

Chapter 1: Introduction

The Persian language has been known to contain an obligatory rule of subject-verb agreement, or “motabeghe-ye feʔl vā faʔel”. However, there are constructions in Persian in which the verb does not agree with the subject. This book investigates two major constructions that appear to induce such a constraint on the verbal agreement pattern in the Persian language.

The first construction contains an inanimate plural subject while the verb appears with the third person singular morphology. Adopting the framework of Distributed Morphology (DM) that has recently been used as a key to capturing several agreement restrictions in other languages, I propose that the restriction caused by animacy in Persian resides in post-syntactic morphology through an “Impoverishment” operation.

The second construction I study contains experiencers and psychological predicates. The nominative experiencer does not induce agreement on the verb and the verb appears in the third person singular form. Adopting the Minimalist Program of Chomsky (1995-2004), I argue that the lack of verbal agreement in Persian psychological constructions is only apparent. Previous studies have proposed that the psychological constructions under review involve compound verbs. However, I provide evidence to show that they do not involve compound verbs. I propose that these constructions have a Tense requirement and involve applied arguments. Furthermore, I propose a new category of strong phase that must be added to the set of strong phases proposed by Chomsky 1999-2004.

1.1. The Notion of Subject

The notion of subject has been extensively explored in the past decades. What is a subject? Do we need a subject in every sentence of every language? Does the notion of subject have fixed syntactic and semantic properties and attributes? How many subjects can a sentence have? Although the notion of subject seems rather straightforward at first glance, linguists from different theoretical frameworks have controversial debates around the notion of “Subjecthood.” This work is an attempt to explore the idea found in linguistic literature that there are no postulated standard criteria for subject and subjecthood, and it is not possible to provide a specific definition or description for subjecthood (Sigurðsson 1992-2004, Landau 2003, Harley 1995, and the references therein). Sigurðsson (2001) suggests that the notion of subject and consequently the Extended Projection Principle (EPP), which states that each clause must have a subject (Chomsky 1981-1995), should be redefined and reduced (which he has done in his more recent papers). However, it is evident that there are one or more NPs in every clause of a language fulfilling certain conditions and relations that are in the “range” of subjecthood (Boeckx 2000, Ura 2000, Sigurðsson 2001, Shibatani & Pardeshi 2001, among others). This book provides further support for the statements below.

“... the set of properties that characterize a canonical subject in a highly “subject prominent” language cannot be linked to a single position- it is simply not the case that there is a one to one mapping of “subject” properties to some universal syntactic position. Even within a given language, many constructions exist in which properties usually attributed to a canonical subject seem to be scattered between two or three NPs.”
(Harley 1995: 14)

“The past 20 years have seen the gradual deconstruction of the notion “subject” (McCloskey 1997, Sigurðsson 2000). Subjecthood is no longer viewed as a package deal; rather, particular subject properties are distributed over separate dimensions (structural positions, case, agreement, EPP, thematic prominence, topicality, etc.).”
(Landau 2003:80)

1.2. Central Problems and Solutions

This book studies two major constructions that pose problems for verbal agreement in Persian: (a) the constructions with inanimate subjects; and (b) the psychological constructions. These two constructions seem problematic because they appear to exhibit the subject-verb agreement restrictions in Persian. It is a common belief that in Standard Persian, verbs agree in Number and Person with the structural subject (Khanlari 1980, Meshkat-al dini 1987, among others). In current minimalist terms, it appears that only Person and Number are among the set of Φ -features that need to be valued in the sense of Chomsky (2001). This is illustrated in (1a-b).

- (1) a. *mæn* *to* *ra* *did-æm*
 I you Acc. saw-1sg
 I saw you
- b. **Farzin væ** **madær-æsh** *be* *Iran* **ræft-ænd**
 Farzin and mother-3sg to Iran went-3pl
 Farzin and his mother went to Iran

In (1.a), the verb *did-æm* contains the 1st person singular morphology that is in agreement with the subject *mæn* (I). The verb in (1.b) *ræft-ænd* agrees with the subject *Farzin væ madær-æsh* (Farzin and his mother) and appears with the 3rd person plural morphology.

However, in the two constructions under study, the subject-verb agreement does not seem to be present and the agreement between the verb and the subject seems to be restricted. A description of the two problematic constructions is provided in 1.2.1, and 1.2.2.

1.2.1. Plural Inanimate Subjects

In (2), the inanimate subject, *in shaye?e-ha* (these rumours), is in the 3rd person plural while the verb *ni-st* (is not) appears in the 3rd person singular and therefore, does not agree with the subject in number.

- (2) In **shaye?e-ha** dorost ni-st
 This rumour-**pl** correct neg-is
 These rumours are not true/correct.

Following a recent body of literature that argues that verbal agreement restrictions do not reside in syntax and provides post-syntactic accounts for such restrictions (Boeckx 2000, Rivero 2004), I propose that the verbal agreement restriction for plural inanimate subjects in Persian does not reside in the Syntax. I suggest a post-syntactic morphological treatment within a Distributed Morphology in which there is a feature mismatch between Number and Animacy resulting in the deletion or impoverishment of the Number feature.

The second construction exhibiting the verbal agreement restriction in Persian involves the psychological predicates.

1.2.2. Psychological Constructions

Experiencers in subject position in the constructions with psychological verbs do not trigger agreement on the verb. For instance, consider the contrast in agreement in (3) and (4)

- (3) **ma** u ra [dust dar-**im**]
 we her/his Acc. friend have-**1pl**
 We like her/him
- (4) **ma** æz u xoš-eman amæd-**Ø**
 we from her/his pleasure-1pl came-**3sg**
 We liked her/him (She/he appealed to us)

Example (3) does not involve any complications; it has a nominative subject *ma* (we) and an accusative object *u* (him/her), and the compound verb *dust dar-im* (friend have-1pl) agrees with the subject. By way of contrast, constructions like example (4) are interesting because they seem to violate the obligatory rule of the subject predicate agreement; the verb *amæd-Ø* (came) is in the 3rd person singular, while the experiencer in the subject position, *ma*, is in the 1st person plural. Constructions such as (4) have been introduced as “Impersonal” or “Subjectless” in the literature, and have not been systematically studied as psychological constructions. Moreover, the properties of the experiencer in constructions such as (4) have never been entirely captured in the previous literature.

I capture the properties of the Persian psychological constructions by proposing that they involve applied arguments. Marantz (1984) was the first to argue that only objects are true arguments of the verb, licensed within the domain of the VP, and that subjects are not in such close connection to the verb. Marantz (1984) argued that unlike objects, subjects are added and therefore ‘extra’ or ‘non-core’ participants in the events described by the verb. He argues that the meaning of the verb (the interpretation of the verb phrase) depends on the choice of the object while the external argument only depends on the structure and does not

affect the meaning of the verb.¹ Kratzer (1996) developed Marantz's insight into a theory of Voice, a syntactic head responsible for licensing the external argument: the external argument is projected as the specifier of Voice, which takes the verbal phrase as its complement. In addition to the objects and the [SpecVoice] subjects, there exists a third kind of argument, the *Applied Argument*, with properties similar to both subjects and objects. An applicative head (Pykkänen 2001, McGinnis 2001, and Cuervo 2003, among others) licenses an applied non-core argument in its specifier and relates it to the category it takes as a complement. Previous studies (Pykkänen 2002, McGinnis 2001) argue that Applicative heads are divided into two kinds: the High Applicative heads which take a vP (an event) complement, and the low applicative heads which take a DP (an individual) complement.

Based on the data from the Persian psychological verbs discussed in this study, I depart from Pykkänen (2002) and McGinnis (2001) who argue that applicative heads can take only a vP or a DP as complement. I propose a new category of Applicative head, the *Super High Applicative* head, which takes a TP (a sentential predication/full proposition) as complement.² In addition, I propose that the *Super High Applicative* projection (SupHighAppLP) is a strong phase with an *EPP/peripheral* feature, which assigns the thematic role of experiencer to the applied argument projected in its specifier. This new category the (SupHighAppLP) is thus to be added to the set of strong phases (CP, vP, and possibly DP) proposed by Chomsky (1999-2004).

The verbal agreement and subject properties have been tied to the nominative licensing in the literature (Chomsky 1982-1995, George & Kornfilt 1981). However, the Persian psychological constructions and the constructions with inanimate subjects discussed in this work provide evidence for the divorce of the nominative licensing and the verbal agreement. Following Haeberli 2002, Pesetsky and Torrego 2001, 2004, and Svenonius 2001, (among others), I propose that Tense is responsible for the nominative licensing.³

The Persian constructions I study in this book provide substantial evidence for the claim that grammatical subjects exhibit distinct behaviours across languages (Boeckx 2000, Sigurðsson 2001, Harley 1995, Landau 2003, Ura 2001, among others,) and that it is not possible to include all the configurations of subject in one definition. Instead, the subject properties (case, agreement, EPP, thematic prominence) are distributed in the "dimensions" (Landau 2003.) Consequently, a hierarchy involving degrees of subjecthood must be established and the concepts of subject and the EPP should be reduced largely (Sigurðsson 2001, Sedighi 2001, and Boeckx 2000, among others.)

1.3. The Theoretical Frameworks

I adopt the Minimalist Program of Chomsky (1995- 2004), which is discussed in Section 1.3.1., as the theoretical framework. The Minimalist Program aims to describe the language faculty in a minimalist way by introducing principles of economy and minimizing/simplifying

¹ This fact might be another reason for the distinct behaviour of subjects.

² See Rivero 2004 for a similar idea on the reflexive clitic *feel-like* constructions in south Slavic/Albanian.

³ This idea is a return to GB in which Tense and not Agreement relates to Nominative case marking.

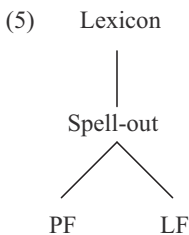
earlier theoretical frameworks. In particular, the Strong Minimalist Thesis states that economy is at the heart of the computational system/language faculty.

In addition to the Minimalist Framework, I adopt Distributed Morphology (DM) (Halle and Marantz 1993, 1994, Marantz 1997, Halle 1997) and assume that the morphological operations and processes can be applied at various points of pre-syntactic, syntactic, and post-syntactic derivations. Some basic concepts of Distributed Morphology are introduced in Section 1.3.2.

1.3.1. The Minimalist Program

Chomsky's Minimalist Program (1992-2004) aims to establish the theory of grammar of human language by postulating only the minimal and necessary assumptions that are required for capturing the conceptual grounds. The Program provides the means to approach "Plato's" problem (Chomsky 1986, 1991), namely the problem of how language is acquired given the limited evidence for its structure. The Minimalist Program tries to establish that there is ultimately 'one' set of universal 'principles' and a finite array of options as to how these principles apply, or 'parameters.' The Minimalist Program has reformulated several assumptions of Generative Grammar as encapsulated in the model known as the Government and Binding Theory. It has also eliminated several unnecessary steps of previous theories such as the deep and surface structure. Some of the basic assumptions of the Minimalist Program are outlined below.

The structure of the grammar shown in (5) consists of lexicon and numeration in which the selected items form the lexicon. The lexicon is defined as a set of LF, PF pairs.⁴ By contrast with the previous stages of the theory, the only levels that remain in Minimalism are the interface levels: PF (the articulatory conceptual level) and LF (the conceptual intentional), since they are the only levels visible to the computation.



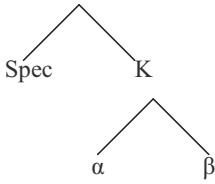
The computational system consists of three recursive operations that build structure: *Merge*, *Move*, and *Agree*.

Merge is the simplest operation. It builds up larger structures from smaller ones. *Merge* takes a pair of syntactic objects (α and β) and forms a more complex one (K). *Merge*

⁴ In his most recent work, Chomsky 2004, it is not clear that LF and PF exist in the strict sense of the term as representation levels. They are accessed or spelled out in chunks as the derivation unfolds/proceeds.

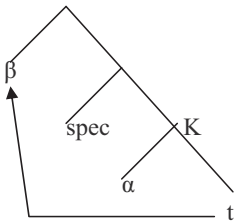
combines a Head (the syntactic object that selects in any merge operation) with a complement. This combined unit is then merged with [Spec] (second merge). Merge corresponds to the *external merge* in “Beyond Explanatory Adequacy,” Chomsky 2001.

(6)



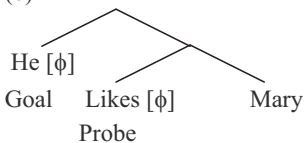
Move takes a structure formed by merge and moves one of the elements of that structure into a c-commanding position. Move is called *internal merge* in “Beyond Explanatory Adequacy,” Chomsky (2001). The structure of Move is given in (7).

(7)



Features are either *interpretable* or *uninterpretable*. Interpretable features have a semantic content or have an effect on semantic interpretation. Uninterpretable features are semantically void of content and must be removed before spell-out. Syntactic operations apply in order to eliminate uninterpretable features. Uninterpretable features must be checked/valued; once checked, they can be deleted (however, they are still accessible unless erased.) In order for features to be checked/valued, they need to match first and then *Agree*. Agree represents an asymmetric relation between a Probe (which has uninterpretable features) and a Goal (which has interpretable features). If features of Probe and Goal match, then Agree eliminates the uninterpretable features on the Probe by ‘valuing’ them and removing them (spelling them out phonologically).

(8)



In (8), the ϕ -features of the verb are uninterpretable, while the ϕ -features of the (pro)nominal are interpretable. The ϕ -features of the verb enter the derivation unvalued. They need to be valued by the relevant ϕ -features of the (pro)nominal. According to the most recent

minimalist assumptions, there is no need for the Goal *He* to move to a specifier position and the Agree can occur while the Goal is in situ via the *Long distance Agree*. The movement of the pronominal to the specifier position in (8) is to satisfy the EPP requirement.

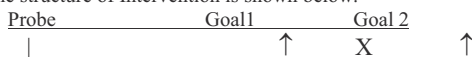
The notion of *EPP* was originally formulated by Chomsky (1981) as the ‘principle P’ defined as “the structural requirement that certain configurations ... must have subjects (Chomsky 1981: 27)”. Chomsky (1982) introduces the term *Extended Projection Principle*, since he believes that the requirement goes beyond the Projection Principle. In (8) the T has an EPP (*OCC*urrence in “Beyond Explanatory Adequacy” and the *Peripheral* in the most recent term) requirement and this requirement is satisfied by the movement of the pronominal *He* into the specifier position of T.

Based on data from Icelandic, which will be discussed in detail in Chapter 2, Chomsky revised his ideas on subjects in *Minimal Inquiries* (2000) and *Derivation by Phase* (2001), suggesting that the correlation between the nominative case and the NP-movement/EPP is only indirect. Among many other distinguished properties, Icelandic accommodates the non-nominative NPs into the [specTP]. Accordingly, Chomsky rejects the feature-based movement, replacing the feature checking via movement with a relation of the *Long distance Agree*: Agree and the elimination of uninterpretable features. He argues that the EPP has nothing to do with feature checking in the sense of Chomsky (1995). Rather, similar to his earliest view, certain functional heads require a specifier. Chomsky (2004) puts further weight on C and states that all operations are determined by the C phase, and T inherits from C. He argues that the Φ -features of C seek a goal that may be raised to the [SpecT]; moreover, the edge-feature of C can seek a goal (maybe a copy of the same goal) and raise it in parallel to the [SpecC].

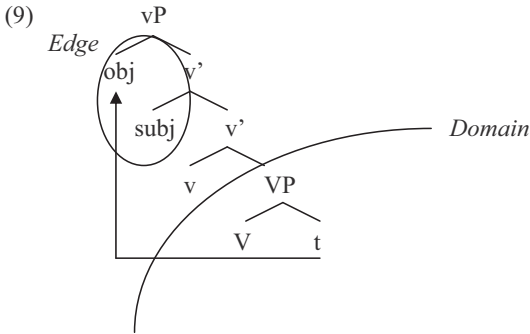
The *Locality* of Agree states that Agree holds between a feature F on X and a matching feature F on Y iff there is no intervening Z[F]. *Intervention* is when Z intervenes between X and Y in [X...Z...Y] while X c-commands Z and Z c-commands Y. Chomsky (2004) also states that in the operation of Agree the probe agrees with the goals in its domain as long as a goal has valued features which blocks further search (intervention).⁵

In this study, I follow the *Phase Theory* (Chomsky 1999, 2000) and assume that syntactic derivations undergo semantic and phonological interpretation in incremental chunks or *phases*. According to this theory, *Phases* are chunks of structure that are semantically and phonologically complete. Weak Phases are not complete propositional structures. Strong phases are phases whose heads may have an EPP/OCC/P feature, are propositional, and can be a target of movement. Strong phases are CP or vP (or possibly DP).⁶ Once a phase is complete, the movement and agreement operations can target a head and constituents in its *edge* (specifiers and adjuncts) but cannot target constituents in its *domain* (complement) as shown in (9).

⁵ The structure of Intervention is shown below.



⁶ vP is the Lexical phase, containing VP. CP is the Functional phase, containing TP.



(McGinnis 2001: 6)

In the structure above, the domain of the head (v) is not accessible to operations outside of the vP , but only the v and its edge (*Head impenetrability condition*). An element that needs to be raised has to move to the edge (either spec or adjunct to vP) of the phase. At every phase, the history of derivation within the previous phase is lost.

In phase theory, the derivation proceeds by phase and is cyclic. That is, the interpretation/valuation and spell-out take place at the phase level, and at the end of each phase the structure is sent to LF and PF for interpretation. Cyclic phases are embedded inside each other; therefore, the meaning of the whole depends on the embedded cycles.

Case licenses the function of an NP in a clause and is divided into two general types: the *Structural/Morphological* case, which is licensed by the functional projections (the usual Nom/subj, Acc/obj case); and the *Inherent/Lexical* case, which is associated with specific lexical items and their theta assignments. Lexical items, verbs, or event nominals assign the lexical case that is checked/valued at the same time that the theta roles are checked/valued (Chomsky 1995, Ura 1996). Examples such as (10) will be further discussed in Chapter 2.

- (10) Calvini liki verkid Icelandic
 Calvin.Dat. like the job.Nom.
 Calvin likes the job

(Harley, 1995:144)

A common assumption is that in (10) the argument in the subject position is the lexically/inherently marked dative while it structurally appears in a nominative (subject) position.

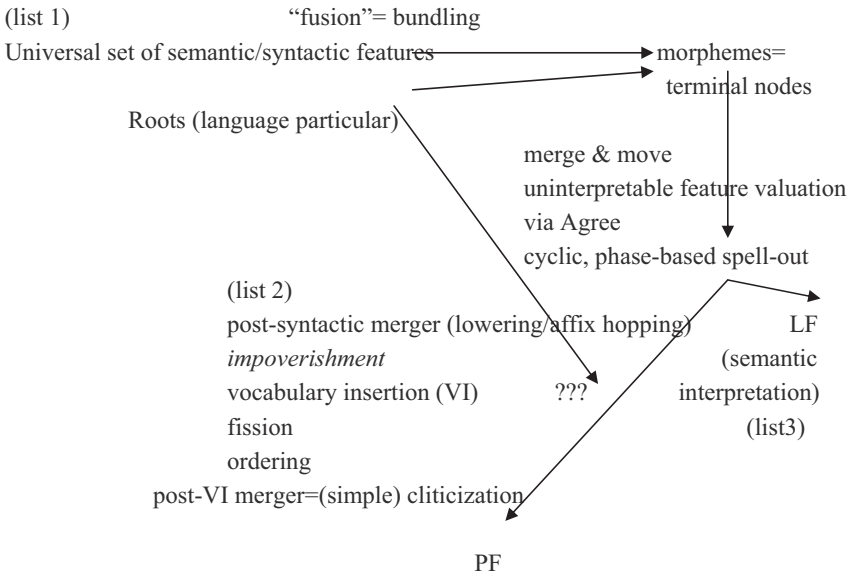
Having provided a background of the Minimalist Program, I will now discuss Distributed Morphology (DM) in which some of the basic assumptions of the Minimalist Program are preserved.

1.3.2. Distributed Morphology⁷

The framework of Distributed Morphology (Halle and Marantz 1993, 1994, Marantz 1997, Halle 1997) views the Morphological Component as a distinct level of grammatical representation, operating on its own principles and mediating between syntax and phonology. Unlike traditional morphology, the morphological principles and processes in DM may apply at various points of the syntactic derivation; therefore, the morphosyntactic information can be changed in the course of the derivation. Only after all morphological operations have been applied at the syntactic and post-syntactic level, morphological structure can insert the vocabulary items. The term *distributed* in DM then means that the job assigned to the Lexicon in earlier theories is *distributed* through various components.

The structure of Grammar according to DM is shown in (11).

(11) Distributed Morphology: The structure of the Grammar



As can be seen in (11) the structure of the grammar is of the classic Y-type, familiar from the Government and Binding framework and the Minimalist Program discussed in the previous section. The sub-set of universal features chosen by a language is combined or bundled into morphemes for the computational system. This is called *Fusion*. Fusion is the pre-syntactic bundling of features to create merge-able nodes for Syntax. These morphemes contain only the features relevant for the computational system of syntax and nothing else. An example of fusion would be the bundling of the Φ -features of the subject with the Tense features into a

⁷ The background information about DM and examples provided in this section, unless otherwise mentioned, is from Marantz's (2003) LSA Summer Institute course material.

single morpheme⁸. This happens, for instance, in Mezquital Otomí (Jensen 1990). In this language, it is not possible to isolate affixes meaning ‘I,’ ‘he,’ ‘you,’ ‘past,’ ‘present,’ and ‘future’. In other words, we cannot separate Person from Tense (12.a). On the other hand, in German, these two sets of features (Person and Tense) are separate (12.b).

- (12) a.

1 st	2 nd	3 rd	
present	<i>dí</i>	<i>gí</i>	<i>bi</i>
future	<i>ga</i>	<i>gi</i>	<i>da</i>
past	<i>dá</i>	<i>gá</i>	<i>bi</i>

Mezquital Otomí
- (Jensen, 1990:50)
- b.

sag-	te-	st	sag-	Null-	st
‘say’	past	2sg	‘say’	pres	2sg

German
- (Halle, 1997: 427)

In (12.a), the manifestation of Person and Number is not separable from Tense. These features are combined or fused together. However, in German (12.b) the set of Φ -features is represented in one morpheme *st* (2nd sg) separate from Tense features *te* (past).

Syntax then generates structures from Roots (language Particular) and the set of universal features by “merge” and “move” and the uninterpretable feature valuation via Agree (as discussed in the previous section, these operations result in the required movements and matching of features when necessary in a minimal way.) Spell-out delivers these syntactically generated structures to the morphophonology and to LF for interpretation in a “phase-based” cyclic way. This means that at the end of each phase the result of the computation is sent for post-syntactic operations as well as interpretation and a new cycle starts.

List (1) is the ‘narrow lexicon’ (Marantz, 1997), which provides the units that syntax operates with. It contains roots of the language as well as the bundles of features. Feature insertion is the insertion of purely morphological features such as case and agreement.

Vocabulary insertion (list 2) is the insertion of the most highly specified vocabulary item that fits. This means that, if English *-en* and *-s* both of which have a [+pl] feature and carry distinct phonological features are competing for insertion into a node from syntax, *-en* will win the competition in the environments of *childr* and *ox*. This is because *-en* carries the contextual feature that limits its insertion to a small set of roots (*childr*, *ox*) and is the more specified vocabulary item. DM comprises a number of morphological operations.

Fission originally proposed by Noyer (1992), is the situation when a single morpheme has not exhausted all its features and may correspond to more than one vocabulary item. Where fission occurs, vocabulary insertion does not stop after a single vocabulary item has been inserted. Rather, it accretes on the sister of the fissioned morpheme (or in a local domain) until all vocabulary items that can be inserted have been inserted. However, the second insertion is always for the less marked vocabulary item from the same list. According to Marantz, Fission allows for multiple vocabulary items to correspond to one syntactic node at a single level of hierarchical structure. An example of fission in Tamazight Berber (Harley &

⁸ Halle (1997:427) introduces English as an example of fusion of the Φ -features (say, said), but the verb *be* (was, were) escapes this generalization.

Noyer, 1999) is given below where the AGR morpheme can appear as one, two, or three separate vocabulary items, realized as prefixes or suffixes.

(13) a. Prefix Conjugation in Tamazight Berber. *dawa* ‘cure’

singular plural
 3m *i-dawa dawa-n*
 3f *t-dawa dawa-n-t*
 2m *t-dawa-d t-dawa-m*
 2f *t-dawa-d t-dawa-n-t*
 1 *dawa-© n-dawa*

b. Vocabulary items in Tamazight Berber

/n-/ ↔ 1st pl
 /t-/ ↔ 2nd
 /-m/ ↔ pl m (2nd)
 /i-/ ↔ sg m
 /-d/ ↔ sg (2nd)

(Harley & Noyer, 1999: 19)

It is argued that in (13) the vocabulary items that correspond to the features in parenthesis cannot be inserted unless the parenthesized feature has been exhausted/discharged. For example, *-m* can be inserted only on a verb to which *t-* (2nd) has already been attached. Parentheses thus denote features that are secondarily expressed by a vocabulary item.

Impoverishment, also called blocking across positions, occurs when one morpheme or vocabulary item deletes the features of another independent morpheme before vocabulary insertion at the other morpheme. Impoverishment will be further explored in section (2.3).

Merger is a binary combination. It trades or exchanges a structural relation between two elements at one level of representation for a distinct structural relation at a subsequent level (Harley & Noyer, 1999). Marantz provides the example of the English possessive ‘s after the Vocabulary insertion. This is shown in (14).

(14) [the queen of England]’s hat → [the queen of England’s] hat

In (14), the clitic ‘s leans over England as opposed to the entire DP that shows the mismatch between word structure and functional structure.

Marantz (1997) describes (List 3) in (14) as the “Encyclopaedia”, that is the list of special meanings of particular roots relative to their system and within a local domain. Marantz argues that in DM all words are built from roots. Therefore, the noun “cat” is the root √CAT with a nominalizing affix, “little n”, which is phonologically realized as zero. The meanings and pronunciation of roots are fixed in the environment of the first category-determining head (n, v, a) to attach to the root and any category higher than that must take and perhaps manipulate the meaning and pronunciation of the lower category head.

Furthermore, Harley & Noyer (1999) state that in DM any given expression requires at least two structural descriptions during its derivation. A morphophonological description, which is the information about the expression’s phonological pieces and a morphosyntactic

description, is the description of an expression's morphemes and their constituents. Therefore, the expression "cats" can be stated as in (15).

- (15) morphosyntactic description: [Root [+plural]]
 morphophonological description: [kat+z]

DM provides the possibility of occurrence of morphological operations at various levels of derivation including post-syntax, which has recently been used as a key factor in capturing several phenomena not previously explained in languages (Bonet 1994, Boeckx 2000, Rivero 2004). For instance, the process of Impoverishment in DM makes it possible for a less marked feature to delete a more marked feature; other theories of linguistics have had a hard time explaining such situations. In section 2.3, I adopt DM to explain the lack of agreement for plural inanimate subjects in Persian by introducing an impoverishment rule deleting number feature in the morphology component after syntax.

1.4. Background Knowledge on Persian

In this section, I provide a summary of some basic attributes of Persian. Persian is a member of the Indo-European family of languages, and within that family, it belongs to the Indo-Iranian subfamily. The specific dialect under investigation, Modern Persian, is spoken in the capital of Iran, Tehran. Examples employed throughout this work may be formal or colloquial variants of Modern Persian.

Many linguists including Dabir Moghadam (1982), Samiiian (1983), Karimi (1989-2005), and Browning and E. Karimi (1994), among others, have argued that the underlying word order in Persian is SOV. An example is provided in (16) where the subject *mæn* (I) is in the sentence initial position followed by the direct object *ketab* (book) and then the indirect object *Sepehr*, with the verb *dad-æm* (gave-1sg) in the sentence final position.

- (16) *mæn ketâb ra be Sepehr dad-æm*
 I book Acc. to Sepehr gave-1sg
 I gave the book to Sepehr

Nonetheless, as scrambling is very common in Persian, example (16) may have grammatical counterparts with different word orders as illustrated in (17). However, as stated in Karimi (2003, 2005) these scrambled forms are not exactly the same and different discourse factors such as Focus and Topic differentiate them.⁹

⁹ See Karimi (2005) for a comprehensive discussion of scrambling in Persian.

- (17) a. mæn dad-æm ketâb ra be Sepehr
 I gave-1sg book Acc. to Sepehr
 I gave the book to Sepehr
- b. mæn be Sepehr dad-æm ketâb ra
 I to Sepehr gave-1sg book Acc.
 I gave the book to Sepehr
- c. dad-æm mæn ketâb ra be Sepehr
 gave-1sg I book Acc. to Sepehr
 I gave the book to Sepehr
- d. ketâb ra mæn dad-æm be Sepehr
 book Acc. I gave-1s to Sepehr
 I gave the book to Sepehr

Persian is a “Pro Drop” or “Null Subject” language; in other words, an overt subject is not required in the clause, and the Φ -features of the null subject are specified for Person and Number through the verbal morphology. Therefore, example (16) has a grammatical variant without an overt subject as illustrated in (18).

- (18) ketab ra be Sepehr dad-æm
 book Acc. to Sepehr gave-1sg
 I gave the book to Sepehr

I adopt the general idea that in pro-drop languages the verbal inflection is rich enough to encode the subject (Alexiadou & Anagnostopoulou 1998). Example (18) also demonstrates that gender is not morphologically encoded in Persian. This can be further seen in (19) where the pronoun “*u*” in the subject position stands for *he* as well as *she*.

- (19) u mæn ra did-Ø
 she/he I Acc. saw-3sg
 S/he saw me

The plural marker on animate NPs is traditionally the suffix *an*,¹⁰ while inanimate NPs acquire the suffix *ha*. Examples are below.

¹⁰ It is however possible for animate DPs to take the suffix *-ha*, such as *doxtær-ha*, but not vice-versa. Persian lexicon is influenced by Arabic and morphological features of Arabic words can be observed in loan words. Some Persian words may exhibit two different plural forms; a Persian plural form and an Arabic plural form that are known as the irregular plural form. This is shown in (I).

- (I) a. loghæt (*word*)
 b. loghæt-ha (*word-pl*)
 c. loghat (*words*)

(I.a) is the singular form; (I.b) is the Persian plural formation; and (I.c) is the Arabic plural formation utilized in Persian.

- | | | |
|------|--|--|
| (20) | a. doxtær- an
girl-pl
girls | b. name- ha
letter-pl
letters |
|------|--|--|

In (20a), *-an* is the plural marker for the animate DP *doxtær* while the plural marker for the inanimate noun in (20b) *name* is *-ha*. This rule is not respected in Modern Persian and the plural marker *ha* is used in both environments.

With respect to case marking, in Persian the nominative case marking is not overtly realized; i.e., the nominative is the unmarked form. Traditionally *ra* was believed to be the accusative marker.¹¹ A rich body of literature has explored the nature of *ra* (Karimi 1996-2005, Ghomeshi 1996, 1997, Dabir Moghaddam 1992, 2005, among many others). Dabir Moghaddam argues that *ra* is the topic marker. Karimi (2003-2005, among others) argues that *ra* is the marker of specific direct objects. Example (19) is repeated below.

- (21) u mæn ra did-Ø
 she/he I Acc. saw-3sg
 She/he saw me

In (21), the pronoun in the nominative form has no nominative marker and appears in the default form. Moreover, the specific direct object, the pronoun in the Accusative. form, appears with *ra*. The suffix *ra* may also be used in other environments such as *temporal adverbs* and *topics*. Examples are in (22).

- (22) a. emshæb ra ba ma bashid
 tonight Acc. with us stay
 Spend tonight with us
- b. Googoosh-ra mæn kheili doost dar-æm
 Googoosh-Acc I a lot friend have-1sg
 As for Googoosh, I like her a lot

In (22.a), *ra* appears on the temporal adverb *emshæb* (tonight) and in (22.b) it attaches to the topic *Googoosh*. Case patterns such as dative and ablative are indicated through prepositions. This is shown below.

¹¹ In colloquial Modern Persian, *ra* may appear as *ro/o*.

Introduction

- (23) mæn ketab ra æz u gereft-æm va be Færzin dad-æm
 I book Acc. from s/he got-1sg and to Farzin gave-1sg
 I got the book from her/him and gave it to Farzin.

In (23), the subject *mæn* (I) appears with no overt nominative case marker and is basically in an unmarked form. The accusative marker *-ra* appears on the direct object *ketab* (book). The indirect object *u* (s/he) is marked ablative by the preposition *æz* (from). The Genitive case is realized by the *Ezafe* construction¹² or by possessor pronouns. This is shown in (24).

- (24) u ketab-e mæn ra be pesær-æsh dad-Ø
 s/he book-*ezafe* I Acc. to boy-3sg gave-3sg
 S/he gave my book to his/her son

In (24), *my book* is expressed by adding the *Ezafe* vowel to the NP *ketab* (book); also, the possessor pronoun *æsh* (his/her) indicates the genitive case of *pesær* (boy).

The pronominal system of Persian is shown in table (25), and each of these full pronouns has a corresponding clitic form, listed in table (26).

(25) Full Pronouns in Persian

	Singular	Plural	
1 st	mæn	<i>I</i>	ma
2 nd	to/shoma	<i>you</i>	shoma/shomaha
3 rd	u/(an)	<i>s/he(it)</i>	anha/ishan
			<i>they</i>

(26) Clitic Pronouns in Persian

	Singular	Plural
1 st	-æm	-eman/emun
2 nd	-æt/et	-etan/etun
3 rd	-æsh/esh	-eshan/eshun

Full pronouns and clitic pronouns in Persian do not display any case distinction. Therefore, the clitic pronouns in (26) may function as *possessor*, *direct object*, or *indirect object* (in which case the clitic is attached to a preposition). Examples are in (27-28).

- (27) mæn ketab-æm ra be-h-eshan dad-æm
 I book-1sg Acc. to-^{h insertion}-3pl gave-1sg
 I gave my book to them

¹² The Ezafe construction will be explained in detail below.

In (27), the clitic *aem* (1sg) is attached to the noun (book) functioning as a possessor. It should be noted that the clitic pronoun *-aem* is homophonous with the verbal agreement *-aem* (1sg). In addition, the clitic pronouns *eshan* (3rd pl) is attached to the preposition *be* (to) functioning as the indirect object. An example of the clitic pronoun functioning as a direct object is provided in (28). The example in (28.a) contains a full pronoun *an* (that), where as (28.b) contains a clitic pronoun *ash* (3sg) on the verb *khord* (ate), and after the subject agreement *aem* (1sg).

- (28) a. **an** ra khord-æm
 that Acc. ate-1sg
 I ate that
- b. khord-æm-**ash**
 ate-1sg-3sg
 I ate it/that

Furthermore, *Negation*, *Tense*, and *Mood* may also be attached to the verb in Persian. Darzi (1995) describes the verbal morphology of the verb in Persian as follows.

- (29) *Verbal morphology of the verb in Persian*
negative marker-mood marker-verbal stem-Pt-Psp-personal ending
(Darzi, 1995:19)

The verb in (30) contains the subject *æm* (1sg), the direct object *æsh* (it), the negation *ne* (not), and the imperfective aspect *mi*.

- (30) ne-mi- bord-æm-eshan
not-ind-took-1st sg-3rd pl
I was not used to taking them

Complex predicates are extremely common in Persian and are replacing simple verbs (Barjasteh 1983, Ghomeshi and Massam 1994, Karimi 1997-2005, Karimi Doostan 1997-2005, Dabir Moghadam 1997, Megerdooomian 2000-2001, Vahedi-Langrudi 1996, Folli, Harley, among many others). A compound verb in Persian consists of a non-verbal constituent and a verbal component/light verb. The non-verbal component may be a noun, an adjective, a prepositional phrase, or a particle. This is shown in (31).

- (31) a. **Noun+ Light Verb (LV)**
kotak zadan/xordan
beating hitting/colliding
to beat, to get beaten
- b. **Adjective+LV**
pahn kardan/shodan
wide doing/becoming
to spread, to widen

c. **Particle+LV**

bala avardan
up bringing
to vomit

d. **PP+V**

be donya amadan
to world coming
to be born

(Karimi 2005: 12)

The subject-verb agreement morphology of complex verbs in Persian always appears on the verbal component and never on the non-verbal constituent.¹³ This is shown in (32).

- (32) mæn u ra [æz dæst dad-æm]
I s/he Acc. from hand gave-1sg
I lost him/her

In (32), the verbal component of the compound verb, *æz dæst dad-æm*, is *dad-æm* (gave-1sg), hosts the agreement clitic.

However, the object may be cliticized either on the nonverbal element of the compound (33.a) or on the verbal element (33.b). If it appears on the verbal component, then it always proceeds the subject agreement.

- (33) a. mæn (u ra) [æz dæst-esh dad-æm]
I s/he Acc. from hand-3g gave-1sg
I lost him/her
b. mæn (u ra) [æz dæst dad-æm-esh]
I s/he Acc. from hand gave-1sg-3sg
I lost him/her

Several studies have attempted to determine the syntactic and semantic roles of the non-verbal element and the light verb. Mohammad and Karimi (1992) argue that the light verb is semantically empty and the non-verbal element contributes to the argument structure of the verb. Vahedi-Langrudi (1996) argues that light verbs are bleached elements and do not correspond to particular thematic roles; moreover, the transitivity of the compound verb is determined by the non-verbal element. Karimi Doostan (1997) argues that the light verb contributes aspectual information but not argument structure. Ghomeshi and Massam (1994) and Barjasteh (1983) consider the non-verbal element as an argument of the verbal element. Ghomeshi and Massam (1994) suggest that complex verbs in Persian are syntactically transparent and are formed in the syntax. Karimi (1997) argues that both components of the

¹³ This issue differentiates constructions like *mæn xosh-æm amæd* (I my-pleasure came or pleasure came to me) which have been considered as compounds by a number of linguists. The fact that the verbal element in these constructions never bears agreement morphology differentiates them from compound verbs. I will discuss these kinds of constructions in detail in chapter 4

compound verb contribute a thematic structure that undergoes a semantic fusion after incorporating at LF.

Traditionally, complex verbs have been considered as a lexical unit since these compounds undergo nominalization and adjectival formation, bear a single stress, and cannot be separated by interveners such as PP's. However, other linguists have argued that compound verbs are visible to syntactic and morphological processes since the two components can be intervened by negation and inflectional affixes, auxiliaries, modals, and emphatic elements (Mohammad and Karimi 1992, Megerdooomian 2001, 2002). In addition, certain (but not all) non-verbal elements can be limitedly modified, gapped (Karimi 1997), or relativized. Based on Hale and Keyser's (1993) approach and the non-lexicalist Distributed Morphology model, and following the previous proposals on compositionality of the Persian compound verbs (Karimi Doostan 1997, Vahedi-Langrudi 1996), Megerdooomian (2001) proposes that the non-verbal element and the light verb decompose to even smaller elements/basic syntactic atoms such as roots and functional elements. Foley, Harley, and Karimi (FHK) (2003) argue that Hale and Keyser's (1993, 2002) model fits perfectly into the Persian compound verbs (see examples therein.) FHK argue that the light verb in compound verbs may have an effect on (a) agentivity/causativity, (b) eventiveness, and (c) duration. Accordingly, Karimi (2005) argues that the light verb and the non-verbal element in the Persian compound verbs are "separately generated and combined in syntax, and become semantically fused at a different level (page 15)." She argues that although the two parts of the compound may be syntactically intervened "limitedly", they semantically behave as a single unit. I will come back to the behavior of compound verbs in Chapter 4 where I discuss the Persian psychological verbs.

Before closing this chapter, I briefly introduce a specific construction in Persian called the *Ezafe* Construction since it might appear in the Persian examples throughout this monograph. *Ezafe* literally means "adding" and comes from a borrowed Arabic root. *Ezafe* is an unstressed enclitic vowel *-e* (*-ye* after vowels) and is semantically vacuous. It appears on lexical items that are linked within adjectival, nominal, and prepositional phrases. The *Ezafe* construction does not appear on verb phrases. Examples of *Ezafe* constructions are illustrated below.

- (34) between a noun and its complement

xordæn-e	ghæza
eating- <i>ezafe</i>	food
The eating of food	

- (35) between a noun and a modifier

doxtær-e	ziba
girl- <i>ezafe</i>	beautiful
Beautiful girl	

(36) between a noun and a possessor

ketâb-e	Soroush/mæn
book- _{ezafe}	Soroush/I
Soroush's/my book	

(37) between an adjective and its complement

abi-ye	roushæn
blue- _{ezafe}	light
Light blue	

(38) between a preposition and its complement

posht-e	dær
behind- _{ezafe}	door
Behind the door	

For a series of different syntactic proposals on the Ezafe construction, see Tabaiian (1979), Samiian (1983), Mahootian (1993), Ghomeshi (1996, 1997), Kahnemuirpour (2000), and Moin Zadeh (2005).

1.5. Outline of the Book

The main purpose of Chapter 2 is to explore the Persian constructions with plural inanimate subjects that exhibit the verbal agreement restrictions. In order to do so, I will discuss several constructions with the verbal agreement restrictions in a cross-linguistic way. I first discuss the 3rd person restriction in Icelandic, which is a restriction raised in presence of non-nominative subjects. I provide a series of proposed analyses for the Icelandic person restriction, most of which have tried to capture the restriction within syntax. Boeckx (2000), however, attempts to capture the restriction within post-syntactic morphology. In order to further establish the place of morphology in explaining verbal restrictions, I introduce the quirky restriction in Spanish (Rivero 2004-2005) and show how this restriction is captured within post-syntactic morphology. The restrictions that lack of animacy causes in Persian will be explored in detail next. I first lay out the situation and then adopt distributed morphology to explain that such agreement does not belong to syntax and is the result of a feature mismatch, which occurs in post-syntactic morphology. I also provide an analysis of Karimi (2005) for the Persian animacy restriction in which the restriction is explained in syntax and interface. I show how this account is not entirely capable of capturing the true nature of the restriction.

Chapter 3 provides the basis for the discussion of the Persian psychological verbs that will be provided in Chapter 4. The aim of Chapter 3 is twofold; (a) to discuss the psychological predicates cross-linguistically and to provide a comprehensive summary of the relevant literature, (b) to introduce a cross-linguistic series of distinct constructions in which an extra argument with some subject properties is added to a complete clause. Namely, I

Chapter 1

discuss the dative/double subject constructions in South Asian languages, Categorical/Broad subjects in Hebrew/Japanese/Arabic, and the dative subject constructions in south Slavic/Albanian. In sum, the discussion of Chapter 3 provides the path to a better understanding of Chapter 4, which concerns the Persian psychological constructions.

Chapter 4 discusses the psychological verbs in Persian. I first introduce the notion of applied arguments (Pylkkänen 2002, McGinnis 2001, and Cuervo 2003) and then argue that the Persian psychological verbs involve applied arguments and applicative heads.

The discussion in Chapter 4 provides an original contribution to the study of the psychological verbs in Persian. I discuss the properties and characteristics of the Persian psychological verbs in great detail. I introduce the previous literature and provide reasons to show that they are not capable of capturing all the properties these constructions exhibit. Unlike previous literature that considers the Persian psychological verbs as compound verbs, I provide evidence such as nominalization to prove that their structure differs from complex predicates. Furthermore, the A and $\bar{A}$ properties of the experiencer will be explored with the conclusion that rather than being a regular subject, the experiencer is an applied argument to a complete clause which contains a Tense requirement. Chapter 5 consists of concluding remarks and implications.