Obligatory Control in Persian

Ivanna F. S. Richardson

A Thesis

in

the Individualized Program

Presented in Partial Fulfillment of the Requirements for

the Degree of Master of Arts (Classics, Modern Languages and Linguistics) at

Concordia University

Montreal, Quebec, Canada

May 5, 2016

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: **Ivanna Richardson**

Entitled: **Obligatory Control in Persian**

and submitted in partial fulfillment of the requirements for the degree of

Master of Arts

complies with the regulations of this University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

_____Chair
Dr. Ketra Schmitt

_____Examiner
Dr. Mark Hale

_____Examiner
Dr. Madelyn Kissock

_____Thesis Supervisor
Dr. Daniela Isac

Approved_____
Graduate Program Director

_____ 20___ _____
Dr. Paula Wood-Adams
Dean of Graduate Studies

Abstract

Obligatory Control in Persian

Ivanna Richardson

The object of study of this thesis is control structures in Persian, more specifically the properties of the empty subject of embedded clauses occurring in Persian control structures. The null subject in these Persian contexts may be one of three different empty categories: pro, NP-trace/copy, or PRO. Obligatory Control contexts are defined by the following properties: they are made up of complex clauses contain a matrix clause and an embedded clause, an argument of the matrix clause must be coindexed with the subject of the embedded clause, and the subject of the embedded clause may be covert or overt, but must crucially be coindexed by the matrix subject. Current theories accounting for Obligatory Control contexts fall into two categories: those that seek to reduce the grammatical mechanisms necessary to account for Obligatory Control contexts and those that seek to account for these contexts by agreement without reducing the grammatical mechanisms.

The Persian data is examined to determine whether the empty subject has the properties of pro, NP-trace/copy, or PRO. The properties of Obligatory Control and the three empty categories are introduced and arguments are outlined for positing a structural subject in Obligatory Control contexts. The current theories accounting for Obligatory Control contexts are described and each is evaluated for its ability to account for the Persian Data. The analysis shows that the Persian Obligatory Control data is best accounted for by positing PRO as the empty subject in those contexts, licensed by agreement.

Acknowledgements

No work of this sort is produced without many, many eyes reading through it first. The effort and dedication that one puts into producing a thesis from the first tiny musings to the final product is dwarfed by the effort, dedication, and support that one receives from supervisors, colleges, friends, and family. The hours and hours of reading and writing are made possible by just as many hours bouncing ideas back and forth, seeking guidance, talking, and laughing. This collaboration is what makes this process possible.

I must first acknowledge how inspiring and supportive the Concordia Linguistics program is. I thank the faculty and staff for many years of learning. I am sad to be leaving after so many years, but so grateful for the time spent with you all.

Second, I must thank each of my language consultants: Mojgan, Adel, and Parissa. Without your help, I would not have had a puzzle to scratch my head over for two years. Thank you for answering my maddening questions with patience and good cheer. Your input made this project possible. Thank you for sharing your beautiful and intricate language with me.

Third, thank you to my committee for guiding me through this process. Mark and Madelyn, you have been so generous with your time and advice. There is nothing I enjoy more than sitting with you and listening to you puzzle through a problem. Watching your minds at work is the best classroom experience I ever had. I must also say thank you to Charles for egging us all on and encouraging us in the must round about and maddening ways possible. We all appreciate your gruff chastisement and advice very much. Dana, how can I thank you enough? Thank you for the hours of study, and for scrutinizing each week's readings. Thank you for your many months of reading my very poor written productions, and for the final push to the finish line, helping me come to the end of this project. Your teaching style—exacting, analytical, and yet so warm and encouraging—will always influence me (I hope someday I can terrorize my students just as kindly as you do).

Finally, thank you to all my friends and family. Sigwan and Matthew, my success is down to your support and camaraderie over these last two years. I will miss our little office with all the laughter, tears, hair-pulling frustration, joyous triumphs, and heartfelt kindness. No one knows it like we do. Cheers to all the new adventures we look forward to in the years to come! To my mother and sister, thank you for your encouragement and patience, especially while listening to me ramble on about Linguistics. T, thank you for having my back and believing in me and my ability abilities, even when I doubted myself. You help me realize my potential.

Contents

1 Introduction **1**

1.1 Object of Study 1

1.2 Goal 1

2 Background **5**

2.1 Features 5

2.1.1 [+/– anaphor] 7

2.1.2 [+/– pronominal] 7

2.2 Features and Types of Empty Categories 8

2.2.1 *pro* 9

2.2.2 WH trace 10

2.2.3 NP trace 11

2.2.4 PRO 11

3 Why posit a subject in OC contexts? **15**

3.1 OC Clauses are CP/TP projections: Arguments in Favour 15

3.1.1 Embedded Infinitive Clause + Complementizer 16

3.1.2 Infinitive Clauses + Conjunction 16

3.1.3 Finite Clauses 17

3.2 Direct Arguments for the EC subject 17

3.2.1 Secondary Predication 17

3.2.2 Floating Quantifiers 19

3.2.3 Partial Control 19

3.2.4 Agreement/Split Control 20

3.2.5 Case-Concord 21

3.2.6 Binding . . . 22
3.2.7 Overt Controllees . . . 23
3.2.8 Expletive Constructions . . . 24

4 Previous Approaches **25**
4.1 Reductionist approaches . . . 26
4.1.1 Hornstein and Boeckx . . . 26
4.1.2 Manzini and Roussou . . . 28
4.1.3 Wurmbrand . . . 28
4.2 Non-reductionist Approaches . . . 29
4.2.1 Agree . . . 29

5 Persian **33**
5.1 General Properties . . . 34
5.1.1 Word Order . . . 34
5.1.2 *little pro* . . . 34
5.1.3 Infinitive/subjunctive . . . 35
5.2 Control Verbs . . . 35
5.2.1 Genuine versus Bogus OC in Persian . . . 36

6 Anaylsis **39**
6.1 Reductionist Approaches to the EC subject . . . 39
6.1.1 Wurmbrand and Persian OC . . . 39
6.1.2 Manzini and Roussou and Perisan OC . . . 42
6.2 Are Persian OC Constructions Movement or Agreement? . . . 43
6.2.1 OC in Persian is not *pro* . . . 44
6.3 Can Movement Account for Persian OC? . . . 46
6.4 An Account of Persian OC based on Agree . . . 49

7 Conclusion **53**

Chapter 1

Introduction

1.1 Object of Study

The object of study of this thesis is control structures in Persian, more specifically the properties of the empty subject of embedded clauses occurring in Persian control structures.

(1) Məzen mi-tun-e [(ke) EC ʃatraŋ bosi be-kun-e].
Məzen DUR-able-3SG [(COMP) EC chess play SUB-do-3SG].
'Məzen can/is able to play chess.'

The empty subject in the embedded clause is marked as EC in example (1). This subject is interpreted as having the same reference as the subject of the matrix clause, *Məzen*. The goal of the thesis is to determine the nature of this empty subject and properties of the clauses in which it is found.

1.2 Goal

The more general aim of this thesis is to explore the properties of empty categories that occur in human languages. Empty categories (hereinafter EC) play an important role in syntactic theory, from the construction of interrogative clauses and structures with an NP that moves, see example (2) and example (3) respectively; to constructions involving a null subject of finite clauses or a null subject of embedded clauses as in examples (4) and (5):

(2) a. What did the dog eat [ ]?

b. The dog ate what?

(3) a. John seems [[ ] to be happy].

b. It seems [John is happy].

(4) a. Habla Espanol
speaks.3sg Spanish
'He/she speaks Spanish.'

b. Ella habla Espanol
she speaks.3sg Spanish
'She speaks Spanish.'

(5) a. John tried to run to the store.

b. John tried [[ ] to run to the store].

In example (2) the interrogative WH-constituent overtly occurs in the CP domain of the clause, however it is linked to the object position, as marked by the empty brackets. Similarly, in example (3), the NP *John* is linked to the subject position of the embedded clause, again marked by brackets in the lower clause. Unlike the EC's in examples (2) and (3), the EC subject in (4) is not linked to another position, it is the covert realization of the pronominal subject. The EC in example (4a) has all the same properties as an overt pronominal subject, save that it goes unpronounced: this can be seen by comparing (4a), with a null subject, and (4b), which has an overt pronominal subject. The null subject in example (4a) and the overt subject in example (4b) both license agreement on the verb, have a theta role, and the same phi features. Example (5) exhibits a third type of EC. This EC can be found in embedded clauses, as in example (5), or as the subject of an independent phrase with an arbitrary reading, as in example (6).

(6) [ ] to dance is magnificent.

The challenge EC's pose for linguistic theory generally, as well as for syntactic theory specifically, is one of acquisition. If a category, or a phrase is phonetically null and a child will never come across auditory stimuli related to the respective EC, then how can acquisition of EC's occur? Chomsky (1981) observes in '*Lectures on Government and Binding*' that these ECs can provide insight into the acquisition process itself.

> "These properties can hardly be determined inductively from observed overt phenomena, and therefore presumably reflect inner resources of the mind. If our goal is to discover the nature of the human language faculty, abstracting from the effects of experience, then these elements offer particularly valuable insights."

Generative linguistic theory posits that human linguistic ability is innate and that we are endowed with the building blocks needed to acquire any human language. These building blocks make up what linguists call the Universal Grammar (UG). In order to explain why it is possible for the acquisition of ECs to take place,

these categories must be posited as real cognitive entities available in UG. The examples above illustrate that although material may be 'missing' from the physical utterance, from the wave form that our articulators produce as speakers of English or Spanish, we can still identify the meaning of the phonetically null material. The observation that Chomsky makes above is that this interpretational fact, that we can still perceive the meaning and existence of material that is not physically present in the input we are given, is crucial evidence of the underlying structural complexity of sentences.

This thesis will focus on a particular type of empty category, the type found in example (5), the null subject of non-finite embedded clauses. In particular, I will offer a detailed description of the properties of this EC, a characterization of the types of clauses that license its presence, and a brief overview of the verbs that select these clauses. I will evaluate the different theoretical analyses of these clauses and evaluate the same types of clauses in Persian to determine which approach best accounts for those patterns.

Chapter 2

Background

2.1 Features

ECs became the locus of attention for generative linguistic research starting in the early 80's. In the *Government and Binding* (henceforth GB (Chomsky, 1981)) framework the distinction between various types of ECs was characterized in terms of two features: [+/– pronominal] and [+/– anaphor]. These features also captured the distribution of overt pronouns and NPs.

Although these features are not currently assumed to be morpho-syntactic features that are active in the syntactic computation, and are no longer posited as motivating features in the grammar, they still capture the descriptive generalization of the distribution of each category. These features in GB relied on a structurally defined domain called, 'the governing category'. This domain delineated the crucial 'space' within which a syntactic element could or could not have a co-referrential relationship with another syntactic element. For example, an anaphor needs an antecedent within its governing category. The theoretical validity of government has been disposed of, however, the domain requirement for having an antecedent or not remains a part of syntactic computational theory.

The properties [+/– anaphor] and [+/– pronominal] together define the distribution of NPs. The interaction of these features are best understood in relation to the notions of **antecedent and c-command**.

An **antecedent** of a NP is a syntactic item that occurs in the same minimal domain as the NP itself, that is coindexed with that NP, and gives that NP its semantic reference. In example (7), *Mary* is the antecedent of *herself*. The two are semantically linked: the NP *Mary* and the reflexive pronoun *herself* share the same referential index.

(7) Mary$_i$ hugged herself$_i$.

The linear order of constituents is not what determines antecedent-hood: it is not the case that any constituent that linearly precedes an NP may be that NP's antecedent, as seen in example (8).

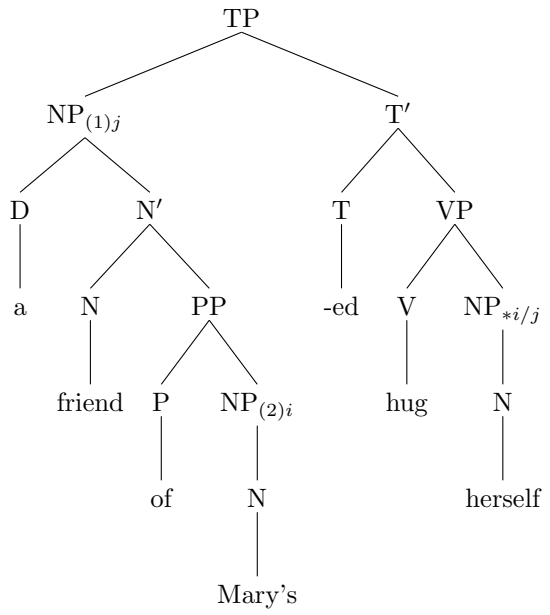
(8) [[A friend$_i$ [of Mary$_{*j}$]] hugged herself$_{*j}$]

In example (8) both *Mary* and *a friend* precede *herself* in the linear order, but only the full constituent *a friend of Mary's* is a possible antecedent for *herself*. Linear order is not enough to determine which NP is the suitable antecedent for *herself*. The choice of antecedent is determined by a structural notion: c-command, defined in (9).

(9) α c-commands β iff α does not contain β and every category that contains α also contains β.

In other words, a constituent α c-commands β when α and β are both dominated by the same phrasal projections and do not contain each other. Looking at the tree in the example (10), *Mary* cannot be the antecedent for *herself*, because *Mary* does not c-command *herself*.

(10) [[A friend$_j$ [of Mary$_{*i}$]] hugged herself$_{*i/j}$]



The NP$_2$ *Mary* is dominated by the NP$_1$ constituent *'a friend of Mary's'* and the NP$_1$ does not dominate the reflexive *herself*. Therefore *Mary* and *herself* are not dominated by the same maximal projections and *Mary* does not c-command *herself*. The NP$_1$ constituent on the other hand, does c-command the reflexive. NP$_1$ is dominated by TP, and the TP node also dominates the reflexive *herself*. The NP$_1$ *'a friend of Mary's'* and *herself* are both dominated by the same projections and neither dominates the other. Therefore NP$_1$ c-commands *herself*. The NP$_1$ is in the correct structural relationship to be the antecedent for *herself*.

2.1.1 [+/− anaphor]

The property [+/− anaphor] defines whether or not an NP must necessarily have a c-commanding antecedent within a minimal structural domain, the NP's governing category. More specifically, the two features are defined below:

- if a constituent A is [+ anaphor], it must have a c-commanding antecedent in its Governing Category.
- if a constituent A is [− anaphor], it cannot have a c-commanding antecedent in its Governing Category (i.e. it must be free in its Governing Category).

These definitions are demonstrated in example (11).

(11) a. Mary$_i$ saw herself$_i$ in the mirror.

b. *Herself saw Mary.

c. Mary$_i$ wants Paul to hug herself$_{*i}$.

In the examples in (11) *herself* is valued as [+ anaphor], which means that *herself* must have a c-commanding antecedent in its governing category. In example (11a) the governing category for *herself* is the full clause *Mary saw herself in the mirror*, and in this clause the NP *Mary* is the c-commanding antecedent for the NP *herself*. This example satisfies the features on *herself* and is grammatical. In contrast, in example (11b), *herself* has no antecedent to c-command it, making this string ungrammatical. In example (11c) the governing category for *herself* is the embedded clause. The example (11c) is ungrammatical, because the matrix subject *Mary* cannot be the c-commanding antecedent for *herself*: *Mary* is not contained within the governing category for *herself*. Therefore *herself* is missing a c-commanding antecedent in the embedded clause, violating the requirements of its feature [+ anaphor].

2.1.2 [+/− pronominal]

The property [+ pronominal] defines when a NP must be free in its governing category and must not have an antecedent within that governing category. The property [− pronominal] defines when a NP must not have an antecedent outside of its governing category. More specifically, the two features are defined below:

- if a constituent A is [+ pronominal], it cannot have a c-commanding antecedent in its Governing Category.
- if a constituent A is [− pronominal], it cannot have a c-commanding antecedent outside of its Governing Category.

The outcome of these features is demonstrated below in examples (12) and 13).

(12) *Mary$_i$ saw her$_i$ in the mirror.

(13) Fred saw Mary$_i$ in the mirror and thought she$_i$ was pretty.

In example (12) the governing category for *her* is the full string *'Mary saw her in the mirror'*. The NP *Mary* is a c-commanding antecedent for *her* and is within the pronoun's governing category. Since [+ pronominal] constituents cannot have a c-commanding antecedent in their governing category, the result is ungrammatical. However in (13) the governing category for *she* is the embedded clause *'and thought she was pretty'*. In this case, *Mary* is a grammatical c-commanding antecedent for *she*, because *Mary* is not part of the governing category for *she*. *Mary* is the subject of the matrix clause and the pronoun is the subject of the embedded clause.

2.2 Features and Types of Empty Categories

The features introduced in Section 2.1 can be used to characterize the various types of empty categories illustrated in Section 1.2. The identity of each EC is described by a specific combination of the [+/– anaphor] and [+/– pronominal] features. I will use the labels provided by the literature for each of the possible featural combinations. Note that with respect to the EC valued [+ anaphor], [+ pronominal] it is debatable whether the label PRO is accurate or not. Thus far I have just referred to it as the EC subject of non-finite clauses in example (5) in Section 1.2. I will come back to this debate in Chapter 4.

Features	Non-Overt
[+ anaphor] [– pronominal]	NP-trace
[– anaphor] [+ pronominal]	*pro*
[– anaphor] [– pronominal]	WH-trace
[+ anaphor] [+ pronominal]	PRO

Table 2.1: Empty Categories

In this thesis the status of the Persian EC subject in embedded non-finite clauses will be shown to be a choice between *pro*, PRO, or a NP-copy. Carefully defining the properties of each EC gives us precise criteria with which to assess the EC subject in Persian.

The following is a description of each type of EC and the interaction of the features describing the ECs distribution.

2.2.1 *pro*

The EC *pro* is the null subject of finite declarative clauses in languages like Spanish, as in example (4) repeated below as example (14). The example (14) explicitly shows *pro* as the subject of the phrase.

(14) *pro* habla Espanol
pro speaks.3sg Spanish
'He/she speaks Spanish.'

As indicated in Table 2.1, *pro* is [– anaphor] and [+ pronominal]. *pro* must be free within its governing category, but it may have an antecedent outside of its governing category. In (14) the governing category for *pro* is the full string, and *pro* does not have any c-commanding antecedents in this clause. In addition to the features [– anaphor] and [+ pronominal], *pro* has the following properties:

- *pro* is not universal across languages.

  (15) a. *pro* habla Espanol
  pro speaks.3sg Spanish
  'He/she speaks Spanish.'

  b. She speaks English.

  c. *speaks English.

  The English examples in (15b) and (15c) show that English may not drop the subject of declarative clauses. The English lexicon does not contain this EC *pro*, but example (15a) shows that the Spanish lexicon does have the EC *pro*.

- It is in contrastive distribution with overt NPs.

  (16) a. *pro* habla Espanol
  pro speaks.3sg Spanish
  'He/she speaks Spanish.'

  b. Ella habla Espanol
  She speaks.3sg Spanish
  'She speaks Spanish.'

  In example (16a) the subject is the covert *pro* and is interpreted as pronominal: *he/she speaks Spanish.* In example (16b) the pronominal subject is overt and the interpretation is: *'She speaks Spanish.'*

- The overt realization of *pro* is pronominal, meaning it has the phi features (person, gender, number), as in example (16b). In example (16a), *pro* has the following phi features: third person, singular, and depending on context, its gender is masculine or feminine.

- From a semantic point of view, the reference of *pro* is recoverable from the context. For example the overt morphological agreement on the verb in Spanish above signals the number of the covert pronominal *pro*. However, some languages that *pro-drop*, such as Mandarin, have no overt morphological agreement on the verb; in this case, the semantic reference is recoverable from the discourse.

2.2.2 WH trace

Syntactic traces are 'placeholders' for items that moved to higher positions in the syntactic tree to check or fulfill featural specifications. In modern minimalist theory these traces are assumed to be copies of the constituent that has been moved. There are a number of theoretical implications that follow from using copies rather than traces, but for the purposes of this discussion I will assume that the two are equivalent, mainly because both are the result of the syntactic operation MOVE.

The first type of copy introduced in Chapter 1, Section 1.2 is the WH-copy, as seen in example (2) repeated below as example (17) with the copy of the interrogative struck out in the object position of the VP.

(17) What did the dog eat [~~what~~]?

In (17) the interrogative NP *what* moves from the object position to the specifier of CP. The result of this movement is the creation of another copy of *what* in the target position. In example (18) (seen in the paragraph below) *who* is generated in the object position, and like *what* in example (17), moves to the specifier of the CP creating a new copy of itself in that position.

These WH copies are valued [– anaphor] and [– pronominal]: they cannot have a c-commanding antecedent within their Governing Category and they cannot have a c-commanding antecedent outside of their Governing Category, either. In other words, WH-copies must be free everywhere. Notice that even though the WH-copy in (18) is coindexed with the higher copy, the higher copy does not count as a c-commanding antecedent for the lower WH-copy. This is because WH constituents are non-referential, and they occupy A bar positions. To show that WH-copies cannot have a referential antecedent, compare the following:

(18) a. Who ~~who~~$_i$ thinks he$_i$ is clever?

b. Who does he$_i$ think ~~who~~$_{*i}$ is clever?

In (18a), *who* and *he* can be the same person, but in (18b) they cannot. In (18a), the WH-copy is not c-commanded by he and therefore *he* does not count as a c-commanding antecedent. The WH-copy is free. In contrast, in (18b), the lower WH-copy is c-commanded by *he* and hence *he* would be the c-commanding antecedent. Coindexing the Wh-copy and *he* results in ungrammaticality (represented by the * before the

index on the WH-copy), which supports the conclusion that the WH-copy cannot have a c-commanding antecedent anywhere.

2.2.3 NP trace

The second type of copy presented in Section 1.2 is the NP copy, as in example (3a) repeated below as (19) with the copy of *John* crossed out in the lower clause.

(19) John seems [~~John~~] to be happy.

In example (19) 'John' moved from the subject position of the embedded clause to the specifier of TP in the matrix clause. The NP copy of *John* is valued [+ anaphor] and [– pronominal], meaning that it must have a c-commanding antecedent within its Governing Category. The Governing Category for the lower copy of the NP that moved is the matrix clause and the c-commanding antecedent for the lower NP copy is the higher NP copy. Notice that the higher NP copy of *John* in the subject position of the matrix clause is in an A position and as such it counts as a c-commanding antecedent for its copy in the subject position of the lower clause.

2.2.4 PRO

The fourth EC is the null subject of non-finite embedded clauses as in example (5) repeated as example (20) below, explicitly marking the empty subject position EC.

(20) [John tries [EC to run]]

This EC is valued [+ anaphor] [+ pronominal], which means it must be both bound and free in its Governing Category. This contradiction was resolved in the GB model by proposing that PRO does not have a Governing Category. PRO is [+ anaphor] in the sense that it must have a c-commanding antecedent and it is [+ pronominal] in the sense that it may also have arbitrary reference.

This division was seen in Section 1.2 (for examples of each interpretation see (5a) and (6) respectively repeated below as (21) and (22)).

(21) [John tries [EC to run]].

(22) [EC to dance] is magnificent.

The structures of the first type are called Control structures and the second type are Arbitrary Control structures. The group, represented by example (21), may be subdivided further into those cases where the EC subject may alternate with an overt disjoint NP and those cases where it cannot.

(23) a. John tried [EC to run].

b. *John tried [Fred to run].

In example (23) the subject of the matrix clause, *John*, is the antecedent or controller of the EC subject. This is called Obligatory Control. In an OC context, if an overt subject is present, it must be coindexed with the controller. As will be discussed in Chapter 3, Section 3.2.7, OC contexts may be further divided into those cases where an overt subject is allowed to alternate with the EC subject, so long as it is semantically linked to the controller, and those cases where the EC subject may not alternate with an overt subject, as in example (23b).

In contrast, in example (24) the EC in the embedded clause may alternate with an overt disjoint NP. As shown in example (24a), the overt disjoint pronoun subject *me* is grammatical and, further, an overt disjoint NP is also grammatical, as in example (24c).

(24) a. Clearly, [EC confessing my crime] was not something they anticipated. (Landau, 2013)

b. Clearly, [me confessing my crime] was not something they anticipated.

c. Clearly, [John confessing his crime] was not something they anticipated.

Those matrix verbs that select an embedded clause with an EC subject that may alternate with an overt disjoint NP are called Non-Obligatory Control (NOC) verbs. The matrix verbs that select an embedded clause whose EC subject cannot alternate with a full NP are OC verbs. NOC environments cover a variety of clause types: subject clauses, extra-posed clauses, and adjunct clauses. The distribution of OC is restricted to complement clauses.[1] This thesis will only focus on Obligatory Control (OC).

OC is, itself, split into two groups, Non-Exhaustive Control, as in example (25), and Exhaustive Control, as in example (26).

(25) The organizer decided [EC to meet right before the parade].

(26) John tried [EC to run].

In (26) the matrix subject must exhaustively control the EC subject: the referent of the EC is co-extensive with the referent of the controller *John*. In contrast, in example (25) the matrix subject only partially controls the EC subject: the set of individuals that the EC denotes includes, but is not limited to, the matrix controller, *the organizer*.

This thesis will focus exclusively on OC structures, the properties of which are the following:

- OC contexts may not have a covert disjoint subject.

[1]There are some cases of OC in adjunct clauses, but only under very restricted circumstances.

- OC contexts may not have an overt disjoint subject.
- OC contexts may allow for overt subjects only if that subject is shared or controlled.

Thus far I have discussed the nature of the EC subject and the clauses that select for it in detail. I have not addressed the verbs that select for the clauses with this EC subject. Landau (2015) divides the cross-linguistic set of verbs that select for these clauses between those that select for [+] semantic tense clauses, as seen in example (27), and those that select for [–] semantic tense, as seen in example (28). The property [+/–] Semantic tense will be discussed in Chapter 4.[2]

(27) a. *Implicative*: avoid, bother, forget
b. *Aspectual*: begin, continue, finish
c. *Modal*: is able, may, must

(28) a. *Factive*: dislike, surprised
b. *Propositional*: believe, declare, think
c. *Desiderative*: decide, promise, want
d. *Interrogative*: ask, know, understand

These verbs and more make up the cross-linguistic set of verbs that select for Control clauses. I will evaluate a set of these verbs and their behaviour in Persian.

[2]The example verbs are excerpts from Landau (2015).

Chapter 3

Why posit a subject in OC contexts?

The following sections will discuss in detail the motivation for positing an empty subject in OC contexts. These arguments are inspired by Landau (2013), who proposed two different types of supporting arguments for the empty subject. The first type of argument focuses on the type of clause in which these empty subjects appear. This first type is an indirect argument for positing this empty subject and focuses on proving that the embedded clauses in OC contexts are full CP/TPs. The second type of argument directly argues for the existence of full subjects in OC contexts. I will briefly outline each type.

3.1 OC Clauses are CP/TP projections: Arguments in Favour

The first set of arguments focus on the status of the embedded clauses in OC contexts, rather than on the subject specifically. To argue that there is a subject position in these embedded clauses, the necessary structure must be available to house that subject. To construct these arguments, it is necessary to first define what a subject is, and then to define where that subject is located and how it behaves. In this case, I will first present arguments that the structure for a subject is present.

For this I assume the standard structural definition and derivation of a subject:

- A subject is generated in the specifier of vP.
- A subject moves to the specifier of TP to satisfy agreement features on T, such as case and phi features.
- An empty subject is not an empty (i.e. implicit) argument: an EC subject fulfills the same specifications as an overt subject.
- A subject satisfies agreement features on T, checks the theta role for the subject position in SpecTP, and checks case.

An argument for a CP/TP projection is an implicit argument for a structurally present subject, with the full specifications of an overt subject. It is also a standard assumption that if a TP is projected, then there must also be a CP projection, satisfying features on the T head.

Much of the literature on the clauses in which the EC subject appears, has focused on the status of infinitive clauses, such as in example (29) below.

(29) [John tries [EC to run]]

Infinitives and other so called 'tenseless' verbal forms (infinitives, gerunds etc) have been argued to be defective. This suggests that they project less than a CP/TP and therefore contain no projection higher than vP. Landau (2013) presents three arguments that these embedded clauses should in fact be treated as full clauses, i.e. as CPs.

3.1.1 Embedded Infinitive Clause + Complementizer

Infinitive clauses can be introduced by complementizers or by phrases that can be independently shown to occupy the SpecCP position:

(30) Mary asked [**which way** to go].

(31) John probeerde [**om** het boek te lezen]
John tried [COMP the book to read]
'John tries to read the book.' *Dutch (Landau, 2013)*

In example (30) from English, a WH phrase introduces the embedded clause. The WH phrase must occupy a position in the CP, therefore the infinitive VP must project as high as a CP in order to accommodate this WH phrase. The clause itself cannot be smaller than a full CP. This is also true of the example (31) from Dutch: the embedded clause is introduced by a complementizer, and again this points to the clause projecting a full CP.

3.1.2 Infinitive Clauses + Conjunction

Infinitive clauses can conjoin with full clauses:

(32) To write a novel and for the world to give it critical acclaim is John's dream.

(33) John expected to write a novel but that it would be a critical disaster. *(Landau, 2013)*

In examples (32) and (33) the infinitive clauses conjoin with full clauses. Although predicates are somewhat free to conjoin with different types of predicates (they can conjoin with things that do not match their

own identity), arguments usually respect type when conjoining. Full clauses will conjoin with full clauses. Examples (32) and (33) show infinitives conjoining with full clauses and therefore these infinitive clauses must also be full clauses, satisfying matching type conditions.

3.1.3 Finite Clauses

Further support for analyzing OC contexts as full CPs comes from languages whose non-finite embedded clauses are actually inflected for agreement. For example, there are infinitive constructions in Portuguese that are marked for person agreement, as seen in (34).

(34) Eu convenci os meninos$_i$ [EC$_i$ a tomar(em) banho].
I convinced-1sg the boys [EC Prep. to.take-3pl bath].
'I convinced the boys to take a bath.' *(Modesto, 2007)*

Landau (2013) suggests that these embedded structures, by the very virtue of the overt agreement in the clause, point to a full CP analysis. The overt agreement points to the presence of a TP with the relevant features that must be checked by a subject.

This last example is the most pertinent for this thesis: the structures under analysis here are from Persian, and in Persian the clauses that contain the EC subject are in the subjunctive mood. The subjunctive is an inflected verbal form, projecting a TP with person agreement features.

3.2 Direct Arguments for the EC subject

The evidence above focused on the status of the clauses that contain empty embedded subjects. Now I turn to the subject itself. Landau (2013) also presents cases of phenomena that support the analysis of the EC subject as a true null subject, independent of a full clausal analysis. The cases below support analyzing these structures as containing a full, structural EC subject. These cases can be split into two types: evidence based on interpretational facts, and evidence based on syntactic structural facts supporting a full subject analysis. The former arguments include facts of secondary predication, floating quantifiers, and partial control. The latter arguments include agreement or split control, case concord, binding, and expletive constructions. I will introduce these arguments in the order mentioned above.

3.2.1 Secondary Predication

Secondary predicates require overt NPs in order to license the predication, as in example (35), and they may not be licensed by implicit objects, as in example (35b), or by implicit agents of passives, as in example (36b).

(35) a. John ate (the meat).

b. John ate *(the meat) raw.

(36) a. He served dinner angry at the guests.

b. *Dinner was served angry at the guests.

Examples (35b) and (36a) show depictive relationships between a subject or object, and an adjective. In example (35a) the argument *the meat* is optional but in example (35b) *the meat* is obligatory. The secondary predicate *raw* cannot be licensed by an implicit object, *the meat* must be overt for the string to be grammatical. In example (36a) the secondary predicate *angry* refers to the overt agentive subject *he*. However, in example (36b) *angry* cannot normally be interpreted as a predicate of the raised subject *dinner*. Only animate individuals can be angry. The string in (36b) does contain an argument that denotes an animate individual, namely the agent, but this argument is implicit. Notice that (36b) is a passive construction, *dinner* is the overt subject, and the utterance has an implicit agentive subject. There was somebody who was serving the dinner, and that person could in principle be serving dinner in an angry way. Landau (2013) argues that the secondary predicate *angry* in (36b) cannot be a predicate of an implicit agent. The ungrammaticality of example (36b) is not caused by the syntactic structure of the phrase itself, but by interpretational facts.

Some analyses of the EC subject under consideration have argued that it is an implicit argument. As in examples (35) and (36), the EC subject should not be able to license secondary predicates if it is an implicit argument. However, in example (37) we see that the EC subject can in fact license secondary predicates.

(37) a. The meat was too chewy [EC to be eaten raw].

b. [EC to serve dinner angry at the guests] is bad manners.

Example (37) shows that secondary predication is accepted in control clauses: for example in (37a) *raw* is a predicate of the EC subject. This would be the end of this argument, except that in example (37a), it could be that the depictive in the embedded clause is licensed by the matrix subject *the meat*, without the necessity of forming a relationship with the EC subject in the embedded clause. The matrix subject already dominates the depictive in (37a), albeit through a clause boundary. Example (37b) supports the control analysis that the EC subject is licensing the depictive *angry*. In (37b) the subject is not dominated by any other structure, making it impossible for *angry* to predicate with anything other than the EC subject. The subject in (37b) is the EC subject with an arbitrary reading. It is necessary to show that the EC embedded subject can also license secondary predicates.

(38) a. *John pleaded [with Mary] cheerfully. (Mary is not cheerful)

b. John pleaded [with Mary$_i$] [EC$_i$ to arrive cheerful]

In example (38a) the secondary predicate *cheerfully* cannot penetrate the prepositional phrase *with Mary*. Contrastively, in (38b) *cheerful* appears to be able to penetrate the prepositional phrase. This difference can be explained if, in example (38b), the EC subject is present in the embedded clause and is able to form a depictive interpretational relationship with *cheerful*. With the presence of the EC subject in example (38b), *Mary* is interpreted as the person who is cheerful, because although the secondary predicate *cheerful* cannot penetrate the prepositional phrase, the EC subject can. The EC subject and *Mary* are coindexed and *cheerful* forms the depictive relationship with the EC subject.

3.2.2 Floating Quantifiers

Floating quantifiers (FQ) offer a similar pattern pointing to the existence of a full EC subject. FQs must have '...a syntactically overt associate' (Landau, 2013) that cannot be implicit, as seen in example (39a), but not in example (39b).

(39) a. They all gained something.

b. *Something has all been gained.

c. They wanted [EC to all gain something]

d. [EC to all gain something], they knew, would be a miracle.

In example (39a) the FQ *all* associates with the overt NP agentive subject *they*. Example (39b) is the passivization of example (39a). The object, *something*, in example (39a) moves to the subject position in example (39b) to fulfill features on T. However, a quantifier may float only when it is associated with an NP with an external theta role, as in (39a). In example (39b) the NP, *something*, does not have an external theta role. The quantifier may not float in example (39b), first, because *something* is not the agentive subject and second, because the agentive subject is implicit and the quantifier cannot associate with an implicit subject. The EC subject in the embedded clause in example (39c) must not be an implicit subject, because the FQ may freely associate with the embedded EC subject. And finally, the EC subject in (39d) must be an agentive subject and not implicit, because the FQ may freely associate with it.

3.2.3 Partial Control

The next piece of evidence hinges on the semantic value of the EC embedded subject. As can be seen in almost all the examples of OC introduced thus far, the EC subject takes its semantic value exclusively from

an argument in the matrix clause. Another possible type of semantic link between the EC subject and a controller NP is one where the EC subject only partially agrees with the controller NP in the matrix clause.

(40) The organizer$_i$ decided [EC$_{i+}$ to meet right before the parade].

In example (40) the matrix clause subject, *the organizer* is only a member of the set that the EC subject denotes. The empty subject means something along the lines of [the organizer + the group of people being organized/running the parade]. This is a crucial piece of evidence for the analysis of this type of subject and these structures. The EC subject is not just a semantic or structural copy of an argument in the matrix clause; like an NP-trace, it denotes a larger set. The referential index for the empty subject is independent from that of its controller. Although the semantic referential index for *the organizer* and the empty subject overlap, these indexes are distinct from one another. This would be impossible if the empty subject was only a copy of its controller. In that case, the index for the EC subject would have to be a full copy of the index for the controller. Given that the embedded subject is semantically independent, it must also be analyzed as syntactically distinct from the matrix controller.

3.2.4 Agreement/Split Control

Plural agreement, as seen between subjects and verbal morphology, is also dependent on syntactically overt NPs. Landau (2013) notes that British English accepts some forms of semantic plurality where the subject is syntactically singular, but the verbal morphology is plural, since it is possible to interpret the subject of the verb as being plural, as in example (41b).

(41) a. *John$_i$ hoped that his uncle$_j$ would be partners$_{i/j}$.

b. *This group is/are partners.

c. John$_i$ proposed to his uncle$_j$ [EC$_{i+j}$ to be partners]

In examples (41a) and (41b), plural agreement on the predicate is ungrammatical. In example (41a) the *partners* is marked with plural morphology, because it is referring to *John and his uncle*. The agreement is either with the implied subject *they*, or *John and his uncle*. In order to make example (41a) grammatical it would be necessary to insert an overt subject to agree with the embedded object: *John hoped that his uncle and he would be partners*. Again, in example (41b) *the group* is an implied plural subject, but neither the verb nor the object may grammatically mark this overtly with plural morphology. The expectation is that if the embedded EC subject is an implicit or semantic element, overt morphological agreement should fail in contexts similar to examples (41a) and (41b). This is not the case in example (41c). In this example the embedded object is marked with overt plural morphology and, unlike example (41a), this string is

grammatical. The embedded subject in (41c) must be a full syntactic subject, distinct from the matrix clause subject or object, otherwise the utterance is considered ungrammatical.

Example (41c) is an example of **split control**, a part of the group of Non-Exhaustive Control verbs. Although neither the subject nor the object of the matrix clause is plural, the object in the embedded clause that appears to refer to *John and his uncle* is perfectly acceptable when plural. The EC subject is capable of encapsulating the set of *[John and his uncle]*. Very simplistically, (41c) shows that the object in the embedded phrase is plural. This object must refer to a plural subject. Therefore the EC subject must be transmitting this plurality to that object. **Split Control** is a key piece of evidence for analyzing the EC subject as the independent syntactic subject PRO. If the subject in (41c) were the trace of *John*, or of *his uncle*, then the EC subject would be singular and would not allow the overt plurality on *partners*. Instead what we see is that the EC subject is semantically linked to two arguments in the matrix clause. The EC subject must be its own independent syntactic constituent. I will return to this discussion in Chapter 5.

3.2.5 Case-Concord

Case-concord refers to morphological agreement between a predicate and the NP to which the predicate applies. Normally the NP and the predicate agree with respect to case. However in examples (42) and (43) the embedded predicates do not agree, with respect to case, with the expected NPs. It appears that the agreement between the matrix NP and the embedded predicate has been blocked somehow. What is blocking the agreement, and from what are the embedded predicates receiving their case?

(42) Ona poprosila ego ne ezdit' tuda odnomu zavtra.
she.nom asked him.**ACC** not to-go there alone.**DAT** tomorrow
'She asked him not to go there alone tomorrow.' (Russian)

In example (42) the predicate *alone* is expected to agree with *him* in the matrix clause with respect to case. Instead of the expected accusative case, *alone* bears dative case.

(43) Olaf hafθi ekki gaman af aθ.vanta einan (i) veisluna.
Olaf.**NOM** had not pleasure of to.lack alone.**ACC** to the.party
'Olaf didn't like being the only one absent from the party.' (Icelandic)

Similarly, in example (43) the predicate *alone* has a quirky accusative case instead of following the usual case concord pattern by agreeing with the matrix subject in the nominative case.

The independence of the embedded dative and accusative case in examples (42) and (43) is surprising if the embedded clause is a small clause, with only a VP projection and with no structural subject. However if these two clauses are analyzed as having a subject that the embedded predicate agrees with, then the

patterns become straightforward. In Russian, embedded EC subjects bear dative case. In example (42) the predicate enters an agree relationship with the EC subject, therefore the embedded predicate bears dative case. In Icelandic, the embedded EC subject bears a quirky accusative case. The predicate *alone* enters an agree relationship with the EC subject and bears accusative case as well. Both these examples show that the embedded predicate has an *independent* case from the matrix subject and object, suggesting that this case must come from something within the embedded clause. This confirms the presence of a full structurally specified EC subject in both these clauses.

3.2.6 Binding

Effects of binding naturally seem to be a venue for positing a structural EC subject, however these arguments must be carefully constructed. Consider example (44). In example (44) the governing category for the reflexive *herself* could be the embedded clause, in which case the EC subject could be the antecedent licensing *herself*. However, the governing category, in which *herself* must be bound, could just be extended to the matrix clause allowing *herself* to find an antecedent in *Mary*.

(44) Mary$_i$ planned [EC$_i$ to buy herself$_{i/*j}$ a new coat].

Only those environments where a matrix clause argument cannot be the antecedent for an anaphor in the embedded clause will provide clear proof of a full structural EC subject. The examples in (45) illustrate this type of evidence. In example (45a) only the EC subject is available to serve as the antecedent for the anaphor. In example (45b) the EC subject is the only available antecedent for *herself*: *Mary's* position is inaccessible to *herself*. The NP *Mary* does not c-command *herself* and cannot serve as the antecedent for the anaphor.

(45) a. [EC$_i$ behaving oneself$_i$ in restaurants] would be necessary.

b. Mary$_i$ realized that it would be useless [EC$_i$ to nominate herself$_i$ for the job].

These examples and others hold, "...as long as one guarantees that the matrix controller is not a potential binder" (Landau, 2013). The EC subject in example (45a) has arbitrary reference and the subject in example (45b) is in an NOC context. However in examples (46a) and (46b), the EC subject is in an OC context. These examples are from Russian and German respectively:

(46) a. John ubedil Mary$_i$ [EC$_i$ navestit svoju$_i$ sestru].
John persuaded Mary EC to-visit SELF's sister
'John persuaded Mary to visit her sister.' (Russian)

b. Sie hat dem Hans$_i$ erlaubt [EC$_i$ sich$_i$ den Fisch mit Streifen vorzustellen.
She has the.DAT John allowed EC SELF the.ACC fish with stripes to-imagine
'She allowed John to imagine what the fish would look like with stripes.' (German)

The Russian *svoji* 'self' is a 'subject oriented' anaphor, an object in the matrix clause cannot be its antecedent. Therefore an EC subject must be present in the embedded clause of example (46a) to act as the antecedent for *svoji*. Similarly, the German *sich* 'self' cannot refer to a dative: it is an accusative and must have a nominative antecedent. The EC subject must be present in the lower clause of example (46b), licensing the presence of the reflexive.

It is important to rule out the possibility that implicit arguments can bind anaphors, otherwise the argument could be made that the anaphors in examples (45) and (46) are being licensed by implicit arguments. Example (47) shows that the reflexive in English cannot be bound by an implicit argument. The reflexive must have an antecedent that is structurally present in the syntactic derivation.

(47) They$_i$ remembered that John talked *(to them$_i$) about each other's$_i$ plans.

The object *each other's* cannot be bound by an implicit argument in example (47). Without an overt realization of *to them*, the string is ungrammatical. However, if the embedded antecedent *to them* is structurally (and overtly) available in (47), then the reciprocal *each other* is licit. In example (48), the reciprocal from example (47) is tested with an EC subject.

(48) a. *John talked with Mary about each other.

b. John$_i$ proposed to Mary$_j$ [EC$_{i/j}$ to help each other$_{i/j}$].

The reciprocal *each other* does not allow for its antecedent to be split, as in example (48a). This is a property of all such reciprocals. However, it appears that in example (48b) the reciprocal is allowing just that, an antecedent of *John and Mary*. Unless a plural embedded EC subject is posited, whose semantic value is *John and Mary*, the string is ungrammatical. With a plural embedded EC subject the reciprocal may be bound grammatically.

3.2.7 Overt Controllees

In some languages ,structures equivalent to the OC contexts presented so far, do in fact have overt subjects, as in example (49). These are cases where the matrix verb is from the set of verbs that select for OC contexts, the embedded subject is still controlled by an argument from the matrix clause, except that, instead of being covert, the subject is an overt NP.

(49) **O hóspede$_i$/nenhum hóspede$_i$** optará por fazer **ele$_i$** o pequeno-almoco todos os dias.
the guest$_i$/no guest$_i$ will.choose up to.do **he$_i$** the breakfast every the days.
The guest/No guest will choose to prepare his breakfast himself everyday. (Portuguese) (Barbosa, 2009 in Landau, 2013)

The matrix verb, *choose*, in example (49) is part of the cross-lingusitic set of verbs that select for OC complements. Crucially the embedded subject in (49) is dependent on the matrix subject. Even though *he* is overt, the matrix clause subject is still the antecedent for the pronoun, and without a controller in the matrix clause, this subject would be ungrammatical. Landau (2013) argues that these contexts which license overt subjects. should also be analyzed as OC contexts. Crucially, when an embedded OC subject is overt, it must still be semantically linked with the matrix clause subject. The embedded subject may not refer to a disjoint subject. Allowing overt subjects in OC contexts, further supports analyzing OC contexts as always containing full structural subjects.

3.2.8 Expletive Constructions

Chomsky (1981) observed that the expletive pronoun *it* may occur in embedded clauses as a subject, as in example (50). In this example (50) the following are conjoined with other clauses: *would be a shame* and *it is illegal*. In example (51) clauses with the empty subject are conjoined with the same clauses from (50): *would be a shame* and *it is illegal*. The results are grammatical. The whole clauses that contain the expletive pronoun may be replaced with whole clauses containing an EC subject, but in example (52) the empty subject may not replace the expletive pronoun itself.

(50) a. For it to be obvious that Bill is lying would be a shame.

b. It is illegal for it to be required that we wear helmets in in class.

(51) a. EC to lie would be a shame.

b. It is illegal [EC to wear helmets in class].

(52) a. *[EC to be obvious that Bill was lying] would be a shame.

b. *It is illegal [EC to be required to wear helmets in class].

This pattern shows that there are interpretive restrictions on the EC subject, namely it cannot replace an expletive pronoun. Landau (2013) argues that whatever the restrictions on the interpretation of the EC subject, they must be based off the existence of a subject that is structurally present. This pattern shows two things: that clauses must have a subject and that the embedded empty subject may not be an expletive pronoun.

Chapter 4

Previous Approaches

Thus far I have provided an overview of the arguments for positing an empty subject in non-finite embedded clauses in OC structures. In this section I will discuss the syntactic status of the empty subject in these structures and I will introduce the theoretical approaches that have been proposed to date.

One type of analysis is to assume that the EC subject of OC contexts is a PRO, as in example (53). This type of theory goes under the name of the *Agree Theory* in the literature.

(53) John$_i$ tries [PRO$_i$ to run]

In (53) the empty subject of the non-finite embedded clause is co-referential with the matrix clause subject. However, it is not always the case that the subject of the matrix clause controls the reference of the embedded subject. As seen in (54), the subject of the embedded clause can also be co-referential with the object of the matrix clause.

(54) [John tells Mary$_i$ [PRO$_i$ to run]]

The alternative analysis assumes that the empty subject of the OC contexts is a copy of a constituent that moves into the matrix clause, similar to the analysis assigned to Raising constructions. This type of theory is called the *Movement Theory* of OC constructions.

(55) a. John seems [~~John~~ to be happy].

b. John tries [~~John~~ to run]

The surface similarity of these two structures and the apparent underlying dissimilarity has been a topic of interest in syntax since Rosenbaum (1967). The Movement (copy) Theory of OC structures, treats raising and OC contexts as both being results of movement. The Agree Theory of OC structures treats raising and OC contexts as the results of different processes.

The division in approach between movement and agreement boils down to a theory internal problem: do we want to posit a 'new' syntactic element, namely PRO, and the mechanisms that must go with it? Or, do we try to develop an analysis that uses the syntactic elements already available in existing syntactic derivations. Alboiu (2007) suggests the difference between the two camps is reduction: the Movement Theory is reductionist, in that it gets rid of PRO and reduces the inventory of syntactic elements, while the Agree Theory is non-reductionists, in that it posits an independent EC, which is specific for OC structures.

In what follows, I will first offer a detailed presentation of reductionist approaches (including, but not limited to, the movement based analyses) and then will go on to present non-reductionist approaches to OC contexts and the empty subject of these embedded clauses.

4.1 Reductionist approaches

The reductionist approach has several different versions, each following from the basic idea that the embedded EC subject is some type of empty element, which is independently needed elsewhere in the grammar, rather than positing a new type of EC that occurs only in these contexts. I will address three of these analyses. The most prominent is the body of work by Hornstein and Boeckx (Hornstein, 1999, 2001, 2003; Boeckx and Hornstein, 2004, 2006, 2007, 2010). Manzini and Roussou (2000) present a slightly modified mechanism accounting for the subject of OC contexts. And finally, Wurmbrand (1998, 2004) proposes by far the most energetically reduced structure for OC contexts, in the sense that, in her proposal, the embedded verb in OC contexts have no syntactic subject at all.

4.1.1 Hornstein and Boeckx

The derivation for a control environment under this analysis looks like (56), which is similar to the derivation of the raising construction in (57).

(56) $[_{CP}\ [_{TP}\ \text{Mary}\theta_{1,2}\ [_{vP}\ \sout{\text{Mary}\theta_{1,2}}\ v\text{+hopes}\ [_{TP}\ \sout{\text{Mary}\theta_1}\ [_{vP}\ \sout{\text{Mary}\theta_1}\ v\text{+win}]]]]]$

(57) $[_{CP}\ [_{TP}\ \text{Mary}\theta_1\ [_{vP}\ \sout{\text{Mary}\theta_1}\ v\text{+seems}\ [_{TP}\ \sout{\text{Mary}\theta_1}\ [_{vP}\ \sout{\text{Mary}\theta_1}\ v\text{+be happy}]]]]]$

The interpretational properties of the embedded subject fall out from the effects of movement under this analysis: the moved copy must be a perfect copy of the embedded subject and it must also be a c-commanding antecedent for the base position. In example (56) the subject *Mary* enters the derivation in the specifier of the vP domain. The TP attracts the subject to its specifier, leaving behind a copy of *Mary* in the position where it entered the derivation. Next, *Mary* moves to the specifier of the vP in the matrix clause and from there *Mary* moves to the NP's final position, the specifier of TP. The derivation in (56) mirrors the raising

structure, save for the fact that *Mary* has two theta-roles in (56) and only one in the raising derivation in example (57).

This difference, that the subject of the OC context has two theta roles and the subject of the raising derivation only has one, at first glance, poses serious problems for this theoretical approach. The theta criterion has always restricted only one theta role for each argument in a derivation. However Hornstein argues that this restriction is stipulative, does not hold under the Minimalist Program, and should be abandoned. Before the advent of the Minimalist Program, theta positions were only available at the D-structure level. Arguments could only enter the derivation in those positions, and therefore those positions were never available as targets for movement (although arguments were free to move out of them): arguments could not move into a theta position once they were already merged into the derivation. Once D-structure had been eliminated, the motivation for restricting theta positions for arguments entering the derivation was no longer a valid restriction, according to Hornstein. The pronunciation site of the subject is determined by Case: only a Case position is pronounced at PF (Nunes, 2004 in Landau, 2013). The advantage of this is that backward control is accepted and possible, where the matrix subject is the covert subject and the embedded subject is the overt copy (Alboiu, 2007). Although this can allow for an argument to have two theta roles, it still does not explain why an argument would have two.

There are several advantages of the movement theory: (i) it is not necessary to posit a new EC for OC structures and the lexical inventory is reduced; (ii) no special grammatical mechanism is needed to account for the coindexation between the EC subject and an NP in the matrix clause; and (iii) the distribution and interpretation of the embedded subject is a result of grammatical movement.

On the other hand, the movement analysis also faces challenges, not the least of which is that this model cannot account for cases of split (example (40) repeated below as (58)) and partial control (example (41c) repeated below as (59))

(58) The organizer decided [EC to meet right before the parade].

(59) John$_i$ proposed to his uncle$_j$ [EC$_{i+j}$ to be partners].

In the case of split control, it is impossible for two arguments to be linked to the same structural position under a Movement theory. In the case of partial control, if the embedded subject is a perfect copy of the matrix controller, the embedded subject must be a full match syntactically and semantically. This does not allow for the referential index of the embedded subject to differ at all from its controller.

This approach also over generates data, as in (60), as discussed in Landau 2013.

(60) *John$_i$ was hated [t$_i$ to live like that].

Example (60) is a case of ungrammatical passivization. The Movement theory allows the subject, *John*, to move out of the embedded clause in example (60) to the matrix clause, accumulating the agentive theta role from the embedded verb and the patient theta role from the matrix clause verb.

4.1.2 Manzini and Roussou

The derivation for a control environment under the analysis by Manzini and Roussou (1998, 2000) (M & R) looks like (61). Crucially, M & R do not posit a full structural subject in the embedded clause.

(61) [John$\theta_{1,2}$ [θ_1 tried [θ_2 read]]]

In (61) *John* is merged directly in the matrix clause, into the specifier of vP, and then moves to the specifier of TP. The NP *John* then attracts both the external theta role from the matrix clause verbal domain and the external theta role from the embedded clause verbal domain. M & R's proposal is that NPs may only merge into their case position. The result for OC contexts is that there is no structural subject in the embedded clause. NPs attract theta roles from verbs (theta roles are assumed to be aspectual features of the verb in this analysis). A NP may attract as many theta roles as are within its minimal domain, up to the next NP. Therefore the interpretational features of the controller in OC contexts are due primarily to the fact that one NP in the matrix clause attracts two theta roles as shown in example (61).

Just as other reductionist approaches, this analysis manages to account for OC facts with machinery that is independently needed, and avoids positing new categories that would be specific to OC contexts. However, the challenges for this proposal follow from the implications about the nature of the embedded subject. As discussed in Chapter 3, Section 3.2, OC contexts must be analyzed as having a full structural subject, but the analysis by M & R does not posit a structural subject in the embedded clause. This analysis, like the Movement theory, cannot account for split or partial control. Both split and partial control require a structurally present subject in order to account for the properties of both situations. Neither split or partial control contexts allow for implicit subjects. Finally, the motivation for a NP to attract more than one theta role is stipulative (Alboiu, 2007).

4.1.3 Wurmbrand

The final reductionist approach to OC contexts that I will discuss is Wurmbrand's (1998, 2004) analysis. The derivation according to this analysis appears in (62).

(62) [$_{CP}$[$_{TP}$ John [$_{vP}$ ~~John~~ [$_{VP}$ tried [$_{VP}$ to eat the cake]]]]].

The embedded clause is reduced completely to a big VP, therefore there is no structure into which a subject could merge in the lower clause. The subject merges directly into the matrix clause, in the specifier of vP, and moves to the specifier of TP. Just as in M & R's analysis, in Wurmbrand's approach there is no structural subject in the embedded clause. The embedded verb is interpreted as sharing an argument with the matrix clause because of semantic rules in this analysis.

The motivation for positing such a reduced structure for these embedded clauses comes from languages that allow scrambling out of the embedded clause, namely German and Italian. These languages allow for arguments from the embedded clause to move to the matrix clause without changing the interpretation of the utterance. This behaviour suggests that the boundary between the matrix clause and the embedded clause in these contexts is porous or weak.

This analysis is advantageous because, again, the grammatical mechanisms are reduced and control is the result of other phenomena. The challenge is, again, that this structure cannot account for split or partial control. Both split and partial control require that the embedded clause have a structural subject that is distinct from the matrix clause argument(s). The embedded clause EC must have its own structure and referential index.

4.2 Non-reductionist Approaches

The agreement analysis of the EC subject of non-finite embedded clauses is the modern iteration of Control Theory. The most prominent advocate for this approach to Control is Idan Landau, who has continuously added to the literature on OC for over a decade (Landau 1999, 2001, 2004, 2008, 2013, 2015).

4.2.1 Agree

The Agree model of OC functions by establishing agreement relationships between the matrix clause arguments and the embedded clause subject. The embedded subject is assumed to be a PRO, which establishes an antecedent relationship with a matrix NP.

4.2.1.1 The Features of the Agree Model

The basic apparatus for the Agree model is the interaction of the features [+/– T] for semantic tense, [+/– Agr] for agreement, and [R] for [+/– R(eferential)].

The [T] feature refers to semantic tense, and it distinguishes between cases where the embedded OC clause can be interpreted with an independent semantic tense, as in example (63), and cases in which the embedded OC clause does not have independent tense, as in (64).

(63) This morning Harry decided [PRO to leave tomorrow].

(64) a. This morning Harry forgot [PRO to leave].

b. *This morning Harry forgot [PRO to leave tomorrow].

In example (63) the matrix clause has a different semantic tense than the embedded clause, as shown by the use of independent time adverbials in the two clauses ('this morning' vs 'tomorrow'). The embedded clause is valued [+T]. In example (64) the embedded clause cannot have an independent semantic tense from the matrix clause, as shown by the ungrammaticality of (64b), where the embedded clause uses an independent time adverbial. The embedded clause in example (64) is valued [–T].

The second feature —[Agr], refers to overt morphological marking of agreement on the verb in an embedded clause, as in example (65) from Hungarian.

(65) Kellemetlen volt Péternek$_i$ [PRO$_i$ későn érkez-ni-e].
unpleasant was Perter.DAT [PRO late arrive-INF-3SG].
'It was unpleasant for Peter [PRO to arrive late].' (Landau 2013)

(66) John tried [PRO to sing].

The verb in the embedded clause in (65) is overtly marked for agreement and is therefore valued [+Agr]. The verb in the embedded clause in example (66) does not have overt agreement marking and is therefore valued [–Agr].

The feature [+/– R] points to the referential nature of NPs or Ts.

4.2.1.2 The Implementation of the Agree Model

PRO is inherently marked as [–R] in this system and therefore it can only be licensed by a T head that is [–R] itself. According to Landau, the T head cannot have a [–R] feature if both its [T] feature and its [Agr] feature are positively valued. Hence, a T head with a [+T] and [+Agr] will not be able to license PRO. PRO is licensed only in the 'elsewhere' cases of the [+T] and [+Agr] feature combination on the T head.

Even if PRO is licensed by the [–R] feature on the embedded T head, its phi features still need to be valued. If the embedded verb is an infinitive, the embedded T head bears [–Agr] features and hence cannot value the phi features on PRO. PRO will thus depend on a matrix element for the valuation of its phi features. The matrix head that is able to value the phi features on PRO is the matrix T head. Given that the matrix T head also agrees with the matrix subject in its phi features, the result is that the matrix subject, the matrix T and PRO will come to share the value of their phi features.

(67) Sbj$_i$ T$_i$ . . . [PRO$_i$ T . . .]

Landau also discusses an alternative analysis, which obtains if the embedded C head also bears phi features. Not all C heads bear phi features, but in case they do, they come to play a role in the valuation of the phi features on PRO. This is because the syntactic position of C is in between the matrix T and PRO.

(68) Sbj_i T_i . . . [C_i PRO_i T . . .]

In this situation, the matrix T will enter an agree relation with the embedded C, rather than directly with PRO, given that C is closer to the matrix T than PRO. Once its phi features are valued from the matrix T, the C head will enter an agree relation with PRO and will value the latter's phi features. The result is similar to the one obtained in (67), with the exception that in (68) the relation between the matrix T and PRO is mediated by C. The situation in (68) is relevant for the derivation of Partial Control. Recall that in Partial Control the controller is a semantic subset of the set denoted by PRO. In Landau's analysis the difference between exhaustive control and Partial Control does not concern the phi features on PRO. The valuation of the phi features on PRO ultimately has the same source in both situations, namely the matrix T. The difference can be explained by positing an additional feature, that Landau calls [MER] (mereological), that Ts and NPs bear, but Cs do not. Thus, if PRO and the matrix T agree, via the intervening C head, the value of the phi features on the matrix T will be transmitted to PRO, but the value of the [MER] feature on the matrix T will not be transmitted, because C does not have a [MER] feature. This allows for the possibility of PRO and the matrix T head (and implicitly the controller) having disjoint values for the [MER] feature. If this is the case, Partial Control occurs.

The advantage of the Agree theory is that the model can account for the cross-linguistics distribution of OC contexts and it can account for both Exhaustive Control and Non-Exhaustive Control. However, the model is not without its own challenges. Landau (2015) points out several challenges this model faces. Here, I will highlight only two of these challenges. The first is that the [R] feature is stipulative: it does not have independent morphosyntactic motivation. The second challenge is that the Agree Model contains PRO, adding additional material to the lexicon and grammar.

The OC Signature

Landau (2013) proposes a definition for OC Control structures, an *OC signature*. This signature is presented below in (69):

(69) *The OC Signature*

In a control construction [....X_i....[$_s$$\text{PRO}_i$....]....], where X controls the PRO subject of the clause S:

a. The controller(s) X must be (a) codependent(s) of S.

b. PRO (or part of it) must be interpreted as a bound variable. (Landau, 2013)

The OC Signature is meant to distinguish OC structures from all others (NOC, NC etc). The controller is identified as X in order to include those instances where PRO is controlled by an implicit argument.

The first condition in (69) can be illustrated in ellipsis contexts. An EC subject in an elided clause may only be controlled by a codependent of the elided clause. More specifically, in (70), the PRO in the elided clause can be controlled by *Sue*, which is a codependent of the elided clause, but not by *Mary*, which is not a codependent of the elided clause.

(70) Mary$_i$ forgot [PRO$_i$ to attend the ceremony] and Sue$_j$ did too ~~[expect [PRO$_{j/*i}$ attend the ceremony]]~~.

The second condition in (69) can be illustrated in the following context from Landau, 2013. Suppose Mary, Rachel, and Jane all play a game. There are two outcomes: one is that they disagree as to who won the game; and two is that they all agree that one discreet individual won the game, Mary for example. These scenarios are described by (71a), (71b), respectively.

(71) a. Mary claims that she (Mary) won the game. Rachel claims that she (Rachel) won the game. And Jane claims that she (Jane) won the game.

b. Mary, Rachel, and Jane claim that Mary won the game.

These scenarios produce different truth conditions for (72a) and (72b). In example (72a), the utterance with the embedded PRO subject may only be interpreted as true in (71a) and false in (71b). However, the embedded pronominal in example (72b) is interpreted as true in (71a) and (71b).

(72) a. Only Mary claims [PRO$_i$ to be the winner].

b. Only Mary claims [that she$_i$ is the winner].

Given that under the scenario in (71a), the winner varies with the claimer, PRO in (72a) gets a bound variable reading: Mary = only x [x claimed x is the winner]. However, the example (71b) is true under both the bound variable reading and a strict read for PRO. The latter is: Mary = only x [x claimed Mary is the winner]. The case with the embedded PRO subject is only true with a the bound variable reading, whereas the case with the pronoun is true with both readings. This property is what is defined under part (69b) of the OC signature: in OC contexts the EC subject is interpreted as a variable (at least partially) bound by the controller. (Landau, 2013)

Chapter 5

Persian

This chapter will present the relevant Persian data for a discussion of the Persian empty subject. I will introduce basic facts of word order in Persian. The data presented in the following sections leads to the resulting facts of Persian: i) the basic word order is SOV; ii) Persian is *pro–drop*, containing the EC *pro* in its lexicon; iii) the Persian verbs presented in Section 5.2 exhibit certain characteristics: they select for embedded phrases with an EC subject that appears to be linked to the matrix subject for its semantic reference, suggesting these may be OC environments; iv) the verbs in the embedded clauses in these contexts are marked with the subjunctive and also person and number agreement agreement. The research questions I will seek to answer with respect to the Persian data are the following:

(I) Is there OC in Persian?

It is possible the Persian data is in fact cases of *pro–drop*, however, if there is at least one instance where it is clearly not *pro–drop*, then the EC subject must be analyzed as either PRO or trace/copy and the grammatical analysis must either be based on an agreement processes or on movement.

(II) If there is at least one instance of OC in Persian then the following questions must be answered:

(i) Which Persian verb/verbs select for OC Control complement clauses?

(ii) What are the properties of the embedded Persian OC Control clauses?

(iii) And finally what is the nature of the Persian embedded EC subject? Is it PRO or trace/copy?

5.1 General Properties

5.1.1 Word Order

The basic word order in Persian is SOV, as seen in example (73). In example (73), the subject is first in the linear order and the object precedes the verb in the string.

(73) Məzen naːn mi-xoradə
Məzen bread is-eating.3SG
'Məzen is eating bread.'

Although the basic word order in Persian is SOV, clausal complements occur to the right of verbal heads, with the verbal head in the initial position, as in example (75).

(74) sima sib-o mi-xad.e
Sima apple-OM DUR-want.3SG
'Sima wants the apple.'

(75) sima mi-xad.e [sib-o be-xor-e].
Sima DUE-want.3sg [apple-OM SUB-eat-3SG].
'Sima wants to eat an apple.'

In example (74) the direct object *sib–o* (apple) appears before the verb, marked with the Object Marker (OM) *–o*. In example (75) the complement clause appears after the verb. In the complement clause the direct object, *sib–o*, appears before the verb and is marked with the OM.

5.1.2 *little pro*

Persian is also a *pro drop* language. Persian allows for *pro* subjects in both the matrix clause and the embedded clause, see examples (76) and (77) below.

(76) a. sima sib-o mi-xad.
Sima apple-OM DUR-want.3SG
'Sima wants the apple.'

b. *pro* sib-o mi-xad.
pro apple-OM DUR-want.3SG
'Sima/she/he wants the apple.'

(77) a. ʒian goft [(ke) Sima mi-ad].
Jian says.PST.3SG [(COMP) Sima DUR-come.3SG]
'Jian said Sima will come.'

b. ʒian goft [(ke) *pro* mi-ad].
Jian says.PST.3SG [(COMP) *pro* DUR-come.3SG]
'Jian said he/she will come.'

In example (76) the subject of the declarative clause can alternate between an overt NP, *Sima*, and *pro*. In example (77) the embedded clause subject can also alternate between an overt NP, *Sima*, and *pro*.

5.1.3 Infinitive/subjunctive

The embedded clauses selected by OC verbs are typically non-finite clauses and often contain infinitive verb forms, like those from the English that we have seen so far, as in example (78).

(78) John tried [PRO **to run** to the store].

However in Persian, there is no infinitive form. Instead, the Persian equivalent of the English infinitive form is the subjunctive, as in example (79).

(79) Sima dust-dar-e [ketab be-xun-e].
Sima friend-have-3SG [book SUB-**read-**3SG]
'Sima likes reading books.'

Subjunctive clauses in general may have overt subjects, as well as an EC subject *pro*, as seen in examples (80a) and (80b).

(80) a. *pro* fekr=mi-kon-æm [(ke) færda nasser bi-ad].
pro thought=DUE-make-1SG [(ke) tomorrow Nasser SUB-come.3SG]
'(I) think that Nasser is coming tomorrow.'

b. *pro* fekr=mi-kon-æm [(ke) færda *pro* bi-ad].
pro thought=DUR-come.3SG]-make-1SG [(ke) tomorrow *pro* SUB-come.3SG]
'(I) think that Nasser/he/she is coming tomorrow.'

Example (80a) shows a subjunctive complement clause with an overt subject while (80b) shows that the same subjunctive clause as that in (80a) can appear with *pro* in the embedded subject position, instead of the overt NP.

5.2 Control Verbs

As expected given the description above, subjunctive complements of OC verbs follow the matrix clause verb directly. The following is a group of verbs that cross-lingustically are classified as OC verbs (cited above in Chapter 2, Section 2.2.4). This paradigm is taken from Ghomeshi (2001) and includes the following 5 verbs: *to be able/can*, in example (81), *to forget*, in example (82), *to try*, in example (83), *to decide*, in example (84), and *to want*, in example (85).

(81) (mæn) mi-tun-æm [(ke) PRO be-r-æm].
(I) DUR-be.able-1SG [(COMP) PRO SUB-go-1SG]
'I am able to go'

(82) (mæn) yad-æm-tæft [(ke) PRO be-r-æm].
(I) memory-1SG-cl.PST.3SG [(COMP) PRO SUB-go-1SG]
'I forgot to go.'

(83) (mæn) sæ'iy-kærd-æm [(ke) PRO be-r-æm].
(I) try.do.PST-1SG [(COMP) PRO SUB-go-1SG]
'I tried to go.'

(84) (mæn) tæsmin-gereft-æm [(ke) PRO be-r-æm].
(I) decision-get.PST-1SG [(COMP) PRO SUB-go-1SG]
'I decided to go.'

(85) (mæn) mi-xa-m [(ke) PRO be-r-æm]
(I) DUR-want-1SG [(COMP) PRO SUB-go-1SG
'I want to go.'

In all the examples (81) – (85), the embedded clause verb is marked with the subjunctive. The embedded clause is also introduced by what appears to be a complementizer, *ke* 'that', which is optional in all these cases. The status of *ke* will be elaborated on in Chapter 6, Section 6.1.1 where I will present two previous studies on *ke*: the first that suggests *ke* is a clitic and the second argues that *ke* is a complementizer. For the moment, *ke* will be glossed as a complementizer. The EC subject in these examples is interpreted as coreferent with the matrix subject and the person and number agreement on the embedded verb match the person and number agreement on the matrix verb.

5.2.1 Genuine versus Bogus OC in Persian

Looking at the data above, it seems that there are 5 potential verbs that could be cases of OC in Persian: *to be able, to forget, to try, to decide, and to want.* However in Persian only two of these verbs actually are cases of genuine OC, *to be able* and *to forget*: these verbs select for embedded clauses with EC subjects that must be coreferent with a NP in the matrix clause. This can be shown by testing the two properties of OC contexts mentioned in Chapter 2, Section 2.2.4, namely the availability of covert and overt disjoint subjects in clauses embedded under OC verbs.

I will first address the availability of disjoint reference for a covert embedded subject. To illustrate the idea of disjoint reference for the covert embedded subject, consider example (86).

(86) (mæn) fekr=mi-kon-æm [(ke) *EC* bi-ad].
(I) thought=DUR-make-1SG [(COMP) *EC* SUB-come.3SG]
'I think he will come.'

In this example, the matrix verb is not an OC verb and the embedded covert subject is not semantically linked to a matrix NP. The covert subject has an independent semantic reference. This can be seen, because the matrix clause subject is in the first person (as marked on the verb) and embedded clause is in the third person (as marked on the embedded verb). Therefore the phi features of the matrix clause subject and the embedded clause subject are different. This is what counts as disjoint reference.

Testing the property of disjoint reference on the five contexts in Section 5.2, we notice that a covert embedded subject with disjoint reference is possible only with some of the verbs listed in Section 5.2, but not with all of them.

(87) *(mæn) mi-tun-æm [(ke) PRO bi-ad].
(I) DUR-be.able-1SG [(COMP) PRO SUB-come.3SG]
'*I am able him to come.'

(88) *(mæn) yad-æm-tæft [(ke) PRO bi-ad].
(I) memory-1sg-cl.PST.3SG [(COMP) PRO SUB-come.3SG]
'*I forgot him to come.'

(89) (mæn) sæ'iy-kærd-æm [(ke) PRO bi-ad].
(I) try.do.PST-1SG [(COMP) PRO SUB-come.3SG]
'*I tried him to come.'

(90) (mæn) tæsmin-gereft-æm [(ke) PRO bi-ad].
(I) decision-get.PST-1SG [(COMP) PRO SUB-come.3SG]
'*I decided him to come.'

(91) (mæn) mi-xa-m [(ke) PRO bi-ad]
(I) DUR-want-1SG [(COMP) PRO SUB-come.3SG]
'I want him to come.'

In examples (87) and (88) disjoint reference on the embedded subject is not grammatical. In examples (89) through (91) the disjoint agreement is accepted as grammatical.[1] This suggests that the verbs in examples (89) through (91) are not cases of genuine OC in Persian, but examples (87) and (88) are.

Overt embedded subjects of OC verbs are subject to the same restriction, namely they cannot have a reference that is disjoint from the reference of a matrix NP.

(92) *ʒian mi-tun-e [(ke) æskan be-r-e]
Jian DUR-be.able-3SG [(COMP) Ashkan SUB-go.3SG]
'*Jian is able Ashkan to go.'

(93) *ʒian yad-eʃ=ræft [(ke) æskan be-r-e]
Jian memory-3SG.CL=go.PST.3SG [(COMP) Ashkan SUB-go.3SG]
'*Jian forgot Ashkan to go.'

[1] In Ghomeshi, 2001 disjoint agreement was questionable in examples (89) and (90). However in this current study, the informants found these examples to be grammatical.

(94) ʒian sæ'iy=kærd [(ke) æskan be-r-e]
Jian try=do.PST.3SG [(COMP) Ashkan SUB-go.3SG]
'Jian tried Ashkan to go.'

(95) ʒian tæsmim=gereft [(ke) æ skan be-r-e]
Jian decision=get.PST.3SG [(COMP) Ashkan SUB-go.3SG]
'Jian decided Ashkan to go.'

(96) ʒian mi-xa-d [(ke) æ skan be-r-e]
Jian DUR-want-3SG [(COMP) Ashkan SUB-go.3SG]
'Jian wants Ashkan to go.'

The verbs in examples (92) and (93) cannot accept an overt disjoint subject in their embedded clauses. The verbs in examples (94) through (96) select for an embedded clause with an overt disjoint NP subject.[2] Thus, the same three verbs that were not genuine OC verbs with respect to the first property (*sæ'iy=kærd* 'try', *tæsmim=gereft* 'decide', and *mi-xa-d* 'want'), are not genuine OC verbs according to this second property, either. Similarly, the verbs *mi-tun-e* 'to be able to/can' and *yad-eʃ=ræft* 'to forget' in (92) and (93) are the only possible OC contexts according to this second property.

The following analysis of Persian OC will focus exclusively on the verbs *mi-tun-e* 'to be able/can' and *yad-eʃ=ræft* 'to forget', the only ones that display genuine OC properties.

[2]In Ghomeshi (2001) the verbs in examples (94) and (95) marginally accepted an overt subject in their embedded clauses. However, the informants in the current study did not find these cases marginal.

Chapter 6

Anaylsis

This chapter will evaluate the data introduced in Chapter 5 against the types of analyses discussed in Chapter 4, and will determine the most appropriate analysis for the Persian OC verbs. I will conclude that the analysis of Persian OC must be an Agree-type analysis and the embedded subject must be PRO.

6.1 Reductionist Approaches to the EC subject

The reductionist approaches, namely M & R and Wurmbrand, where the interpretational effects of the EC subject are analyzed as being semantic rather than structural, cannot account for the OC structures in Persian. These two approaches will be assessed and discarded in the next sections 6.1.1 and 6.1.2.

6.1.1 Wurmbrand and Persian OC

Wurmbrand's proposal is that the embedded clause projects no higher than a VP. This type of analysis has been attempted in Ghomeshi (2001), who proposes a modified version of Wurmbrand's analysis.

6.1.1.1 Ghomeshi (2001)

Ghomeshi (2001) proposes that the genuine OC verbs, like *to be able* and *to forget*, select for only a vP complement. Ghomeshi cites Wurmbrand's analysis of restructuring verbs in German to support her analysis of Persian. One of the motivations for Wurmbrand's analysis (OC contexts project only a VP) is the fact that in languages like German, objects can scramble out of the embedded clause, without changing the interpretation of the clause as a whole. Ghomeshi argues that genuine OC verbs in Persian also allow for object scrambling with little or no interpretational changes, just as in German, as in example (97).

(97) a. Askan **ketab–o** mi-tun-e [(ke) EC **ketab–o** be-xun-e].
Ashkan **book**–OM DUR-be.able-3SG [(COMP) EC **book**–OM SUB-read-3SG]
'Ashkan is able to read the book.' (Ghomeshi, 2001)

b. Askan **ketab–o** yad-eʃ=raeft [(ke) EC **ketab–o** be-xun-e].
Ashkan **book**–OM memory-3SG.CL=go.PST.3SG [(COMP) EC **book**–OM SUB-read-3SG]
Ashkan forgot to read the book. (Ghomeshi, 2001)

Further, Ghomeshi shows that the verb *to be able* gives rise to a parasitic interrogative interpretation as in example (98). Ghomeshi argues that this parasitic interrogative reading is evidence that the lower clause lacks a WH-feature slot, which is traditionally posited in the C domain. This, she argues, is similar to the clitic climbing effects that can be seen in Italian. In languages like Italian, clitics may climb out of clauses embedded under OC verbs to the matrix clause without changing the overall interpretation of the sentence.

(98) Askan **tʃi** mi-tun-e [(ke) EC **tʃi** be-xun-e].
Ashkan **what** DUR-be.able-3SG [(COMP) EC **what** SUB-read-3SG]
'What is Ashkan able to read?' (Ghomeshi, 2001)

In contrast, verbs that select for non OC contexts do not allow for object scrambling and do not have a parasitic interrogative reading. Rather, in non OC contexts object scrambling induces interpretational differences, and movement of a WH-constituent out of the embedded clause induces interrogative scope effects.

(99) Ashkan **ketab–o** fehr=mi.kon-e [(ke) EC ~~**ketab–o**~~ be-xun-e].
Ashkan **book**–OM thought=DUR-do-3SG [(COMP) EC **book**–OM SUB-read-3SG]
'Ashkan thinks he will read the book.'

(100) Ashkan **tʃi**$_1$ fehr=mi.kon-e [(ke) EC ~~**tʃi**~~$_2$ be-xun-e].
Ashkan **what** thought=DUR-do-3SG [(COMP) EC **what** SUB-read-3SG]
1. 'What does Ashkan think he will read?'

2. 'Ashkan is thinking what to read'

In example (99) when the object is scrambled to the matrix clause, it induces discourse or pragmatic effects. An object scrambled out of an non OC context must be interpreted as contrastive. In example (98) scrambling the WH-element to the matrix clause induces the wide scope reading (1.) and when the WH is left in the embedded clause the WH-element induces narrow scope (2.). The behaviour of object scrambling in Persian and the parasitic interrogative reading constitute Ghomeshi's first argument for the vP analysis. Second, Ghomeshi crucially argues that the particle *ke* is not a complementizer, arguing that *ke* is enclitic on the matrix verb. This allows the embedded clause under *to be able* and *to forget* to be reduced to vP.

6.1.1.2 Against Gomeshi/Wurmbrand

There are several reasons why an analysis along the lines proposed in Wurmbrand/ Gomeshi won't work for Persian. First, Persian embedded clauses show evidence for structure beyond a VP. The verb in the embedded clause is inflected for phi features (person and number, specifically) and the clause is itself introduced (optionally) by the complementizer *ke*, as shown in example (101). This suggests that the clause must project as high as a CP.

(101) Məzen mi-tun-e [(ke) PRO ʃatraŋ bosi be-kun-e].
Məzen DUR-able-3SG [(COMP) PRO chess play SUB-do-3SG]
'Məzen can/is able to play chess.'

The particle *ke* shows robust evidence that it cannot be analyzed as a clitic. Darzi (2008) shows that when *ke* is omitted and the temporal adverb appears at the end of the matrix clause, just before the embedded clause, the utterance is ambiguous.

(102) maen fekr+mi-keard-aem **haemiʃe** ke u to-ra daer moqabel-e digaeran
I thought=DUR-do.PST-1SG **always** that she/he you-OM in front-EZ others
saerzaeneʃ+be-kon-e.
blame=SUB-do-3SG
'I always thought that she/he would blame you in front of others.'

(103) maen fekr+mi-keard-aem ke **haemiʃe** u to-ra daer moqabel-e digaeran
I thought=DUR-do.PST-1SG that **always** she/he you-OM in front-EZ others
saerzaeneʃ+be-kon-e.
blame=SUB-do-3SG
'I thought that she/he would always blame you in front of others.'

(104) maen fekr+mi-keard-aem **haemiʃe** u to-ra daer moqabel-e digaeran
I thought=DUR-do.PST-1SG **always** she/he you-OM in front-EZ others
saerzaeneʃ+be-kon-e.
blame=SUB-do-3SG
'I always thought that she/he would blame you in front of others.'

'I thought that she/he would always blame you in front of others.'

In example (102) the adverb appears before *ke*, modifying the matrix verb; in example (103) the adverb appears after *ke* modifying the embedded verb; and in example (104) *ke* is omitted and the adverb appears between the matrix clause and the embedded clause, rendering the utterance ambiguous.

Darzi (2008) argues that if *ke* is a clitic in control contexts, then even when the adverb appears after *ke* it should still be able to modify the matrix clause verb, giving rise to an ambiguous reading. However example (105) shows that this interpretation is not available under these circumstances.

(105) maen mi-tun-aem ke **haemiʃe** u to-ra daer moqabel-e digaeran
I DUR-be.able-1SG that **always** she/he you-OM in front-EZ others
saerzaeneʃ+be-kon-e.
blame=SUB-do-3SG

'I thought that she/he would always blame you in front of others.'

'*I always thought that she/he would blame you in front of others.'

Therefore the particle *ke* cannot be a clitic on the matrix verb and should be analyzed as a complementizer, delineating the seperation between the matrix clause and the embedded clause. Others have argued for some version of this analysis as well. For example, Ilkhanipour (2014) argues that *ke* is some type of subordinate clause marker (following Hashemipour (1989)) and that it should be analyzed as a Fin head, or some other finer grained head in the C domain. Nevertheless, the conclusion is that this particle cannot be enclitic on the verb and must be analyzed, at the very least, as part of the C domain.

Given the evidence presented above, Persian embedded clauses of OC verbs cannot be analyzed as VPs/vPs.

6.1.2 Manzini and Roussou and Perisan OC

The proposal by Manzini and Roussou would treat the Persian data by merging the subject directly into the matrix clause and attracting the theta roles from both the matrix clause and the embedded clause, as illustrated in example (106).

(106) a. Məzen mi-tun-e [(ke) ʃatraŋ bosi be-kun-e].
Məzen DUR-able-3SG [(COMP) chess play SUB-do-3SG]
'Məzen can/is able to play chess.'

b. Məzen$_{1,2}$ [θ_1 mi-tun-e [(ke) θ_2 ʃatraŋ$_3$ bosi be-kun-e θ_3].

In example (106) the embedded verb *bosi be-kun-e* 'play-do' has two theta roles to disperse: the internal theta role to *ʃatraŋ* 'chess' and the external role to the matrix subject *Məzen*. The matrix verb *mi-tun-e* 'can/be able' disperses its one theta role to the matrix subject.

Although we can reduce the grammatical mechanism by using this model, it is undesirable because it cannot account for the Persian OC contexts. Besides the problem of motivating the attraction of theta features to the matrix subject, it is not clear how or why the external theta feature on the embedded T can trigger the person agreement on the subjunctive verbs in these contexts, given that there is no structural embedded subject in this analysis. M & R claim that OC contexts are cases of structure sharing, and that the embedded clause is not a full CP and forms one clause with the matrix clause. This cannot be the case for Persian, since the embedded clauses are optionally introduced by a complementizer in Persian. A clause introduced by a complementizer, as seen in Chapter 3 in Section 3.1.1 and just above in Section 6.1.1, is a full CP.

Therefore Persian OC complements cannot be examples of structure sharing.

6.2 Are Persian OC Constructions Movement or Agreement?

Having discarded both Wurmbrand's and Manzini & Roussou's approaches as possible analyses of the Persian data, we are left with three possibilities: the empty subject of embedded OC contexts is (i) a *pro*, (ii) a PRO, or (iii) an NP-copy.

The following will outline what we would expect to see in each case. In order to decide on the status of the EC subject under OC verbs in Persian, I will use the following working criteria:

- ***pro*** – If the EC in question is *pro* it should behave in the following ways:

  (i) it should show contrastive distribution with overt NPs;

  (ii) it should be able to allow for disjoint agreement on the embedded verb.

- **NP–copy** – If the EC subject is a Trace/Copy, it will:

  (i) be in complementary distribution with overt NPs;

  (ii) not allow for disjoint agreement on the embedded verb;

  (iii) not allow for split or shared agreement;

  (iv) allow the embedded NP to be overt, but the controller to be covert, see example (107).

- **PRO** – If the EC subject is PRO it will:

  (i) be in complementary distribution with overt NPs;

  (ii) not allow for disjoint agreement on the embedded verb (provided it is not a case of split or partial control);

  (iii) allow for split control.

First I will show that the subject may not be *pro*, thereby discarding this type of analysis. Next, I will distinguish between whether the Hornstein type reductionist analysis positing movement or Landau's agreement analysis is the best analysis for Persian. In order to decide between the movement analysis and the agree analysis, I will use the following criteria:

- **Phasal status**.

  Given that the NP shared between the embedded clause and the matrix clause in OC contexts is an argument, we expect its movement out of the embedded clause to be possible only if the embedded

clause is non-phasal (i.e. has a reduced structure). If the embedded clause turns out to be a phase, then agreement is the appropriate analysis. The phasal status of the embedded clause will be evaluated based on the following properties:

(i) Is the embedded clause headed by a lexical complementizer? If yes, then the clause is phasal. If no, then the clause not phasal.

(ii) Is the T temporally saturated and phi complete? If yes to both of these, then the clause is phasal. If no, then the clause is not phasal.

(iii) Can the embedded domain value case on the subject? If yes then the clause is phasal. If no then the clause is not phasal.

- **Overtness of the shared argument.**

  Does the shared argument show up as an overt copy in the embedded clause, as in example (107)?

(107) (**Victor**) incearca (**Victor**) [sa cinte (**Victor**) la trombon (**Victor**)].
(Victor.NOM) try.PRES.3SG (Victor.nom) [PRES sing.3SG (Victor.NOM) la trombon (Victor.NOM)]
'Victor is trying to play the trombone.'

  If the embedded subject can appear as an exact copy of the shared argument, then the appropriate analysis is compatible with a movement analysis. If it cannot, then agreement is the appropriate analysis.

- **Partial control.**

  Is there evidence of partial control? If there is no evidence of partial control, then the appropriate analysis is movement. If there is evidence of partial control then agreement is the appropriate analysis.

6.2.1 OC in Persian is not *pro*

The EC subject in OC contexts in Persian is not *pro*. First, this EC subject may not have an overt disjoint NP realization, as shown in (108a) and (108b), nor may the embedded verb show disjoint agreement marking, as shown in (109a) and (109b)).

(108) a. *ʒian mi-tun-e [(ke) aeskan be-r-e]
Jian DUR-be.able-3SG [(COMP) Ashkan SUB-go-3SG]
'*Jian is able Ashkan to go.'

b. *ʒian yad-eʃ=raeft [(ke) aeskan be-r-e]
Jian memory-3SG.CL=go.PST.3SG [(COMP) Ashkan SUB-go-3SG]
'*Jian forgot Ashkan to go.'

(109) a. *(maen) mi-tun-aem [(ke) EC bi-ad].
(I) DUR-be.able-1SG [(COMP) PRO SUB-come.3SG]
'*I am able him to come'

b. *(maen) yad-aem-taeft [(ke) EC bi-ad].
(I) memory-1SG-CL.PST.3SG [(COMP) PRO SUB-come.3SG]
'*I forgot him to come.'

Further, examples (110) and (111) show that even when the embedded subject is prosodically emphasized, an exact copy is not an acceptable alternate for the EC subject.

(110) a. *ʒian$_i$ mi-tun-e [(ke) ʒian$_i$ be-r-e]
Jian$_i$ DUR-be.able-3SG [(COMP) Jian$_i$ SUB-go-3SG]
'*Jian is able Jian to go.'

b. ʒian$_i$ mi-tun-e [(ke) PRO$_i$ xodesh$_i$ be-r-e]
Jian$_i$ DUR-be.able-3SG [(COMP) PRO$_i$ self$_i$ SUB-go-3SG]
'Jian is able himself to go.'

(111) a. *ʒian$_i$ yad-eʃ=raeft [(ke) ʒian$_i$ be-r-e]
Jian$_i$ memory-3SG.CL=go.PST.3SG [(COMP) Jian$_i$ SUB-go-3SG]
'*Jian forgot Jian to go.'

b. ʒian$_i$ yad-eʃ=raeft [(ke) PRO$_i$ xodesh$_i$ be-r-e]
Jian$_i$ memory-3SG.CL=go.PST.3SG [(COMP) PRO$_i$ self$_i$ SUB-go-3SG]
'Jian forgot himself to go.'

The analysis of *xodesh*, in (110) and (111), is some what unclear (see for more discussion (Ghomeshi, 1996 and Anagnastopoulou 2013)). However, the data above suggests that the correct analysis of *xodesh* should be an anaphor. Although, a full copy of the subject is ungrammatical in both (110a) and (111a), *xodesh* is grammatical in both environments, as shown in examples (110b) and (111b). Therefore *xodesh* cannot be a pronominal, because if it were, it should be able to alternate with an overt referential NP like *Jian*. This is not the case and thus *xodesh* must be analyzed as anaphor. This directly eliminates an analysis of the EC subject of clauses embedded under OC verbs as *pro*, because *pro* must be able to alternate with an overt referential NP or pronoun. Examples (110) and (111) show that neither an overt referential NP or pronoun is grammatical even if they are semantically linked to the matrix clause controller.

6.3 Can Movement Account for Persian OC?

The analysis via movement treats the embedded subject just like the raising subject in English. The subject is merged into the syntactic structure in the embedded clause as the spec of vP. It then moves to the spec of TP to satisfy agreement features on T and then is attracted to the matrix clause to satisfy agreement features on vP and T. This derivation in Persian would follow the path in example (112).

(112) a. Məzen mi-tun-e [(ke) ʃatraŋ bosi be-kun-e].
Məzen DUR-able-3SG [(COMP) chess play SUB-do-3SG]
'Məzen can/is able to play chess.'

b. $[_{CP}$ $[_{TP}$ Məzen $[_{vP}$ ~~Məzen~~ $[_{VP}$ mi-tun-e]]]$[_{CP}$(ke)$[_{TP}$ ~~Məzen~~$[_{vP}$~~Məzen~~ $[_{VP}$ʃatraŋ bosi be-kun-e]]]]].

The following discussion will show that the analysis via movement does not work for Persian OC complements.

- **Phasal status**. Given that the NP shared between the embedded clause and the matrix clause in OC contexts is an argument, we expect movement out of the embedded clause to be possible only if the Persian embedded clause is non-phasal (i.e. has a reduced structure). If the embedded clause in Persian turns out to be a phase, then agreement is the appropriate analysis.

  The first property of Phases is that they are headed by overt lexical complementizers. As Darzi (2008) argues, Persian OC contexts are headed by an (optionally) overt lexical complementizer, as in example (113).

  (113) Məzen mi-tun-e [**ke** PRO ʃatraŋ bosi be-kun-e].
  Məzen DUR-able-3SG [**COMP** PRO chess play SUB-do-3SG]
  'Məzen can/is able to play chess.'

  As I outlined in 6.1.1, the particle *ke* must be analyzed as an overt complementizer.

  The second property of phases is that they must be temporally saturated, meaning they allow for independent tense, and they must also be phi complete. The embedded clauses in Persian OC contexts are not temporally independent, but they are phi complete, as seen in example (114).

  (114) *Məzen diruz yad-eʃ=raeft [(ke) PRO faerda ovos be-xun-e]
  Məzen yesterday memory-3SG.CL=go.PST.3SG [(ke) PRO tomorrow song SUB.sing.3SG]
  '*Yesterday Məzen forgot to sing the song tomorrow.

  The independence of the tense is tested in example (114) by the overt temporal adverb *yesterday* in the matrix clause and *tomorrow* in the embedded clause. The presence of these two adverbs renders the utterance ungrammatical. Therefore the embedded clause must be dependent on the matrix clause

for its semantic tense. The embedded clause in example (114) is phi complete, marked with person and number.

- **Overtness of the shared argument.** Persian does not allow the shared subject to occur overtly anywhere other than in the subject position of the matrix clause. Compare the example (115) from Romanian, that does allow the the subject this freedom, to the examples in (116) from Persian, where the shared argument is only grammatical in the matrix clause.

(115) (**Victor**) incearca (**Victor**) [sa cinte (**Victor**) la trombon
(Victor.NOM) try.PRES.3sg (Victor.NOM) [SUB sing.3SG (Victor.NOM) la trombon
(**Victor**)].
(Victor.NOM)].
'Victor is trying to play the trombone.'

(116) a. Məzen mi-tun-e [(ke) PRO ʃatraŋ bosi be-kun-e].
Məzen DUR-able-3SG [(COMP) PRO chess play SUB-do-3SG]
'Məzen can/is able to play chess.'

b. mi-tun-e Məzen [(ke) PRO ʃatraŋ bosi be-kun-e].
DUR-able-3SG Məzen [(COMP) PRO chess play SUB-do-3SG]
'*Məzen can/is able to play chess.'

'Məzen can/is able to play chess?'

c. *mi-tun-e [(ke) Məzen ʃatraŋ bosi be-kun-e].
*DUR-able-3SG [(COMP) Məzen chess play SUB-do-3SG]
'Məzen can/is able to play chess.'

d. mi-tun-e [(ke) ʃatraŋ bosi be-kun-e Məzen]
DUR-able-3SG [(COMP) PRO chess play SUB-do-3SG Məzen]
'*Məzen can/is able to play chess.'

In example (115), from Romanian, the shared subject may appear in any one of the positions in bold, however it may only appear in one of those positions at a time. This contrasts with the examples in (116). In example (116a) the subject is grammatical in the specifier of TP in the matrix clause. In example (116b) the subject is ungrammatical in the specifier of vP of the matrix clause, though it is grammatical with an interrogative reading. The subject in the the specifier of TP in example (116c) is also ungrammatical and there is no interrogative interpretation. And finally, in example (116d) the subject in the specifier of vP in the embedded clause is ungrammatical.

- **Partial control.** Persian displays partial control, as defined as a semantically plural EC subject controlled by a semantically singular NP, as shown in example (117).

(117) ʒian$_i$ mi-tun-e [(ke) PRO$_{i+}$ be-meeting bi-ad].
Jian$_i$ DUR-able-3SG [(COMP) PRO$_{i+}$ be-meeting SUB.come-3SG]
'Jian is able to meet.'

In example (117) The matrix subject *Jian* controls the embedded subject, however the embedded subject is semantically plural. There are several people meeting and *Jian* is able to be one of them. Crucially, although there is singular marking on the embedded verb (*bi-ad*), the subject of the embedded clause is still interpreted as plural.

We can see that Persian OC contexts are phasal embedded clauses, as they do not allow the overt copy of the shared subject to appear anywhere other than the specifier of the matrix TP. Finally, there are partial control contexts in Persian. Therefore, a movement analysis of Persian OC is not appropriate.

Further, one of the motivating factors for a movement analysis is to reduce the mechanism for OC to the same mechanism that produces raising. However, as discussed in Ghomeshi (2001), Persian lacks subject-to-subject raising and instead seems to exhibit scrambling or left-ward extraposition in structures similar to English raising constructions. Examples (118a) introduces one potential raising verb and examples (118b) and (118c) illustrate the behaviour of this verb.

(118) a. lazem=budaen
necessary=be
'to be necessary' [1]

b. lazem-e [(ke) beacc-ha be-r-ean meadresse].
necessary-3SG [(COMP) child-PL SUB-go-3PL school]
'It is necessary that the children go to school.'

c. beacce-ha lazeam-e/*eand [(ke) be-r-aen maedresse].
child-PL necessary-3SG/*3PL [(COMP) SUB-go-3PL school].
'It is necessary that the children go to school.'

In example (118b) the verb *lazem=budaen* selects an embedded clause with an overt subject. In example (118c) the subject moves to what could be the subject position (the specifier of TP) of the matrix clause. However the evidence suggests that the subject does not actually move to the specifier of TP. First, the matrix clause verb is always marked with the 3rd person singular agreement. When the embedded subject is plural and moves to the intial position in the string, the agreement on the matrix verb does not change. Ghomeshi (1996) argues that Persian does not have an expletive pronoun, but has 'expletive agreement' (3rd person). Second, the subject of the embedded clause may appear overtly in the embedded clause as well as in the matrix clause (though only one overt copy is permissible at a time). Movement of the embedded

[1] The verb *to seem* (a more traditional raising verb) exhibits the same behaviour as *to be necessary*

subjects is therefore not obligatory, as it is in English. Finally, subjects are not the only constituents that can move into the matrix clause, as seen in examples (119a–119d).

(119) a. momken=nist [ke haemsaye-ha-ye ma jom'e saeb-ha xune be-mun-ae].
possible=neg.be.3SG [COMP neighbour-PL-EX us Friday night-PL home SUB-stay-3PL]
'Its impossible for our neighbours to stay at home Friday nights.'

b. [haemsaye-ha-ye ma] momken=nist [ke jom'e saeb-ha xune be-mun-ae].
[neighbour-PL-EX us] possible=neg.be.3SG [COMP Friday night-PL home SUB-stay-3PL]
'Its impossible for our neighbours to stay at home Friday nights.'

c. [jom'e saeb-ha] momken=nist [ke haemsaye-ha-ye ma xune be-mun-ae].
Friday night-PL possible=neg.be.3SG [COMP neighbour-PL-EX us home SUB-stay-3PL]
'Its impossible for our neighbours to stay at home Friday nights.'

d. [xune] momken=nist [ke haemsaye-ha-ye ma jom'e saeb-ha be-mun-ae].
home possible=neg.be.3SG [COMP neighbour-PL-EX us Friday night-PL SUB-stay-3PL]
'Its impossible for our neighbours to stay at home Friday nights.'

In examples (119b–119d) the embedded subject, embedded direct object, and embedded temporal adjective each can be extra-posed above the matrix clause verb. This suggests the process these elements under go is not raising, but extraposition to a target position in the CP domain, possibly focus. I will conclude that Persian OC contexts cannot be subsumed under the same analysis as Persian raising.

6.4 An Account of Persian OC based on Agree

Recall that the Agree model of OC functions by establishing an agreement relationship between a matrix clause argument and the embedded clause subject. The embedded subject is assumed to be a PRO, which establishes an antecedent relationship with a matrix NP. Recall also that Tense heads that are marked as [+ T], [+Agr] are not able to license PRO, as this feature combination is not compatible with a [– R] feature on PRO. Persian T heads in embedded OC contexts are [– T] (given that they are not able to license their own independent temporal adverb), and [+ Agr] (given that the embedded verb is subjunctive and inflected for phi features). This specific feature content of the embedded T has consequences for the way in which control configurations are derived in Persian.

I will start with exhaustive control, illustrated in (120).

(120) **Exhaustive Control**

Məzen mi-tun-e [(ke) PRO bosi be-kun-e].
Məzen DUR-able-3SG [(COMP) PRO play SUB-do-3SG]

'Məzen can/is able to play.'

The most important question to determine, is the precise mechanism of control, i.e. the way in which PRO comes to be coindexed with a matrix NP.

(121) Sbj_i T_i . . . [PRO_i T . . .]

In English, PRO is coindexed with the matrix subject because the matrix T serves double purpose: to check the phi features on the matrix subject and to check the phi features on the embedded subject as well. The relation between PRO and the matrix subject is thus mediated by the matrix T.

This analysis cannot be extended to Persian, however, because in Persian, the phi features of PRO are checked by the embedded T, which, unlike the English embedded T, has [+Agr] features (being subjunctive). In Persian, therefore, the matrix subject and PRO check their phi features from independent probes. The question is how PRO and the matrix subject come to share the same index in Persian? I propose that this is because the tense feature on the embedded T head is anaphoric, and therefore the embedded T must be coindexed with the matrix T head. Given this coindexation between the two T heads, the two subjects will derivatively be coindexed, as well.

(122) Sbj_i T_i . . . [PRO_i T_i . . .]

Partial control in Persian will follow a similar derivational path. The analysis proposed in Landau (2013) for partial control relied on the intuition that the relationship between the matrix T and PRO is mediated by the embedded C head, which also potentially bears phi features.

(123) Sbj_i T_i . . . [C_i PRO_i T . . .]

The C head allows the transmission of the phi features from the matrix T head to PRO, but its presence blocks the transmission of the [MER] (mereological) feature. This allows the matrix subject and PRO to bear independent [MER] features and partial control obtains.

This analysis cannot be extended as such to Persian, as the matrix T head does not enter an agree relation with PRO. PRO checks its phi features against the embedded T, not against the matrix T head. However, I proposed above that in order to account for exhaustive control, an agree relation obtains in Persian between the matrix T head and the embedded T head. Landau's analysis for partial control can thus be implemented in a similar way in Persian, but with C blocking the relation between the two T heads, rather than the relation between the matrix T and PRO.

(124) Sbj_i T_i . . . [C_i PRO_i T_i . . .]

The relevant feature of C, that intervenes and blocks the agree relation between the matrix T and the embedded T, is the tense feature on C, rather than its phi features. Otherwise, the account for partial

control is the same in Persian, as in English. The C head mediates the relation between the two T heads, but blocks the transmission of the [MER] feature (given that C does not bear a [MER] feature itself. The two T heads can thus have independent [MER] features, and indirectly, the matrix subject and PRO will be allowed to have independent [MER] features.

6.4.3 Persian OC signature

Can the Persian data be subsumed under the OC signature definition? The components of the OC signature are reproduced below in example (125).

(125) *The OC Signature*

In a control construction [....X_i....[$_s$$PRO_i$....]....], where X controls the PRO subject of the clause S:

a. The controller(s) X must be (a) codependent(s) of S.

b. PRO (or part of it) must be interpreted as a bound variable. *(Landau, 2013)*

In Persian OC contexts a controller from the matrix clause (X) must control PRO in an embedded clause. The controller must be a codependent of the clause which contains PRO, just as in English. Finally, the Persian PRO must also be interpreted as a bound variable. Consider the following examples (126) and (127).

(126) Fahrt Ashkan yad-eʃ=raeft [PRO dars be-de].
Only Ashkan memory-3SG.CL=go.PST.3SG [PRO subject SUB-give]
'Only Ashkan forgot to teach a class.'

(127) Fahrt Ashkan yad-eʃ=raeft [ke un boiad dars be-de].
Only Ashkan memory-3SG.CL=go.PST.3SG [that he must subject SUB-give]
'Only Ashkan$_i$ forget that he$_{i/j}$ should teach.'

The EC subject, PRO, gives rise to the bound variable reading: Ashkan = only x [x forgot x to teach]. The overt pronoun, on the other hand, gives rise to both the bound variable reading and the strict reading: Ashkan = only x [x forgot Ashkan to teach].

Chapter 7

Conclusion

In this thesis I discussed Persian Obligatory Control contexts. The following properties were identified for these contexts:

- Clauses embedded under OC verbs in Persian are **subjunctive clauses**

- **Clauses embedded under OC verbs in Persian are full CPs.** This is against Ghomeshi (2001) who argues that the embedded subjunctives should be analyzed as vP. Two arguments were discussed for the CP status of these clauses: (i) these clauses are optionally introduced by a complementizer (*ke*), as shown in (128) below, and (ii) the T head of the subjunctive embedded clauses is phi complete, as shown by the agreement morphology on the embedded verb in example (130) and (129) below.

(128) (mæn) mi-tun-æm [(ke) EC be-r-æm].
(I) DUR-be.able-1SG [(COMP) PRO SUB-go-1SG]
'I am able to go'

(129) (mæn) yad-æm-tæft [(ke) EC be-r-**æm**].
(I) memory-1SG-CL.PST.3SG [(COMP) EC SUB-go-**1sg**]
'I forgot to go.'

- **Anaphoric Tense:** Persian subjunctive clauses embedded under OC verbs are not independent temporally. Their T must be interpreted as dependent on the matrix T. This is shown in (130).

(130) *Məzen diruz yad-eʃ=raeft [(ke) faerda PRO ovos be-xun-e]
Məzen yesterday memory-3SG.CL=go.PST.3SG [(ke) PRO tomorrow song SUB.sing.3SG]
'*Yesterday Məzen forgot to sing the song tomorrow.

- **Exhaustive and Partial Control:** Persian OC contexts exhibit both exhaustive control, as in (131), and partial control, as in (132).

(131) Mzen$_i$ mi-tun-e [(ke) EC$_i$ ʃatraŋ bosi be-kun-e].
Mzen DUR-able-3SG [(COMP) EC chess play SUB-do-3SG]
'Məzen can/is able to play chess.'

(132) ʒian$_i$ mi-tun-e [(ke) EC$_{i+}$ be-meeting bi-ad].
Jian$_i$ DUR.able.3SG [(COMP) RC$_{i+}$ be-meeting SUB.come-3SG]
'Jian is able to meet.'

- **Overt embedded subjects are ungrammatical.** Overt subjects are not grammatical in subjunctive clauses embedded under OC verbs in Persian, regardless of whether their reference is shared with a matrix NP (133) or they have disjoint reference (134).

(133) *ʒian$_i$ yad-eʃ=raeft [(ke) ʒian$_i$ be-r-e]
Jian$_i$ memory-3SG.CL=go.PST.3SG [(COMP) Jian$_i$ SUB-go-3SG]
'*Jian forgot Jian to go.'

(134) *ʒian yad-eʃ=raeft [(ke) aeskan be-r-e]
Jian memory-3SG.CL=go.PST.3SG [(COMP) Ashkan SUB-go-3SG]
'*Jian forgot Ashkan to go.'

While the unavailability of overt subjects with disjoint reference is to be expected under an analysis, which posits PRO as the subject of the embedded clauses of OC verbs, the unavailability of overt subjects with shared reference is unexpected. According to Landau (2013) overt subjects in embedded OC contexts are expected to be grammatical, as shown in (135), on condition their reference is shared or controlled by a matrix NP. Persian does not allow this type of overt subject, as in example (133) repeated below as (136).

(135) *pro* decidiu ir **so ele/ *so o Joao** ao mercado.
pro decided to.go **only he/ only the John** to.the market
i. 'He/John is the only one who decided to go to the market.'
ii. 'he decided for it to be the case that only that only he/*John goes to the market. (Portuguese) (Barbosa, 2009 in Landau, 2013)

(136) *ʒian$_i$ yad-eʃ=raeft [(ke) ʒian$_i$ be-r-e]
Jian$_i$ memory-3SG.CL=go.PST.3SG [(COMP) Jian$_i$ SUB-go-3SG]
'*Jian forgot Jian to go.'

The reason overt embedded subjects in Persian will never be licensed in OC contexts, may be due to the fact that T is phi complete, but its Tense feature is anaphoric. Following Pesetsky and Torrego

(2000) nominative case is a valued [uT] feature on NP subjects. The [uT] feature on the subject NP may not be valued or checked by a defective (or anaphoric) T. Therefore Persian T cannot license Case on an overt subject and Persian OC contexts will never license overt subjects.

Several views were considered for an analysis of Persian OC contexts (various reductionist analyses, including a movement analysis, as well as non-reductionist analyses). A movement analysis for these contexts was ruled out for the following reasons:

- The Persian OC contexts are phasal: these contexts exhibit an overt complementizer, *ke*; although the embedded OC clauses are saturated temporally, relying on the matrix clause for their tense interpretation, these clauses are phi complete, exhibiting overt person and number agreement morphology on the verb;
- Persian OC contexts do not allow the shared argument to occur overtly in the embedded clause;
- Persian has instances of partial control contexts.

A movement analysis is unable to account for these properties, therefore an analysis based on agreement was posited for Persian OC.

The analysis I proposed for OC contexts in Persian posits a PRO as the subject of the subjunctive embedded clauses and has PRO coindexed with a matrix NP. I showed that the Persian OC contexts exhibit the characteristics outlined in the OC Signature. In particular, the structure of Persian OC contexts is [$X_{controller/i}$.....[PRO_i....]], and

1. The controller must be codependent of the embedded clause.
2. PRO is interpreted as a bound variable.

For cases of exhaustive control I proposed that PRO is coindexed with the matrix subject as a consequence of the fact that the embedded T head is anaphoric and dependent on the matrix T. Given that each of these T heads is in its turn coindexed with its respective subject, the two subjects end up coindexed.

For cases of partial control I proposed that the agree relation between the matrix T and embedded T is mediated by the C head, following Landau (2004). This 'mediation' is made possible by the fact that the C head bears a Tense feature. The C allows for the transmission of tense between the two Ts, but blocks the transmission of the [Mer] (Mereological) feature from the matrix T to the embedded T. This allows the two heads to have different valuations for that particular feature and gives rise to partial control.

Bibliography

Alboiu, Gabriela (2007). Moving forward with romanian backward control and raising. In *New horizons in the analysis of control and raising*, pp. 187–211. Springer.

Anagnastopoulou, E and Everaert, Martin (2013). Identifying anaphoric dependencies. *Diagnosing Syntax*.

Boeckx, Cedric and Hornstein, Norbert (2003). Reply to "control is not movement". *Linguistic inquiry*, **34**(2), 269–280.

Boeckx, Cedric and Hornstein, Norbert (2004). Movement under control. *Linguistic inquiry*, **35**(3), 431–452.

Boeckx, Cedric and Hornstein, Norbert (2006). The virtues of control as movement. *Syntax*, **9**(2), 118–130.

Boeckx, Cedric, Hornstein, Norbert, and Nunes, Jairo (2010). *Control as movement*. Volume 126. Cambridge University Press.

Chomsky, Noam (1981*a*). A note on non-control pro. *Journal of Linguistic Research*, **1**(4), 1–11.

Chomsky, Noam (1981*b*). Principles and parameters in syntactic theory. *Explanation in linguistics: The logical problem of language acquisition*, **32**, 75.

Chomsky, Noam (1993). *Lectures on government and binding: The Pisa lectures*. Number 9. Walter de Gruyter.

Chomsky, Noam and Lasnik, Howard (1977). Filters and control. *Linguistic inquiry*, **8**(3), 425–504.

Darzi, Ali (2001). Nonfinite control in persian.

Darzi, Ali (2008). On the vp analysis of persian finite control constructions. *Linguistic Inquiry*, **39**(1), 103–116.

Darzi, Ali and Motavallian, Rezvan (2010). The minimal distance principle and obligatory control in persian. *Language Sciences*, **32**(4), 488–504.

Ghomeshi, Jila (2001). Control and thematic agreement. *The Canadian journal of linguistics*, **46**(1-2), 9–40.

Ghomeshi, Jila and Ritter, Elizabeth (1996). Binding, possessives, and the structure of dp. In *PROCEEDINGS-NELS*, Volume 26, pp. 87–100. UNIVERSITY OF MASSACHUSETTS.

Haegeman, Liliane (1997). *Elements of grammar*. Springer.

Hashemipour, Margaret Marie (1989). *Pronominalization and control in modern Persian*. University of California, San Diego.

Hornstein, Norbert (1999). Movement and control. *Linguistic inquiry*, **30**(1), 69–96.

Hornstein, Norbert (2000*a*). *Move! A minimalist theory of construal*. Volume 10. Wiley-Blackwell.

Hornstein, Norbert (2000*b*). On a-chains: A reply to brody. *Syntax*, **3**(2), 129–143.

Hornstein, Norbert (2003). On control. *Minimalist syntax*, 6–81.

Hornstein, Norbert, Nunes, Jairo, and Grohmann, Kleanthes K (2005). *Understanding minimalism*. Cambridge University Press.

Ilkhanipour, Negin (2014). On the cp analysis of persian finite control constructions. *Linguistic Inquiry*, **45**(2), 323–331.

Karimi, Simin (2008). Raising and control in persian. *Aspects of Iranian linguistics*, 177–208.

Landau, Idan (1999). *Elements of control*. Ph.D. thesis, Massachusetts Institute of Technology.

Landau, Idan (2001). *Elements of control: Structure and meaning in infinitival constructions*. Volume 51. Springer Science & Business Media.

Landau, Idan (2004). The scale of finiteness and the calculus of control. *Natural Language & Linguistic Theory*, **22**(4), 811–877.

Landau, Idan (2006). Severing the distribution of pro from case. *Syntax*, **9**(2), 153–170.

Landau, Idan (2008). Two routes of control: Evidence from case transmission in russian. *Natural Language & Linguistic Theory*, **26**(4), 877–924.

Landau, Idan (2013). *Control in generative grammar: A research companion*. Cambridge University Press.

Landau, Idan (2015). *A two-tiered theory of control*. Volume 71. MIT Press.

Manzini, M Rita and Roussou, Anna (2000). A minimalist theory of a-movement and control. *Lingua*, **110**(6), 409–447.

Modesto, Marcello (2010). What brazilian portuguese says about control: remarks on boeckx & hornstein. *Syntax*, **13**(1), 78–96.

Pesetsky, D and Torrego, E (2000). T-to-c movement: Causes and consequences. kenstowicz (ed.) ken hale: A life in language.

Rizzi, Luigi (1997). The fine structure of the left periphery. In *Elements of grammar*, pp. 281–337. Springer.

Roussou, Anna (2001). Control and raising in and out of subjunctive complements. *Comparative syntax of Balkan languages*, 74–104.

Taleghani, Azita H (2008). *Modality, aspect and negation in Persian*. Volume 128. John Benjamins Publishing.

Wurmbrand, Susanne (1998). *Infinitives*. Ph.D. thesis, Massachusetts Institute of Technology.

Wurmbrand, Susi (2002). Syntactic versus semantic control. In *Studies in comparative Germanic syntax: Proceedings from the 15th workshop on comparative Germanic syntax*, Volume 53, p. 93. John Benjamins Publishing.

Wurmbrand, Susi (2004). Two types of restructuring—lexical vs. functional. *Lingua*, **114**(8), 991–1014.