



National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file - Votre référence

Our file - Notre référence

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

**INTERROGATIVE CONSTRUCTIONS
IN PERSIAN**

M.A THESIS

SUPERVISOR: PROFESSOR P. HIRSCHBUHLER

UNIVERSITY OF OTTAWA

SHAHLA RAGHIBDOUST



National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file *Voire référence*

Our file *Notre référence*

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-89678-7

Canada



UNIVERSITÉ D'OTTAWA
UNIVERSITY OF OTTAWA

This thesis treats the various constructions of interrogativity in colloquial modern Persian, within the Government and Binding framework. Chapter 2 presents the basic properties and the various methods of forming yes-no questions. I argue that changes in word-order may not be considered as a strategy to indicate an interrogativity, and that consequently, the movement of the verb from its position in I to C in Persian, unlike languages such as English, is disallowed. I also propose the possibility of assigning a mood phrase (MP) position to the yes-no particles.

Chapter 3 gives a detailed survey of the movement processes of Wh-words which, *prima facie*, appear to be a syntactic movement. However, exploring the more complicated data, we subsequently arrive at the conclusion that this movement, by and large, patterns with an optional topicalization process, and has nothing to do with the syntactic movement to SPEC CP. I propose that the availability of question particles in a number of languages, including Persian, correlates with the lack of syntactic Wh-movement. Persian extraction patterns are argued to resemble those of Nordic languages in which no structural constraint is imposed. Subjacency, therefore, is unable to explain the extraction phenomenon in this language. To give a reasonable account of the extraction rules in Persian Dominance, in turn, is proposed as a convincing condition.

Chapter 4 concentrates on the fact that multiple Wh-fronting in Persian is radically different from extensively-studied languages, such as Bulgarian and Romanian, even though these languages manifest resemblance with respect to a number of properties. Furthermore, it is shown that in multiple Wh-fronting languages, the Wh-phrases are morphologically complex, and need to satisfy a licensing requirement independent of clausal typing. This morphological characteristic is absent, however, in Persian. Analysis of the preceding factors leads me to propose that multiple Wh-fronting in Persian also results from adjunction of Wh-phrases to IP, in other words, topicalization.

ACKNOWLEDGEMENTS

I would like to take this opportunity to express my gratitude and appreciation to the many people who have had a part in my graduate studies and in the writing of this thesis. First of all, it has been my great fortune to have had as my thesis supervisor Dr. Paul Hirschbuhler to whom I express my deepest thanks for enormous generosity in offering me his time and assistance. His inspired perception, insights and suggestions, have benefited me tremendously. I appreciate especially his amazing patience in listening to my half-cooked (and even crude) analyses. I am also grateful to all my teachers at the University of Ottawa from whom I have profited and to a number of fellow students with whom I have had the pleasure of studying.

The pursuit of my studies, and consequently the completion of this thesis would not have been possible without the emotional and intellectual support of my relatives. I am greatly indebted to my mother for all her love, encouragement and devotion. I am thankful to my father who was so anxious for the academic improvement of his children. My special thanks to my brothers Shahram and Ali Reza, my sister Shima and my sister-in-law Sangeeta, for their constant moral support over the years. I owe a great deal to my husband Fereidoun without whose invaluable and enduring support, this research could never have been undertaken. I would like to dedicate this thesis to him. And finally, I thank my dear sons Amir Hossein and Iman who understood their mother's situation and gave completely from their hearts.

TABLE OF CONTENTS

	PAGE
CHAPTER 1.....	1
Introduction.....	1
Theoretical framework.....	2
CHAPTER 2.....	4
Yes-no questions.....	4
Description of yes-no questions in main clauses.....	4
Yes-no questions with "aya".....	9
A-not-A questions.....	14
Description of yes-no questions in subordinate clauses.....	15
Notes.....	25
CHAPTER 3.....	26
Description of Wh-questions.....	26
Direct questions.....	26
Typology of Wh languages.....	28
Embedded questions.....	31
Cheng's (1991) analysis.....	33
Surface position of Wh-topic.....	39
Characterizing Wh-movement.....	40
The gap feature.....	43
Bridges.....	44
Subjacency Condition.....	45
Extractions out of embedded questions.....	47
Extractions out of complex noun phrases.....	58
Extractions out of relative clauses.....	59
Extractions out of noun complements.....	71
Notes.....	74
CHAPTER 4.....	75
Multiple Wh-fronting	75
Historical review of the literature.....	75
Multiple Wh-fronting languages.....	78
Multiple Wh-extraction from a clause.....	83
Wh-islands.....	86
Constituent structure of fronted Wh-words....	88
Wh-words.....	91
Summary.....	95
Toward an account.....	96
Wh/indefinites in multiple fronting languages.....	97
Multiple Wh-fronting structure in Persian....	100
Notes.....	103
Bibliography.....	105

Chapter 1

Introduction

This thesis deals with the study of question structures in modern Persian, within recent Government and Binding theory. My main focus will be on colloquial informal Persian as spoken by people in their ordinary daily contacts. Since the colloquial register displays a more natural and spontaneous form of speech, the processes and changes taking place in the language could be traced in a clearer way. Our examination, therefore may lead to a more profound and reasonable account of the issues under investigation. Though, it should be pointed out that informal and formal Persian do not differ substantially from each other with regard to the relevant topic, as many of the syntactic changes occurring primarily in colloquial Persian have been transferred to the formal form.

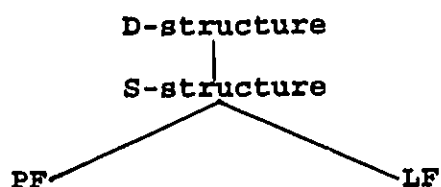
A few scholars have taken a look at some aspects of question formation in Persian from the traditional, structural and generative-transformational points of view, but as far as can be ascertained, practically no comprehensive and systematic treatment of the interrogatives within GB syntax has yet been made.

The aims of the present study are twofold: a) To propose a description and account of the basic characteristics of Persian interrogatives on the basis of recent hypotheses. b) To represent the essential points which have been made by others, compare them with the findings in this research, and thus provide a critical analysis. Although I am mainly concerned with questions, I will

consider examples involving other types of constructions such as relative clauses whenever necessary.

Theoretical framework

The syntactic framework which I adopt is that of GB which assumes a Y-model of grammar:



Phrase structures are in accordance with X-bar theory and rules of lexical insertion. The map from D-structure to S-structure is provided by transformations which have the general form "Move alpha". S-structure feeds into the phonological as well as the semantic component of the grammar. On the semantic side, a level of logical form is assumed at which, among other things, expressions corresponding to logical operators are assigned scope. This involves movement of quantified NPs and unmoved Wh from argument positions to appropriate operator positions. LF is thus derived from S-structure through transformations in the same sense as S-structure is derived from D-structure.

Chomsky (1977) has argued that the question construction shares the following properties with the relative construction:

- a) It involves a Wh-word in COMP.
- b) It has a gap.
- c) The relation between COMP and the gap obeys Subjacency.
- d) It appears to be unbounded.

On the other hand, questions differ from relatives in a number of respects:

- a) Questions can contain more than one Wh-word.
- b) There are questions that do not involve movement.
- c) Questions involve the full range of Wh-expressions, while relatives involve only Wh NPs and Wh PPs; In other words, relatives

modify nouns, and questions do not.

All of these properties lead to a general theory of "Move Wh", initiated by Chomsky (1973), and further developed in Chomsky (1977a).

Having made explicit the broad theoretical assumptions behind this work, I will proceed focussing on three major types of Persian interrogative structures: yes-no questions, Wh-questions and multiple Wh-fronting structures in the following three chapters respectively.

Chapter 2

YES-NO QUESTIONS

This section presents a study of the basic properties of yes-no questions in Persian.

I will start by describing the different ways of making yes-no questions, as opposed to declaratives.

At the beginning, yes-no questions will be examined in main clauses and the various types of constructions that express them will be described. I will then survey yes-no questions in subordinate clauses.

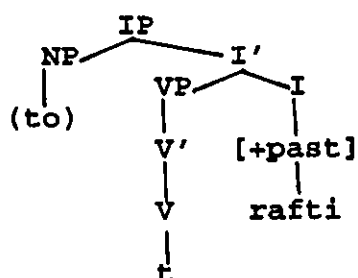
1. Description of yes-no questions in main clauses

As a starting point, let us consider the following declarative sentences and their relevant tree diagram representations:

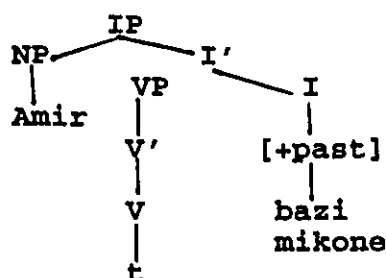
- (1)
- a. to/pro raft-i.
you/pro went-2S
"you went."
 - b. amir bazi mi-kon-e.
Amir play CM-do-3S
"Amir is playing."
 - c. sanaz ketab-ro xund.
Sanaz book -DO read
"Sanaz read the book."

(CM and DO stand for continuous marker and direct object respectively).

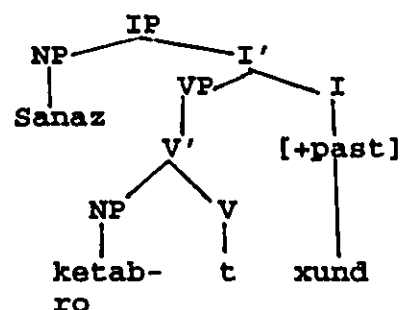
(2a)



(2b)



(2c)



The examples above suffice to illustrate that Persian is a SOV language, i.e. that its VP is head-final. In sentence (1a), the intransitive verb *rafti* "went" assigns the agent theta-role to the pronominal NP *to*, the external argument, and in the case of its absence to the non-overt *pro*. As Persian is a pro-drop language, the pronominal subject could, therefore, be deleted freely. Similarly, the compound intransitive verb *bazi mikone* "is playing" takes the NP *Amir* as its external argument. Sentence (1c) contains the transitive verb *xund* "read" which takes the internal argument NP *ketab*, and the NP *Sanaz* corresponds to its external argument. Following the discussion in Vahedi (1992) I adopt the view that in Persian, verb raises to I, leaving a trace.

Now consider the interrogative forms of the above sentences:

- (3) a. *to/pro raft-i ?*
 you/pro went-2S
 "did you go?"
- b. *amir bazi mi-kon-e ?*
 Amir play CM-do-3S
 "is Amir playing?"
- c. *sanaz ketab-ro xund ?*
 Sanaz book -DO read
 "did Sanaz read the book?"

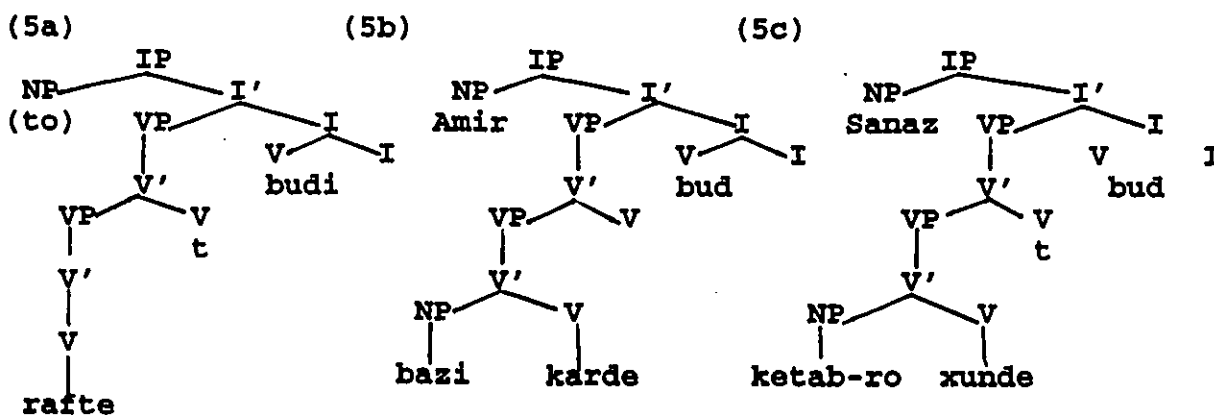
Sentences (3a-3c) indicate the simplest form of yes-no questions in the colloquial form of the Persian language.

The comparison of the declarative sentences (1) with their interrogative counterparts (3) reveals that originally there are

no differences between the positions of the syntactic elements of the subject and the verb in the above declarative and interrogative sentences. Changes in word order on their own are never sufficient as a strategy to indicate an interrogativity in Persian. On this basis, it seems reasonable to argue that the tree representations of these declarative and interrogative forms would be identical. The difference just lies in the intonation contour with no grammatical changes involved. That is to say, the utterance of the declarative sentence is with a falling pitch, while the question form is pronounced with a sharply raising one. So, unlike all verbs in French, and auxiliaries and modals in English, the movement of the verb from its position in I to C is not allowed in Persian interrogative constructions. This also holds for the auxiliary verb as illustrated in examples (4) below:

- (4) a. to/pro rafte bud-i ?
 you/pro gone was-2S
 "had you gone?"
- b. amir bazi karde bud ?
 Amir play done was
 "had Amir played?"
- c. sanaz ketab-ro xunde bud ?
 Sanaz book -DO read was
 "had Sanaz read the book?"

The relevant syntactic representations of the examples (4) are as follows, assuming that the auxiliary bud is the head of a VP and that it is limited to raising to I in questions:



It should be pointed out that the auxiliary as a categorically independent projection might project an aspectual phrase outside the projection of VP. In other words, the base position of the auxiliary, could be in an ASPP, rather than VP (see Ouhalla (1990) for a more detailed discussion). Once the auxiliary has moved to I, and contrary to what happens for the movement of the auxiliary in English, Persian does not allow the auxiliary verb to move from the I position to the position dominated by C through head-to-head movement. Perhaps it is the case that the movement of the auxiliary to the C node is illicit in Persian for some reason to be determined.

Two possibilities present themselves:

- a) The complementizer position does not exist in Persian. If this is the case, it implies that there is no CP.
- b) The complementizer position exists, but it is inaccessible to I (and V) in the syntax, although one could speculate that it is available as a landing site for these elements at LF.

Clearly something like a complementizer position exists in Persian, as demonstrated by the fact that embedded tensed clauses are introduced by the complementizer *ke*. It would then appear at the present time that the second possibility is probably the one to be explored. I will thus assume that some factor prevents I to C in direct questions.

It is worth mentioning that when the subject NP appears in the sentence, its placement could be either the preverbal or, to a lesser degree, the postverbal position of the sentence. In agreement with Karimi (1989) and disagreement with Hashemipour (1989), therefore, I believe that the subject NP can be preceded by the verb.¹ The postverbal position of the subject may be observed in the non-interrogative as well as the interrogative forms, as is shown in the following examples:

Preverbal subject:

- (6) a. to rafte bud-i ?
you gone was-2S
"had you gone?"

- b. sanaz ketab-ro xund ?
 Sanaz book -DO read
 "did Sanaz read the book?"

Postverbal subject:

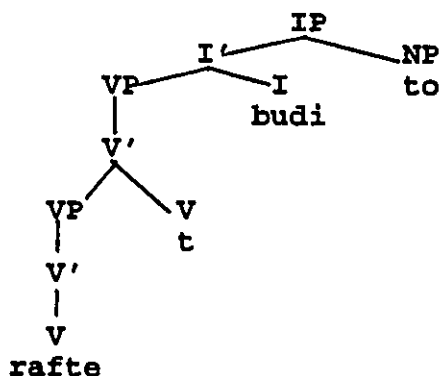
- (7) a. rafte bud-i to ?
 gone was-2S you
 "had you gone?"

- b. ketab-ro xund Sanaz ?
 book -DO read Sanaz
 "did Sanaz read the book?"

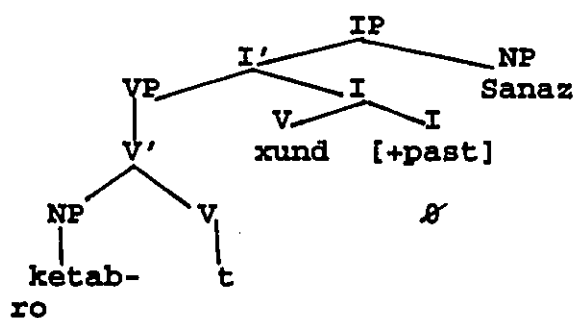
So, without the intonation characteristics of the questions, examples (6a-b) and (7a-b) also have the assertion reading. From the discussion above, we therefore conclude that intonation plays a large part in interrogativity in Persian.

I assume that the syntactic representations of the examples (7) would be as in (8). In our proposed tree diagram (8a), the auxiliary in accordance with the standard assumption, is generated inside the VP predicates:

(8a)



(8b)



Notice that the postverbal subject phenomenon is related to the free word order characteristic of different syntactic categories of Persian. Therefore in cases such as (6-7) where we see subject-verb inversion, it is not the verb that moves, the subject position has been changed from preverbal to postverbal position of the verb, as the tree diagrams (8) indicate. Furthermore, this claim could be confirmed by cases such as (7b) which contain a direct object, and where the relative order of direct object and inflected verb is

the same as in (6b).

Yes-no questions with "aya"

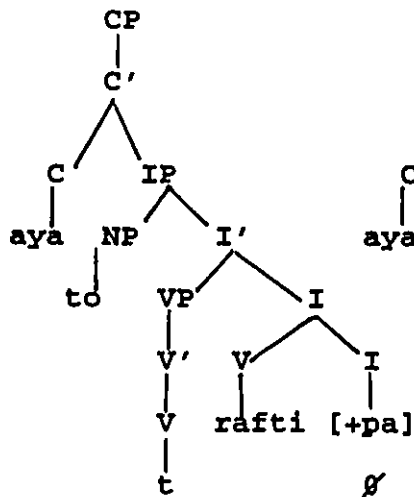
There are certain question structures which, in addition to a raising tone, also have interrogative elements. One variation for this type can be found in formal style of Persian. Consider for instance (9). (In the glosses below, INT.M=Interrogative marker):

- (9) a. aya to/pro raft-i ?
 INT.M you/pro went-2S
 "did you go?"
- b. aya amir bazi mi-kon-e ?
 INT.M Amir play CM-do-3S
 "is Amir playing?"
- c. aya sanaz ketab-ro xunde bud ?
 INT.M Sanaz book -DO read was
 "had Sanaz read the book?"

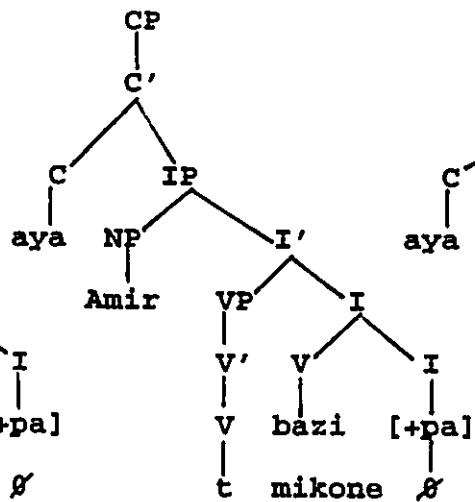
Aya is an interrogative marker which is most frequently used in literary and formal yes-no questions of Persian, and is not usually observed in informal speech. No specific meaning can be attributed to this phrase. Aya seems to represent more emphasis on the question and lacks a concept as other Wh-elements like čī "what", koja "where", key "when", čera "why" and četor "how" exhibit. A parallel interrogative phrase does not exist in English.

The important point concerns the placement of aya in the sentence structure of Persian. As in most cases of direct questions, aya occupies the sentence-initial position and is followed by the overt or covert subject (examples 9a-c), it would first be natural to assign aya to the position dominated by C in main clauses, as the syntactic representations (10) illustrate:

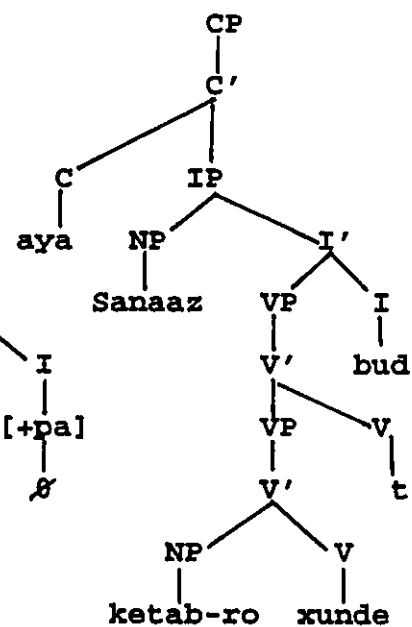
(10a)



(10b)



(10c)



If this were correct, then with respect to the question of why there is no inflected verb movement to the clause initial position, we would have or might have to say that although C is present, its properties are incompatible for it to be a landing site for I (i.e. the inflected verb). Additional data from indirect questions suggest, however, that the position occupied by *aya* should not be identical with C.

Furthermore, a more precise examination shows that this interrogative lexeme, is also allowed in other positions of the sentence, and does not lead to ungrammaticality. Syntactic freedom of word order in Persian permits permutation of the constituents of a sentence. The following acceptable examples indicate that *aya* appears in different positions freely and its occurrence is not restricted to a specific syntactic position:

- (11) a. *amir aya bazi mi-kon-e ?*
 Amir INT.M play CM-do-3S
 "is Amir playing?"
- b. *sanaz aya ketab-ro xunde bud ?*
 Sanaz INT.M book -DO read was
 "had Sanaz read the book?"

c. sanaz ketab-ro aya xunde bud ?
 Sanaz book -DO INT.M read was
 "had Sanaz read the book?"

d. amir dust-eš-ro dar park aya molaqat kard?
 Amir friend-her-DO in park INT.M meet did
 "did Amir meet his friend in the park?"

Although aya may occur in different positions, in my opinion, the sentence which is introduced by aya represents the most traditional and formal form of yes-no questions. In other words, in a question with the particle aya, the unmarked order involves aya being in first position.

In addition, it should be pointed out that the use of aya in postverbal position, i.e. in final position, although controversial, seems to be possible:

(12) a. amir bazi mi-kon-e aya ?
 Amir play CM-know-3S INT.M
 "is Amir playing ?"

b. sanaz ketab-o xunde bud aya ?
 Sanaz book-DO read was INT.M
 "had Sanaz read the book?"

Further evidence for this claim is provided by the example below which is a sample of the Iran's contemporary poetry:

(13) xoda-ya to ba an bozorgi dar an asman-ha čenin
 God-oh you with that magnitude in that sky-PL such
 arezu-yi be-din kučaki ra tavn-i baravard aya ?
 wish-a to-this smallness DO can-2S grant INT.M
 "Oh God! You with that magnitude, in those skies, could you grant such a small wish?"²

Also, from the sentences below, we find out that even when there is aya in a clause, the movement of NP subject to postverbal position is not blocked. So, the examples (14) would be as grammatical as (11):

(14) a. aya bazi mi-kon-e amir ?
 INT.M play CM-do-3S Amir
 "is Amir playing?"

- b. aya ketab-ro xunde bud sanaz ?
 INT.M book-DO read was Sanaz
 "had Sanaz read the book?"
- c. aya dust-eš-o dar park molaqat kard amir ?
 INT.M friend-his-DO in park meet did Amir
 "did Amir meet his friend in the park?"

From the sentences above, we can make the deduction that a fixed placement can not be assigned to this question morpheme. Aya as a free-position marker may appear everywhere in a sentence, and from a semantic point of view, displays more emphasis on the question.

Meanwhile, there are two additional alternative ways of constructing yes-no questions which involve either the phrase *mage* or *hič*.³ Although *mage* and *hič*, unlike *aya*, belong primarily to informal Persian, their syntactic behaviour resembles that of *aya*. That means, *mage* and *hič* may occur in all possible syntactic positions in which *aya* does. Thus, we should find the arrangement of the grammatical data in (15):

- (15) a. mage to raft-i ?
 INT.M you went-2S
- b. hič to raft-i ?
 INT.M you went-2S
- c. to raft-i mage ?
 you went-2S INT.M
- d. to raft-i hic ?
 you went-2S INT.M
 "did you go?"
- e. mage sanaz ketab-o xunde bud ?
 INT.M Sanaz book-DO read was
- f. hič sanaz ketab-o xunde bud ?
 INT.M Sanaz book-DO read was
- g. sanaz mage ketab-o xunde bud ?
 Sanaz INT.M book-DO read was
- h. sanaz hic ketab-o xunde bud ?
 Sanaz INT.M book-DO read was

- i. sanaz ketab-o mage xunde bud ?
 sanaz book-DO INT.M read was
- j. sanaz ketab-o hič xunde bud ?
 Sanaz book-DO INT.M read was
- k. sanaz ketab-o xunde bud mage ?
 Sanaz book-DO read was INT.M
- l. sanaz ketab-o xunde bud hič ?
 Sanaz book-DO read was INT.M
 "had sanaz read the book?"

It should be pointed out that the above statement regarding the syntactic similarity of the interrogative words in question constructions is not tantamount to saying that all such questions in Persian have an identical semantic interpretation. The choice of one or the other question phrases depends, in fact, on the attitude or the intention of the speaker. *Hič*, compared to *mage*, seems to function more or less semantically like *aya*. A question involving either *aya* or *hič* indicates that the speaker is not sure about the answer and he is inquiring about something to which the answer can be yes/no. However, an utterance containing *hič* appears to be more emphasized than one containing *aya*. With respect to *mage*, the matter is different, as the speaker actually feels certain about the answer and he merely puts the question to the listener, say, with the intention of having it confirmed. Thus, concerning (15 e/g/i), for example, the speaker implies that: "I know that Sanaz had not read the book, but I want to have it confirmed". Due to the different semantic properties, *mage* and *hič* do not behave similarly in those structures which involve *ya na*. The appearance of *mage* is incompatible with *ya na*:

- (16) a.* *mage to raft-i ya na ?*
 INT.M you went-2S or not
 "did you go or not?"
- b.* *mage sanaz ketab-o xunde bud ya na?*
 INT.M Sanaz book-DO read was or not
 "had Sanaz read the book or not?"

Hič, on the other hand, appears to act in a similar way to *aya*.

Nonetheless, as stated earlier, its semantic function can not be taken to be exactly the same as that of *aya*:

- (17) a. *hič to raft-i ya na?*
 INT.M you went-2S or not
 "did you go or not?"
- b. *hič sanaz ketab-o xunde bud ya na ?*
 INT.M Sanaz book-DO read was or not
 "had Sanaz read the book or not?"

A-Not-A Questions

To make our discussion in this part more complete, it should be mentioned that in all cases of yes-no questions, the interrogative sentence can end in the phrase *ya na* "or not", whether "*aya*" is present or not. In other words, the occurrence of these phrases in a clause does not depend on each other. Any of them can appear in the question without imposing any obligation on the appearance of the other one. Persian like Mandarin Chinese (see Li and Thompson (1984), and Huang (1981/82), therefore has grammaticized the disjunction of an affirmative clause and its negative counterpart as a device to form a question.⁴ Semantically, *ya na* makes more explicit the alternative typical of a yes-no question which results in more emphasis. Under no condition, is the appearance of *ya na* obligatory. We shall see in the following subsection that the occurrence of *ya na* in indirect questions compared with its occurrence in direct interrogatives, is more desirable and required. Consider the sentence (18) below which are taken to be a variation of the examples (3) and (9):

- (18) a. *to raft-i ya na ?*
 you went-2S or not
- b. *aya to raft-i ya na ?*
 INT.M you went-2S or not
- c. *amir bazi mi-kon-e ya na ?*
 Amir play CM-do-3S or not
- d. *aya amir bazi mi-kon-e ya na ?*
 INT.M Amir play CM-do-3S or not

e. sanaz ketab-o xunde bud ya na ?
 Sanaz book-DO read was or not

f. aya sanaz ketab-o xunde bud ya na ?
 INT.M Sanaz book-DO read was or not

The only context in which the appearance of *ya na* is permitted, is the interrogative clause. *Ya na* is not used in declarative sentences, as the following ungrammatical examples indicate:

(19) * amir xořhal-e ya na.
 Amir happy-is or not
 "Amir is happy or not."

Saha (1984) in an argument about different methods of creating question forms in various languages, has pointed out that "using a disjunctive question (A-not-A question) could be viewed as one of the techniques of constructing yes-no questions in some languages. For example in Chinese, it is possible to ask questions whose structure may be represented in English by the example "he is or isn't a leader?":

(20=82) ni xihuan tade chenshan bu xihuan ?
 you like his/her shirt NEG like
 "do you like his/her shirt?"

Note that while *aya* lacks a fixed position in syntactic representation, this is not the case with *ya na*. The only possible position in which *ya na* is placed, is the last position of the clause, as the grammatical examples (21) illustrate:

(21) a. sanaz aya ketab-o xunde bud ya na ?
 Sanaz INT.M book-DO read was or not
 b.*sanaz aya ketab-o ya na xunde bud ?
 c.*sanaz aya ya na ketab-o xunde bud ?
 "had Sanaz read the book or not?"

2. Description of yes-no questions in subordinate clauses

In this section, I will attempt to shed light on the various characteristics of yes-no questions in subordinate clauses. Consider the following examples:

(22)

- a. man ne-mi-dun-am (ke) aya (to) raft-i ya na.
 I NEG-CM-know-1S (that INT.M (you) went-2S or not
 "I wonder/don't know whether you left."
- b. amir ne-mi-dun-e (ke) aya pedar dust-eš-o mi-bin-e
 Amir NEG-CM-know-3S(that)INT.M father friend-his-DO CM-see
 ya na.
 or not
 "Amir wonders/doesn't know whether father will see his
 friend or not."
- c. sanaz ne-mi-dun-e (ke) aya dar emtehan-a-š
 Sanaz NEG-CM-know-3S (that) INT.M in exam-PL-her
 movaffaq mi-š-e ya na.
 successful CM-become-3S or not
 "Sanaz wonders/doesn't know whether she will pass her exams
 or not."

(22a-c) exhibit the most formal types of the subordinate clauses in Persian in which the interrogative phrase *aya* is used in the initial part of the embedded clause.⁵ Likewise, regarding the clause-final position of *ya na*, an A-not-A question may not only be a direct question as shown earlier by (15), it may also be indirect. The appearance of the conjunction *ke* which is normally placed under C, indicate that *aya* should be assigned a different position, other than C. In order to consider a position for the interrogative marker *aya*, I would like to hypothesize, following Bhatt and Yoon (1991), that the morpheme *aya* functions as a mood-marker. According to them, all types of clauses, including matrix clauses should possess a COMP node. Based on the recent logic of giving each functional feature its separate projection (Pollock (1989), Chomsky (1989)) and the usual assumption about crosslinguistic uniformity, they claim that the category COMP should be dissociated into a category that indicates clause-type, or Mood, and for subordinate clauses, a category of subordinators. The lexical complementizer system of some languages such as English conflates/lexicalizes the two categories, while in languages like Korean, Japanese and Kashmiri, they are kept apart. Hence, Bhatt & Yoon believe that the clause-type or mood-marking is required universally in all clauses. However, as just indicated, languages

may differ in the way in which these distinctions are marked. Given these analyses, *aya* seems to function as a mood-marker. Although Persian unlike languages like Korean and Kashmiri, does not systematically possess separate lexical categories of mood-markers, concerning *aya* we can take this account and provisionally assign *aya* to the mood phrase (MP) position.

Meanwhile, the examples (23) indicate the free-position characteristic of *aya* in embedded clauses. *Aya* in the data below is preceded by the subject, direct object and prepositional phrase respectively:

(23)

- a. man ne-mi-dun-am (ke) to aya raft-i ya na.
I NEG-CM-know-1S (that) you INT.M went-2S or not
"I wonder whether you left or not."
- b. amir ne-mi-dun-e (ke) pedar dust-eš-o aya mi-bin-e
Amir NEG-CM-know-3S (that) father friend-his-DO INT.M CM-see-3S
ya na.
or not
"Amir doesn't know whether father will see his friend or not."
- c. sanaz ne-mi-dun-e (ke) dar emtehan-a-š aya movaffaq
Sanaz NEG-CM-know-3S (that) in exam-PL-her INT.M successful
mi-š-e ya na.
CM-become or not
"Sanaz doesn't know whether she will pass her exams or not."

In addition, the phrase *aya* could be inserted between the two elements of the compound verb. Examples (24) shows that *aya* occurs after the NP element of the compound verb and is followed by the lexical verb:

- (24) sanaz ne-mi-dun-e (ke) dar emtehan-a-s movaffaq aya
Sanaz NEG-CM-know-3S (that) in exam-PL-her successful INT.M
mi-s-e ya na.
CM-became-3S or not.
"Sanaz doesn't know whether she will pass her exams or not."

Let us now consider the occurrence of *aya* in postverbal position, next to the phrase *ya na*, as shown in (25):

(25)

- a.? man ne-mi-dun-am (ke) (to) raft-i aya ya na.
I NEG-CM-know-1S (that) (you) went-2S INT.M or not
- b.? amir ne-mi-dun-e (ke) pedar dust-eš-o mi-bin-e
Amir NEG-CM-know-3S (that) father friend-his-DO CM-see-3S
aya ya na.
INT.M or not
- c.? sanaz ne-mi-dun-e (ke) dar emtehan-a-š movaffaq
Sanaz NEG-CM-know-3S (that) in exam-PL-her successful
mi-š-e aya ya na.
CM-become-3S INT.M or not

The examples in (25) are marked as not being perfect. Yet they don't seem to be totally unacceptable. Perhaps, the lower degree of acceptability of such cases could be attributed to the non-syntactic property of intonational feature. By that I mean that the adjacency of the elements *aya* and *ya na* yield to a heavy and undesirable intonation.

Given the variety of positions that *aya* can occupy, it is not clear that in those examples where *aya* occurs in clause-initial position, it should be considered as dominated by a category like MP. At this time, I have no specific proposal to make with regard to this problem.

The phrase *ya na*, on the other hand, appears in the final position of the same clause. As shown by the parentheses, there is no necessity for the occurrence of the complementizer *ke* in this construction: in both cases of the absence or appearance of the *ke*, the sentence would be equally acceptable and grammatical. In fact, *ke* seems to appear more frequently in formal style and is usually preferred not to be present in colloquial speech. However, I think that the function of *ke* in distinguishing the formal speech from the informal one, is not distinctive and significant: the choice of *ke* is in fact arbitrary in both formal and informal forms.

Actually the formal characteristic of the sentence is mostly dependent on the appearance of the *aya*, as was the case for the main interrogative clauses as well. That is to say, in formal

Persian, the sentence is required to be initiated with *aya*, and in turn, the use of *aya* in colloquial form seems to be relatively odd for native speakers. The phrase *aya* could be used optionally, but its absence makes the context more informal. Hence, *aya* could be referred to as an indicator of the formal and literary Persian interrogative structures.

The following examples in which *aya* is absent, are the informal counterparts of the sentences (22), which in turn, represent the most common form of the indirect questions in present Persian:

(26)

- a. man ne-mi-dun-am (ke) (to) raft-i ya na.
I NEG-CM-know-1S (that) (that) went-2S or not
"I wonder whether you left."
- b. amir ne-mi-dun-e (ke) pedar dust-eš-o mi-bin-e
Amir NEG-CM-know (that) father friend-his-DO CM-see-3S
ya na.
or not
"Amir doesn't know whether father will see his friend or not."
- c. sanaz ne-mi-dun-e (ke) dar emtehan-a-š movaffaq
Sanaz NEG-CM-know-3S (that) in exam-PL-her successful
mi-š-e ya na.
CM-become-3S or not
"Sanaz doesn't know whether she will pass her exams or not."

Ya na, as well as the complementizer *ke*, can show up in both formal and informal subordinate clauses. In contrast with *aya*, the use of the phrase *ya na* in such clauses is more or less obligatory, and in the case of its absence, the sentence though almost grammatical, seems to be incomplete. As illustrated in examples (27):

(27)

- a.? man ne-mi-dun-am (ke) aya (to) raft-i.
I NEG-CM-know-1S (that) INT.M (you) went-2S
"I wonder whether you left."
- b.? amir ne-mi-dun-e (ke) aya pedar dust-eš-o
Amir NEG-CM-know-3S (that) INT.M fatherfriend-his-DO
mi-bin-e .
CM-see-3S
"Amir doesn't know whether father will see his friend."
- c.? sanaz ne-mi-dun-e (ke) aya dar emtehan-a-š movaffaq

Sanaz NEG-CM-know-3S (that) INT.M in exam-PL-her successful
 mi-š-e ya na.
 CM-become-3S or not
 "Sanaz doesn't know whether she will pass her exams."

Finally, as shown by the sentences (28), both of the phrases *aya* and *ya na* are deleted from the subordinate clauses simultaneously. In these contexts, as we have seen in constructions (27), the speaker is expected to make his statement complete by adding the phrase *ya na* to the sentence. Nevertheless, it could not be claimed that the omission of the two elements *aya* and *ya na* renders the sentence fully unacceptable. Its adequacy, compared with the previous cases, is more restricted, but not rare, with the proper intonation:

(28)

a.? man ne-mi-dun-am (ke) (to) raft-i.
 I NEG-CM-know-1S (that) (you) went-2S
 "I wonder whether you left."

b.? amir ne-mi-dun-e (ke) peadr dust-eš-o mi-bin-e.
 Amir NEG-CM-know-3S (that) father friend-her-DO CM-see-3S
 "Amir doesn't know whether father will see his friend."

c.? sanaz ne-mi-dun-e (ke) dar emtehan-a-š movaffaq
 Sanaz NEG-CM-know-3S (that) in exam-PL-her successful
 mi-š-e ya na.
 CM-become-3S or not
 "Sanaz doesn't know whether she will pass her exams."

Thus, in the last two cases (27,28), the indirect question is characterized by a strong stress on the last syllable of the verb. Such a stress is absent in sentences (22,26). With respect to (28), an even stronger stress is required, since in (27), the subordinate clause is introduced by the lexeme *aya* which itself is a question phrase, and represents the interrogative characteristic of the clause, whereas there is no lexical indicator of the subordinate clause being a question in (28). So, the structures will be the same as that of the subordinate declarative clauses. In order to differentiate between these two identical constructions, the indirect question must be characterized by a different intonation, otherwise it would be indistinguishable from a subordinate declarative structure. As a result, the acceptability

of sentences such as (27,28) has much to do with intonation. If read with normal intonation, they usually sound bad as indirect questions, but properly emphasized they become relatively acceptable.

The summary of the conclusions of this section, with regard to the formal versus informal specification, could be proposed within the following schema:

(29)	Subordinate clauses	
	Formal	Informal
	a. aya.....ya na	
	b.ya na
	c.? aya.....	
	d.??

Coming back to the phrase *ya na*, it appears to play a more significant role than the marker *aya* in acceptability of the Persian subordinate clauses. While, the cases containing *ya na* - with or without *aya*- are considered fully grammatical, a relative acceptability holds for those structures lacking it. This seems to be the opposite of what occurs in English, where the appearance of "or not" is optional, whereas the absence of the complementizer "whether/if" leads to ungrammaticality, as shown by(30):⁵

- (30) a. I wonder whether/if you left or not.
 b. I wonder whether/if you left.
 c.* I wonder you left or not.

Generally speaking, the occurrence of *aya* is related straightforwardly to the appearance of *ya na*, while the opposite case is not true. That is to say, the adequacy of *ya na* does not necessitate the occurrence of *aya*.

We will now turn to another question. Remember that we have adopted the view that Persian is SOVI. However, we see that a clausal complement, for example an indirect question, appears on the right of the inflected verb. In order to account for this, we assume that a clausal complement is extraposed on the right of the inflected verb. The question that should be answered is where the

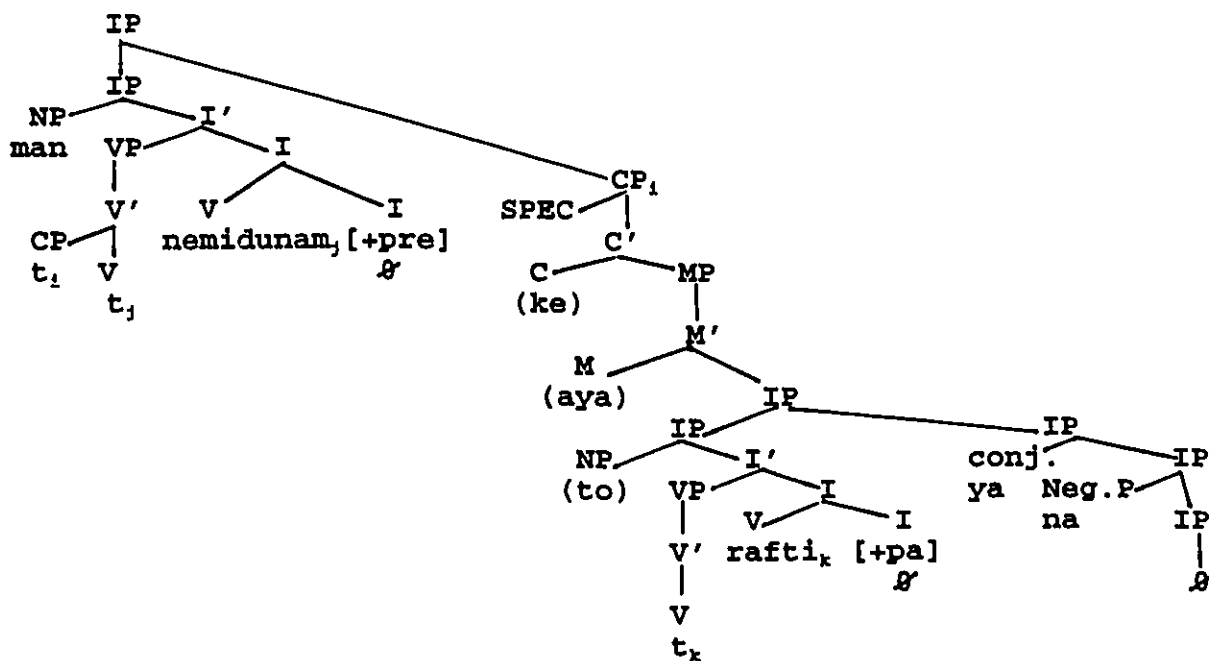
clause which is affected by movement is moved, i.e. its landing site. It might be assumed that the clause is extraposed under I or IP. But, as mentioned earlier, subject in Persian can be postposed to the verb. Thus, in relation to the above sentences of indirect questions, the subject intervenes between the verb of the main clause and the complement subordinate clause. This is shown by the examples (31):

(31)

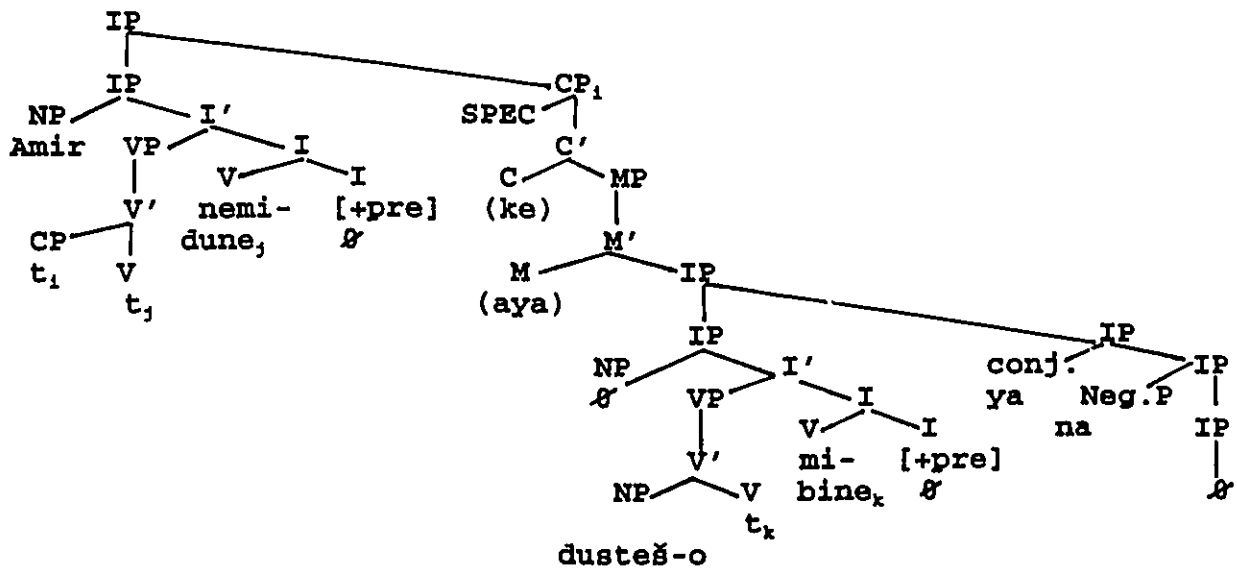
- a. ne-mi-dun-am man (ke) (aya) (to) raft-i ya na.
 NEG-CM-know-1S I (that) (INT.M)(you) went-2S or not
 "I wonder whether you left."
- b. ne-mi-dun-e amir (ke) (aya) pedar dust-eš-o
 NEG-CM-know-3S Amir (that)(INT.M) father friend-his-DO
 mi-bin-e ya na.
 CM-see-3S or not
 "Amir doesn't know whether father will see his friend or not."
- c. ne-mi-dun-e sanaz (ke) aya dar emtehan-a-š movaffaq
 NEG-CM-know-3S Sanaz (that) INT.M in exam-PL-her successful
 mi-š-e ya na.
 CM-become-3S or not
 "Sanaz doesn't know whether she will pass her exams or not."

We have assumed that when the subject is postverbal, I' and the subject are permuted: ($[_{IP} NP I'] \rightarrow [_{IP} I' NP]$). Given this, we will suggest that the extraposed clause, rather than moving to I' or IP, is adjoined to IP. In our proposed tree diagrams below, the phrase ya na, will be also assumed to behave as an IP projection, producing the structures like (32):

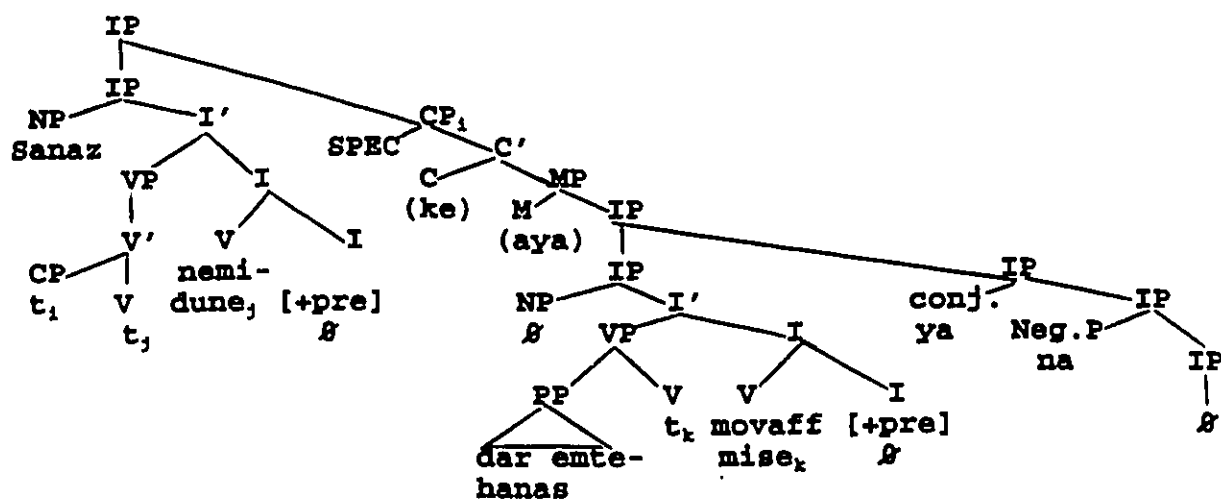
(32a)



(32b)



(32c)



Having looked at the different methods of forming yes-no questions in both main and embedded clauses, we will now examine the various characteristics of Wh-questions in the following chapter.

NOTES

1- Hashemipour (1989, p.267) argues that the subject and object NPs must precede the verb, and they never occur in post-verbal position. She gives the following examples and claims that (71c,d) are ungrammatical, which I think are perfectly well-formed and usual:

- (71) a. moin ketab-o tond xund.
 Moin book-DO fast read:3S
 "Moin read the book fast."
 b. moin tond ketab-o xund.
 c.* ketab-o tond xund moin.
 d.* moin tond xund ketab-o.

2- See shafii-ye Kadkani (M. Sereshk), 1347/1968: Az zaban-e barg.

3- Magar is the phonological variant of mage which is only observed in formal and literary Persian.

4- The difference between these two languages is that in Mandarin, the disjunctive morpheme haishi "or" is generally absent, whereas in Persian, ya, the disjunctive element is always present (For details, see Li & Thompson, 1981, 535-541).

5- With respect to the distribution, aya differs from its counterparts in English "whether/if" in that, the former may appear in both matrix and embedded questions, while the latter can always appear in embedded questions and they never appear in matrix yes-no questions:

* whether/if John is sleeping.

Chapter 3

Description of Wh-questions

1. Direct questions

We have already touched upon the various kinds of yes-no questions in main clauses. In this section, we give a general survey of Wh-questions. In this relation, the process of the movement of Wh-words and the different variants of Wh-interrogatives will be described.

At first, consider the typical instances (1) which illustrate the Wh-question in main clauses:

- (1)
- a. ki gerye mi-kard ?
 who cry CM-did
 "who was crying?"
 - b. to ki-yo/ro davat kard-i ?
 you who-DO invitation did-2S
 "who did you invite?"
 - c. sanaz či-yo-ro xeili dust dar-e ?
 Sanaz what-Do very like has-3S
 "what does Sanaz like very much?"
 - d. yasaman ba ki sobat mi-kon-e ?
 Yasaman with who talk CM-do-3S
 "to who is Yasaman talking?"
 - e. amir hafte-ye dige koja mi-r-e ?
 Amir week-EZ next where CM-go-3S
 "where will Amir go next week?"

The Wh-constituent *ki* "who" is the external argument of the verb in sentence (1a). *Ki* is the direct object of the verb *davat kardi* and part of the prepositional complement of the verb *sobat mikone* in (1b,d) respectively. Likewise, in (1c), *či* "what", the Wh-constituent functions as the direct object of the verb *dust dare* and (1e) represents the adjunct Wh-phrase *koja* "where" in the

position of a place adverb.

Let us now compare the above examples with their declarative counterparts in (2):

- (2)
- a. bačče gerye mi-kard.
 baby cry CM-did
 "the baby was crying."
 - b. to mehmun-a-ro davat kard-i.
 you guest-PL-DO invitation did-2S
 "you invited the guests."
 - c. sanaz arusak-o xeili dust dar-e.
 Sanaz doll-DO very like has-3S
 "Sanaz likes dolls very much."
 - d. yasaman ba sanaz sobat mi-kon-e.
 Yasaman with Sanaz talk CM-do-3S
 "Yasaman is talking with Sanaz."
 - e. amir hafte-ye dige tehran mi-r-e.
 Amir week-EZ next Tehran CM-go-3S
 "Amir will go to Tehran next week."

As the reader can verify, there is no structural difference depending on whether it is a subject or a non-subject that is questioned. The NPs in declarative structures are replaced by the Wh-constituents in interrogative phrases, and no obligatory word-order shifts to signal interrogative formation is observed. In other words, Persian Wh's are not fronted to sentence initial position, but remain in the same syntactic slots as do non-question words with the same grammatical function. Similar to what happened with respect to yes-no questions, the intonation is the only distinctive feature that sets the Wh-questions apart from the declaratives, besides the choice of the lexical items. Thus, the surface structure and the deep structure of the Wh-phrases will be identical and unlike English, the inflected verb is found under I in Persian interrogatives at surface structure and is not moved to the position dominated by C.

On the basis of examples like (2), we would think that Persian is an in-situ language, and we would expect the scopal properties to pattern like other in-situ languages, such as Chinese (see Li &

Thompson (1984) and Cheng (1991)).

A proper understanding of Persian questions leads us to have a brief review on current typology of Wh languages.

Typology of Wh languages

As a matter of fact, languages display considerable diversity in question formation strategies. In general, as both Cheng (1991) and Srivastav (1991), among others, have pointed out, two language types are recognized with respect to question formation strategies. In some languages, e.g. English, Wh questions show S-structure movement of the Wh-word, in contrast to those like Mandarin Chinese, in which the Wh-constituent stays in-situ at surface structure. The first is exemplified by English (3=Srivastav's 1a), the second by Chinese (3=Srivastav's 1b):

- (3=1a) a. what did Lisi buy?
 (3=1b) b. Lisi mai-le sheme?
 Lisi bought what
 "what did Lisi buy?"

Within the Government and Binding framework, question formation is characterized as the movement of a Wh-phrase to SPEC of COMP. A Wh-trace remains in the launching site. (3a), thus, is analyzed as the result of S-structure movement of "what" from an A position to A-bar position, namely to SPEC of CP. In contrast, (3b) appears to lack such movement. However, languages do not clearly divide into +/-Wh-movement types at all levels of representation. According to Huang's (1982) analysis, the difference between languages such as English and Chinese, is not in the presence or absence of Wh-movement, but in the level at which this movement takes place. Based on this argument, Wh expressions in any language are quantificational expressions which have to move to SPEC of CP in order to ensure interpretation. For languages, like English, this movement occurs at S-structure, whereas it happens at LF in Chinese. Given this account, the examples (3a,b), though distinctive at S-structure, are identical at LF:

(3) c. [_{CP} what [_{IP} Lisi buy t]

there is also another Wh-type of language which has been referred to by Cheng and Mahajan (1990). That is, some languages like Egyptian Arabic, Bahasa Indonesia, etc. are said to have the option of either leaving the Wh-word in-situ or having it in a clause-initial position. This type will be discussed in some detail in what follows.

Turning back to Persian, it is not immediately clear which of these patterns the interpretation of Persian interrogatives fits into. It was shown by sentences (1) that Persian Wh-words do not undergo fronting at S-structure. However, this is not the only mode of Wh-construction in this language. The examples in (4) alert us against taking Persian in terms of Chinese, that is to say, as an in-situ language:

- (4) a. ki-yo/ro to davat kard-i ?
 who-DO you invitation did-2S
 "who did you invite?"
- b. ci-yo/ro sanaz xeili dust dar-e ?
 what-DO Sanaz very like has-3S
 "what does Sanaz like very much?"
- c. ba ki yasaman sobat mi-kon-e ?
 with who Yasaman talk CM-do-3S
 "to who is Yasaman talking?"
- d. koja amir hafte-ye dige mi-re ?
 where Amir week-EZ next CM-ro
 "where will Amir go next week?"

In (4), the interrogative NPs *ki*, *či*, *ba ki* and *koja* have moved to the clause-initial position, and show that the movement of the Wh-elements is not blocked at S-structure. Persian at first glance, thus looks like "optional movement languages" because it allows typical Wh-in-situ's as well as fronting of Wh-words. Although, the in-situ Wh-words in examples like (1) may be taken by some speakers as more common and acceptable, I believe that sentences (4) are also completely unmarked and grammatical. Notice that (1) can, in addition, have word order variants of the kinds represented

below. Thus, we can have the following structures of the typical examples (1a,b,d). V1 and V2 represents the two constituents of the compound verb ([NP+V]) which can be separated from each other by inserting an NP in Persian:

- (5a) 1. gerye ki mi-kard ? (V1 S V2)
 cry who CM-did
 2. gerye mi-kard ki ? (V S)
- (5b) 1. (to) davat ki-yo/ro kard-i (S V1 DO V2)
 you invitation who-DO did-2S
 2. (to) davat kard-i ki-yo/ro ? (S V DO)
 3. ki-yo/ro davat kard-i to ? (DO V S)
 4. ki-yo/ro davat to kard-i ? (DO V1 S V2)
 5. davat kard-i (to) ki-yo/ro ? (V DO S)
 6. davat kard-i (to) ki-yo/ro ? (V S DO)
- (5c) 1. yasaman sobat ba ki mi-kon-e ? (S V1 IO V2)
 Yasaman talk with who CM-do-3S
 2. yasaman sobat mi-kon-e ba ki ? (S V IO)
 3. ba ki sobat mi-kon-e yasaman ? (IO V S)
 4. ba ki sobat yasaman mi-kon-e ? (IO V1 S V2)
 5. sobat mi-kon-e ba ki yasaman ? (V IO S)
 6. sobat mi-kon-e yasaman ba ki ? (V S IO)

To judge from our data, it appears that the word order possibilities allowed with non-questions are possible with questions too. Hence, a Wh-constituent is not confined to a specific syntactic position. The free placement characteristic of the question phrases is also observable in Hindi, as Mahajan (1990) has noted. Hashemipour (1989) claims that question words can not be fronted to COMP position at S-structure, and states that: "colloquial Modern Persian does not have syntactic Wh-movement in question formation...". So, according to her judgements, the sentences below (adapted from her) are ungrammatical, while they are not for me and the informants that I have consulted with:

- (6=35b) a. ĩi maryam dad be doxtar-e ?
 what Maryam gave to girl-DF
 "what did Maryam give to the girl?"
- b. be ki maryam ketab-o dad ?
 to who Maryam book-DO gave
 "who did Maryam give the book to?"

The question which arises with Persian is, thus, whether the fronting in these cases involves actual fronting of the Wh-phrases to SPEC of C or whether it involves another process. In other words, is it a syntactic Wh-movement language, or a Wh-in-situ one?

2. Embedded Questions

To examine this issue, let us see how Persian Wh behaves in embedded contexts. Consider the embedded questions as (7) in which verbs like *danestan* "know" and *porsidan* "ask", subcategorize for embedded questions (see Huang, (1982)):

(7)

- a. *man ne-mi-dun-am (ke) ki in goldan-o šekaste ?*
I NEG-CM-know-1S (that) who this vase-DO broken
"I wonder/don't know who has broken this vase."
- b. *amir ne-mi-dun-est (ke) baradar-eš ci xaride bud.*
Amir NEG-CM-know-3S (that) brother-his what bought was
"Amir wondered/didn't know what his brother had bought."
- c. *madar ne-mi-dunest (ke) sanaz ba ki be ketabxune*
mother NEG-CM-knew (that) Sanaz with who to library
rafte bud.
gone was
"the mother wondered/didn't know with whom Sanaz had gone to the library."
- d. *yasaman porsid (ke) dust-eš ki-yo molaqat karde.*
Yasaman asked (that) friend-her who-DO meet done
"Yasaman asked who her friend has met."
- e. *ayda az mo?allem-eš porsid (ke) hamkelasi-ha-s*
Ayda from teacher-her asked (that) classmate-PL-her
koja raft-and.
where went-3PL
"Ayda asked her teacher where her classmates had gone."

(7) exemplifies the status of Persian Wh-questions in embedded clauses. The verbs *danestan* "know" and *porsidan* "ask" are two-place predicates which assign their theta roles to the subject NP, *man*, *Amir*, *madar*, *Yasaman* and *Ayda* respectively, and the internal theta roles are assigned to the interrogative clauses which they govern. From the data above, we can see that the question word remains in the embedded clause in the same position as a non-

question argument occurs. Hence, the Wh-in-situ in Persian is like Mandarin Chinese in that it is allowed in both matrix and embedded clauses. Thus, it is not like French which only allows in-situ in matrix questions (see Rizzi (1989) on this issue). However, as has already been indicated, Persian differs from Chinese in this respect in that it permits the possibility of fronting a Wh besides leaving it in base syntactic position in a main clause. The question to address here is whether Wh-movement holds for embedded questions as it did for main Wh-interrogatives or not. In contrast with Hashemipour (1989) who argues against the possibility of S-structure movement of Wh-elements in Persian embedded clauses, I am going to confirm the existence of such movement in Persian syntax. The behaviour of Wh-words in the contexts below, therefore provides further evidence for the optionality of either moving Wh-phrases to the sentence-initial position or keeping them in-situ:

(8)

- a. amir ne-mi-dunest (ke) ĉi baradar-eš xaride bud.
Amir NEG-CM-knew (that) what brother-his bought was
"Amir wondered/didn't know what his brother had bought."
- b. madar ne-mi-dunest (ke) ba ki sanaz be ketabxune rafte.
mother NEG-CM-knew (that) with who Sanaz to library gone
"the mother wondered/didn't know with whom Sanaz had gone to the library."
- c. yasaman porsid (ke) ki-yo dust-eš molaqat karde.
Yasaman asked (that) who-DO friend-her meet done
"Yasaman asked who her friend has met."
- d. ayda az mo?allem-eš porsid (ke) koja hamkelasi-ha-š
Ayda from teacher-her asked (that) where classmate-PL-her
raft-and.
went-3S
"Ayda asked her teacher where her classmates had gone."

The data in (8) which are perfectly well-formed and normal in colloquial Persian, thus confirms strongly Riemsdijk and Williams's (1986) statement: "in many languages the rule corresponding to Wh-movement is truly optional in the sense that failure to move still results in normal Wh-question interpretation".

We turn now to the question raised above: is Persian an in-situ language like Chinese, or does it also pattern as a Wh-

movement language like English?

In order to examine this matter, we will approach this question via recent work by Cheng (1991).

Cheng's (1991) analysis

Cheng explores a generalization which holds among languages employing in-situ Wh-words. Based on this generalization, she proposes the Clausal Typing Hypothesis, which states that all clauses must be typed at S-structure. In general, her argument rests on this assumption that having in-situ Wh-words excludes the possibility of having syntactic Wh-movement of the Wh-words. On this basis, she provides an interesting account which stems from different principles: the Principle of Economy of Derivation (Chomsky (1989)), the Earliness Principle (Pesetsky (1989)) and The Lateness Principle.

Based on Cheng's argument, syntactic Wh-movement serves to type a sentence as interrogative. Languages which do not have syntactic Wh-movement, have another way to type a clause as interrogative, namely by the use of question particles. Furthermore, assuming the Principle of Economy of derivation, no language uses both ways to type a Wh-question.

An extensive examination of various languages leads her to propose that some in-situ languages have overt yes-no particles and overt Wh-particles. In some languages, such as Mandarin, the yes-no particle is overt, while the overt Wh-particle is used only optionally. She thinks that in the cases when the overt Wh-particle is not used in Mandarin, a non-overt counterpart (\emptyset) is used. Besides, some languages like Iraqi Arabic and Turkish (and also Persian) have only overt yes-no particles. There is never any overt marking in Wh-question in these languages. Concerning Persian, we have already discussed that it possesses a way of forming yes-no questions by means of some morphologically overt particles such as *aya*, *mage* and *hič*. In contrast, Wh-questions do not have any specific Wh-particle. Cheng assumes that in these

languages, although we do not see a Wh-particle in Wh-questions, a non-overt one is present. Hence, the languages which have overt yes-no particles, also have Wh-particles, be it overt or non-overt. So, the use of in-situ Wh-phrases correlates with the particular manner in which yes-no questions are formed. Since, there is no language which uses an overt Wh-particle and which lacks an overt yes-no particle, it appears to be the case that if a language has an overt particle, it must be found in yes-no questions. This suggests that there is an implicational relationship between yes-no particles and Wh-particles: the presence of overt yes-no particles in a given language implies the presence of Wh-particles (overt or non-overt). Regarding this analysis, she gives this general hypothesis:

(9=6) In-situ languages have Wh-particles, languages with Wh-particles are in-situ languages.

It appears that a language with a Wh-particle is a language without syntactic Wh-movement. Thus, the presence of a Wh-particle serves the same purpose as syntactic Wh-movement. The particle *ne*, for instance, in (10) (Cheng's 7) in Mandarin Chinese is in fact present to indicate that the clause type of the sentence is a Wh-question:

(10=7) *hufei mai-le na-yi-ben-shu* (*ne*)
 Hufei buy-ASP which-one-cl-book QWh
 "which book did Hufei buy?"

As the Wh-question clause type of a sentence is identified by these particles, she calls them "typing particles". Following Bach (1970) and Bresnan (1972), Cheng assumes that typing particles are generated in the complementizer position C. So, a Wh-question in languages with in-situ Wh-words, has a typing particle in C to type the sentence as a Wh-question. It is then assumed that languages may have such non-overt particles. This would be the case for Persian. In other words, an overt or covert Wh-particle in C indicates the [+Wh] feature of a clause. But in languages such as English, there is no question morpheme or [+Wh] feature

base-generated in COMP or C. However, after a Wh-word moves to SPEC of CP, C acquires the [+Wh] feature from the Wh-word in its SPEC position through SPEC-head agreement. In other words, movement of a Wh-element into SPEC of C ensures that the C has the [+Wh] feature.

As a result, the following three predictions follow from her discussion:

- a. No language has yes-no particles (and thus Wh-particles), and also syntactic Wh-movement.
- b. No language has the option of using either a Wh-particle or syntactic Wh-movement of Wh-words to type a sentence as a Wh-question.
- c. No language fronts more than one Wh-word for clausal typing.

In general, there are two ways to type a clause as a Wh-question: 1) by a Wh-particle and 2) by Wh-movement.

It should however be noticed that the occurrence of a Wh-word in a clause-initial position is not a good diagnostic of whether or not a language has syntactic Wh-movement, as Wh-words can be in clause-initial as a result of scrambling or some other operation. Cheng's proposal, therefore contradicts the Baker's (1970) analysis of Wh-movement. To determine whether a language has a Wh-movement rule, Baker states the following hypothesis:

- (11) Only languages which position their particles for yes-no questions in clause-initial position permit a movement rule for questioned constituents. (Baker, 1970, p.207)

However, there are languages such as Persian and Hindi (See Mahajan (1990)) which have initial yes-no particles but no syntactic Wh-movement. Thus, it appears that the position of yes-no particles has nothing to do with whether or not a language has syntactic Wh-movement. A language either allows Wh-words to stay in-situ or has syntactic Wh-movement, and no language alternates between the use of in-situ Wh-words and syntactic Wh-movement. Therefore, not all cases of apparent fronting involve movement of the Wh-words to SPEC of CP.

Given this analysis, I would like to follow Cheng in assuming

that Persian like languages such as Egyptian Arabic, Bahasa Indonesia and Palauan which exhibit both movement and in-situ Wh, is basically an in-situ language. Yet the crucial point to consider is that Cheng's proposal with respect to languages like Persian in which the C position is filled by the complementizer, appears to pose some problem. Given the Clausal Typing Hypothesis which requires that every clause be typed at S-structure, the question arises as how the embedded clauses in (12) can be typed as [+Wh] in Persian:

- (12) a. man ne-mi-dun-am ke amir aya umade ya na.
 I NEG-CM-know-1S that Amir INT.M came or not
 "I wonder whether Amir came."
 b. sanaz ne-mi-dun-e ke amir koja raft.
 Sanaz NEG-CM-know-3S that Amir where went
 "Sanaz wonders where Amir went."

The Clausal Typing Hypothesis requires that the embedded CP be typed at S-structure. Thus, in (12), the embedded C has to bear the [+Wh] feature. The only way for the embedded C to acquire the [+Wh] feature is through the existence of an overt yes-no particle in (12a) and a non-overt Wh-particle in (12b), assumed to be generated in C. Given the fact that this position is occupied by the complementizer ke, the [+Wh] requirement of the embedded clause cannot be satisfied, and the Clausal Typing Hypothesis remains an unresolved problem in Persian embedded clauses, unless the particle is in a position such as the MP discussed in the preceding chapter.

It was proposed above that Persian and languages such as Egyptian Arabic are alike in that both are basically in-situ languages. Nonetheless, Persian differs from those languages in some respects. Thus, the same analyses, which satisfactorily explain the properties of question-formation in the relevant languages, do not hold for Persian and are unable to account for the Wh-question characteristics of this language.

With respect to Egyptian Arabic and Bahasa Indonesia, Cheng suggests that the apparent fronting of Wh-words is an instance of a cleft sentence in the case of arguments and topicalization in the

case of adjuncts, and is not the same as the fronting of Wh-words in English which involve movement of the Wh-words to SPEC of C. In Persian, wh-questions do not resemble relative clauses and cleft sentences, unlike what is the case in Egyptian Arabic, where the fronting of Wh-arguments has striking similarities with what is found in relative clauses as well as in clefts. Considering the characteristics of Wh-question themselves, it appears that Persian, unlike Egyptian Arabic, does not show any difference in movement of Wh-arguments and Wh-adjuncts to the clause-initial position. Wh-phrases seem to act in a way similar to that of Egyptian Arabic's adjuncts and resemble topicalization of an NP. In Egyptian Arabic, the complementizer "illi" shows up in relatives and clefts as well as in Wh-questions.¹ The fronted Wh-argument, as Cheng (1991, p.42) states, is originally base-generated as a subject of a cleft sentence, and no movement of the Wh-argument is involved, as shown in (13a-c). Examples from Egyptian Arabic are from Wahba (1984):

- (13) Relative clause and clefts
- a. il-ragil illi mona shaafit-uh.
 the-man that Mona saw-him
 "the man that Mona saw."
- b. (dah) muhamad illi gih.
 (this) Mohammed that came
 "it is Mohammed that came."
- Wh-question
- c. miin illi mona darabit-uh?
 who that Mona hit-him
 "who did Mona hit?"

The fronting of Wh-adjuncts, however, is not like relativization. Instead, it looks like topicalization in that none of them contains "illi":

- (14) Wh-question
- a. ma'a miin mona raabit il-qahirah?
 with whom Mona went to Cairo
 "with whom did Mona go to Cairo?"

Topicalization

- b. fi-l-shari'dah mona kaanit bitdawwar 'ala sha??ah.
 on-the-streetDEM Mona was looking for apartment
 "on the street, Mona was looking for an apartment."

In Persian fronting of Wh-phrases can not be argued to be parallel to clefts and relativization. Furthermore, arguments and adjuncts do not differ with each other with regard to fronting. The complementizer *ke* is missing in both Wh-fronting and topicalization, while readily available in relative clauses and clefts, as exemplified by the following examples:

(15) relative clause

- a. ketab-i ke amir be dust-eš dad.
 book-the that Amir to friend-his gave
 "the book that Amir gave to his friend."

Cleft

- b. in ketab-i-ye ke amir be dust-eš dad.
 this book-the-EZ that Amir to friend-his gave
 "this is the book that Amir gave to his friend."

Topicalization

- c. ketab-i-ro, amir be dust-eš dad.
 book-the-DO, Amir to friend-her gave
 "a book, Amir gave to his friend."

- d. ba madar-eš ayda be iran raft.
 with mother-her Ayda to iran went
 "with her mother. Ayda went to Iran."

Wh-question

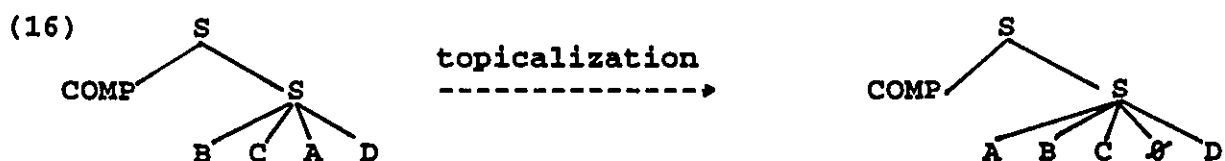
- e. ci-ro amir be dust-eš dad ?
 what-DO Amir to friend-his gave
 "what did Amir give to his friend?"
- f. ba ki, ayda be iran raft ?
 with who, Ayda to Iran went
 "with whom, Ayda went to Iran?"

Given the fact that Wh-questions pattern with topicalization, as opposed to relativization and clefts, it is normal to investigate the extent of the parallelism, and see whether the movement of the wh-phrase in question is not simply an instance of topicalization, given that movement of the Wh-phrase is not obligatory. If this were the case, movement of the Wh-phrase

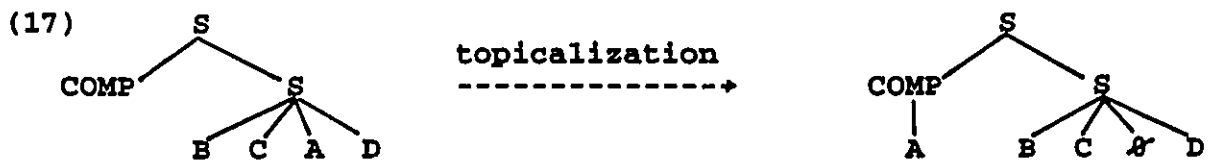
should not be interpreted as an indication of the existence of optional Wh-movement in the language as distinct from topicalization. The evidence which supports this approach is that Wh-elements, like any other syntactic phrases, can appear everywhere in the clause freely, as discussed earlier. However, following Chomsky (1977), I will still hold to the view that topicalization is an operation which is a type of Wh-movement operation. This assumption will not be inconsistent with Cheng's hypothesis either.

SURFACE POSITION OF WH-TOPIC

There are at least two possible accounts of the surface position of a topicalized element in the literature. In the first analysis, topicalization is treated as an operation of XP adjunction to IP. This is illustrated in (16), taken from Lowenstamm (1977):



The second analysis, on the other hand, depends on the assumption that topicalization is an instance of syntactic movement to SPEC CP. The derivation of this account is diagrammed in the figure that follows:



Though the second account has also been argued to be a convincing alternative (see Lowenstamm (1977), Authier (1992) for analyses along these lines), with respect to Persian Wh-topic, I will adduce some evidence to propose that the first account is to be favored.

The evidence for the proposal that the fronted Wh-phrase in Persian results directly from adjunction to IP, involves those structures in which the Wh-phrases are preceded freely by the complementizer *ke* in embedded clauses, as shown in examples (8). Consider (8a,b) below, repeated here for ease of reference:

(18=8)

- a. *amir ne-mi-dunest ke ĉi baradar-eš xaride bud.*
 Amir NEG-CM-knew that what brother-his bought was
 "Amir wondered/didn't know what his brother had bought."
- b. *madar ne-mi-dunest ke ba ki sanaz be ketabxune rafte.*
 mother NEG-CM-knew that with who Sanaz to library went
 "the mother wondered/didn't know with whom Sanaz had gone to the library."

Since the Wh-elements *ĉi* and *ba ki* follow the complementizer *ke* and precede the subjects *baradareš* and *Sanaz* in the above examples, they obviously are not in SPEC CP. This fact leads us to propose that Wh-phrases in Persian are adjoined to IP.³

CHARACTERIZING WH-MOVEMENT

Now, let us review the relevant characteristics of Wh-movement in English according to Chomsky (1977).

Chomsky proposes a cluster of properties which can be taken as a diagnostic for the rule of Wh-movement. In later versions of the theory, this rule is seen as an instance of the more general rule "move α " (cf. Chomsky 1982, p.33). These general characteristics are:

(19=49)

- "a.it leaves a gap.
 b.Where there is a bridge, there is an apparent violation of subjacency, propositional-island condition (PIC), and specified subject condition (SSC).
 c.it observes complex noun phrase constraint (CNPC).
 d.it observes Wh-island constraint."

The set of above properties, as Chomsky demonstrates, "follow[s], on the theory outlined, from the assumption that Wh-movement moves a phrase (implying (a)), observes SSC, PIC, and subjacency (implying (c) and (d)), and is permitted from COMP-to-COMP under "bridge" conditions (implying (b)). These properties

are taken to constitute an indication that the rule of Wh-movement is involved when observed in construction in which there are no overt signs of Wh-movement such as Wh-pronouns. On the basis of this diagnosis, constructions such as the comparative construction, the cleft construction, etc. are reanalyzed as involving Wh-movement.

Topicalization, yields this configuration too. Chomsky proposes a structure like (20), generated by the phrase structure rules $S'' \rightarrow \text{Top } S'$, and $S' \rightarrow \text{COMP } S$:

(20) [_{S''} Top [_{S'} COMP [_{S...}]_S]_{S'}]_{S''}

"Move Wh" in such a structure must take place. Furthermore, the rule deleting Wh from COMP can apply to delete the Wh-word in PF. The topicalization construction, thus, can be said to arise from the movement of an NP (or PP) to COMP. Since the movement can be iterated from COMP to COMP, and since each movement is subject to the subadjacency condition, the construction manifests the Wh-diagnostics, exactly like a construction with an overt Wh-word. Under such an analysis, therefore, topicalization is expected to involve the same properties found in relatives and in direct and indirect questions.

So, as illustrated in the examples below, it follows that topicalization should be accounted for by the general rule of Wh-movement, and has the properties of (19=49):

- (21=63) a. this book, I really like.
 b. this book, I asked Bill to get his students to read.
 c.*this book, I accept the argument that John should read.
 d.*this book, I wonder who read.

The impossibility of the sentences such as (21c,d) and also (22) below can, thus, be explained by essentially the same lines of argument that accounts for the Wh-island constraint:

- (22=75) a.*to whom the books did John give away (to whom did the books John give away).
 b.*whom the books did John give away to.
 c.*the boy to whom the books John gave away.
 d.*the boy whom the books John gave away to.

These data indicate that topicalization is impossible within relatives or questions in English. But as the following example shows, it is possible out of that-clauses:

(23=77) I believe that the books, John gave away to some friends.

Based on Chomsky's assumption, (22c-d) for example, would have the following underlying structure:

(24) the boy [_S, COMP [_S, [Top the books] [_S, COMP John gave away which to whom]

Wh-movement can apply to "which" in the embedded sentence, placing it in the internal COMP position, and leaving a trace. However, movement of "whom" to the higher COMP is blocked, since it has to cross the internal COMP which is already occupied by "which" under the Wh-movement analysis of topicalization.

Turning to Persian, we will see that there seems to be evidence for Wh-movement also, although the evidence is not as straightforward as in English. While the examples (21c-d) and (22a-d) are ruled out in English, many of them are well-formed in Persian, which means that topicalization can escape many of the island conditions:

(25) a. in ketab-o, (man) vaqeʔan dust dara-am.
this book-DO, I really like have-1S
"this book, I really like."

b. in ketab-o, (man) az amir xast-am ke danešju-ha-š
this book-DO, I from Amir wanted-1S that student-PL-his
bexun-nand.
read-3PL
"this book, I asked Amir to get his students read."

c. *in ketab-o, (man) in matlab-o qabul dar-am ke
this book-DO, I this argument-DO accept have-1S that
sanaz bayad bexun-e.
Sanaz should read-3S
"this book, I accept the argument that Sanaz should read."

d. in ketab-o, (man) ne-mi-dun-am ki xund-e.
this book-DO I NEG-CM-know-1S who read-3S
"this book, I wonder who read."

e. man fekr mi-kon-am ke ketab-o amir be čand-te
 I thought CM-do-1S that book-Do Amir to some
 az dust-a-š dade.
 from friend-PL-his gave

"I believe that the books, Amir gave away to some friends."

While extraction out of some complex NPs, specifically noun complements is excluded, as shown by the ungrammaticality of (25c), the grammaticality of (25d) shows that Persian is not subject to the Wh-island condition. The unacceptability of sentences such as (25c), thus, requires an explanation for the asymmetry between the Wh-island constraint and the complex NP constraint in Persian. Meanwhile, the contrast between (25c) and (26), which is semantically close to it, ensures that the account for (25c) should be structural:

(26) in ketab-o, man qabul dar-am ke sanaz bayad bexun-e.
 this book-DO, I accept have-1S that Sanaz should read-3S
 "this book, I accept that Sanaz should read."

This turns out to be somewhat reminiscent of the situation described by Rizzi (1982) for Italian. Based on this similarity, one might assume that in Persian just as for Italian, in the terms of the theory discussed at the time by Rizzi, NP and CP are bounding nodes. However, we will see in the following sections that Persian extraction patterns actually are more complicated than they look.

Before going into a more detailed examination of the structures above, I am going to proceed considering the four diagnostic properties of Wh-movement, stated in (19), in relation to Persian Wh-topicalized configurations. In so doing, we will discover to what extent Chomsky's syntactic characterization of Wh-movement is applicable to the Persian topicalization facts.

A. THE GAP FEATURE

Syntactic Wh-movement holds between a preposed constituent and a gap. Consider, for instance, the English example (27) in which this dependency is indicated by coindexing the preposed constituent with the empty position:

(27) who₁ do you think that she saw --₁ ?

Topicalization in Persian also conforms to this characterization, and leaves a gap, as shown by the examples such as (28a-b) and also (28c-d) in which the presence of the resumptive pronoun or a postposed clitic pronoun leads to ungrammaticality:

- (28) a. č_i₁, mi-g-e (ke) --₁ dust dar-e ?
 what, CM-say-3S (that) like has-3S
- b. ketab₁, mi-g-e (ke) --₁ dust dar-e.
 book, CM-say-3S (that) like has-3S
 "book, he/she says that he/she likes"
- c. *ketab, mi-g-e (ke) un-o dust dar-e.
- d. *ketab, mi-g-e (ke) dust-eš dar-e.

Topicalization, thus contrasts with left-dislocated structures where an initial NP is related to an anaphoric pronoun or NP within the clause:

- (29) ketab-o, mi-g-e (ke) xarid-i-š.
 book-DO, CM-say-3S (that) bought-2S-it
 "the book, he/she says that you bought it."

B. BRIDGES

In the presence of a bridge verb, the topicalized Wh-phrase may be moved out of tensed clauses, across specified subjects and through several successive cycles which results in apparent violating Propositional Island Constraint (PIC), Specified Subject Condition and Subjacency respectively, as illustrated in (30):

- (30) č_i madar mi-g-e ke sanaz fekr mi-kon-e ke be
 what mother CM-say-3S that Sanaz thought CM-do-3S that to
 nazar mi-res-e to dašte baš-i.
 look CM-reach-3S you had be-2S
 "what does mother say that Sanaz thinks that (it) seems that
 you have?"

Thus far we have seen that the two properties of the Wh-movement rule are respected in Persian. Now, let's turn to the two other diagnostic features which require a brief review of the

notion of Subjacency.

C. Subjacency Condition

In this part, I will sketch an analysis of Subjacency, and present Persian data in order to address the issue of what type of restrictions this phenomenon is subject to.

Within generative grammar various attempts have been made to delineate the restrictions on extractions, and a major thrust of the syntactic research over the past decade has been to formulate abstract principles from which island constraints could be deduced. Ross (1967) proposed a number of structural configurations that are immune to extractions. Relative clauses, embedded Wh-questions, sentential subjects are among those structures that do not permit extraction unless a pronoun occurs in the position of the extraction site. Ross's major contribution was in bringing about the realization that a general explanatory system of conditions on transformations could be formulated. Although the particular formulations of the constraints have changed since then, the predictions have remained more or less the same, and has led much subsequent work to concentrate on generalizing and unifying these constraints, resulting in the advancement of several hypotheses about linguistic universals. One of the most significant among these attempts is the principle of Subjacency proposed in Chomsky (1973).

The Subjacency Condition is a powerful principle by which a wide variety of structures that were previously analyzed in terms of a disparate set of conditions such as the Complex Noun Phrase Constraint and the Sentential Subject Constraint are unified. Chomsky, therefore, has tried to replace all of Ross's constraints with a set of conditions on the application of syntactic rules. One crucial fact that has come to light is that this principle can be extended to seemingly conflicting and different material in other languages which involve adding to, or varying the set of bounding nodes for Subjacency. So, it has been shown (cf. Chomsky

(1973, 1977, 1981, among others) that a number of conditions on movement rules are predicted by Bounding Theory (The Subjacency Principle).

This principle constrains the application of transformations to the domain of one cyclic node or to the domain of two adjacent cyclic nodes. So, Subjacency states that no rule can involve X and Y in the structure:

(31) ...X...[α ...[β ...Y... (or: ...Y...] β ...] α ...X...)

Where α and β are bounding nodes

By appealing to Subjacency, the violation of the Complex NP Constraint, for instance, is blocked in the following manner, represented in (32):

(32)
* $[_s \text{ who}_1 \text{ did you believe } [_{NP} \text{ the claim } [_s \text{ that } [_s \text{ John invite --}_1]]]]$

The sentence is ungrammatical, since the movement of the constituent "who" to the front would cross two cyclic nodes, NP and S. However, according to this interpretation of the Subjacency Principle, the grammatical sentences like (33) would be ruled out.
(33)

$[_{s_1} \text{ which prize}_1 \text{ did you say } [_{s_2} \text{ that Mary didn't think } [_{s_3} \text{ that John had won --}_1]]]$

To account for such structures, it has been proposed by Chomsky that the Wh-movement rule applies in a successive cyclic way. The Wh-element, on each cycle, moves into the COMP position which thus becomes an "escape position". The successive cyclic COMP-to-COMP movement of (33) can therefore be illustrated in (34):

(34)
 $[_{s_1} \text{ COMP you said } [_{s_2} \text{ COMP Mary didn't think } [_{s_3} \text{ COMP John had won which prize }]]]$

Bounding nodes were at first taken to be S and NP. However, after a distinction between S and S' was made by Bresnan, there was some uncertainty on determining S or S' as the bounding node. Chomsky (1977) discusses both possibilities and settles for assuming S as the bounding node. As a result, island constraints are not as uniform across languages, as they first were thought to be. This

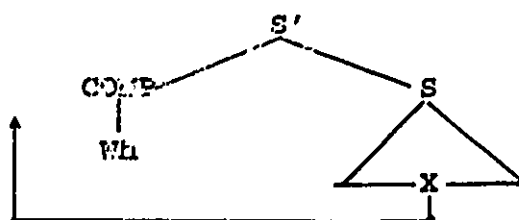
fact has led some linguists, notably Chomsky, to the conclusion that island constraints must be formulated as very general schemata which leave unspecified certain parameters, the actual value of which may vary from language to language.

In the subsequent section, I will deal with the extraction possibilities in Persian, and show that island constraints in this language are far less restricted than in languages such as English. I will introduce evidence from the grammar of Persian that supports Rizzi's (1982) approach to Wh-island violation, and argue that there are reason to believe that Rizzi's analysis of Subjacency is quite plausible. However, I will also explore the dissimilarities found in Persian and Italian as well as the characteristics which are common to both Scandinavian languages and Persian with respect to extraction patterns. In general, using Subjacency Condition as a diagnostic, I will examine whether Wh-topicalization formation in Persian is indeed a result of movement or not, by looking at extractions out of embedded questions in (1) and at extractions out of complex noun phrases in (2) below.

1. Extractions out of embedded questions

According to Ross (1967), a clause introduced by a Wh-pronoun is an island. So, in the following abstract structure, an element X can not be extracted out of (35), because: a) The Tensed S Condition (TSC) and Specified Subject Condition (SSC) will be violated. b) Given the assumption that COMP can not be filled more than once, the movement of X to the COMP position will be blocked, since this position is already filled:

(35)



Although, most of the discussion on constraints on movement patterns so far has been based on facts from English or similar

languages, i.e. in languages where extractions are limited to one per sentence, the more recent approaches have shown that the constraint according to which embedded constituent questions and yes/no questions count as Wh-islands, may not be regarded as a universal explanation any more. It is now a well-known fact that in several languages, violation of the Wh-island constraint is freely allowed. This is exactly what we find in Persian. That is, in Persian, as in Italian, it is quite possible to Wh-move from within indirect questions. This contrasts with the situation in some dialects of English (Ross (1967)) in which Wh-movement is assumed to be subject to the wh-island constraint. Consider now the examples in (36) which show that it is possible to relativize into an indirect question:

(36)

- a. tablo-i₁ ke ne-mi-dunest-i ki ---₁ xaride, axiran
 tableau-the that NEG-CM-knew-2S who bought recently
 be sergat rafte.
 to theft gone
 "the tableau that you didn't know who bought, has been
 recently stolen."
- b. hediye-i₁ ke ne-mi-dunest aya sanaz --₁ dust dar-e,
 gift-the that Neg-CM-knew whether Sanaz like hve-3S
 keyli jaleb bud.
 very interesting was
 "the gift that she/he didn't know whether Sanaz likes, was
 very interesting."
- c. dust-et₁ ke be-heš ne-mi-dun-am či --₁ gofte bud-i
 friend-your that to-him NEG-CM-know-1S what told was-2S
 saxt negaran bud.
 very worried was
 "your friend to whom I wonder what you had told, was very
 worried."
- d. diruz film-i-ro₁ did-am ke yad-am rafte ki --₁
 yesterday movie-a-DO saw-1S that memory-1S gone who
 kargardani karde.
 direction did
 "yesterday, I saw a movie that I have forgotten who
 directed."

Furthermore, unlike what is the case in Italian, with interrogative pronouns extraction out of a Wh-question also gives acceptable results in Persian. Extraction from indirect questions with

interrogative pronouns leads to mixed results in Italian (cf. Rizzi, 1982, p.51).³ Persian like Spanish (see Torrego (1984)) exhibits no such asymmetry between interrogative and relative pronouns in Wh-movement from indirect questions:

(37)

- a. *či-yo/ro₁ ne-mi-dunest-i ki --₁ qarz gerefte ?*
 what-DO NEG-CM-knew-2S who loan taken
 "what didn't you know who has borrowed?"
- b. *če kasi-ro₁ amir ne-mi-dun-e ki --₁ davat kard ?*
 what person-DO Amir NEG-CM-know-3S who invitation did
 "which person doesn't Amir know who invited?"
- c. *ba če kasi₁ fekr mi-kon-i ki --₁ be ketabxune rafte ?*
 with who thought CM-do-2S who to library gone
 "with whom do you think who has gone to the library?"
- d. *kodum film-o₁ yad-eš rafte ki--₁ kargardani karde ?*
 which movie-DO memory-3S gone who direction did
 "which movie has he forgotten who directed?"

As can be seen from the examples above, there is no restriction against applying the same process more than once in Persian. In view of this fact, both relativization and question formation out of an embedded question are acceptable. Notice that in the process of relative clause and question formation as represented in (36) and (37), a preposed constituent is related to a structural gap. Although one might occasionally hear a resumptive pronoun instead of a gap in sentences with only two levels of embedding, as in (36), (37), the general consensus among speakers of Persian is that a gap is much more preferable. This is shown by the marginal acceptability of the following typical examples:

(38)

- a. *?tablo-i ke ne-mi-dunest-i ki un-o xaride, axiran*
 tableau-the that NEG-CM-knew-2S who it-DO bought recently
 be serqat rafte.
 to theft gone
 "the
- b. *?či-yo ne-mi-dunest-i ki un-o qarz gerefte ?*
 what-DO NEG-CM-knew-2S who it-DO loan taken
- c. *?kodum film-o yad-es rafte ki un-o kargardani karde ?*
 which movie-DO memory-3S gone who it-DO direction did

The comparison of the data in (36) and (37) on one hand, and (38) on the other, thus, shows that gap and resumptive pronouns are not actually in free alternation in Persian. It seems to me that the overriding factor determining the occurrence of resumptive pronouns is the length and complexity of the intervening lexical material, in other words distance. This is illustrated in (39) where the extraction site is embedded more than two sentences down and a resumptive pronoun is inserted:

(39)

a.diruz film-i-ro₁ did-am [S' ke ne-mi-dun-am [S' aya
 yesterday movie-a-DO saw-1S that NEG-CM-know-1S whether
 to yad-et mi-yad [S'ki--₁ un-o₁ kargardani karde ya na]]
 you memory-2S CM-come who it-DO direction done or not
 "yesterday I saw a movie that I wonder if you remember who
 that directed."

This indicates that resumptive pronouns are accepted in islands when the distance between the filler and the pronoun is sufficiently large. Though still some speakers, even at this level of embedding may find the pronoun a bit redundant, they also believe that using a pronoun in constructions like (39) sounds much more natural than the ones in (38). However, we will see later on that the ungrammatical sentences involving island violations will not be improved by the insertion of the resumptive pronouns in Persian, and argue that the generalization that resumptive pronouns and gaps are in complementary distribution does not hold for Persian. Hence, it is complexity per se that matters, and not the complexity arising from violations of syntactic islands, whether the locus of extraction is inside an island or not. The evidence for this claim is provided by sentences such as (40) below in which the extraction site is embedded under a series of bridge verbs which do not create islands. The length and amount of embedding requires a resumptive pronoun to appear optionally:

(40)

so?al-at-e emtehani-ro₁ mo?allem goft [_S.ke keyli
 question-PL-EZ exam-DO teacher said that very
 qeyr-e ?adi-ye [_S'ke ma?lum šode [_S.ke faqat čand ta az
 unusual-is that clear become that only few number of
 danešamuz-a-yi ke emtehan mi-dad-and tunest-and --₁ / un-a-ro₁
 student-PL-the that exam CM-gave-3PL could-3PL it-PL-DO
 javab beđ-and.

answer give-3PL

"the exam questions, the teacher said that it was very unusual that it has appeared that only a few of the students who gave the exam could answer."

The examples discussed so far crucially involve movement of more than one constituent on the same cycle, and show that Persian allows for a longer distance between the extracted constituent and the gap related to it than English does. Thus, Subjacency, as given in (31), appears to be inadequate to cover facts about extraction patterns in Persian. Assuming for the typical example (37b) the structure given in (41) (with irrelevant movements and traces omitted for simplicity):

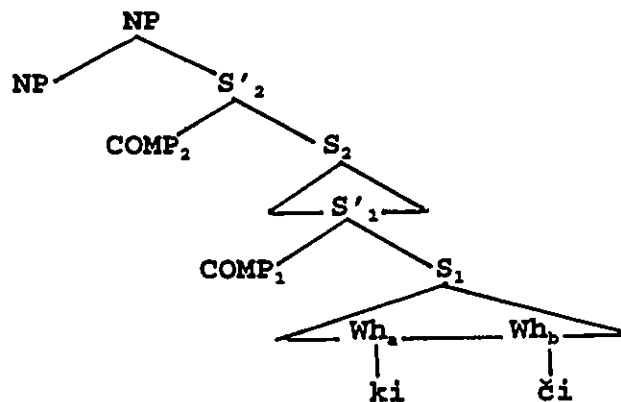
(41)

a. [_{CP} [če kasi-ro]_{IP} Amir ne-mi-dune-e [_{CP} [ki [_{IP} --₁ davat kard]]]]]]

It is clear that, if the bounding nodes in Persian were NP and S(=IP), namely exactly those that Chomsky (1973) suggests for English, then the example, in (36) and (37) should violate Subjacency: the movement crosses two bounding nodes, i.e. the two IPs. Thus, the grammaticality of sentences in (36) and (37) may mean either that Wh-movement in Persian is simply not subject to Subjacency, or that Subjacency is operating but some other factor makes it inapplicable. Rizzi (1982) argues that the latter must be the case in Italian. To maintain the universality of Subjacency in the face of such violations of the Wh-island, Rizzi puts forward a parametrization of Subjacency: languages such as Italian (cf. also Sportiche (1981) for French), unlike English, have S' not S as bounding node, thus allowing extraction from what in English would constitute a Wh-island. If I extend to Persian the type of analysis proposed for Italian, the derivation of sentences (36) and

(37) is possible only at the cost of violating some conditions on rules, or changing the syntax of the complementizer. Given the plausible base structure shown in (42):

(42)



sentence (37a), for instance, could be derived according to one of the following three solutions:

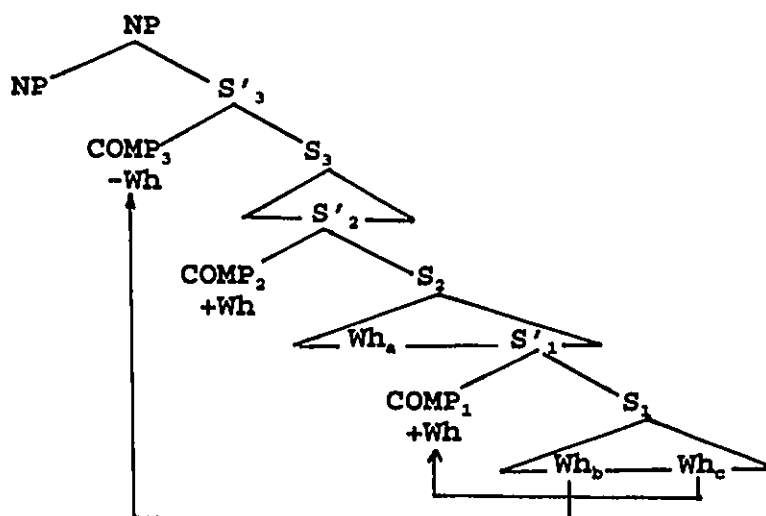
(43)

- a. Strict Cyclicity Condition is violated.
- b. A COMP may be "doubly filled" at the end of a cycle as long as it is not doubly filled at S-structure.
- c. Not S, but S' is a bounding node for the Subjacency Condition in Persian.

Rizzi rejects a solution that involves violating the Strict Cycle Condition. According to this solution, Wh_b first moves to COMP₁, then vacates it to move to COMP₂. Wh_a then moves into COMP₁, violating the Strict Cycle Condition. Since eliminating the Strict Cycle Condition would in effect destroy the foundation upon which the Subjacency Condition is built, this solution is not a viable one. According to (43b), both Wh_a and Wh_b may move into COMP₁ at the first cycle. At the second cycle, Wh_b moves on to COMP₂ without violating Subjacency. On the other hand, based on (43c) solution, Wh_a is moved into COMP₁ at the first cycle, while Wh_b is moved directly into COMP₂ from its base position at the second cycle without violating Subjacency, because only one S'-boundary is crossed. Hypothesis (43c) makes it possible for the pronoun to skip one COMP containing a Wh-phrase, but never more than one, since then two S'-nodes would be crossed. By contrast, hypothesis

(43b) does not discriminate in such cases, because the COMPs can be doubly filled on every cycle. In order to investigate which of these two hypotheses hold for Persian, let us, following Rizzi, consider the D-structures of the following type which provide the crucial test:

(44)



Under solution (43b), both Wh_b and Wh_c are extractable, so long as in S-structure they occupy different COMPs. On the contrary, solution (43c) blocks the derivation of any grammatical surface structure from (44), since two boundaries intervene between the source position and the target position. Sentences of this type are systematically unacceptable in Italian, whereas they turn out to be fully grammatical in Persian. This shows that the prediction of hypothesis (43c) is correct, and therefore preferable over the hypothesis (43b) in Italian, but not in Persian. Both the Italian example and the data from Persian are presented below for ease of comparison:

(45)

Italian

*questo incarico, che non so proprio chi possa avere indovinato a chi affidero, mi sta creando un sacco di grattacapi.

"this task, that I really don't know who might have guessed to whom I will entrust, is getting me into trouble."

(46)

Persian

a. in ketab, ke ne-mi-dun-am čand nafar momken-e xabar
 this book, that NEG-CM-know-1S how many possible-is news
 dašte baš-and ke dar bare-ye či nevešte šode, bara-m
 had be-3PL that about-EZ what written become for-me
 xeyli jaleb-e.
 very interesting-is.

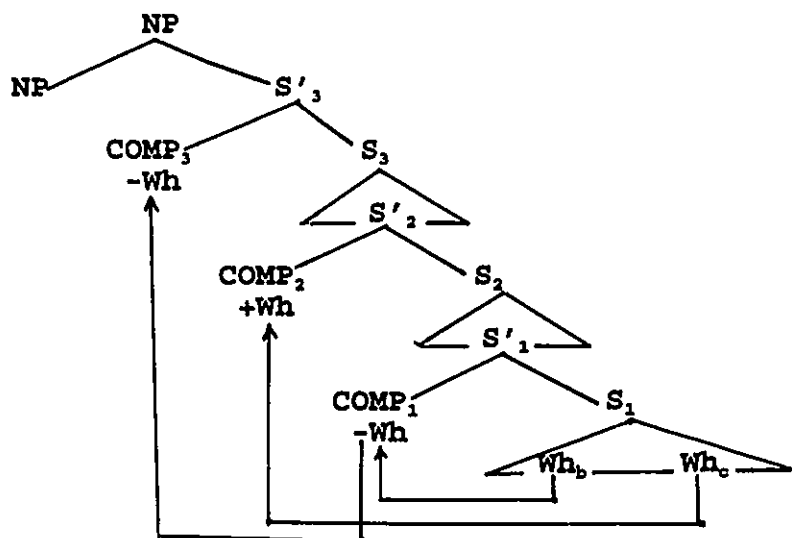
"this book which I wonder how many people might know about what
 has been written, is very interesting for me."

b. sanaz, ke vaqe'an ne-mi-dun-am ki xabar dar-e koja ba
 Sanaz that really NEG-CM-know-1S who news have-3S where with
 maman-eš rafte, diruz xeyli xošhal bud.
 mother-her gone yesterday very happy was
 "Sanaz, that I really wonder who knows where has gone with her
 mother, was very happy yesterday."

The above sentences have exactly the structure (44) as their D-structure. The fact that (46) is grammatical means that the solution (43c) should be rejected. If S' is assumed to be a bounding node for Subjacency in Persian, then (46) should be ruled out. As shown, this claim, however, is not consistent with the facts in this language.

There is still a further type of structure which confirms strongly our analysis, and can be roughly represented as follows:

(47)



In (47), S'_3 is a relative clause, S'_2 and S'_1 , are interrogative and declarative clauses respectively. Again unlike other solutions, solution (43c) predicts that it is not possible to

derive a well-formed surface structure in Italian. The relative pronoun moves into COMP₁ at the first cycle, and interrogative pronoun moves into COMP₂ at the second cycle, and at the third cycle the relative pronoun moves from COMP₁ to COMP₂, hence violating Subjacency. The validity of this prediction is illustrated by the typical unacceptable Italian example in (48) which as before contrasts with the grammatical sentences in (49), indicating the inadequacy of solution (43c) for Persian:

(48)

Italian

*il mio primo libro, che credo che tu sappia a chi ho dedicato, mi e sempre stato molto caro.

"my first book, which I believe that you know to whom I dedicated, has always been very dear to me."

(49)

Persian

a. avvalin ketab-eš ke fekr mi-kon-am (ke) mi-đun-i
 first book-his/her that thought CM-do-1S (that) CM-know-2S
 be ki taqdim karde, hamiše bara-š xeyli ?aziz bud.
 to who dedication done always for-him/her very dear was
 "his/her first book, which I think (that) you know to whom
 he/she has dedicated, has always been very dear to him/her."

b. sanaz, ke be-heš mi-đun-am či-yo/ro fekr mi-kon-i (ke)
 Sanaz, that to-her CM-know-1S what-DO thought CM-do-2S (that)
 mi-xa-m bed-am, baradar-zade-ye man-e.
 CM-want-1S give-1S, brother-born-EZ I-is
 "Sanaz, to whom I know what you believe that I want to give, is
 my niece."

All the examples considered so far, therefore illustrate that Wh-island violations in both Persian and Italian, at first glance, appear to basically involve similar characteristics. However, the more complicated constructions as in (45-46) and (48-49) provide us with the evidence that these languages indeed diverge sharply. So, while Subjacency Condition may explain satisfactorily the extraction phenomenon in languages such as Italian, it is unable by itself to cover facts in Persian. This is similar to what is found in some Nordic languages such as Swedish (see, Engdahl (1982, 1986)) and Danish (Erteschik-Shir (1982)).

The only solution which turns out to exist for the above Persian sentences, would thus be to relax the prohibition against

doubly filled COMPs, thereby letting two Wh-phrases move into COMP on the same cycle.⁴ However, I assume with Engdahl (1982, 1986) that this could be an interesting solution, provided that there are reasons to believe that there is a fixed number of escape positions in COMP for a given language. An obvious question at this point is hence whether it is sufficient to provide for two gaps in a sentence, or if any number of extractions could be possible. Let us try a sentence with three gaps that would normally require three escape positions in COMP. Such a sentence naturally gets quite complex, since we need at least three levels of embedding to construct it. We will start off with (50), a double extraction involving a relativization out of an indirect question:

(50)

tu in kelas danešamuz-i-ye ke ba-haš, ne-mi-dun-am
 in this class student-a-is that with-him/her NEG-CM-know
 kodum mo?allem₁ --₁ mi-xa-d --₁ dar mored-e masa?el-e
 which teacher CM-want-3S about-EZ issues-EZ
 ejtema?i sohbət kon-e.
 social talk do-3S

"in this class, there is a student to whom I wonder which teacher wants to talk about social issues."

In (51), we topicalize dar morede masa?ele ejtema?i "about social issues" out of the embedded question, which creates three gaps inside this clause:

(51)

dar mored-e masa?el-e ejtema?i_x tu in kelas danešamuz-i-ye
 about-EZ issues-EZ social in this class student-a-is
 ke ba-haš, ne-mi-dun-am kodum mo?allem₁ --₁ mi-xa-d
 that with-him-her NEG-CM-know-1S which teacher CM-want-3S
 --₁ --_x sohbət kon-e.
 talk do-3S

"about social issues in this class there is a student to whom I wonder which teacher wants to talk."

Although a sentence such as (51) is a very complex sentence, it is not in fact structurally wrong in Persian. Nonetheless in an ordinary conversation, it would sound highly odd, mainly because it requires a very special context to be motivated. But, if the essential factor, namely the required context is available, then the occurrence of a sentence as in (51) will be plausible. However, to quote Engdahl (1982) "to let the grammaticality of a

sentence depends on whether it is easy or hard to think of a context that makes the sentence plausible would be a serious methodological mistake". So, it seems to be exceedingly hard to establish a clear cut-off point in the grammar between grammatical and ungrammatical multi-extraction sentences.

Given the fact that with respect to extraction patterns, reapplying the same process is not restricted in Persian, we might expect to have constructions of the type where triple extractions is possible. This prediction holds for Persian as well as Swedish. Consider the structure in (52) which represents a question out of question, creating a two-gap structure:

(52)
 ba kodum danešamuz, qarar bud (ke) kodum mo?allem,
 with which student arrangement was (that) which teacher
 --₁ --₂ dar mored-e ketab-ha-ye jaleb sohbat kon-e ?
 about book-PL-EZ interesting talk do-3s
 "with which student it was supposed to which teacher talks about interesting books?"

By topicalizing dar morede ketabhaye jaleb "about interesting books", we will have the following triple extraction structure:

(53)
 dar mored-e če ketab-i_k bud ke ne-mi-dunest-i ba kodum
 about-EZ which book-the was that NEG-CM-knew-2S with which
 danešamuz, qarar bude (ke) kodum mo?allem,₁ --₁ --₂ --_k
 student arrangement was (that) which teacher
 sohbat kon-e?
 talk do-3S
 "about which book was it that you didn't know to which student it was supposed to, which teacher talks?"

(53), similar to the case in (51), is extremely hard to process, since it violates conversational norms on how much of the information need be directly expressed in the question. However, under some specific context, an anxious person who is trying to make himself perfectly clear might phrase the question in this way. Perhaps as Engdahl proposes accurately, the overall impression of complexity arises from the fact that three fillers must be kept available simultaneously when the parser reaches the most embedded clause, which obviously places quite a strain on the processor. This analysis accounts for the high markedness of constructions

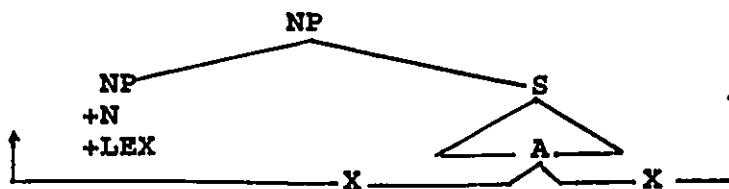
such as (51) and (53) without putting any limit on the number of extractions in the grammar. Therefore, I also assume that it appears to be in principle misguided to set a numeric limit n , on the number of permissible extractions, such that sentences with n extractions would be grammatical, while those with $n+1$ extractions would be ruled out. Recall that although, sentences like (51) and (53) are not normally produced in speech, each part of them are formed according to the rules of the grammar. Hence, it seems to be unreasonable to exclude the possibility of generating them by a restriction in the grammar. This contrasts with languages such as English in which the constraint against multi-gap sentences appear to be a grammatical constraint.

In sum, this argument has shown that the case for multiple extractions in Persian is quite parallel to that of Swedish. That is extraction patterns in none of these languages are subject to the structural restrictions alone. Consequently, following the suggestion that Engdahl (1982) has made for Swedish, I take the view that the factors that influence acceptability of extraction out of embedded questions in Persian, have to do primarily with sentence processing, discourse requirements and with the availability of a pragmatically plausible context. Clearly, these factors have little to do with the formal aspect of the grammar.

2. Extractions out of Complex Noun Phrases

We have seen ample evidence that the Wh-island constraint is not applicable in Persian. We will now consider the status of the Complex NP Constraint (CNPC). Ross's structural definition of the CNPC, as presented in (54), subsumes two cases: a relative clause construction, and a noun-complement construction:

(54)



In Chomsky (1973), it is shown that the effect of the CNPC also is a consequence of the more general Subjacency Condition together with the lack of a COMP node in NP. Therefore, the derivation of a sentence like:

(55) *it is a book that I haven't met anyone who likes.

is blocked by the Subjacency Principle, as indicated in (56):

(56)
 ...that I haven't met [NP anybody [S' who [S t likes which]]]

While some languages do not allow extractions out of either of the constructions in (54), others allow extractions out of noun complements but not out of relative clauses. There are also languages such as Swedish which show good extractions out of both types (cf. Engdahl (1982)). So, contrary to what was assumed earlier, the constraints against movement from NPs can not hold as universal restrictions on syntactic rules any more, even though of course they could be valid for those specific languages for which they were originally proposed (particularly English).

In the following sections, I will examine the properties of each of the relevant constructions in Persian. We will see how differently the two categories behave with respect to extraction patterns. Given this fact, it will be argued that Persian extraction facts can not be captured by a unified and straightforward account. Let us consider each of these structures in turn.

Extractions out of Relative Clauses

Extraction out of most relative clauses is prohibited in Persian, as illustrated in the following set of examples:

(57)

- a. *in un film₁-e ke amir doxtar-i-ro ke --₁ dust dar-e
 this that movie-is that Amir girl-the-Do that like have-3S
 mi-šnas-e.
 CM-know-3S
 "it is that movie that Amir knows the girl who likes."

- b. *in no? hava₁-st ke man ketab-i be doxtar-i ke --₁
 this kind weather-is that I book-a to girl-a that
 tarjih mi-d-e dad-am.
 preference CM-give-3S gave-1S
 "it is this kind of weather that I gave a book to a girl who
 prefers."
- c. *un šokolat₁-e ke bačče-i ke --₁ dust dar-e be
 that chocolate-is that child-a that like have-3S to
 forušgah umad.
 store came
 "it is that chocolate that a child that likes came to the
 store."

The data in (57) represents cases which I have no doubt that all speakers of Persian find absolutely ill-formed. Notice that inserting a resumptive pronoun also does not ameliorate the above sentences with the unacceptable extractions. This shows that resumptive pronouns can not indiscriminately patch-up ill-formed sentences in Persian:

(58)

- a. *in un filmi-ye ke Amir doxtar-i-ro ke un-o dust dar-e mi-
 šnas-e.
- b. *in no? hava-st ke man ketab-i be doxtar-i ke tarjih mi-d-e
 dad-am.
- c. *un šokolat-e ke bačče-i ke dust dar-e be forušgah umad.

This property that structures as in (57) are not remedied by resumptive pronoun is also observed in the cases of Danish (Erteschik-Shir (1982), and Swedish (Engdahl (1982), Zaenen & Maling (1982)). For Persian, therefore the generalization that resumptive pronouns and gaps are in complementary distribution does not hold, since I have found no cases of unacceptable extractions which show a noticeable change in their status by using a resumptive pronoun.

Although the examples in (57) indicate the ill-formed cases, it is not difficult to find instances with extractions out of relative clauses which sound possible. Similarly, this property is also found in Danish (Erteschik-Shir & Lappin (1979), Erteschik-Shir (1982)). That is while extraction from structures such as (57) is blocked in Danish, it is permitted from a limited number of

relative clauses. The possibility of extractions out of a rather small subset of relative clauses in Danish is shown in (59) (=2 Erteschik-Shir (1982)), and (60) represents the equivalent structures in Persian:

(59=2] Danish

- a. Det₁ er der mange der kan lide --₁.
that there are many who like
- b. Det₁ kender jeg mange der kan lide --₁.
that know I many who like
- c. Det₁ har jeg mødt mange der har gjort --₁.
that have I met many who have done
- d.?Det₁ har jeg spurgt mange der har gjort --₁.
that have I asked many who have done
- e.*Det₁ har jeg drillet mange der har gjort --₁.
that have I teased many who have done
- f.*Det hus₁ kender jeg en mand som har købt --₁.
that house know(stressed) I a man who has bought

(60) Persian

- a. un-o₁ keyli-ha hast-and ke --₁ dust dar-and.
that-DO many-PL is-PL that like have-3PL
- b. un-o₁ keyli-ha-ro mi-šnas-am ke --₁ dust dar-and.
that-DO many-PL-DO CM-know-1S that like have-3PL
- c. un-o₁ keyli-ha-ro diđ-am ke --₁ anjam dad-and.
that-DO many-PL-DO saw-1S that performance gave-3PL
- d.?un-o₁ az keyli-ha porsid-am ke --₁ anjam dad-and.
that-DO from many-PL asked-1S that performance gave-3PL
- e.*un-o₁ keyli-ha-ro ke --₁ anjam dad-and azyat
that-DO many-PL-DO that performance gave-3PL teasing
kard-am.
did-1S
- f.*un xuna-ro₁ mard-i-ro ke --₁ xaride mi-šnas-am.
that house-DO man-a-DO that bought CM-know-1S

Both (59) and (60) exhibit two identical hierarchies in which the degree of acceptability gradually decreases. However, notice that speakers in both languages may differ as to the cut-off point. The data indicate that extractions out of the same structure thus can

produce either fully acceptable sentences, as in (59/60a-c), or questionable (59/60d) and even ill-formed sentences, as in (59/60e-f). So, it seems reasonable to question which way to state the generalization for Persian? Either we should say that all extractions out of relative clauses are ungrammatical and that acceptable sentences such as (59/60a-c) happen to be felicitous for probably some non-syntactic reason, or should it be assumed that extractions in general are possible and that the ungrammatical cases are infelicitous for some non-grammatical reason. Clearly, the more and the less acceptable cases do not show any syntactic distinction. This leads us to assume that extractions out of relative clauses in Persian are not blocked by any rule in the grammar, and so is not predictable by any structural constraint as well as with lexical properties of the matrix verb. Such lexical variation, however, does not work for extractions out of indirect questions. A descriptively appropriate explanation must therefore lie outside the domain of syntax. In the spirit of Erteschik-Shir(1982), I propose that the pragmatic notion of Dominance is needed to explain the extraction facts in Persian. In Erteschik-Shir & Lappin (1979), Dominance has been defined as:

DOM: A constituent *c* of a sentence *S* is dominant in *S* if and only if the speaker intends to direct the attention of his hearer(s) to the intension of *c*, by uttering *S*.

We can deduce that Dominance, according to this definition is a discourse property assigned to a constituent in a context of conversation. This also entails that not all constituents of a sentence may be dominant. In order to identify the dominance possibilities of embedded and matrix sentences, Erteschik-Shir suggests a procedure which involves placing the entire complex sentence in a context of direct discourse and denying first the matrix sentence, and then the embedded sentence. If denying the complement clause does not turn out to be possible, then the possibility of interpreting the complement as dominant will be excluded. Let us consider the following example, given by Erteschik-Shir, which shows that not only the matrix clause can be

naturally interpreted as dominant, but also the dominant interpretation of the complement is possible. In view of this fact the dominant constituent, as Erteschick-Shir points out, is an indicator of a possible future topic of conversation in the hearer's immediate response:

- (61=4) Speaker A: I know many people who like that.
 Speaker B: a. That's a lie...you don't
 b. That's a lie...nobody likes that.

In contrast, example in (62) shows the reverse case in which the dominant interpretation of the complement is forbidden:

- (62=5) Speaker A: Peter knows the girl who likes the movie.
 Speaker B: a. That's a lie...he doesn't
 b.*That's a lie...she doesn't.

The tests of Dominance as in (61) and (62) work in exactly the same way in Persian as in English and Danish, and triggering the same results. Given the fact that some but not all relative clauses can be dominant, Erteschick-Shir (1973) proposes that this distinction is the basic factor which motivates the possibility of extraction. In other words, the condition on extraction stems directly from the pragmatic source of dominance, and is formalized as follows:

Dominance Condition: An NP can be extracted out of a clause iff the clause may be interpreted as dominant.

This Condition predicts that, the relative clause in (61) should allow extraction, while the one in (62) should not. The examples in (60b) and (57a) illustrate that this prediction matches neatly with the facts in Persian. The same correlation also obtains between the results of the dominance test and between the extraction facts in other sentences.

A further point which Erteschick-Shir raises concerns the factors such as the definiteness of the head noun of the relative clause or the verb in the matrix which play a role in dominance and hence in extraction.⁵ It should be noted that extraction out of most relative clauses is blocked. The reason for this could simply be that the speaker employs a relative clause in order to merely

enable the hearer to pick out the referent of the head NP, not to persuade him to focus on the content of the relative clause. Obviously, such a clause gives rise to a non-dominant interpretation. However, if the matrix serves only to introduce the head of the relative clause, then the relative clause is interpreted dominantly. This is shown in the example below in which the response (a), indicates that the whole sentence may be interpreted as dominant. (b) demonstrates that the relative clause by itself can be dominant. However, the unacceptability of the (c) response refers to the fact that the matrix clause by itself cannot be dominant:⁶

(63) Speaker A: xeyli-ha hast-and ke un-o dust dar-and.
 many-PL is-PL that it-DO like have-3S
 "there are many who like it."

Speaker B: a. eštebah-e, kasi nist.
 wrong-is, someone isn't
 "that's not true, there isn't."

b. eštebah-e, hič-kas un-o dust nadar-e.
 wrong-is, nobody it-DO like NEG-have-3S
 "that's not true, nobody likes it."

c. *eštebah-e, xeyli-ha nist-and.
 wrong-is, many-PL isn't-3PL
 "That's not true, there aren't many."

As a result, any matrix which is interpreted as introducing the head of the relative clause, allows the relative clause to be interpreted as being the dominant part of the sentence. Interestingly enough, the more the matrix gets complex, the less likely such an interpretation will be, as the hierarchy in (60) illustrate this fact.

The choice of the matrix verb apparently also influences the acceptability (cf. Allwood (1976, 1982), Engdahl (1986)). Some speakers who cannot accept (64a) and (64b), can accept (64c) and (64d):

(64)

a. *be in gol-a be doxtar-i fekr mi-kon-am
 to this flower-PL to girl-a thought CM-do-1S
 ke ab mi-d-e.
 that water CM-give-3S

"these flowers, I am thinking about a girl who is watering."

- b. *be in gol-a doxtar-i-ro davat mi-kon-am
to this flower-PL girl-a-DO invitation CM-do-1S
ke ab mi-d-e.
that water CM-give-3S
"these flowers, I invite a girl who is watering."
- c. be in gol-a doxtar-i-ro mi-šnas-am ke ab mi-d-e.
to this flower-PL girl-a-DO CM-know-1S that water CM-give-3S
"these flowers, I know a girl who is watering."
- d. be in gol-a doxtar-i-ro mi-bin-am ke ab mi-d-e.
to this flower-PL girl-a-DO-DO CM-see-1S that water CM-give-3S
"these flowers, I see a girl who is watering."

However, as Erteschik-Shir argues this also follows from the definition of dominance. That is the complexity of the verb, draws the attention of the hearer, and blocks a dominant reading of the embedded clause.

The nature of the subject should be regarded as another factor playing a role in dominant interpretation. It has been argued that the interpretation of the matrix as introducing the rest of the sentence will be harder when the matrix subject introduces another figure into the discourse:

(65)

?piyaderavi-ro₁ amir keyli-ha-ro mi-šnas-e ke --₁ dust dar-and.
hiking-Do Amir many-PL-DO CM-know-3S that like have-3PL
"hiking, Amir knows many who like."

However, the sentence will be improved if the subject is stressed and is interpreted contrastively. (65), thus would be appropriate in a context where the speaker is asked whether he knows anyone who likes hiking. If the response is negative, he may add that, in contrast, Amir does. The contrastive interpretation of Amir accounts to some extent for the extraction possibility:

(66)

piyaderavi-ro Amir keyli-ha-ro mi-šnas-e ke dust dar-and.

Erteschik-Shir & Lappin (1981) have argued that "what is dominant in cases of contrast is not the stressed constituent as such, but rather its identification as a member of a set with which it is contrasted". This means that a contrasted element in the matrix

makes it possible for another element in the embedded clause to be paired with it as a dominant pair. On the other hand, if the contrastive reading of a stressed element in the matrix is not permitted, then the matrix should be interpreted as dominant, and extraction out of the embedded clause will be impossible. Clearly, cases of contrast are not easily identifiable, since the set to which the element belongs is usually not specified in the sentence. Cases of contrast along with enumeration have been argued to operate on finite sets, and are referred to as cases of "restrictive dominance". The important characteristic of both cases is that they can apply to relations among objects.

One more point which appears to be related to the argument of the interaction of restrictive dominance with the regular dominance is the issue of the definiteness of the head noun of the relative clause. Generally speaking, extraction out of relative clauses with definite heads are less acceptable than extractions with indefinite or even generic heads in languages such as Swedish, Danish. (67) illustrate that this also holds for Persian:

(67)

- a. Amir-o hic-kas ne-mi-šnas-am ke -- dust daste base.
- b.? Amir-o ye mo?allem mi-šnas-am ke -- dust dare.
- c.?? Amir-o mo?allem-i-ro mi-šnas-am ke -- dust dare.

Amir no one don't know-I that likes.
 Amir a teacher know-I that likes.
 Amir the teacher know-I that likes.

In his discussion of which factors influence the acceptability of sentences with extractions in Swedish, Andersson (1982) demonstrates that it is not the definiteness of the head NP per se, which is the cause for the acceptability judgement. It is rather whether or not a definite NP is felicitous in a given sentence, and in a given discourse. The central point is hence that if the sentence or discourse requires a definite NP, then extraction out of the relative clause would be possible, provided that the extraction is also motivated by the structure of the discourse. This claim is supported by the example below which in comparison with (67) shows the opposite pattern in the ranking of the

sentences:

(68)

a.?? un tasvir-o hic kas-o ne-mi-šnas-am ke -- kešide baše.

b.? un tasvir-o ye mard mi-šnas-am ke -- kešide.

c. un tasvir-o mard-i-yo mi-šnas-am ke -- kešide.

that picture-DO no one don't know-I that has painted

that picture a man know-I that has painted.

that picture the man know-I that has painted

To get the case in (c), according to Andersson's analysis, we require a context where a number of pictures are correlated with a number of artists. This interestingly means that we have restrictive dominance. In other words, there are a finite set of pictures and a finite set of artists, that allow us to express the correlation of these two sets by means of extraction in (c). On the other hand, since no one is not a good candidate to be inserted in an enumerative list, (a) sounds odd. To summarize, the dominant interpretation of the relative clause and hence extraction will be more feasible, when the head of the relative clause is indefinite. In the case of definite head noun, the dominant interpretation and consequently extraction will be possible only if the context allows for a contrastive or enumerative reading.

In general, the conclusion we may draw from the argument in this part is that it would be inappropriate to capture the extraction processes in Persian with constraints based on syntactic structure which need to be explicitly formulated in a grammar for this language. Instead I suggested that apparent restrictions on extractions out of relative clauses are due to limitations which may be analyzed in terms of the pragmatically based principle of Dominance.

With respect to this type of extraction, we have seen that Danish and Persian behave similarly: in both languages the possibility of extractions is closely related to the possibility of dominance. Such a similarity may lead one to assume that dominance assignment is also the key factor determining the possibility of extraction in other constructions in both languages. This, however, is not the case. In what follows, we will argue that

concerning extraction out of embedded questions, Danish and Persian show distinct possibilities:

Extraction is possible out of most embedded questions in Danish. However, there are a few cases such as (69) which prevent extraction:

(69)

- a. *Hende₁ bestemte peter hvem der besøgte --₁.
her decided Peter who visited
- b. *Eksamenen₁ var det klart hvordan peter havde klaret --₁.
the exam was it clear how Peter had done

Here again, as Erteschik-Shir suggests, the possibility of extraction is dependent on the dominance factor. By applying dominance, she shows that the extraction out of an embedded question will be possible, if it can be interpreted as being dominant. On the contrary, in sentences such as (69), the embedded question may not be dominant, hence the impossibility of extraction (for a more detailed discussion of this issue, see Erteschik-Shir, 1982. p.181).

In Persian, on the other hand, the equivalents of the sentences in (69) are well-formed:

(70)

- a. un-o₁ amir tasmim gereft ke ki --₁ bebin-e.
that-DO Amir decision took that who see-3S
- b. emtehan-o ma?lum bud ke četor amir dade bud.
exam-DO clear was that how Amir given was

Although the results of the dominance test appear to be parallel for speakers of both languages, the acceptability of (70) versus the unacceptability of (69) indicates that dominance and extraction are not in a direct correlation in all cases of embedded questions in Persian; otherwise we would have the sentences in (70) as ungrammatical rather than grammatical. Relative clauses thus contrast with embedded questions in Persian in that the same notion of dominance correctly explains extraction out of the first one, but not the latter. The fact that extraction is freely allowed in all cases of embedded interrogatives, therefore shows that dominance, unlike in relative clauses, does not play the

determinant role in these structures.

At this point, it remains to show that if it is the case that dominance explains the Persian extraction facts with relative clauses, why is it then that the same account does not apply to embedded questions? In order to establish my explanation, I would like to propose with Erteschik-Shir that when there is a case in which a certain structure is not homogeneous in its dominant properties, as in the cases of the relative clauses and embedded questions in both Danish and Persian, the situation can be resolved in the following ways:

(71)

- a. The extraction facts accord with the dominant condition.
- b. Extraction is blocked in all cases of the structure in question. i.e., a syntactic constraint is imposed.
- c. Extraction is possible in all cases, i.e., the dominance constraint is violated and no syntactic constraint is imposed.

Presumably, Danish represents the solution in (a) in both relative clauses and embedded interrogatives. Languages such as English, exhibit the solution in (b). It should be noted that in English, the evidence of the role of dominance is found to distinguish the English cases as well, yet there is no tight correlation between the dominance condition and extraction. The fact that (72a,b), and the like, despite the general impossibility of extraction out of relatives in English, are preferable to all other cases, indicates that dominance is still operative in this language:

(72)

- a. this is the kind of weather_i that there are many people who like --_i.
- b. ?this is the kind of weather_i that I know many people who like --_i.
- c. *this is the kind of weather_i that I am familiar with many people who like --_i.
- d. *this is the kind of weather_i that Peter knows the girl who likes --_i.

Similarly, the examples in (73), unlike the unacceptable sentences in (74), are cases where the dominance test would indicate that the embedded questions can be used dominantly, and hence they should be acceptable:

(73)

- a. this book, I am not sure whether or not I should read.
- b. this is something that I don't know what I should do about.

(Kuno & Robinson (1972))

(74)

- a.*this book₁ Peter decided who should read --₁.
- b.*this is something₁ that it is quite clear what I should do about --₁.

While the test results are more or less identical for speakers of the two languages, extraction out of both relatives and embedded questions is considered much less acceptable for most speakers of English than it is in Danish. The marginality of the extraction cases in English indicates that a process of grammatization has occurred to prevent extraction in all cases of these structures. This, therefore means that the situation in English may be explained by solution (b).

With respect to relatives in Persian, it is clear that solution (a) is favoured. So, similar to Danish, the possibility or impossibility of extraction is straightforwardly related to the dominance condition. Concerning embedded questions, we have seen above that extraction does not always accord with the dominance, so that even non-dominant interrogatives allow extraction. This means that solution (c) is the one to be chosen in these cases, and that no syntactic constraint is imposed. There is therefore a noticeable difference in selecting each of the choices in (71a-c) between languages which follows from the interaction of dominance condition and syntax in each language.

The question to be answered now is why languages like Persian and Danish on the one hand, and languages such as English, on the other, differ in the choice of tactic to resolve cases where dominance is not assigned homogeneously to a certain syntactic structure. The answer, as pointed out by a number of linguists (Allwood (1982), and Erteschik-Shir & Lappin (1979)) among others, should be related to the fact that Persian and Danish employ topicalization process with much more frequency than English.⁷ The fact that Persian speakers are so thoroughly accustomed to topicalization, facilitates the perception of the connection

between the extracted element and the extraction site, as indicated in (75):

(75) A...[...B...] (A=extracted element, B=extraction site)

As the same tactic is employed in the perception of sentences which exhibit extractions out of relatives and embedded interrogatives, these sentences turn out to be more acceptable for those speakers, i.e. Persian and Danish speakers who use (75) more extensively.

Extraction out of Noun Complements

In the previous sections, we have discussed that structurally based constraints such as the Wh-island and the Complex NP Constraint are neither necessary nor sufficient to account for the extraction processes in Persian. Given this fact, it was argued that while a merely structural account of extraction out of relative clauses and embedded interrogatives appears to be totally appropriate for languages like English, it is not adequate enough to work in Persian. Instead I suggested that apparent restrictions on extraction patterns are due to processing and pragmatic limitations which hold for all languages and which need not be explicitly incorporated in the syntax of a particular language. Consequently, the notion of Subjacency does not seem to be relevant to a proper characterization of movement rules in Persian. We have seen that even with the parametric approach to the formulation of restrictions on movement, Persian extraction facts do not fall under the pattern expected if they were the result of a syntactic movement rule. However, looking at the second category of CNPC, i.e. extraction out of Noun Complements, such a general judgement turns out to be somewhat hasty. Since extractions out of Noun Complement structures are equally ill-formed in Persian as in English, as shown in the examples below:

(76)

- a. *bara-ye un šoql hič kas in emkan-ro tasavvor
 for-EZ that job no one this possibility imagination
 ne-mi-kard ke ali taqaza mi-kon-e.
 NEG-CM-did that Ali application CM-do-3S
 "for that job no one imagined this possibility that Ali would
 apply."
- b. *un kargar-o hame be xater-e in haqiqat ?asabani bud-and
 that worker-DO all to sake-EZ this fact angry was-3PL
 ke reyis-e karkune exraj karde.
 that chairman-EZ factory expelling done
 "that worker, all were angry over this fact that the chairman
 of the factory has expelled."
- c. *dust-et ke in šaye?e-ro šenide bud-am ke bimar
 friend-your that this rumor-DO heard was-1S that sick
 šode, diruz be xune-ye ma umad.
 become yesterday to home-EZ we came
 "your friend, about whom I had heard this rumor that has got
 sick, came to our home yesterday."
- d. *ki-yo to in xabar-o bavar ne-mi-kard-i ke dar emtehan
 who-DO you this news-DO belief NEG-CM-did-2S that in exam
 mardud šode ?
 failed become
 "who didn't you believe this news that has failed the exam?"
- e. *yasaman-o ke to in vaqeiyat-o qabul na-dašt-i
 Yasaman-DO that you this reality-DO acceptance NEG-had-2S
 ke keyli bahuš-e...
 that very smart-3S
 "Yasaman whom you didn't accept the reality that is very
 smart..."

The ungrammatical status of the data above indicates that extractions out of Noun Complement structures are less susceptible to pragmatic and contextual constraints than extraction out of relative clauses. This means that Persian does not show the same tendency as other languages like English and Swedish do: In Swedish, extractions out of Noun Complements, unlike extraction out of relatives, are perfectly grammatical in all cases (cf. Engdahl (1982)). Furthermore, extractions out of certain Noun Complement structures are acceptable in English, whereas extractions out of tensed relative clauses in general are impossible, as illustrated by the contrast between (77) and (78):

(77)

*who₁ did John make [_{NP} a claim [_s that John saw --₁]]
 (Chomsky 1977: fn 10)

(78)

*who₁ did Harry make the claim₁ that John had suggested --₁ to
 --₁ ?

So, the general assumption that Noun Complement constructions are inherently easier to process, does not hold for Persian. We could perhaps propose that this restriction is also pragmatically motivated, but it has been grammaticized to a more noticeable degree than the previous cases. On the basis of this analysis, I suggest that the situation in extractions with Noun Complements may be resolved in the following way, as stated before in (71):

(71)

b. Extraction is blocked in all cases of the structure in question, i.e., a syntactic constraint is imposed.

All the cases of the extractions considered in this chapter, thus show that the extraction patterns are distributed in a very different manner in Persian which itself raises a number of open questions. Moreover, the issue of how an originally pragmatically based constraint turns into a grammatical constraint is intriguing, and deserves further research.

Notes

1- A similar analysis, with some modification holds for Bhasa Indonesia as well. However, I have restricted my analysis to Egyptian Arabic.

2- Lasnik & Saito (1984) also have made the same point with respect to Polish.

3- This, as Rizzi states is to be related to the fact that in Italian, non-echo questions with more than one Wh-constituent are often ill-formed, or at least highly marginal. This characteristic does not hold for Persian. So, question formation out of an embedded question may be expected to be fine.

4- A similar approach has been suggested for Hebrew which allows multiple extractions, by Reinhart (1980).

5- On the basis of these factors, linguists such as Engdahl (1981b, and 1982) have argued that extractions out of relative clauses are more marked than extractions out of embedded questions in some Scandinavian languages including Swedish. The acceptability of extractions out of relative clauses, for instance, vary with the definiteness or indefiniteness of the head noun as well as with lexical properties of the matrix verb. Such lexical variation, however, does not work for extractions out of indirect questions.

6- The matrix by itself can be dominant, only if the NP quantifier *keyli-ha* "many" is stressed and is interpreted contrastively.

7- Andersson (1982) also mentions the same distinction between Swedish and Norwegian, on the one hand and English, on the other.

Chapter 4

MULTIPLE WH-FRONTING

This chapter offers an explanation of multiple wh-fronting, the characteristic of which is that all of the Wh-words occur sentence-initially. Before going into the major line of our argument, I will give a brief review of the literature regarding the topic in question. In this review, I will present some of the hypotheses originating in the early 1970s. Though it has been shown by more recent treatments that not all of the points proposed in these hypotheses can be taken as acceptable and appropriate, however in order to provide a conspicuous perspective of this issue, I have found it reasonable to start with this preliminary discussion.

Historical review of the literature

It has been assumed by several linguists (e.g. Chomsky (1964,1971), Baker (1970), Bresnan (1970) and Bach (1971)) that languages fall into two categories with respect to Wh-question movement:

1. Languages which do not have Wh-question movement (e.g. Japanese).
2. Languages which have Wh-question movement (e.g. English).

Furthermore, based on examples of English, it was assumed that if there is more than one Wh-word in the clause, only one of them can be moved to the clause-initial position. This was believed to be a language universal. However, as Wachowicz (1974) showed, there are languages like Polish and Russian which are counterexamples to the hypothesis that at most one occurrence of Wh-movement is allowed in a single clause, as all the Wh-elements in those

languages must occur at the beginning of the clause. Thus, more than one Wh-constituent can move.

Multiple question movements can not be explained within Baker's (1970) Q-morpheme hypothesis either. This hypothesis is based on the data given in Greenberg (1966, pp. 81-83) In the languages that Baker considered, only one question word could be moved to sentence-initial position. To account for these facts, Baker proposed the following mechanism:

1. There is an abstract Q-morpheme which in the SVO and the VSO languages is placed at the beginning of the sentence, and in the SOV languages at the end of the sentence.
2. If we view the Q-movement transformation as a replacement of the Q-morpheme, then there can only be a single replacement of the Q-morpheme.
3. Morphemes such as "whether" and "if" are lexical realizations of the Q-morpheme.

It follows from the above assumptions that only one question-movement per simple sentence is possible. That means that after a single movement has taken place, the Q-morpheme has been replaced, and as there is no Q-morpheme left after the replacement has occurred, it would be impossible to perform another Wh-movement. Wachowicz (1974, p. 157) presented three sources of counterexamples to Baker's version of the Q-morpheme:

1. Languages that perform more than one question-movement.
2. The SVO or the VSO languages that do move the Wh-phrases to the sentence-initial position, but retain their yes-no particles.
3. SOV languages that do not move their wh-words to the beginning of the sentence, but in spite of that, do not retain their question particles in Wh-questions.

In addition, it is well-known that German is a SOV language, and as Wh-movement is leftward in German, its Q-morpheme should be on the left of the clause.

Kuno and Robinson (1972) who attempted to explain question-movement in language-specific terms, also noticed that the constraints on question-movement are similar to the constraints on topicalization. In English, for example, only one question-movement in a simple sentence is possible, and one can topicalize

only once. Therefore, the following sentences (borrowed from Wachowicz), where either double question movement (1) or double topicalization (2,3) has occurred, are all ill-formed:

- (1=29) *what to who(m) did Herbert give 0 0 ?
 (2=30) *this teddy-bear to Monica Herbert gave 0 0 ?
 (3=31) *to Monica this teddy-bear Herbert gave 0 0 ?

Based on this observation, Kuno and Robinson suggested a double dislocation constraint that prohibits dislocation of more than one constituent from its base position. According to this constraint, questions like (1=29) are possible in languages that can have declarative orders like (2=30) or (3=31). This is true of Persian:

- (4) či-ro be ki pedar-bozorg dad ?
 what-DO to who grandfather gave
 "what did grandfather give to who(m)?"
- (5) in ?arusak-o be sanaz pedar bozorg dad.
 this doll-DO to Sanaz father grand gave
 "this doll the grandfather gave to Sanaz."

Although the double dislocation constraint could be regarded as a constraint on possible orders in languages, it is unable to explain why many languages do not permit a succession of several Wh-words. For instance, as is pointed out by Wachowicz, the Finnish sentences (6) and (7) where the SOV orders are good in a declarative structure but unacceptable in a question, can not be accounted for consistent with this constraint. Notice that Finnish has an unmarked surface order SVO:

- (6=34) herbert teddy-karbun antoi.
 Herbert teddy-bear gave
 "it was Herbert who gave the teddy-bear."
- (7=35) *kuka mitakin antoi ?
 who what gave
 "who gave what?"

As the examples above illustrate, Finnish can dislocate twice in declaratives, and yet it can move only one Wh-word per clause to the clause-initial position. This shows that although constraints on question-movement sometimes seem to be related to general

constraints on word order in a given language, Kuno and Robinson's double dislocation constraint can not explain facts about Wh-question movement in a general sense either. Thus, in order to provide an insightful explanation of the multiple Wh-fronting phenomenon, a different analysis is needed.

Multiple Wh-fronting languages

Different studies have shown that languages do not behave identically in the extent to which they place Wh-phrases in a clause-initial position. Languages such as English, normally place one and only one Wh-word in the specifier position of CP (8a). Other languages like Chinese and Japanese have all Wh-elements in-situ at S-structure (8b). There are also languages like French which can follow either the English or the Chinese pattern (8c) in direct questions:

- (8) a. what did you give to whom?
- b. ni xiang-zhidao lisi weisheme mai-le sheme ? (Huang 1982)
 you wonder Lisi why bought what
 "what do you wonder why Lisi bought (it)?"
- c. qu' as- tu donné à qui ?= tu as donné quoi à qui?
 what have you give to whom you have given what to whom

Turning to Persian, we see that although it is basically a Wh-in-situ language, with respect to multiple Wh-question, it does not function strictly like other in-situ languages such as Chinese. Persian also has, like the Slavic languages, the property of being able to move all Wh-elements to a clause-initial position at S-structure. Consider the examples from some Slavic languages (Bulgarian, Polish) and a Romance language (Romanian) in (9), and the data from Persian as exhibited by (10):

- (9) Bulgarian (Rudin 1986)
 a. koj kogo vidjal ?
 who whom saw
 "who saw whom?"

Romanian (Rudin 1988)

- b. cine cu ce merge ?
 who with what goes
 "who goes by what (i.e. means of transportation)?"

Polish (Wachowicz 1974)

- c. co komu monika dala ?
 what to whom Monica gave
 "what did Monica give to whom?"

(10) Persian

- a. ki ki-yo did ?
 who who-DO saw
 "who saw whom?"
- b. či-yo/ro be ki amir dad ?
 what-DO to who Amir gave
 "what did Amir give to whom?"
- c. ki ba ki sobat mi-kard ?
 who with who talk CM-did
 "who was talking with whom?"
- d. či-yo/ro be ki mo?allem pišnahad kard ?
 what-DO to who teacher proposal did
 "what did the teacher propose to whom?"

The above simple multiple Wh-questions (9) and (10) in which all of the Wh-words are fronted, look extremely similar. However, we will see in what follows that although multiple Wh-questions in Persian appear to be very similar to languages like Romanian and Bulgarian, on closer inspection they turn out to differ significantly in a number of ways.

It is worth noticing that in Slavic languages, if the second Wh-element does not front, the sentence is still well-formed but it can only be interpreted as an echo question, as in the following typical examples of Bulgarian and Polish:

(11) Bulgarian (echo reading only)

- a. koj e vidjal kogo ?
 who saw whom
 "who saw who?"

Polish (echo reading only)

- b. co monika kumo dala ?
 what Monica to whom gave
 "what did Monica give to whom?"

The echo question in Persian, however, does not necessarily arise from the occurrence of the second Wh-word in situ, and is straightforwardly dependent on a particular stress on a wh-phrase. So, in (10), for example, either of the Wh-words or both of them could involve an echo reading if produced with a more emphatic intonation, irrespective of the syntactic position in which they appear.

It should be pointed out that Persian does not manifest obligatory multiple Wh-fronting. Hence, in contrast to Bulgarian and Romanian, etc., all of the Wh-words may also remain in-situ:

- (12) a. amir čī-yo be ki dad ?
 Amir what-DO to who gave
 "what did Amir give to whom?"
- b. moʔallem čī-yo be ki pišnahad kard ?
 teacher what to who proposal did
 "what did the teacher propose to whom?"
- c. pedar ki-yo be ki moʔarrefi kard ?
 father who-DO to who introduction did
 "who did the father introduce to whom?"

In addition, the formation of Persian multiple Wh-question allows another possibility in that the first Wh-phrase moves to the clause-initial while the second one does not undergo movement and stays in situ:

- (13) a.? čī-yo amir be ki dad ?
 what-DO Amir to who gave
- b.? čī-yo moʔallem be ki pišnahad kard ?
 what-DO teacher to who proposal did
- c.? ki-yo pedar be ki moʔarrefi kard ?
 who-DO father to who introduction did

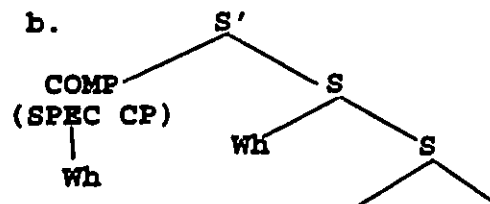
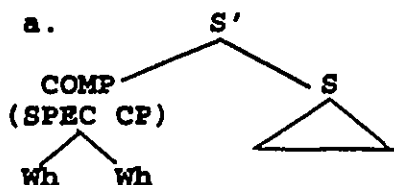
The different processes employed in (10), (12) and (13) indicate that there is no strong restriction imposed on Persian multiple question formation. However, all the Persian speakers are in agreement that a multiple question with the Wh-phrases in situ (i.e. 12) -contrary to the forms in (10) and (13) which are almost marginally accepted- is the most well-formed and usual

construction, as is the case with the Persian single questions discussed earlier. It should be noted that with respect to the structures (10) and (13), the judgements of my informants were not always the same in all cases. But overall, many of the native speakers including myself, consider the multiple Wh-questions in (10) to be more legitimate. This indicates that the multiple Wh-questions containing two adjacent Wh-elements give rise to a more acceptable context than ones involving separate Whs. Clearly, the structure in (10) is the one which I am primarily concerned with here.

As noted before, Slavic languages as well as Persian multiply front Wh-words. The apparent similarities between these languages leads us to question whether this similarity manifested by the multiple Wh-fronting goes beyond surface resemblance or not. In other words, can the multiple Wh-fronting phenomenon in the relevant languages be accounted for by a unified analysis? To explore this issue, we have to turn to more complex multiple Wh-constructions in order to provide further evidence. We will argue that despite the superficial similarities observed in all these languages, the nature of Persian multiple Wh-fronting is absolutely different. So, the analyses proposed for the Slavic and Romance languages fail to capture this construction in Persian.

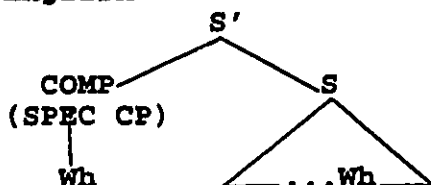
The fronted Wh-phrases in multiple fronting languages do not necessarily land in the same place. Thus, these languages can be divided into different types in terms of where the words are at S-structure. There are analyses in which the Wh-phrases are all in the same S-external position: COMP or specifier of CP (Comorovski 1988, Rudin 1982). Other scholars suggestions have been based on the hypothesis that some or all of the Wh-elements are in an S-internal position, perhaps adjoined to S (Reinhart 1982, Toman 1981, Lasnik & Saito 1984). The two possible structures are given schematically in (14) with non-multiple-fronting structure in (15) for comparison:

(14) Multiple wh-fronting languages

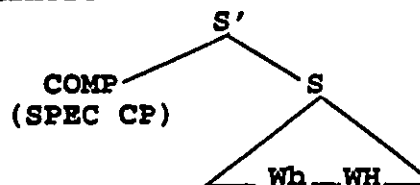


(15) Non-multiple wh-fronting languages

a. English



b. Chinese



As Rudin (1988) argues, both (14a) and (14b) exist at S-structure in different languages: in some multiple Wh-fronting languages all Wh-words are in SPEC CP at S-structure, while in others only one Wh-phrase is in SPEC CP, and the others are adjoined to IP. She labels them as [+multiply filled SPEC CP] ([+MFS]) languages and [-multiply filled SPEC CP] ([-MFS]) languages. The former includes Bulgarian and Romanian. The latter includes Serbo-Croatian, Polish and Czech. With respect to Persian multiple Wh-fronting, we will see that none of the representations in (14) could be regarded as an adequate structure. We will argue that Persian differs from other languages in several respects. These differences ultimately cluster in a way which leads to the conclusion that there is another type of multiple fronting language with quite different S-structure for multiple questions.

The [+MFS] languages, according to Rudin's argument differ from the [-MFS] languages in four ways:

1. [+MFS] languages can extract multiple Wh-words from a clause.
2. [+MFS] languages can violate Wh-islands.
3. Fronted Wh-words in [+MFS] languages can not be separated.
4. Fronted Wh-words in [+MFS] languages appear in a fixed order.

I will now discuss the above differences in some detail. To do so, I will examine the differences with respect to Slavic and Romance languages as well as Persian in order to find out how far

multiple Wh-fronting behaviour in Persian is characteristic of [+MFS] or [-MFS] languages.

1- Multiple Wh-extraction from a clause

Rudin notes that Bulgarian and Romanian ([+MFS] languages) can extract all of the multiple Wh-phrases from a clause, whereas in Polish, Serbo-Croatian and Czech [-MFS] languages- all but one of the Wh-words must remain in situ:

(16) Bulgarian

- a. koj kude mislis [ce e otisul - -] ?
 who where think-2S that has gone
 "who do you think (that) went where?"

Romanian

- b. cine cui ce ziceai [ca i -a promis - - -] ?
 who to whom what said-2S that to him has promised
 "who did you say promised what to whom?"

Serbo-Croatian

- c. *ko sta zelite [de vam kupi - -] ?
 who what want-2P. to you buy-3S
 "who do you want to buy you what?"

Polish¹

- d. *co komu maria chce, zeby janek kupil ?
 what to whom Maria wants that Janek buy
 "what does Maria want Janek to buy for whom?"

As demonstrated by the sentences in (16), the Bulgarian-type languages allow more than one wh-phrase to be fronted from an embedded clause, and the Polish-type languages do not. According to Rudin, sentences such as (16a) and (16b) support the contention that Bulgarian-type languages allow multiple adjunction to SPEC of CP in the syntax. (16c) and (16d), on the other hand, support the claim that the Polish-type languages do not permit multiple Whs to adjoin to SPEC CP except at LF. Therefore, a syntactic configuration like (17) is just possible in Bulgarian-type languages:

- (17) [_{CP} Wh₁ Wh_j ... [_{CP} [_{SPEC CP} t₁ [t_j]] ... e₁ ... e_j ...]]

Concerning Persian multiple extraction of Wh-words from a

clause, we will see that it is neither an absolute Bulgarian-type language nor a Polish-type one. However, the Wh-constituents in Persian multiple questions like Bulgarian-type languages may move up to the matrix initial position. Thus, as (18) indicates, long movement of multiple Whs is possible in this language:

- (18) a. ki koja fekr mi-kon-i (ke) - - rafte ?
 who where thought CM-do-2S (that) went
 "who do you think (that) went where?"
- b. ki či-yo be ki goft-i (ke) - - - qol dade ?
 who what-DO to who said-2S (that) promise gave
 "who did you say promised what to whom?"
- c. či-yo vase ki amir mi-xa-d (ke) sanaz - - bexar-e ?
 what-DO for who Amir CM-want-3S (that) Sanaz buy-3S
 "what does Amir want Sanaz to buy for whom?"
- d. ki-yo be ki dust-et mi-xast (ke) - - mo?arrefi
 who-DO to who friend-your CM-wanted(that) introduction
 bokon-e ?
 do-3S
 "who did your friend want to introduce to whom?"

Even Wh-phrases that originate in different clauses, can move up together, either to the initial position of the higher subordinate clause which is fully normal, as in (19),

- (19) goft-i (ke) ki či-yo - fekr mi-kon-e (ke) (to) -
 said-2S (that) who what-DO thought CM-do-3S (that) (you)
 xarid-i ?
 bought-2S
 "who did you say thinks you bought what?"

or to the matrix clause initial-position, which is only a marginally accepted form:

- (20) ?ki či-yo goft-i (ke) - fekr mi-kon-e (ke) (to) -
 who what-DO said-2S (that) thought CM-do-3S (that) (you)
 xarid-i ?
 bought-2S

The relative ungrammaticality of (20), compared with (19), shows that Persian has a restriction on how many clauses the wh-words may cross, and that multiple wh-movement across more than one CP boundary may not be taken as a normal construction.

Hence, in Persian movement of more than one Wh-phrase out of the clauses in which they originate is more or less normal, but of course not the most preferred structure. Furthermore, Persian, like Polish-type languages in which all but one of the Wh-words must stay in situ, allows the Wh-words not to appear adjacently. However, for most speakers this type of multiple question seems somewhat deviant, as the following sentences illustrate:

- (21) a. ?ki fekr mi-kon-i (ke) - koja rafte ?
 who thought CM-do-2S (that) where went
- b. ?ki goft-i (ke) - či-yo be ki qol dade ?
 who said-2S (that) what-DO to who promise gave
- c. ?či-yo amir mi-xa-d (ke) sanaz - vase ki bexar-e ?
 what-DO Amir CM-want-3S (that) sanaz for who buy-3S
- d. ?ki-yo dust-et mi-xast (ke) - be ki mo?arrefi
 who-DO friend-your CM-wanted (that) to who introduction
 bokon-e ?
 do-3S

Lastly, the sentences with two Wh-words fronted out of the embedded clause and one Wh-word left behind are also relatively acceptable in Persian, contrary to Polish-type languages. The evidence for this can be found in (22):

- (22) a. ?ki či-yo goft-i (ke) - - be ki qol dade ?
 who what-DO said-2S (that) to who promise gave
 "who did you say promised what to whom?"
- b. ?ki či-yo mi-dun-i (ke) - - az ki qarz gerefte ?
 who what-DO CM-know-2S (that) from who borrow took
 "who do you know borrowed what from whom?"

In sum, Wh-words extraction in a multiple question is not restricted in a unique manner. So, concerning this process, Persian make use of all the possibilities available for Slavic and Romance languages. Nonetheless, as pointed earlier, the different structures differ from each other with respect to the degree of acceptability.

2-Wh-islands

Given subadjacency as the explanation of Wh-islands, it is, based on Comorovski's (1986) proposal, predicted that a language that allows multiple Wh-elements in COMP at the level at which Wh-movement occurs do not respect any form of the Wh-island constraint. Given this hypothesis, Rudin maintains that Bulgarian-type languages can violate Wh-islands but Polish-type languages can not. In (23a) relativization out of indirect questions is shown to be permitted, and in (23b) movement of interrogative Wh-element out of an embedded Wh-question is fairly acceptable with heavy Wh-phrases:

(23) Bulgarian

a. vidjah enda kniga, kojato₁ se cudja [koj znae [koj
saw-1S a book which wonder-1S who knows who
prodava -₁]]
sells

"I saw a book which I wonder who knows who sells (it)."

b. ?koja ottezi knigi se cudis koj znae koj prodava ?
which of these books wonder-2S who knows who sells

"which of these books do you wonder who knows who sells?"

Rudin notes that the asymmetry between questioning and relativization is similar to what is found in some Scandinavian languages. Following Lie (1982), she assumes that some semantic or pragmatic constraints make it more complex to question than to relativize out of a question. Further support for the claim that Bulgarian-type languages can violate wh-islands is provided by examples such as (24) below from Romanian in which questioning out of an indirect question is acceptable:

(24) pentru care clauzai vrei sa afli cine nu a decis
for which paragraph want-2S to learn who not has decided
inca ce va vota -₁ ?
yet what will-3S vote

"for which paragraph do you want to learn who has not decided yet what he will vote?"

The situation in [-MFS] languages is reversed. Thus, neither interrogative nor relative Wh-movement is legitimate from inside a Wh-clause. Sentences in (25) are examples from Serbo-Croatian:

- (25) a. *sta si me pitao ko moze da uradi ?
 what have-2S me asked who can to do
 "what did you ask me who can do?"
- b. *...osoba, koja sam ti rekao gde (ona) zivi ...
 individual who have-1S you told where he lives
 "...the individual who you asked me where (he) lives..."

Turning now to Persian, we see that it behaves like a [+MFS] language. It has already been argued that Persian shares with Bulgarian-type languages the property of extracting multiple Wh-words out of an embedded clause. Therefore, we can make the hypothesis that it also does not respect Wh-island. This prediction holds, as we shall now see.

Examples (26a,b) illustrate the fact that relativization from inside a Wh-question is freely allowed. Recall that Persian lacks a relative Wh-pronoun, and in this case it simply employs the relative marker *ke*. Moreover, (26c) shows that even extraction from several interrogative Wh-clauses, though not elegant, is almost well-formed in this language:

- (26)
- a. so?al-i ke to ne-mi-dunest-i četor javab bed-i
 question tha you NEG-CM-knew-2S how answer give-2S
 moskel bud.
 difficult was
 "the question which you didn't know how to answer was difficult."
- b. še?r-i-ro ke u mi-dunest ki sorude xundam.
 poetry-the-DO that he/she CM-knew who composed read-1S
 "I read the poetry which he knew who had composed (it)."
- c. ketab-i-ro xund-am ke ne-mi-dun-am ki porsid (ke)
 book-the-DO read-1S that NEG-CM-know-1S who asked (that)
 ki mi-foruš-e.
 who CM-sell-3S
 "I read a book which I wonder who knows who sells (it)."

Furthermore, movement of interrogative Wh-words out of an embedded question is also felt to be grammatical. Although, as in Bulgarian and Romanian, it seems to be more difficult to construct acceptable sentences with interrogative Wh-movement than with relative pronoun. The examples in (27) illustrate this point:

(27)

- a. čī-yo az-am porsid-i (ke) ki extera karde ?
 what-DO from-me asked-2S (that) who invention did
 "what did you ask me who invented?"
- b. baraye kodum dust mi-dun-i (ke) ki tasnim gerefte
 for which friend CM-know-2S (that) who decision took
 čī-yo bexar-e ?
 what-DO buy-3S
 "for which friend, do you know who has decided what to buy."

In short, Persian allows Wh-movement of both interrogative Wh-phrases and relativization from inside an indirect question, and like [+MFS] languages does not respect Wh-islands. However, notice that the lack of Wh-island effects, as Rudin points out, "is not due to an absence of subjacency effects in general in these languages". As argued in the preceding chapter, Persian obeys Complex NP Constraint in the case of noun complementation, as indicated in (28):

(28)

- *nevisande-yi ke to in haqiqat-o bavar ne-mi-kard-i
 writer-the that you this fact-DO belief NEG-CM-did-2S
 ke mardom-e ziyadi dust-eš dar-and, jayze-ye nobel gereft.
 that people-EZ many like-him have-3PL prize-EZ Nobel took
 "the writer whom you didn't believe the fact that many people
 like (him) got the Nobel prize."

3-Constituent structure of fronted Wh-words

The third source of evidence for the claim that [+MFS] and [-MFS] languages differ rests on the difference in constituent structure of the Wh-word sequence. This argument is based on the possibility or impossibility of interrupting the Wh-phrases with parentheticals, clitics, adverbs, etc. That is to say, in Bulgarian-type languages nothing may intervene between the fronted Wh-words, while in Polish-type languages, various types of material can split the fronted Wh-word sequence. This clearly supports the idea that the fronted Wh-words do form a single constituent in the former group, but not in the latter, as can be seen in (29) and (30):

(29) Adverbs

a. zavisi ot tova, koj kogo pruv e udaril. (Bulgarian)
 depends on this who whom first has hit
 "it depends on who hit whom first."

b. *zavisi ot tova, koj pruv kogo e udaril.
 depends on this who first whom has hit

c. ko je koga prvi udario ? (Serbo-Croatian)
 who has whom first hit
 "who hit whom first?"

d. ko je prvi koga udario ?
 who has first whom hit

("Je" is a clitic auxiliary attached to the first Wh-element.)

(30) Parentheticals

a. *koj, spored tebe, kakvo e kazal ? (Bulgarian)
 who according to you what has said
 "who, in your opinion, said what?"

b. kto wedlug ciebie komy co dal ? (Polish)
 who according to you to whom what gave
 "who in your opinion gave what to whom?"

The data above indicates that while in Bulgarian-type languages, all of the Wh-phrases form an indivisible unit, there is a constituent break between the first and second Wh-words in Polish-type languages. This is predicted on the assumption that all of the fronted Whs in [+MFS] group are in SPEC CP, whereas in [-MFS] type, only the first Wh is in SPEC CP as a result of movement.

It remains to be determined to which type of languages Persian belongs. The different data illustrates that the characteristics of Persian Wh-words constituent structure appear to be more or less amenable to the same descriptive generalization proposed for [-MFS] languages; That means the fronted Wh-elements in this language can be separated by adverbs and parentheticals, as indicated in (31) and (32) which are the equivalent translations of the examples (29a,b) above respectively:²

(31) Adverbs

a. bastegi be in dar-e ke ki ki-yo avval zade.
 relation to this has-3S that who who-DO first hit
 "it depends on who hit whom first."

- b. bastegi be in dar-e ke ki avval ki-yo zade.
relation to this has-3S that who first who-DO hit

(32) Parentheticals

- a. ki či-yo, be nazar-e to gofte ?
who what-DO to opinion-EZ you said
"who, in your opinion, said what?"
- b. ki, be nazar-e to, či-yo gofte ?
who to opinion-EZ you what-DO said
- c. ki či-yo be ki, be nazar-e to dade ?
who what-DO to who to opinion-EZ you gave
"who, in your opinion, gave what to whom?"
- d. ki, be nazar-e to, či-yo be ki dade ?
who to opinion-EZ you what-DO to who gave

In (32d), the parenthetical follows the first Wh-word of the fronted Wh-sequence. However, unlike Polish, the occurrence of the parenthetical after the whole group of Wh-phrases (cf. 32c) as well as the second one of three is not blocked either, as given in (33). This also holds for the other [-MFS] languages such as Czech in which parentheticals can be placed anywhere in the Wh-phrase sequence. Nonetheless, it should be added that the structure involving unseparated Whs, as in (31a, 32a and 32c) are usually considered to be far better than those with separated whs (31b, 32b and 32d) by native speakers:

- (33) ki či-yo, be nazar-e to be ki dade ?
who what-DO to opinion-EZ you to who gave

On the basis of the examples in (31-33), we are led to conclude that Persian Wh-elements can be split up freely, and therefore do not form a unique constituent. Yet, this does not provide any conclusive evidence that the first Wh-word in Persian, like in [-MFS] languages, is in SPEC CP, as will be discussed later on.

Finally, intonation may be considered as a plausible evidence too. Cichoki (1983) notes that Polish exhibits an intonation break between the first and second Wh-phrases, but not between the second and third or third and fourth, indicating that the first Wh-word is

an independent constituent, while the second and third Wh-words together form a constituent.³ Persian, on the other hand, has an intonation break between all elements of a Wh-sequence. Thus, if we continue the line of analysis initiated by Cichocki, we could propose that the second and third Wh-elements in Persian do not turn out to be a single constituent, contrary to what is said to be the case for Polish.

4-Wh-word order

The fourth argument deals with the order of the fronted Wh-words. As Rudin shows, fronted Wh-words in the Bulgarian-type languages have a fixed order; In particular, they involve the order NOM>ACC>DAT, and Wh-adverbs (adjuncts) are preceded by non-adjuncts (Rudin 1985), as illustrated by examples (34). By contrast, the Wh-word order in Polish-type languages is free. So, accusative and nominative Wh-words, for instance, can occur in either order, as in (35) below:

- (34) Bulgarian
- a. koj kogo vizda ?
 who whom sees
 "who sees whom?"
 - b. *kogo koj vizda ?
 - c. koj kude e otisul ?
 who where went-3S
 "who went where?"
 - d. *kude koj e otisul ?
- (35) Polish
- a. kogo komu przedstawiles ?
 whom to whom introduced-2S
 "whom did you introduce to whom?"
 - b. komu kogo przedstawiles ?
- Czech
- c. kdo kdy koho pozval, nevim.
 whom when whom invited I don't know
 "who invited who when, I don't know."
 - d. kdy kdo koho pozval, nevim.
 - e. koho kdy kdo pozval, nevim.

Persian involves a relatively fixed order. So, it could be

treated on a par with [+MFS] languages. The data in (36) illustrates that the order of Wh-phrases is restricted at least to subject>non-subject, and that arguments should be followed by adjuncts. The point to note is that some speakers do not have a strict ordering. Thus, they might also use the deviant orders, albeit with some difference in emphasis and appropriate contexts. However, no Persian speaker doubts that the cases such as (36a,c,e,g) below should be taken as the only normal and perfectly acceptable Wh-word orders:

- (36) a. ki ki-yo dide ?
 who who-DO saw
 "who saw who?"
 b.*ki-yo ki dide ?
- c. ki koja mi-r-e ?
 who where CM-go-3S
 "whi is goinh where?"
 d.*koja ki mi-r-e ?
- e. ki-yo be ki mo?arrefi kard-i ?
 who-DO to who introduction did-2S
 "whom did you introduce to whom?"
 f.*be ki ki-yo mo?arrefi kard-i ?
- g. ki ki-yo key davat karde, ne-mi-dun-am.
 who who-DO when invitation did NEG-CM-know-1S
 "who invited who when, I don't know."
 h.*ki key ki-yo davat karde, ne-mi-dun-am.
 i.*key ki ki-yo davat karde, ne-mi-dun-am.

Before going to the more detailed discussion of the above differences, let's consider the S-structures for [+MFS] and [-MFS] languages, as proposed by Rudin:



(37a) is the structure of [+MFS] languages in which all the Wh-words are adjoined in SPEC of CP, and (37b), following Lasnik and Saito (1984), is assumed to be the structure of [-MFS] languages

which involve fronting one wh-word to SPEC of CP, and adjoining the other Wh-words to IP. The semantic differences in extraction possibilities for Wh-words, island effects, the position of clitics, parentheticals, etc. within the fronted Wh-phrase sequence, and strictness of multiple Wh-word order in the two types of languages, as Rudin argues, can be accounted for by the S-structures in (37): multiple extraction from an embedded clause is possible if adjunction to SPEC CP is allowed. In contrast, if adjunction to SPEC CP is not allowed, as in (37b), then subadjacency will be violated. Similarly, the possibility of multiple adjunction to SPEC CP permits a Wh-word to be extracted from an indirect question by passing through the embedded COMP. So, Wh-island violations would be voided. Furthermore, given (37a) for the [+MFS] languages, parentheticals will not intervene between Wh-elements. However it is possible for parentheticals to appear between the first and the rest of Wh-elements, given (37b). Lastly, with respect to word order, the adjunction pattern in (37a), as Rudin proposes, can account for the fixed ordering of Wh-phrases. The restrictions on Wh-word order, according to her is reminiscent of the Superiority effect in non-multiple-fronting languages like English. That means in multiple questions involving a subject Wh-phrase and an object Wh-phrase, the movement of the subject must take place prior to the movement of the object. This is demonstrated in the following English example:

- (38) a. I wonder who did what?
 b.*I wonder what who did?

This subject-object asymmetry is often explained by the ECP. That is at LF, e_i is properly governed in (39a), but not (39b), the LF representations of (38a,b) respectively:

- (39) a. ...[what_j [who_i]] e_i did e_j
 b.*...[who_i [what_j]] e_i did e_j

The fact that [+MFS] languages turn out to manifest the same structure for multiple questions at S-structure that non-multiple languages have at LF, leads Rudin to propose that restriction on

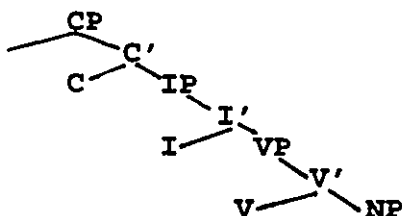
word order in [+MFS] languages falls under the ECP as well. She adopts the split ECP theory proposed in Aoun, Horstein, Lightfoot and Weinberg (1987):

- a. An empty element must be governed by a lexical head at PF.
- b. An A anaphor (e.g. a trace of a wh-word) must be A bound in its domain at LF.

Based on this, the Wh-word which is substituted into SPEC CP first triggers SPEC-head agreement (Chomsky 1986), and C having the same index as the first Wh-word can function as a head-governor. A subject Wh, thus has to move to SPEC CP before an adjunct, so that the former can be head-governed by the C co-indexed with the subject, and thereby satisfying the ECP. However this hypothesis, as Cheng (1991) convincingly argues is not capable of explaining the order [subject+adjunct]: adjunct must be head-governed too. So, why is its pre-subject movement to SPEC CP blocked? On the other hand, within Rudin's analysis, we face another problem: the order [adjunct+object] is expected to be good, contrary to the facts in Bulgarian. In short, Rudin's account easily handles [subject+object] order, but is not generalized to other types of order.

We have seen that the Wh-word order in Persian, similar to [+MFS] languages, is constrained. Both of these languages manifest the same c-command relationship between the subject and the object as well as that between arguments and adjuncts. We argued above that given Rudin's assumptions, the ordering of multiply fronted Wh-words can not be fully accounted for. Hence, I continue to assume, along the lines of Cheng (1991), that the ordering facts in both Persian and [+MFS] languages can be captured by the Principle of Economy of Derivation (Chomsky 1989). According to Chomsky, Superiority violations may be accounted by Economy of Derivation. Therefore in a structure such as (40),

(40)



movement from α to γ is a shorter derivation than movement from β to γ . Thus, the movement of β to γ will be ill-formed by Economy of Derivation. With respect to the fixed order of the [subject+object], if the object moves first, the output will be ruled out by Economy of Derivation. The movement of subject is shorter, and thus the order [subject+object] is favoured. Meanwhile, the order [argument+adjunct] could also be accounted for along this line, if we assume with Larson (1988) that adverbs are generated in the most embedded positions. Namely, given a uniform rightward binary branching structure, an adjunct is generated in the most embedded position, and so, is asymmetrically c-commanded by the subject and the object. By the Principle of Economy of Derivation, thus, the movement of the adjunct is predicted to take place after the movement of the subject and the object.⁴

Summary

To sum up the discussion so far, [+MFS] languages and [-MFS] languages as well as Persian, display different systematic properties in a number of ways, summarized in the following table:
(41)

	[+MFS]	Persian	[-MFS]
-Multiple Wh extraction from a clause	+	+	-
-Wh-island violations	+	+	-
-Parentheticals, adverbs after first Wh	-	-/+	+
-free NOM/ACC Wh-word order	-	-	+

As indicated in (41), concerning Wh-extraction from a clause, Wh-island violations and Wh-word order, Persian and [+MFS] languages pattern together. The only difference is that parentheticals, adverbs, etc. are not allowed to occur after the first Wh-element

in [+MFS] languages, whereas it is not totally ruled out in Persian. However recall that the most unmarked and well-formed construction, according to the judgements of the majority (if not all) of my informants, involves the presence of these materials in a post-Wh-word position. Given this fact, we can maintain the assumption that, by and large, Persian manifests a behaviour similar to that of Bulgarian-type languages with respect to the properties in question. In other words, it shows a [+MFS]-like setting for the characteristics cited in (41). So, one might expect the same S-structure holding for [+MFS] languages, to be attributed to Persian as well. Prima facie, there seems to be no reason to challenge this assumption. Yet, various considerations lead us to propose a different account for multiple Wh-fronting phenomenon in Persian. We will see that although both Bulgarian-type languages and Persian move all Wh-phrases to the beginning of the clause, the clause-initial landing sites and the resulting S-structures for multiple Wh-constructions are fundamentally different. In the following discussion, we will therefore lay out a number of facts which help to exactly characterize the S-structure of Persian multiple Wh-fronting:

Toward an account

As has already been discussed in detail (cf. ch.3), Persian is basically a Wh-in-situ language, although its Wh-phrases are free to appear everywhere, including initial position in a clause. We have argued that fronting of a Wh-word does not involve an obligatory syntactic movement, i.e. to SPEC CP in this language. Instead, Wh-word fronting was shown to be an instance of optional topicalization process. By contrast, the [+MFS] and [-MFS] languages exhibit obligatory syntactic Wh-movement. So, the other choices which are available for Wh-question formation in Persian, are not available in these languages. The same hypothesis is also extended to the multiple Wh-fronting in the relevant languages. That means fronting of Wh-constituents in both [+MFS] and [-MFS]

languages results from an obligatory syntactic operation, whereas that of Persian is again an optional movement.

Wh/indefinites in multiple fronting languages

A further crucial factor which is worth exploring concerns the question of what multiple fronting languages have in common which leads to the multiple fronting of Wh-words. The multiple Wh-fronting languages, as discussed below, share some common property which is absent from other languages. My purpose in this examination is to explore whether Persian shares the same feature with other multiple Wh-fronting languages or whether it behaves similarly to non-multiple Wh-fronting ones in lacking such a property. The account proposed here is based mainly on Cheng's (1991) analysis who assumes that the properties of Wh-phrases themselves in the multiple Wh-fronting languages contribute to the requirement of fronting all Wh-words.

The interrogative reading of the Wh-words in multiple fronting languages take a bare/core form, and the indefinite reading of the Wh-words are derived from the bare form with certain particles which are either prefixed or suffixed to the bare form. Consider, for example, the interrogative and the indefinite forms of Polish in the chart below:

(42)

Polish			

kto	who	ktos	someone
gdzie	where	gdzies	somewhere
kiedy	when	kiedys	sometime
jaki	what sort of	jakis	some sort of

The data from Polish, Bulgarian, etc. indicate that the multiple fronting languages always have an affix in the case of indefinite reading, while non-multiple fronting languages have no morphological alternations, as shown in (43):

(43)	Mandarin Chinese		
	shei	who	anyone
	sheme	what	anything

Persian differs from the multiple fronting languages in that there is no affix attached to the Wh-words in the indefinite reading. In other words, the derivation of the indefinite NPs from Wh-words is not possible through affixing in this language. Instead, the two different readings have distinct morphological manifestations, as indicated in (44):

(44)	Persian			
	ki	who	kasi	someone
	koja	where	jayi	somewhere
	key	when	vaqti	sometime
	ce no?i	what sort of	no?i	some sort of

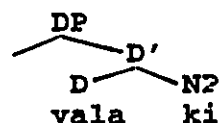
Thus, the morphological characteristics found in all multiple fronting languages can not be attributed to Persian.

The morphological properties of Wh/indefinite in [+MFS] and [-MFS] languages led Cheng to put forward an interesting argument concerning the obligatory feature of fronting Wh-elements. She explores the internal structure of Wh-words and suggests that the multiple fronting nature of Wh-words in these languages is a result of licensing requirement which is independent of typical Wh-movement of these elements. Hence, the Wh-phrases front for reasons of licensing and not for the purpose of clause typing. Let us have a closer look at this argument:

She extends Nishigauchi's (1990) analysis of Japanese Wh-words -which itself is based on Heim's (1982) theory of indefinite NPs- to the multiple fronting languages in order to account for the Wh-indefinite alternation.⁵ She argues that since the interrogative reading of a Wh-phrase can alternate with an indefinite reading of a Wh-phrase if a certain affix is present, the null hypothesis is that the Wh-words do not have inherent quantificational force. The affixes, therefore can be treated as determiners. The Hungarian

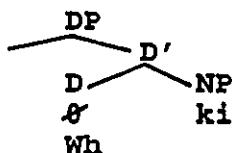
"valaki", for instance, which is made up of vala the prefix, and ki "who", is thus a DP. "Ki" the core, is the part without any inherent quantificational force, and "vala" is the part contributing the existential quantificational force which binds the core, as represented in (45):

(45)



With respect to what contributes interrogative force to a bare Wh-word, she suggests that when the core is interpreted as an interrogative, a null determiner is occupying the D position which has interrogative force: [D 0 [+Wh]].⁶ Consider (46):

(46)



On this basis, the multiple fronting of Wh-words in these languages can be attributed to a licensing requirement. That is the [D 0 [+Wh]] must be licensed by a C which is marked [+Wh]. Since multiple fronting languages do not have a C yes-no particle, and thus no Wh-particle, one Wh-word needs to move to SPEC of C to type the clause as a Wh-question. The [D 0 [+Wh]] in SPEC of C is licensed by the C by being in a SPEC-head agreement relationship. Besides, other Wh-phrases in the sentence also undergo fronting, because the [D 0 [+Wh]] in the Wh-phrases has to be in a local relationship with the [+Wh] C in order to be licensed. Hence, obligatory fronting of all Wh-words in multiple fronting languages is not triggered by clausal typing hypothesis. Movement of one and only one Wh-word to the SPEC CP suffices to meet the clausal typing requirement. So, these languages do not serve as a counterexample to the requirement in question, as was put forth in chapter three. The conclusion to draw from this argument is that multiple fronting of Wh-words is due to the inherent properties of Wh-words. The Wh-elements in these languages are morphologically complex, and need

to satisfy a licensing requirement independent of clause typing.

Turning back to Persian, we have seen that Wh-words do not show the same morphological characteristics found in other multiple fronting languages. Consequently, the clause typing may not be expected to be explained under the same analysis proposed for languages such as Bulgarian and Polish. Though all Wh-phrases in Persian may undergo fronting, this fronting stems neither from licensing requirement nor from clausal typing. Clausal typing is satisfied through another strategy which is specific of the wh-in-situ languages, including Persian. It was argued before that Persian has an overt yes-no particle and a non-overt Wh-particle at the sentential level which contribute interrogative force. Thus, movement to SPEC of C is not needed to type the clause as a Wh-question. We have also discussed how the availability of question particles correlates with the lack of syntactic Wh-movement. Languages either employ question particles or syntactic Wh-movement to type a clause as a Wh-question. According to the Principle of Economy of Derivation, no language has the option of alternating between the two methods of clausal typing. This in turn gives rise to the assumption that there are no languages with optional movement of Wh-constituents.

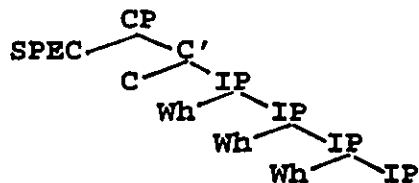
Given this fact, we can maintain that the similarities found in the multiple Wh-fronting structure of Persian and the relevant languages does not go further the superficial aspect, and therefore is not supported by any internal evidence. We can easily see that it is insufficient to take only the surface presentation into account. Hence, the structure of multiple Wh-fronting in [+MFS] and [-MFS] languages (cf.37), as proposed by Rudin, is definitely awkward for Persian. The facts observed in Persian can not be accommodated with her analysis for those languages.

Multiple Wh-fronting structure in Persian

Following the reasoning developed in this part and applied to Persian multiple Wh-fronting construction, I suggest a structure

for this language in which all of the Wh-phrases are adjoined to IP, producing a tree like (47):

(47)



As shown by (47), the clause-initial landing sites for multiple Wh-phrases in Persian turns out to be substantially distinct from what is claimed for [+MFS] and [-MFS] languages in that no Wh-element is allowed to move to SPEC CP. Thus, the fronting of Wh-words in multiple questions should be analyzed on a par with the movement of a Wh-word in single Wh-questions. That means multiple fronting of Wh-phrases results from topicalization as well. The argument in this chapter, therefore, reinforces our view concerning the topicalized nature of the Wh-movement in Persian, as proposed earlier.

The other crucial piece of evidence which confirms our structure in (47) can be given by clauses containing a complementizer together with fronted Wh-words. The following example illustrate that the complementizer *ke* can co-occur with and precede Wh-words:

- (48) mo?allem fekr mi-kard ke či-yo az ki bepors-e.
 teacher thought CM-did that what-DO from who ask-3S
 "the teacher was thinking what to ask whom."

No other combination of Wh-phrases and the complementizer would be possible in Persian. Therefore, the combination of the type "Wh+ke+Wh" may not be found, as indicated in the pair of structures below :

- (49) a. [_{SPEC CP}] [_{COMP COMP}] [_{IP Wh}] [_{IP Wh}] [_{IP...}]
 b. * [_{SPEC CP Wh}] [_{COMP COMP}] [_{IP Wh}] [_{IP...}]

The generalization emerging from the representations above and from the lack of syntactic Wh-movement in general, is that the Doubly Filled COMP Filter, as stated in (50) (adopted from Lasnik & Saito

(1984)) is perfectly respected in Persian:

(50=164) *[comp XY]

That means no COMP in Persian c l ever contain several Wh-elements or a complementizer and a Wh-phrase at S-structure. The absence of syntactic Wh-movement blocks the possibility of violating this filter.

The particular properties of [+MFS] and [-MFS] languages, according to Rudin, is attributed to a difference in the setting of a single parameter, given in (51):

(51=70) **The Multiply-Filled SPEC CP Hypothesis**
 SPEC CP can contain multiple Wh-words at S-structure in Bulgarian and Romanian ([+MFS]) languages, but only one in Serbo-Croatian, Polish and Czech ([-MFS]) languages.

Clearly, this parameter is not consistent with the facts in Persian. So, due to the characteristics of multiple Wh-fronting in this language, I propose a modified form of (51), stated in (52):

(52) SPEC CP contains no Wh-word at S-structure in Persian

It should be noted that (52) holds for both multiple Wh-questions and single Wh-questions. Regarding the fact that the level at which Wh-movement applies varies with respect to different languages, Persian therefore resembles the Chinese-type languages in having only LF Wh-movement.

This concludes our discussion in this chapter. It was argued that despite the apparent similarities, the multiple Wh-fronting construction in Persian is systematically different from that of the [+MFS] and [-MFS] languages. We have seen that multiple Wh-questions share properties of single Wh-questions. Fronting of Whs results from topicalization. Hence, in addition to the two well-studied [+MFS] and [-MFS] languages, Persian-type structure, may also be regarded as another type of multiple wh-fronting structure.

NOTES

1- In Polish, any extraction from a finite clause is normally ungrammatical. So, in this example, a subjunctive embedded clause is used.

2- The position of clitics in Persian can not be taken as a further evidence to determine the constituent structure of the Wh-words. Meanwhile, the examples with clitics in [+MFS] languages are not strong arguments either, because as Rudin points out, clitics in both Bulgarian and Romanian are proclitic to the verb. Pronouns in Persian may be represented as clitics by postposing to the different syntactic categories of V, P and N:

- (1) a. diruz did-am-eš.
yesterday saw-1S-him/her
I saw him/her yesterday."
- b. amir ketab-am-o qarz gereft.
Amir book-my-DO borrow took
"Amir borrowed my book."
- c. amir be-m hedye-yi dad.
amir to-me gift-a gave
"Amir gave me a gift."

The appearance of a clitic following a Wh-phrase is acceptable in a single Wh-clause, as in (2):

- (2) a. či-š-o šekast-i ?
what-his/her-DO broke-2S
- b. koja-š dard mi-kard ?
where-his/her pain CM-did

However, it gives rise to an ungrammatical sentence when placed between the first Wh-word and the second in a multiple wh-question:

- (3) *či-š-o be ki dad-i ?
what-his/her-DO to who gave-2S

3- This line of argument can not be extended to all [-MFS] languages, as no other analysis on intonation in the other multiple fronting languages has been made yet.

4- With respect to [-MFS] languages Cheng, unlike Rudin, argues that most Polish speakers seem to have strict ordering. If so, then the strict ordering among fronted Wh-words in this group of languages is also expected to fall under an Economy of Derivation account. Nonetheless, given the structure in (37a,b), there is a major difference between the two groups: all the fronted Wh-words in [+MFS] languages compete for the same position (SPEC CP), while in [-MFS] languages, one Wh-word appear in SPEC of C and the others adjoin to IP. If only movement to the same position is constrained by the Principle of Economy of Derivation, then Wh-word movement in the [-MFS] languages will not be constrained by this principle.

5- Nishigauchi's (1990) argument is based on the assumption that Japanese Wh-words are like indefinite NPs in that they do not have inherent quantificational force, and thus always need a binder.

6- Note that languages like Polish, Hungarian, etc. do not have Wh-particles. So, the core can not receive interrogative force from a Wh-particle, as in the case of languages such as Japanese.

Bibliography

- Allwood, J., 1982, "The Complex NP Constraint in Swedish", in E. Engdahl and E. Ejerhed (eds.), 15-32.
- Andersson, L. G., 1982, "What is Swedish an Exception to? Extractions and island Constraints", in E. Engdahl and E. Ejerhed (eds.), 33-45.
- Anward, J., 1982, "Basic Swedish", in E. Engdahl and E. Ejerhed (eds.), 47-75.
- Aoun, J., Hornstein, N. and Sportiche, D., 1981, "Some Aspects of Wide Scope Quantification", *Journal of Linguistic research* 1.3. 69-95.
- Authier, J. M., 1992, "Iterated CPs and Embedded Topicalization", *Linguistic Inquiry*, 23.2, 329-336.
- Bhatt, R., and Yoon, J., 1991, "On the Composition of COMP and Parameters of V-2", *WCCFL...* 42-52.
- Cheng, L. L-S., 1991, "On the Typology of Wh-Question", Ph.D Dissertation, MIT Press.
- Chisholm Jr, W.S. (ed.), 1984, "Typological Studies in Language: Interrogativity", John Benjamins Publishing Company Amsterdam/Philadelphia.
- Chomsky, N., 1977, "On Wh-Movement", in P.W. Culicover, T. Wasow and A. Akmajian (eds.), *Formal Syntax*, New York: Academic Press.
- Chomsky, N., 1986a, "Barriers", MIT Press, Cambridge, Massachusetts.
- Christensen, K. K., 1982, "On Multiple Filler-Gap Constructions in Norwegian", in E. Engdahl and E. Ejerhed (eds.), 77-98.
- Chung, S., and McCloskey, J., 1983, "On the Interpretation of Certain Island Facts in GPSG", *Linguistic Inquiry*, 14, 704-713.
- Comorovski, I., 1986, "Multiple Wh-Movement in Romanian", *Linguistic Inquiry*, 17, 171-177.
- Comrie, B., 1987, "The World's Major Languages", New York, Oxford University Press.

- Ejerhed, E., 1982, "The Processing of Unbounded Dependencies in Swedish", in E. Engdahl and E. Ejerhed (eds.), 99-151.
- Engdahl, E., 1982, "Restrictions on Unbounded Dependencies in Swedish", in E. Engdahl and E. Ejerhed (eds.), 151-174.
- Engdahl, E., and Ejerhed, E. (eds.), 1982, "Readings on Unbounded Dependencies in Scandinavian Languages", University of Umea.
- Engdahl, E., 1986, "Constituent Questions, the Syntax and Semantics of Questions with Special Reference to Swedish", D. Reidel Publishing Company, Dordrecht, Holland.
- Erteschik-Shir, N., 1982, "Extractability in Danish and the Pragmatic Principle of Dominance", in E. Engdahl and E. Ejerhed (eds.), 175-192.
- Grimshaw, J., 1986, "Subjacency and the S/S' Parameter", *Linguistic Inquiry*, 17, 364-369.
- Hajati, A., 1977, "Ke-Constructions in Persian: Descriptive and Theoretical Aspects", Ph.D Dissertation, University of Illinois at Urbana.
- Hashemipour, M., 1989, "Pronominalization and Control in Modern Persian", Ph.D Dissertation, University of California, San Diego.
- Hong, Y.-Y., 1989, "The Empty Category Principle: Antecedent Government", Ph.D Dissertation, Ewha Womans University, Seoul, Korea.
- Huang, C.-T. J., 1981/1982, "Move Wh in a language without Wh Movement", *The Linguistic Review* 1, 369-416.
- Jones, M. A., 1983, "Getting 'Tough' with Wh-Movement", *Journal of Linguistics*, 19, 129-159.
- Karimi, S., 1989, "Aspects of Persian Syntax, Specificity and the Theory of Grammar", Ph.D Dissertation, University of Washington.
- Kuno, S., and Robinson, J., 1972, "Multiple Wh-Questions", *Linguistic Inquiry* 3, 463-487.
- Lahiri, U., 1991, "Embedded Interrogatives and Predicates that Embed Them", Ph.D Dissertation, MIT.
- Lasnik, H., and Saito, M., 1984, "On the Nature of Proper Government", *Linguistic Inquiry* 15.2.
- Law, P. S., 1990, "Effects of Head Movement on Theories of

- Subjacency and Proper Government", Ph.D Dissertation, MIT.
- Lie, S., 1982, "Discontinuous Questions and Subjacency in Norwegian", in E. Engdahl and E. Ejerhed (eds.), 193-203.
- Lowenstamm, J., 1977, "Relative Clauses in Yiddish: A Case for Movement", *Linguistic Analysis*, 3, 3.
- Mahajan, A. K., 1990, "The A/A-Bar Distinction and Movement Theory", Ph.D Dissertation, MIT.
- McDaniel, D., 1989, "Partial and Multiple Wh-Movement", *NLLT* 7, 4, 565-605.
- Rivero, M. L., 1977, "On Topicalization and Wh-Movement in Spanish", *Linguistic Inquiry* 9, 513-517.
- Rivero, M. L., 1980, "On Left-Dislocation and Topicalization in Spanish", *Linguistic Inquiry* 11, 363-395.
- Rizzi, L., 1978, "Violations of the Wh Island Constraint and the Subjacency Condition", Montreal working papers in Linguistics 11. Reprinted in L. Rizzi, 1982, "Issues in Italian Syntax", Foris, Dordrecht.
- Rizzi, L., 1990, "Relativized Minimality", MIT Press, Cambridge, Massachusetts.
- Rizzi, L., 1990b, "Speculation on Verb Second", ms.
- Rochemont, M., 1989, "Topic Islands and the Subjacency Parameter", *Canadian Journal of Linguistics*, 34, 145-170.
- Rudin, C., 1988, "On Multiple Questions and Multiple Wh-Fronting", *NLLT* 6.4, 445-502.
- Samian, V., 1983, "Structure of Phrasal Categories in Persian: An X-bar Analysis", Ph.D Dissertation, University of California, Los Angeles.
- Sells, P., 1984, "Syntax and Semantics of Resumptive Pronouns", Ph.D Dissertation, University of Massachusetts at Amherst.
- Soheili-Isfahani, A., 1976, "Noun Phrase Complementation in Persian", Ph.D Dissertation, University of Illinois at Urbana-Champaign.
- Sportiche, D., 1981, "Bounding Nodes in French", *the Linguistic Review*, 1, 219-264.
- Srivastav, V., 1991, "Wh Dependencies in Hindi and the Theory of Grammar", Ph.D Dissertation, Cornell University, Ithaca, New

York.

- Srivastav, V., 1992, "Subjacency Effects at LF: The Case of Hindi Wh", *Linguistic Inquiry*, 22.4, 762-771.
- Taraldsen, K. T., 1982, "Extraction from Relative Clauses in Norwegian", in E. Engdahl and E. Ejerhed (eds.), pp
- Van Riemsdijk, H., and Williams, E., 1986, "Introduction to the Theory of Grammar", MIT Press.
- Wachowicz, K. A., 1974, "Against the Universality of a Single Wh-Question Movement", *Foundations of Language*, 11, 155-166.
- Windfuhr, G. L., 1978, "Persian Grammar: History and state of its study", The Hague: Mouton Publishers.
- Zaenen, A., and Maling, J., 1982, "The Status of Resumptive Pronouns in Swedish", in E. Engdahl and E. Ejerhed (eds.), 223-230.