type of INFs which take their complements under the VP-node to the left of the verb as type-1 INFs. They function similarly to nouns, but have certain restrictions in distribution. We assign them the features [+N, -V, +tense, -Agr]. Let us now consider (73 d-e) where the complements appear under the projection of NP in (70). This is justified by the observation that NPs are head initial in the language, and nouns take EZ before their complements. They show the same pattern of distribution as derived nominals in section (3) above. The order of complements in (73 d-e) between the DO-NP and the IO-PP is fixed, due to EZ, while it is not so in (73 a-b). This is expected as IO-PPs do not need case, but DOs do. The bracketed structure of (73 d) is shown in (75) below:

(75) [NP [TP [VP bor-] -d] -an -e ketab baraye amir]
EZ.

Thus we see the INFs in Farsi pattern both with nominal and finite clauses. The order of complements and adjuncts in each case patterns with the respective (derived) nominals and finite clauses. Notice (76), the extension of (73 b) above:

(76) diruz ketab ra baraye amir be madrase bordan...
yesterday book-DOM for Amir to school INF-to take

The same rules of topicalization discussed in chapter (I) apply in (76). DOs and IOs may freely exchange position when the former is specific, but be madrase may not. The same is true for (73 d) as we note from (77):

(77) bordan-e diruz-e ketab baraye amir be madrase...
EZ EZ

Another property of word order in INFs is that prepositional phrases may appear on either side of it, when the DOs are on one or the other as we note from (78):

(78) baraye amir bordan-e ketab be madrase
for amir INF-TO take-EZ book to school
taking the book to school

The IO in front of the INF P in (78) seems topicalized, or it may also denote the subject (agent) of the action of the verb, which is

quite normal, and expected, as prepositional phrases are not dependent on the head noun for case reasons, hence are freer than NP-complements which are.

- (79) ketab ra bordan baraye amir be madrase
 book DOM INF-to take for Amir to school

We characterize the second type of INFs which take their complements to the right under the NP-node and appear in EZ construction as type-2 INFs, and assign them the features [+N, -V, -tense, -Agr]. However, remember that they are not really [-tense], due to the function of negation, but they are so for their difference of distribution in case positions. INFs with nonspecific bare NPs are also worth mentioning, where the bare-NPs have a generic interpretation. We had proposed that they receive inherent case to the left of the verb, under government. Notice examples below:

- (80) a. u baraye bache-ha ketab mi-khan-ad
 he for children-pl book read
 He reads book for children.
- b. baraye bache-ha ketab khand-an
 for children-pl book INF/to read
 reading book for children
- c. khand-an-e ketab baraye bache-ha
 reading book for children

(80 a-b) indicate that ketab "the book" is a generic NP which receives inherent case from the verb to the left. In (80 c) it has appeared to the right, hence appearance of EZ. It seems to me that direction of inherent case is also to the left of the verb. ketab in (80 c) may still convey a generic meaning. These we may also label as group-2 INFs. Below we study the distribution of the two types of INFs

Group-1 INFs can not appear in positions where oblique and ACC case are assigned, but type-2 INFs may:

- (81) a. u ba [khand-an-e name-ye ma] mitavan-ad befahmad
 he with INF/to read-EZ letter we can understand

 ke ...
 that
 By reading our letter he can understand that ...
- b.* u ba [name-ye ma ra khand-an] mitavan-ad befahmad
 ke ...
- (82) a. man baraye [defa kard-an az ali] vaght lazem daram
 I for defend INF/to make from ali time need have

I need time to defend ALI.

- b. * man baraye [az ali defa kard-an] vaght lazem daram
I for from ali defend INF/to make time need have

In (81 & 82) the INFs follow the prepositions, ba, and baraye respectively, and as INFs in Farsi function as NPs, they must receive oblique case from the prepositions, and the structures must be all right. But only the (a) sentences with type-2 INFs are acceptable and the (b) sentences with type-1 INFs are not. The arguments of the (a) sentences appear under the NP node, but the ones of the (b) sentences under the VP/TP node. The first group (type-1 INFs) are governed by the verb and receive structural case from the verb, but the second set (type-2 INFs) do not. We characterized the first with the feature [+ tense] in order to differentiate them from the second which were [- tense]. Now, let us consider more examples:

- (83) a. man [dastgir kardan-e amir] ra az television did-am
I arrest INF/to make amir DOM from TV saw
I watched the (scene of) arresting Amir from TV.
- b.* man [amir ra dastgir kard-an] (ra)* az television didam

INF P in (83) occupies the direct object position. The (a) which is type-2 INF is grammatical, but (b), type-1 INF is not. The appearance of ra in (a) is a sign of specificity for the DO-INF in (a). Thus we conclude at this stage that type-2 INFs are not case resistant and may appear in oblique and structural case positions, but type-1 INFs are. Let us now consider (84) and (85) below, where both types of INFs may appear in subject position:

- (84) a. [dastgir kardan-e amir] lazem bud
arrest INF/to make amir necessary was
arresting Amir was necessary.
- b. [amir ra dastgir kardan] lazem bud
amir DOM arrest INF/to make necessary was
- (85) a. dar in vaght [neveshtan-e yek name-ye ahvalporsi]
at this time INF/to write-EZ one letter-EZ greetings
ba?es mishavad ke ...
cause become that ...
causes that ...
At this time, writing a letter of greetings causes that...
- b. dar in vaght [yek name-ye ahvalporsi neveshtan]
ba?es mishavad ke ...

IN both (84) and (85) the INFs appear in the subject position where NOM case is assigned, that is both types of INFs may appear in this position. Type-2 INFs could appear in case positions, thus their occurrence in NOM case position was predictable, because their complements are under the NP node, and in fact are governed by it. But in (b) sentences the type-1 INFs which could not appear in [-NOM] case positions, have apparently received NOM case from the Agr. What might the explanation for this be? I suppose that type-1 INFs differ from finite clauses in feature specification of their Agr. Their complements are under the VP node. Their arguments are governed and case marked by the verb. Thus if they appear in oblique or ACC case positions, they may again be governed and case marked by the respective heads, that is Ps and Verbs. We may remind the UNLIKE CATEGORY CONDITION of Hoekstra (1984), as:

UNLIKE CATEGORY CONDITION

At s-structure, no element of [αN , βV] may govern a projection of [αN , βV]

Thus our type-1 INFs do not appear in oblique and ACC case positions because of UCC. The reason it appears in NOM case position then may be that typ-1 INFs do not contain Agr, hence they may be governed by the Agr and receive NOM, as they lack this feature, and are not governed by it.

Our examples (80) where a bare, generic DO, (which we characterized as type-2 INFs), appeared next to an INF, exhibits both functions, as we see below:

- (86)
- a. [khandan-e kitab] khub ast
INF/to read-EZ book good is
reading/To read book is good.
 - b. [kitab khandan] khub ast
book INF/to read good is
 - c. man baraye [kitab khandan] be kitab-khane miravam
I for book INF/to read to library go
I go to the library to read books.
 - d. man baraye [khandan-e kitab] be kitab-khane miravam
I for INF/to read-EZ book to library go
I go to the library to read books.
 - e. man [khandan-e kitab] ra dust daram
I INF/to read-EZ book DOM like have
I like to read/reading books.
 - f. man [kitab khandan] ra dust daram
I book INF/to read DOM like have

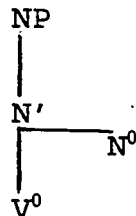
(86 a-b) indicates that INF P is in the subject position, in (86 c-d) after a preposition and receives oblique case, and finally in (86 e-f) it is in DO position. ra further indicates that INFs are always specific in Farsi.

The structure of INFs in this section complies with Pollock's (1989) position that the order of the elements in INFs and finite clauses obey totally the same principles, and they differ only in the feature specification of [Agr] in Farsi, in his terms in [+finite], [-finite]. In Farsi we might as well distinguish two types of INFs, which differ in feature specification [+/- tense]. [+tense] INFs may not receive ACC, and oblique case, but they may receive NOM case; [-tense] INFs may appear in all case positions.

4-2-GERUNDS

Gerunds are formed from a verbal stem (the present stem) and a nominal suffix. I assume the verbal element has the status of a V^0 , which as a lexical entry is subcategorised by the nominal affix; in other words there is no V^{max} in gerunds. The complements of the head verb can only appear under the projection of the nominal affix, only to its right, as (87) below:

(87)



Thus there is no VP and TP in the structure of the gerunds in Farsi. The two frequent gerund suffixes are -i, and (-e) sh :

(88)

INF	present stem	Gerund
parvarandan to breed	parvar-	parvar-esh breeding
negah-dashtan maintain	negah-dar-	negah-dar-i maintaining
shoridan rebel	shor-	shor-esh rebellling

The NP-complements of gerunds appear to the right after the EZ construction. The lack of TP in the structure of gerunds in Farsi makes them incompatible with negation which is sensitive to the presence of tense. To conclude we consider them as derived nominals, since they show all their characteristics, notice examples below:

- (89) a. parvarash-e gusfand
breeding-EZ sheep
Breeding a cattle of sheep.
- b.* na-parvarash-EZ gusfand
Neg-breeding-EZ sheep

4-3-EXTENDING THE SYNTACTIC FEATURE

I conclude this by the following diagram indicating categorial distinctions among the phrasal categories we have been studying:

- (90) Finite clause [+N, -V, +tense, +Agr]
Derived nominals [+N, -V, -tense, ---]
& gerunds
- Type-1 INF [+N, -V, +tense, -Agr]
Type-2 INF [+N, -V, -tense, -Agr]

Categories which carry [+tense] feature may not appear in [-NOM] case positions, but categories which contain [-tense] position may appear in all case positions. Categories containing [+Agr] may not appear in preverbal positions. If we amend (90) with categories bearing only [+N, +V] features, we will have (91), adopted from Stowell (1981) which apply to Farsi as well.

- (91) a. [-N] categories may assign case.
b. [+N] categories may not assign case.
c. [-N] categories may not be assigned case.
d. [+N] categories may be assigned case.

(91 a-b) are uncontroversial from our discussion so far, but (91 c-d) deserve more consideration. As far as (91 c) is concerned, we argued that prepositional phrases in the structure of NP and/or EZ construction were not real prepositional phrases. We grouped those PPs as compound NP-adverbs of location that enter into theta relation with head nouns, thus must be preceded by EZ. So (91 c) is correct as our group-2 Ps do not appear in case position. For (91 d) we may agree that nouns do receive case in Farsi. Stowell argues that adjectives may receive case in agreement with a head noun, i.e., German, Russian. Furthermore, many adjectives can function as nouns in Farsi, too. They may be pluralized and take determiners, and function as/or replace nouns. Thus, for Farsi we might assume that the EZ which intervenes between a head noun and the attributive adjective exemplifies a kind of agreement similar to agreement between adjectives and nouns in French and Dutch. Given these assumptions, which I assume to have been well motivated throughout our discussions in this chapter, we will maintain (91 d) which claims that [+N] categories may be assigned case, in Farsi as well. Thus Samiian's argument (1983) that non-verbal categories (N, P, A) form a natural class in contrast with verbal categories, and the fact that verbs do not appear in EZ construction, but

adjectives, nouns, prepositions do, is not correct. Her argument was based on the claim that some prepositional phrases appear in EZ construction, which we showed to be wrong. Thus we conclude that the syntactic features [+/- nouns, +/-verb] do hold for Farsi as well.

FOOTNOTES OF CHAPTER III

1- NPs which appear in subject position in a configuration like (6) always agree in person and number with the Agr on the verb, as well as being governed by it. For government we assume the notion of M-command and not C-command to be relevant. I will also argue, in the section on pro-drop that the lack of that trace filter in Farsi is not due to extraction from post-verbal position, rather the [NP, IP] subject position is governed by the verb when the verb raises to Agr, contrary to Italian where the subject is extracted from the postverbal position, and the verb governs it to its right.

2- The last one, that is sar may be controversial as a bare-NP adverb. It means "head", but sar-e can mean "on the top of ..., at a place", i.e., u sar-e kar ast "He is at work". However, apart from its category label, it patterns with the other prepositions, and does not refute our arguments.

CHAPTER IV

VERBAL STRUCTURE

1-INTRODUCTION

Farsi is a SOV language, with the auxiliaries following the verb and the tense and agreement markers following the rightmost element. In other words the tense and agreement suffixes attach to the lexical verb in the absence of an auxiliary and in its presence to the auxiliary. The agreement suffixes agree in person, and number with the subject. Because of this rich agreement property which holds in all finite clauses, the language allows null subjects. The only auxiliary in the active structures and Pseudo-passives is the different forms of the auxiliary and/or copula verb budan "to be" and its older form astan "to be" which is only used as an auxiliary for the present perfect in Farsi. In the description of the verbal system of Farsi both in older traditional method and the structural approaches of modern years, it is normally assumed that Farsi has two verb stems which are derived from the infinitive form, from which the different conjugations of the verb are derived. There are, however certain rules to derive these stems with a number of irregular derivations for the present stem, as the examples below show:

(1)	Infinitive	past stem	present stem
a.	khordan "to eat"	khord	khor
b.	didan "to see"	did	bin (Irg)
c.	khand-an "to read"	khand	khan
d.	nevesht-an "to write"	nevesht	nevis (Irg)
e.	kard-an "to do"	kard	kon
f.	resand-an "to cause reach, to give"	resand	resan

The suffix an is known as the infinitive marker; if dropped we get the past stem, and by deleting the last phoneme of the past stem we normally get the regular present stem. I do not recall any linguistic analysis in which the presence of a tense morpheme (TP), in Farsi, have been recognized. I believe that there is a tense morpheme in Farsi infinitives and past stems, namely, the last phoneme of the past stem. In the verbs raftan "to go", the past

stem is raft, and in didan "to see" the past stem is did. I believe that the last phonemes of the past stem -t and -d (or the stylistic variant of them -it, and -id) are tense morphemes (Tm) and accordingly the infinitives in Farsi do contain a tense morpheme. I further assume that the present stem contains an abstract tense morpheme. This view is justified on the basis of the function of negative marker -na in Farsi which according to Zanuttini's proposal (1991) can be analyzed as type one negative morpheme, neg^1 , whose appearance requires the presence of a tense morpheme, and the Farsi negative marker patterns with her type one neg-markers exactly. This point is also verified on the basis of the behaviour of infinitives in Farsi. They behave both as a noun and a finite verb with regard to the complements they take. In Farsi finite verbs take their complements and specifier on the left but nouns take their complements only to the right. Now infinitives can potentially take their complements both to the left behaving as a verb, or to the right behaving as a noun, as we discussed extensively in chapter III. On the other hand this view might be controversial in the study of verbal system of Farsi since to my knowledge, no clear indication and recognition of a separate TP has been made in the verbal system of the language. There exist no bound morphemes as raf and di, in the language, which according to our view count as past stems in the language. However, we believe that past stems and the tense morpheme count in many cases as a single, inseparable morpheme in the structure of the verbal system in the language, it is actually the latter which might count as a lexical entry in the lexicon. We claim that assuming a TP has nice consequences in the description and explanation of the language, as we noticed in chapter III, in the discussion of INFs and gerunds. Furthermore this assumption considers Farsi on par with other languages such as French, English, Italian, etc which are presented as having two separate morphemes for the tense and Agr even though these do not necessarily constitute concrete morphemes. Thus I will take this for granted in my study of the verbal system in Farsi, though it does have some unpleasant consequences for my analysis of future tense. We adopt the proposal of Pollock (1989) that Infl consists of two separate nodes Tense and Agr, each projecting its own maximal projection, but contrary to him we place AgrP above TP in Farsi. I further adopt Ouhalla (1990) who (contrary to standard belief which proposes that AUXs are generated inside the VP predicate) suggests that AUXs are categorically different from lexical verbs and project to a separate maximal projection outside the VP. He suggests that AUXs belong to the class of ASP(ectual) elements and thus project an ASPP, rather than under VP. This we will show not to be only a theoretical position, but there are differences of function and distribution between instances where an AUX takes a VP complement (perfect tenses, where AUXs select the past participles (ppr) of the main verbs), and verbs/modal-like verbs that take a VP-complement in the language.

The agreement markers on the verb or auxiliary indicate singular and plural in three persons as (2) below indicates. Third

person singular suffix for the tenses formed with the past stem is a zero morpheme, but for the third person present stem is -ad as we see below:

(2)		singular	plural
	1st	-am	-im
	2nd	-i	-id
	3rd	0/-ad	-and

Furthermore, I conclude this introductory section by saying that there is no real modal in the Farsi verbal system.

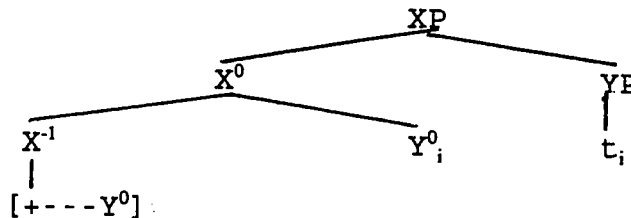
2-SIMPLE TENSES

Example (3) below characterizes simple past tense, which takes no AUX. It is formed from the past stem (ps) of the verb, the tense morpheme, and the Agr suffixes:

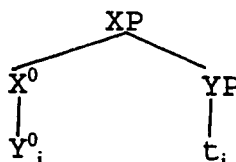
- (3) ma be madrase raf-t-im
 we to school ps/go-Tm-1S
 We went to school.

As we already proposed the existence of a separate past stem as raf is controversial, but if we study the structure of verbal system in the language, we notice that whenever there is a VP projection there is also a TP above it, independent from the fact that there might or might not be an AgrP present, i.e., INFs, and finite clauses. In order for the verb to receive tense and Agr properties the V^0 , that is raf- in (3), raises to t^0 and Agr^0 , thus illustrating an instance of head-to-head movement. Roberts (1991) proposes the following types of head to head-movement, which is an extension of the theory of Noun/Verb Incorporation (NI/VI) of Baker (1988). We will study verb incorporation (VI) in more detail in the section on future and the causatives in the next section:

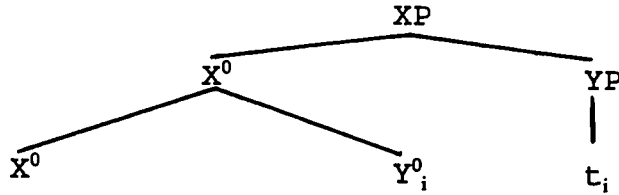
- (4) a. Incorporation by Selectional Substitution



- b. Incorporation by Free Substitution



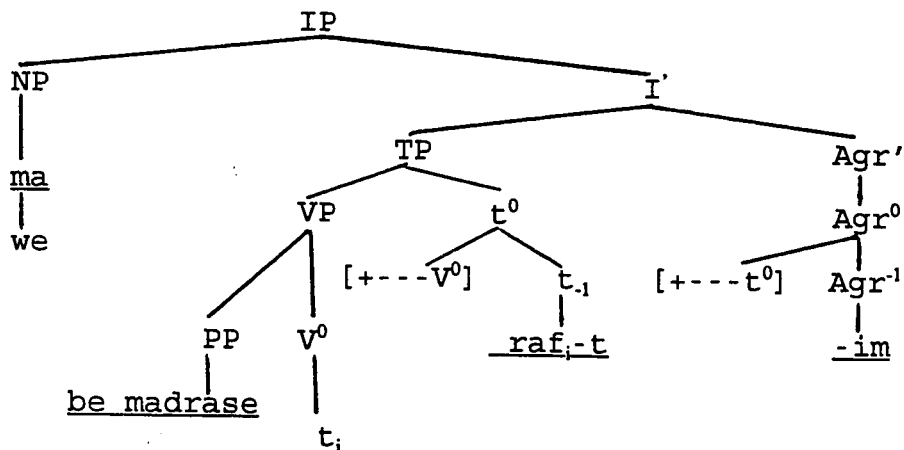
c. Incorporation by Adjunction (Chomsky-Adjunction)



In our study of the verbal system of Farsi we are mostly concerned with the first type of head-to-head movement (4 a), which is also on par with the PROJECT ALPHA system, and the theory of the lexicon we adopted in chapter I. Roberts (1991) explicates (4 a) as follows: "The incorporation host morphologically subcategorizes for the incorporee V^0 : hence a structural slot is created for the incorporee at D-structure as a function of the lexical properties of the incorporation host." This function causes raising of the incorporee and is of the type of substitution which leads to creation of an amalgamated head. In such a configuration the head of the incorporation host remains the head of the compound, but the index of the incorporee percolates up so that the trace left behind by its raising can properly satisfy ECP. In (4 a) the index of the Y^0 percolates up to X^0 , head of the compound formed by the raising, so that its trace satisfies ECP, by percolation convention of Lieber (1983), and Di Sciullo and Williams (1987). However Roberts in his analysis does not confine this kind of raising or incorporation by substitution to the types that lead to a morphological amalgamation of the incorporee and the incorporation host only, rather he extends it to compound-types as well, in his footnote (2, P 213), where he argues that selection by substitution can be extended to compounding of two heads which does not necessarily make a morphological compound, but leads to a simple juxtaposition of the selector and the selectee. We will encounter cases of this latter type of substitution of the compound-type in our study of future and perfect tenses in the next sections.

Now for our sentence (3) above the Phrase Marker may be as (5) below:

(5)



In the first step of the movement the V^0 raises to t^0 , and makes an amalgamated compound, that is raf-t; then the newly-formed compound raises to Agr^0 by selectional substitution as in the previous step, leading to raf-t-im.

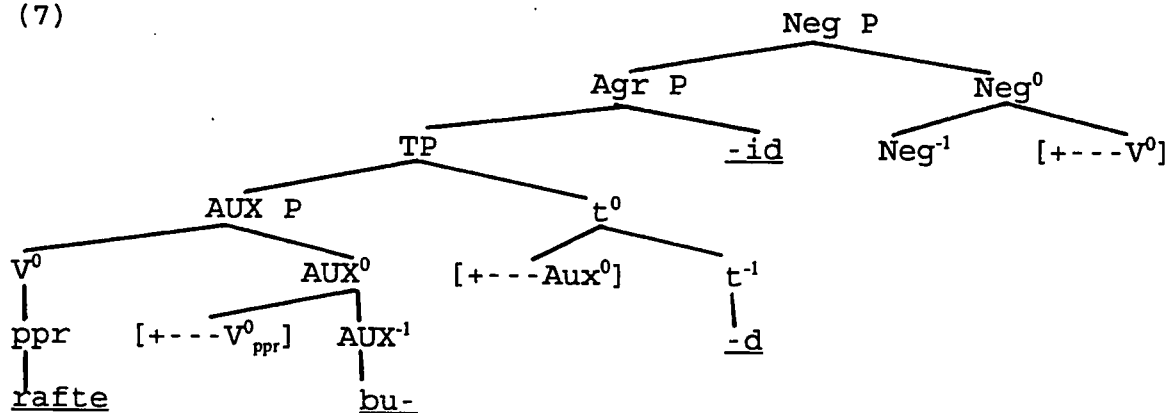
Now let us consider (6) below, where an AUX is present in the structure of the verbal system, that is the appropriate form of the AUX/copula bud-an "to be". I am also suggesting here that the only AUX in Farsi is bud-an. The verb shod-an "to become", the so-called AUX of the passive constructions does not count as an AUX, (see the section on passive). The AUX bud-an as I suggested before projects an ASPP outside the VP. We will see in the section on future and negation that in fact there should be a structural difference between the lexical verbs and AUXs in Farsi. The three different forms of bud-an that function as AUX in the language all subcategorize for the past participle (ppr) of the main verb. Past participle is formed by affixing the morpheme -e to the past stem + TP which in many cases functions as an independent adjective. Thus I assume that the past participle is available as an independent lexical entry in the lexicon. It also contains a Tm in it which attracts the NegP in Farsi.

- (6)
- a. shoma be madrase rafte bu-d-id
you to school ppr/gone ps/AUX-Tm-2Pl
You had gone to school.
 - b. shoma be medrase na-rafte bu-d-id
you to school Neg-ppr/gone ps/AUX-Tm-2Pl
You had not gone to school.
 - c. Shayad be madrase rafte bash-and
perhaps to school ppr/gone prs/be-3Pl
Perhaps they (might) have gone to school.
 - d. shayad be madrase na-rafte bash-and
perhaps to school Neg-ppr prs/be-3Pl
Perhaps they (might) have not gone to school.
 - e. u be madrase rafte (?ast¹)
he to school ppr/gone prs/is
He has gone to school.
 - f. u be madrase na-rafte (?ast)
he to school Neg-ppr/gone prs/is
He has not gone to school.

As we notice from the examples (b-d-f) the negative prefix must always attach to a lexical verb dominated by a tense morpheme. It never attaches to an AUX which in our examples absorbs the tense and Agr features of the sentence. Had we not assumed the AUXP outside the projection of the VP as an independent projection, ASPP, there would be no way for the Neg P to distinguish between the lexical main VP and the AUX. As they would be both under the

same projection and both carrying the same features. Below I will explain the relevant movements in the verbal system of the sentence (6 b) which characterizes all the other examples in (6), and all the verbs with an AUX in Farsi.

(7)



AUXP has an empty slot in its lexical entry for the V^0 , more specifically for the past participle of the lexical verb, by selectional substitution of the compound type which triggers raising of the V^0/ppr to AUX , as shown in the first step of the movement in (7). All the other projections in (7) above, seem to have similar selectional properties of the morphological type as part of their lexical entries which cause the raising of the previous step of derivation, depending on whether a projection level is present in the course of derivation or not. In the first step of derivation V^0/ppr raises to AUX^0 and substitutes the empty slot already available for it, (see Roberts (1991)). The trace left behind by the first step of movement can be properly antecedent governed, as the index of the raised V^0 percolates up to AUX^0 , which is now the head of the newly formed compound. In the second step the newly-formed compound raises to t^0 , where the head of the compound², that is the AUX bu-, receives the tense features as desired. The ECP is satisfied similarly to the previous step. The third step of movement is to Agr^0 , which also morphologically selects the AUX/V. The fact that Agr^0 may only attach to the AUX and not to the V/ppr may be accounted for by assuming that it is the head of the previous step of movement, that is the AUX, which counts as the head of the compound and must receive the Agr features. The last step of movement involves raising to Neg^0 which subcategorizes only for a lexical verb dominated by a TP. Now if we had not distinguished between AUX P and VP, the Neg would not be able to differentiate between them. But as it is, the lexical features of V^0 percolate up in the V-I system of the extended projection of the verbal system, as proposed by Grimshaw (1990). Negation properly distinguishes between V^0 with the feature specifications of [+V, -N, +L, -F] and AUX with the features [+V, -N, -L, +F], and always selects the former set of features which must also contain a tense morpheme (Tm) as well; (notice that both AUX and V/past participle contain a tense morpheme). Every step of

must raise to t^0 to acquire tense, thus it is incompatible with subjunctive/imperative moods, indicated by be-, which is absent whenever tense projection is present there.

3 - FUTURE

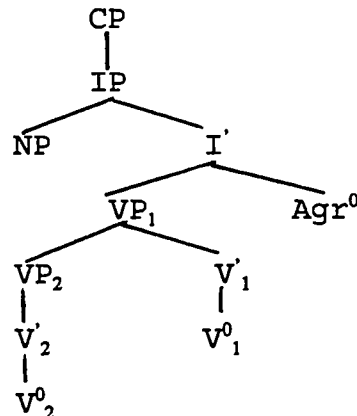
Sentence (10 a) below shows a typical example of future tense in Farsi and (10 b) is its negated form. In (10) there are two lexical verbs. The first one is the present stem khah- from the infinitive khastan "to want". It can be used both as a main verb, taking its own NP complements, or it can be used similarly to modals in English to form future in Farsi. It is the second usage that we are concerned with here. khah in (10) is the present stem and the agreement suffix -am is attached directly to it, so it is finite. The second verb is the past stem avard, from the infinitive avardan "to bring". It does not take agreement suffixes in the future and is used only in one form, and thus is non-finite. The direct object, ketab ra "the book DOM", and the indirect object, baraye to "for you" are NP and PP complements of avard and not khah.

(10) a. man ketab ra baraye to khah-am avard
 I the book DOM for you prs/buy-1st-S ps/bring
 I will bring the book for you.

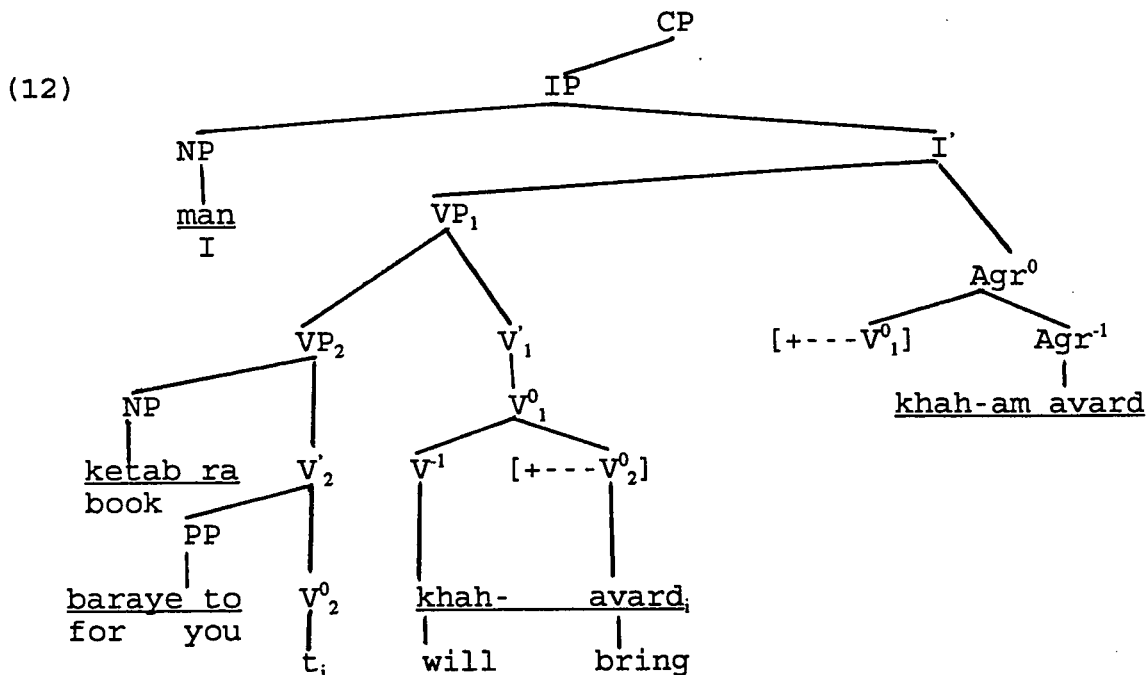
(10) b. man ketab ra baraye to na-khah-am avard
 I the book DOM for you neg-prs-1st-S ps/bring
 I will not bring the book for you.

To analyze the structure of future as explicated in (10 a-b), I would like to follow Li (1990) according to whom we may say that matrix (finite) verb khah-am lexically subcategorizes for a non-finite past tense verb. The tree diagram (11) further explicates this point. It is essential for my analysis here to adopt the point that the past stem of the subcategorized verb, avard, is strictly selected by the matrix verb, and the past stem functions as a single, atomic unit. In other words it is not made up of two separate morphemes, a verbal stem and a past morpheme, rather a single lexeme.

(11)



(11) is the type of verb incorporation (VI) where a matrix verb takes a VP complement and causes the head of its complement verb to raise or incorporate into the higher one through a syntactic head to head movement, $\text{move } \alpha$. The trace left behind by this kind of movement is subject to the ECP. Comparing (11) with the Farsi examples (10), VP_2 is the subcategorized past stem, avard together with its complements, and VP_1 is the matrix finite verb khah-am. I believe that the matrix verb has a morphologically subcategorized slot for the embedded verb, and thus the V^0 of the embedded verb raises to the V^0 of the matrix verb and substitutes the empty slot already available for it. The first step of movement that I will explain for the future tense in Farsi is that of the compound type of incorporation explicated above. Before explaining the structure of (10) based on the above theoretical background, I would like to return to the point I made in section (1). I argued that the existence of a phonetically realized past morpheme for the past stem in Farsi but an abstract one for the present stem could have an unpleasant consequence for my analysis of future tense. I would like to emphasize that the past stem of the embedded verb in (10) above is selected as a whole amalgamated compound in future tense and past morpheme is not separable from its verbal element. Taking this for granted, together with (11) and ((4) in section (2)) I present the tree diagram (12) for (10 a):



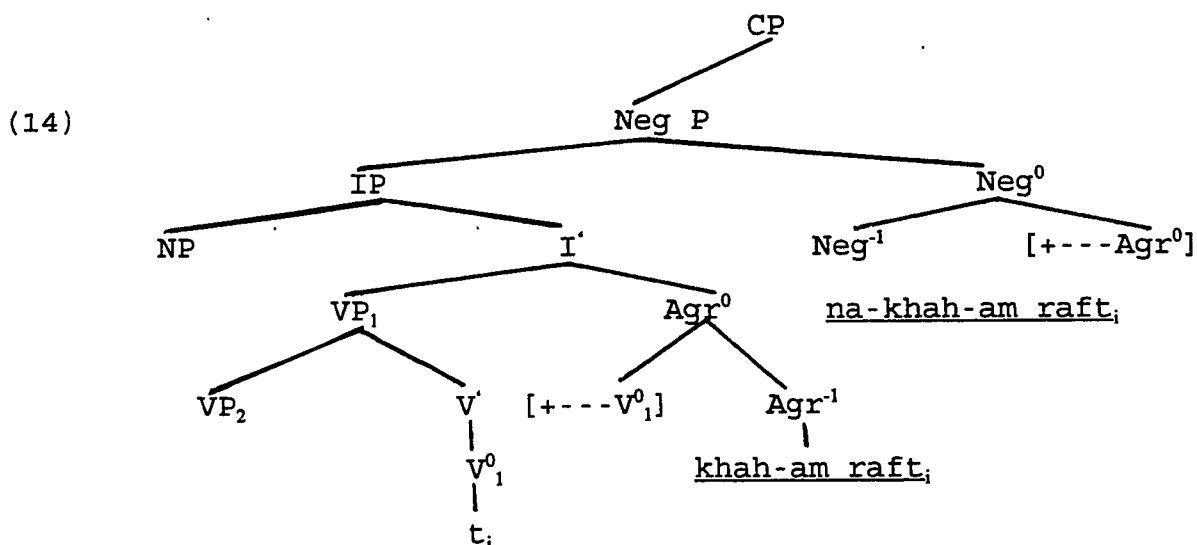
In (12) as the diagram must clearly show the matrix verb VP_1 , khah the present tense of the modal like element similar to will in English, has an empty selected slot for past stem, V^0 of the embedded verb, of the type of compounding as explicated above in (11). This triggers raising of V^0_2 in the first step of movement, so

that it substitutes for the empty slot already available for it. Now the index of the raised verb percolates up to the head of the compound just formed by the percolation principle of Di sciullo and Williams (1987) so that the trace left behind can satisfy ECP. Furthermore it is the head of the matrix verb, V_1^0 , which counts as the head of the compound. Remember that as the structure after the first step of movement clearly shows this step of raising is obligatory, because the head of the embedded verb, V_2^0 , follows the head of the matrix verb, it thus can not be assumed to be left behind. In the second step of the movement, V_1^0 , the head of the newly formed-compound which is morphologically subcategorized by the bound morpheme, -am, (Agr^0), raises, incorporating into Agr^0 that already has an available empty slot for it. Interestingly enough, as desired, the agreement morpheme affixes to the head of the raised compound, that is V_1^0 , and we have the surface structure khah-am avard, and not khah- avard-am, which would be incompatible with our analysis. Like the first step of movement, the index of the raised element percolates up and the trace left behind satisfies the ECP. The type of chain we have by this raising is $\langle V, V, Agr \rangle$ which is normally assumed to be a proper chain. While if we had assumed that there were a separate past tense morpheme with the past stem of the subcategorized verb, that is V_2^0 , then our chain would be as $\langle V, T, V, Agr \rangle$. This type of movement where a functional head intervenes between the lexical heads is assumed to be improper movement and violates the ECP, (see Li (1990)). Li also argues that: "Though a verb morpheme may form a compound with Neg(ation), it is impossible for a VI compound to "contain" a neg that exclusively negates the embedded verb." This prediction comes true in the construction of future tense in Farsi where the Neg⁰ (a prefix na-) attaches to the lexical verb of the matrix verb and never to the head of the embedded one. Li also predicts that T(ense) or Infl is not possible in the embedded verb of a VI compound. However, Farsi data above shows that his assumption is better change to the absence of a finite embedded verb in a VI compound, although we have neglected the presence of an independent head for tense in our data. Returning to sentence (10b) repeated here as (13), we observe that the negative prefix na- can only attach to the lexical head of the matrix clause and not to the embedded clause, which is compatible with Li's prediction that a VI compound can not contain a negative morpheme only negating the embedded clause.

- (13) man ketab ra baraye to na-khah-am avard
 I the book DOM for you neg-prs/will-1st-S ps/bring
 I will not bring the book for you.

Following Pollock (1989) Li assumes a functional projection for negation. We also follow them to assume a functional projection for negative in Farsi. Farsi negative marker is of type-one negation, Zanuttini (1991), that always selects a complement that contains a tense Phrase (TP). Thus I assume that Neg morpheme in Farsi, too

has a morphologically selected slot of the type ((4a) in section (2)), that is incorporation by substitution as proposed by Roberts (1991), which triggers raising of the compound formed under Agr⁰ to neg⁰. Interestingly enough neg⁰ being a prefix can only attach to the lexical stem of the matrix verb, a desired result as we see in the surface structure na-khah-am raft. The tree diagram (14) shows this last step of movement:



In (14) Agr⁰ in the last step of movement raises to Neg⁰ where there is already an empty slot available for it. If following Li we assume that Agr⁰ is not a position to/from which a theta role is assigned, then its trace left behind by raising will not be subject to the ECP. Even if we assume that its trace does count as a variable, the ECP is still satisfied by feature percolation principle of Di sciullo and Williams (1987). Distribution of negation will indicate in the section on passive that there is no real passive constructions in Farsi, where Neg morpheme does not attach to the (past participle) lexical head of the main verb but to the head of the so called passive auxiliary, while Neg never attaches to an auxiliary in Farsi. Assuming na- and -am as heads of functional categories, and adopting Grimshaw's definition of Extended Projection according to which functional categories take unique functional or lexical complements and functional categories always appear higher up in the projection structure than the lexical ones, we may conclude that the three movements explicated in the structure of future tense in Farsi are examples of head to head movement that fall within the Extended Projection of the V-I system and observes principles of head to head movement including ECP. Before closing this section I would like to present some defective, modal-like elements which show parallel distribution and function in one of their functions. They are (1) bayad "must", (2) bayasti "must", (3) bayasti "must", all stylistic variants, and (4) mitavan "one can". They can take both a CP or a VP

complements. I will deal with the latter type here. They have lost most of their modal-like properties and are used in third person singular with a small Pro subject which can denote an impersonal subject meaning "one", as the examples below clearly show:

- | | | | |
|---------|-------------------|--|-------------------|
| (15) a. | Pro bayad raft | | |
| | Pro must ps | | "one must go" |
| b. | Pro na-bayad raft | | |
| | Pro neg- must ps | | "one must not go" |
| c. | Pro bayad na-raft | | |
| | Pro must neg-ps | | "one must not go" |

(15 a-b) pattern exactly with our analysis of future, and the Neg morpheme also functions accordingly. (15 c) however deviates from our analysis, it has a different interpretation and is much less common than (15 b). The other two modal-like elements also pattern as (15) and I will not further study them here.

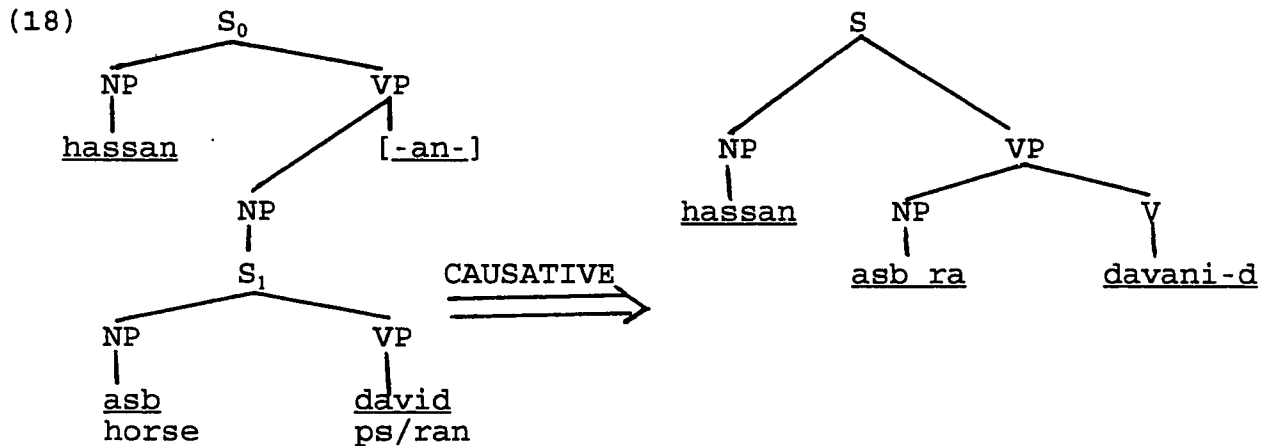
4-CAUSATIVES

In this section I will study the structure of one type of causatives that have morphological realisation in Modern Farsi. I will discuss the study of causatives by Moyne & Carden (1974) and the one of Dabir-Moghaddam (1986). After explaining shortcomings of their analysis, I will present a new analysis of causatives based on Baker (1988) and Li (1990), which can properly accommodate the structure of morphological causatives in Farsi. Causative formation in the traditional grammar books has been accounted for as a lexical process taking place in the lexicon. According to this approach the causative morpheme -an (or its stylistic variant -ani, which is only used for the past stem), is added to the present stem of certain transitive and intransitive verbs in order to derive the present stem of a causative verb to which the past stem morpheme or the infinitive marker is then added as example (16), below shows:

- | | | | |
|------|---------------------|--------------|--|
| (16) | Non-causative | present stem | causative (Infinitive) |
| a. | paridan
"to fly" | par | par- <u>an</u> -dan/par- <u>ani</u> -dan
"cause to fly" |
| b. | khordan
"to eat" | khor | khor- <u>an</u> -dan/khor- <u>ani</u> -dan
"cause to eat" |
| c. | davidan
run | do (Irreg) | dav- <u>an</u> -dan/dav- <u>ani</u> -dan
"cause to run" |

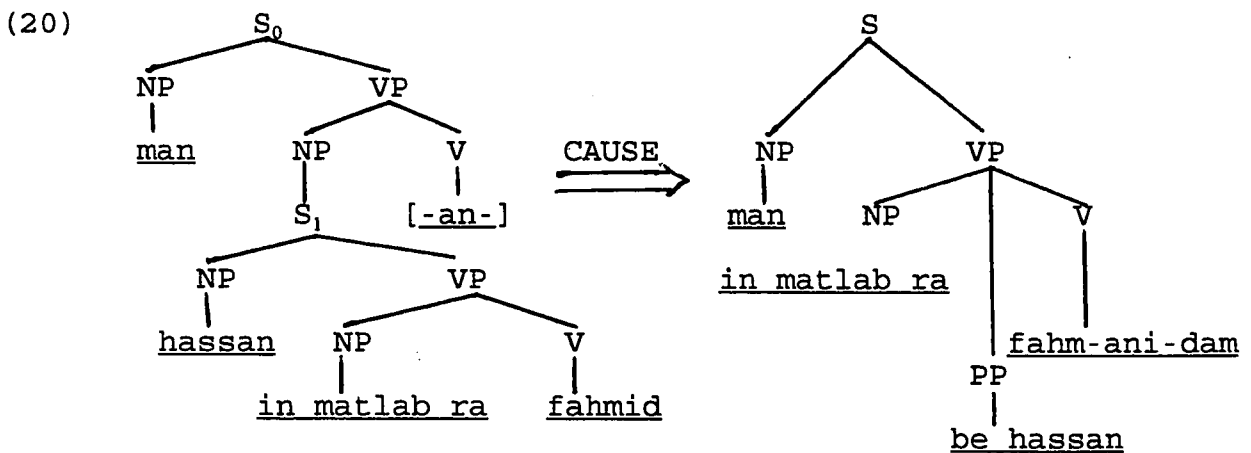
Moyne and Carden (1974) suggest an analysis for Farsi causatives based on strict subcategorization and selectional restrictions where a higher S(entence) selects a lower one which is in turn dominated by an NP. In other words a higher S selects an NP complement which contains an S, as (17) below and its tree diagram (18) taken from Moyne & Carden (1974) show:

- (17) hassan asb ra dav-ani-d
 Hassan the horse DOM prs cause-ps
 Hassan caused the horse to run



According to Moyne & Carden (1974) in (18) the subject of the non-causative, embedded S_1 , becomes the object of the causative structure on the right of the tree (18). For a transitive non-causative such as (19 a) and its causative equivalent (19 b) they suggest the tree diagram (20):

- (19) a. Hassan in matlab ra fahmid
 hassan this subject DOM ps-3rd-S
 Hassan understood this subject.
 b. man in matlab ra be hassan fahm-ani-d-am
 I this subject DOM to hassan prs-cause-ps-1st-S
 I caused Hassan understand this subject.



In (20) on the left hassan is the subject of S_1 , which turns to a dative position and the NP in matlab ra, the direct object, becomes the direct object of the causative sentence on the right. Moyne & Carden (1974) combine the verb of the embedded clause with that of

the causative morpheme using various transformations based on the generative-transformational theory of the late sixties and early seventies. However as they claim themselves they can give "no convincing explanations for the asymmetry" between (18) and (20) where in the former the subject of the embedded clause becomes the direct object of the causative equivalent while in the latter the direct object of the embedded clause remains the direct object of the causative sentence and the subject demotes to a dative position. They propose some arbitrary transformations, as they claim themselves, to demote the subject of S_1 in (20) to a dative position, what they call a "raising-to-dative" which holds independently in Farsi. Analysing in a pre-GB theory they are not concerned with the ECP and other constraints in the movement of elements in the syntactic component. As in the generative grammar of the early seventies the verb counts as the head of a sentence, even if they were to account for the constraints on head-to-head movement, they would have no problem in doing so as the movement would be a case of raising from a lower A-position, that is the V^0 of the embedded clause, to the V^0 of the matrix clause, another A-position and as the chain $\langle V, V \rangle$ is a proper chain their analysis would have no problem. But given the fact that Infl/Agr is taken as the head of a clause in GB model the only way to save their analysis is to delete the S_1 of the embedded sentence in (18) and (20) and the NP immediately dominating it, and substituting instead a VP. This would also justify the asymmetry they failed to explain. Otherwise their analysis will face the problem of an improper movement and would fail by the ECP, because the embedded Infl/Agr would count as a functional head (A'-position) that intervenes in the raising between two A-positions and leads to ECP violation.

Dabir-moghaddam (1979) (I quote from Dabir-Moghaddam(1986)) proposes a similar syntactic analysis for Farsi causatives in which the NP dominating the embedded S_1 in (18) and (20) above is left out and inserts a semantic element CAUSE instead of the causative morpheme -an- under V^0 of the matrix verb. He then amalgamates the two sentences with a verb-raising transformation. He accounts for the asymmetry in the grammatical functions of the two types of causatives discussed above, (18) and (20), with recourse to accessibility hierarchy of the grammatical relations proposed by Keenan and Comrie (1977):

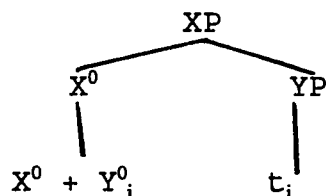
(21) subject > direct object > indirect object > complement object

In an example like (18), when the "verb-raising" transformation leads to the amalgamation of the matrix causative and the embedded verb, the subject of S_1 takes the direct object position of the resulting causative construction because the subject position is already filled by the subject of the matrix verb; but in (20) the subject and the direct object positions of the resulting causative clause are already filled with the subject of the matrix verb and the direct object of the embedded clause respectively, thus the subject of the embedded sentence can only occupy the indirect

object position according to the accessibility hierarchy (21). As the indirect object is always preceded by a preposition in Farsi, he introduces the preposition *be* "to" in the structure in (20). His analysis faces the same type of problems pointed out above with respect to Moyne & Carden's analysis. Dabir-Moghaddam (1986) proposes a lexicalist approach for the causative constructions in Farsi, following Jackendoff (1975). He divides causative constructions in Farsi into root, morphological, and auxiliary ones. He subdivides the root ones into subdivisions based on their morphological form. He adopts Jackendoff's view (1975) that the lexicon consists of "lexical entry(ies)" and "redundancy rules". Lexical entries consists of phonological, syntactic and semantic representations; and the redundancy rules indicate the systematic, rule-governed relations between lexical entries by means of three types of redundancy rules: (1) lexical, (2) semantic, and (3) syntactic redundancy rules. He keeps the causative and non-causative verbs as separate entries in the lexicon and relates them with the three kinds of redundancy rules mentioned above. Most of his non-morphological causatives, that is, the auxiliary causatives fall under compound transitive verbs and their unaccusative counterpart which can not be handled by syntactic rules of the head-to head movement type. The root causatives are verbs which match transitive verbs and their ergative counterparts, which do not fall under incorporation and syntactic compounding. As far as morphological causatives are concerned, I believe he misses the generalization made by the type of syntactic incorporation discussed by Baker (1988).

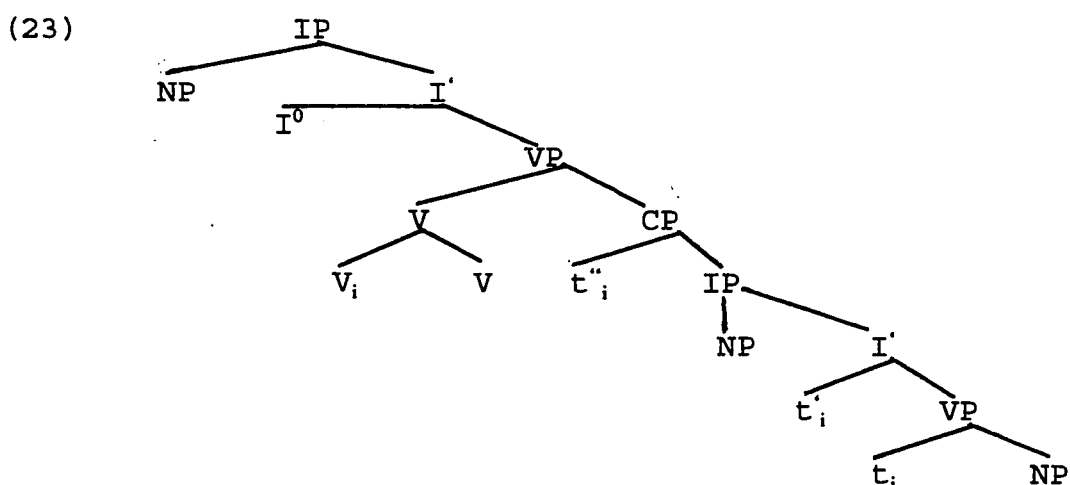
Baker (1988) presents a theory of syntactic incorporation where it is assumed that the operations that permit deriving morphologically complex words from basic, bound/free elements is a variant of the rule "move α " that applies in the cases of head to head movement. Thus many cases of morphological amalgamation of morphemes are considered as instances of head to head movement that apply in the syntactic component, and are subject to syntactic rules and constraints, in particular to the ECP, as (22) shows:

(22)

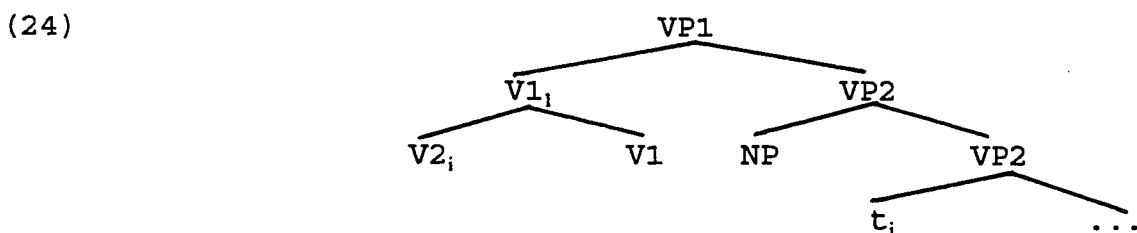


The particularity of this analysis is that it can account for certain morphological and compound word-formations on the basis of syntactic terms and conditions which are independently motivated and required, most importantly the Empty Category Principle (ECP). Baker also argues for cases of verb incorporation (VI) where certain modal-like verbs and causative morphemes can take a clausal complement. He attributes this to UTAH and the projection principle. He argues that VI is a process of head-to-head movement where the V^0 of the embedded clause raises to V^0 of the matrix verb

as the diagram (23) shows:



In (23) t_i occupies the D-structure position of an embedded verb which raises to the head position of the matrix verb by first moving to I and then to C of the embedded clause, which are both A-bar positions, and then it moves to the upper V, an A-position. Li (1990) correctly argues that VI can never involve an embedded clause or Infl. In other words the embedded verb can never incorporate through an intervening Infl and C^0 of the embedded clause because they function as A-bar positions (functional heads) that block such a raising by violating principle C of the binding theory as well as ECP. He suggests that VI always, cross-linguistically, involve modal-like verbs and causative morphemes subcategorizing for an embedded VP and not an IP/CP, as (24) below shows:



In (24) t_i raises in one step to the head of the main verb and the trace left behind can properly satisfy ECP and there is no violation of the principle C of the binding theory. The chain produced by the first type of movement, (23), is $\langle V, C, I, V \rangle$ which is normally assumed an improper movement, because it is a movement from A-position to A-bar position and back again to A-position, violating principle C of binding theory and ECP. But the chain produced by (24) is $\langle V, V \rangle$ which is a proper chain respecting principle C and the ECP. LI's analysis is however more sophisticated but it clearly indicates that VI involves a matrix verb selecting a VP complement and never a CP one. Having shown the superiority of Li's analysis

and the previous analyses of causatives in Farsi, I will discuss the causative constructions in Farsi on the basis of Li's analysis, reminding that such a process is already available in the construction of future tense in Farsi, as I discussed before. There are two types of causatives in Farsi:

Type (1): The embedded verb is intransitive and its subject (external argument) functions as the object of the whole sentence (cf (18)) as (25) and (26) below show:

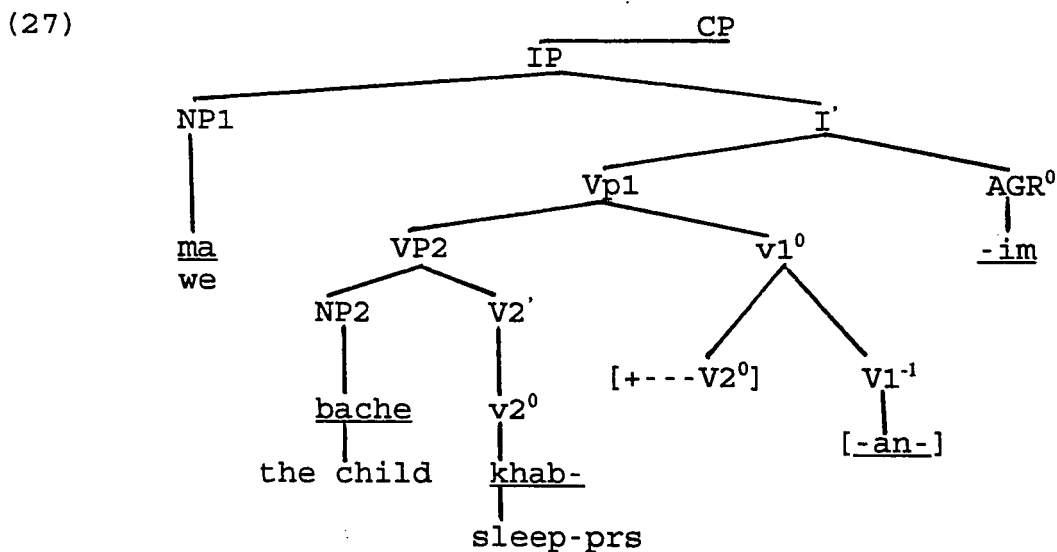
(25) a. bache khabid
the child ps-3rd-S "The child slept"

b. ma bache ra khab-and-im
we the child DOM prs-cause-ps-1st-Pl
We made the child sleep.

(26) a. ali khand-id
ali ps-3rd-S "Ali laughed"

b. ma ali ra khand-and-im
we ali DOM prs-cause-1st-Pl
We made Ali laugh.

The English translations seem to be literal and the causative verbs in (25) and (26) behave as transitive verbs. (27) shows the structure of (25 b):



I assume that $V2^0$ is specified for case, as it is a lexical verb that subcategorizes for bache, but can not assign case to it because $V2^0$ is dominated by $V1$, the causative, which is unspecified for case and prevents $V2^0$ from assigning case to its argument⁵, (see Li (1990) for more detail). This seems to me to be a correct assumption since $V1$ only has one external argument which is

assigned NOM case by Agr⁰. V1⁰ triggers verb raising so that V2⁰ moves to the empty slot already available for it by morphological subcategorization of V1⁰ for V2⁰, which is an instance of selection by substitution as discussed by Roberts (1991) and adopted already in my discussion of the future tense. V1⁰ now acquires case [+C] from V2⁰ by feature percolation of Di Sciullo and Williams (1987). Verbs in Farsi may assign Acc case to NPs. This is indicated by ra, DOM, after NP2. The trace left behind by the first step of the movement is properly governed. The second step in which the new compound moves to Agr⁰ is as I explained before for the future tense.

Type (2): The embedded verb functions as the object of the whole sentence, as in (28) and (29) below:

- (28) a. ali ghaza ra khord
ali the food DOM ps-3rd-S
Ali ate the food.
- b. ali ghaza khord
ali food ps-3rd-S
Ali ate the food.
- c. man ghaza ra be ali khor-an-d-am
I the food DOM to ali prs-cause-ps-1st-S
I caused Ali to eat the food.
- d. man ali ra ghaza khor-an-d-am
I ali DOM food prs-cause-ps-1st-S
I caused Ali to eat the food.
- (29) a. bache lebas ra pushid
the child the clothes DOM ps-3rd-S
- b. bache lebas pushid
the child the cloth ps-3rd-S
The child put on the clothes.
- c. man lebas ra be bache push-an-d-am
I the cloth DOM to the child prs-cause-ps-1st-S
I caused the child put on the clothes.
- d. man bache ra lebas push-an-d-am
I the child DOM the cloth prs-cause-ps-1st-S
i caused the child put on the clothes.

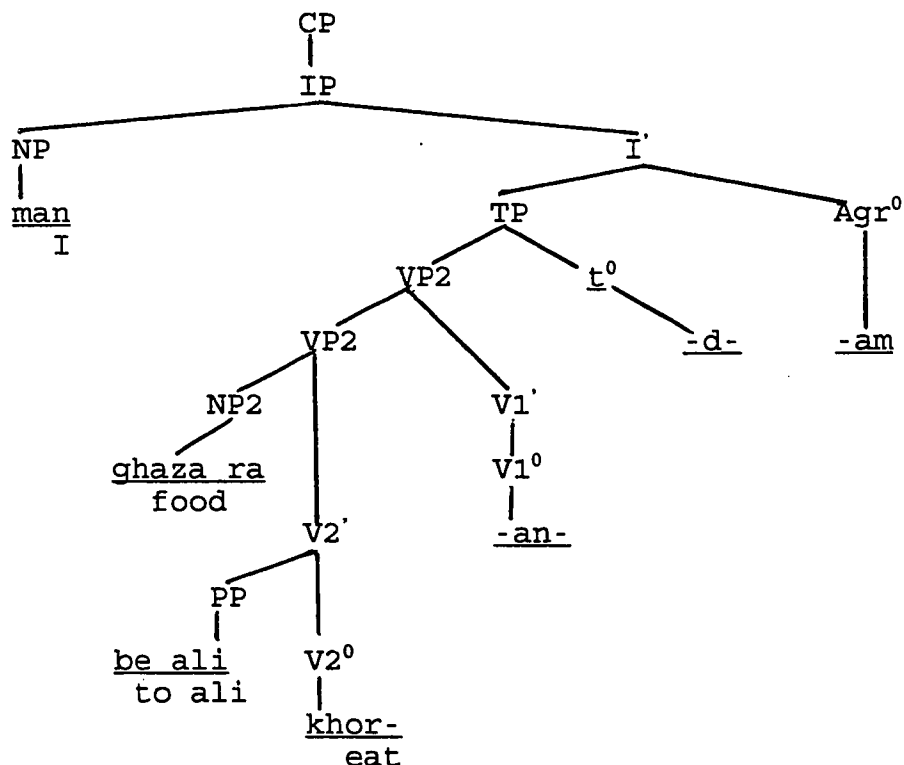
I would like to emphasize that in example (28 b-d) and (29 b-d) we are concerned with non-specific DO-NPs which must stay adjacent to the verb, and according to our arguments in chapters II & III receive inherent case from the verb. In the (b) examples the verbs ghaza khord "ate the food" and lebas pushid "put on the cloth" seem to have made a complex (or compound) verb with the non-

specific NPs, and appears to function as a single unit. They fall under type (1) causatives discussed above. When their respective verbs raise to the causative morpheme which has a morphologically selected slot for them, they acquire the property to assign ACC case as well, (see footnote 4). They appear to produce a situation similar to double objects in English. The verb in its base position assigns inherent case to the non-specific DO, (and is not required to be dominated by a TP in this case), but assigns ACC case when it amalgamates with the causative morpheme, dominated by a TP. In the (a) examples we have transitive verbs where the direct object functions as the direct object of the causative counterparts in the (c) examples and the subject receives an oblique case from a prepositions. The subject can also receive a Gen(itive) case from the Ezafe construction (EZ) (equivalent to of in English) in colloquial speech, as (30) shows:

- (30) Pro [_{NP} lebas-e bache]-ro push-an-d-i?
 Pro [_{NP} the cloth-EZ the child]-DOM prs-cause-ps-2nd S
 Did you make the child put on the cloth?

The diagram (31) shows the structure of (28 c)

(31)=(28 c)



In (31) the causative morpheme takes a VP complement, VP2 in our example which lacks tense/Agr features (is non-finite). Now VP2 in (31) is equivalent to its non-causative counterpart (28 a), except that it (VP2) does not have Agr features and is thus non-finite.

Being non-finite, its external argument, ali, can not receive NOM case. Neither can ghaza receive ACC case from $V2^0$, because the $V1^0$ which dominates it is not specified for case⁵. Now $V1^0$ causes verb raising as explained before so that $V2^0$ moves up, and amalgamates with the causative morpheme. Now its case features percolate up to $V1^0$, and it acquires the property to assign case. The question is why the direct object of the non-causative counterpart, that is ghaza, and not ali, can only receive accusative case and ali must either receive oblique or Gen case. Adopting a less technical and less sophisticated analysis than Li we may argue the ghaza inherits its accusative case from its non-causative counterpart, (28 a). Only ghaza and not ali can carry ra the DOM. ghaza has identical thematic relations with both the causative, and the non-causative verbs. I would like to adopt recent literature on theta-grid according to which the argument structure and theta roles of the predicates are arranged in the theta-grid in conformity with the thematic hierarchy and hence in the argument structure. As such they are mapped into syntax in such a way that prominence in the thematic hierarchy and in the theta-grid directly corresponds to syntactic prominence. In the argument structure of khordan the head of VP2, ghaza occupies a closer position to the head than ali. This is part of the lexical entry of the verb khordan "to eat". The fact that ghaza can make a complex unit with the verb as we see in the (a-d) examples further supports this idea. The theta role assigned to ghaza is assigned direct but the theta role of ali is compositional through $\langle V2^0, \text{ghaza} \rangle$. Verbs assign structural case to their internal arguments and in their absence to the external ones in exceptional cases. Taking these theoretical points into consideration, it is ghaza that must receive accusative case, and it does so as is shown by the DOM, ra. As ali, the subject of the embedded verb can not receive NOM case either, it is demoted to a lower syntactic position. It can either receive oblique case from a preposition as (16) shows, or Gen case as (15). Recalling Fukui's (1986) theory of projection, it is the external argument that lacks case/Kase in the IP/VP system and requires a case assigner in order to project into syntactic structure.

Li predicts that the embedded VP can not contain a negative element or tense. We may add that the embedded VP can not be finite. As our discussion of future tense in Farsi showed the embedded VP may be tensed but must be non-finite. As far as his claim about negation is concerned, it is true for Farsi data, because in Farsi Neg morpheme is of the type that takes a tensed-complement, Negp-1, as discussed by Zanuttini (1991). But in a language like English that has two types of negations: the preverbal negative marker that takes a tense-complement as n,t which always precedes the verb; and the post verbal negative marker as not which takes non-finite, non tensed complement, Neg², and may follow or precede the verb, there is no reason that post verbal negation should be prohibited from occurring with the embedded VP, as our examples of English clearly indicates:

- (32) a. I made him go.
 b. I made him not go.
 c. * I made him n,t go.
 d. * I made him don,t go.

(32b) shows that the embedded VP may contain a negative marker if the negation is of the post verbal type that takes a non-tensed/non-finite complement. Negation in Farsi is of type 1 so it can take only a tensed clause as complement. Thus it complies with Li's prediction as far as causative structures in Farsi are concerned, but for future tense, which I argued to have a tensed but yet non-finite embedded VPs, we may either change Zanuttini's type one negation slightly to apply to finite clauses and/or VPs in Farsi, or claim that it is not possible to have two negation markers for the same proposition in the language. The following Farsi data is yet worth considering where in (33 b) the matrix verb carries the negation but in (33 c) the embedded tensed but non-finite verb has the negation marker. All examples are in future tense.

- (33) a. man ketab ra khah-am avard
 I the book DOM prs-1st-S ps
 I will bring the book.
- b. man ketab ra na-khah-am avard
 I the book DOM neg-prs-1st-S ps
 I will not bring the book.
- c.* man ketab ra khah-am na-y-avard
 I will the book DOM prs-1st-S neg-ps
 I will not bring the book.

(33 c) is marked as ungrammatical because it is never used and is very marked and odd if it were used. Yet it does not seem to me to be totally unacceptable. I remember hearing sentences similar to it from the speeches of the late leader of our country in recent years, but it seems to be very archaic. (33 c) is not as ungrammatical as (34) below where the Neg marker is attached to the auxiliary and not to the verb:

- (34) ** Man ketab ra avarde na-bud-am
 I the book DOM ppr neg-ps-1st-S
 I had not brought the book

Comparing these two with ((15 c) section (3)) which is grammatical may lead us to suggest the following amendment to Li's proposal: the embedded VP may be tensed but must be non-finite, one capable of containing a type two negation, Neg², but not a type one negation, Neg¹ (which may be possible in very marked, and archaic situations as examples (15 c (3)) and (33 c) show).

4-PASSIVES

Passive constructions are rarely used in Farsi. They are not common. I will argue that passive constructions in Farsi, in particular the past participle of the main lexical verbs (PP) in passive constructions, function like compound verbs which do not raise to T⁰/Agr⁰ while lexical verbs do so in order to assign ACC case to their arguments. The neg marker and aspect marker (Asp) which head their own functional projections in my analysis, and prefix in Farsi only to the lexical verbs, do not function so in passive but attach to the so called passive auxiliary shodan "to become".

Moyne (1974) correctly claims that the equivalent of an English sentence like (35) is simply expressed in active in Farsi:

(35) John was seen by Mary.

(36) Mary John ra did
 Mary John DOM ps/see-3rd-S
 Mary saw John.

the by-phrase constructions in English are equivalent to instrumental ones in Farsi. Moyne (1974) names the so-called passive constructions in Farsi, pseudo-compounds which have instrumental agents, if any, otherwise agentless. His study is mainly based on a transformational approach of the Aspects model, but the conclusions he makes exactly match the generalizations I would like to make on the basis of the verb raising, and VP structure. Farsi has a very rich and productive verb compounding in which nouns, adjectives, PPs, and prepositions combine with certain verbs such as kardan "to do", shodan "to become" which is also assumed to be the Aux of passive (similar to to be in English), gardidan "to turn, to become", zadan "hit", etc... Thus many intransitive compound can be made transitive, unaccusative, and ergative by changing the verbal element of the compound, as the examples below show:

- (37) a. dastgir kardan (tr)
 arrest INF/do "to arrest"
 b. dastgir shodan (intr)
 arrest INF/become
- (38) a. shekast dadan (tr)
 defeat INF/give "to defeat"
 b. shekast khordan (intr)
 defeat INF/eat "to be defeated"
- (39) a. ziyad kardan (tr)
 increase INF/do "to increase"

- b. ziyad shodan (intr)
increase INF/become "to become increased"

The (a) examples are all transitive verbs taking a direct object, but the (b) examples are their unaccusative equivalents that promote their internal arguments at S-S to the subject position for case reason because the relevant verbs are incapable of assigning ACC case to them as below:

- (40) a. polis ali ra dastgir kard
police ali DOM arrest ps/did-3rd-S
The police arrested Ali.
- b. ali dastgir shod
ali arrest ps/became-3rd-S
Ali was arrested.

To make (40 a) passive we will have (41) which is not in use at all, but is used in some eastern dialect of Iran and the kind of Farsi used in Afghanistan.

- (41) ali dastgir karde shod.
ali arrest ppr/done ps/became-3rd-S
Ali was arrested.

In Farsi the Neg (na-), progressive (PRG) (mi-) and the imperative/subjunctive (be-) prefixes always attach to the lexical verbs both in simple and compound verbs. The non-verbal element can separate from the verb and in many cases head their own phrases and take modifiers. Thus for (40 a-b) we can have (42 a-b) respectively:

- (42) a. poles ali ra dastgir na-kard /na-mi-kard
police ali DOM arrest Neg-ps-3rd-S /Neg-PRG-ps-3rd-S
The police did not arrest Ali/would not arrest Ali.
- b. ali dastgir na-shod /na-mi-shod
ali arrest Neg-ps-e-S /Neg-PRG-ps-3rd-S
Ali was not arrested/would not be arrested.

Given these facts, to passivize a sentence like (43 a) with a simple verb we will have (43 b) with the negative (43 c):

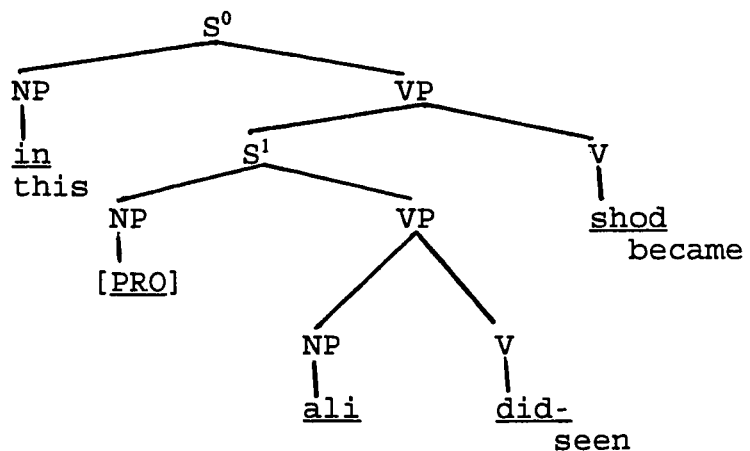
- (43) a. man ali ra mi-did-am
I ali DOM PRG-ps/saw-1st-S
I would see/ was seeing Ali.
- b. ali dide mi-shod
ali ppr/seen PRG-ps/became-3rd-S
Ali would be seen/ was being seen.

- c. ali dide na-mi-shod
 ali ppr/seen Neg-PRG-ps/became-3rd-S
 Ali would not be seen/ was not being seen.

Agent is not used in passive constructions in Farsi. As we notice from the above examples, the Neg and PRG markers attach to the passive verb shod "became", and not to the past participle of the lexical verbal stem which normally receives these affixes in active constructions. They pattern with the compound verbs (37-42). The only difference they have with the compounds is that the non-verbal elements of the compounds can in many cases separate from the verbal elements and act as direct objects or adjective phrases and head a nominal or adjectival projection but the past participles of the passives can not do so in any cases. That is why, I assume, that Moyne (1974) calls them pseudo-compounds and pseudo-passives. The past participles in Farsi are called esme maful "objective noun". They can in many cases be used as modifying adjectives. I suggest that past participles in Farsi make compound verbs with shodan on par with the other compound verbs. Moyne (1974) suggests the deep structure (44 b) for a sentence like (44 a). He also names passives "inchoative structures". His reason for proposing (44 b) is that such sentences are usually agentless but do have subjects. He thus assigns a complex structure similar to the causative structure proposed by Moyne & Cardan (1974) for passives:

- (44) a. ali dide shod
 ali ppr/seen ps/became-3rd-S
 Ali was seen.

- (44) b.

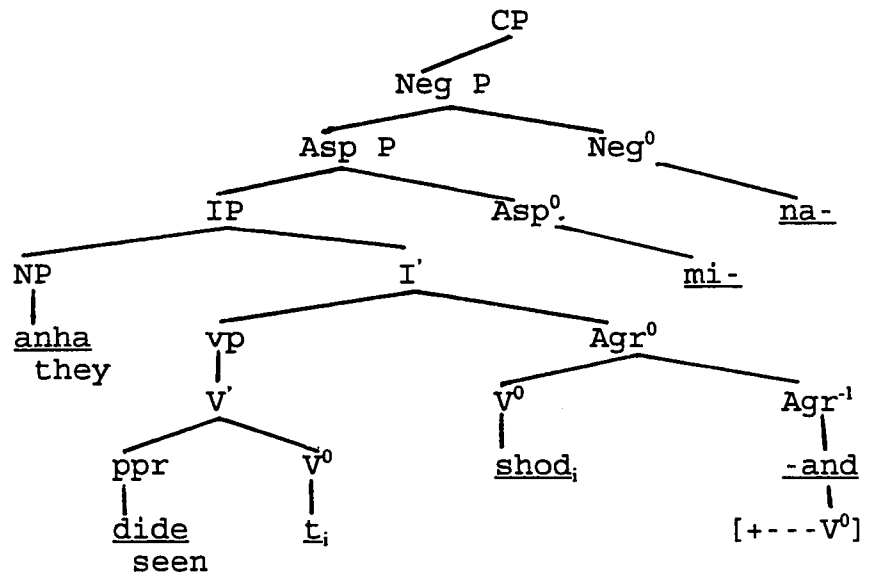


In (44 b) PRO is the absent agent, in is an expletive. It looks very similar to our analysis of future and causatives where a matrix verb took a VP complement if we leave out the node S¹. I see no reason for such an analysis, and assuming in "this" which functions as a pronoun and adjective is an unmotivated one. He suggests that the lexical verb did- the past stem of didan "to see" is in the D-structure position and when it raises to V of the higher verb, it turns into a ppr, and PRO moves to replace the

expletive in. The problem with his analysis is that if PRO raises to higher NP, then it must receive nominative case from the Agr⁰ of the matrix verb, and because the passive verb can not assign ACC case, ali remains caseless and violates case filter. I assign an active construction to the so-called passives in Farsi on par with simple and compound active verbs. A sentence like (45 a) will have the structure (45 b) in my proposal:

- (45) a. anha dide na-mi-shod-and
 they ppr/seen Neg-PRG-ps-3rd-pl
 They were not being seen.

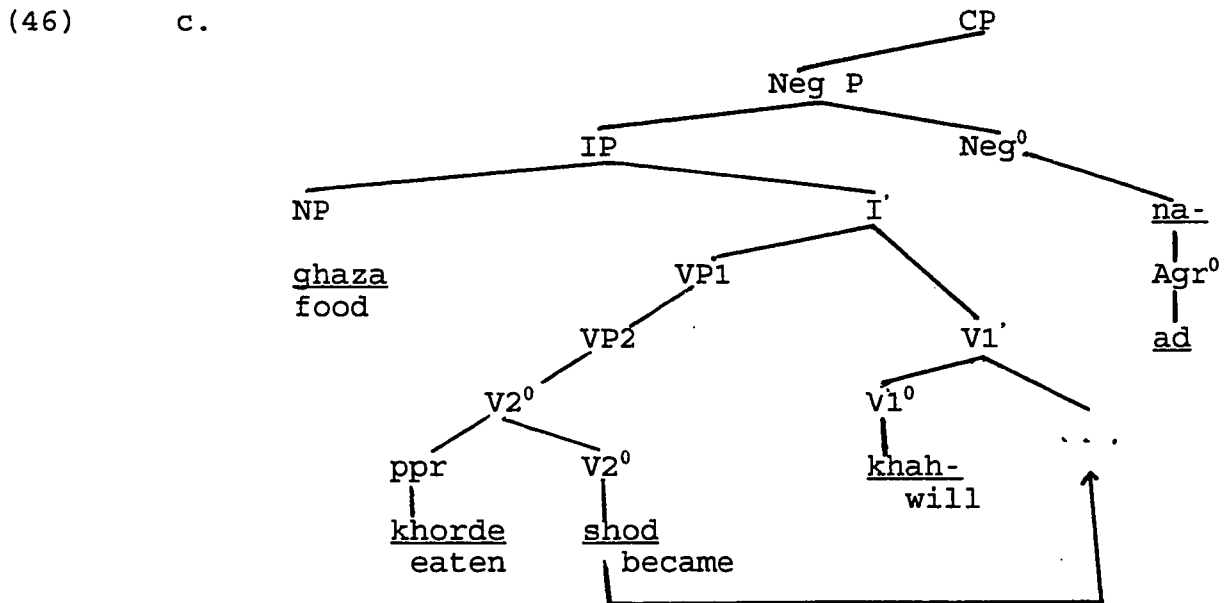
(45) b.



In (45 b) shod is the head of the compound verb dide shod(-and) as the configuration clearly indicates. V⁰ of the compound, shod, raises to agr⁰, by which it is morphologically selected (see Roberts (1991) and our discussion of future and causatives.), leaving the ppr, dide, behind. Its trace can be properly antecedent governed. In the second step of the movement the compound formed under Agr⁰ raises to Asp⁰ (morphological selection as before) where mi- prefixes to the lexical element bearing the features [+V, -N]. In the last step of the movement Neg⁰ triggers raising of the compound formed in the previous step, and attaches to mi-shod, giving na-mi-shod. My analysis of passive constructions in comparison with the equivalent active constructions explained in sections (2 & 3), indicates that shod the Ps of the so called passive auxiliary verb shodan, is the head of the compound verb dide shod-and in (45 b) and patterns with active simple and compound constructions. The past participle in passive constructions is a mere adjective and not the lexical V⁰ of the passive compound. Had it been the head, it should have raised to T⁰/Agr⁰, since verb raising to these positions is obligatory, but it (the past participle of the passive verb) should not assign case because it is not a lexical head; and

the compound verb behaves as a transitive verb. It should have also received Neg^0 and Asp^0 prefixes. The theme in my proposal is base generated as an external argument under the subject position, and receives NOM by spec-head agreement. To conclude the passive construction, notice the example (46 a) and its passive equivalent (46 b) and the tree diagram (46 c). The passive counterpart is not common but completely acceptable and grammatical.

- (46) a. ali ghaza ra khah-ad khord
 ali the food DOM prs-3rd-S ps/eat
 Ali will eat the food.
- b. ghaza khorde na-khah-ad shod
 the food PP neg-prs-3rd-S ps/become
 The food will not be eaten.



(46) is an instance of future in Farsi. As discussed before, the matrix verb khah- subcategorizes for an embedded VP. The head of VP2 in our example is shod. As the surface structure of (46) clearly indicates the ppr of the compound is left behind and does not raise, and only the head of the compound shod has raised to $V1^0$. Had khorde shod raised together to $V1^0$ we would have khah-ad khorde shod and not khorde khah-ad shod. This clearly indicates that we are dealing with an active construction. The lexical verb (ppr khorde) need not raise because it is not a head verb any more; because it patterns with adjectives and has formed a compound. There is no direct object in (46 b-c), because direct object (in other words the internal argument) has been base generated under the NP-subject position. Even if we assume that the NP-direct object raises from its base position to the NP-subject position for case reasons, our analysis still holds. The trace left behind by the first step of movement is properly antecedent governed. In the

second step of movement the compound $V1^0$ ($V1^0 + V2^0$, khah shod) raises to Agr^0 where the affix attaches to the head of the compound, that is $V1^0$ khah leading to khah-ad shod, (see our discussion on future where this process of raising is fully discussed, and (4 a) in section (2) in particular). In the third step the compound under Agr^0 raises to Neg^0 where the negative prefix attaches to khah the lexical head of the compound leading to na-khah-ad shod. The suffix should obligatorily attach to $V1^0$, which counts as the lexical head of the extended projection in the future construction, bearing the features $\langle +V, -N, F_0 \rangle$. The fact that the Neg-P can not attach to $V2^0$ in (46) reminds us again of the prohibition of Neg-P in embedded VPs (with our own revision that the Neg-P must count as type one Negatives, $Negp-1$). The movement chain we have in (46 b-c) is e.g., $\langle V1, V2, Agr, Neg \rangle$. This kind of chain starting from lexical heads and leading to functional ones, is a proper chain as discussed before, (see also Li (1990)). The movements described are of the type head-to-head movement which fall under the extended projection of V-I system in Farsi, and thus are compatible with Grimshaw's (1991) theory.

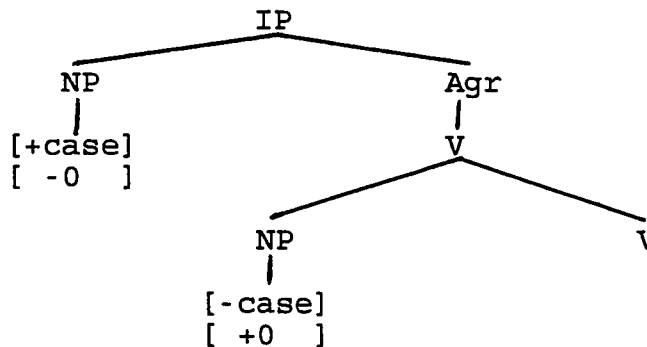
Thus I suggest that all verbal complexes formed with the verb shodan be considered as cases of ergative verbs which are formed in the lexical component. Ergativity being used to refer to constructions in which the verb does not assign a theta-role to its subject, although it may select a complement. According to the ergative analysis the NP-subject ali and ab "water" in (47 a-b) are base generated in the object position, as the internal argument of the verb, similar to the transitive examples (47 c-d), but are projected to the subject position by Burzio's generalization which claims that a verb fails to assign structural case if it fails to assign external theta role. Thus I assume that the external argument projects to the subject position in order to pass the case filter:

(47)

- (a) ali ghargh shod
ali drowned became
Ali became drowned/drowned.
- (b) ab garm shod
water hot became "The water became hot"
- (c) anha ali ra dar darya ghargh kardand
they ali DOM in sea drowned made
They made Ali drown/drowned (Ali) in the sea.
- (d) man ab ra garm kardam
I water hot made
I made the water hot.

The configuration (47 e) below indicates the status of ergative verbs, including all the complex verbs whose verbal element is shodan in Farsi:

(47 e)



We may conclude that all the verbal complexes formed from a N/ADJ/ppr + kard-an "to do" are transitive verbs in Farsi. By replacing the verbal element kard-an with shod-an we get the ergative equivalents. We may further complement our proposal by claiming that the past participles of simple transitive verbs may form ergative complex verbs with shodan if there is no semantic restriction. My contention in this section has been to prove that this process belongs to the lexicon and not to the syntax.

5-SUBJUNCTIVES & IMPERATIVES

Subjunctives and imperatives are formed by adding the prefix be- to the present stem of the verb, as examples below indicate:

- (48) a. shayad pro be madrase be-rav-am
perhaps pro to school Subj-prs/go-1S
Perhaps I go/will go to school.
- b. bayad pro ketab ra be-yar-i
must pro book DOM Subj-prs/bring-2S
You must bring the book.
- c. ketab-o bi-yar
book-DOM Imp-prs/bring
bring the book.

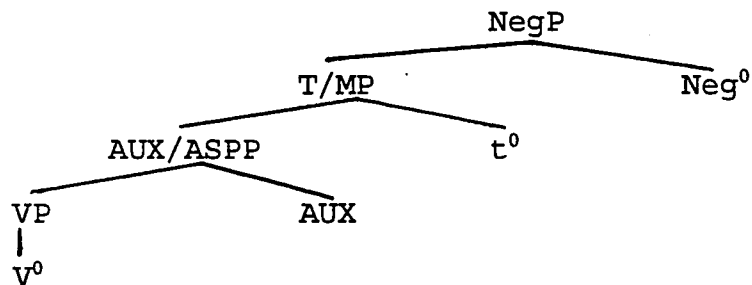
Subjunctives and imperatives are incompatible with negation, na- which takes a tensed VP as its complement, as explicated before. In the example below the Subj/Imp marker be- must drop before affixing the na- of negation:

- (49) a. shayad be madrase na-(*be-)-rav-im
perhaps to school Neg-(*Subj)-prs-1S
perhaps we will not go to school.
- b. ketab-o na-(*be-)-yar
book-DOM Neg-(*Imp)-prs
Do not bring the book.

In both examples in (49) the presence of be- causes ungrammatical-

ity as indicated by the star. In the section on negation I will argue that negation in Farsi is the type of Neg p-1 as proposed by Zanuttini (1991). It is a preverbal negation and depends on the presence of a functional projection of tense (TP) over the VP. What I would like to suggest based on this assumption, is that tense in Farsi occurs at the same level of projection as Mood (subjunctive, Imperative, and Indicative). Now a non-indicative mood precludes the concurrent expression of tense on the same level. be- being the typical sign of non-indicative mood (subjunctive & imperative) projects at the same level as tense and when the V^0 of the prs raises to TP/MP (mood projection) there is no abstract tense morpheme, rather be-. It attaches to the present stem of the verb. Thus there is no abstract tense morphemes with these two moods. Thus I am suggesting that in Farsi TP/MP are morphosyntactically represented in the same head positions. See Mitchel (1991) for a similar situation in Finnish. Based on the discussion of AUXP/ASPP I suggest that these are embedded within T/MP in Farsi. While the former is always concurrent with the presence of TP, MP precludes occurrence of both TP and AUXP/ASPP, see (50) below. There seems to be a loose end in my discussion. There is a verbal construction in Farsi called Mazi-ye eltezami "Past Subjunctive", as (51) below, which is formed from the present stem of the AUX budan "to be", (that is bash-), and the ppr of the lexical verb.

(50)



(51) a. anha bayad be madrase rafte bash- and
 they must to school ppr/gone prs/be
 They must have gone to school.

b. anha (na-)bayad be madrase (na-)rafte bash- and

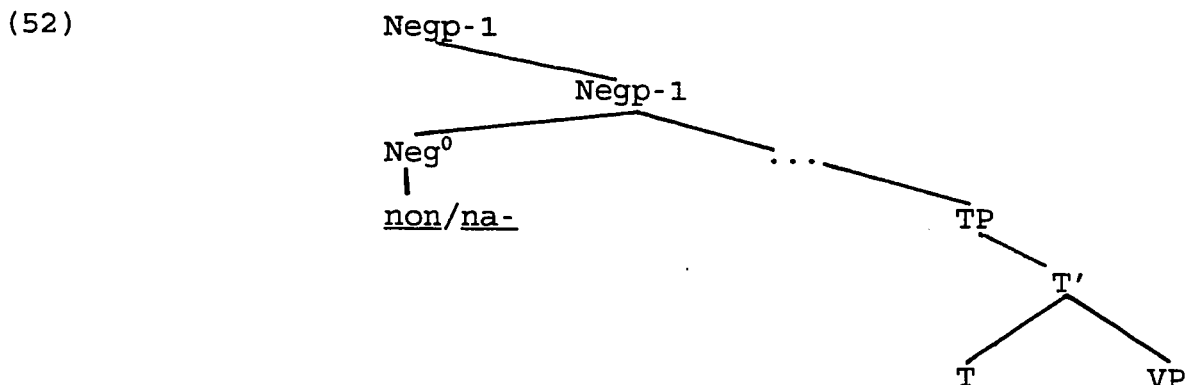
My suggestion based on the distribution of negative marker, na-, which attaches to ppr of the main verb, and the lack of subjunctive marker be- in (51) is that this kind of constructions do not count as real subjunctives. As we argued before the ppr contains a Tm. The AUX, that is bash- in (51), lacks the subjunctive marker be, and counts as the prs stem of the AUX bud-an "to be" which receives the abstract Tm when it raises to TP. As a result tense is present in this construction, and constructions like those in (51) do not function as subjunctive mood in Farsi. It is exactly on a par with Future, except for the case that the higher verb, bash, selects a past participle and not a past stem, since bash is a real

AUX according to our analysis while the main, higher of the future is a modal-like verb.

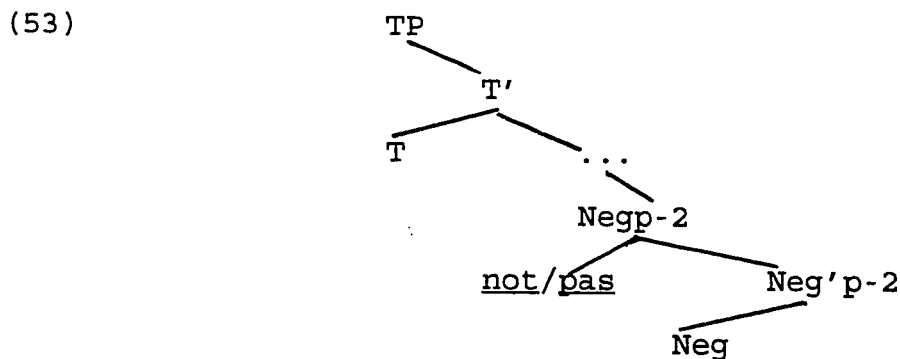
6-NEGATION

I have discussed the functions of negation in every section of this chapter and in previous chapters as well in some detail. In this section I will argue that negation in Farsi is of the type that takes a VP/TP as its complement. Zanuttini (1991) proposes two types of negations as below:

(A)-Preverbal NegP, or Negp-1, which strictly depends on the presence of a functional projection of tense (TP). Negp-1 can only appear if TP is present in the verbal system. In other words Negp-1 either directly selects a TP as complement, or selects a VP that contains a TP. On the surface structure it normally precedes the finite/tensed verb as in Farsi na-, Italian non, English n't, and French ne, as (52) below shows:



(B)-The postverbal NegP, Negp-2, which does not show similar dependencies on the presence of TP, hence appears after the finite verb, cooccurring mainly with non-finite elements such as French pas, English not, as (53) shows:



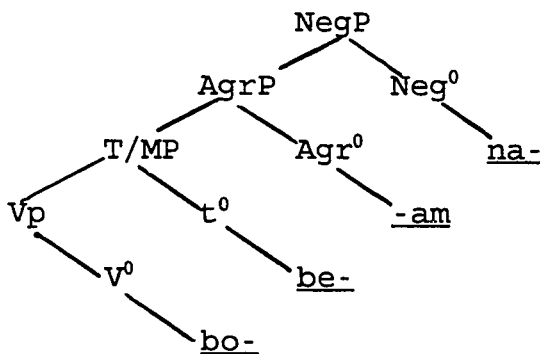
In this section I will explain why negation in Farsi is incompatible with subjunctive/imperative, and certain derived nominals and adjectives which contain a verbal root, but occurs

with some others which contain a TP/VP. Notice (54) below:

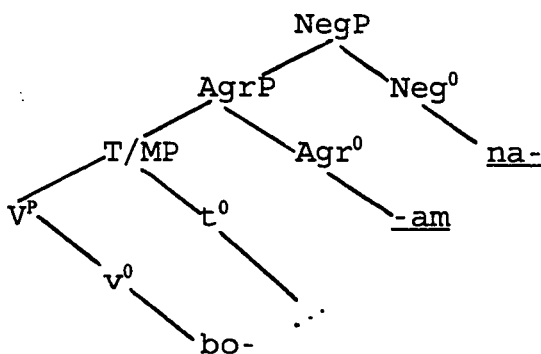
- (54) a. bo-ro
Imp-prs/go "go"
- b. na-ro
Neg-prs/go "do not go"
- c.* na-bo-ro
- d. shayad pro be-rav-am
perhaps pro Subj-prs/go-1S
Perhaps I will go.
- e. shayad pro na-rav-am
- f.* shayad pro na-be-rav-am

I follow Pollock (1989) in assuming the existence of an independent Functional Projection (FP), NegP for negation in Farsi: it is of the type 1 Negp-1 above. We already argued that be- is a sign of Subj/Imp Mood which projects at the same level as tense, thus precluding the presence of tense. Therefore when be- is present, there is no tense on the verb, but when be- is absent the abstract Tm is there. V⁰ the prs of the verb raises to t⁰ and receives it, on its way to Agr⁰ and Neg⁰, as (55) below indicates:

- (55) a.



- b.



Remind also that the prs of the verb is not unique to Subj/Imp moods, rather it is used for present indicative, as well. Thus when be- is not there, the prs stem raises to t^0 and receives the abstract Tm, so the selectional property of Neg is satisfied. As a general conclusion I may say that negation takes a tensed VP complement and always attaches to the lexical verb. The point confirms our previous argument that AUXP must be outside of VP. Had it been inside VP, we could not understand why Neg only attaches to the lexical verb and not to AUX. Negation also helped to show that there is no real passive in Farsi. We also noticed that the so-called "past subjunctive" can not be considered a real one, in other words the "past subjunctive" does contain a Tm as evidenced by the absence of be- and presence of AUX and TP.

Let us now consider infinitives, gerunds, and ppr as (56) below shows:

- (56) a. forukh-t-an
ps/sell-Tm-INF "to sell"
- b. na-forukh-t-an-e kala
Neg-ps/sell-Tm-INF-EZ goods
Not to sell goods. / Not selling goods

(56 a-b) indicate that there is a Tm in INFs in the language, which is absorbed by the Neg. Now consider the ppr which is formed from the past stem of the verb plus the Tm, plus the morpheme -e, as (57) below indicates:

- (57) a. forukh-t-an----> forukh-t---->forukh-t-e
ps/sell-Tm-INF sold
- b. pusi-d-an----> pusi-d----> pusi-d-e
ps/decay-Tm-INF decayed

As past participle contains a tense morpheme they must be compatible with negation as (58) below confirms:

- (58) a. mive-ha-ye pusid-e va na-pusid-e ro joda kon
fruit-PL-EZ decayed and neg-deccayed DOM separate make
Separate the decayed fruits from the non-decayed ones.
- b. ?ash-e na-khord-e , dahan-e sukht-e
soup-EZ Neg-ppr/eaten mouth-EZ ppr/burned
Has not eaten the soup, but burned the mouth (a proverb)

Thus past participles in Farsi may be used as adjectives as well, and as such may be negated since they contain a tense morpheme.

Let us now consider (59) below with a gerund, where a V^0/VP is dominated by an NP, but there is no tense morpheme:

- (59) a. forush-e kala
prs-EZ goods "selling goods"
- b.* na-forush-e kala
Neg-prs-EZ goods
- c. bin-esh-e siyasi
prs-GRm-EZ political "political observation"
- d.* na-bin-esh-e siyasi

Forush is the present stem (prs) of the verb forukhtan "to sell" which also functions as a gerund. There is no abstract tense morpheme in its structure, thus (59 b) is bad. This also indicates that recognition of an abstract tense morpheme in this thesis is a correct one. (59 c) is formed from prs+ -esh (GRm). It can not be used with Neg as (59 d) shows. Let us finish this section with a problematic example:

- (60) a. kharid-an-e kala
INF/to buy-EZ goods "to buy goods"
- b. na-kharid-an-e kala
Neg-INF/to buy-EZ goods
- c. khar-id-e kala
ps/buy-Tm-EZ goods "buying goods"
- d.* na-khar-id-e kala

(60 a-b) are good and compatible with our discussion so far. (60 c-d) seem to have the same verbal structure as (60 a-b). It seems to have been composed of ps+ Tm, but why then it is incompatible with Neg. It particularly seems to have the same tense morpheme, (-i)d, as (60 a-b). I assume that kharid is not a tensed NP, but a non-tensed gerund, with an independent lexical entry. It patterns with gerunds in Farsi, and can make a compound verb/noun as in kharid kardan "to do shopping" and kharid-o-forush "buying and selling". It is not (most probably) a tensed verbal-compound of the form ps+ Tm, rather it functions as an independent gerund as we notice by its parallelism with a gerund above, forush. It might have been formed from the past stem (ps) of the verb and a gerund forming morpheme, (-i)d, which happens to be homophonous with the tense morpheme.

7- THE STRUCTURE OF IP

Based on our arguments so far, I suggest (61) below as the schematic structure of IP in Farsi. In doing so I take the Mirror Principle of Baker (1988) into consideration according to which the order of affixes reflects the syntactic derivation of the words. The lower Wh/QP is the X max of a WhP/QP to which a quantifier and

Footnotes of chapter IV

(1) - ?ast "is" is derived from the AUX ?astan "to be". It is used only for the present perfect. It is mainly used as a copula in present tense.

- a. hava sard ?ast
 weather cold is "The weather is cold"
- b. have sard_e
 weather cold-is "The weather is cold"
- c. hava sard bud
 weather cold was "The weather was cold"

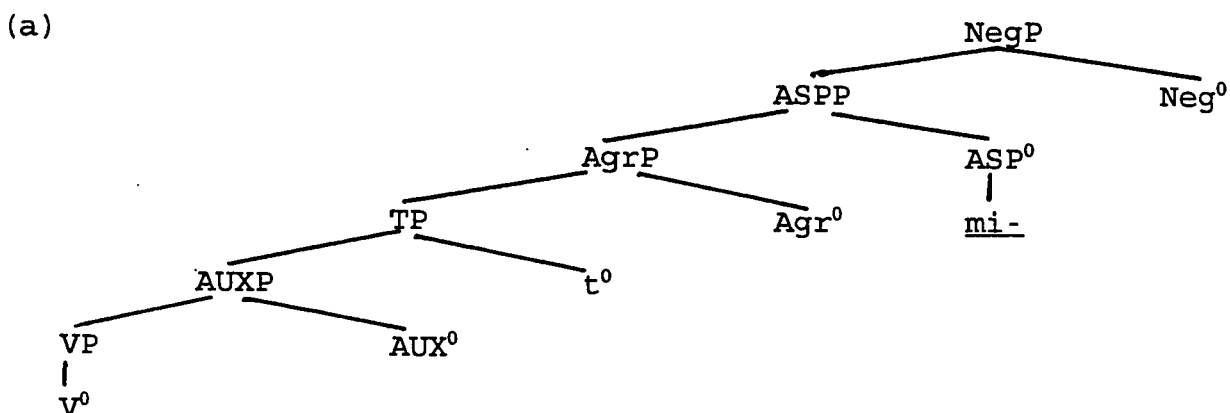
While ?astan is used as the present form of the copula budan or/and ?astan, budan may be used in the past tense only. Whether we take ?astan as a separate AUX from budan or as an archaic variant of it which is also used as the AUX of the present perfect tense does not affect our arguments in this theses. The other forms of ?astan are as below:

- d. rafte-?am (-am)
 ppr/gone-AUX "I have gone"
- e. rafte-?i (-I) "you have gone"
- f. rafte ?ast (-e) "He has gone"
- g. rafte-?im (-im) "We have gone"
- h. rafte-?id (-id) "You have gone"
- i. rafte-?and (-an) "They have gone"

The forms in the parentheses are the colloquial ones.

(2) - The suggestion that the past participle of the main verb moves up to the AUXP and forms a compound with the auxiliary can be accounted for by the observation that in verb preposing to IP initial position and/or to COMP (according to Karimi (1989)) in Farsi, it is the whole verbal complex that is the main verb together with the auxiliary and/or modal like elements like khastan "to want" (which functions as a modal to form the future tense in the language) which move to the initial position of the clause, and not just the main or the auxiliary. Furthermore, in our system where the NegP is located high up in the tree it is necessary for the main verb to move up to NegP through AUXP.

(3) - It may also be possible to assume that mi- projects to its own independent maximal projection below the NegP but above the AgrP as (a) below shows:



In this system we can not account for incompatibility of AUXP with ASPP, that is mi. There is a special clausal construction where a modal-like verb, dasht-an "to have", which carries its own TP/AGRP selects another CP/IP. This subordinate CP/IP contains a verb which must obligatorily contain a mi- in its verbal system.

- (a) man dasht-am ba ajale mi-raft-am
 I ps/had-1S with hurry ASP-ps/went-1S
 I was going hurriedly.
- (c) anha dar-and ba ajale mi-rav-and
 They prs/have-3PL with hurry ASPP-prs/go-3PL
 They are going hurriedly.

The constructions with dashtan never occur with embedded CPs which contain an AUX. These observations, I assume, indicate that incompatibility between AUXP and ASPP (mi-) may be attributed to the identical levels of projection.

(4) - bo- is the phonologically conditioned equivalent of be-, the subjunctive/imperative prefix.

(5) This does not seem to me to be an arbitrary condition since the causative morpheme not being a lexical element lacks the capacity to assign accusative case, and the NOM case is only assigned by the Agr which is associated with it. The fact that it prevents the embedded verb from assigning case may be attributed to a property of verbs in Farsi which we hinted upon in chapter III. It seems that verbs in Farsi may assign case only when they are dominated by and/or associated with a TP. We noticed this property in the study of infinitives and gerunds in chapter III. We noticed that in all the finite verbs there is a TP associated with the verb. If this argument is correct, then we see that the embedded verbs in these constructions are not directly associated with a TP, hence may not assign case to their arguments. But when they raise to V1, that is the causative morpheme, they become directly associated with the TP which dominates the causative morpheme, and hence the new

amalgamated compound. They may now assign case to their respective arguments.

BIBLIOGRAPHY

- Alberston, S. (1991) "La Loi Tobler-Musafia en Ancien Francias et en Ancien Italian", M. A. Thesis, university of Ottawa.
- Baker, M. (1988) "Incorporation: A Theory of Grammatical Function Changing", University of Chicago Press, Chicago, Illinois.
- Baker, M. & Hale, K. (1990) "Relativized Minimality and Pronoun Incorporation", LI 21, 289-297.
- Belletti, A. (1988) "the Case of Unaccusatives", LI 19, 1-34.
- Bhatt, R. & Yoon, J. (1990) "On the Composition of COMP and Parameters of V-2", WCCFL 8, 42-52.
- Chomsky, N. (1965) "Aspects of the Theory of Syntax", The MIT Press, Cambridge.
- Chomsky, N. (1970) "Remarks on Nominalization", in R. Jacobs and P. Rosenbaum (eds.) readings in English Transformational Grammar, Waltham, Mass.
- Chomsky, N. (1981) "Lectures in Government and Binding", Foris Publications, Dordrecht, Holland.
- Dabir-Moghaddam, M. (1986) "Sakhthaye Sababi in Farsi" (Causative Constructions in Farsi), Journal of Iranian Linguistics 5, 13-76.
- Enc, M. (1991) "The Semantics of Specificity", LI 22, 1-25.
- Fukui, N. (1986) "A Theory of Category Projection and its Application", Doctoral dissertation, MIT, Cambridge.
- Grimshaw, J. (1990) "Extended Projection", Brandis University.
- Hashemipour, P. (1989) "Pronominalization and Control in Modern Persian", Phd. dissertation, University of California, San Diago.
- Hoekstra, T. (1984) "Transitivity: Grammatical Relation in Government & Binding Theory", Foris Publications.
- Karimi, S. (1989) "Aspects of Persian syntax, Specificity, and the Theory of Grammar", Phd dissertation, University of Washington, U. M. I.
- Larson, R. K. (1985) "Bare-NP Adverbs", LI 16, 595-621.
- Li, Y. (1990) "X⁰-Binding and Verb Incorporation", LI 21, 399-426.
- Mitchell, E. (1991) "Evidence from Finnish for Pollock's Theory of IP", LI 22, 373-379.
- Moyne, J. (1974) "The So-Called Passive in Persian", Foundation of Language 12, 249-267.
- Moyne, J. & Carden, G. (1974) "Subject Reduplication in Persian", LI 5, 205-249.
- Noonan, M. (1989) "Operator Licensing and the Case of French Interrogatives", in the proceedings of the English WCCFL 8, 315-330.
- Ouhalla, J. (1990) "Sentential Negation, Relativised Minimality and the Aspectual Status of Auxiliaries", The Linguistic Review 7, 183-231.
- Pollock, J.-Y. (1989) "Verb Movement, Universal Grammar, and the Structure of IP", LI 20, 365-424.
- Rivero, M-L. (1990) " The Location of Non-Active Voice in Albanian

- and Modern Greek", LI .
- Rivero, M-L. (1992) "Adverb Incorporation and the Syntax of Adverbs in Modern Greek", University of Ottawa, to appear in Linguistics and Philosophy.
- Rivero, M-L. (1991) "Long Head Movement and Negation: Serbo-Croatian vs Slovak and Czech",
- Rizzi, L. (1986) "Null Objects in Italian and the Theory of Pro", LI 17, 501-557.
- Rizzi, L. (1990) "Relativized Minimality", Linguistic Inquiry Monograph.
- Roberts, I. (1990) "Excorporation and Minimality", LI 22, 209-218.
- Samian, V. (1983) "Origins of Phrasal Categories in Persian, an X-bar Analysis", Phd dissertation, UCCLA.
- Speas, M. J. (1990) "Phrase Structure in the Natural Language", Kluwart Academic Publishers, Dordrecht.
- Stowell, T. (1981) "Origins of Phrase Structure", Phd dissertation, MIT.
- Stowell. T. (1983) "Subject Across Categories", The Linguistic Review 2, .