

The Empty Noun Construction in Persian

by

SAEED GHANIABADI

**A Thesis submitted to the Faculty of Graduate Studies of
The University of Manitoba in partial fulfilment
of the requirements of the degree of**

DOCTOR OF PHILOSOPHY

**Department of Linguistics
University of Manitoba
Winnipeg**

Copyright © 2010 by Saeed Ghaniabadi

Abstract

This dissertation explores, within the general framework of Distributed Morphology, the licensing conditions of empty nouns in Persian, a Western Iranian language, and the issues that arise within this context for the distribution of plural marking and the insertion of the Ezafe vowel. With respect to the licensing of the empty noun, the proposal made in this thesis is along the lines of those that link ellipsis to information structure (e.g. Rooth 1992a, 1992b; Gengel 2007, among others). It is suggested that the Empty Noun Construction (ENC) is derived through the interaction between the following two information-structural features: (i) the E-feature, which ensures that the head noun is identical with its counterpart in the antecedent and specifies the head noun for non-pronunciation; (ii) the focus feature F, which specifies the remnant modifier as an element which is in some kind of contrastive relationship with its corresponding element in the antecedent. The interaction between these two features is implemented in the syntax in a phase-based derivation.

Plural marking and Ezafe insertion in the ENC are accounted for within an articulated derivational model of PF (Embick & Noyer 2001; Embick 2003 *et seq.*; Pak 2008). It is proposed that the displacement of the plural marker in the ENC is motivated by the non-pronunciation of the head noun and is handled early in the PF derivation by Local Dislocation operation. Adopting Pak's (2008) model of syntax-phonology interface, the rule responsible for the insertion of the Ezafe linker *-e* is argued to be a phonological rule that applies at the Late-Linearization stage to connect [+N] heads to their following modifiers/complements.

Acknowledgements

First and foremost, I would like to express my sincere and deepest gratitude to Dr. Jila Ghomeshi, my supervisor. She has been the backbone of this entire process, and I could not have done it without her. Whenever I was feeling lost or doubting my research, I would schedule an emergency meeting with Jila because I knew that she would set me back on the right track by putting things in perspective and reminding me why my work was worthwhile. Besides shaping this dissertation in content and form, Jila mentored me intellectually, professionally, and personally. She is a model that I aspire to, not only as an academic but also as a person, and I thank my lucky stars that I have had the opportunity to work with her.

I would also like to thank the other core members of my thesis committee, Dr. Kevin Russell and Dr. Chris Wolfart, for their valuable contributions to the thesis. Kevin was my phonology teacher. I learned many things in his phonology course including Optimality Theory. He supervised my first generals paper on m/p-echo reduplication in Persian which I wrote within the framework of OT. I am also thankful to him for providing constructive comments on my thesis and raising provocative questions that define further directions for my future research. My interest in working on plural marking in Persian started in Number Seminar with Professor Wolfart. Chris helped me shape my ideas into a presentable linguistic form. The course paper I wrote for that seminar served as the basis for the ideas pursued in my second generals paper on noun ellipsis and plural marking in Persian. The ideas in that paper were used as the building block of the present thesis.

I am very grateful to Dr. Melanie Soderstrom and Dr. David Embick who joined my thesis committee later. Melanie contributed to the defense with thoughtful questions and comments. It was a great honor and privilege to have David as my external examiner. Most of the proposals I have made in this thesis are based on David's extensive theoretical contributions to Distributed Morphology. I could not have wished for a better person than David to read my work. He brought a very constructive perspective to the thesis. The influence of his report, comments, questions, and discussions goes far beyond the present thesis.

My gratitude goes to Dr. Rob Hagiwara, my other linguistics teacher, with whom I had Intonation/Prosody Seminar and Field Methods. I particularly enjoyed the structure of his Field Methods course and learned a lot about language documentation. I would also like to thank other linguistics people at the University of Manitoba, Dr. Terry Janzen and Dr. David Pentland for their kindness and help throughout my time there. Thank you to Debbie Spindler, the administrator, and Arden Ogg, the technician, for their help when I most needed it.

A big thank you to my first linguistics teacher Dr. Mohammad Dabir-Moghaddam to whom I owe much of what I learned about linguistics during my undergraduate and

Master's studies in Iran. His passion for linguistics and his encouragement along the way made me more motivated to pursue linguistics at the graduate level. I should also thank Dr. Kourosh Safavi, my other long-term linguistics teacher in Iran, who was a constant source of inspiration, support, and encouragement. He was fun and made linguistics fun to study.

Parts of this thesis have been presented in various conferences. I would like to thank the audiences at the 2007 ICIL2 meeting; the Mass/Count Workshop at the University of Toronto; the University of Manitoba Linguistics Colloquium; and the 2009 CLA conference at the Carleton University for their comments and discussions.

I gratefully acknowledge financial support from the I.R. of Iran's Ministry of Science, Research, and Technology for the first four years of my graduate work and numerous travel grants from the University of Manitoba Faculty of Graduate Studies, Faculty of Arts, Graduate Students' Association, UMSU, and Alumni Association.

Finally, I would like to thank my parents, who have supported me my entire life and are undoubtedly excited that I can finally say "I'm done". To my mother, Aliye Saleki, and my father, Abbasali Ghaniabadi, thank you for making me the person that I am today. I am also greatly thankful to my parents-in-law, Naser Amini Salehi and Sedighe Afchangi, for their constant help and support throughout the time we were out of Iran and specifically during the three months before I completed my Ph.D. program. Thank you to my dearest sons Sajjad and Sina for being such good boys for me and their mom. I apologize to you for not being with you when you needed my presence and company. In the end, thank you most to my dear wife, Behjat Amini Salehi. You have been there every step of the way with me, and your unfailing love, support, and sacrifices have made me get through my graduate studies. This humble thesis is dedicated to you.

Table of Contents

Abstract	i
Acknowledgements	ii
Table of contents	iv
List of abbreviations	ix
Chapter 1: Setting the scene	1
1.0. Introduction.....	1
1.1. Theoretical framework.....	4
1.1.1 Distributed Morphology.....	4
1.1.2 Phase Theory and cyclic spell-out.....	9
1.1.2.1 What constitutes a phase?.....	10
1.1.2.2 A derivational model of PF.....	12
1.1.2.3 Local Dislocation.....	17
1.2. Main proposals in brief.....	19
1.3. Background: The Persian noun phrase structure in brief.....	22
1.3.1 Prenominal position.....	23
1.3.2 Postnominal position: the Ezafe domain.....	26
1.4. The syntax of Persian noun phrase.....	29
1.4.1 Functional projections.....	29
1.4.2 The Ezafe Domain.....	31
1.5. Outline of the thesis.....	36
Chapter 2: Previous accounts of Empty Noun Constructions	40
2.0. Introduction.....	40
2.1. The pro-based accounts.....	41
2.2. The PF-deletion accounts.....	48
2.2.1 Focus Condition on ellipsis.....	49
2.3. Empty nouns and Distributed Morphology.....	56

2.4.	Summary and relevance to dissertation	58
Chapter 3: The Empty Noun Construction in Persian.....		61
3.0.	Introduction.....	61
3.1.	Where do we get empty nouns?.....	62
3.1.1	Prenominal position.....	63
3.1.2	The Ezafe domain.....	67
3.2.	The structural status of the empty noun.....	71
3.3.	Towards an analysis.....	74
3.3.1	Information-structural configuration of the ENC.....	75
3.3.1.1	The E-feature	75
3.3.1.2	The Focus feature F	83
3.3.2	Proposal.....	94
3.3.2.1	Focus Condition on the ENC	94
3.3.2.2	Licensing properties of the ENC.....	96
3.3.2.3	The ENC at the interfaces.....	99
3.3.3	Explanation of the data.....	101
3.4.	Summary	116
Chapter 4: Plural marking and the ENC		118
4.0.	Introduction.....	118
4.1.	Distributional properties of nouns and plural marking in Persian	119
4.2.	Plural marking and empty nouns	125
4.3.	The syntax of plural formation	134
4.3.1	Definiteness/number syncretism	137
4.3.2	Persian plural marker: a head or a modifier?	141
4.3.3	Proposal.....	143
4.3.4	What category hosts <i>-hâ</i> ?.....	150
4.4.	Plural marking and the ENC in Persian	151
4.4.1	<i>-hâ</i> mobility: a case for Local Dislocation	154
4.4.1.1	<i>-hâ</i> mobility within the <i>nP</i> phase	155

4.4.1.2	- <i>hâ</i> mobility across phase boundary	169
4.4.1.3	- <i>hâ</i> mobility in the presence of two potential hosts	174
4.4.1.4	- <i>hâ</i> mobility and numerals	177
4.5.	Summary	179
Chapter 5: The Ezafe construction revisited.....		180
5.0.	Introduction.....	180
5.1.	Syntactic properties of nominal modifiers in the Ezafe domain.....	181
5.1.1	Background literature.....	181
5.1.1.1	Samiian (1983) and Ghomeshi (1997).....	182
5.1.1.2	Samvelian (2006).....	184
5.1.2	Proposal: mixed approach.....	185
5.1.3	Distinguishing nominal modifier types: further evidence.....	190
5.1.3.1	Compounding.....	190
5.2.	Accounts of the presence of Ezafe.....	202
5.2.1	Ezafe as a Case-marker (Samiian 1994; Larson & Yamakido 2005)	205
5.2.2	Ezafe insertion at PF (Ghomeshi 1997)	206
5.2.3	Syntactic movement triggers Ezafe (Kahnemuyipour 2000, 2006).....	208
5.2.4	Ezafe as phrasal affix (Samvelian 2006).....	212
5.3.	Proposal: Ezafe insertion at the syntax-phonology interface.....	214
5.3.1	A theory of syntax-phonology interface (Pak 2008).....	216
5.3.2	Sensitivity of Ezafe insertion to morphosyntactic structure.....	218
5.3.3	Towards establishing the Ezafe insertion domain.....	221
5.3.3.1	The <i>nP</i>	223
5.3.3.2	The phrasal modifiers in ModP/PossP	230
5.3.4.	EIR in Persian noun phrases: the complete picture.....	251
5.4.	Ezafe insertion in the ENC	256
5.4.1	- <i>hâ</i> mobility and Ezafe insertion	264
5.5.	Summary	268

Chapter 6: Conclusions and future work	269
6.0. Introduction.....	269
6.1. Summary of proposals	269
6.2. Future directions	275
6.2.1 Classes of licensing adjectives in the ENC	275
6.2.2 Partitive constructions and empty nouns.....	278
6.3. Concluding remarks	281
References	283

List of Abbreviations

1	first person	IMP	imperative
2	second person	IND	indefinite
3	third person	LD	Local Dislocation
<i>aP</i>	adjective phrase	LF	Logical Form
ADJ	adjectival suffix	Mod _H	head modifier
CardP	cardinal phrase	Mod _{ph}	phrasal modifier
CL	Clitic	ModP	modifier phrase
CLP	classifier phrase	NEG	negative
CLS	Classifier	NUM	numeral
COMP	Comparative	OM	object marker
DEF	definite	PF	Phonological Form
DEM	demonstrative	PIC	Phase Impenetrability Condition
DemP	demonstrative phrase	PossP	possessor phrase
DM	Distributed Morphology	PP	prepositional phrase
DP	determiner phrase	PP _{LOC}	locative prepositional phrase
DUR	durative	PL	plural
E-feature	ellipsis feature	PL _{DEF}	definite plural
EIR	Ezafe Insertion Rule	PP	past participle
ENC	Empty Noun Construction	PreN	prenominal modifier
EZ	Ezafe	PRS	present
F-feature	Focus feature	PST	past
H _n	head noun	Q	quantifier

QP	quantifier phrase	SUP	superlative
REL	relative marker	SUPORD	super-ordinal
r_F	F-specified remnant modifier	SuprP	superlative phrase
SBJ	subjunctive	VI	Vocabulary Item
SG	singular	\sqrt{P}	Root phrase
SG _{DEF}	definite singular		

Chapter 1

Setting the Scene

1.0. Introduction

This thesis is concerned with the study of the Empty Noun Construction (henceforth ENC) in Persian and the consequences it has for plural marking and Ezafe insertion¹. Like many other languages, Persian, a Western Iranian language, allows noun phrases with phonetically unexpressed head nouns (a.k.a. *empty nouns*). There are two conditions that must be met for the non-pronunciation of the head noun: (i) there must be a linguistic or discourse antecedent (shown in (1)a and (1)b respectively), and (ii) there must be at least one prenominal or postnominal modifying element in the noun phrase. Throughout this thesis, empty nouns in the illustrative examples of the ENC are struck through; nominal modifiers are in bold only in the examples in (1).

(1) a. *Empty nouns with prenominal modifiers*

hamiše avvalin jomle [**saxt-tar-in** ~~jomle~~]-e.
always first sentence hard-COMP-SUP ~~sentence~~-is
‘The first sentence is always the hardest.’

¹ Ezafe, realized as *-e*, is a linking element that connects a nominal head to its following modifier(s)/complement. It is discussed in Chapter 5 of this thesis.

b. *Empty nouns with postnominal modifiers*

[A shopkeeper's response to a customer who asks for green pens]

[~~ædkâ~~ **sabz**] na-dâr-am, [~~ædkâ~~ **âbi**] mi-xây?
~~pen~~ green NEG-have-1SG ~~pen~~ blue DUR-want.2SG

'I don't have a green [one], do you want a blue [one]?'

One interesting and overlooked issue about the ENC in Persian is the fact that the plural specification of empty plural nouns must be overtly expressed on one of the immediately neighboring modifiers. The host to the stranded plural marker *-hâ* can be an attributive adjective immediately following the empty noun. In the absence of the attributive adjective, *-hâ* can attach to its immediately preceding modifier. This is illustrated in (2):

	QUANTIFIER	NOUN	-PL	ADJ	POSSESSOR
(2) a.	behtar-in	dânešju	-hâ -ye	javân-e	dânešgâh
	best-SUP	student	-PL-EZ	young-EZ	university

'the best young students of the university'

ENC: behtar-in ~~dânešju~~⟨-hâ⟩ javân-**hâ**-ye dânešgâh
 'the best young [ones] of the university'

Q	N	-PL	POSSESSOR
b. behtar-in	dânešju	-hâ-ye	dânešgâh

‘the best students of the university’

ENC: *behtar-in-hâ-ye* ~~*dânešju*~~(-hâ) *dânešgâh*

‘the best [ones] of the university’

We see in the above examples that the absence of an overt head noun causes the plural marker to move onto the immediately adjacent modifier. Two further observations can be made. When the noun is not pronounced in (2)a, the Ezafe vowel also disappears but when the noun is not pronounced in (2)b, an Ezafe vowel appears on the pronominal modifier. A proper treatment of Ezafe must account for these facts.

This dissertation explores the licensing conditions of empty nouns in Persian and the issues that arise within this context for the distribution of plural marking and the insertion of the Ezafe vowel. As such, it seeks to provide answers to the following fundamental questions: (i) How does the interaction between having antecedents and the presence of a nominal modifier play a role in the licensing process of empty nouns in Persian? (ii) How does the non-pronunciation of the head noun affect the distribution of plural marking and what is its consequence for the syntax of the Persian noun phrase? (iii) What is the explanation for the behavior of the Ezafe vowel in the context of the ENC and how does it contribute to our overall understanding of the nature and function of Ezafe in the Persian noun phrase?

The example sentences/phrases I use in the present study are taken from Conversational Persian. There are also places where I draw on examples from Formal Persian; reference to such examples will be pointed out where relevant. The naturally occurring data that I have used in some places are taken from Persian weblogs². Such examples are accompanied by reference to the sources from which they are taken.

1.1. Theoretical framework

This thesis is situated within the general theoretical framework of Distributed Morphology (DM hereafter—Halle & Marantz 1993, 1994; Harley & Noyer 1999; Embick & Noyer 2006, and related work), combined with Phase Theory (Chomsky 1999 *et seq.*) and cyclic (multiple) spell-out (e.g. Chomsky 2000). In what follows, I will present a brief overview of the main tenets of these theories.

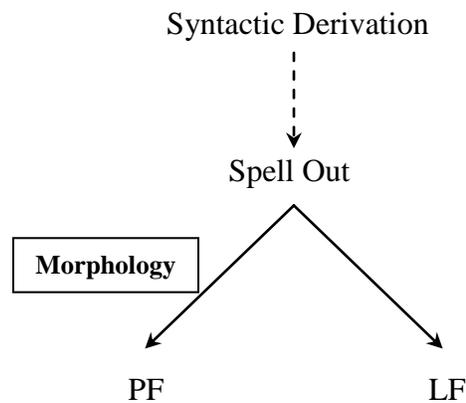
1.1.1 Distributed Morphology

DM proposes an architecture of grammar in which word structure and phrase structure are derived from a single generative system, namely syntax. The syntax involves operations Merge, Move, and Agree that generate syntactic structures. At spell-out, the syntactic output is handed over to the PF and LF branches in a cyclic (phase-based) manner. At this stage, the derivation is input to further operations. At LF, syntactic

² Persian weblogs provide sentences in a spoken discourse context. These sentences represent conversational dialect in writing. They include properties such as shortened verbal stems and frequent use of attached pronominal forms that are not part of the standard formal grammar (cf. Megerdooomian 2008).

derivations are subject to operations that affect semantic interpretation regardless of phonological form. At PF, morphology interprets the output of syntax: syntactic structure undergoes morphological operations that affect phonological form, not interpretation. These PF operations readjust the minimal syntax-morphology mismatches that arise from language-specific requirements. All of these operations follow the principles of syntax, i.e. hierarchical principles and locality conditions. The schema in (3) is the model of grammar presented in Embick & Noyer (henceforth E&N—2006):

(3) *The Grammar*

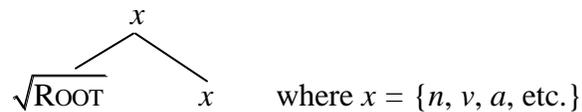


In DM, the basic units of syntactic derivation are *morphemes*: abstract terminal nodes with complexes of syntactico-semantic features. Morphemes, according to E&N (2006), are of two types: *Abstract/Functional morphemes* and *Roots*. Functional morphemes are composed exclusively of non-phonetic features drawn from Universal Grammar, such as [past] or [pl], or the feature (or features) that make up the determiner node D of the English definite article *the*. Roots, as language-specific combinations of

sound and meaning, make up the open-class or *lexical* vocabulary. They include items such as $\sqrt{\text{CAT}}$, $\sqrt{\text{OX}}$, or $\sqrt{\text{SIT}}$.

As a general assumption, following E&N (2006), I presume that Roots cannot appear (cannot be pronounced or interpreted) without being *categorized*; Roots must always be categorized in certain local relations with category-defining functional heads (*n*, *v*, etc.; cf. Marantz 1996, 2001). This is because Roots are claimed not to possess or contain any grammatical features (Marantz 1997, Harley & Noyer 1999 and references therein). Thus, if we take $\sqrt{\text{ROOT}}$ as a variable for a member of the so-called *lexical categories* such as Nouns, Verbs and Adjectives, $\sqrt{\text{ROOT}}$ has to be in a local relationship with the category-defining functional head *x* to be categorized. On this view, the category of $\sqrt{\text{ROOT}}$ is a function of the functional category that takes $\sqrt{\text{ROOT}}$ as a complement.

(4) *Root categorization*



Within DM, all phonological expressions are subject to Late Insertion, i.e. insertion after syntax. Elaborating, functional morphemes such as [pl] or [past] are supplied postsyntactically with phonological features (including the phonological null element or zero - \emptyset -) in a process called *Vocabulary Insertion*. Vocabulary Items (VIs), shown schematically in (5), consist of phonological exponents (Ph1, Ph2 ...) paired with their syntactico-semantic features (F1, F2 ...) or context of insertion. There is usually a set of VIs specified for insertion at a particular terminal node. VIs of a given functional

morpheme compete for insertion into that particular functional terminal node. Such competition is subject to Halle's (1997) *Subset Principle* which requires the VI with the largest subset of the features present on the terminal node be inserted (the schematic representation of VIs is adopted from Poot & McGinnis (2005)).

- (5) /Ph1/ ↔ [F1] [F2]
 /Ph2/ ↔ [F2] [F3]
 NULL ↔ [F1]
 /Ph3/ ↔ [F3]
 /Ph4/ ↔ []

In the formation of plural nouns in English, for example, Vocabulary Insertion provides phonological features to the functional [pl] morpheme, which has syntactically merged with a noun ($\sqrt{\text{ROOT}} - n$). The VI in (6) formally expresses the addition of the default phonological exponent of the English plural /-z/ to the node bearing the feature [pl]:

- (6) /-z/ ↔ [pl]

We note that besides -z in (6), there are other phonological exponents that compete for insertion into the node with the feature [pl], viz. $-\emptyset$ (as in *moose- \emptyset*) and *-en* (as in *ox-en*).

- (7) $-\emptyset$ ↔ [pl] / { $\sqrt{\text{MOOSE}}$, $\sqrt{\text{SHEEP}}$, ... } _____
 /-en/ ↔ [pl] / { $\sqrt{\text{OX}}$, $\sqrt{\text{CHILD}}$, ... } _____

According to the Subset Principle, the winner of the competition is the most specified one that meets its contextual feature. For instance, in the context of \sqrt{OX} , *-en*, and in the context of \sqrt{MOOSE} , $-\emptyset$ spell out the feature [pl] and thus take precedence to the less specified *-z*.

Unlike VIs inserted into functional morphemes, VIs for Roots are not inserted in competition. Rather, their insertion, subject to conditions of *licensing*, can occur freely (Harley & Noyer 1999: 5). H&N define licensers as category-defining functional morphemes, i.e. little *x*, that appear in certain structural relations to the Root where the VI is inserted. For instance, VIs like *table*, *tree*, *robin*, *dog* or any other VI we normally call a ‘noun’ can be inserted into a Root morpheme that is syntactically merged with little *n*.

Vocabulary Insertion is part of the operations comprising (*PF*) *Spell-Out* that take a representation from the syntactic component to its phonological form. *Morphology* in (3) refers to the module of grammar responsible for spell-out. In the default case, the PF representation of an expression parallels its syntactic representation. However, language-specific requirements may cause certain syntax-morphology mismatches to arise. These requirements involve PF operations that may add or delete features or terminal nodes (e.g. Fission, Impoverishment), or modify structural relations between two elements (Merger).

In semantic interpretation, under DM, LF does not represent meaning but is merely “a level of representation which exhibits certain meaning-related structural relations, such as quantifier scope” (Harley & Noyer 1999: 4). As suggested by Marantz (1997), the derivation as a whole is subject to semantic interpretation. The other component of grammar that interfaces with LF in semantic interpretation is *the Encyclopedia*. The

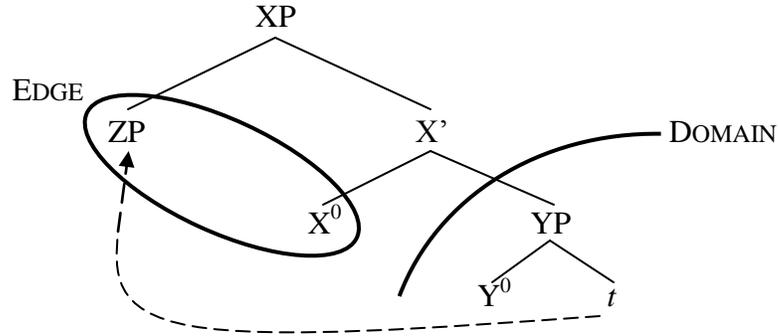
Encyclopedia assigns noncompositional meanings to VIs (sometimes in the context of other VIs). In other words, the Encyclopedia lists any expression (even a single word or subpart of a word) whose meaning is not wholly predictable from its structural representation. The term *idiom* is used to characterize such expressions (Marantz 1996, 1997). Accordingly, the Encyclopedia can contain only Roots or conventional idioms (e.g. *kick the bucket*), not functional morphemes³.

1.1.2 Phase Theory and cyclic spell-out

As noted earlier, within DM, spell-out delivers formed syntactic structures to the interface levels in a cyclic or phase-based manner. This view is based on Chomsky's (1999 *et seq.*) proposal that syntactic derivations proceed in chunks (*phases*). A phase refers to a locality domain for semantic and phonological processing, which is identified with certain designated syntactic nodes (for Chomsky, *v*P and CP). When a phase is complete, it is sent off to the interfaces (LF and PF) for interpretation. Chomsky's (2000) Phase Impenetrability Condition (PIC) constrains accessibility of later syntactic operations to the *domain* (complement) of a phase head once it is transferred to the interfaces. Syntactic operations outside the phase can only target the head and other elements in the *edge* (specifiers and adjuncts) of the phase.

³ Idioms such as *kick the bucket* are syntactically constructed objects that are listed in the Encyclopedia as corresponding to non-compositional meanings. Although the individual elements inside these syntactically complex objects maintain their compositional meanings, their combination, according to Marantz (1996), involves no compositional semantics. Thus, in the Encyclopedia, the definite determiner *the*, a functional morpheme, does not have the semantics of a functional morpheme in combination with other elements in the phrasal idiom *kick the bucket*.

(8)



The constituents contained in the internal domain (complement) of X^0 in the above structure are not accessible to syntactic operations outside of XP. Material inside the domain (e.g. ZP) can be accessible to operations in the higher structure by moving to the edge of the phase.

1.1.2.1 What constitutes a phase?

The question of what constitutes a phase has been subject to theoretical and empirical controversies (see e.g. Munakata 2009, Beek 2008, among others). Chomsky (2000) suggests that only CP and (transitive) vP correspond to phases because they are propositional: vP has a full-fledged argument structure and CP has a clausal status since it contains tense and force. Munakata (2009) questions the idea that CP and vP, as two quite distinct syntactic objects, should be regarded as having the same conceptual property, i.e. propositionality. He makes clear that according to Chomsky (2000), “the phase status of v^*P [i.e. transitive vP—SG] is attributed to its full-fledged argument structure, which is a conceptual matter [...] and clearly motivated by the selectional requirements of lexical heads, whereas the phase status of CP arises from the presence of tense and force, which are not so obviously conceptual in nature and cannot be motivated

by the selectional requirements of lexical heads” (Munakata 2009: 53). Munakata is thus suggesting that propositionality is not a crucial factor in determining phasehood. This view is reinforced by the fact that although DP (Svenonius 2004) and PP (Abels 2003) have also been proposed as phases, they are not propositional in any sense.

Beek (2008) provides empirical evidence against the claim that phases are propositional objects. The property of being a phase, according to her, cannot systematically be related to a characteristic of the head. A phase is rather defined by locality and edge effects. Beek shows that there is variation in the phase structure of sentences. She demonstrates that the three phase heads C, P and v are not phase heads in all environments. For each of the heads, it is possible to find cases in which the head does not define a phase. For instance, in explaining transparency effects in Dutch infinitival complement clauses, Beek argues that the absence of locality boundaries between the embedded infinitival clause and the matrix clause results in transparency effects. Thus, if the infinitival clause projects a locality boundary (i.e. defines a phase), no transparency effects are observed. If there is no locality boundary (phase head) between the matrix verb and the infinitive, then transparency effects are obtained.

In addition to phase heads mentioned so far (i.e. CP, transitive v P, DP, and PP), other heads have been introduced in the literature as phases. Marantz (2001, 2007) argues that category-defining functional heads—*little x* heads (a , n , v)—are phase heads in that they initiate spell-out (see also Embick 2008: §1.4). Svenonius (2004) introduces n as a phase head in the nominal structure, parallel to v in the clausal structure. Adopting Beek’s (2008) view on phases, I will demonstrate in this thesis that in the ENC, n defines a locality boundary when the empty noun is licensed in the n P domain. Whenever the

empty noun is licensed in the domain of a higher head, nP and its higher constituent form a single phase.

So far we have briefly reviewed a version of Distributed Morphology which relies on the notion of phases and cyclic spell-out. This is the main framework within which licensing empty nouns, mobile plural marking, and Ezafe insertion in the ENC are accounted for in this thesis. However, since the non-pronunciation of the head noun, movement of the stranded plural marker, and insertion of the Ezafe vowel are handled by independent operations on the PF branch, we need to have a brief overview of an articulated model of PF derivation proposed in E&N (2001, 2006), Embick (2007), and Pak (2008).

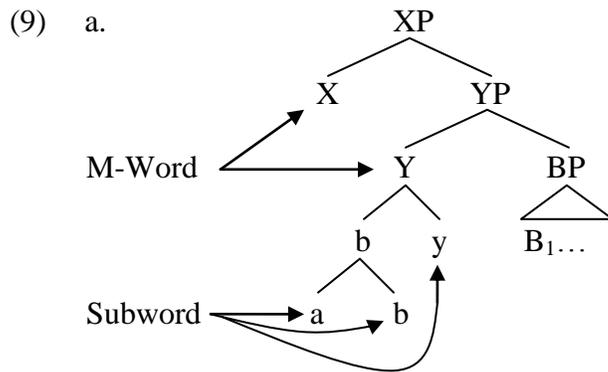
1.1.2.2 A derivational model of PF

As noted earlier, within PF, syntactic structures may be subject to several processes. Language-specific requirements may cause certain syntax-morphology mismatches to arise. These requirements involve PF operations that may add or delete features or terminal nodes (e.g. Fission, Impoverishment), or move terminal elements (Morphological Merger). Some of these operations apply before the insertion of phonological exponents (Vocabulary Insertion) and some apply after it. Morphological Merger, for instance, can be identified as applying prior to Vocabulary Insertion (Lowering) or after or simultaneous with it (Local Dislocation)⁴. In this section, we focus

⁴ For a thorough investigation on Lowering within DM, see Skinner (2009). Local Dislocation has been discussed in E&N (2001) and in more detail in Embick (2003).

on operations that apply after or simultaneous with Vocabulary Insertion as they are relevant to the current study. Before discussing these operations, an important distinction between the syntactic objects that are targeted by PF rules is in order.

E&N (2001) introduce two distinct types of syntactic objects that post-syntactic operations like Lowering and Local Dislocation manipulate in Morphology: *Morphosyntactic Words* (M-Words) and *Subwords* (note that phrases and intermediate nodes (like the higher b in (9)a) are not significant for PF operations).



b. M-Word: (Potentially complex) head not dominated by further head-projection

(cf. Chomsky (1995) “ H^{0max} ”)

c. Subword: Terminal node within an M-Word (i.e. either a Root or a feature bundle)

Additionally, E&N (2001) propose a typing assumption according to which M-Words may only undergo morphological transformations (e.g. Lowering, Local Dislocation) with M-Words, and Subwords with Subwords. As a representative example of Local Dislocation, E&N discuss the case of the Latin conjunctive enclitic *-que* ‘and’ that

attaches to the first head of the second conjunct (Subwords of the second conjunct are separated by ‘.’).

- (10) vivimus vig.e.mus-**que**
live.1PL.PRS flourish.1PL.PRS-and
‘We live and flourish’

Given E&N’s assumption that Local Dislocation occurs under string-adjacency, we see that the rule affixing *-que* to the first head of the second conjunct is *type-sensitive*. Accordingly, the conjunctive enclitic *-que*, itself an M-Word, locally dislocates only with its string-adjacent M-Word *vigemus*, not the Subwords inside it. For this reason, interpolation of *-que* with the Subwords inside the complex head *vigemus* is not possible.

- (11) *Impossible interpolated forms*

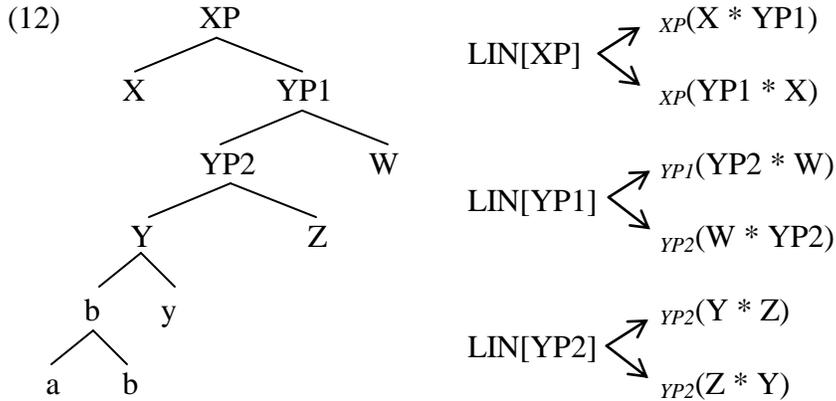
*vig-**que**.e.mus

*vig.e-**que**.mus

Returning to post-Vocabulary Insertion operations, the question to ask is what representations are input to these operations. Within the framework I adopt in this thesis, these operations are assumed to apply to statements of linearization. By hypothesis (E&N 2001, 2006; Embick 2003; Pak 2008), syntactic structures are converted into linear structures at Vocabulary Insertion through a set of *linearization procedures* so that they

can be input to performance systems. E&N (2001, 2006), following Sproat (1985), decompose Linearization into the two steps below:

A. STRUCTURAL ADJACENCY: The Linearization procedure LIN applies to each branching node of a structure and generates a statement of left-adjacency (*) between its daughters, in accordance with language-specific headedness of syntactic categories. Thus, for [X YP], the *-operator may generate either (X * Y) or (Y * X), where * is read as ‘is left-adjacent to’. This kind of linear procedure establishes relations among larger syntactic objects, i.e. M-Words (maximally complex heads like X, Y, Z) and phrases (XP, YP). The output *-statements of the following hypothetical structure, adapted from Pak (2008), are illustrated below:



Since * relation is established only between complex heads (M-Words) and phrases, PF must impose further order on the syntactic terminals so that they are instantiated in real time. For this to happen, we need statements that relate all the M-Words contained within the structure and all the Subwords (e.g. a, b, y in (12)) contained within M-Words. In the *-statement (X * YP1), for instance, the M-Word X is stated to be left-adjacent to the phrase YP1. This means that X appears to the left of the first

element of YP1. It is however not encoded in this statement what M-Word(s) contained within YP1 X is adjacent to. This kind of linear information is provided by a second operation called *Concatenation*.

B. CONCATENATION: This operation takes each *-statement as input and produces a corresponding concatenation statement between the peripheral M-Words. Concatenation statements are significant in the theory of Local Dislocation in that they structurally determine the environment in which adjunction under adjacency occurs. Embick (2003) formalizes the effect of this concatenation procedure as follows (\frown represents concatenation of M-Words):

$$(13) \text{ For } X(P) = [W_1 \dots W_n] \text{ and } Y(P) = [K_1 \dots K_n], \text{ where } W_i, K_i \text{ are M-Words, } (X(P) * Y(P)) \rightarrow (W_n \frown K_1)$$

The *-statements in (12) can thus be converted into concatenation statements as follows:

$$(14) \text{ a. } {}_{YP2}(Y * Z) \rightarrow Y \frown Z$$

$$\text{ b. } {}_{YP1}(YP2 * W) \rightarrow Z \frown W$$

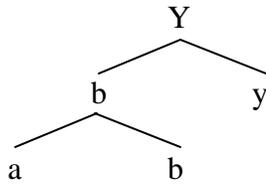
$$\text{ c. } {}_{XP}(X * YP1) \rightarrow X \frown Y$$

The M-Word Y in the structure in (12) is internally complex, i.e. it contains more than one Subword. In order to obtain complete linearization, the same procedures involved in linearizing M-Words (i.e. adjacency and concatenation) must apply to derive the linear order of Subwords (represented with \oplus). The effect of applying Concatenation

to Subwords, formalized in Embick (2003), is shown in (15) and the concatenation procedures applied to the Subwords of Y are illustrated in (16):

(15) For $X = [W_1 \dots W_n]$ and $Y = [K_1 \dots K_n]$, where W_i, K_i are Subwords, $(X * Y) \rightarrow (W_n \oplus K_1)$

(16) *Structure*



Linearization

- a. ADJACENCY (*): $((a * b) * y)$
- b. CONCATENATION (\oplus): $a \oplus b, b \oplus y$

Having discussed the procedures involved in linearizing syntactic structures, I now show that post-Vocabulary Insertion operations apply to the output of linearization procedures. For this purpose, I will briefly discuss Local Dislocation, a type of Morphological Merger operation that applies to Concatenation statements in a cyclic manner.

1.1.2.3 Local Dislocation

Given the system of Linearization above, Embick (2003: 13) proposes that Local Dislocation (henceforth LD) “is defined structurally in terms of Concatenation statements. Formally, the operation is one of adjunction under adjacency.” Concentrating on the Concatenation statement in (14)c, it meets the structural description for the LD rule that adjoins X to Y. Since adjunction is hierarchical in nature, Embick (2003)

suggests that the LD rule (i) deletes the Concatenation statement and (ii) introduces the hierarchical information that X is adjoined to Y. As such, X becomes a Subword in the derived structure.

$$(17) X \sim Y \begin{cases} \nearrow [[Y]X] \\ \searrow [X[Y]] \end{cases}$$

Since the adjunction of X to Y results in a hierarchical structure, we need the linear process to apply again and calculate * and \oplus -statements. The recalculation of the linear statements reveals the two effects that LD operations on M-Words⁵ bring about: (i) reordering; (ii) string-vacuous dislocation. In both cases the moving M-Word becomes a Subword but in the first case, additionally, the order of the two elements is reversed. These two effects are illustrated in the linear statements below:

(18) a. *Inversion*

$$*: (((a * b) * y) * X)$$

$$\oplus: a \oplus b, b \oplus y, y \oplus X$$

b. *String-vacuous LD*

$$*: (X * ((a * b) * y))$$

$$\oplus: X \oplus a, a \oplus b, b \oplus y$$

E&N (2001) propose that LD operations are restricted to apply within a particular domain such that the linear precedence and adjacency relationship between the elements in that domain is not disrupted. Embick (2003) suggests that such locality domains are

⁵ I am restricting my discussion to only LD operations on M-Words due to the irrelevance of LD operations on Subwords to the analysis that I present in this dissertation. See Embick (2003, §3.3.2) for a discussion of challenges that LD of Subwords pose for the theory of Local Dislocation.

provided by Concatenation statements within a given cycle of PF computation. As such, LD applies only to a pair of string-adjacent elements within the same PF cycle (phase); string-adjacent elements outside the domain cannot influence the LD operation and are thus inactive.

In her model of syntax-phonology interface, Pak (2008) discusses the possibility of phonological rules applying to Concatenation statements—as well as other kinds of linearization statements—as their domain. Apart from Adjacency (*) and Concatenation (\sqcap) operators that produce binary statements (i.e. that see two M-Words at a time), Pak introduces a linearization operation that generates statements containing an n -ary chain of M-Words internal to a given spell-out domain, namely *Chaining*. She further discusses *Late-Linearization* rules that apply across spell-out domains at the end of the derivation. I will review Pak’s model of syntax-phonology interface in more detail in Chapter 5 (§5.3.1). The underlying idea in Pak’s (2008) model is that the domains for the application of phonological rules “are automatically made available during the PF derivation as linear order is established over larger and larger objects; the phonology simply applies to the structure as it happens to exist at a given stage in PF” (p. 29). Accordingly, she predicts that the PF rules that apply to smaller, partially linearized syntactic objects operate earlier in the derivation than those that apply to fully linearized Chains at the end of the derivation.

1.2. Main proposals in brief

In this section I present a brief overview of the main proposals of this thesis. In so doing, I provide answers to the main questions of this study (posed in the introduction)

concerning the licensing of the empty noun, displacement of the plural marker *-hâ* in the ENC, and Ezafe insertion.

THE ENC: With respect to the licensing of the empty noun, my proposal is along the lines of those that link ellipsis to information structure (e.g. Rooth 1992a, 1992b; Gengel 2007, among others). I argue that the non-pronunciation of the head noun is linked to the information-structural properties of the ENC. Under this view, I propose that the ENC is derived through the interaction between the following two information-structural features: (i) the E-feature, which ensures that the head noun is identical with its counterpart in the antecedent and specifies the head noun for non-pronunciation; (ii) the focus feature F, which specifies the remnant modifier as an element which is in some kind of contrastive relationship with its corresponding element in the antecedent. The interaction between these two features is implemented in the syntax in a phase-based derivation. I suggest that within DM, the non-pronunciation of the head noun occurs at Vocabulary Insertion, where pronunciation of abstract morphemes, or the lack thereof, is determined.

I propose that the semantic and syntactic conditions on licensing the empty noun in Persian require the derivation to proceed in a phase-based manner. This assumption follows from the cyclic nature of licensing the empty noun which requires the empty noun (the E-specified head noun) and its licensor (the F-specified modifier) to be in the same cyclic (phase) domain, i.e. the *nP* domain. The *nP* phase is the domain where F-specification is assigned to an attributive adjectival head modifier in the context of an E-specified head noun. This head modifier is the only licensor to which the focus feature F is assigned. I assume that pronominal modifiers that license the empty noun when there is

no modifier inside the *nP* domain are not subject to F-specification process; rather, they are *inherently* F-specified. This proposal is shown to account for all instances of the ENC in Persian.

MOBILE PLURAL MARKING: I argue that the displacement of the plural marker in the ENC (cf. examples in (2)) cannot happen at syntax because *-hâ* is category sensitive and does not select for categories other than nouns. I suggest that this phenomenon can receive a plausible explanation if we assume that *-hâ* mobility takes place postsyntactically. On this view, the displacement of the plural marker is assumed to be motivated by the PF-imposed language-specific requirement for plural marking to appear on phonologically realized heads. I propose that *-hâ* mobility in the ENC is handled by Local Dislocation because the phenomenon is Vocabulary sensitive and involves linear precedence and adjacency. I show that the LD operation moves the stranded plural marker onto the nominal modifier that licenses the empty noun. Given that the empty noun is licensed in a phase-based manner, I propose that the LD operation for *-hâ* mobility applies in the same cycle in which the empty noun is licensed, moving the stranded plural marker onto the same remnant (head) modifier that licenses the empty noun.

In accounting for both the licensing of the empty noun and *-hâ* mobility, I show that only head, not phrasal, modifiers license the empty noun and host the stranded plural marker in the ENC. One important implication of this analysis is to provide further evidence in support of the proposal that postnominal modifiers in the Ezafe domain belong to two domains of modification: inner (*nP*) domain for head modifiers and outer domain for phrasal ones.

EZAFE INSERTION: Adopting Ghomeshi's (1997) view that Ezafe is a linking element inserted at PF, I propose that Ezafe insertion is determined at the interface of syntax and phonology. Within Pak's (2008) model of syntax-phonology interface, I argue that Ezafe insertion rule (EIR) is a phonological rule that applies at the Late-Linearization stage at PF. The rule inserts the Ezafe vowel onto an overt [+N] head and spreads it rightward onto the following [+N] M-Words that are modifiers/complements of the nominal head. I further argue that in the PF derivation of ENC's, EIR and the Local Dislocation (LD) operation responsible for *-hâ* mobility follow strict ordering such that LD applies early at Concatenation but EIR at Late-Linearization.

In this section, I summarized the main proposals made in this thesis. In the following section, I will present a background on the Persian noun phrase as the empirical focus of this thesis.

1.3. Background: The Persian noun phrase structure in brief

The structure of a prototypical Persian noun phrase has been thoroughly described in various sources, particularly in Samiiian (1983) and Ghomeshi (1996). I assume that (19) is the order of elements that appear in prenominal and postnominal positions of a Persian noun phrase:

(19) **DEM Q NUM CLS H_{noun} n a aP PP Possessor**

1.3.1 Prenominal position

Demonstratives, interrogatives, superlatives, super(lative)-ordinals⁶, and quantifiers are the elements that can appear in the prenominal position.

(20) a. <i>Demonstratives</i>	b. <i>Interrogatives</i>	c. <i>Superlatives</i>
in/un ⁷ ketâb	kodum ⁸ ketâb	boland-tar-in šab
this/that book	which book	long-COMP-SUP night
‘this/that book’	‘which book’	‘the longest night’
d. <i>Super-ordinals</i>	e. <i>Quantifiers</i>	
panjom-in ketâb	har ketâb	
fifth-SUP book	every book	
‘the fifth book’	‘every book’	

⁶ Persian ordinals are formed by adding the suffix *-om* or its phonological variant *-vom* to cardinal numbers: *yek* ‘one’ → *yek-om* ‘first’ (the suppletive alternant *avval* is more common); *do* ‘two’ → *do-vom* ‘second’. Super-ordinals are formed by adding the suffix *-in* to the ordinals. Unlike ordinals, super-ordinals precede the noun and do not take Ezafe

- | | | |
|---------------------|---|----------------------|
| (i) ketâb-e avval | → | avval-in(*-e) ketâb |
| book-EZ first | | first-SUP book |
| ‘the first book’ | | ‘the first book’ |
| (ii) ketâb-e se-vom | → | se-vom-in(*-e) ketâb |
| book-EZ three-ORD | | three-ORD-SUP book |
| ‘the third book’ | | ‘the third book’ |

I have chosen the term super(lative)-ordinals for this kind of ordinal because of the following similarities to superlatives: (a) both are morphologically marked with the same morpheme, i.e. *-in*; (b) both occur preminally; (c) the meaning contrast between ordinals and super-ordinals seems to correspond to a great extent to that between ‘positive’ adjectives and ‘superlatives’.

⁷ *Un* ‘that’ is the colloquial pronunciation of *ân*.

⁸ *Kodum* ‘which’ is the colloquial pronunciation of *kodâm*.

Among the above elements, only demonstratives can occur with superlatives, as shown in (21). Other than that, co-occurrence does not happen between the other pronominal elements.

(21) *Demonstratives + Superlatives*

mi-dunest-am in beh-tar-in lahze-hâ âxar-in lahze-hâ-st.
 DUR-knew-1SG this good-COMP-SUP moment-PL end-SUP moment-PL-is
 ‘I knew these best moments were the last moments.’

[www.khosravanameh.blogfa.com/post-33.aspx]

In (19), NUM represents cardinal numbers such as *yek* ‘one’, *do* ‘two’, *se* ‘three’, etc. that follow the pronominal elements listed in (20). The co-occurrence possibilities of NUM with other pronominal elements are shown in (22).

(22) a. *Demonstratives + NUM*

in se ketâb
 this three book
 ‘these three books’

b. *Interrogatives + NUM*

kodum se ketâb
 which three book
 ‘which three books’

c. *Superlatives + NUM*

beh-tar-in se tim
 good-COMP-SUP three team
 ‘the best three teams’

d. *Super-ordinals + NUM*

dovvom-in se emtiyâz
 second-SUP three point
 ‘the second three points’

e. *Quantifiers + NUM*

har se ruz

every three day

‘every three days’

Classifiers (CLS) in Persian never appear independently; they always occur with numerals. Following Samiiian (1983), I assume that classifiers are of three types: true classifiers, measure nouns, and group nouns. True classifiers are used with count nouns while measure nouns and group nouns are used with mass nouns.

(23) TRUE CLASSIFIERS

-tâ ‘unit’, used with all count nouns

nafar ‘person’

jeld ‘unit’, used for books

pors ‘unit’, used for meals

GROUP NOUNS

daste ‘bunch’

goruh ‘group’

fenjun ‘cup’

qâšoq ‘spoon’

MEASURE NOUNS

metr ‘metre’

kilo ‘kilogram’

litr ‘litr’

-tâ is the default classifier that simply functions to individuate the *nP*. It can optionally replace all true classifiers modifying count nouns (24)b; it however cannot co-occur with

them, as shown in (24)a'. It can co-occur with all group nouns to reinforce its individuation function (25). It should also be mentioned that, as a general rule, classifiers need not always appear (cf. Bateni 1969, Mahootian 1997, Gharib et al. 1994).

(24) a. se jeld ketâb a'.* se-tâ jeld ketâb
 three UNIT book
 'three books'

b. se-tâ ketâb

(25) a. se daste gol b. se-tâ daste gol
 three bunch flower
 'three bunches of flowers'

1.3.2 Postnominal position: the Ezafe domain

Unlike the prenominal elements, the constituents that occur after the head noun are linked with the Ezafe vowel; for this reason, the postnominal position is referred to as *the Ezafe domain* in the literature (see for example Samiian 1983, Ghomeshi 1997; the Ezafe domain is discussed in Chapter 5). The postnominal modifiers can be attributive nouns

(*n*), bare adjectives (*a*), *a*P_s, PP_s, and Possessors, all connected to the head noun (*H_n*) by the Ezafe vowel in a fixed order, as demonstrated in (26) and exemplified in (27)⁹:

(26) PRENOMINAL POSITION [*H_n* *n* *a* *a*P *PP* *Possessor*]

The Ezafe Domain

(27) a. *kif-e_n* [čarm]-e *a*[siyâh]-e *a*P[por az pul]-e *PP*[dâxel-e mâšin]

bag-EZ leather-EZ black-EZ full of money-EZ inside-EZ car

‘the black leather bag full of money inside the car’

b. *kif-e_n* [čarm]-e *a*[siyâh]-e *a*P[por az pul]-e *Possessor*[in mard]

bag-EZ leather-EZ black-EZ full of money-EZ this man

‘this man’s black leather bag full of money’

Possessors can be expressed in the Ezafe domain as *n*P_s, DemP_s, proper nouns, independent pronouns, or pronominal enclitics. I use *Possessor* in this thesis as a cover term for the categories that can be possessors (possessors are in italics).

⁹ Relative clauses (RC) in Persian also occur in postnominal position; however, they are (arguably) not connected to the head noun via Ezafe. Consider the following example:

(i) mard-i *RC*[ke savâr-e dočarxe bud]...
 man-?? that riding-EZ bicycle was
 ‘the man who was riding the bike...’

As can be seen in the above example, the relative clause is introduced by the suffix *-i* attached to the head noun. The nature of this suffix is still puzzling to Persian linguists. Ghomeshi (1996) and Samiian (1983) assume that the relative marker *-i* is the same morpheme that marks indefiniteness. Kahnemuyipour (2006), however, suggests that it is an allomorph of the Ezafe vowel. Since the dominant view in the literature is that suggested by Ghomeshi and Samiian, I will not discuss RCs in this thesis as part of the Ezafe construction.

(28) *Possessor categories*

a. *nP Possessor*

kif-e *mard-e javân*

bag-EZ man-EZ young

‘the young man’s bag’

b. *DemP*

kif-e *in dânešju*

bag-EZ this student

‘this student’s bag’

c. *Proper name*

kif-e *Sinâ*

bag-EZ Sina

‘Sina’s bag’

d. *Independent pronoun*

kif-e *u*

bag-EZ s/he

‘her/his bag’

e. *Pronominal enclitic*

kif-*aš*

bag-CL.3SG

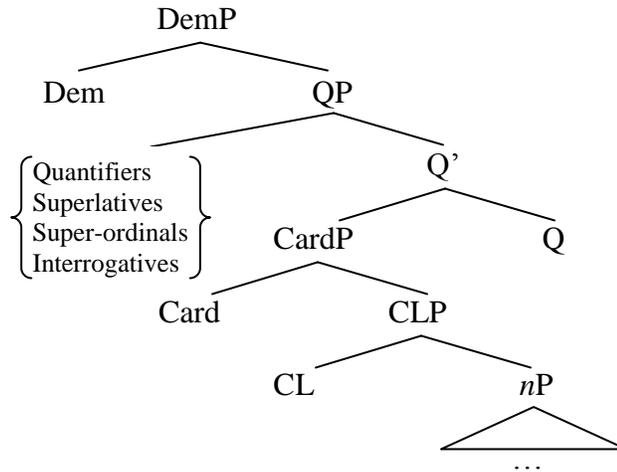
‘her/his bag’

I will not discuss pronominal possessors and proper names in this dissertation for two reasons: (i) the behavior of pronominal possessors with respect to the ENC is no different from non-pronominal possessors; (ii) pronominal possessors and proper names seem to behave like DPs, but this is not relevant to this study and I will use *Possessor* for all.

In this section, we surveyed the internal structure of the Persian noun phrase, discussing the pronominal and postnominal elements, their order of occurrence, and their co-occurrence restrictions. These elements and their order are schematized in (29):

among others, I propose the following syntactic configurations for the functional projections that dominate the NP in prenominal position:

(30) *Functional Projections*



In proposing this structure, I follow Alexiadou et al’s (2007: Ch.1, section 4) suggestion that the interpretation of the nominal projection is determined at two different levels. The highest layer of the nominal projection encodes discourse/pragmatic aspects of its interpretation and the lower layer expresses determination, i.e. definiteness, indefiniteness and so on. Under this view, demonstratives (i.e. DemP) encoding concepts such as familiarity, referentiality and deixis appear in the highest layer of the nominal projection¹⁰. I assume, along the lines of Ghomeshi (2003), that the Q layer is where indefiniteness is expressed. The non-cooccurring elements subsumed under Q in (29) are

¹⁰ Alexiadou et al. (2007) observe that in some languages like Greek, demonstratives may precede or follow the head noun. They argue that in such languages, prenominal demonstratives are interpreted deictically and postnominal or post-adjectival demonstratives are interpreted anaphorically. We shall not discuss the detail of their argument here as all interpretations of demonstratives in Persian are obtained in the same position, i.e. the highest layer in the nominal projection.

generated in the SpecQP. I reserve the head Q for the quantitative indefinite marker *-i* (see Ghomeshi 2003: section 4). Numerals and classifiers are further assumed to head their own functional projections: Cardinal Phrase (CardP) and Classifier Phrase (CLP), respectively.

1.4.2 The Ezafe Domain

With reference to the empirical observations presented in this thesis and the descriptive generalizations made in Samiiian (1983), Ghomeshi (1997), and Samvelian (2006), I assume that nominal modifiers in the Ezafe domain are distinct in terms of category and structural position. The attributive noun (*n*) and adjective (*a*) are head modifiers and are positioned in the *nP* domain along with the head noun; the transitive adjective (*aP*), the PP, and the possessors are phrasal modifiers and occur in a separate domain. I argue in Chapter 3 that the head modifiers in the *nP* domain are adjuncts to the uncategorized Root (cf. (32) below) because, among other things, they can be left stranded when the nominalized Root is not pronounced. I consider this as being consistent with Ghomeshi's (1997) proposal that nouns in Persian are non-projecting.

I propose in Chapter 4 that the *nP* domain, as the lowest layer of the nominal projection is where definiteness is determined. I argue that this follows from my proposal that in Persian, definiteness and number are syncretic. Elaborating, Ghomeshi (2008) proposes that in definite noun phrases, bare nouns are construed as bearing “a (phonetically) null [singular] feature and a null [definite] feature” (p. 93). In colloquial

Persian, the stressed suffix *-é* (*-á* before consonants), which optionally attaches to bare nouns, instantiates both the number and the definiteness feature (ibid. p. 94)¹¹.

(31) a. ali **ketâb-á-ro** xunde
Ali book-DEF.SG-OM read.PST.3SG
‘Ali has read the book

b. **ketâb-é** gerun bud
book-DEF.SG expensive was.3SG
‘The book was expensive.’

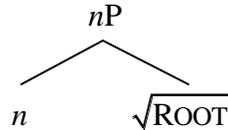
When the plural marker appears in such contexts, I suggest that it realizes both a [+pl(ural)] feature and a [+def(inite)] feature. In indefinite noun phrases, the occurrence of the plural marker is interpreted as instantiating only the number feature ([+pl]). I conclude from this argumentation that definiteness in Persian noun phrases is expressed via number markers. Applying Wiltscho’s (2008) diagnostics on the categorial status of plural markers to Persian (cf. §4.3.2), I propose that number markers in general are base-generated in the syntax as adjuncts to the nominalized Root. Based on this analysis according to which the definiteness of a noun phrase is determined at the point of the derivation at which number is assigned, I propose that Persian does not have a DP

¹¹ I have modified Ghomeshi’s gloss for the suffix *-é* so it reflects that the morpheme bears both the number and the definiteness feature.

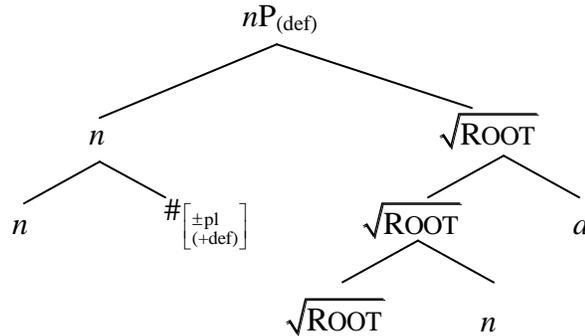
projection. Given our assumptions so far, the syntactic derivation of the *nP* can be represented as follows:

(32) *The nP derivation*

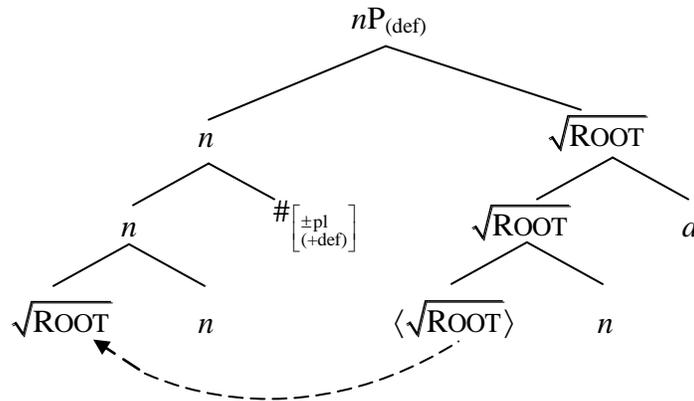
- a. Merger of an acategorial Root with the category-defining *n*:



- b. Base-adjunction of head modifiers to Root and the number node to *n*:



- c. Categorization:



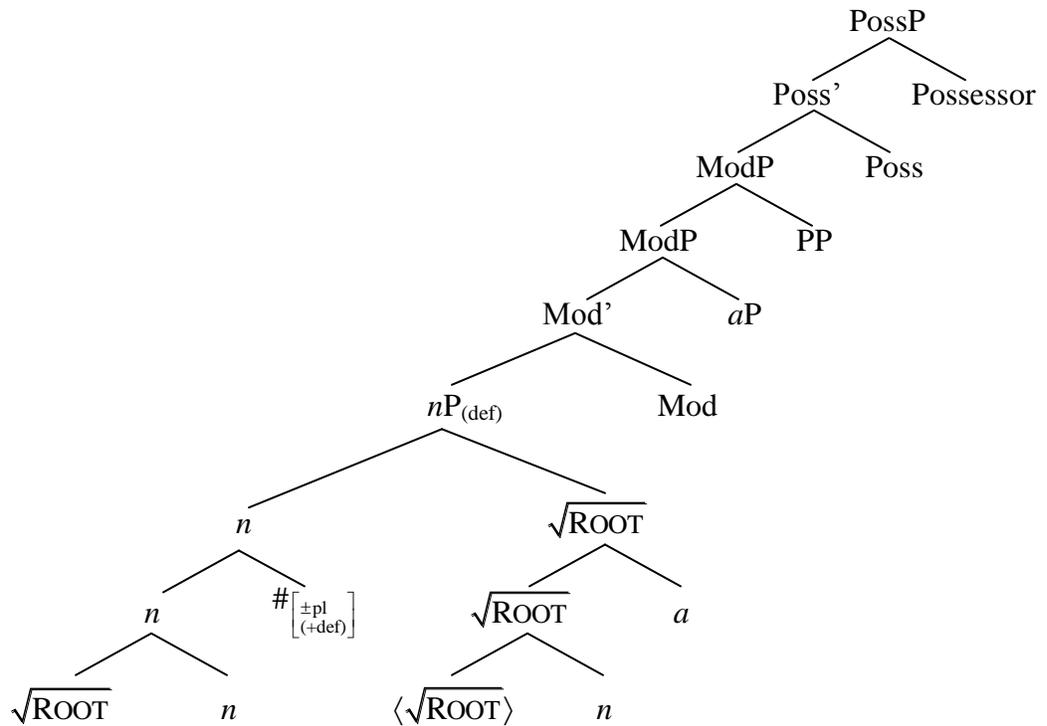
The structures in (32) represent how the noun and its head modifiers are derived in Persian within DM. The *nP* structure consists of a Root projection ($\sqrt{\text{ROOT}}$) and a

category-determining functional projection (*n*). Given our assumption that the modifiers inside the *nP* are heads, they are base-generated with the Root in a head-adjunction structure. The terminal Root node then head-adjoins to the category-forming functional head to become a noun. I adopt Adger's (2003) angled bracket notation $\langle \rangle$ to represent the moved constituent in the tree structure.

I now turn to projections that house phrasal modifiers. As noted earlier, I suggested that phrasal modifiers (viz. *aP*, *PP*, and *Possessor*) are generated outside the *nP* domain. Three motives are presented in this thesis that support this proposal: (i) only head modifiers can license the empty noun (Chapter 3); (ii) only head modifiers can host the stranded *-hâ* in the ENC (Chapter 4); and (iii) only head modifiers can enter into compounding with the head noun (Chapter 5). The fact that phrasal modifiers are not subject to any of the operations that target head modifiers indicates that they cannot be in the same domain with them. Given our assumption that *DP* does not exist in Persian, phrasal modifiers are generated within *ModPs*/*PossPs*. I propose *ModP* as an arbitrary projection that hosts phrasal modifiers other than possessive noun phrases (*Possessors*). *Possessors* are generated in the specifier of *PossP*. *ModP* is simply a place-holder for the maximal projection in the specifier of which phrasal modifiers are merged. The identity of this projection and the exact nature of its head remain to be explored in future research. I have not included the *PossP* as a constituent within *ModP* because I believe the head *Poss* is a Case assigner and that the specifier of the *PossP* is the only argument

position available in the Persian noun phrase¹². The structure of the Ezafe domain is schematically represented in (33):

(33) *The structure of the Ezafe domain*



So far in this chapter we have looked at the major contexts in which empty nouns can appear in Persian. More specifically, I have shown that empty nouns are possible with all prenominal elements, and that only attributive adjectives allow empty nouns in the Ezafe domain. Relying on the descriptive generalizations and other proposals, I have proposed syntactic representations which reflect properties of various elements inside the Persian noun phrase (cf. (30), (32), and (33)). The structural split between the head and

¹² See Chapter 5, ft. 88 for further explanation.

phrasal modifiers in the Ezafe domain is among the significant properties of the proposed structure of the Persian noun phrase. The assumptions concerning the syntax of the Persian noun phrase underlie the proposals set forth in this thesis. To end this chapter, I now give an outline of the chapters that follow in this dissertation.

1.5. Outline of the thesis

In Chapter 2, I review the two main approaches within which noun ellipsis has been discussed and analyzed in the literature, namely the *pro*-based approach, and the PF-deletion approach. The empty noun in the *pro*-based approach is analyzed as *pro*, subject to licensing and identification by a properly governing head specified for strong agreement. As such, the empty noun is assumed to lack internal structure. Contrary to this view, the PF-deletion approach assumes that elliptical constructions contain the same syntactic structure as non-elliptical ones. At interfaces, a complete mapping between the syntactic structure of the ellipsis site and its semantic representation occurs at LF and the terminal node for the head noun is not pronounced at PF. I adopt the PF-deletion approach in this thesis to explain the licensing conditions of the ENC. Under this approach, ellipsis phenomena are linked to information-structural features, contrastive focus in particular.

Chapter 3 lays out the main proposal of this dissertation concerning the licensing of the empty noun in Persian. I begin with a survey of the distribution of empty nouns in the Persian noun phrase. I establish that the empty noun is structurally a head because it allows for the stranding of the *nP*-internal head modifiers. To explain the empirical facts about the behavior of the empty noun in Persian, I present a phase-based account of the

ENC within the framework of Distributed Morphology. I argue that the licensing of the empty noun is the result of the interaction between the information-structural features E (specified on the empty noun) and F (assigned to one of the remnants). I show that my analysis accounts for a wide range of data involving empty nouns.

In Chapter 4, I present a detailed investigation of plural marking in Persian and its displacement in the ENC. I begin with a background on the distributional properties of nouns and plural marking in Persian. This sets the scene for introducing the distributional properties of plural marking in the ENC. The chapter continues with an analysis of the syntax of plural formation. I propose that definiteness and number are syncretic in Persian and that number markers are modificational rather than heading a functional projection. Turning to plural marking in the ENC, I argue that the displacement of the plural marker is motivated by a morphophonological requirement that *-hâ* be spelled out on a phonologically realized head. This requirement renders *-hâ* mobility Vocabulary-sensitive and thus a PF phenomenon. I propose that the PF operation responsible for handling the movement of the stranded *-hâ* onto an overt host is Local Dislocation (LD). This is due to the fact that only nominal modifiers immediately adjacent to the empty noun are the possible hosts. In Chapter 4, I expand the phase-based analysis of the ENC proposed in Chapter 3 so that it includes an account of *-hâ* mobility. I argue that *-hâ* mobility in the ENC can be derived from LD operations that apply to Concatenation statements generated over M-Words in a given cycle of PF derivation. I further suggest that the LD rule that adjoins the stranded *-hâ* to an overt adjectival head modifier applies on an earlier PF cycle than the one in which a prenominal modifier hosts the stranded *-hâ*. Thus, the application of LD to a Concatenation statement involving the stranded *-hâ*

and the attributive adjectival modifier bleeds the application of the procedure which attaches *-hâ* to an apparently string-adjacent prenominal modifier that is outside the LD domain. Finally, I show that this analysis can account for a wide variety of structures in which *-hâ* mobility occurs.

Chapter 5 incorporates the consequences of the analyses in Chapter 3 and 4 into a revised theory of nominal modification in the Ezafe construction. Ultimately, the purpose of this chapter is to account for the distribution of the Ezafe vowel in the ENC. I begin by investigating the syntactic properties of nominal modifiers in the Ezafe domain. Reviewing the syntactic accounts of these modifiers in the previous literature, I discuss the empirical inadequacies of these accounts and present my own proposal. I argue for a categorial and structural split between postnominal modifiers; namely, attributive nouns (*n*) and adjectives (*a*) are head modifiers and positioned in the *inner* nominal domain (*nP*) while other phrasal modifiers (*aPs*, *PPs*, and *Possessors*) are generated in the *outer* nominal domain (*ModP/PossP*).

Another important issue addressed in Chapter 5 is Ezafe insertion. Adopting Pak's (2008) model of syntax-phonology interface, I argue that the Ezafe vowel as a linker is inserted at the Late-Linearization stage when all the spell-out domains in a nominal structure are chained together. I show that the Ezafe insertion rule (EIR), unlike other PF operations that apply within a given cycle of PF derivation, applies across spell-out domains. As a Late-Linearization rule, EIR merges the *nP* spell-out domain with the *ModP/PossP* spell-out domain to form the modification domain where the Ezafe vowel connects the head noun to its following head and phrasal modifiers. I present EIR in this

thesis as further evidence in support of Pak's claim that any type of syntactic information is visible to phonological rules at all levels of their applications in the PF derivation.

Chapter 6 provides a brief summary of the thesis and discusses some areas of further research that will continue the current project.

It is my hope that this study will not only serve to provide new insight into the inner structure of the ENC in Persian, but also contribute to a fine-tuned understanding of two seemingly disparate aspects of the Persian noun phrase; namely, plural marking and Ezafe insertion.

Chapter 2

Previous Accounts of Empty Noun Constructions

2.0. Introduction

The Empty Noun Construction (ENC) is an understudied topic in Persian linguistics. In Persian traditional grammars, ellipsis¹ in general is basically defined as an anaphoric phenomenon in which a linguistic element is elided due to either the presence of a corresponding element in the co-text or the conception of its referent in the situational context (cf. Khanlari 1990, Gharib et al. 1994, Ahmadi-Givi 1985, Farshidvard 1969, among others). Traditional descriptions aside, there are two linguistics Master's theses by Kavousi Nejad (1992, 1998) and Fazlali (1994) that specifically focus on ellipsis in Persian. Both works approach ellipsis from a Hallidayan functional perspective in which ellipsis is viewed as a process contributing to the text cohesion. Among the cases of ellipsis they discuss in their works, viz. nominal, verbal, and clausal ellipsis, ellipsis in NP is discussed in the most detail. Although they discuss in detail the pre- and postnominal modifiers that can possibly appear in the ENC, they do not address the nature of the elided noun and the syntax of its modifiers. As accomplishing this is the main objective of this dissertation, I will not review the details of these studies. The focus of this chapter will instead be on the general literature on ellipsis.

¹ I refer to ellipsis in a pretheoretical way as an umbrella term for all kinds of phonologically null constituents.

Winkler (2006) considers studies on ellipsis as crucially revolving around three core questions: (i) how is ellipsis represented? (ii) how is ellipsis interpreted? (iii) how do information structure and focus contribute to the licensing and interpretation of ellipsis? In the literature, there are two main approaches to these questions each of which make different predictions: the *pro*-based accounts and the *PF-deletion* accounts. In this chapter, I will address both of these approaches in turn. In reviewing the *PF-deletion* accounts, I will particularly address the works that have linked ellipsis phenomena to information-structural features, most importantly, contrastive focus. I will then review the *DM*-based analyses of nominal ellipsis that will serve as a background to my analysis of the *ENC* in Persian in the ensuing chapter. I wrap up the chapter with a summary of the main ideas of the reviewed works and their relevance to this dissertation.

2.1. The *pro*-based accounts

Most of the analyses of empty nouns are built on the *pro*-based approach in which the rich morphology manifested on the remnant modifiers is the licensing mechanism for ellipsis (Lobeck 1991, 1992, 1995, Kester 1996, Sleeman 1996). Lobeck proposes that empty nouns are on a par with empty pronominals (*pro*)², licensed and identified by a properly governing head specified for strong agreement. She defines strong agreement as overtly realized morphological agreement in a productive number of cases. In English, for instance, plural demonstratives, quantifiers, and numerals are considered by Lobeck as being specified for the strong agreement feature [+Plural] on the grounds that number

² See Panagiotidis (2003: §1) for a detailed review of the *pro*-based approach to noun ellipsis.

agreement of the noun with these elements results in overt realization of the plural marker -s on the head noun (1)a. In the same vein, [+Possessive] is introduced as a strong agreement feature on D that is realized as 's and checked through Spec-head agreement relation with the possessor (1)b. Therefore, these elements meet the licensing and identifying conditions on empty nouns in English.

- (1) a. John calls on these students because he is irritated with [DP those_[+Plural] [students_[+Plural]]].
- b. Although John's friends were late to the rally, [DP Mary's_[+Poss] [friends]] arrived on time.

Lobeck further argues that elements not specified for strong agreement fail to license and identify empty nouns. The singular demonstrative in (2) is therefore not grammatical because it lacks strong agreement feature [+Plural].

- (2) *Although John doesn't like this new air conditioner that he got at K-mart, he likes [DP that_[-Plural] new [air conditioner_[-Plural]] that Mary bought at Sears].

Lobeck (1995) later suggests that quantifiers and numerals are further specified for [+Partitive] feature. This way she explains why *one* and *each* can license and identify empty nouns although they are [-Plural]. Thus, Lobeck (1991, 1992, 1995) proposes that the head of the DP must be specified as [+Poss], [+Plural], and/or [+Partitive] to license and identify empty nouns. Adjectives, (in)definite determiners, singular demonstratives,

and the quantifier ‘every’ cannot be licensors of empty nouns as they are either [–Plural] or [–Partitive] or both.

Unlike in English, adjectives are inflectionally marked for agreement in Germanic and Romance languages. In analyzing adjectives that allow empty nouns in such languages, Kester (1996) adopts Lobeck’s (1995) theory of ellipsis and Cinque’s (1994) assumption that APs are phrasal specifiers of functional heads in the DP. She claims that in Germanic and Romance languages, adjectives bearing morphological inflection are in a Spec-head agreement relation with a functional head specified for strong agreement. This strong agreement licenses the empty noun (*pro*) in these languages. The nominal elided constituent is identified with a lexical antecedent in the discourse.

Sleeman (1996), however, argues that there are other languages like French, Italian, and Spanish in which empty nouns do not seem to be licensed by inflection. She shows in the following example from French, that noun ellipsis is not possible even though the adjective *intéressante* is inflected.

(3) [*In the morning an interesting lecture and some less interesting ones were given*]:

*Malheureusement je n’ai pas entendu l’intéressante.

Unfortunately I neg have neg heard the interesting

‘Unfortunately, I have not heard the interesting one.’

[Sleeman 1996: 14.6]

Furthermore, Sleeman provides contrasting evidence that non-inflecting adjectives in French sometimes allow noun ellipsis.

(4) De ces robes, je préfère la vert foncé.

of these dresses, I prefer the green deep

‘Of these dresses, I prefer the deep green one.’

As can be seen in the above example, composed color adjective *vert foncé* ‘deep green’ does not inflect, but allows noun ellipsis. This observation poses an empirical problem for the approaches that assume that noun ellipsis is licensed only by nominal modifiers that carry rich morphological agreement.

The other evidence against the *pro* analysis of empty nouns is the fact that noun ellipsis is allowed even in languages lacking morphological agreement on nominal modifiers. Ntelitheos (2004) observes that in Malagasy, prenominal modifiers do not exhibit overt agreement with the nominal head that they modify; however, empty nouns are allowed to appear in the DP.

(5) Hitan’i Koto ny alike mainty ary Rasoana ny ~~alike~~ volontany.

saw Koto Det dog black and Rasoana Det dog brown

‘Koto saw the black dog and Rasoe saw the brown one.’

[Ntelitheos 2004:33.31a]

In (5), the adjectival modifier *volontany* ‘brown’ is not morphologically marked for agreement with the head noun; however, noun ellipsis is allowed.

Persian falls within the same type of languages in which nominal modifiers lack markers of morphological agreement with the head noun but this does not affect the licensing of empty nouns.

- (6) Sânz sag siâh-a-ro did o Simâ qahve'i-a-ro.
Sanaz dog black-DEF-OM saw.3SG and Sima brown-DEF-OM
'Sanaz saw the black dog and Sima [saw] the brown one.'

Based on similar evidence in English, Sleeman (1996) argues that the only relevant feature for licensing noun ellipsis is [+Partitive], thus dispensing with Lobeck's other features [+Poss] and [+Plural]. In her analysis, the empty noun is licensed by a [+Partitive] element properly governing *pro*.

For Sleeman, partitivity means *properly or improperly included within a set*. She assumes, for instance, that *some* and *all* both have a partitive meaning: *some* denotes a set that is *properly* included in another set, while *all* involves *improper* inclusion because it denotes a set that is identical to the superset. Although Sleeman's usage of partitivity in the sense of *proper or improper inclusion in a set* is the same as Enç's (1991), she departs from Enç's view that partitives always have a *specific* interpretation. She uses partitivity to refer to *non-specifics* as well. Consider the following Sleeman's examples in which the same partitive element (*two*) has specific (7) and non-specific (8) interpretations:

- (7) Several children entered my room. Two girls began to talk to me.

(8) - What have you done today? - I have bought two books.

[Sleeman 1996: 34. 75-76]

In (7), *two girls* has specific interpretation because it denotes a set that is included in a superset present in the domain of discourse, i.e. *several children*. However, *two books* in (8) does not denote information that is mentioned in the domain of discourse and thus has non-specific interpretation. Sleeman argues that although this noun phrase is non-specific, it denotes a subset that is formed out of the superset denoting the *kind*. She concludes from this argument that partitive elements are “potentially but not necessarily specific” (p. 34; see her Chapter 2, §8 for more detail).

Relating this to noun ellipsis, Sleeman claims that in French and in some other languages, the licensing of the empty noun is only possible if the empty noun occurs within a *specific* DP. She argues that specificity is required because DPs containing empty nouns must be associated with another DP in the context, otherwise the empty noun cannot be interpreted.

Building on the concept of proper or improper inclusion and its major role in the licensing of empty nouns, Sleeman makes a distinction between elements that always imply the inclusion in a set and elements that do not. She calls the first group *D-partitives* and the second group *N-partitives*. D-partitives (inherent partitives) are elements like quantifiers, superlatives, and ordinals that do not themselves denote properties and can therefore only be used as attributes to form a subset. These elements, according to Sleeman, allow empty nouns in all Romance languages and in English. N-partitives (non-inherent partitives), on the other hand, refer to elements that denote a property themselves

and can therefore be used as predicates. N-partitives denote a set themselves, which is not necessarily a subset. Sleeman considers certain adjectives of quality as N-partitives. The attributive use of N-partitives results in the intersection of the properties denoted by the adjective and those denoted by the noun. The intersection forms a subset of a superset denoted by the noun. According to Sleeman, supersets in noun ellipsis constructions are simple nouns referring to *kinds*. This is illustrated in the following example from Sleeman (1996: 45.109):

(9) Je n'aime pas ces robes vertes. Je prends la bleue.

I neg like neg these dresses green. I take the blue

'I do not like the green dresses. I will take the blue one.'

In this example, Sleeman shows that the superset is not formed by the DP *ces robes vertes* 'the green dresses', but simply by the noun denoting the kind, i.e. *robes* 'dresses'. She generalizes that "noun ellipsis always implies the inclusion of a subset in a previously established set denoting a kind" (p. 46).

Sleeman shows that in French a relatively small number of adjectives of quality are N-partitives. Other adjectives of quality, which do not allow empty nouns, are neither D-partitives nor N-partitives. In her analysis of noun ellipsis in English, Sleeman seems to suggest that D-partitivity vs. N-partitivity determines where the implementation of *one*-support is obligatory and where it is optional. Consider the following examples:

(10) *D-partitives*

- a. I will take these four. [Sleeman 1996: 51.129]
- b. This is the third (one). [Sleeman 1996: 51.130]
- c. Of these boys, I prefer the eldest (one). [Sleeman 1996: 51.131]

(11) *N-partitives*

- a. I only attended the interesting *(one). [Sleeman 1996: 51.127]
- b. I will take the small *(one). [Sleeman 1996: 51.135]

As can be seen in the above examples, *one*-support is optional with D-partitives and obligatory with N-partitives.

I now turn to the PF-deletion approach which has been adopted less predominantly in analyzing nominal ellipsis.

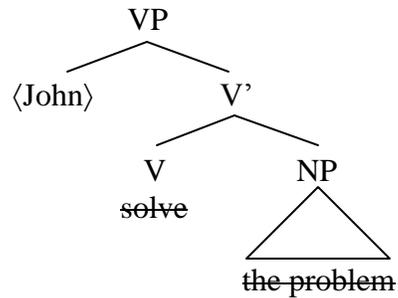
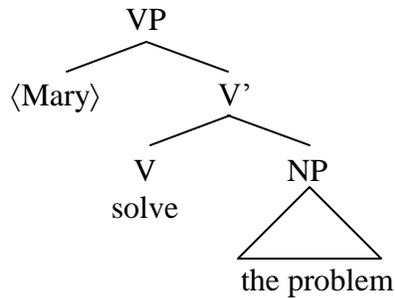
2.2. The PF-deletion accounts

The basic idea of the PF-deletion accounts (Ross 1969, Hankamer & Sag 1976, Chomsky 1995, Merchant 2001, among many others) is that the ellipsis site involves a fully-merged syntactic structure which is interpreted at LF and the phonological content of which is deleted at PF. Under this approach, the derivation of elliptical constructions proceeds identically to that of non-elliptical ones at syntax. When the output is delivered to the interfaces, a straightforward mapping between the syntactic structure of the ellipsis site and its semantic representation occurs at LF and the deletion of the given/redundant material at PF. This is illustrated in the following partial structure for VP Ellipsis.

(12) Mary will solve the problem and John will ~~solve the problem~~ too.

Mary will...

and John will... too



PF-deletion accounts, according to Winkler (2006), are based on the assumption that the syntactic structure of the ellipsis site has the abstract features of phonetic and phonological representation, but that the grammar contains a means of blocking its pronunciation.

2.2.1 Focus Condition on ellipsis

PF-deletion applies to redundant/given material, i.e. information recoverable from linguistic or pragmatic context. The deletion of the given material at PF is licensed under semantic and syntactic identity with corresponding parts of the antecedent clause—i.e. the sentence that contains the referent (antecedent) of the missing material. The remnants are associated in the recent literature with discourse functions, especially the expression of contrastive focus. As stated in the following hypothesis, the given or redundant material, which is subject to PF-deletion, motivates the expression of focus on the remnants (cf. Rooth 1992a, 1992b, Merchant 2001, among others):

(13) *Contrastive Focus Hypothesis:*

Given or redundant information licenses contrastive focus on the remnants.

The term *contrastive focus* evokes a choice of a subset from a suitable set of alternatives. This is reflected in Rooth's (1992a) focus analysis within the framework of Alternative Semantics. Braver (2007) has informally phrased the basic idea behind the Rooth's proposal as follows: if an expression *a* is contrastive, a membership set $M = \{\dots, a, \dots\}$ is generated, the members of which are determined both contextually and ontologically. For instance, the sentence in (14) reads as John having introduced no one other than Bill (among the members of *M*) to Sue.

(14) John introduced BILL to Sue

$M = \{\text{Bill, Carl, Mark}\}$

[Braver 2007: 68.134, 135]

Applying Rooth's proposal to N-drop in Spanish, Braver (2007) claims that nominal ellipsis brings about the contrast between the remnant modifiers and the modifiers of the antecedent. This is shown in (15), where the focus expression can be replaced by other suitable alternatives in (16).

(15) El [estudiante] de fonología [de Japón] y el [] de sintaxis

the student of phonology of Japan and the of syntax

'The phonology student from Japan and the syntax student (from Japan)'

(16) M={The syntax student from Japan, the syntax student from Mexico, the phonology student from Finland...}

[Braver 2007: 69.138, 139]

It has been argued in Rooth (1992a, 1992b) that there are requirements of parallel contrastive foci in elliptical processes. His (1992b) analysis of parallelism between VP ellipses and their fully spelled out but phonologically reduced counterparts is based on two central observations: (i) both types of constructions have the pragmatic function of expressing redundant information and (ii) redundant information in both constructions licenses contrastive focus. Rooth (1992b) shows in the following examples that the presence of a focus feature is correlated with prominence assignment. Both sentences in (17) license a contrastive focus reading of their subject *Mary* (capitals indicate prosodic prominence/focus).

(17) a. First John came up with a good idea, then [MARY]_F did [_{VP} *e*].

b. First John came up with a good idea, then [MARY]_F [*came up with a good idea*].

[Rooth 1992b: 4.4]

On the basis of such data, Rooth (1992b: 14) concludes that the “link between ellipsis and contrastive focus is tenable.”

Merchant (2001) follows the same line of reasoning, initiated by Rooth (1992a, 1992b) and pursued by Schwarzschild (1999), arguing that VP Ellipsis obeys a focus condition, which basically states that a VP can be elided if it is GIVEN in a context. A

constituent is GIVEN for Merchant if it is not in focus and if it has a particular type of salient antecedent in the context. Merchant (2001) specifically proposes that ellipsis be able to occur when the givenness relationship is established between the two conjuncts. He dubs this e(llipsis)-GIVENness. This condition is imposed on the elided constituent by what he calls the E-feature. The E-feature is defined in Merchant's (2001) analysis such that it has a particular semantics as well as a phonological effect triggering deletion. To put it informally, e-GIVENness is semantically a mutual entailment condition holding between the antecedent XP and the elided XP. That is, it is not just enough for the elided XP to be entailed by the antecedent XP, but the antecedent XP must be entailed by the elided XP. Phonologically, the E-feature, in Merchant's definition, triggers phonological deletion of the complement of the head on which it is placed. The E-feature instructs the grammar to delete the lexical material positioned below the head bearing the E-feature. The deletion takes place at the PF level.

Similar conditions for nominal ellipsis are reflected in Giannakidou & Stavrou's (1999) Contrast Condition of Nominal Ellipsis which states that "a nominal subconstituent α can be elided in constituent β only if the remnant of β is not identical to the corresponding part of the antecedent γ of α ." Under this view, the contrastivity of the remnant nominal modifier with its corresponding modifier in the antecedent DP is the licensing mechanism for having empty nouns. The following examples from Greek show that this is borne out (taken from Ntelitheos (2004: 36.34)).

(18) a. i Maria forese to [BLE]_F fustani ke i Eleni forese to [PRASINO]_F ~~fustani~~.
 the Maria wore-3SG the blue dress and the Eleni wore-3SG the green ~~dress~~
 ‘Maria wore the blue dress and Eleni wore the green one.’

b. *i Maria forese to ble fustani ke i Eleni forese to ble ~~fustani~~.
 the Maria wore-3SG the blue dress and the Eleni wore-3SG the blue ~~dress~~
 ‘Maria wore the blue dress and Eleni wore the blue one.’

We see in (18) that contrastivity is required not only for licensing nominal ellipsis but also for F(ocus)-marking of the remnant modifiers. (18)b is ungrammatical because the modifier in the second conjunct does not contrast with the one in the antecedent DP and thus no contrastive focus condition can apply.

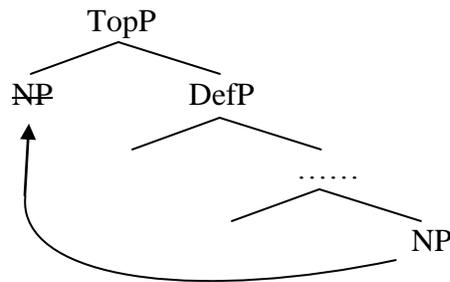
The general argument that most PF-deletion accounts share is that ellipsis is a derivational process that involves movement of the remnant followed by phonological deletion of the redundant material (e.g. Ross 1969, Merchant 2001). Merchant (2001), for instance, argues that sluicing in the following example is the outcome of the obligatory movement of the highest wh-phrase to SpecCP and deletion of TP at PF.

(19) Toby met someone at the party. Guess [_{CP} who [_{TP} Toby met ~~(who)~~ at the party]].

The link between deletion and focus has been made explicit in recent works by, for instance, Ntelitheos (2004) and particularly Gengel (2007) who adopt Jayasleen’s (2001) focus movement approach. In his analysis of nominal ellipsis, Ntelitheos (2004) follows

Johnson (2001) in assuming that ellipsis is licensed by topicalization. He assumes that nominal ellipsis involves two movement operations: (i) The DP-internal constituent NP, carrying redundant information, moves to the specifier of a nominal topic projection and gets deleted. This is schematically shown in the following structure.

(20) *Nominal Ellipsis as NP-topicalization*



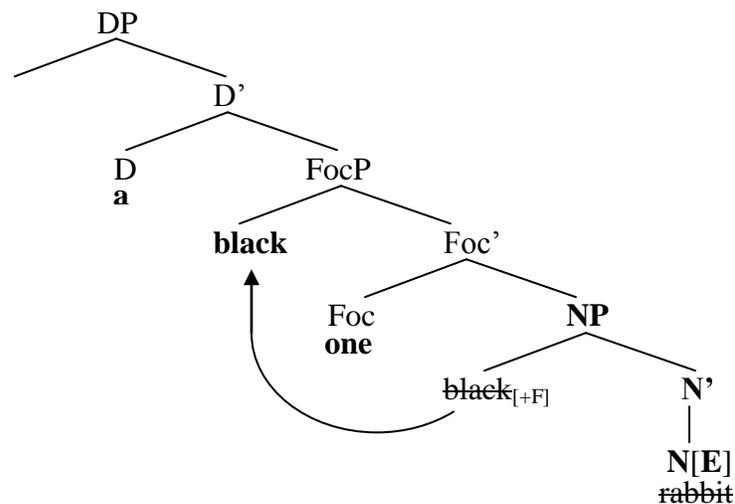
[Ntelitheos 2004: 9]

(ii) The remnant modifier, carrying new information and thus being focused, moves to the specifier of a nominal focus projection. Nominal ellipsis, according to Ntelitheos, is therefore the result of NP-topicalization and modifier focalization.

Gengel (2007) proposes an analysis of Pseudogapping, subsequently extended to other types of elliptical constructions, adopting the focus movement approach. For her, ellipsis in general and Pseudogapping in particular, is the outcome of the interaction between Merchant’s (2001) E-feature, and the Focus feature, as reflected in the works of Rooth (1992a, 1992b) and Schwarzschild (1999). To elaborate, she claims that ellipsis derives from interaction between the E-feature, which specifies a particular portion of the syntactic structure for phonological deletion, and the Focus feature through focus movement. Focus movement, in Gengel’s definition, “only obtains if it is necessary to

move contrastive material out of the phrase that is specified for deletion” (p. 288)³. In analyzing nominal ellipsis, Gengel (2007) assumes that the adjectival remnant moves to SpecFocP because it is focused and thus cannot remain inside the NP, which is specified for deletion. As can be seen in the following structure, the E-feature is specified on N which, as suggested by Gengel, could also be inserted directly on the head of the focus projection, to delete the complement.

(21) I saw a BLACK one yesterday.



Gengel (2007: 212) claims that the motivation for movement in all types of ellipsis is not only due to the presence of a [+contrastive] or [+focus] feature but to the presence of this particular feature within the deletion domain specified by the E-feature.

So far we have reviewed two major approaches to ellipsis, namely *pro*-based and PF-deletion. In reviewing the latter approach, the emphasis has been on works that have

³ Gengel (2007: 227) explains that in deaccented structures, there is no need to move the focused remnants out of their base positions.

particularly linked ellipsis to contrastive focus. Now we turn to reviewing the work that has proposed a PF-deletion account within the framework of Distributed Morphology.

2.3. Empty nouns and Distributed Morphology

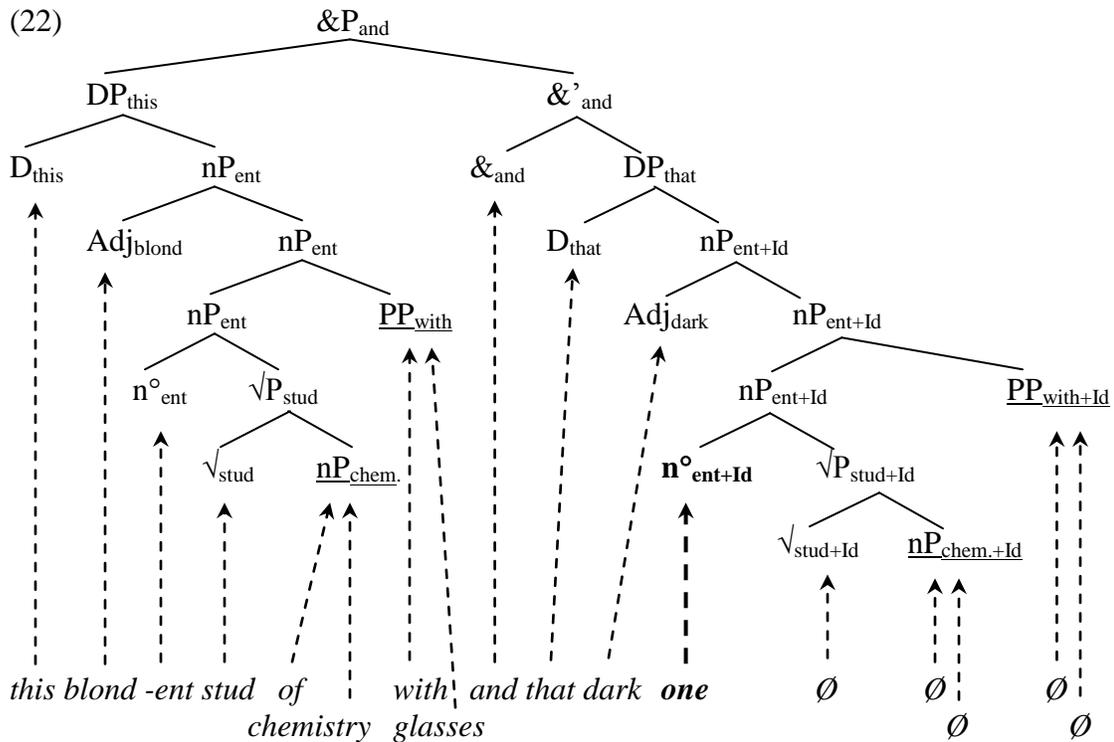
Distributed Morphology (DM), to the best of my knowledge, has not been widely used as the theoretical framework for analysis of ellipsis in general and nominal ellipsis in particular. In her DM account of *one*-replacement in English, Harley (2005: 14) treats ellipsis as “a failure to be pronounced” caused by “a failure of Vocabulary Insertion”. Under this view, the terminal nodes of ellipsis structures are all present at spellout, but are specified with a special feature that blocks the normal process of Vocabulary Insertion⁴. Harley calls this feature [+Id]. This feature ensures that a given terminal node in an ellipsis structure is formally *identical* and in an appropriate licensing relationship to its corresponding node in the antecedent clause. Harley assumes that this feature is bundled with other features that make up these nodes in the Numeration.

Recall from Chapter 1 that in DM, feature bundles representing individual terminal nodes in the syntax are realized at Vocabulary Insertion. Vocabulary Items specified for these feature bundles compete to realize these nodes. The winner of the competition is the Vocabulary Item that is more specified than other competing items.

Based on Harley’s approach, terminal nodes in ellipsis structures are specified with the [+Id] feature and thus require special Vocabulary Items. Assuming that most

⁴ This view is reminiscent of PF-deletion (e.g. Merchant 2001) that claims that the entire structure associated with the elided material is present throughout the derivation, but its identity with a salient antecedent licenses its non-pronunciation.

Vocabulary Items are specified as [-Id], she suggests that only few Vocabulary Items in English will contain the [+Id] specification: the null morpheme \emptyset (mainly), and *one*, specified for ([+count]) n° nodes. Thus, she proposes that in a *one*-replacement situation like (22), *one* blocks the insertion of \emptyset into the [+Id]-specified n° node because *one* is featurally more specific. The null morpheme \emptyset in turn blocks the insertion of [-Id] Vocabulary Items (phonologically overt items) into other [+Id]-specified nodes in the representation of conjoined NP *this blond student of chemistry with glasses and that dark one* in (22). In schematically representing the structure of this NP, Harley uses subscripts to represent the feature bundles associated with each terminal node, and the arrows to represent insertion of the winning Vocabulary Items into each node:



[Harley 2005: 15.26]

Kornfeld & Saab (2004)'s account of nominal ellipsis in Spanish is another example of analyzing ellipsis within DM. Unlike Harley (2005), they do not assign any special feature to the terminal nodes in ellipsis structures. They simply consider ellipsis “as a phenomenon of non-insertion of phonological features into terminal nodes, under strict identity of lexical and formal features” (p. 188). Based on empirical evidence from Spanish, they refine some ideas related to morphological identity, which I will not discuss here.

The common idea between the above-discussed two DM-based approaches to ellipsis is that the non-pronunciation of terminal nodes in ellipsis structures is decided at Vocabulary Insertion. In Harley's approach, Vocabulary Insertion is sensitive to the [+Id] feature specified on these terminal nodes and realizes them as \emptyset . In Kornfeld & Saab's approach, under strict identity, Vocabulary Insertion does not apply in ellipsis structures, although it is not clear from their account how this strict identity is established.

2.4. Summary and relevance to dissertation

In this chapter, we reviewed the accounts of ENC's within two main approaches to ellipsis in the literature: (i) the *pro*-based and (ii) the PF-deletion approaches. The proponents of the *pro*-based approach (e.g. Lobeck 1995) analyze the empty noun as *pro*, subject to licensing and identification by a properly governing head specified for strong agreement. Sleeman (1996) shows that this claim proves to be the case for some languages with inflection (e.g. the Germanic languages); however, there are languages like French, Italian, and Spanish in which empty nouns are not licensed by inflection. She proposes instead that *pro* is licensed and identified by a properly governing head specified with the

feature [+Partitive]. We mentioned that Persian falls within the type of languages that lack inflectional markers of agreement on nominal modifiers and yet allows empty nouns. As will be shown in the next chapter, the empty noun in Persian is an *n* head and thus cannot be *pro* because *pro* is assumed in the literature to represent a nominal phrase. *pro* is also assumed to lack internal syntactic structure. In Chapter 4, we will see that the requirement for the overt realization of the plural marker in Persian indicates that noun ellipsis in Persian requires internal structure and therefore cannot be *pro*.

The PF-deletion approach, as a reaction to the *pro*-based (non-deletion) approach, assumes that elliptical constructions contain the same syntactic structure as non-elliptical ones. At interfaces, a complete mapping between the syntactic structure of the ellipsis site and its semantic representation occurs at LF and the deletion/non-pronunciation of the phonological content of the relevant terminal nodes at PF.

The PF-deletion approach links ellipsis phenomena to information-structural features, contrastive focus in particular. Within this approach, which I adopt in this dissertation, ellipsis applies at PF to redundant/given material under (i) semantic and syntactic identity with corresponding material in the antecedent clause; and (ii) contrastivity of the remnant with its antecedent. In my feature-based account of the ENC in Persian in Chapter 3, I use Merchant's (2001) *E-feature* to specify the nominal root that goes missing after categorization. This feature also ensures that the nonpronounced material is in some sense identical with its corresponding antecedent (cf. §3.6.1.1). The contrastivity of the remnant modifiers with their corresponding antecedents is ensured by the *F(ocus) feature*, as reflected in the works of Rooth (1992a, 1992b) and Schwarzschild (1999). I will argue that the F-feature on remnant modifiers not only encodes

contrastivity, but it also entails that the F-specified elements are partitive, in the sense defined in Sleeman (1996). Partitivity, according to Alexiadou & Gengel (2008), is a prerequisite for contrastivity. Thus, the partitive feature fulfils a similar role in the identifying of elements as the focus feature because in both cases, reference is made to a set. Building on these assumptions, I propose, along the lines of Gengel (2007), that the ENC in Persian is the outcome of the interaction between the E-feature and the F-feature. This proposal will be elaborated in the next chapter.

Chapter 3

The Empty Noun Construction in Persian

3.0. Introduction

In Persian noun phrases¹, the head noun can be left unpronounced when there is an attributive adjectival head modifier in the Ezafe domain. If there is no such modifier, the empty noun is licensed by a prenominal modifier.

(1) a. Sajjâd pirhan-e âbi pušid, Sinâ ~~pirhan~~ a[qermez].

Sajjad shirt-EZ blue wore.3SG Sina ~~shirt~~ red

‘Sajjad wore a blue shirt, Sina a red one.’

b. Sajjâd behtarin durbin-o xarid, Sinâ ^{Supr}[badtarin] ~~durbin~~ -o.

Sajjad best camera bought.3SG Sina worst ~~camera~~-OM

‘Sajjad bought the best camera, Sina the worst.’

In this chapter I will investigate the syntactic and semantic properties of the Empty Noun Construction (ENC) in Persian. The goal is to provide an analysis of the ENC within the framework of Distributed Morphology (Halle & Marantz 1993, and subsequent

¹ I use the term *noun phrase* in this dissertation in a theory-neutral way to include all various nominal categories such as QP, CardP, *n*P, etc. I do not assume the DP-hypothesis because, as I discuss in the next chapter, the definiteness feature in Persian is syncretic with the number feature and this renders the assumption of DP untenable.

related works). The underlying claim is that the non-pronunciation of the head noun is linked to the information-structural properties of the ENC. Under this view, the ENC is derived through the interaction between the following two information-structural features: (i) the E-feature (adapted from Merchant 2001), which ensures that the head noun is identical with its counterpart in the antecedent and specifies the head noun for non-pronunciation; (ii) the focus feature F, which specifies the remnant modifier as an element which is in some kind of contrastive relationship with its corresponding element in the antecedent. The interaction between these two features² is implemented in the syntax in a phase-based derivation. It will be shown that the proposal offered here accounts for all instances of the ENC in Persian.

At the outset, we explore the contexts in which the empty noun is allowed in Persian. Following Ghomeshi (1997), I adopt the idea that nouns in Persian are non-projecting and can be modified by head-adjoined elements. I then show that empty nouns allow the stranding of head-adjoined modifiers. In §3.3, I motivate the information-structural features E and F; present my proposal for the analysis of the ENC; and apply the proposal to the Persian data. §3.4 highlights the main points of this chapter.

3.1. Where do we get empty nouns?

Noun ellipsis, according to Sleeman (1996: 13), “is the omission of a noun that can be recovered from the linguistic or extra-linguistic context.” For instance, *Peter’s* in (2) is

² I first came across the idea of ellipsis being the outcome of an interaction between the E-feature and the focus feature F in Gengel (2007). The mechanism of deriving the elliptical construction, however, is totally different in Gengel’s work compared to mine.

used instead of *Peter's camera*, because *camera* can be recovered from the linguistic context, i.e. *my own camera*.

(2) My own *camera*, like Peter's ~~*camera*~~, is Japanese.

In this section, we survey the possibility of having empty nouns in Persian with the noun phrase-internal constituents introduced in Chapter 1.

3.1.1 Prenominal position

The categories in the Persian noun phrase that appear prenominally all allow empty nouns, subject to a certain amount of category-internal idiosyncrasies and exceptions. Examples (3) to (5) show demonstratives, interrogatives, and superlatives allow empty nouns:

(3) *Demonstratives*

in ketâb-o be-de be man, [un ketâb]-o be-de be Simâ.
this book-OM IMP-give.2SG to I [that ~~book~~]-OM IMP-give.2SG to Sima
'Give this book to me, that [one] to Sima!'

(4) *Interrogatives*

kodum ketâb barâ man-e, [kodum ~~ketâb~~] barâ šomâ?
which book for I-is which ~~book~~ for you
'Which book is for me, which [one] for you?'

(5) *Superlatives*

gâh-i to fekr=mi-kon-i kar-i ke anjâm=mi-d-i
time-IND you thought=DUR-do-2SG work-IND that implement=DUR-do-2SG
[dorost-tar-in kâf]-e.
correct-COMP-SUP work-is

‘Sometimes you think what you do is the rightest.’

[bkm86.persianblog.ir/1385/3/]

Super-ordinals seem to occur with empty nouns only in formal Persian. In colloquial Persian, the adjectival form of ordinals, which is formed by adding the suffix *-i* to the ordinal, tend to appear instead.

(6) a. *Formal Persian*

dar in fasl faqat čâhâr šekast-râ pazirofte=bud-am va in
in this season only four defeat-OM accepted=was-1SG and this
[panjom-in šekast] bud.
fifth-SUPORD defeat was

‘In this season, I had suffered only four defeats and this was the fifth.’

[www.sharghnewspaper.ir/830515/sport2.htm]

b. *Colloquial Persian*

naqqâši-hâ-t hama-š qašang-e vali [naqqâši sevvom-i]³
painting-PL-CL.2SG all-CL.3SG beautiful-is but painting third-ADJ
/*sevvom-in naqqâši] binazir-e.
/ third-SUPORD painting unique-is
‘Your paintings are all beautiful, but the third [one] is unique.’

Most of the simple quantifiers and almost all the compound quantifiers can occur with empty nouns. I use the terms *simple* and *compound* quantifiers to make a morphological distinction between quantifiers that consist of one word and those that involve the compounding of a universal quantifier with words like *no* ‘kind’ and *gune* ‘type’ that denote kinds.

(7) *Simple quantifiers*

pedar kami pul be man va [kami ~~pu~~] ham be Sinâ dâd.
father a little money to I and a little ~~money~~ also to Sina gave.3SG
‘[My] father gave a little money to me and also a little to Sina.’

³ According to my intuition, in this context, the adjectival ordinal with the pronominal enclitic attached to it— i.e. *sevvom-i-š* [third-ADJ-CL.3SG]—is the most acceptable possibility. I presume that *sevvom-i-š* is the phonologically reduced form of *sevvom-in-aš* [third-ADJ-SUP-CL.3SG], which is an instance of a partitive construction. In (6)b, the pronominal enclitic is the pro-form for *naqqâši* ‘painting’. This raises the question whether there is an empty noun here or not. I leave aside this question for future research on empty nouns in Persian partitive constructions (cf. §6.2.2 for further details).

(8) *Compound quantifiers*

man bâ xošunat, [har=no xošunat] ke bâš-e, moxâlef-am.
I with violence every=kind violence that be-3SG opposed-am
‘I am against violence, whichever kind it may be.’

Numerals and classifiers—both true classifiers and group nouns—can also appear with empty nouns, as exemplified in (9):

(9) a. *Numerals + true classifiers*

Sinâ [noh šâxe [roz]] xarid: [se šâxe [roz sefid]], [se
Sina nine branch rose bought.3SG three branch rose white three
šâxe [roz zard]], o [se šâxe [roz qermez]].
branch rose yellow and three branch rose red
‘Sina bought nine branches of rose: three white, three yellow, and three red.’

b. *Numerals + group nouns*

har ruz bâyad [čâr qâšoq [šarbat]] bo-xor-am: [do qâšoq [šarbat]]
every day must four spoon syrup SBJ-eat-1SG two spoon syrup
sobh o [do qâšoq [šarbat]] šab.
morning and two spoon syrup night

‘I have to drink four spoons of syrup everyday: two spoons in the morning and two spoons at night.’

3.1.2 The Ezafe domain

As schematically shown in §1.4.2, the Ezafe domain consists of two head (i.e. *n* and *a*) and three phrasal modifiers (i.e. *aP*, *PP*, and *Possessor*). Of the two head modifiers, only attributive adjectives, simple and derived, license empty nouns; the phrasal modifiers do not appear with empty nouns.

Attributive nouns do not occur with empty nouns presumably because they cannot maintain their attributive function in the absence of the head; that is, they themselves may be mistakenly construed as the head.

- (10) a. [kif-e_n[čarm]] gerun-e. ⇒ b. * [kif-e_n[čarm]] gerun-e.
bag-EZ leather expensive-is ‘Leather bags is expensive.’
‘Leather bags are expensive.’

As reflected in the translation of (10)b, *čarm* is not interpreted as an attributive noun modifying *kif*; rather, it is interpreted as a nominal head. For this reason, speakers commonly use the derived adjectival forms of attributive nouns in the ENC. In so doing, the adjectival suffix *-i* is added to the nominal stem.

- (11) [~~kif-e~~_a[čarm-i]] gerun-e.
leather-ADJ
‘Leather [ones] are expensive.’

Languages behave differently with regard to the extent that adjectives allow noun ellipsis. Some languages, such as English and Italian, are more restrictive and some, like most Germanic languages, are more permissive⁴. Persian falls into the more permissive type of languages in this respect, in that all kinds of adjectives can seemingly occur in the noun ellipsis construction⁵.

(12) A: [xodkâr-e sabz] dâr-i?
 pen-EZ green have-2SG
 ‘Do you have a green pen?’

B: [~~xodkâr~~ sabz] na-dâr-am, [~~xodkâr~~ âbi] mi-xâ-y?
 pen green NEG-have-1SG pen blue DUR-want-2SG
 ‘I don’t have a green [one], do you want a blue [one]?’

⁴ See Sleeman (1996) for an account of how adjectives in English, Dutch, Italian, and Spanish behave differently from French in allowing empty nouns.

⁵ Not all adjectives co-occur with empty nouns. The adjectives licensing empty nouns bear certain semantic properties that distinguish them from those that disallow empty nouns. Sleeman (1996) reports that, contrary to French and Sardinian, the licensing of empty nouns is generally impossible with adjectives of quality in languages like English, (some varieties of) Italian, and Spanish. These two groups of languages, though, share the possibility of combining empty nouns with quantifiers and pronouns. Accordingly, she proposes the following hierarchy of classes of licensing elements in the ENC in terms of the extent to which they are more likely to license empty nouns:

Quantifiers, Pronouns, Superlatives, Ordinals > Color adjectives > big-small; good-bad; old-young, new > All other adjectives of quality

Persian tends to behave like French and Sardinian (i.e. more permissive languages), however I will leave an exploration of the semantic properties of Persian adjectives licensing empty nouns for future research (cf. §6.2.1 for more details).

Phrasal modifiers in the Ezafe domain, viz. *aP*, *PP*, and Possessor, do not license empty nouns⁶, as shown in (13) to (14).

(13) *Adjective Phrase*

a. *nP*[kešvar-hâ-ye *aP*[negarân-e afzâyeš-e qeymat-e naft]] ettelâ'iyye-i
 country-PL-EZ worried-EZ increase-EZ price-EZ oil statement-IND
 sâder=kard-and.

issue=did-3PL

‘The countries worried about the increase in oil prices issued a statement.’

b. **nP*[~~kešvar-hâ~~ *aP*[negarân-e afzâyeš-e qeymat-e naft]] ettelâ'iyye-i sâder=
 kard-and.

‘*Concerned of the increase in oil prices issued a statement.’

⁶ The only exception to this generalization is *aPs* containing comparative adjectives. Consider the following example:

(i) kif-hâ-ye *aP*[*a*[kamtar *PP*[az sad dolâr]]-emun tamum=šod-e.
 bag-PL-EZ less than hundred dollar-CL.1PL end=became-PP.3SG
 ‘Our less-than-hundred-dollar bags are sold out.’

As can be seen in (i), the basic comparative *aP* consists of a comparative adjective and a *PP*. The prepositional complement in the comparative *aP* can be scrambled to the front of the comparative adjective, as shown in (ii).

(ii) kif-hâ-ye *PP*[az sad dolâr] kamtar]]-emun tamum=šode.

Both of the basic and scrambled comparative *aPs* allow empty nouns and can take the default plural suffix *-hâ* (cf. Chapter 4 for an analysis of plural displacement in the Persian ENC).

(iii) a. ~~kif~~ [kamtar [az sad dolâr-*i*]]-â-mun tamum=šode.
 -ADJ-PL-CL.1PL

b. ~~kif~~ [az sad dolâr] kamtar]]-â-mun tamum=šode.
 ‘Our less-than-hundred-dollar [ones] are sold out.’

This issue will not be addressed in this thesis as its analysis requires research beyond the scope of this study.

(14) *Prepositional Phrase*

a. n_P [kafš-â-ye PP [tu(-ye) vitrin-e maqâze]] xeyli qašang-e.
shoe-PL-EZ inside(-EZ) window-EZ shop very beautiful-is
'The shoes inside the window shop are very beautiful.'

b. * n_P [~~kafš-â~~ PP [tu(-ye) vitrin-e maqâze]] xeyli qašang-e.
'*(The ones) inside the window shop are very beautiful.'

(15) *PossP*⁷

a. n_P [kif-e $PossP$ [$Possessor$ [in xânom]]]-o man peydâ=kard-am.
bag-EZ this lady -OM I found=did-1SG
'I found this lady's bag.'

b. * n_P [~~kif~~ $PossP$ [$Possessor$ [in xânom]]]-o man peydâ=kardam.
'*I found this lady.'

So far, we have described the possibility of empty nouns appearing with noun phrase-internal constituents. Let us now turn to the structural properties of the empty noun.

⁷ Recall from Chapter 1 (§1.4.2) that Possessors are generated in the specifier of PossPs. We will motivate this assumption further in §5.1.2 and §5.3.3.2.

3.2. The structural status of the empty noun

I assume that the empty category in the Persian ENC corresponds to the head noun. This assumption is motivated by the fact that the empty noun allows for the stranding of the *nP*-internal attributive adjective, the modifier that is head-adjoined to the nominal Root and licenses the empty noun.

A common instance of the ENC contains a missing head noun with an attributive adjective as the remnant.

- (16) Sajjâd sib-e sabz-o xord o *nP*[~~sib~~ qermez]-o be Sinâ dâd.
Sajjad apple-EZ green-OM ate.3SG and ~~apple~~ red -OM to Sina gave.3SG
'Sajjad ate the green apple and gave the red one to Sina.'

Attributive adjectives can appear recursively in the noun phrase. In such cases, more than one adjective may go missing along with the head noun:

- (17) Sinâ kif-e čarm-i-ye kučik-o bâ *nP*[~~kif~~—~~čarmi~~ bozorg]
Sina bag-EZ leather-ADJ-EZ small-OM with ~~bag~~—~~leather~~ big
avaz=kard.
exchange=did.3SG
'Sina exchanged the small leather bag with a big one.'

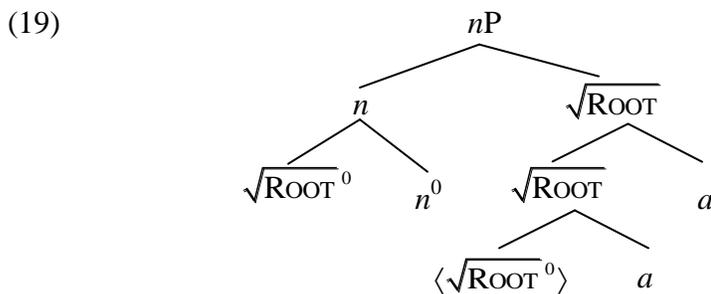
The head noun can also be elided with all its adjectival modifiers if there is a prenominal modifier outside the *nP*.

- (18) Sinâ in kif-e čarm-i-ye kučik-o bâ *DemP*[un *nP*[~~kif~~—čarmi
 Sina this bag-EZ leather-ADJ-EZ small-OM with that ~~bag~~—leather
~~kučik~~] avaz=kard.
 small exchange=did.3SG

‘Sina exchanged this small leather bag with that one (= small leather bag)⁸.’

We see, then, that at least one adjective must be pronounced unless there is a pronominal modifier in which case it can serve as the remnant and all the adjectives can go missing.

The structure that I have proposed in Chapter 1 for the Ezafe domain suggests that the empty category is not a maximal projection. In this configuration, *nP*-internal modifiers are heads that adjoin to the Root of the head noun to form a complex head. The *nP* structure with recursive adjectives looks like the following in the ENC (I use the ‘⁰’ diacritic to distinguish terminal head nodes from non-terminal ones):



As I will discuss in §3.3.1, the stranded adjectives should be informationally salient; otherwise, they will not be pronounced in the ENC. Given the configuration above, I

⁸ In Persian and English, the elided part in (18) can be interpreted in other ways, given the context.

assume that the empty category is always a head regardless of whether the head noun goes missing with or without its attributive modifiers. Thus, this generalization automatically extends to sentences like (18) in which the whole *nP* is not pronounced and the empty noun is licensed by a prenominal modifier, i.e. a demonstrative, quantifier, superlative, or a super-ordinal.

Attributive adjectives that are the possible remnants in the ENC are intransitive, in the sense that they do not take complements. This is the reason that I assume, following Ghomeshi (1996, 1997), that they, along with the attributive nouns, are heads (see also Samiian 1983, chapter 2 for more empirical evidence). Adjectives that take complements, i.e. transitive adjectives, also modify nouns; however, they cannot occur with the empty noun because they are phrases.

- (20) *~~barâdar~~ _{AP}[âšeq-e futbâl]-e man] bâz rafte estâdiom.
~~brother~~ enthusiastic-EZ soccer-EZ I again gone stadium
 ‘My mad-about-soccer [one] has gone to the stadium again.’

As outlined in Chapter 1, we class transitive adjectives together with PPs and Possessors as the nominal modifiers that are generated in the specifier of ModP and PossP. I claim that head modifiers, including attributive adjectives, occur within *nP*, in the same domain as the head noun and share its head status. The head status of *nP* modifiers follows from the assumption that the nominal Root is a non-projecting head to which the *nP* modifiers

are head-adjoined⁹. The fact that head modifiers and not phrasal modifiers license the empty noun supports a two-domain theory of nominal modification according to which head and phrasal modifiers belong to two different syntactic domains. Other factors reinforcing the plausibility of this assumption will be introduced and discussed in Chapter 5.

3.3. Towards an analysis

In this section, I propose an analysis of the ENC in Persian within the theoretical framework of Distributed Morphology (Halle & Marantz 1993, and subsequent related work). According to DM, the phonological expressions of abstract Roots and morphosyntactic features are supplied postsyntactically at Vocabulary Insertion. In this sense, following Harley (2005) and Kornfeld & Saab (2004), I argue that ellipsis is a PF phenomenon within DM, in which the pronunciation of abstract morphemes, or the lack thereof, is determined at Vocabulary Insertion. What determines the non-pronunciation of the head noun in the ENC, I propose, is the interaction between the information-structural features assigned to the empty noun and its remnant. This way, my proposal is along the lines of those that link ellipsis to information structure (e.g. Rooth 1992a, 1992b; Gengel 2007, among others).

⁹ The idea of nouns being non-projecting in Persian was first discussed in Ghomeshi (1997).

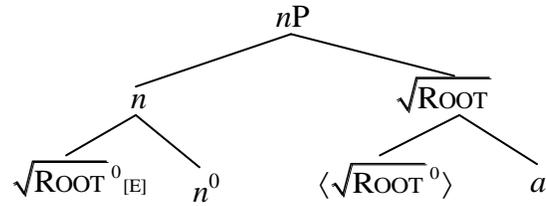
3.3.1 Information-structural configuration of the ENC

In the review of focus conditions on ellipsis in the previous chapter, we saw that there are two principal requirements in licensing ellipsis: (i) ellipsis applies to the material that has already been introduced in the discourse context; (ii) the remnant evokes some sort of contrastive relationship with the antecedent. Merchant's (2001) E-feature specifies the relevant syntactic structure that goes missing and ensures that the nonpronounced material is in some sense identical with its corresponding element in the antecedent clause. The focus feature F, on the other hand, is assigned to the remnant, which represents new or contrastive information. Let us now explore how these two features (E & F) are implemented in the ENC in Persian.

3.3.1.1 The E-feature

We established in §3.2 that the nominal Root of a noun phrase in Persian is non-projecting and that the attributive modifiers are heads adjoined to it (cf. (19)). We also showed that non-pronunciation (ellipsis) targets the *n* Root and allows the head modifier adjuncts to be left stranded. This observation was taken as evidence in support of our general assumption about the head status of the *n* Root in the Persian noun phrase structure. Building on these assumptions, I propose that the E-feature specifies that the nominal Root is to be skipped for the purposes of pronunciation. Following the categorization assumption in DM (e.g. Marantz 1997, 2001, among others), I propose that at syntax, the uncategorized $\sqrt{\text{ROOT}}$ moves to adjoin to the category-defining functional head *n* to become a noun. The E-feature is assigned to the Root that is sister to the *n* head.

(21)



The relevant generalization can be stated as follows:

(22) *E-feature specification in the Persian ENC*

The E-feature is specified on the Root that is sister to the categorizing n^0 head.

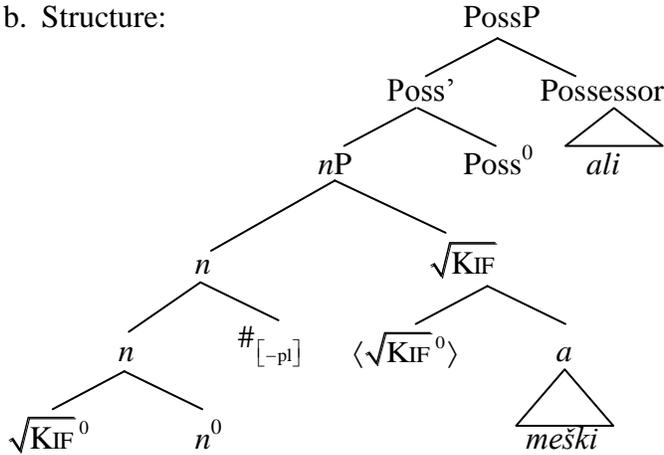
I assume that at Morphology the E-feature blocks the Vocabulary Insertion from inserting the phonological features of the head noun, i.e. $\sqrt{\text{ROOT}}^0_{[E]} - n^0$.

We observed in examples (17) and (18) above that in addition to the head noun, one or more head modifiers can also be left unpronounced. This raises the question as to what mechanism is involved in the non-pronunciation of the head modifiers in the ENC. As I will argue shortly, the sister Root to the categorizing n^0 head can be a complex Root containing the terminal Root node and one or more head modifiers. The E-specification of the complex Root is interpreted at PF as non-pronunciation of the constituents contained in it, namely, the terminal Root and the head modifier nodes.

Recall that the nominal Root and the head modifiers inside the nP form a head-adjunction structure. This structure consists of an uncategorized Root to which attributive adjectives are head-adjoined. The uncategorized Root undergoes head movement to the category-forming n head to become a noun. At PF, the nominalized Root and the head modifiers are spelled out with Ezafe linking them together. This process is shown below:

- (23) a. kif-e meški-e ali
 bag-EZ black-EZ Ali
 ‘Ali’s black bag’

b. Structure:

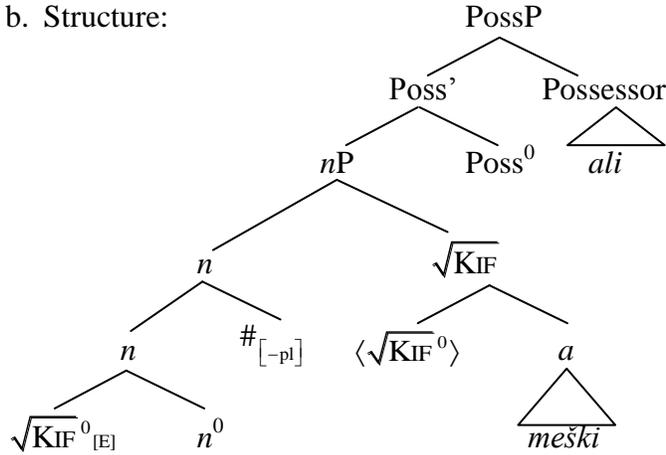


- c. PF: kif-e meški-e ali

In the ENC, the sister Root to the nominalizing n^0 head is specified with the E-feature, in accordance with (22). At PF, this specification results in the non-pronunciation of the nominalized Root only.

- (24) a. ~~kif~~ meški-e ali
 ~~bag~~ black-EZ Ali
 ‘Ali’s black one’

b. Structure:



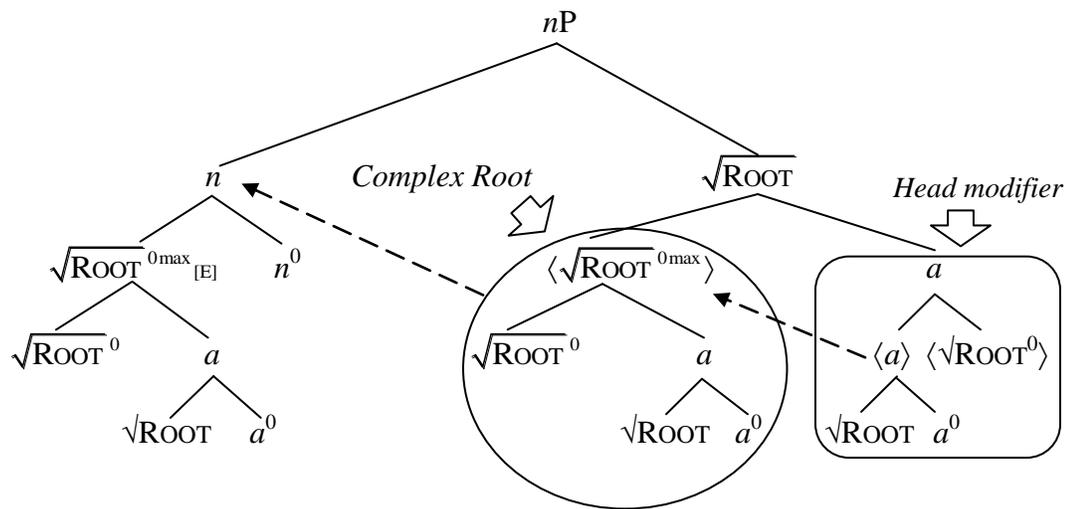
c. PF: meški-e ali

In cases where the head noun is missing along with its head modifier (i.e. attributive adjective), I propose that the modifier has undergone incorporation into the uncategorized Root.

Recall that within DM, adjectives as lexical categories are assumed to be internally complex, involving Roots that become adjectives by merging with the head a^0 (Marantz 2001). I propose, along the lines of Harley (2008), that the modifier-Root incorporation inside the nP involves first the categorization of the adjectival Root and then adjunction to the nominal Root. The resulting structure is a complex Root which adjoins to the n head for categorization. As such, the complex Root node is sister to the nominalizing n^0 and is thus specified with the E-feature. The E-specification of the complex Root is interpreted at PF as non-pronunciation of the constituents contained within the complex Root, i.e. the Root and the categorized adjective (head modifier). In the context of having more than one head modifier, the complex Root would consist of more than one adjectivized Root. I assume that the complex Root so formed is an incorporation structure in the sense that the meaning of the head modifier(s) becomes part of the meaning of the

Root¹⁰. The modifier-Root incorporation process is schematically shown below; (I use the $\langle 0_{\max} \rangle$ diacritic to mark complex Roots formed by incorporation):

(25) *Modifier-Root incorporation inside the nP*



Let us now consider how this proposal works out in the more complex examples in

(26) and (27).

(26) Sinâ kif-e čarm-i-ye kučik-o bâ ENC[kif—čarmi bozorg]

Sina bag-EZ leather-ADJ-EZ small-OM with ~~bag—leather~~ big

avaz=kard.

exchange=did.3SG

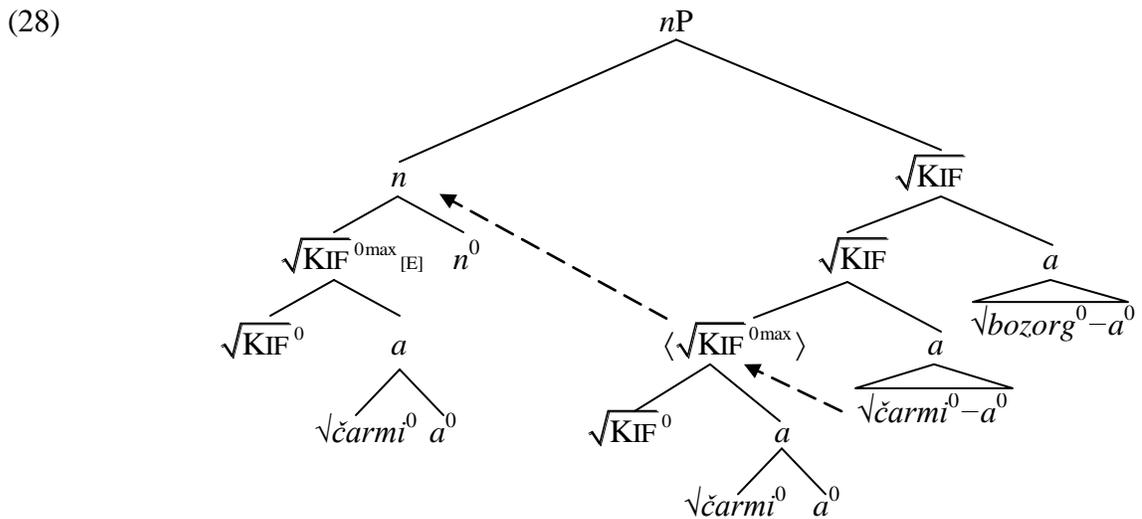
‘Sina exchanged the small leather bag with a big one.’

¹⁰ *Syntactic compounding* is the term I use to refer to this kind of incorporation. See §5.1.3.2 for more discussion on this type of compounding.

- (27) Sinâ in kif-e čarm-i-ye kučik-o bâ DemP[un nP[kif—čarmi
 Sina this bag-EZ leather-ADJ-EZ small-OM with that bag-leather
 kučik] avaz=kard.
 small exchange=did.3SG

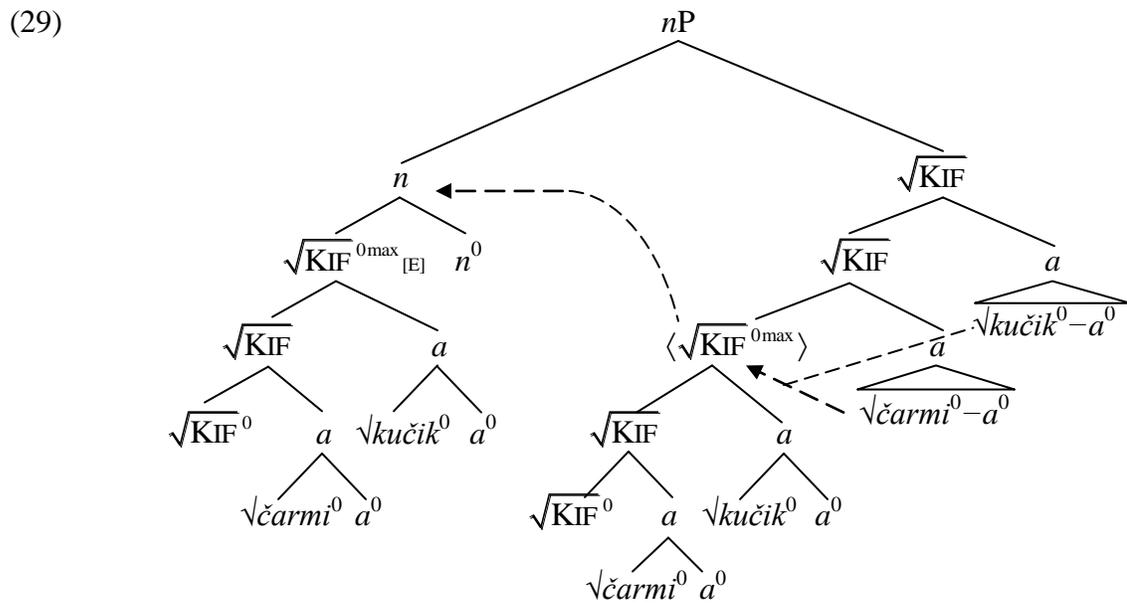
‘Sina exchanged this small leather bag with that [one].’

The derivation of the ENC in (26) can be syntactically represented as follows:



In the above derivation, the adjectivized Root *čarmi* ‘of leather’ first incorporates into the nominal Root *kif* ‘bag’ forming a complex Root. The complex Root then adjoins to the categorizing *n* head, where it is specified with the E-feature. The presence of the E-feature on the complex Root node sister to the n^0 head indicates that the phonological exponents of the terminal nodes contained within it will not be inserted at Vocabulary Insertion.

The derivation of the ENC in (27) involves incorporation of all the head modifiers into the nominal Root. The resulting complex Root is then categorized by adjoining to the n head. The subsequent E-specification of the complex Root in the sister position of the n^0 head specifies that the Root and all its head modifiers will skip pronunciation at PF. As the bracketed structure in (27) shows, the empty noun in this context is licensed by the prenominal demonstrative. This allows all the the head modifiers inside the nP domain to be left unpronounced. For conciseness, I have excluded the superstructure of the nP in (29).



The above account for the non-pronunciation of head modifiers in the ENC further supports our claim that the empty noun in the Persian ENC is a head. We have seen that the non-pronunciation of the nominalized Root along with its head modifiers does not affect the head status of the empty noun because in such cases the E-feature is specified on a complex Root comprising of the Root and the modifiers.

Having proposed the syntactic implementation of the E-feature, I will now turn to its semantics.

Semantics of the E-feature

Recall from Chapter 2, §2.2.1, that Merchant’s definition of the E-feature involves a particular semantics, in addition to its phonological effect, namely non-pronunciation. There we discussed that the semantics of the E-feature imposes a mutual entailment condition between the antecedent and the elided material. Merchant’s proposal for the semantics of the E-feature was intended to account for cases of ellipsis in the verbal region. I assume that the semantics of the E-feature in the ENC involves an *identity of sense* condition on the empty noun and its antecedent. In making this assumption, I am following Grinder and Postal (1971) who propose that ellipses are Identity of Sense Anaphora (cited in Lobeck 1995: 35)¹¹. Looking at the example (30) below, we see that the empty noun and its antecedent do not refer to the same entity, although they share the same sense.

- (30) Sajjâd sib-e sabz-o xord Sinâ [~~sib~~ qermez]-o.
Sajjad apple-EZ green-OM ate.3SG Sina ~~apple~~ red -OM
‘Sajjad ate the green apple, Sina the red [one].’

¹¹ Also see Williams (1997) for considering the notion of deletion as an instance of anaphora. For him, deleted and destressed materials are *nonintrinsic* anaphors. Such anaphors are “*contextually defined or identified*” (p. 594).

The empty noun and its antecedent in (30) share the meaning of *sib* ‘apple’; however, they refer to different tokens of the type *sib*. Thus, we can conclude that although the empty noun is in an anaphoric relationship with its antecedent, it involves sloppy reference¹².

The identity of sense condition has bearing on the semantic interpretation of the empty noun. Within DM, semantic interpretation takes place at LF. I propose that at LF, the E-feature, as its semantic function, ensures that this anaphoric relationship between the empty noun and its antecedent is established.

So far I have established that the E-feature specifies only the head noun for non-pronunciation, ensuring that the empty noun is connected to an antecedent via identity of sense relationship. For obtaining the full derivation of the ENC, we need to account for the feature specification of the remnants and their interaction with the part specified for the E-feature. I will pursue this issue in the ensuing section.

3.3.1.2 The Focus feature F

The E-feature was shown in the previous section to be sensitive to anaphoric relations in the ENC. The anaphoricity incorporated in the E-feature is to establish a relationship of *identity* between the empty noun and its antecedent. It is this identity relationship that allows non-pronunciation of the given/old material. Contrary to the identity condition imposed on the empty category, the remnant has been argued to be involved in a

¹² For an instance of an anaphoric element in Persian involving sloppy reference, see Ghomeshi’s (2006) analysis of *mâl* ‘(lit.) property’ that is used in possessive constructions as a place holder for the possessee phrase when it is recoverable from the context.

relationship of *non-identity* with the corresponding elements in the antecedent clause (e.g. Giannakidou & Stavrou 1999; Gengel 2007¹³, among others). I adopt Williams' (1997) term *disanaphora* to describe this kind of non-identity relationship. I assume that this disanaphoric relationship is established by the focus feature F that specifies the remnant for *contrastiveness* of some sort. I will show that contrastiveness can manifest itself in contexts in which the relationship between the remnant and the antecedent is that of partitivity, information focus, and emphasis.

As we saw in the previous chapter, contrastiveness was given as the principal condition on licensing empty nouns to ensure the non-identity of the ellipsis remnant with its corresponding element in the antecedent clause (cf. Giannakidou & Stavrou's (1999) Contrast Condition of Noun Ellipsis; Rooth 1992a & b). Rooth (1992b) observes that in sentences where two elements are in contrast, the contrasting elements are prosodically stressed, while the invariant material is prosodically reduced. Rooth considers this phenomenon as being particularly more perceptible in ellipsis and concludes that ellipsis is linked to contrastive focus. Gengel (2007) claims that Rooth's observation is applicable to all instances of elliptical constructions. Below is her example of NP ellipsis in which the remnant is contrastively focused.

(31) I bought the red coat, and you bought the BLUE (one).

[Gengel 2007: 152.59]

¹³ Gengel (2007) adopts the non-identity (contrastiveness) requirement from Kíss (1998) and Szendrői (2001) who assume that overt focus movement in Hungarian requires the moved element to be contrastive.

Considering examples of the ENC in Persian, we observe that the remnant—prenominal (32) or attributive adjectival modifiers (33)—is optionally marked with prosodic prominence¹⁴ when it is in contrast with its counterpart in the antecedent—i.e. when it is focused. The optionality of using focus accent in Persian implies that formal marking of contrastive focus is not necessary for licensing the empty noun¹⁵. Rather, it is important for the focused element to be in contrastive/disanaphoric relationship with its antecedent¹⁶.

(32) mâh-e piš man do-tâ ketâb xund-am, Sinâ [haft]_F-tâ
 month-EZ before I two-CLS book read.PST-1SG Sina seven-CLS
 ketâb .
 book

‘Last month, I read two books, Sina seven.’

¹⁴ Sadat-Tehrani (2007: Ch. 6) considers a (contrastively) focused constituent in Persian as phonologically involving the same Accentual Phrase (AP) pattern as a non-focused one, i.e. (L+)H*. However, he shows that a focused AP has phonetically more pitch excursion and longer duration than the non-focused counterparts (see the original work for further details).

¹⁵ See Alexiadou and Gengel (2008) for a similar observation in Dutch and German.

¹⁶ Focus in this sense is reminiscent of *pragmatic focus* as defined in Lambrecht (1994):

Focus: The semantic component of a pragmatically structured proposition whereby the assertion differs from the presupposition (p. 213).

The focus of the proposition expressed by a sentence in a given utterance context, is seen as the element of information whereby the presupposition and the assertion *differ* from each other. The focus is that portion of a proposition which cannot be taken for granted at the time of speech. It is the *unpredictable* or pragmatically *non-recoverable* element in an utterance. The focus is what makes an utterance into an assertion (p. 207).

- (33) *sos-e sefid-o az ~~sæs~~ [qermez]_F bištar dust=dar-am.*
 sauce-EZ white-OM from ~~sauce~~ red more friend=have-1SG
 ‘I like white sauce more than red’

Recall from §2.2.1 that, within Rooth’s (1992a) theory of Alternative Semantics, a contrastively focused element represents a subset from the set of its suitable alternatives. The F-specification of the remnants in the above ENC signals that the focused element contrasts with a previously mentioned member of the same set of alternatives. Thus, in (33), the focused expression ~~sæs~~ *qermez* ‘red ~~sauce~~’ belongs to a set of alternatives that includes the antecedent *sos-e sefid* ‘white sauce’: $M = \{\text{white sauce, red sauce, etc.}\}$.

Given the above discussion, the contrastiveness requirement of the F feature requires the remnant modifier in the ENC to be included in a set. Reference to a set also characterizes partitive elements licensing empty nouns (Sleeman 1996). This implies, as suggested by Alexiadou & Gengel (2008: 15), that partitivity is a “prerequisite for contrastiveness, which, in the ellipsis constructions, is optionally expressed via contrastive focus”. This also confirms Sleeman’s (1996) proposal that noun ellipsis is linked to partitivity.

Recall from the previous chapter that Sleeman (1996: 33) argues that the elements licensing empty nouns have a partitive interpretation. Partitivity, according to Sleeman, requires that an element be properly or improperly included within a particular set. She divides partitive elements into inherent (D-partitives) vs. non-inherent partitives (N-partitives). Prenominal elements are inherent partitives and adjectives of quality are non-inherent partitives. The occurrence of each of these elements with the head noun creates a

subset of the set that the noun refers to. As we have established in §3.1, prenominal modifiers and attributive adjectives are elements that license empty nouns in Persian. Let us now see how the presence of these elements in the ENC induces partitivity:

- (34) Sajjâd ye âb+sib xarid, [ye+kam]_F ~~âb+sib~~ ham be
 Sajjad one water+apple bought.3SG one+little ~~water+apple~~ also to
 Sinâ dâd.
 Sina gave.3SG
 ‘Sajjad bought an apple juice and gave a little to Sina, too.’

In (34), the prenominal compound quantifier *ye kam* ‘a little’ is an inherent partitive element that forms a subquantity of the antecedent clause *ye âb+sib* ‘an apple juice’. The contrastive/partitive connection between the ENC and its antecedent can be expressed by the partitive construction *ye kam az âb+sib* ‘a little of the apple juice’. Let us now see how partitivity is established in the ENCs containing attributive adjectives.

- (35) sos-e sefid-o az sôs [qermez]_F bištar dust=dar-am.
 sauce-EZ white-OM from sauee red more friend=have-1SG
 ‘I like white sauce more than red’

[<http://marshal-modern.blogspot.com/2008/08/marshal-modern-sandalie-daghe-marshal.html>]

Attributive adjectives, as mentioned earlier, are non-inherent partitives. Non-inherent partitives (N-partitives), according to Sleeman (1996: 35), “do not necessarily form

subsets because they denote properties themselves,” and hence “do not license or only marginally license empty nouns.” The subset is however formed as a result of the intersection between the denotation of the attributive adjective and that of the noun. Thus, in (35), the intersection of the properties of *qermez* ‘red’ and the noun *sos* ‘sauce’ forms a subset of the *kind* denoted by the noun. Sleeman (1996: 46) claims that “noun ellipsis always implies the inclusion of a subset in a previously established set denoting a *kind* [emphasis is mine]”. Under this view, *sos-e sefid* ‘white sauce’ in the antecedent clause and *sos-e qermez* ‘red sauce’ in the ENC are distinct subsets which share the noun *sos* ‘sauce’ denoting a kind as their superset. What Sleeman claims about the superset being a kind-denoting noun seems to correspond to my assumption that there is an identity of sense relationship between the empty noun and its antecedent.

The remnant modifier in the ENC is not always in contrast with a corresponding element in the antecedent, nor is it always prosodically prominent. There are contexts in which the remnant modifier simply adds new properties to the empty noun while there is no counterpart in the antecedent. I consider the remnant in such contexts as being *informationally focused*. Information focus is usually identified by question/answer pairs. The part that provides the inquired information is the information focus.

- (36) A: xune xarid-i?
house bought-2SG
‘Did you buy a house?’

B: âre, balaxare tunest-am ye ~~xune~~ [vilâ-yi]_F be-xar-am.’
 yeah finally could-1SG one ~~house~~ villa-ADJ SBJ-buy-1SG
 ‘Yeah, I could finally buy a villa [one].’

In (36)B, the remnant modifier *vilâ-yi* lacks any corresponding element in the antecedent; it is, as information focus is defined, “the new information in relation to the topic [i.e. *xune* ‘house’—SG]” (Gundel & Fretheim 2004: 181).

Choi (2006: 22) cites Zhu (1997) as suggesting that “every semantic (information) focus is inherently contrastive because there always exist some entities that are alternatives to the information focus. However, it is the speaker’s decision whether to particularly draw the addressee’s attention to it by explicitly marking it as contrastive or not.” Although information focus and partitivity do not share the prosodic prominence with contrastive focus, their properties seem to overlap in that they somehow involve inclusion in a set. For this reason, I use the focus feature F to represent information focus and partitivity as properties associated with contrastiveness. In so doing, I follow Gengel (2007: 168) who notes that the focus feature figures prominently in the semantic derivation of elliptical structures and it is therefore less economical to add features that are similar in kind.

So far we have seen that the role of the F-specified remnant modifier in the ENC is to create some kind of distinction between the empty noun and its antecedent. However, the focused remnant modifier in the ENC can be identical in form to its counterpart in the antecedent when emphasis is intended.

(37) age barâ Sajjâd kâpšen-e čarm-i be-xar-im barâ Sinâ ham bâyad
 if for Sajjad jacket-EZ leather-ADJ SBJ-buy-1PL for Sina also must
~~kâpšen~~ [ČARM-I]_F be-xar-im.
~~jacket~~ leather-ADJ SBJ-buy-1PL

‘If we buy a leather jacket for Sajjad, we must buy a leather [one] for Sina too.’

Since the redundant material is not expected to be pronounced in the ENC, I suggest that the prosodic prominence of the remnant modifier *čarmi* ‘of leather’ in (37) is not informationally significant. Rather, it marks emphasis of the antecedent modifier. This claim is supported by the fact that the presence of the remnant modifier is optional. Thus, the whole object of the main clause in (37) can be left unpronounced and the subject *Sina* can instead be focalized.

(38) age barâ Sajjâd kâpšen-e čarm-i be-xar-im barâ [SINÂ]_F ham
 if for Sajjad jacket-EZ leather-ADJ SBJ-buy-1PL for Sina also
 bâyad _{nP}[~~kâpšen~~ — ~~čarm-i~~] be-xar-im.
 must ~~jacket~~ ~~leather-ADJ~~ SBJ-buy-1PL

‘If we buy a leather jacket for Sajjad, we must buy [one] for Sina too.’

Looking back at (37), the question arises as to why the focus feature F is assigned to the emphatic remnant while it is considered to be informationally insignificant. To answer this question, I present cross-linguistic evidence from languages in which the overt focus marker is also an emphatic marker. Corver & Koppen (2005), for instance,

show that in some dialects of Dutch, the focus marker is also associated with emphatic force. In the dialect of Grave, for example, when the attributive *aP* is emphasized, the *e*-affix (a focus marker) appears on the adjective.

- (39) a. et choe:t kient
 the good child
- b. et choej-e kient
 the good-E child
 ‘the GOOD child’

[Corver & Koppen 2005: 7.21]

I use this evidence to argue for assigning the focus feature *F* to emphatic remnants.

In regard to the occurrence of the identical modifier in the ENC and the antecedent clause, a constraint is in order: The remnant modifier can be identical to its counterpart in the antecedent if and only if the ENC is not embedded in a coordinate structure. The ungrammaticality of (40) shows that the identity of modifiers in the ENC and the antecedent clause not only rules out the possibility of having an empty noun (40)a but also does not license focus marking of the remnant (40)b.

- (40) a. *Sajjâd kâpšen-e čarm-i xarid o Sinâ ~~kâpšen~~ čarm-i.
 Sajjad jacket-EZ leather-ADJ bought.3SG and Sina ~~jacket~~ leather-ADJ
 ‘Sajjad bought a leather jacket and Sina [bought] a leather [one].’

b. *Sajjâd kâpšen-e [ČARM-I]_F xarid o Sinâ kâpšen [ČARM-I]_F.

As we saw in the previous chapter, Ntelitheos (2004) reports a similar observation in Greek based on which he argues for the contrastiveness requirement for licensing empty nouns and focus marking.

(41) *i Maria forese to ble fustani ke i Eleni forese to ble ~~fustani~~.
the Maria wore-3SG the blue dress and the Eleni wore-3SG the blue ~~dress~~
'Maria wore the blue dress and Eleni wore the blue one.'

[Ntelitheos 2004: 36.34b]

A comparison of the facts in (37), (38), and (40) reveals that the requirement for the obligatory presence of focus stress implicitly suggests that the two conjuncts are in a contrastive relationship. This is particularly evident in (38) where the non-pronunciation of the whole nominal phrase results in the shift of focus stress onto the contrasting subject *Sinâ*. Looking more deeply at (37), where there is an emphatic remnant, we find two kinds of relationships with the antecedent: *identity of sense* and *non-identity of reference*. As far as the semantic interpretation is concerned, the intersection of the meanings of the adjective *čarmi* 'of leather' and the noun *kâpšen* 'jacket' in both conjuncts produces the same intension. However, the empty noun and the antecedent pick out two different objects in the real world. Technically speaking, they refer to two different tokens of the same type. Here is where I suggest that the empty noun and the antecedent are in a non-identity of reference relationship. Building on this argument, I

propose that, like other relationships holding between the remnant modifier and the antecedent (i.e. partitivity and information focus), emphasis also triggers contrastiveness. Specifically, the emphatic remnant modifier induces non-identity of reference with its counterpart in the antecedent and is thus marked with the focus feature F, as predicted.

In this section, I argued for the existence of two kinds of relationships:

- (42) i. a relationship of identity of sense between the empty noun and its antecedent
- ii. a relationship of contrastiveness between the remnant modifier and its counterpart in the antecedent.

I assumed that in deriving the ENC, the E-feature specifies the sister Root to the categorizing n^0 head for non-pronunciation in the syntax, and ensures that the semantic relationship in (42)i is established at LF. The focus feature F is also assumed to specify the remnant modifiers licensing empty nouns at syntax, and ensure that the relationship in (42)ii is implemented at LF. I demonstrated that although there are four kinds of relationships, viz. partitivity, contrastiveness, information focus, and emphasis, holding between the remnant modifier and its counterpart, contrastiveness (non-identity) is somehow implied by each of them. Thus, I concluded that the focus feature F can be assigned to the remnant that is involved in any kind of the above relationships with the antecedent. Now that the feature specification of the empty noun and the remnant is determined, I turn to my proposal for a DM analysis of the ENC in Persian.

3.3.2 Proposal

My proposal for the derivation of the ENC in Persian is bipartite. Part of it accounts for the link between non-pronunciation of the head noun and information structure, the other part addresses the syntactic conditions on licensing the empty noun. Assuming the framework of Distributed Morphology (Halle & Marantz 1993, and subsequent related work) and a phase-based syntactic theory (Chomsky 2000 *et seq.*), I propose that these two aspects conspire to yield a picture of the ENC where the derivation is computed in a phase-based manner.

3.3.2.1 Focus Condition on the ENC

We explored earlier the information-structural motivation for non-pronunciation of the given/old material in the ENC. In light of the discussion in §3.3.1, I propose that the following condition clarifies and regulates the environment in which the empty noun is licensed in terms of information structure:

(43) *Focus Condition on the ENC in Persian*

A nominal subconstituent α can be left unpronounced in constituent β iff β has a salient antecedent δ that is mentioned or implied in the previous text or discourse, such that

- (i) α is in identity-of-sense relationship with the corresponding subpart of δ ;
- and
- (ii) The remnant of β is disanaphoric to the corresponding part of δ .

Given our assumptions so far, the given/old material is specified with the E-feature, and the nominal modifiers that are new or contrastive are specified with the focus feature F. As discussed in §3.3.1.1, although the Root of the head noun always represents the given/old material, it may be specified with the E-feature as a complex Root containing the head modifier. This results in non-pronunciation of the nominal Root and its head modifiers at PF. In cases where a simple Root node or a complex Root containing the head modifier is specified with the E-feature, non-pronunciation occurs when there is an F-specified head modifier inside the *nP* or a prenominal modifier in the superstructure of the *nP*.

Assuming Distributed Morphology, the Root of the head noun and the head modifier(s) adjoined to it (i.e. attributive adjectives) are represented in the syntax as terminal nodes containing bundles of semantic, morphological, and syntactic features with no phonological content. At this level, I assume that the information structural features E and F are assigned to their relevant terminal nodes in the *nP* domain. The E-feature specified on the sister Root to the nominalizing head requires one of the head modifiers to be specified for the focus feature F. The F-specified terminal node cannot be contained within an E-specified complex Root node because the E-specified portion of the *nP* constitutes the given/old material and it cannot include any element that is part of the new or contrastive information. Recall that in the absence of a head modifier (i.e. attributive adjective) inside the *nP* domain, a prenominal modifier in the superstructure of the *nP* allows empty nouns in Persian. I propose that since prenominal modifiers are inherently partitive (Sleeman 1996), i.e. involve contrastiveness, they are inherently specified with the focus feature F.

The (non-)identity requirement for the E-specified and F-specified elements in the ENC, I maintain, is an LF constraint which is satisfied at the interface with the syntax through checking the required relationship between the relevant terminal nodes in the ENC and their corresponding nodes in the antecedent clause. All things being equal, we expect that the application of the Focus Condition in (43) results in the E-specified constituent not being pronounced at PF and the F-specified terminal node being spelled out by an overt exponent distinct from that of its counterpart in the antecedent clause. The exception to the latter prediction is where the contrastive relationship between the F-specified modifier and the antecedent is that of emphasis, in which case the exponents of the modifier and the antecedent are identical. Thus, I assume that the terminal nodes for head modifiers in the *nP* that are not specified with the E or F feature will be pronounced, as is normally the case for all the other nodes. The pronunciation of these unspecified head modifiers is regardless of their relationship with their antecedents. In the ensuing section, I will discuss in more detail the syntactic conditions governing the licensing of empty nouns in Persian.

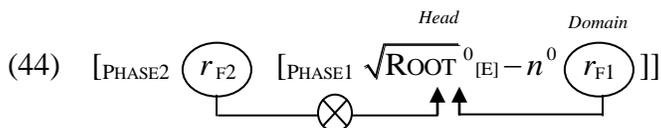
3.3.2.2 Licensing properties of the ENC

Assuming that F-specified remnant modifiers are the empty noun licensors in the ENC, we have seen that the licensing process is sensitive to locality effects. Specifically, the empty noun (abbreviated here as $\sqrt{\text{ROOT}_{[E]}}$) is licensed by the F-specified element inside the same domain. To put it informally, the empty noun is licensed by the presence of an attributive adjective inside the *nP*—call this remnant modifier r_{F1} . The inherently F-specified modifier, i.e. prenominal modifier, comes into play only when there is no r_{F1}

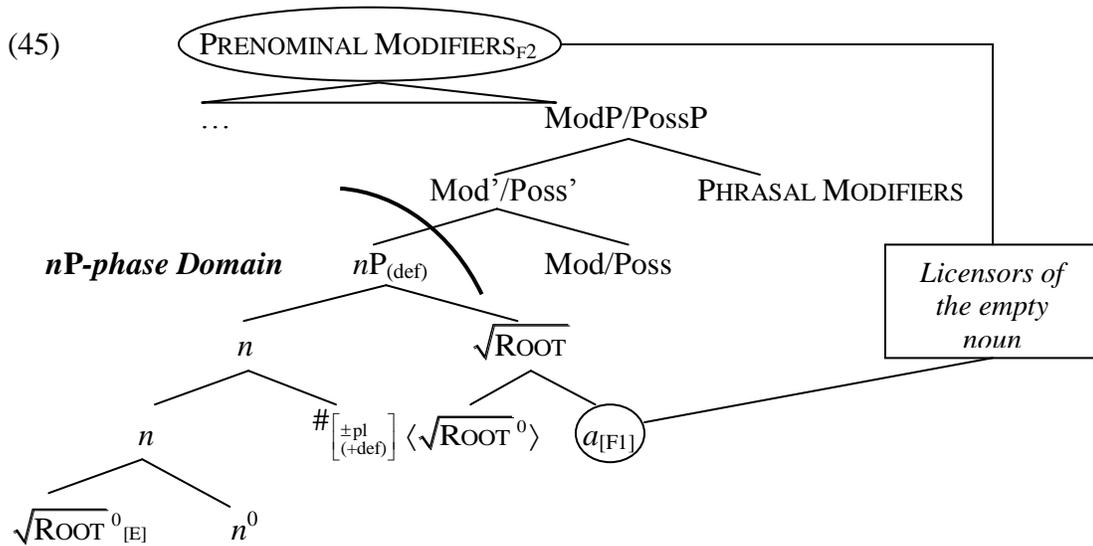
inside the nP ; hence we call this modifier r_{F2} in the ENC. Now the question is what syntactic properties might underlie this licensing process?

I propose that the licensing properties of the empty noun in Persian fall out naturally from Chomsky's (1999 *et seq.*) phase-based approach to syntactic derivations. According to the phase theory, syntactic derivations proceed in chunks (*phases*). As discussed in Chapter 1, a phase refers to a locality domain for semantic and phonological processing. When a phase is complete, it is sent off to the interfaces (LF and PF) for interpretation. Chomsky's (2000) Phase Impenetrability Condition (PIC) constrains accessibility of later syntactic operations to the *domain* (complement) of a phase head once it is complete. Syntactic operations outside the phase can only target the head and *edge* (specifiers and adjuncts) of the phase.

Marantz (2001, 2007) argues that category-defining functional heads—*little x* heads—are phase heads in Chomsky's sense, in that they initiate spell-out (see also Embick 2008: §1.4). Svenonius (2004) introduces n as a phase head in the nominal structure, parallel to v in the clausal structure. In my analysis, I adopt the view that n is a phase head and propose that it defines a locality domain in which the empty noun is licensed by r_{F1} . I argue that the reason why the empty noun cannot be licensed by both r_{F1} and r_{F2} is due to a PIC effect. To elaborate, when the licensing requirement for the empty noun can be satisfied within the domain of the nP , the presence of a licenser outside the phase domain cannot affect the licensing process. This is schematically shown in (44):



When there is no r_{F1} in the domain of the nP , however, the empty noun is licensed by r_{F2} in the domain of a higher structure. Given our assumption in §3.3.2.1 that the specification of the E-feature on the n Root requires the presence of an F-specified head modifier, I propose that the n head defines a phase only when it contains r_{F1} ; otherwise, it is part of the derivation of the whole ENC. As illustrated in the following syntactic configuration, r_{F1} and $\sqrt{\text{ROOT}}_{[E]}$ are within the nP phase and r_{F2} is outside the phase domain. In the absence of r_{F1} , the structure in (45) is computed at once.



Building on the discussion above, I propose that the licensing of the empty noun follows the conditions in (46). Since head movement of an uncategorized Root to a category-defining head n is part of normal derivation of nominal phrases, I assume that the following Licensing condition applies after categorization:

(46) *Licensing condition on the empty noun in Persian*

The empty noun $\sqrt{\text{ROOT}_{[\text{E}]}}$ is licensed by an F-specified remnant modifier r_{F} in a phase-based manner, where the phase contains r_{F} .

3.3.2.3 The ENC at the interfaces

Recall that phases, upon completion, are sent off to PF and LF for interpretation. I propose that at LF, the E-feature and the focus feature F ensure that the Focus Condition on the ENC in (43) are established. That is, the E-feature ensures that the node that it specifies is in an identity of sense relationship with the antecedent. The focus feature F, on the other hand, ensures that the remnant modifier is in a non-identical (contrastive) relationship with its counterpart in the antecedent. If the empty noun is licensed in the $n\text{P}$ phase by $r_{\text{F}1}$, I assume that the non-identity condition must be satisfied by this modifier. In such cases, LF is not sensitive to the contrastiveness of $r_{\text{F}2}$, if present in the next higher phase. Where the empty noun is licensed by $r_{\text{F}2}$, the modifier must satisfy the contrastivity requirement for remnant modifiers. It is, in fact, this interaction between the $\sqrt{\text{ROOT}_{[\text{E}]}}$ and remnant modifiers that results in interpreting the empty noun and its antecedent as two distinct subsets of the kind denoted by the noun.

The input to the PF component of a DM model of grammar is a syntactic structure which contains all the featural information essential for syntactic derivation and semantic interpretation (see, for instance, Embick & Noyer 2006). At PF, the terminal nodes containing syntactico-semantic feature bundles are mapped onto their phonological exponents through a process called *Vocabulary Insertion*. The mapping process involves

competition of Vocabulary Items (VI), specified for certain feature bundles or context of insertion, to realize individual terminal nodes in the syntax. VIs can only compete to realize nodes whose feature bundles are equally or more fully specified than they themselves are. The most fully specified matching VI wins the competition for each terminal node.

Given that the information-structural and syntactic licensing conditions of the empty noun are satisfied before entering PF, I propose that at this interface level, the E-feature blocks the Vocabulary Insertion from inserting the phonological exponent of the terminal node(s) representing the head noun¹⁷. This way, this special feature blocks the normal process of Vocabulary Insertion¹⁸. This view is essentially compatible with the basic idea of the PF-deletion accounts, in which elliptical structures are assumed to maintain their internal structure in the course of derivation (see §2.2). Defining the E-feature as bearing an instruction for non-pronunciation of terminal nodes at Vocabulary Insertion implies that ellipsis does not involve syntactic process of deletion, as similarly suggested by Merchant (2004: 671) and Kornfeld & Saab (2004).

The licensing of the empty noun in my account of the ENC so far has been attributed to the interaction between the properties of the E-feature and those of the focus feature F. Specifically, we have established that the presence of the E-feature on the head noun requires F-specification of the remnant modifier. Relying on the same argument, I propose that the non-pronunciation of the E-specified head noun at Vocabulary Insertion

¹⁷ By the head noun I also mean the *n* head that contains a complex E-specified Root consisting of a daughter Root and a head modifier *a* (see e.g. (28) and (29)).

¹⁸ See Harley (2005: 14) for a similar analysis of ellipsis using the feature [+Id] to specify terminal nodes which are “exact equivalents of other nodes in an appropriate licensing relation”.

intersects with the pronunciation of the F-specified remnant modifier. That is, the E-specified head noun can be left unpronounced only when an F-specified head modifier gets spelled out.

The important implication this analysis has for syntactic derivation and spell-out is that the n head is included in the spell-out domain of the phase which contains r_F (i.e. the F-specified remnant modifier). Thus, when the empty noun is licensed by r_{F1} inside the nP , the n head is a phase head which is spelled out along with its complement. This is contra previous models under which only the complement of the phase head undergoes spell-out (Chomsky 1999 *et seq.*). Where the empty noun is licensed by a prenominal modifier (r_{F2}) in the next higher phase, I assume that the nP is spelled out along with its superstructure. In other words, the whole derivation of the ENC undergoes spell-out. In such cases, the derivation of the nP cannot be sent off to PF because the empty noun does not meet its information-structural and syntactic licensing conditions. This, I suggest, is why the spell-out of the nP is delayed until the whole ENC is constructed.

Having presented and developed an analysis of the ENC in Persian, it is now time to investigate the workings of this proposal with reference to the Persian data.

3.3.3 Explanation of the data

The two core assumptions underlying the proposal developed in this chapter are repeated in (47):

(47) *Assumptions*

a. *Focus condition on the ENC in Persian*

A nominal subconstituent α can be left unpronounced in constituent β iff β has a salient antecedent δ that is mentioned or implied in the previous text or discourse, such that

- (i) α is in identity-of-sense relationship with the corresponding subpart of δ ;
and
- (ii) The remnant of β is non-identical to the corresponding part of δ .

b. *Licensing Condition on the empty noun in Persian*

The empty noun $\sqrt{\text{ROOT}}_{[E]}$ is licensed by an F-specified remnant modifier r_F in a phase-based manner, where the phase contains r_F .

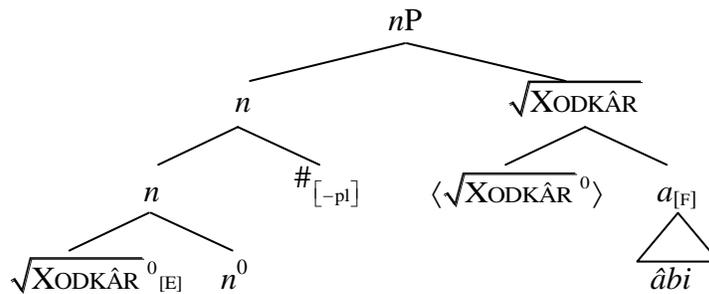
Let us begin with the basic instances of the ENC: the empty noun with an attributive adjectival modifier (48); and the empty noun with a prenominal modifier (49). The examples used here are all out-of-the blue (context-neutral) sentences.

- (48) Sajjâd xodkâr-e qermez dâr-e, Sinâ ~~xodkâr~~ $r_{F1}[\hat{a}bi]$
Sajjad pen-EZ red have-3SG Sina ~~pen~~ blue
'Sajjad has a red pen, Sina a blue one.'

- (49) Sajjâd DO-tâ xodkâr dâr-e, Sinâ $r_{F2}[SE-tâ]$ ~~xodkâr~~
 Sajjad two-CLS pen have-3SG Sina three-CLS pen
 ‘Sajjad has two pens, Sina three.’

(48) is an example of the ENC in which the head noun is generated in the syntax in the same phase as its modifier. The node representing the head noun contains the E-specified Root and the one representing the modifier is F-specified, in accordance with the requirements of Focus Condition in (47)a. Structurally, the empty noun is licensed by an r_{F1} inside the nP . Thus, based on the condition in (47)b, the nP is a phase because it contains r_{F1} . The following partial configuration illustrates the condition where the empty noun in (48) is licensed:

- (50) Sajjâd xodkâr-e qermez dâr-e, Sinâ...



The tree diagram in (50) thus represents a complete nP phase since the conditions in (47) are satisfied. As such, it is sent off to LF and PF for interpretation. At LF, the requirements in the Focus Condition are checked. At PF, the presence of the E-feature on the n Root blocks the insertion of the phonological exponent of the node representing *xodkâr* ‘pen’. For this to happen, the F-specified modifier must be phonologized, as

predicted. After Vocabulary Insertion and Linearization (cf. §1.1.2.2), the concatenated elements are subject to the PF-operation *Chaining*, where all information in a linear sequence is represented (Embick 2007, 2008 and Pak 2008). I propose that the prosodic prominence of the focused modifier is assigned to it after Chaining, at the end of the PF derivation¹⁹.

A caveat is in order, though. Not all remnant modifiers end up having focus stress in the ENC. As the sentences used in this dissertation are context-neutral, the focus stress overlaps with the default phrasal stress. I will not pursue a systematic account of the prosody of the ENC in Persian here, as it is beyond the scope of this thesis. However, in my account of the data, I will refer to this factor where relevant.

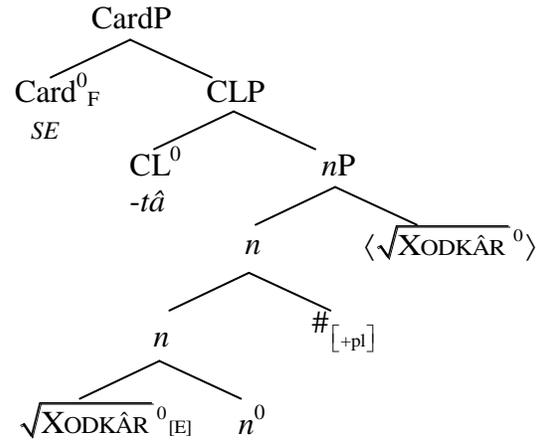
The possibility of licensing the empty noun with a prenominal modifier is exemplified in (49), repeated below.

- (51) Sajjâd DO-tâ xodkâr dâr-e, Sinâ $r_{F2}[\text{SE-tâ}]$ ~~xodkâr~~
 Sajjad two-CLS pen have-3SG Sina three-CLS pen
 ‘Sajjad has two pens, Sina three.’

(51) is an example of a situation where there is no r_{F1} in the *nP* domain. At this point, the empty noun cannot be licensed and thus the *nP* does not constitute a phase.

¹⁹ Pak (2008) terms PF operations that apply after all constituents in a derivation are chained together as *Late-Linearization* rules. I assume that stress assignment in Persian instantiates such rules. In Chapter 5, I propose that Ezafe insertion in Persian is a Late-Linearization rule.

(52) Sajjâd DO-tâ xodkâr dêr-e, Sinâ...



The nP in the above ENC is not a phase because it does not contain an r_{F1} and as a result does not meet the information-structural and syntactic licensing conditions. Thus, the computation of the immediately higher structure continues. At this stage, the nP is computed as a part of the root node, where the empty noun is licensed by r_{F2} . Upon licensing the empty noun, the derivation is sent off to the interfaces. If our analysis is correct that the licensing of the empty noun takes place as a result of the interaction between the E and F features specified on the n Root and the remnant modifier, it follows that the semantic and phonological interpretation of the empty noun in (52) should occur in the same spell-out domain as that of the prenominal licensor. I suggest that in such contexts where the empty noun is licensed by a prenominal modifier, the licensing modifier inherits the [+N] feature of the empty noun. This feature inheritance, I propose, causes widening of the nP domain to include the projection of the prenominal modifier

An interesting observation about this example is that although both of the modifiers induce contrastiveness with their counterparts in the antecedent, only r_{F1} bears the focus/primary stress. The ungrammaticality of (54) shows that r_{F2} cannot be stressed when r_{F1} is present in the ENC²¹.

- (54) *Sajjâd DO-tâ xodkâr-e qermez dâr-e, Sinâ r_{F2} [SE_F-tâ] ~~xodkâr~~ r_{F1} [âbi]
 ‘*Sajjad has TWO red pens, Sina THREE blue.’

I consider the two sentences (53) and (54) as evidence corroborating my phase-based proposal for the derivation of the ENC in Persian. Under this approach, the derivation of (53) starts out with the nP phase. Since both conditions in (47) can be satisfied within this phase, the derivation of the phase proceeds to completion. The phase is then shipped off to the interfaces for semantic and phonological interpretation. Recall our assumption earlier that the spell-out domain of the nP phase includes the head noun and its complement. After spell-out, the phase cannot be accessed by operations in the higher phases (due to PIC). Given this, I assume that the semantic and phonological interpretation of the ENC in (53) is complete by the time the derivation gets to the higher phase. The prenominal modifier in that phase (i.e. r_{F2}) can thus have no role in licensing the empty noun.

The same argument holds for the impossibility of assigning focus stress to r_{F2} when there is a licensing modifier inside the lower phase. I assume that focus/primary stress

²¹ Context-neutral sentences like (54) may obtain different grammaticality judgements when considered within certain contexts.

assignment to the empty noun licensing modifier occurs in the same phase where the empty noun is licensed and interpreted at the interfaces. This explains the ungrammaticality of (54) in which r_{F2} is stressed in a phase different from one in which the empty noun has been licensed.

If this analysis is correct, I predict that violation of the Focus Condition (47)aii should not yield ungrammaticality for r_{F2} in (53) because all the requirements of the ENC have been satisfied in the lower phase. As shown in (55), this prediction is borne out.

(55)	Sajjâd do-tâ xodkâr-e QERMEZ dâr-e, Sinâ	r_{F2} [do-tâ]	xodkâr	r_{F1} [ÂBI].
	‘Sajjad has two RED pens, Sina two BLUE.’			

Violation of the Focus Condition (47)aii, however, leads to ungrammaticality when it happens inside the nP phase in the course of licensing process. In (56), the sentence is ungrammatical because r_{F1} does not satisfy the non-identity requirement in (47)aii and hence the derivation crashes in the same phase. Notice that even satisfying this condition by r_{F2} in the next phase cannot rescue the ungrammaticality of this sentence.

(56)	*Sajjâd do-tâ xodkâr-e qermez dâr-e, Sinâ	r_{F2} [se-tâ]	xodkâr	r_{F1} [QERMEZ].
	‘Sajjad has two RED pens, Sina three RED.’			

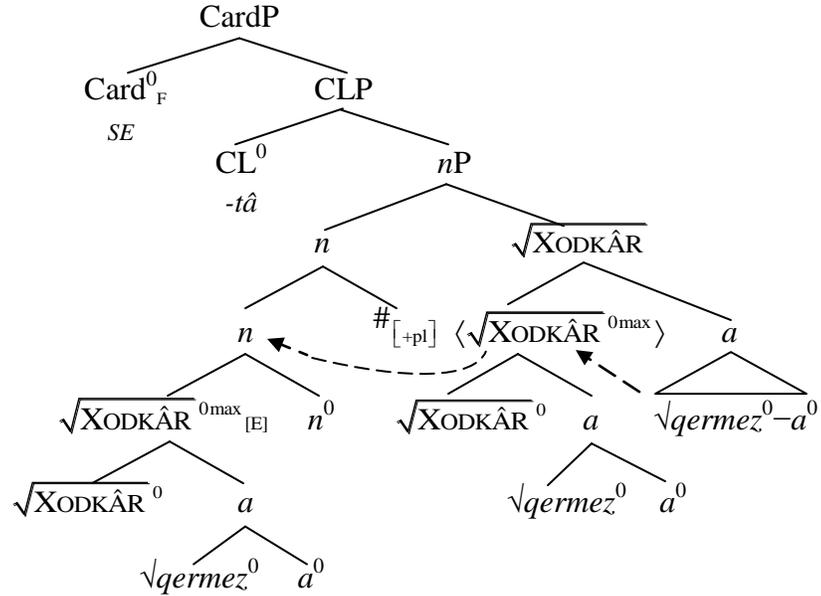
Recall from §3.3.2.1 that the E-specification of the Root requires one of the head modifiers inside the nP to be F-specified. In cases like (56) where there is only one head modifier inside the nP phase, that modifier has to be F-specified. This entails that the F-

specified modifier has to be interpreted at LF as being in a disanaphoric relationship with its counterpart in the antecedent, and is phonologically realized at PF by a distinct exponent from that of its antecedent. Since none of these conditions applies to the head modifier in the above ENC, the derivation crashes. The derivation however can be rescued if the terminal node representing *qermez* ‘red’ in the ENC is incorporated into the nominal Root and thus becomes part of the complex Root that adjoins to the nominalizing *n* head. In that case, the E-feature assigned to the complex Root specifies both the terminal Root node and the head modifier for non-pronunciation; that is, it specifies them as being part of the given/old information. As such, the derivation will proceed just as predicted for cases where there is no r_{F1} in the *nP* phase (see my account for (52)). Specifically, the empty noun is licensed by the inherently F-specified numeral (r_{F2}) in the prenominal position. The sentence (56) will therefore be grammatical if *QERMEZ* is not pronounced.

- (57) Sajjâd DO-tâ xodkâr-e qermez dâr-e, Sinâ _{ENC}[r_{F2} [SE-tâ] ~~xodkâr~~ {*QERMEZ*}].
 ‘Sajjad has TWO red pens, Sina THREE.’

The syntactic derivation of the above ENC can be shown as follows:

(58)



Notice that the stress placement in (57) follows the same pattern we observed in (49): r_{F2} bears the focus stress. The comparison of (54) and (57) demonstrates that the prosodic properties of the ENC is computed at spell-out in a phase-based manner. Thus, if all the requirements of licensing the empty noun are satisfied within the nP -phase, stress placement applies to this phase at PF spell-out. If not, stress placement occurs at PF spell-out of the higher phase. This is what we see happen in (57); the empty noun is not licensed inside the nP domain and the derivation continues in the next higher phase where the presence of r_{F2} satisfies the requirements of licensing the empty noun. The derivation is then shipped off to PF and LF. Stress placement occurs at PF and the contrastive r_{F2} receives focus stress.

If the empty noun is licensed by r_{F2} , the Focus Condition in (47)aii—the non-identity requirement—must also be satisfied by this modifier. This is confirmed by (59), where the identity between r_{F2} and its counterpart in the antecedent yields ungrammaticality.

- (59) *Sajjâd DO-tâ xodkâr-e qermez dâr-e, Sinâ_{ENC} [_{r_{F2}} DO-tâ] xodkâr [QERMEZ].
 ‘*Sajjad has TWO red pens, Sina TWO.’

The last issue to resolve concerns the examples of the ENC that contain more than one attributive adjectival modifier.

- | | | | | | | | |
|------|--------|----------|------------|--------|---------|----------------------------|--|
| | | | | | | the <i>nP</i>-phase | |
| (60) | Sajjâd | pirhan-e | râh+râh-e | QERMEZ | dâšt, | pirhan | (_{r_{F1}})[sâde-ye] |
| | Sajjad | shirt-EZ | striped-EZ | red | had.3SG | shirt | plain-EZ |
| | | | | | | | |
| | | | | | | | (_{r_{F1}})[MEŠKI]. |
| | | | | | | | black |
- ‘Sajjad had a red striped shirt, Sina a plain black one.’

As far as the derivation of such sentences is concerned, the same licensing mechanism that accounts for the derivation of typical instances of the ENC is also at work here. The presence of more than one modifier in the *nP* is allowed as long as the Focus Condition in (47)iii is satisfied. To elaborate, in (60), one of the terminal nodes for the head modifiers has to be F-specified. Since both of the head modifiers are in contrastive relationship with their counterparts in the antecedent, the terminal node for any of them can be specified with the F feature. The two head modifiers can also meet the licensing condition (47)b. That these possible instances of *r_{F1}* are in the *nP* domain indicates that they can license the empty noun in the *nP* phase. Accordingly, the *nP* phase is complete and ready to be interpreted semantically and phonologically at the interfaces.

One significant question that arises regarding the licensing process of cases like (60) is which one of the two head modifiers is the F-specified modifier (r_{F1}) licensing the empty noun? Let us assume, for the sake of argument, that the the closest head modifier to the *n* Root *pirhan* ‘shirt’ is r_{F1} . If this assumption is correct, the order of attributive adjectives is expected to be restricted, contrary to fact. As can be seen in (61), reversing the order of adjectives does not affect the grammaticality.

						the <i>n</i>P-phase	
(61)	Sajjâd	pirhan-e qermez-e	RÂH+RÂH	dâšt,	Sinâ	pirhan	(r_{F1})[meški-ye]
	Sajjad	shirt-EZ red -EZ	striped	had.3SG	Sina	shirt	black-EZ
							(r_{F1})[SÂDE].

plain

‘Sajjad had a striped red shirt, Sina a black plain one.’

This observation entails that licensing the empty noun is not sensitive to the distance of the licensor, but rather to its F-specification. A fact that strongly supports this claim is that in ENC_s involving more than one attributive adjective, unspecified head modifiers can intervene between the empty noun and its F-specified head modifier (r_{F1}).

						the <i>n</i>P-phase	
(62)	Sajjâd	pirhan-e qermez-e	RÂH+RÂH	dâšt,	Sinâ	pirhan	[qermez]-e
	Sajjad	shirt-EZ red -EZ	striped	had.3SG	Sina	shirt	red-EZ
							r_{F1} [SÂDE].

plain

‘Sajjad had a striped red shirt, Sina a red plain one.’

Besides showing that the empty noun and its licensor need not be adjacent, it also demonstrates that unspecified head modifiers can be realized as phonologically identical to their counterparts in the antecedent.

That adjacency between the empty noun and its licensor is irrelevant to the licensing process also finds support in stress placement of the *nP* phase at the PF interface. We see in examples (60) and (61) that the *nP* exhibits the default stress pattern regardless of ordering of the adjectives; i.e. stress falls on the rightmost word. This confirms our earlier assumption that focus and primary stress overlap in the ENC context-neutral examples in Persian in which the empty noun is licensed by r_{F1} .

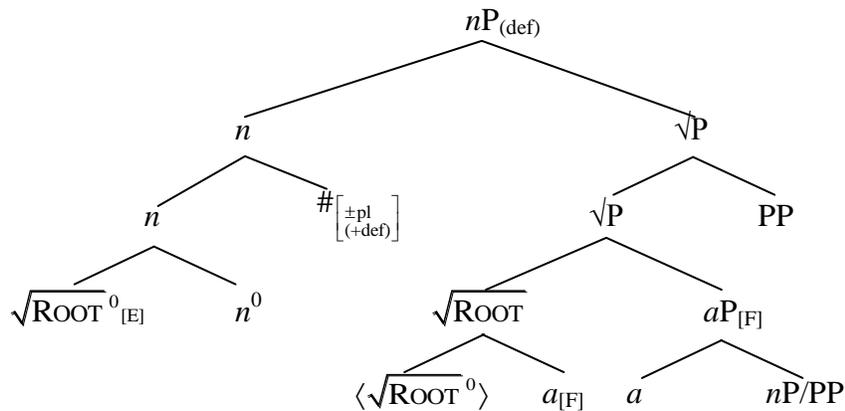
The other implication of this observation is for the status of attributive adjectives as head adjuncts in the structure of noun phrases in Persian (cf. Chapter 1, §1.4.2). I propose that the status of these head modifiers as adjuncts is evidenced by their ordering flexibility. The comparison of (60) and (61) shows that the free ordering of adjectives in the *nP* domain is not determinant for licensing the empty noun—and presumably any other syntactic operations (cf. §5.2.3).

I would like to close this chapter by pointing out one significant implication of a phase-based approach to the ENC for the structural configuration of the noun phrase in Persian. Recall from Chapter 1 that I proposed a distinction between nominal modifiers of the Ezafe domain in terms of their head status and structural positions. I suggested that the modifiers inside the *nP* domain are all heads adjoined to the uncategorized Root of the head noun (following Ghomeshi 1997). I further proposed that phrasal modifiers are generated outside the *nP* domain in the specifiers of ModP/PossP. At the end of §3.2, I argued that the structural split between head modifiers and phrasal modifiers in the Ezafe

domain is motivated by the empty noun licensing conditions in Persian. Specifically, the fact that the empty noun is licensed only by head modifiers and not phrasal ones indicate that these modifiers belong to two distinct domains of modification, i.e. nP and ModP/PossP, respectively.

I propose that the distinction between head vs. phrasal modifiers in the Ezafe domain falls out naturally from a phase-based analysis of the ENC. Recall that the principal licenser of the empty noun, i.e. attributive adjective, is a head modifier in the nP -phase. Only modifiers of this kind can be specified with the focus feature F. The absence of such modifiers would result in the empty noun being licensed by a prenominal modifier. Let us suppose, for the sake of discussion, that phrasal modifiers are generated with the head modifiers in the same phase. As such, the structural configuration of the nP would look like the following:

(63)



Assuming (63), the head modifier a and the phrasal modifier aP can both meet the empty noun licensing conditions in (47). This is not desirable because it cannot reflect the fact that the empty noun cannot be licensed in contexts in which the only existing modifier in

the *nP* domain is *aP*. The pair of examples below shows this context (64) in contrast with the one in which the empty noun is possible with head modifiers (65).

(64) * man bâ xod-am be irân ye čamedun-e *aP*[por az ketâb]
 I with self-CL.1SG to Iran one suitcase-EZ full of book

mi-bar-am	o	ye		čamedun	<i>aP</i> [por az lebâs] _F .
DUR-take-1SG	and	one		suitcase	full of clothes

‘I will take with myself to Iran a suitcase full of books and one full of clothes.’

(65) man bâ xod-am be irân ye čamedun-e *a*[bozorg] mi-bar-am
 I with self-CL.1SG to Iran one suitcase-EZ large DUR-take-1SG

				the <i>nP</i>-phase	
o	ye			čamedun	<i>a</i> [kučik] _F .
and	one			suitcase	small

‘I will take with myself to Iran a large suitcase and a small one.’

In (64), the only F-specified modifier in the *nP* domain is the *aP* *por az lebâs* ‘full of clothes’; hence it meets the licensing condition (47)b according to which the *n* head defines a phase. This *aP* is also in contrastive relationship with its counterpart in the antecedent, i.e. *por az ketâb* ‘full of books’. It therefore satisfies the Focus Condition (47)aii and is F-specified. The ungrammaticality of this sentence can thus be used as supporting evidence for the plausibility of my proposal for distinguishing *inner* (*nP*) vs.

outer (ModP/PossP) modifiers. It also implies that a phase-based analysis of the ENC in Persian can best explain the necessity of assuming the distinction between *inner* vs. *outer* domains of nominal modification.

3.4. Summary

In this chapter I have provided a novel account of the Empty Noun Construction in Persian using a combination of Distributed Morphology and Phase Theory of syntax. I have argued that ellipsis process is a PF phenomenon within DM, in which pronunciation of abstract morphemes or the lack thereof, is determined at Vocabulary Insertion. What determines the non-pronunciation of the head noun in the ENC, I propose, is the interaction between the information-structural features E and F assigned to the empty noun and its remnants. This way, my proposal is along the lines of those that link ellipsis to information structure (e.g. Rooth 1992a, 1992b; Gengel 2007, among others).

I have proposed that the semantic and syntactic conditions on licensing the empty noun in Persian require the derivation to proceed in a phase-based manner. This assumption follows from the cyclic nature of licensing the empty noun which requires the empty noun (the E-specified head noun) and its licenser (the F-specified modifier) to be in the same cyclic (phase) domain, i.e. the *nP* domain. The *nP* phase is the domain where F-specification is assigned to an attributive adjectival head modifier in the context of an E-specified head noun. This head modifier is the only licenser to which the focus feature F is assigned. I assume that pronominal modifiers that license the empty noun when there is no modifier inside the *nP* domain are not subject to F-specification process; rather, they are *inherently* F-specified. We will see throughout the remainder of this dissertation that

the phase-based approach to the derivation of the ENC plays a key role in accounting for various structural properties of noun phrases in Persian.

Chapter 4

Plural Marking and the ENC

4.0. Introduction

This chapter provides a Distributed-Morphology based analysis of *-hâ* displacement in the ENC. In the presence of a phonologically null noun in Persian noun phrases, the plural marker *-hâ* moves onto the right-adjacent attributive adjective (1a). In the absence of an attributive adjective, the plural marker is instead attached to its immediately preceding nominal modifier (1b).

- | | | | | | |
|--|---|----------|-----|----------|-----------|
| | Q | <i>n</i> | -PL | <i>a</i> | Possessor |
|--|---|----------|-----|----------|-----------|
- (1) *behtarin dânešju -hâ-ye javân-e dânešgâh*
best student -PL-EZ young-EZ university
'the best young students of the university'

a. *behtarin dânešju(-hâ) javân-hâ-ye dânešgâh*
'the best young ones of the university'

b. *behtarin-hâ-ye dânešju(-hâ) dânešgâh*
'the best ones of the university'

In presenting my analysis of *-hâ* displacement in this chapter, I first explore the

noun phrase-internal constituents that allow empty nouns and host the plural marker. Next, I will discuss the syntax of plural formation in Persian proposing a novel analysis of plural marking and its link to definiteness. A detailed analysis of *-hâ* mobility in the ENC will conclude the chapter.

4.1. Distributional properties of nouns and plural marking in Persian

This section starts with a brief introduction to general properties of nouns and their distribution in Persian. This sets the stage for a detailed discussion of plural marking in the Persian noun phrase in the remainder of the section.

Persian allows nouns to occur bare in argument positions. Bare nouns¹ are construed as either definite or kind-referring². Indefinite noun phrases are marked with the indefinite enclitic *-i*, which yields quantitative reading of a noun (Ghomeshi 2003). Definite bare nouns in direct object positions (2)b are distinct from indefinite (3)b and kind-referring (4)b ones in that they must be followed by the Case marker *-râ* (*-ro* after vowels and *-o* after consonants in colloquial Persian).

- | | | | | | | | | | |
|-----|----|-----------------------------|-------|-------------|--|----|-----------------------|---------|--------------|
| (2) | a. | ketâb | ru-ye | miz-e. | | b. | Sinâ | ketâb-o | xund. |
| | | book | on-EZ | table-CL.is | | | Sina | book-OM | read.PST.3SG |
| | | ‘The book is on the table.’ | | | | | ‘Sina read the book.’ | | |

¹ For a thorough description and analysis of plural marking and bare nouns in Persian, see Ghomeshi (2003) and (2008), respectively.

² In this thesis, I am using kind-referring interchangeably with generics.

- (3) a. ketâb-i gom=šod. b. ketâb-i xarid-am.
 book-IND lost=became.3SG book-IND bought-1SG
 ‘A book was lost.’ ‘I bought a book.’
- (4) a. ketâb dust-e mâ-st. b. Sinâ ketâb xund.
 book friend-EZ we-CL.is Sina book read.PST.3SG
 ‘Books are our friends.’ ‘Sina read books.’

Nouns in Persian are most commonly marked with plural by adding the suffix *-hâ* to their simple forms³ (in spoken Persian, *-hâ* is pronounced *-â* when attached to consonant-final stems, as exemplified in (5)b):

- (5) a. xune-hâ ‘houses’ b. ketâb -â ‘books’
 qu-hâ ‘swans’ gol-â ‘flowers’
 dânešju-hâ ‘students’ kârgar -â ‘workers’

Definite, indefinite, and kind-referring noun phrases can have plural marking. Like definite singular nouns, definite plural nouns are required to appear with *-râ* when they are in direct object position, as in (6)b.

³ In formal language, the tendency is to use *-hâ* for inanimate nouns and *-ân* for animate ones. Also words of Arabic origin follow Arabic rules of plural formation yielding a variety of ways of marking plurality. In informal language, all these different plural forms may alternatively appear with *-hâ*.

(6) *Definite*

- a. ketâb-â ru-ye miz-e. b. Sinâ ketâb-â-ro xund.
book-PL on-EZ table-CL.is Sina book-PL-OM read.PST.3SG
'The books are on the table.' 'Sina read the books.'

(7) *Indefinite*

- mâh-e gozašte ketâb-â-ye xub-i xund-am.
month-EZ past book-PL-EZ good-IND read.PST-1SG
'Last month, I read some good books.'

(8) *Kind-referring*

- ketâb-â dust-â-ye xub o qâbel-e+e'temâd-i hast-and.
book-PL friend-PL-EZ good and able+reliance-IND is-3PL
'Books are good and reliable friends.'

<http://sardargomestan.persianblog.ir/1386/3/>

Plural marking does not occur in indefinite noun phrases introduced by numerals+classifiers whether they are in subject (9) or direct object position (10).

- (9) se-tâ gonješk ru deraxt nešast-e=bud-an.
three-CLS sparrow on tree sat-PP=was-3PL
'Three sparrows were sitting on the tree.'

- (10) *se-tâ gonješk did-am.*
 three-CLS sparrow saw-1SG
 ‘I saw three sparrows.’

However, noun phrases involving numerals+classifiers can optionally co-occur with plural marking when they are definite. Plural marking in this context functions as a definite marker, as its presence induces an inclusive reading for the noun phrase which contrasts with indefinite cardinality reading. Specifically, the plural marker *-hâ* in (11) triggers a reading which involves reference to the three books as a whole, not as a set of three individual items. In this sense, *-hâ* induces a universal quantifier reading, i.e. all the three books. This is empirically supported by the fact that the quantifier *har* ‘every’ can occur in this noun phrase for emphasis, yielding the same interpretation (12).

- (11) *se-tâ ketâb(-hâ) ru-ye miz-e.*
 three-CLS book-PL on-EZ table-CL.is
 ‘The three books are on the table.’

- (12) *har se-tâ ketâb(-hâ) ru-ye miz-e.*
 every three-CLS book-PL on-EZ table-CL.is
 ‘All the three books are on the table.’

The fact that plural marking in such contexts is optional raises the question as to how the (in)definiteness reading of noun phrases like *se-tâ ketâb* ‘three books’ in (13) is

determined when the plural marker is absent. I show in the remainder of this section that, in the absence of plural marking, the definite or indefinite readings associated with noun phrases with numerals+classifiers are distinguished by prosody.

- (13) se-tâ ketâb ru-ye miz-e.
 three-CLS book on-EZ table-CL.is
 ‘Three books/ the three books are on the table.’

Definiteness is instantiated in Persian by both prosodic phrasing and prominence. In indefinite noun phrases with numerals+classifiers, the numerals and the head noun are equally stressed while in corresponding definite noun phrases only the head noun is stressed (PROMINENT syllables are marked by small caps).

- | | |
|---------------------------|--------------------|
| (14) a. <i>Indefinite</i> | b. <i>Definite</i> |
| SE-tâ keTÂB | se-tâ keTÂB |
| three-CLS book | three-CLS book |
| ‘three books’ | ‘the three books’ |

The prosodic level that reflects definiteness, I assume, is the Phonological Phrase. I take numerals+classifiers as forming a Phonological Phrase with other nominal constituents following them. Building on Nespor and Vogel’s (2007: Chapter 6) parameters for defining the domain of Phonological Phrases, I presume that a Phonological Phrase in Persian noun phrases is composed of the head and everything that

falls under its maximal projection. Accordingly, prosodic (in)definiteness marking of a noun phrase with numerals+classifiers (i.e. a CardP) applies to the head Card and whatever constituents under it.

Having proposed that a noun phrase with a numeral+classifier is a Phonological Phrase (ϕ), I propose that an indefinite noun phrase appears as consisting of two Phonological Words (ω) each with a PROMINENT syllable. The two Phonological Words are the combination of numeral+classifier and the head noun. In contrast, definite noun phrases with numerals+classifiers involve a Phonological Phrase which consists of one Phonological Word, with the prominent syllable in the rightmost element. I suggest that this phonological phrasing pattern is a prosodic correlate of definiteness in noun phrases involving numerals⁴.

(15) a. *Indefinite*

b. *Definite*

[[SE-tâ]_ω [keTÂB]_ω]_φ

[[se-tâ keTÂB]_ω]_φ

‘three books’

‘the three books’

With this general overview of the distributional properties of nouns and plural marking in the Persian noun phrase, let’s now turn to a discussion of plural marking in the Empty Noun Construction (ENC).

⁴ Sadat-Tehrani (2007) treats similar cases in Persian as instances of the interaction between information structure and prosody. Based on his approach, (15)a, as the subject noun phrase in (13), is used as part of the new information. It can be an answer to the question “What is on the table?” When the subject noun phrase in (13) is given/old information, (15)b is used as part of the answer to the question “Where are the three books?”

4.2. Plural marking and empty nouns

In the ENC in Persian, the plural specification on the head noun has to be spelled out on postnominal modifiers—specifically, attributive adjectives (16). When there is no postnominal modifier, pre-head elements host the plural marker *-hâ* (17). In the presence of more than one attributive adjective in the Ezafe domain, the closest one to the empty noun is marked with plural, as shown in (18).

(16) $DemP[Dem[in/un\ nP[n[ketâb-\hat{a}-ye\ a[sefid]]]]]$

this/that book-PL-EZ white

‘these/those white books’

\Rightarrow [in/un [~~ketâb~~⟨-â⟩ [sefid-â]]]

‘these/those ~~book~~⟨-PL⟩ white [ones]’

(17) $DemP[Dem[in/un\ nP[ketâb-\hat{a}]]] \Rightarrow [in-\hat{a}/un-\hat{a}\ [~~ketâb~~⟨-\hat{a}⟩]]$

this/that book-PL

‘these/those [ones]’

‘these/those books’

(18) $DemP[Dem[in/un\ nP[n[ketâb-\hat{a}-ye\ a[sefid]-e\ a[qadimi]-ye\ PP[tu-ye$

this/that book-PL-EZ white old inside-EZ

ja:be]]]]]

box

‘these/those old white books in the box’

⇒ [in/un ~~ketâb~~⟨-â⟩ [sefid-â]-ye [qadimi]-ye [tu-ye ja:be]]
 ‘these/those old white [ones] in the box’

Notice that in (18), the plural marker cannot appear with any other modifiers except the closest one to the empty noun. This is shown in (19):

(19) * in/un ~~ketâb~~⟨-â⟩ [sefid]-e [qadimi-**hâ**]-ye [tu-ye ja:be]
 * in/un ~~ketâb~~⟨-â⟩ [sefid]-e [qadimi]-ye [tu-**hâ**-ye ja:be]

There are various ways of marking plurality in Persian (cf. ft. 3). This study is restricted to plural marking with *-hâ* and the appearance of this suffix with the constituents in the ENC. In fact, only *-hâ* and *-ân* can appear on noun phrase-internal constituents in the ENC. Consider the following example in which *-ân* moves onto the adjacent modifier:

(34) hamvatan-**ân**-e aziz-e sâken-e xârej az kešvar-e mâ
 compatriot-PL-EZ dear-EZ resident-EZ outside of country-EZ we
 ‘the dear fellow-countrymen residing outside of our country’

⇒ ~~hamvatan~~⟨-ân⟩ aziz-**ân**-e sâken-e xârej az kešvar-e mâ
 ‘the dear ones residing outside of our country’

In cases where the plural noun is formed via Arabic rules, only *-hâ* is possible on the alternative host in the ENC. This implies that *-hâ* in such contexts emerges as the default plural marker. In (35), *asâtid* ‘professors’ is the broken plural of *ostâd* ‘professor’.

(35) *dowlat asâtid-e qadimi-ye dânešgâh-o bâznešast mi-kon-e.*
 governmentprofessors-EZ senior-EZ university-OM retired DUR-do.3SG
 ‘The government retires the senior university professors.’

⇒ *dowlat asâtid qadimi-hâ-ye dânešgâh-o bâznešast mikone.*
 ‘?the government retires the senior university ones’

In this section we investigate further which modifiers can host the stranded (default) plural marker in the ENC.

Among the pre-head elements, only demonstratives and the categories generated in the specifier of QP (viz. quantifiers, superlatives, super-ordinals, and interrogatives) can host the plural marker; numerals and classifiers cannot. Demonstratives, exemplified in (16), were shown to take plural marking in the absence of the head noun. Similarly, interrogatives and superlatives can appear with the plural marker, as shown in (20) and (21):

(20) *Interrogatives*

az [ketâb-â-ye čâmski] kodum-â ketâb(-â) be fârsi
of book-PL-EZ Chomsky which-PL book(-PL) to Persian
tarjome=šod-and?
translation=became-3PL

‘Of Chomsky’s books, which [ones] have been translated into Persian?’

(21) *Superlatives*

...hamiše [ahmaqâne-tar-in kâr-hâ] [jâleb-tar-in-hâ kâr(-hâ)]
always stupid-COMP-SUP work-PL interesting-COMP-SUP-PL work(-PL)
hast-and!
is-3PL

‘... the most stupid works are always the most interesting [ones]!’

[A comment by one of the visitors of *Ashkan-e Khajenoori*’s weblog at

http://khajehnoori.com/2005/11/blog-post_15.html]

In Persian, it is not common to use the plural marker with super-ordinals. The only super-ordinal that is commonly used with the plural marker is *avvalin* ‘first’⁵.

⁵ I was not able to find a single example of pluralized super-ordinals of two to ten in my Google search. Even *avvalin-hâ* is mostly used with an arbitrary reading, referring to the first events happened, things found, or people’s important achievements. In this sense, the word *avvalin-hâ* can be taken as the equivalent of ‘firsts’ in English.

(22) *Super-ordinals*

... va çe hâl-i dâšt šenidan-e sedâ-ye Yaqmâyi ke az
and what joy-IND had.3SG listen.INF-EZ voice-EZ Yaghmayi that from
[avval-in-hâ sedâ<-hâ>] bud ke mojavvez peydâ=kard barâ-ye
first-SUPORD-PL voice<-PL> was that permit found=did.SG for-EZ
entešâr.

publication

‘...and how enjoyable it was to listen to Yaghmayi’s voice that was among the first
which received a publication permit.’

[From *Gardoone Adabi* weblog at

<http://maroufi.malakut.org/gardoone/archives/004627.shtml>]

In the ENC, quantifiers occurring with count nouns tend to appear with plural marking. This is shown in the following table and exemplified in (24). Quantifiers used with mass nouns or those with neutralized mass/count distinction never appear with the plural marker.

(23) *The possibility of plural marking quantifiers in the ENC*

	<u>QUANTIFIER</u>	<u>TYPE OF COMPLEMENT</u>	<u>GLOSS</u>
*	besyâri-hâ	N _{count}	‘many+PL’
✓	bištar-i-hâ	N _{count}	‘most+PL’
✓	xeyli-hâ	N _{count}	‘many+PL’
✓	ba:zi-hâ	N _{count}	‘some+PL’

*	kami-hâ	N _{mass}	‘a little+PL’
*	qadri-hâ	N _{mass}	‘somewhat+PL’
*	hame-hâ	N _{mass/count}	‘all+PL’
*	tamâm-â	N _{mass/count}	‘all+PL’

(24) az [hamkelâsi-hâ-ye dânešgâh-am], $\left(\begin{array}{l} \text{bištar-i-**hâ**} \\ \text{xeyli-**hâ**} \\ \text{ba:zi-**hâ**} \end{array} \right) \left(\begin{array}{l} \text{hamkelâsi} \langle -hâ \rangle \\ \text{hamkelâsi} \langle -hâ \rangle \\ \text{hamkelâsi} \langle -hâ \rangle \end{array} \right)$
of classmate-PL-EZ university-CL.1SG most/many/some-PL classmate $\langle -PL \rangle$
raft-and xârej.
went-3PL abroad
‘Of my university classmates, most/many/some have gone abroad.’

It should be mentioned that the quantifiers exemplified in (25) are interpreted as having an arbitrary antecedent when there is no specific linguistic or extra-linguistic antecedent involved. In this case, they mean ‘most/many/some *people*’ respectively.

(25) $\left(\begin{array}{l} \text{bištar-i-**hâ** } e_{arb} \\ \text{xeyli-**hâ** } e_{arb} \\ \text{ba:zi-**hâ** } e_{arb} \end{array} \right) \text{mi-tars-and} \quad \text{e:terâz=kon-and.}$
most/many/some-PL DUR-fear.PRS-3PL protest=do.PRS-3PL
‘Most/many/some people are afraid to protest.’

Numerals and classifiers can never occur with the plural marker in the ENC.

- (26) in se-tâ ketâb ⇒ * in se-tâ-**hâ** ketâb
 this three-CLS book
 ‘these three books’

As can be seen in the example above, the head noun following the numeral+classifier is not marked with plural either.

As noted earlier, the occurrence of the plural marker with numerals+classifiers is possible when the noun phrase is definite. Consider the following example:

- (27) in se-tâ jurâb-**â**-ro barâ Sinâ xarid-am.
 this three-CLS sock-PL-OM for Sina bought-1SG
 ‘I bought these three [pairs of] socks for Sina.’

The appearance of *-hâ* on the head noun in (27), however, does not mean that in the absence of the head noun it can move onto the numeral+classifier (28)a, or the demonstrative *in* ‘this’ (28)b. When the noun is not pronounced, the plural marker cannot appear in such cases (29).

- (28) a. * in se-tâ-**hâ** jurâb(-hâ)-ro ... b. * in-**â** se-tâ jurâb(-hâ)-ro ...
 (29) in se-tâ jurâb-ro ...

As shown in (30), the stranded plural marker in such circumstances can be displaced only when there is an attributive adjective adjacent to the head noun; otherwise, it does not appear in the ENC.

- (30) in se-tâ jurâb-â-ye qermez-o barâ Sinâ xarid-am.
 this three-CL sock-PL-EZ red-OM for Sina bought-1SG
 ‘I bought these three [pairs of] red socks for Sina.’

⇒ in se-tâ ~~jurâb~~⟨-hâ⟩ qermez-â-ro barâ Sinâ xarid-am.
 ‘I bought these three red ones for Sina.’

As for indefinite noun phrases, even in the presence of the attributive adjective, the empty noun is not allowed. This is regardless of whether or not the noun phrase involves numerals+classifiers and is illustrated by the ungrammaticality of (31) and (32):

- (31) * Sinâ ~~jurâb~~⟨-hâ⟩ arzun-â-yi xarid.
 Sina ~~sœk~~⟨-hâ⟩ cheap-PL-IND bought.3SG
 ‘Sina bought some cheap ones.’

- (32) * Sinâ se-tâ ~~jurâb~~⟨-hâ⟩ arzun-â-yi xarid.
 three-CLS
 ‘*Sina bought some three cheap ones.’

contrastiveness condition on remnant modifiers and is therefore not spelled out, the next closest overt adjective takes over the same properties and acts as the host of the stranded plural marker. For instance, in (34), *qadimi* ‘old’ can take the plural marker in the absence of *sefid* ‘white’.

(34) [in/un [~~ketâb-sefid~~(-â) [qadimi-â]-ye [tu-ye ja:be]]]

‘these/those old [ones] in the box.’

As we established in §3.1.2, none of the phrasal modifiers in the Ezafe domain licenses empty nouns⁶ and consequently appears with the plural marker. Having surveyed the possible hosts of the stranded plural marker in the ENC, we showed in this section that the plural marker moves onto the modifier that licenses the empty noun and can occur with plural marking. Additionally, it was shown that the displacement of the plural marker is constrained by a locality condition according to which the plural marker can attach only to string-adjacent licensing modifiers. In the following section, we will look at the syntactic properties of plural marking in Persian and the role they play in motivating *-hâ* mobility in the ENC.

4.3. The syntax of plural formation

Ghomeshi (2003) argues that number marking in Persian and English involves two different functional structures: D/QP (DP as Determiner Phrase and QP as Quantifier

⁶ See footnote 6 in chapter 3 for an exception to this generalization.

Phrase) in Persian and NumP in English. She assumes DP as the projection in which plural marking triggers a definite reading in Persian. The co-occurrence of plural marking and the indefinite enclitic *-i* rules out the definiteness effect of plural marking. Yet, the fact that Persian, unlike English, allows such a co-occurrence indicates that the indefinite enclitic has a different nature compared to the indefinite determiner in English *a/an*. Ghomeshi argues that the indefinite enclitic *-i* “occurs in contexts in which English *any/some/no* occur, and that its presence can result in a specific reading for a noun.” (p. 64). Specifically, she proposes that *-i* is a quantitative indefinite determiner which heads a Quantifier Phrase. The difference between Persian and English in having different functional structures for number marking is further supported by the evidence that bare nouns in Persian are allowed to appear in argument positions while English arguments must minimally be NumPs. Ghomeshi argues that this observation explains why bare singular count nouns receive a mass reading in English but not in Persian.

(35) *Persian*

a. [NP sag] did-am.
 dog saw-1SG

‘I saw dogs.’

English

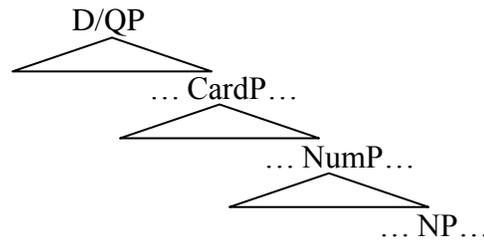
b. I saw [NumP mass dog].

[Ghomeshi 2003: 71. 72]⁷

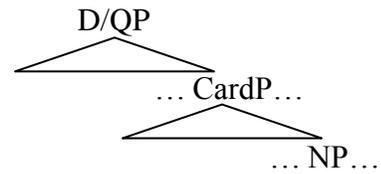
Ghomeshi concludes that the difference in the interpretation of Persian and English bare count nouns in argument position amounts to assuming that Persian lacks a NumP projection.

⁷ For consistency, I have adjusted the original glosses.

(36) a. *English*



b. *Persian*



[Ghameshi 2003: 71.71]

In the same vein, Ghomeshi argues that plural marking in Persian is licensed only on noun phrases contained within D/QP. The evidence she provides in support of this proposal is the fact that plural nouns are construed as definite when there is no overt marker of indefiniteness. The noun phrase, under this circumstance, will be contained within a DP headed by a null D specified for definiteness (37)a. In cases where the indefinite enclitic *-i* co-occurs with the plural marker, as in (37)b, the noun phrase will be contained within a QP. She proposes that plural marking is licensed in the following configurations:

(37) a. [DP [NP N-PL] Ø_{DEF}]

b. [QP [NP N-PL]-i]

[Ghameshi 2003: 67.56c&d]

Although Ghomeshi (2003) acknowledges that the differences in number marking across languages are due to different ways that comparable morphemes correspond to the syntactic structure, she does not explicitly elucidate how this correspondence takes place in Persian. This is the issue that I will address in this section.

4.3.1 Definiteness/number syncretism

The description of the possibilities of plural marking in Persian in §4.2 and Ghomeshi's (2003) proposal in (37) suggest that the plural marker *-hâ* marks plurality for definite and indefinite noun phrases. However, in absence of any overt definiteness marking, *-hâ* has the added value of marking definiteness. This is particularly evident in the case of definite noun phrases involving numerals+classifiers. In such cases, the plural marker simply functions as a definite marker. As will be discussed in §4.4.3, plural marking cannot co-occur with numerals+classifiers in indefinite noun phrases because numerals are inherently specified with the feature [pl]; the presence of the plural marker is thus rendered redundant and not economical.

Ghaniabadi (in prep.) argues that plural marking on a certain class of mass nouns also induces a definiteness reading. There I present a set of mass nouns like *electricity*, *water*, *snow*, *ice*, *sand*, *fog* which can optionally be pluralized in episodic contexts without inducing coerced (i.e. count) readings. I call this class of mass nouns *non-coerced mass nouns*. Plural marking on these mass nouns results in definite descriptions involving inclusiveness. That is, the reference of these nouns is to the totality of the mass that satisfies the description. I take this semantic characteristic as evidence to argue that plural marking on non-coerced mass nouns functions as a definite marker⁸. Consider the following examples:

⁸ Inclusiveness, as the semantic property of plural marking on non-coerced mass nouns in Persian, is the feature that Lyons (1999) attributes to definite plural and mass nouns (see Ghaniabadi, in prep., for more details).

(38) barq(-â) ke mi-r-e, aksolamal-e šomâ či-ye?
electricity(-PL) that DUR-go-3SG reaction-EZ you what-CL.is

‘What is your reaction when the power shuts off?’

<http://www.cloob.com/club/post/show/clubname/tabrizlilar/topicid/1646777/wrapper/true>

(39) ru mâse(-hâ) či nevešt-i?
on sand(-PL) what wrote-2SG

‘What did you write on the sand(*s)?’

In (38), *barq-â* ‘the power’ refers to the whole power supply of the location the speaker is talking about. *Mâse-hâ* ‘the sand’ in (39) refers to the whole area of the sand on which something is written. In analyzing these examples, I have noted that what the definite article does in English translations of plural non-coerced mass nouns, plural marking does in Persian.

In indefinite noun phrases, however, plural marking on examples of non-coerced mass nouns does not yield non-coerced readings; it rather results in count, i.e. coerced, readings, as shown in (40). I take this as an indication that *-hâ* in the context of indefinite noun phrases only functions as a number marker.

(40) zemestun-â barf-â-ye sangin-i mi-bârid.
winter-PL snow-PL-EZ heavy-IND DUR-fell.3SG

‘In winters, it snowed heavily.’ (lit. some heavy snows fell.)

<http://www.tiknews.org/display/?ID=70669&page=30>

In (40), *barf-â*, the plural form of ‘snow’, refers to snowfalls.

The following table presents a summary of what has been discussed so far about the possibilities of plural marking in Persian:

(41) *Plural marking on Persian noun phrases*

		<i>-hâ</i>		
		DEFINITE	INDEFINITE	KIND
COUNT		✓	✓	✓
MASS	<i>Coerced</i>	✓	✓	✓
	<i>Non-Coerced</i>	✓	*	*
NUMERALS+CLASSIFIERS		✓	*	n/a

According to the table in (41), plural marking can occur in all kinds of count noun phrases and coerced mass noun phrases. However, in non-coerced mass noun phrases and noun phrases involving numerals+classifiers, plural marking is possible only when the noun phrase is definite. If we assume that the plural marker *-hâ* in definite noun phrases plays the role of a definite marker as well as a number marker, it follows that plural marking in indefinite noun phrases only marks number. Thus, we can argue that in indefinite noun phrases, plural marking is not possible with mass nouns because they are not generally marked with number, unless a count reading is intended (see Ghaniabadi, in prep.). In explaining why in indefinite noun phrases numerals+classifiers do not co-occur with plural marking, I propose that numerals are inherently specified with [+pl] and thus prevent the spell-out of the plural marker, which bears the feature [+pl] lower in the structure.

Considering the fact that Persian uses a number marker to mark noun phrases for definiteness, I propose that *definiteness is syncretic with number* in Persian. Thus, I suggest that *-hâ* functions as a definite/number marker when its number feature is bundled with the definite feature, and as a plural marker when it is specified only with the number feature. As such, in definite noun phrases *-hâ* is a portmanteau morpheme with a feature bundle [+def, +pl], but in other contexts, *-hâ* is a homophonous morpheme which realizes plural number. This is shown schematically in (42):

(42) *-hâ* : plural and definite plural marker



These two plural morphemes have different singular correlates. The singular of the plural marker is the unmarked noun but the singular of the definite plural marker can optionally be realized as a stressed vowel *-é* or its allomorph *-á* (before consonants).

(43) ketâb(-é) ru miz-e.
 book(-SG_{DEF}) on table-CL.is
 ‘The book is on the table.’

Based on this discussion, I propose a number system for Persian in which number markers and definite number markers enter into different system of contrast. This number system is shown in (44).

(44) *Number system in Persian*

	NUMBER	DEFINITE NUMBER
Singular	Unmarked	-é (SG _{DEF})/ -∅
Plural	-hâ (PL)	-hâ (PL _{DEF})

Based on this number system, definite noun phrases are distinguished from other types of noun phrases by the feature composition of their number markers. If the feature [def] is bundled with a value of the feature [pl], the number marker triggers a definite reading. If not, the noun phrase is ambiguous between being indefinite and kind-referring.

Having established that plural marking in Persian, in addition to marking plurality, can function as a definite marker, we now need to determine its categorial nature and how it is accommodated in the Persian syntax.

4.3.2 Persian plural marker: a head or a modifier?

Building on distributional properties of plural markers in English and Halkomelem, Wiltschko (2008) proposes that plural markers across languages can be either a projecting functional head (as in English) or a modifier adjoined to a head (as in Halkomelem). She proceeds to suggest that plural markers that merge as heads have the syntax of heads and modifying plural markers will have the syntax of adjuncts. She puts forward a number of criteria as diagnostics for distinguishing the categorial status of plural markers as heads vs. modifiers. According to Wiltschko (2008), modificational plural makers “[...] do not trigger agreement, their absence is not associated with a specific meaning but instead is

truly unmarked, [...] and do not allow for form-meaning mismatches” (p. 688). Let us now apply these three diagnostics to the Persian plural marker⁹.

First, plural marking in Persian does not trigger agreement. This is shown in (45) where demonstratives do not agree with the head noun in number¹⁰.

- (45) in(*-â) pesar-â mi-dov-and.
this(*-PL) boy-PL DUR-run-3PL
‘These boys are running.’

Second, in Persian, unlike English, the unmarked (singular) noun, e.g. *ketâb* ‘book/books’, is compatible with a plural interpretation; therefore, Persian has what Corbett (2000) calls *general number*.

Third, Wiltschko predicts that form-meaning mismatches arise if plural marking is inflectional (i.e. a syntactic head). She attributes the form-meaning mismatches in examples of *pluralia tantum* in English to the head status of plural marking in this language. Persian does not have any cases of *pluralia tantum*. Nouns instantiating

⁹ Wiltschko (2008) introduces other criteria as diagnostics for modificational plural markers. I do not use those diagnostics because they are too complicated and not fine-grained enough to capture the choice of the plural marker in Persian (see Wiltschko’s work for the details of those diagnostics).

¹⁰ Another instance of lack of agreement in Persian is constructions with inanimate plural subjects. In such constructions, the verb does not agree with the subject in number, e.g.

barg-â rixt.
leaf-PL fell.3SG
‘Leaves fell.’

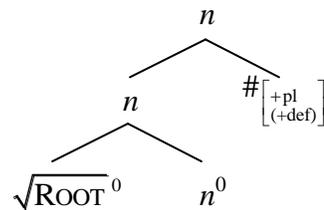
I do not consider this as viable evidence in support of the modificational status of the plural marker in Persian because I believe lack of agreement in such contexts is triggered by “the degree of volition or control of the verbal action by the subject” (Sedighi 2005: 188). According to Sedighi, agenthood, not subjecthood, is responsible for verbal agreement in Persian. Therefore, because inanimate plural subjects are not agents, they do not agree with verbs (see the original work for further details).

pluralia tantum in English, e.g. *scissors*, *pants*, *glasses*, etc., always appear singular in Persian. When pluralized, they are construed as referring to non-atomic sets of pairs.

- | | | |
|---------------------|------------------|--------------------|
| (46) a. qeyči-hâ | b. šalvâr-â | c. eynak-â |
| scissors-PL | pants-PL | glasses-PL |
| ‘pairs of scissors’ | ‘pairs of pants’ | ‘pairs of glasses’ |

Building on the observations above, I propose that the Persian plural marker is a modifier adjoined to the head noun and not a projecting functional head:

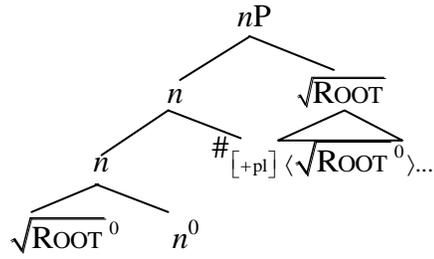
(47) *Persian plural marker as a modifier*



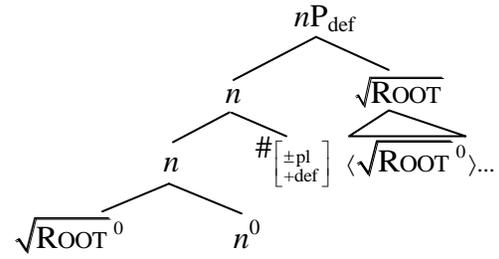
4.3.3 Proposal

Based on the generalizations that we have made about plural marking and definiteness in Persian so far, I propose, within DM, that the syntactic structures in (48) are input to Morphology.

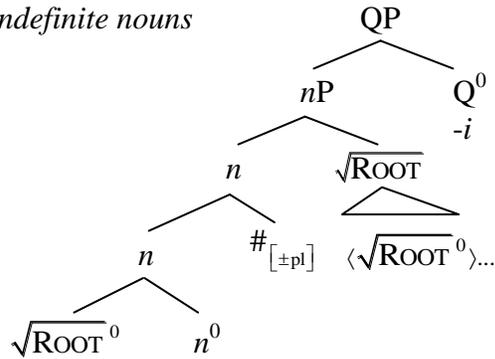
(48) a. *Kind-referring nouns*



b. *Definite nouns*



c. *Indefinite nouns*



The phonological exponents of morphemes are inserted at Vocabulary Insertion. Assuming that PL and PL_{DEF} are two independent morphemes with different feature specifications, it is expected that Vocabulary Insertion assigns different Vocabulary Items (VIs) to each. As we discussed earlier, the homophonous *-hâ* is the phonological exponent for both of these two morphemes, as schematically shown in (49) and (50).

(49) *VI for the definite plural marker in Persian*

$$-hâ \leftrightarrow \begin{bmatrix} +pl \\ +def \end{bmatrix}$$

(50) *VI for the plural marker in Persian*

$$-h\hat{a} \leftrightarrow [+pl]$$

(51) *VI for the definite singular marker in Persian*

$$-\emptyset / -\acute{e} \leftrightarrow \begin{bmatrix} -pl \\ +def \end{bmatrix}$$

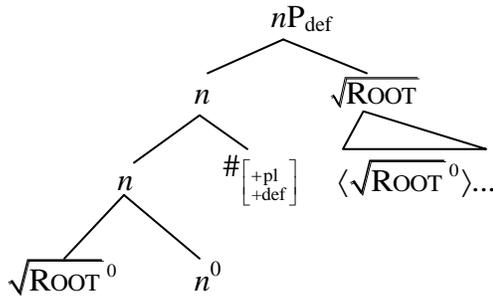
In (49), we are dealing with a portmanteau (syncretic) plural morpheme with the feature bundle [+def, +pl] which occurs only in definite noun phrases. In (50), we have the VI for the plural marker which can occur in contexts other than definite noun phrases. The VI for the definite singular is represented in (51). This morpheme can have either a null exponent or be realized as *-é*.

We discussed earlier two types of noun phrases in which *-hâ* induces a definite reading: noun phrases with non-coerced mass nouns and noun phrases involving numerals+classifiers. Let us examine how the above proposal can account for the occurrence of *-hâ* in these two environments.

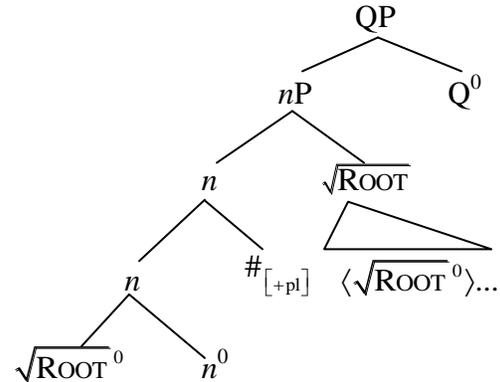
As we showed in §4.4.1, plural marking on mass nouns like *barq* ‘power’, *mâse* ‘sand’, and *barf* ‘snow’ results in a non-coerced reading only when it occurs in definite noun phrases. In indefinite noun phrases, however, the occurrence of *-hâ* results in a coerced (count) reading (cf. (40)). The difference in the interpretations of *-hâ* can be accounted for if we assume that the node for *-hâ* in definite and indefinite noun phrases has different morphosyntactic feature specifications. I propose that the non-coerced reading of mass nouns is obtained when *-hâ* is inserted for PL_{DEF} , while the coerced

reading is obtained when *-hâ* is inserted for PL. The following syntactic structures represent the contexts in which plural marking occurs with mass nouns¹¹:

(52) a. *Non-coerced reading*



b. *Coerced reading*



where $\sqrt{\text{ROOT}} = \{ \sqrt{\text{BARQ}}, \sqrt{\text{M\^A} \text{SE}}, \sqrt{\text{BARF}}, \text{etc.} \}$

‘power’ ‘sand’ ‘snow’

As we have seen, noun phrases involving numerals+classifiers do not occur with plural marking unless they are definite, in which case the portmanteau plural morpheme (PL_{DEF}) can optionally occur. From this follows that these noun phrases never occur with the non-portmanteau PL. Compare the following examples¹²:

¹¹ I have no explanation to offer for the correlation between being non-coerced vs. coerced and being definite vs. indefinite.

¹² See §4.2 for the prosodic distinction between definite and indefinite noun phrases involving numerals+classifiers.

As illustrated above, I propose that the head of CardP is inherently specified with the feature [+pl] and that this feature prevents the spell-out of the plural marker, which also bears the feature [+pl] lower in the structure. In other words, in such contexts, the realization of the feature [+pl] inherently specified on numerals renders the spell-out of another [+pl] feature redundant. This instantiates *redundancy avoidance* and can be stated as the following economy constraint, similarly proposed by Ortmann (2000)¹³:

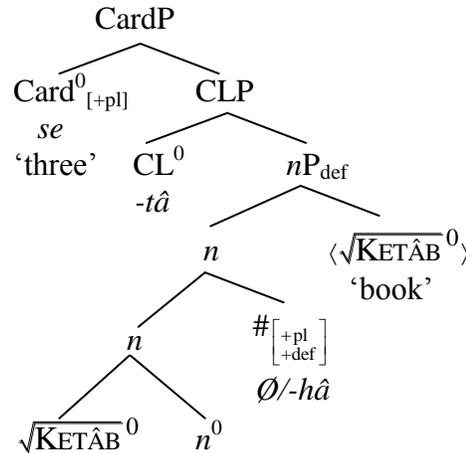
(55) *Principle of Avoiding Redundant Plural Marking*

There is no more than one realization of [+pl] within the noun phrase.

Our discussion so far has provided the answer to the first question. The second and third questions are concerned with plural marking in the definite noun phrases with numerals+classifiers. This kind of noun phrase is generated as a CardP that takes a definite *nP* as its complement.

¹³ Ortmann's (2000) *Principle of Economic Plural Marking* (PEPL) accounts for the disallowance of several noun-phrase internal realizations of plural in Hungarian and many other languages including Persian. He calls these languages as *Type Hungarian* and compares them with languages of *Type English*, in which several overt realizations of plurality are possible. To capture the difference between these two language types, he proposes an Optimality-Theoretic analysis in which plural marking in each of these language types occurs under a different constraint ranking. In Type Hungarian languages, Ortmann suggests, PEPL is undominated (see the original work for more details).

(56)



The node that is realized as *-hâ* in (56) is characterized by a feature bundle containing the features [+pl] and [+def]. I suggest that in this context, *-hâ* spells out [+def], not [+pl]. In other words, the co-occurrence possibility of plural marking and numeral is motivated by a morphological requirement that [+def] be spelled out. This is the answer to the second question.

The third question concerning the optionality of plural marking in the presence of numerals+classifiers is perhaps the most troubling question for which I cannot provide an answer within DM. However, I am going to propose an account inspired by constraint ranking system of Optimality Theory (Prince and Smolensky 1993) which addresses all the issues surrounding the occurrence of the plural marker in noun phrases with numerals+classifiers in Persian.

I propose that the appearance of the plural marker *-hâ* in noun phrases involving numerals+classifiers is determined by economy through the interaction between the following constraints:

- (57) (i) [+pl] must be phonologically realized.
(ii) [+def] must be spelled out.
(iii) Avoid redundancy in number marking.

In indefinite noun phrases, the presence of the plural marker results in ungrammaticality because economy disallows the co-occurrence of the plural marker and numeral+classifier. Thus, I propose that in deriving indefinite noun phrases with numerals, constraint (iii) outranks constraint (i). In definite noun phrases, the plural marker is optional because constraints (ii) and (iii) are tied. In other words, both possibilities are equally economical.

4.3.4 What category hosts *-hâ*?

Based on the proposal in the previous section, we expect to see plural marking only on nouns. However, we have already seen in the ENC that the plural marker *-hâ* attaches to attributive modifiers and some pronominal modifiers. This may lead us to hypothesize that plural marking in Persian can attach to nouns, adjectives, and some pronominal modifiers. However, this generalization is not borne out because plural marking cannot appear on adjectives (58)b or pronominal modifiers (58)c when the head noun is present. This observation entails that the plural morpheme in Persian is category sensitive, in the sense that it attaches only to nouns.

by a prenominal modifier in the next higher phase. We established that the derivation does not trigger spell-out unless the empty noun meets its required Licensing and Focus Conditions. The non-pronunciation of the head noun is obtained at PF when the E-feature specified on the nominalized Root ($\sqrt{\text{ROOT}}_{[E]} - n$) blocks the Vocabulary Insertion from inserting the phonological exponent of the head noun. This phenomenon has a morphophonological consequence for the ENC that contains the plural marker¹⁴ *-hâ*. When Vocabulary Insertion applies and the exponents are linearized, the presence of a phonologically empty noun leaves the exponent of PL/PL_{DEF} (*-hâ*) stranded, in need of an overt host. Relying on our description of plural marking in the ENC (see §4.3), the plural marker is not unpronounced with the empty noun; rather, it appears on the phonologically realized heads adjacent to it. Elaborating, *-hâ* will be hosted by the right adjacent attributive adjective (60)a; if there is none, the prenominal modifier (e.g. demonstrative, quantifier, superlative) is the alternative host (60)b.

Q	<i>n</i>	-PL _{DEF}	<i>a</i>	Possessor	
(60)	behtarin	dânešju	-hâ -ye	javân-e	dânešgâh
	best	student	-PL-EZ	young-EZ	university
	‘the best young students of the university’				

a. *behtarin* ~~*dânešju*~~ $\langle -hâ \rangle$ *javân-hâ*-ye *dânešgâh*
└──────────────────┘ ↑

‘the best young ones of the university’

¹⁴ I use *plural marker* from now on to refer to both the definite and indefinite morphemes.

b. *behtarin-hâ-ye dânešju(-hâ) dânešgâh*
 ‘the best ones of the university’

I suggest that this behavior follows from the following morphophonological condition:

(61) *Condition on plural marking in Persian (imposed at PF)*

-hâ must be spelled out on a phonologically realized head.

The above condition is in fact a condition on the implementation of the constraint (57)i requiring the phonological realization of the feature [+pl] in Persian noun phrases. As such, the condition in (61), on the one hand, sets up a morphological well-formedness condition whereby PL/PL_{DEF} must be spelled out and, on the other hand, specifies that the suffixation of the plural marker is dependent upon the phonological realization of the host, which by itself implies that plural marking happens after Vocabulary Insertion. This morphophonological condition is consistent with the general requirement for affixes to be spelled out on an *overt* host. We argued in §4.3 that the phonologically realized head that hosts the stranded *-hâ* is the adjacent head-adjoined attributive adjectival modifier, if there is one. As this operation is *Vocabulary sensitive*, i.e. sensitive to the phonological exponent of the VI, its application is subject to *The Local Dislocation Hypothesis* which states that “[i]f a movement operation is Vocabulary sensitive, it involves only string-adjacent items” (Embick & Noyer 2001: 566). The mechanics of this movement

operation constitutes the discussion of the next section. We will discuss the cases where there is no head-adjoined attributive adjectival modifier in §4.5.1.2.

4.4.1 *-hâ* mobility: a case for Local Dislocation

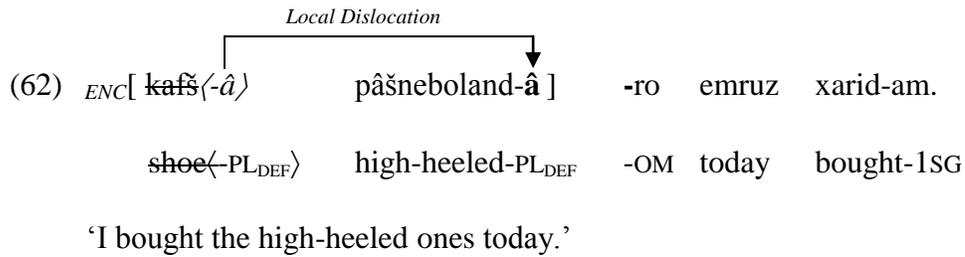
The fact that the displacement of *-hâ* in the ENC is Vocabulary sensitive makes the case for the LD operation to apply. As discussed in §1.1.2.3, I assume, following Embick (2003, 2007), that the input to LD is a Concatenation statement e.g. $X \sim Y$ generated by the Linearization process. The Concatenation statement provides the locality condition for LD to apply within a given cyclic domain (phase) at PF and adjoin string-adjacent elements of the same type. String-adjacent elements outside the domain in which LD applies cannot influence the operation and are thus inactive. In this section I will show that *-hâ* mobility in the ENC can be derived from LD operations that apply to Concatenation statements generated over M-Words in a given cycle of PF derivation. I will argue that the LD rule that adjoins the stranded *-hâ* to an overt adjectival head modifier applies on an earlier PF cycle than the one in which a prenominal modifier hosts the stranded *-hâ*. Thus, the application of LD to a Concatenation statement involving the stranded *-hâ* and the attributive adjectival modifier bleeds the application of the procedure which attaches *-hâ* to an apparently string-adjacent prenominal modifier that is outside the LD domain.

As discussed in Chapter 3, in the course of the incremental derivation of the ENC, spell-out occurs at the point where the empty noun (i.e. E-specified *n* Root) meets its Licensing and Focus Conditions. At PF, the syntactic output is subject to Vocabulary Insertion and Linearization, among other operations. Building on the discussion in

Chapter 1, I assume that Linearization takes place in a series of steps. I illustrate this below with two ENC_s containing the PL/PL_{DEF} morpheme.

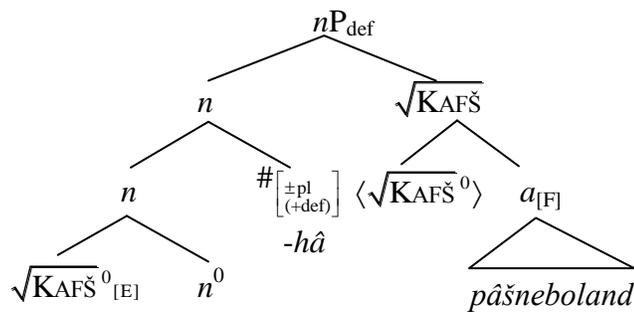
4.4.1.1 *-hâ* mobility within the *nP* phase

I begin with an ENC with an attributive adjective as its remnant modifier.



The ENC in (62) is an *nP* with the syntactic representation in (63). PF receives this representation and applies Linearization procedures to it.

(63) *Syntactic Structure*



The key properties of this structure can be noted as follows:

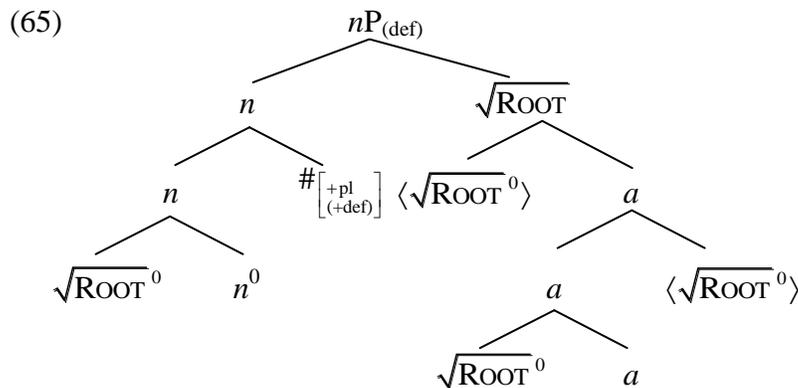
- The uncategorized Root is assumed to head-move to the nominalizing head, where it is specified with the E-feature. As a result, the n head becomes a maximally complex head (M-Word).
- The PL_{DEF} morpheme is base-generated as an adjunct to the complex n , forming part of its extended head-projection.
- The F-specified attributive adjectival modifier is a base-generated adjunct to the uncategorized n Root. The modifier consists of a non-projecting a Root, i.e. one that does not take complements. I consider this head modifier as an M-Word.

Before discussing the procedures involved in linearizing the structure in (63), an explanation is in order as to the morphological type of the attributive adjectival modifiers base-generated as adjuncts to the n Root. I distinguish these elements from those head-adjoined to the nominalizing head n via syntactic head movement. Recall from Chapter 1 that according to Embick & Noyer (2001) and Embick (2003, 2007), there are two types of objects in the morphological structure: M-Words and Subwords. Let us review the definitions again.

- (64) a. M-Word: (Potentially complex) head not dominated by further head-projection
(cf. Chomsky (1995) “ H^{0max} ”)
- b. Subword: Terminal node within an M-Word (i.e. either a Root or a feature bundle)

[Embick 2003: 307.2]

The M-Word status of the highest n in the structure of the nP is uncontroversial, as its constituent elements are terminal nodes internal to this complex head. In (63), for instance, the highest n is an M-Word containing $\sqrt{\text{KAFS}}^0_{[E]}$, n^0 , and PL_{DEF} as its Subwords. The M-Word status of the head-adjoined attributive adjectival modifier, however, is less obvious¹⁵. On the one hand, this constituent looks like a Subword because it is a part of a structure which resembles a complex head. On the other hand, given the DM theory of category determining morphology (Marantz 2001) adopted in this thesis, the adjectival head modifier is internally complex, involving the projections of a Root and a category-defining functional head a , and as such looks more like an M-Word. I take the position that head-adjoined adjectival modifiers in nP are M-Words and have the structure shown in (65)¹⁶.



¹⁵ This issue arises in any language in which attributive adjectives are assumed to be head-adjoined to nouns (see e.g. Sigurðsson 1993 for an analysis of the Icelandic DP).

¹⁶ See also the discussion on modifier-Root incorporation in §3.3.1.1 where I assume the same internal structure for adjectival head modifiers.

The difference in behavior between Root-adjoined adjectival modifiers and elements adjoined to the nominalizing head *n* can be explained more explicitly with reference to compounds in Persian.

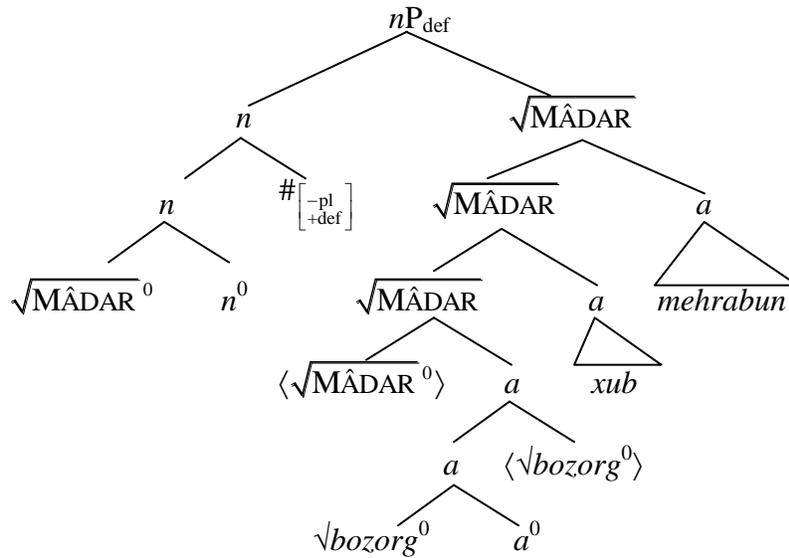
As I further develop in Chapter 5 (§5.1.3.1), compounding in Persian applies within the *nP* domain, where it takes one or two head modifiers and combines them with the nominal Root. Compounds are syntactically represented under *n* and head modifiers under $\sqrt{\text{ROOT}}$. As such, head modifiers adjoined to $\sqrt{\text{ROOT}}$ are part of the noun phrase and hence are M-Words while those compounded with the nominal Root are part of the head noun (i.e. the complex *n* head) and are Subwords. Consider the following examples in which the attributive adjectival modifier *bozorg* ‘big’ and the noun *mâdar* ‘mother’ in (66)a are combined together in (66)b forming the compound *mâdar-bozorg* ‘grandmother’:

<i>n</i>	<i>a</i>	<i>a</i>	<i>a</i>
(66) a. <i>mâdar-e</i>	<i>bozorg-e</i>	<i>xub-e</i>	<i>mehrabun</i>
mother-EZ	big-EZ	good-EZ	kind
‘the big good kind mother’			

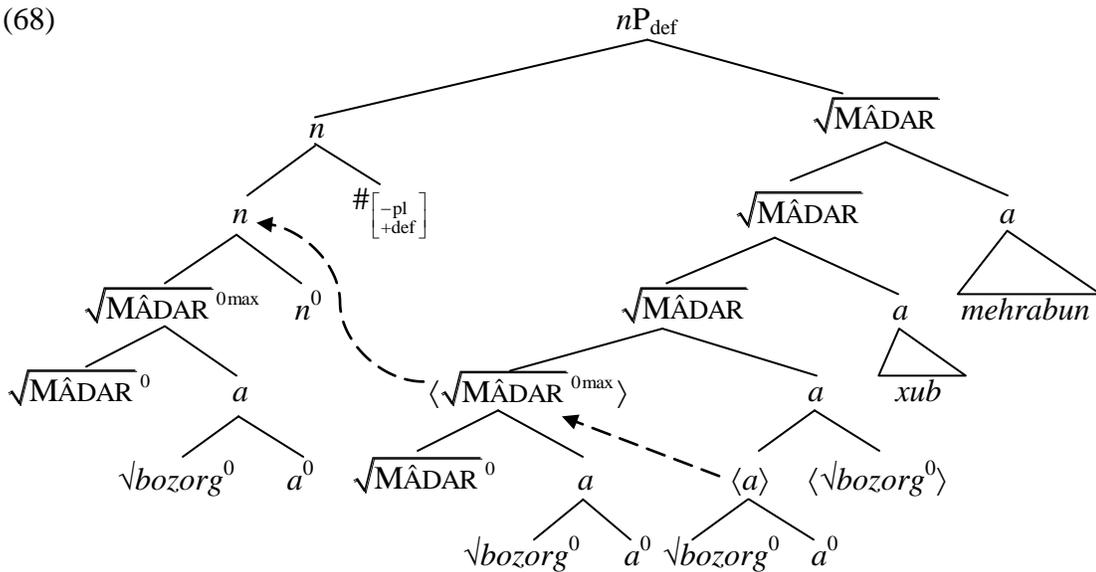
<i>n</i>	<i>a</i>	<i>a</i>
b. <i>mâdar+bozorg-e</i>	<i>xub-e</i>	<i>mehrabun</i>
‘the good kind grandmother’		

The following configurations represent the status of the adjectival modifier *bozorg* ‘big’ before and after compounding with the nominal Root $\sqrt{mâdar}$ ‘mother’ (see the detailed analysis in §5.1.3.1)¹⁷:

(67)



(68)



¹⁷ Following our convention in §3.2 and §3.3.1.1, I use the ‘⁰’ diacritic for terminal nodes and the ‘^{0max}’ diacritic for complex Roots.

Building on the assumption above that attributive adjectival modifiers are internally complex and are thus M-Words, I propose, following Harley (2008), that in compounding, the categorized modifier incorporates into the nominal Root before the resulting structure is adjoined to the nominalizing *n* head via head movement. As can be seen in (68), the adjectivized *bozorg* first incorporates into the uncategorized Root $\sqrt{mâdar}$. The incorporated structure then adjoins to *n* via syntactic head movement and becomes part of that M-Word. Therefore, *bozorg* ‘big’ as a head modifier in (67) is under Root and thus is an M-Word but as an element of a compound in (68) is under *n* and hence is a Subword.

Two empirical facts can be used to distinguish the M-Word status of Root-adjoined head modifiers from the Subword status of constituents of compounds under *n*: (i) stress assignment and (ii) position of plural marker.

STRESS ASSIGNMENT: In the *nP*, the head-adjunction structure headed by Root is distinguished from the one headed by *n* in terms of the stress assignment. The former, containing attributive adjectival modifiers, is assigned phrasal-level stress while the latter is assigned word-level stress. Accordingly, in the above examples, *bozorg* ‘big’ as a head modifier under Root receives stress within the prosodic domain of the noun phrase. When it appears as a part of the compound *mâdar-bozorg* ‘grandmother’ under *n*, it receives stress in the prosodic domain of the complex head *n*, which is a word-stress domain.

Assuming that in the Persian *nP*, the head noun forms a phonological phrase with its following modifiers, the non-compounded string in (66)a is a phonological phrase (ϕ), with the noun and each of its following adjectival modifiers being a phonological word (ω).

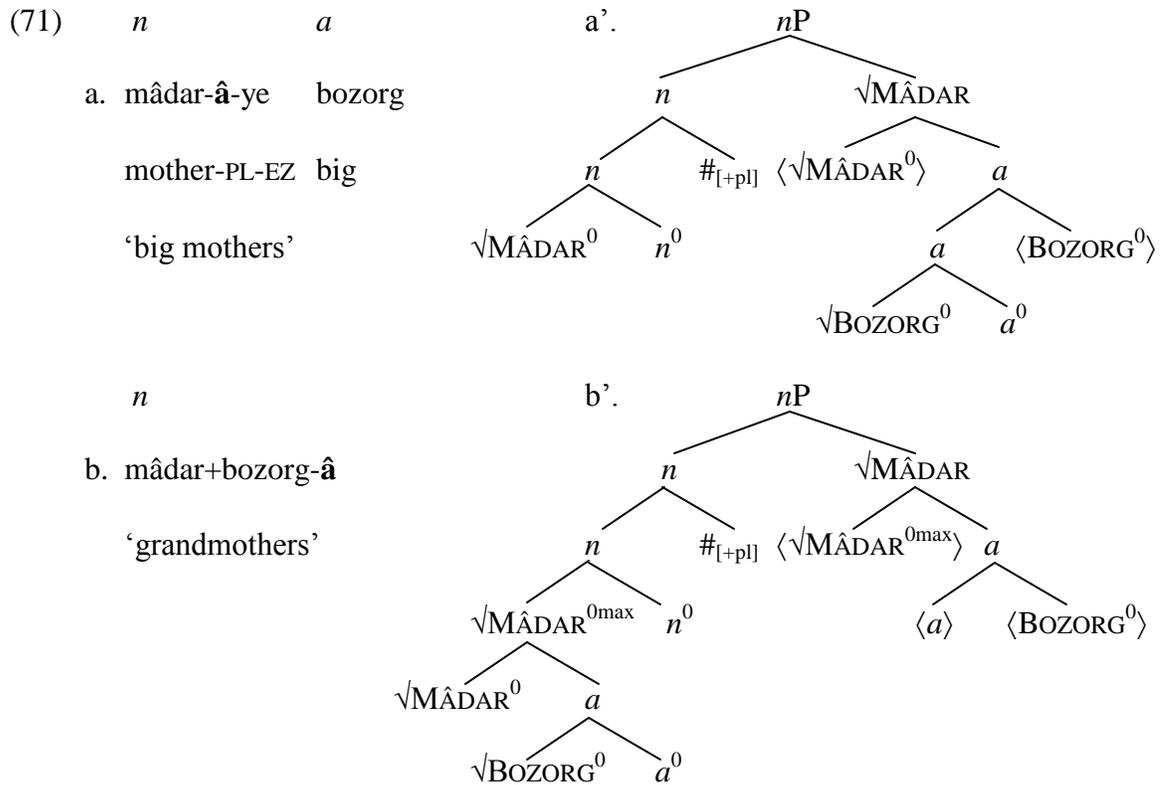
(69) [[mâDÀR]_{ω-e} [boZÒRG]_{ω-e} [XÙB]_{ω-e} [mehraBÚN]_ω]_φ

Notice that in (69) *mâdar* ‘mother’ and *bozorg* ‘big’ are assigned word stress independently. At the phrasal level, the primary stress is assigned to the last word in the *nP* and the preceding words are assigned secondary stress. In the syntax, this corresponds to the derivation where *bozorg* does not incorporate into the Root and the Root alone head-moves to the category-defining functional projection. As such, *bozorg* would be a Subword of the first Root-adjoined M-Word *a*. When *mâdar* and *bozorg* are compounded, the two form a single phonological word at PF and thus take word stress¹⁸. At the phrasal level, the whole compound receives a secondary stress, as shown in (70):

(70) [[mâdar+boZÒRG]_{ω-e} [XÙB]_{ω-e} [mehraBÚN]_ω]_φ

POSITION OF PLURAL MARKER: The position of plural morphology in the Persian *nP* is another criterion to determine the morphological type of attributive adjectives when head-adjoined to Root vs. when being part of a compound under *n*. As established in §4.3, the plural marker is attached outside the category-defining head *n* as an adjunct. When a noun is followed by an adjectival head modifier, as in (69), the plural marker appears on the noun and Ezafe connects it to the following modifier. However, when an attributive adjective appears as an element of a compound under *n*, as in (70), the plural marker *-hâ* can only appear on the whole compound. This contrast is shown in (71):

¹⁸ This is consistent with Embick’s (2003: 306) suggestion that Subwords are normally expected to have *word-level phonology*.

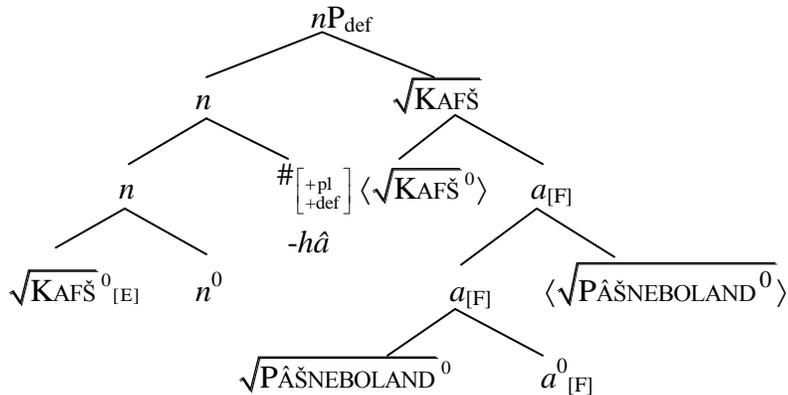


This observation implies that only M-Words can host the plural marker. Thus, the appearance of the plural marker on *mâdar* in (71)a shows that the noun (i.e. the complex n head) is an independent M-Word from its following attributive adjectival modifier *bozorg*. Additionally, the occurrence of the Ezafe linker with both the noun and the head modifiers in (69) corroborates that the complex n head and the Root-adjoined attributive adjectival modifiers are M-Words¹⁹. The fact that the plural marker cannot appear inside the compound in (71)b (**mâdar-â-bozorg*) is indicative of the Subword status of elements head-adjoined to n .

¹⁹ A detailed analysis of Ezafe insertion is presented in Chapter 5.

Having established the M-Words in the nP , we now look into how the structure in (63) is converted into linear order. Given our assumptions so far, the structure in (63) can be modified in the following manner.

(72) *Syntactic Structure*



As discussed in Chapter 1, Linearization involves two stages, namely Adjacency (LIN) and Concatenation. Below I briefly review and apply these procedures to the structure in (72):

(STEP 1) ADJACENCY: The structural adjacency or LIN procedure applies to each node X in the structure and generates a set of left-adjacency (*) statements between the two daughters of X , in accordance with the language-specific headedness of syntactic categories. For instance, the fact that Persian is head-initial determines that (P * DP) is the possible left-adjacency relationship and hence rules out (DP * P). To be more explicit, I expand the node a in (72) in the following *-statements to include the Root and the categorizing head a^0 .

(73) *Adjacency* *

$$\text{LIN}[a_{[F]}] \rightarrow (\sqrt{\text{P}\check{\text{A}}\text{SNEBOLAND}^0} \oplus a_{[F]}^0) * \langle \sqrt{\text{P}\check{\text{A}}\text{SNEBOLAND}^0} \rangle$$

$$\text{LIN}[\sqrt{\text{KAF}\check{\text{S}}}] \rightarrow \langle \sqrt{\text{KAF}\check{\text{S}}^0} \rangle * a_{[F]}$$

$$\text{LIN}[nP_{\text{def}}] \rightarrow ((\sqrt{\text{KAF}\check{\text{S}}^0}_{[E]} \oplus n^0) \oplus -h\hat{a}) * \sqrt{\text{KAF}\check{\text{S}}}$$

At this level of Linearization process, the structural adjacency relationship between heads (M-Words) and phrases are represented in *-statements. Notice that these statements include elements that are not pronounced, either because they are traces/copies, or because they are internally complex—i.e. phrases that still contain *-statements. In addition to traces/copies, we have the E-specified noun Root that is not pronounced. The non-pronunciation of the noun Root is due to its E-feature blocking the Vocabulary Insertion from inserting the Root’s phonological exponent (cf. §3.3.1.1). I assume, though, that the PL_{DEF} morpheme adjoined to the E-specified noun Root is pronounced here because (i) it must be spelled out under a language-specific requirement (cf. (61) above); and (ii) it meets its context of insertion—as was formulated in (49).

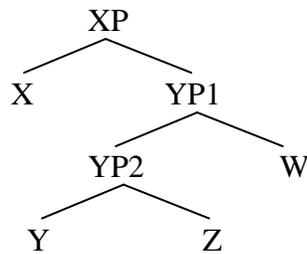
In our *-statements in (73), there are two overt M-Words— $((\sqrt{\text{KAF}\check{\text{S}}^0}_{[E]} \oplus n^0) \oplus -h\hat{a})$ and $(\sqrt{\text{P}\check{\text{A}}\text{SNEBOLAND}^0} \oplus a_{[F]}^0)$ —in two separate statements, which need to be further linearized. We need other steps in order to determine how they are linearized with respect to each other.

(STEP 2) CONCATENATION: The last *-statement in (73) tells us that the M-Word $((\sqrt{\text{KAF}\check{\text{S}}^0}_{[E]} \oplus n^0) \oplus -h\hat{a})$ is left-adjacent to an internally complex element $\sqrt{\text{KAF}\check{\text{S}}}$. This statement however does not give us information as to what M-Word inside this internally

complex element $((\sqrt{\text{KAF}\hat{S}}^0_{[E]}\oplus n^0)\oplus-h\hat{a})$ is adjacent to. This information is provided in the second stage of Linearization process by an operation called Concatenation, which generates statements of left-adjacency between M-Words only. Pak (2008: 217-218) proposes that there are two separate steps involved in Concatenation: *Head-left Concatenation* and *Phrase-left Concatenation*.

- “*Head-left Concatenation* begins with a *-statement whose left-hand member is an overt M-word X (rather than a phrasal category or null head), and searches within its right-hand member until it identifies the overt M-word Y that appears immediately to the right of X.

Hypothetical structure:



$$\{(X * YP1), (YP2 * W), (Y * Z)\} \rightarrow X \frown Y, Y \frown Z$$

In other words, Head-left Concatenation identifies pairs of M-words X, Y where (i) X is left-adjacent to Y, and (ii) X c-commands Y.

- *Phrase-left Concatenation* begins with a *-statement whose left member is a phrasal category; it then searches within the left-hand member until it identifies the right-most M-word, and within the right-hand member until it identifies the left-most M-word. In other words, it identifies pairs of M-words X, Y where (i) X is left-adjacent to Y, and (ii) X *does not* c-command Y.

$$\{(X * YP1), (YP2 * W), (Y * Z)\} \rightarrow Z \frown W$$

The *-statements in (73) are subject to Head-left Concatenation in that they contain statements that begin with overt M-Words. Pak (2008) suggests an algorithm for deriving the Concatenation statements within a PF cycle that I follow here. This algorithm starts from topmost node in the structure.

(74) *Concatenation* \frown

a. Begin with $\text{LIN}[nP_{\text{def}}] \rightarrow ((\sqrt{\text{KAFŠ}}^0_{[\text{E}]} \oplus n^0) \oplus -h\hat{a}) * \sqrt{\text{KAFŠ}}$

$\sqrt{\text{KAFŠ}}$ is internally complex. Locate *-statement for $\sqrt{\text{KAFŠ}}$.

b. $\text{LIN}[\sqrt{\text{KAFŠ}}] \rightarrow \langle \sqrt{\text{KAFŠ}}^0 \rangle * a$

$\langle \sqrt{\text{KAFŠ}}^0 \rangle$ is a silent element; a is internally complex. Locate *-statement for a .

c. $\text{LIN}[a_{[\text{F}}]] \rightarrow (\sqrt{\text{PĂŠNEBOLAND}}^0 \oplus a^0_{[\text{F}]}) * \langle \sqrt{\text{PĂŠNEBOLAND}}^0 \rangle$

$(\sqrt{\text{PĂŠNEBOLAND}}^0 \oplus a^0_{[\text{F}]})$ is an overt M-Word. Concatenate $((\sqrt{\text{KAFŠ}}^0_{[\text{E}]} \oplus n) \oplus -h\hat{a})$ and $(\sqrt{\text{PĂŠNEBOLAND}}^0 \oplus a^0_{[\text{F}]})$.

d. $((\sqrt{\text{KAFŠ}}^0_{[\text{E}]} \oplus n^0) \oplus -h\hat{a}) \frown (\sqrt{\text{PĂŠNEBOLAND}}^0 \oplus a^0_{[\text{F}]})$

Now that linear precedence and adjacency has been established between the overt M-Words in the statements in (74), an LD rule in terms of \frown can be formulated to the effect that the M-Word representing the phonologically null noun and an overt PL/PL_{DEF} morpheme (-h \hat{a}) adjoins the string-adjacent M-Word representing the attributive adjectival modifier.

(75) *Persian PL/PL_{DEF} Local Dislocation in the ENC (Preliminary)*

$$n((\sqrt{\text{ROOT}}^0_{[E]} \oplus n^0) \oplus -h\hat{a}) \sim a_{[F]}(\sqrt{\text{ROOT}}^0 \oplus a^0_{[F]}) \rightarrow [[a_{[F]}] n]$$

where $n(\sqrt{\text{ROOT}}^0_{[E]})$ is unpronounced and $a_{[F]}(\sqrt{\text{ROOT}}^0)$ is phonologically realized

As discussed in Chapter 1, LD is formulated in terms of Concatenation statements (cf. Embick 2003). Since LD is an operation which involves adjunction under adjacency, it deletes the input Concatenation statement and introduces the hierarchical information that the two M-Words involved in the Concatenation statement are adjoined to each other. For the derivation of *pâšneboland-â* (high-heeled-PL_{DEF}) ‘the high-heeled ones’ in (62), the Concatenation statement in (74)d meets the environment for the application of (75): the E-specified noun Root is silent and the M-Word *n* is string-adjacent to a phonologically realized M-Word *a*. In order for *-hâ* to attach to the phonologically realized *a*, we need to apply the LD rule in (75) and adjoin $((\sqrt{\text{KAFŠ}}^0_{[E]} \oplus n^0) \oplus -h\hat{a})$ to $(\sqrt{\text{PÂŠNEBOLAND}}^0 \oplus a^0_{[F]})$.

$$(76) \quad n((\sqrt{\text{KAFŠ}}^0_{[E]} \oplus n^0) \oplus -h\hat{a}) \sim a_{[F]}(\sqrt{\text{PÂŠNEBOLAND}}^0 \oplus a^0_{[F]}) \rightarrow [[a_{[F]}] n]$$

Due to the hierarchical nature of adjunction, the application of this LD rule results in a syntactic structure which needs the linearization process to apply again and calculate its * and \oplus -statements. The recalculation of the linear statements of this structure shows two effects: (i) the order of the two M-Words is reversed; and (ii) the moving M-Word, i.e. $n((\sqrt{\text{KAFŠ}}^0_{[E]} \oplus n^0) \oplus -h\hat{a})$, becomes a Subword. These effects are illustrated in (77):

(77) *Linearization of the output of PL/PL_{DEF} Local Dislocation*

a. Structure: $[[a_{[F]}] n]$

b. PF

i. Adjacency: $((a_{[F]}) * n)$

ii. Concatenation: $((\sqrt{P\hat{A}\check{S}NEBOLAND} \oplus a_{[F]}^0) \oplus (\sqrt{KAF\check{S}_{[E]}} \oplus n^0 \oplus -h\hat{a}))$

So far we have seen that an LD rule takes two M-Words and creates a single M-Word. The moving M-Word n housing $-h\hat{a}$ has become a Subword inside the derived structure linearized in (77). Following Pak (2008: 233) who argues that morphemes with null exponents are deleted after the internal content of an M-Word has undergone Vocabulary Insertion and Linearization, I suggest that the resulting M-Word from the application of the LD rule in (76) does not contain phonologically null elements after Vocabulary Insertion and Linearization have taken place. Thus, the M-Word can be represented as follows:

(78) $p\hat{a}\check{s}neboland \oplus -h\hat{a}$

At the final stage of the derivation, the output of LD operation is input to what Embick (2003, 2008) calls *Chaining*²⁰. Chaining is the final PF operation in the process of linearizing the syntactic structure whereby concatenated elements are *chained* into a

²⁰ In Pak's (2008) more articulated model of PF derivation that we will adopt in Chapter 5 for the analysis of Ezafe insertion in Persian, Chaining is introduced as the final stage of the derivation of a PF *cycle*. The Chaining of all the spell-out domains of phases (cycles) involved in a derivation takes place at what Pak dubs *Late-Linearization*.

linear representation. At this stage, all information in a linear sequence is represented. In the case of (78), for instance, I propose that the phonologically conditioned allomorphy shown by the Persian PL/PL_{DEF} morpheme *-hâ* is handled at Chaining. To elaborate, *-hâ* is pronounced fully when it follows a vowel-final word; it appears as *-â* when preceded by a consonant-final word. This allomorphy can be formalized as follows:

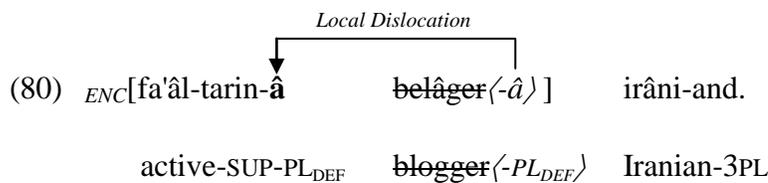
$$(79) \text{ PL/PL}_{\text{DEF}} \leftrightarrow -h\hat{a} / \text{V} \oplus \text{ ______}$$

$$\text{ PL/PL}_{\text{DEF}} \leftrightarrow -\hat{a} / \text{C} \oplus \text{ ______}$$

Thus, the context of PL/PL_{DEF} allomorphy in (79) determines that at Chaining, *pâšneboland*⊕*-hâ* should be represented as *pâšneboland-â*.

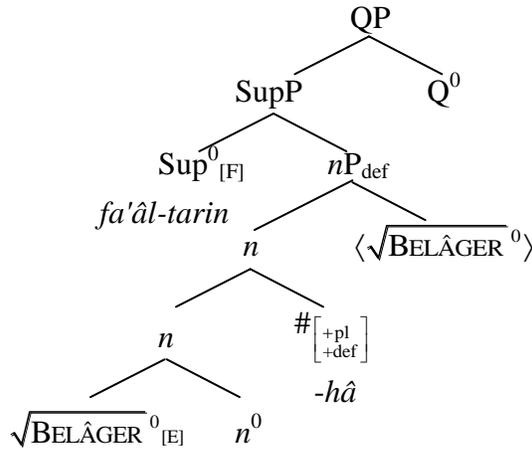
4.4.1.2 *-hâ* mobility across phase boundary

I now consider an example of the ENC in which the empty noun is licensed by a prenominal modifier outside the *nP* phase. This occurs when there is no attributive adjectival modifier in the *nP* cycle of the ENC derivation. In such circumstances, the stranded *-hâ* can simply lean over to its adjacent prenominal modifier in which case the adjunction of *-hâ* does not involve reordering.



I assume that the ENC *fa'âl-tarin-â belâger<-â>* ‘the most active ones’ has the following syntactic configuration:

(81) *Syntactic Structure*



As we discussed in Chapter 3, when there is no F-specified attributive adjectival modifier inside the *nP* phase to license the empty noun, the spell-out is delayed until the whole noun phrase is constructed and the Licensing requirements of the empty noun are met. Given this assumption, I suggest that the input to PF computation for the ENC in (80) is the whole structure in (81). Based on this structure, the empty noun is licensed by an inherently F-specified superlative adjective that c-commands the E-specified Root inside the *nP* phase. At this point, the derivation is ready to be sent off to PF and LF for interpretation.

Given the rules and procedures described earlier, the PF derivation of the ENC *fa'âl-tarin-â belâger<-â>* ‘the most active ones’ starts out with applying Vocabulary Insertion and Linearization procedures to the hierarchical structure in (81). Assuming that

the whole QP constitutes the PF cycle of derivation, the Linearization procedures apply and generate * and \frown - statements in the following manner:

(82) PF

$$i. * : \text{LIN}[nP_{\text{def}}] \rightarrow ((\sqrt{\text{BELÂGER}}_{[\text{E}]}^0 \oplus n^0) \oplus -h\hat{a}) * \langle \sqrt{\text{BELÂGER}}^0 \rangle$$

$$\text{LIN}[\text{SupP}] \rightarrow \text{Sup}_{[\text{F}]}^0[\text{fa}'\hat{a}l\text{-tarin}] * nP_{\text{def}}$$

$$\text{LIN}[\text{QP}] \rightarrow \text{SupP} * Q^0$$

ii. \frown : *Head-left Concatenation*

$$a. \text{Begin with } \text{LIN}[\text{SupP}] \rightarrow \text{Sup}_{[\text{F}]}^0[\text{fa}'\hat{a}l\text{-tarin}] * nP_{\text{def}}$$

nP_{def} is internally complex; Locate *-statement for nP_{def} .

$$b. \text{LIN}[nP_{\text{def}}] \rightarrow ((\sqrt{\text{BELÂGER}}_{[\text{E}]}^0 \oplus n^0) \oplus -h\hat{a}) * \langle \sqrt{\text{BELÂGER}}^0 \rangle$$

$((\sqrt{\text{BELÂGER}}_{[\text{E}]}^0 \oplus n^0) \oplus -h\hat{a})$ is an overt M-Word. Concatenate

$\text{Sup}_{[\text{F}]}^0[\text{fa}'\hat{a}l\text{-tarin}]$ and $((\sqrt{\text{BELÂGER}}_{[\text{E}]}^0 \oplus n^0) \oplus -h\hat{a})$.

$$c. \text{Sup}_{[\text{F}]}^0[\text{fa}'\hat{a}l\text{-tarin}] \frown ((\sqrt{\text{BELÂGER}}_{[\text{E}]}^0 \oplus n^0) \oplus -h\hat{a})$$

The Concatenation statement $\text{Sup}_{[\text{F}]}^0[\text{fa}'\hat{a}l\text{-tarin}] \frown ((\sqrt{\text{BELÂGER}}_{[\text{E}]}^0 \oplus n^0) \oplus -h\hat{a})$ in (82) is generated via Head-left Concatenation because the overt M-Word $\text{Sup}_{[\text{F}]}^0[\text{fa}'\hat{a}l\text{-tarin}]$ is left-adjacent to the M-Word $n((\sqrt{\text{BELÂGER}}_{[\text{E}]}^0 \oplus n^0) \oplus -h\hat{a})$ and also c-commands it. The morphophonological requirement that $-h\hat{a}$ must be spelled out on a phonologically realized head (cf. (61)) triggers the M-Word $((\sqrt{\text{BELÂGER}}_{[\text{E}]}^0 \oplus n^0) \oplus -h\hat{a})$ —whose noun Root is devoid of phonological exponent—to adjoin the M-Word $\text{Sup}_{[\text{F}]}^0[\text{fa}'\hat{a}l\text{-tarin}]$. The

environment, however, does not match the one defined for the Persian PL/PL_{DEF} Local Dislocation rule in (75). That is, the M-Word *n* is not left-adjacent to the modifier, but rather the modifier is left-adjacent to the M-Word *n*. Thus, the affixation of the stranded *-hâ* to the prenominal modifier does not result in reordering. In order to accommodate this string-vacuous Local Dislocation, I propose a revised version of the LD rule in (83):

(83) *Persian PL/PL_{DEF} Local dislocation in the ENC (Final Version)*

$$(i) \quad n((\sqrt{\text{ROOT}}^0_{[E]} \oplus n^0) \oplus -h\hat{a}) \frown a_{[F]}(\sqrt{\text{ROOT}}^0 \oplus a^0_{[F]}) \rightarrow [[a_{[F]}] n]$$

$$(ii) \quad \text{PreN}^0_{[F]} \frown n((\sqrt{\text{ROOT}}^0_{[E]} \oplus n^0) \oplus -h\hat{a}) \rightarrow [[\text{PreN}^0_{[F]}] n]$$

Where (i) $n(\sqrt{\text{ROOT}}^0_{[E]})$ is unpronounced and $a_{[F]}(\sqrt{\text{ROOT}}^0)/\text{PreN}^0_{[F]}$ are phonologically realized; (ii) $\text{PreN}^0_{[F]}$ (inherently F-specified prenominal modifiers) include demonstratives, quantifiers, superlatives, superordinals, and interrogatives.

(83)i&ii illustrate two different Concatenation statements to which LD applies in order for the stranded *-hâ* to find a phonologically realized host to attach to. In (83)i, the input Concatenation statement to the LD rule is generated in the course of linearizing the *nP* phase, where the empty noun is concatenated with the attributive adjectival modifier. The output of the application of the LD rule to this \frown -statement results in the reversal of the order of the two M-Words involved. In addition, the moving M-Word becomes a Subword. The input to the LD rule in (83)ii, however, is a Concatenation statement that is generated in the last PF cycle of deriving the ENC. In this statement, the M-Word representing one of the prenominal modifiers mentioned above is concatenated with the

M-Word n . The application of the LD rule to this \frown -statement does not arise the reversal of the linear order, although the moving M-Word becomes a Subword.

Having completed the formulation of the PL/PL_{DEF} LD rule, we now see that the Concatenation statement $\text{Sup}_{[F]}^0[\text{fa}'\hat{\text{a}}\text{l-tarin}] \frown ((\sqrt{\text{BEL}\hat{\text{A}}\text{GER}}_{[E]}^0 \oplus n^0) \oplus -h\hat{\text{a}})$ meets the condition for the application of the string-vacuous LD, as formulated in (83)ii. With reference to the PF derivation in (84), we see the M-Word with the stranded $-h\hat{\text{a}}$ adjoins to the M-Word for the prenominal superlative adjective $\text{fa}'\hat{\text{a}}\text{l-tarin}$ ‘the most active’. As established earlier, the resulting adjunction structure is further subjected to recalculation of its Linearization statements and application of Chaining.

(84) PF

a. *String-vacuous LD*

Input: $\text{Sup}_{[F]}^0[\text{fa}'\hat{\text{a}}\text{l-tarin}] \frown ((\sqrt{\text{BEL}\hat{\text{A}}\text{GER}}_{[E]}^0 \oplus n^0) \oplus -h\hat{\text{a}})$

Output: $[[\text{Sup}_{[F]}^0] n]$

b. *Recalculation of Linearization statements*

i. $*$: $((\text{Sup}_{[F]}^0) * n)$

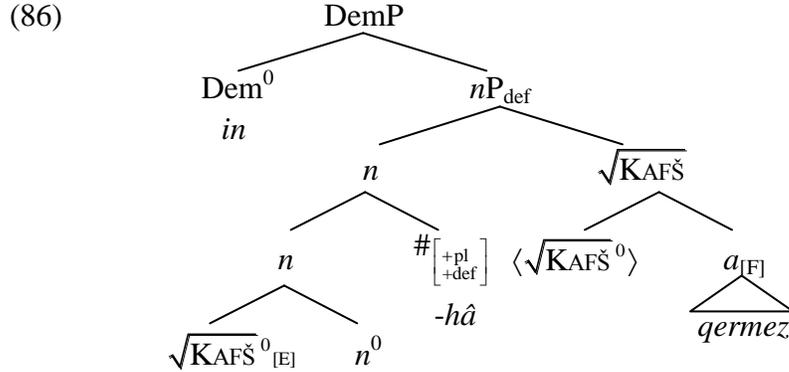
ii. \frown : $(\text{Sup}_{[F]}^0[\text{fa}'\hat{\text{a}}\text{l-tarin}] \oplus ((\sqrt{\text{BEL}\hat{\text{A}}\text{GER}}_{[E]}^0 \oplus n^0) \oplus -h\hat{\text{a}}))$

iii. Chaining: $\text{fa}'\hat{\text{a}}\text{l-tarin-}\hat{\text{a}}$

So far in this section we have accounted for the possibility of prenominal modifiers hosting the stranded $-h\hat{\text{a}}$ when the empty noun is not licensed by an adjectival head modifier inside the nP phase. The question that arises here is: Does the LD rule in (83)i apply to the Concatenation statement $n \frown a$ when the M-Word a is the left-most M-Word

higher structure is computed. Given the model of PF derivation adopted in this dissertation, I present my analysis of (85) below.

I assume that (86) represents the syntactic structure of the ENC in *kafš qermez-â* ‘these red ones’.



In the first stage of the derivation, the PF receives the *nP* structure and processes it in the manner illustrated in (87):

(87) a. Structure: $nP_{def}[n[\sqrt{KAFŠ}^0_{[E]}n^0]_{\sqrt{KAFŠ}}[\langle\sqrt{KAFŠ}^0\rangle a_{[F]}]]$

b. PF

i. *: $LIN[\sqrt{KAFŠ}] \rightarrow \langle\sqrt{KAFŠ}^0\rangle * (\sqrt{QERMEZ}^0_{[F]} \oplus a^0_{[F]})$

$LIN[nP_{def}] \rightarrow ((\sqrt{KAFŠ}^0_{[E]} \oplus n^0) \oplus -hâ) * \sqrt{KAFŠ}$

ii. \frown : *Head-left Concatenation*

a. Begin with $LIN[nP_{def}] \rightarrow ((\sqrt{KAFŠ}^0_{[E]} \oplus n^0) \oplus -hâ) * \sqrt{KAFŠ}$

$\sqrt{KAFŠ}$ is internally complex. Locate *-statement for $\sqrt{}^{0max}$

b. $LIN[\sqrt{KAFŠ}] \rightarrow \langle\sqrt{KAFŠ}^0\rangle * (\sqrt{QERMEZ}^0_{[F]} \oplus a^0)$

$\langle \sqrt{\text{KAFŠ}}^0 \rangle$ is a silent element. $(\sqrt{\text{QERMEZ}}^0_{[\text{F}]} \oplus a^0)$ is an overt M-

Word. Concatenate $((\sqrt{\text{KAFŠ}}^0_{[\text{E}]} \oplus n^0) \oplus -h\hat{a})$ and $(\sqrt{\text{QERMEZ}}^0_{[\text{F}]} \oplus a^0)$.

$$\text{c. } ((\sqrt{\text{KAFŠ}}^0_{[\text{E}]} \oplus n^0) \oplus -h\hat{a}) \frown (\sqrt{\text{QERMEZ}}^0_{[\text{F}]} \oplus a^0)$$

iv. $PL_{DEF} LD$

$$\text{Input: } n((\sqrt{\text{KAFŠ}}^0_{[\text{E}]} \oplus n^0) \oplus -h\hat{a}) \frown a_{[\text{F}]}(\sqrt{\text{QERMEZ}}^0_{[\text{F}]} \oplus a^0)$$

$$\text{Output: } [[a_{[\text{F}]}] n]$$

v. *Recalculation of Linearization statements*

$$\text{i. } * : ((a_{[\text{F}]} * n)$$

$$\text{ii. } \frown : (((\sqrt{\text{KAFŠ}}^0_{[\text{E}]} \oplus n^0) \oplus -h\hat{a}) \oplus (\sqrt{\text{QERMEZ}}^0_{[\text{F}]} \oplus a^0_{[\text{F}]}))$$

$$\text{iii. Chaining: } qermez-\hat{a}$$

With the above PF structure for the nP_{def} cycle, the derivation continues the creation of the ENC in (86) by PF processing the highest nominal projection, namely DemP. Given the application of LD in the previous cycle, the M-Word Dem^0 (*in* ‘this’) cannot undergo Local Dislocation with the M-Word n because the structural description for the application of LD is not met. In fact, the complex n ceases to exist as an M-Word at this stage of the derivation because the LD rule has turned it into a Subword after adjoining it to the M-Word $a_{[\text{F}]}$ in the nP_{def} cycle. This explains why **in-â qermez* is not a possible string. Given that there is no overt M-Word but $\text{Dem}^0[in]$ within the DemP cycle, the M-Word Dem^0 is simply concatenated with the spell-out domain of the nP_{def} (notationally shown as $\text{PF}_{nP_{def}}$): **Dem[in] \frown PF $_{nP}$ [qermez-â]**. This statement is eventually subject to Chaining, which clusters these two M-Words as *in qermezâ* ‘these red ones’.

I suggest that Persian does not allow the plural marker to spell out the feature [+def] when the noun is specified with the E-feature. As discussed in Chapter 3, the semantics of the E-feature in the ENC involves identifying the empty noun from its previous mention in the discourse. The E-feature specified on the nominal Root therefore characterizes the empty noun as being semantically (or pragmatically) definite²¹. If this is correct, then the realization of the plural marker as the spell-out of the feature [+def] would be rendered redundant. It is therefore reasonable to conclude that *-hâ* can appear in the ENC just to spell out [+pl]; the feature [+def] can be realized only when there is an overt noun. Given this, I propose the following modification to our set of constraints on plural marking in Persian (the modified part is in italics):

- (89) (i) [+pl] must be phonologically realized.
(ii) [+def] must be spelled out.
(iii) Avoid redundancy in number/*definiteness* marking.

Thus, I propose that Persian does not allow redundant definite marking as it disallows redundant number marking. Redundancy avoidance therefore prevents the insertion of *-hâ* as a marker of definiteness in the ENCs involving numerals+classifiers. As such, the question of why it does not undergo string-vacuous LD would be irrelevant.

²¹ See Lyons (1999) for a distinction between semantic/pragmatic definiteness vs. definiteness as a grammatical concept.

4.5. Summary

This chapter has offered a Distributed-Morphology based analysis of plural marking in the ENC in Persian. I have presented a novel analysis of plural marking in Persian arguing that number and definiteness are syncretic. Building on Wiltschko's (2008) diagnostics on the categorial status of plural markers, I have suggested that the plural marker in Persian is base-generated at syntax as an adjunct to the nominalized Root ($\sqrt{\text{ROOT}}^0-n^0$). The displacement of the plural marker in the ENC, however, cannot happen at syntax because *-hâ* is category sensitive and does not select for categories other than nouns. I have suggested that this phenomenon can receive a plausible explanation if we assume that *-hâ* mobility takes place postsyntactically. On this view, the displacement of the plural marker is assumed to be motivated by the PF-imposed language-specific requirement for plural marking to appear on phonologically realized heads. As such, this movement operation is Vocabulary sensitive and involves linear precedence and adjacency. Adopting Embick & Noyer's (2001, 2006) and Embick's (2003) approach to Linearization procedures and Local Dislocation, I have presented a detailed analysis of *-hâ* mobility in Persian. I have argued that there is a correlation between being a licenser of the empty noun and being the host of the stranded *-hâ* in the ENC. In other words, the host of the stranded *-hâ* must be identical with the licenser of the empty noun. One important implication of this analysis is to provide further evidence in support of the proposal that postnominal modifiers in the Ezafe domain belong to two domains of modification: inner (*nP*) domain for head modifiers and outer domain for phrasal ones. I will pursue further argumentation in the next chapter to establish that such an assumption is necessary for an adequate account of nominal modification in Persian noun phrases.

Chapter 5

The Ezafe Construction Revisited

5.0. Introduction

In this chapter I will incorporate the consequences of the analyses in Chapters 3 and 4 into a revised theory of nominal modification in the Ezafe construction. Nominal modifiers that follow the head noun constitute the Ezafe domain, the domain in which modifiers are connected to the head noun via the Ezafe vowel *-e*. There are two different views in the literature about the category of nominal modifiers in the Ezafe construction in Persian. One view maintains that modifiers in the Ezafe domain are all heads (e.g. Ghomeshi 1997), the other claims that they are all phrases (e.g. Samvelian 2006). Reviewing briefly the relevant literature on the Ezafe construction, I will propose a hybrid approach assuming that there are both head and phrasal modifiers in the Ezafe domain that are distinguished in terms of their structural positions. Head modifiers are generated as head-adjoined to the nominal Root and phrasal modifiers as specifiers of ModP (i.e. *aP* and *PP*) and PossP (i.e. Possessor).

In addition to conditions on licensing the empty noun and *-hâ* mobility, discussed in Chapters 3 and 4, further support for assuming a structural and categorial split between modifiers in the Ezafe domain comes from compounding and syntactic compounding. I will discuss these phenomena in sections 5.1.3.1 & 5.1.3.2.

The issue that I will address in the final section of this chapter concerns the longstanding question of how Ezafe appears in Persian noun phrases. There are three

views on the function of Ezafe in the literature: (i) Ezafe is a Case marker (Samiian 1983, Larson & Yamakido 2005); (ii) Ezafe is inserted at PF to identify constituenthood (Ghomeshi 1997); and (iii) Ezafe is a phrasal affix (Samvelian 2006). Building on the PF derivation of *-hâ* mobility proposed in chapter 4 (cf. 23), I argue, along the lines of Ghomeshi (1997), that Ezafe Insertion is a PF phenomenon that occurs at the final stage of derivation, after all other PF operations have applied. I propose that Ezafe Insertion occurs at the interface of phonology and syntax where the Ezafe vowel *-e* is encliticized to a head that bears the feature [+N] to link it to its following modifier(s)/complement(s).

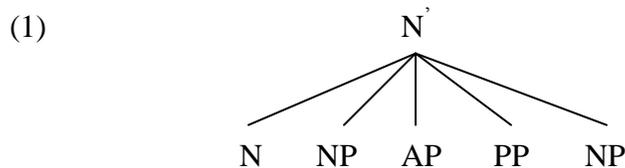
5.1. Syntactic properties of nominal modifiers in the Ezafe domain

5.1.1 Background literature

Subsequent to Samiian's (1983) exhaustive account of the Ezafe construction within X-bar theory, Ghomeshi (1997) and Samvelian (2006) have contributed to the characterization of the syntactic properties of the Ezafe domain with their different approaches. While Samiian and Ghomeshi assume that the modifiers in the Ezafe domain behave as heads, Samvelian considers these modifiers as phrases. In this section, I will specifically concentrate on those aspects of the accounts presented by Ghomeshi and Samvelian that are relevant to the main concern of the current study.

5.1.1.1 Samiian (1983) and Ghomeshi (1997)

Ghomeshi (1997) primarily builds her analysis on Samiian's (1983) observations and analysis. In Samiian's analysis, the nominal modifiers occurring in the Ezafe domain are all *theoretically* assumed to be phrasal and below N', as shown in (1):



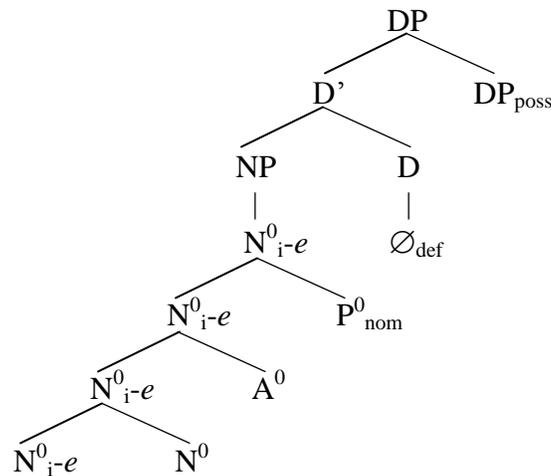
In the examples that Samiian provides, however, all nominal modifiers, except the possessive NP, are shown to occur simply as bare heads, with no complements. To accommodate this observation into her analysis, Samiian formulates a filter barring the modifiers between the head noun and the possessive NP from taking complements. She rejects the idea that these modifiers are heads rather than phrases for certain theoretical reasons. The most important reason she offers is that heads, according to the classic X-bar principles (Jackendoff 1977), are not allowed to take non-maximal projections as complements.

Ghomeshi (1997) proposes a modification to the classic X-bar theory in order to accommodate the possibility of adjoining heads to one another. She postulates that projection in the syntax is not obligatory and that “the phrasality of a constituent in the syntax depends not only on whether a head projects, but on whether that head itself is selected by a projecting element” (p. 372). She further suggests that “while certain X⁰ categories may be inherently non-projecting in syntax, they may still appear as XPs, provided they are selected by a projecting head” (p. 733-4). For her, an inherently non-

projecting head is one that will never appear with filled specifier and complement positions.

The fact that the head noun constitutes an X^0 -adjoined structure with its modifiers in the Ezafe domain leads Ghomeshi to conclude that nouns in Persian are inherently non-projecting. This entails that (a) nouns never appear with filled specifier and complements; (b) they cannot be dominated by NP unless selected by a projecting element; and (c) the Ezafe vowel never attaches to phrasal categories. Based on the above assumptions, Ghomeshi suggests the following syntactic configuration for the structure of the Ezafe construction in Persian DPs:

(2) *The Ezafe construction in Persian DP* (Ghomeshi 1997: 780.90)



By assuming the above syntactic configuration for the Ezafe construction, Ghomeshi captures Samiiian's restrictions on elements that appear between the head noun and possessive DP without resorting to a filter.

5.1.1.2 Samvelian (2006)

Samvelian's work is an attempt to present a new analysis of the Ezafe construction in Persian that can be extended, with slight modifications, to other members of the West Iranian languages. She builds her analysis on the problems of the previous accounts by Samiian and Ghomeshi. Most importantly, she questions the validity of the constraints on the nominal attributive modifiers primarily put forward by Samiian and subsequently adopted by Ghomeshi. She presents several examples to show that those restrictions do not hold and that the generalizations made on the basis of the data are thus ill-founded. The following examples particularly support Samvelian's challenge of Samiian's view that modifiers in the Ezafe domain cannot take complements and hence be phrasal⁵⁹.

(3) *Phrasal modifiers in the Ezafe domain*

a. *Adjective Phrases*

- (i) ... N''[N'[N[čašm-ân-e A''[A[negarân-e N''[forurixtan-e divâr-e
eye-PL-EZ worried-EZ crumble-EZ wall-EZ
xâne-aš]]]]]...

house-CL.3SG

'...his eyes worried about the crumbling of the wall of his house...'

[Samvelian 2006: 12.31]

⁵⁹ The labeled bracketing is mine. Also, in glossing Samvelian's examples, I use my conventions to prevent any confusion.

(ii) ... N''[N'[M[niru-hâ-ye N''[qazzâq-e A''[mostaqar dar qazvin]]]]]...
 force-PL-EZ Kozak-EZ deployed in Ghazvin

‘...Kozak forces deployed in Ghazvin...’

[Samvelian 2006: 12.34]

b. *Prepositional Phrases*

N''[N'[M[sekke-hâ-ye P''[tu-ye jib-e šalvâr-aš]]] ...
 coin-PL-EZ in-EZ pocket-EZ trousers-CL.3SG

‘the coins in his trousers’ pocket...’

[Samvelian 2006: 13.39]

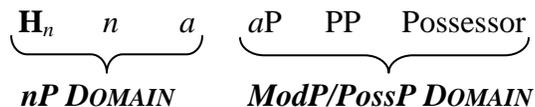
By establishing that phrasal modifiers can appear in the Ezafe domain, Samvelian shows that Ghomeshi’s claim that nouns are inherently non-projecting and as a result do not license phrasal modifiers cannot be maintained.

5.1.2 Proposal: mixed approach

So far we have seen that the nominal modifiers preceding the possessive modifier in the Persian noun phrase are assumed in the literature to be either heads (Samiian 1983, Ghomeshi 1997) *or* phrases (Samvelian 2006). Building on the observations and discussions in Chapters 3 and 4, I propose a mixed approach according to which nominal modifiers in the Ezafe domain must be distinguished in terms of categorial status and structural position.

With respect to categorial status, I assume that there are both head *and* phrasal modifiers in the Ezafe domain. In consistence with the empirical facts reported in Samiian (1983) and Samvelian (2006), I consider the attributive noun and adjective as heads because they never take complements. Other modifiers that do take complements are therefore considered as phrases. With respect to structural position, I follow Ghomeshi (1997) in assuming that head modifiers are adjoined to the head noun. Specifically, I propose that *nP* defines a syntactic domain in which only head modifiers can be generated. This restriction suggests that phrasal modifiers must be generated outside this domain. Given our assumption that DP does not exist in Persian (see Ch. 3: ft. 1), phrasal modifiers are generated within ModPs/PossPs. As noted in Chapter 1, I propose ModP as an arbitrary projection that hosts phrasal modifiers other than possessive noun phrases. Possessive noun phrases are generated in the specifier of PossP. This structural distinction in fact defines two domains of nominal modification in the Ezafe construction: *nP*_(def) for head modifiers and ModP/PossP for phrasal ones. Notice in (4) that head modifiers always precede the phrasal ones, while both maintain their strict inner ordering.

(4) *Domains of modification*



The idea of assuming two distinct domains of modification inside the noun phrase is not without precedent in the literature. Larson and Takahashi (2007) posit the following outer modifier/inner modifier contrasts in English:

(5)	<p style="text-align: center;"><u>OUTER</u></p> <p>visible</p> <p style="text-align: center;"><i>TEMPORARY PROP</i>⁶⁰</p> <p>Thursday</p> <p style="text-align: center;"><i>DEICTIC</i></p> <p>beautiful</p> <p style="text-align: center;"><i>INTERSECTIVE</i></p>	<p style="text-align: center;"><u>INNER</u></p> <p>visible stars</p> <p style="text-align: center;"><i>ENDURING PROP</i></p> <p>Thursday lecture</p> <p style="text-align: center;"><i>GENERIC</i></p> <p>beautiful dancer</p> <p style="text-align: center;"><i>NON-INTERSECTIVE</i></p>
-----	--	--

They propose that NP and DP constitute two distinct domains of modification and that the inner vs. outer contrasts in the examples above correspond to NP- vs. DP-modification, respectively (p. 13):

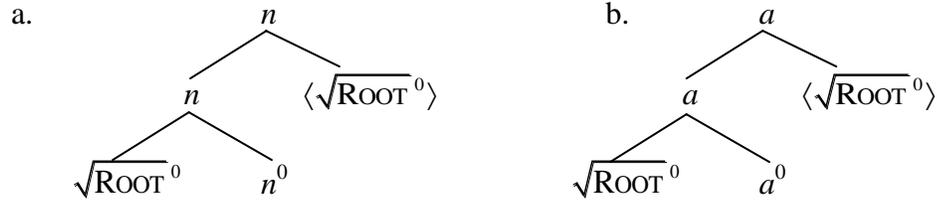
$$(6) \quad [_{DP} \quad D \quad \beta \quad [_{NP} \quad \alpha \quad N \quad] \quad \beta \quad] \quad (\alpha = \text{NP-modifier}; \beta = \text{DP-modifier})$$

I suggest that the Larson and Takahashi’s proposal nicely fits our observation of the behavior of nominal modifiers in the Ezafe domain. In their account of English nominal modifiers, the semantic properties determine the structural positions of nominal modifiers. In my account, however, the syntactic properties determine both the categorial

⁶⁰ Prop is the short form of ‘proposition’.

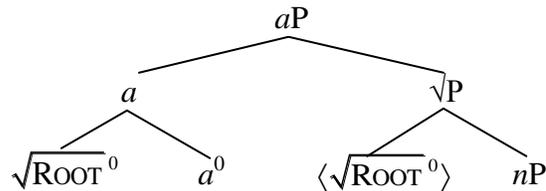
In (7), *n* and *a* in the inner domain of modification represent the head modifiers. I consider the attributive noun and adjective as heads because their derivation involves Roots that never select for complements.

(8) *Internal structure of head modifiers*



Contrary to the head modifier *a*, the phrasal modifier *aP* in the outer domain involves a Root that selects for a complement.

(9) *Internal structure of the phrasal modifier aP*



The analyses of the ENC and *-hâ* mobility in Chapters 3 and 4 imply that the process of licensing the empty noun and the application of the PF operations involved in displacing the stranded *-hâ* are not only sensitive to the categorial status of the nominal modifiers but also to their position in the structure of the noun phrase. As suggested in Chapter 3, the sensitivity of the empty noun licensing process and *-hâ* mobility to categorial and structural distinction between these modifiers is captured by assuming a

phase-based derivation. In an incremental cyclic derivation of the Persian ENC, the head modifiers are computed in the *nP* phase, where they license the empty noun and host the stranded *-hâ*. The phrasal modifiers which cannot be subject to these grammatical operations are outside this domain and thus computed in the next phase of derivation.

Apart from licensing the empty noun and *-hâ* mobility, there are two processes that can provide further evidence in favor of assuming a two domain theory of nominal modification⁶²: compounding and syntactic compounding. I will discuss these two operations in the ensuing section.

5.1.3 Distinguishing nominal modifier types: further evidence

5.1.3.1 Compounding

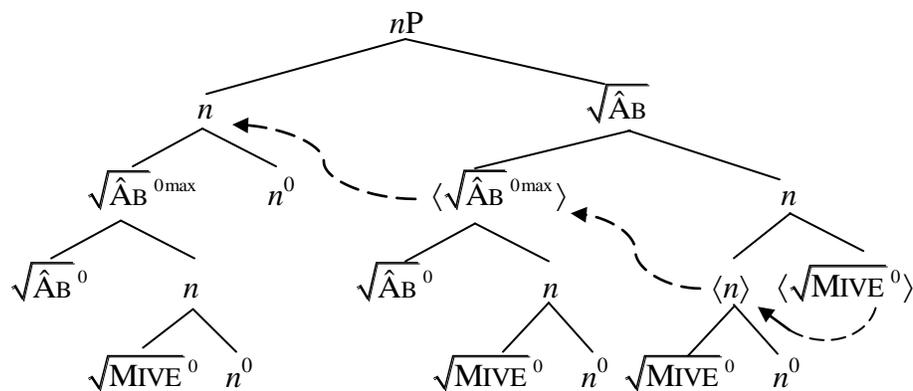
The semantic relationship between the head noun and its attributive modifier in the *Ezafe* domain is reflected in some endocentric compounds in Persian. What is of particular interest for our discussion is that only head modifiers can undergo compounding and not phrasal ones (\Leftrightarrow = correspondence).

- (10) *Ezafe Construction* *Compounding*
- a. $n_P[n[\hat{a}b-e \quad n[mive]]] \quad \Leftrightarrow \quad n[\hat{a}b+mive] \quad \text{'fruit juice'}$
- water-EZ fruit

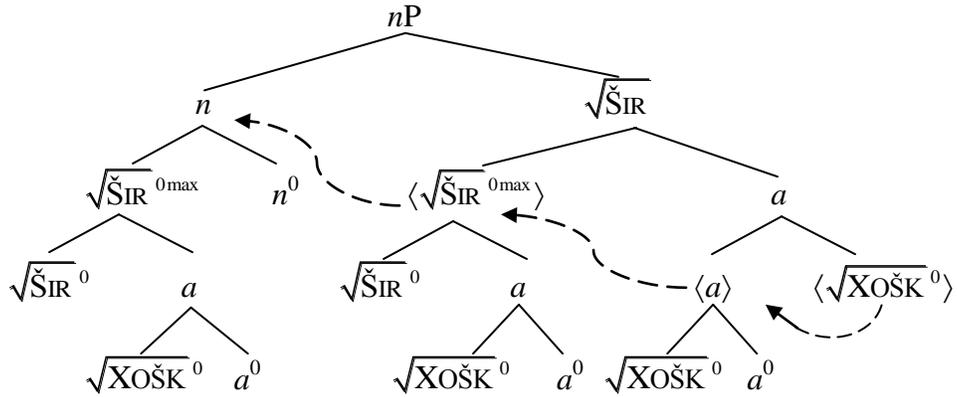
⁶² The structural difference between the nominal modifiers, particularly adjectives, has been widely studied cross-linguistically and various syntactic and semantic motivations have been proposed (see Cinque 1994, Bouchard 2002, and Alexiadou 2001, 2004, among others). I by no means intend to present an exhaustive account of such a distinction in Persian, although the issue is certainly worthy of investigation.

derived from the same underlying structure. This underlying structure does not contain phrasal modifiers; thus, compounding involves combining head modifiers with nouns. I follow Harley (2008) in treating compounds as incorporation structures. Under this view, deriving compounds such as those in (10), whose elements are in modificational relationship (Harley's *primary compounds*), involves incorporation of the modifying element with the uncategorized Root prior to the complex Root being categorized by the category-defining head. Recall from §4.5.1.1 that in the nP domain, base-generated Root-adjoined modifiers are M-Words and hence are visible to syntactic operations. This explains why incorporation of the head modifiers into the uncategorized Root, and the syntactic head movement of the resulting incorporated Root node into the categorizing head are possible. It is crucial in Harley's approach that the modifying element be in a direct sisterhood relationship with the uncategorized Root. In Persian noun phrases, this condition applies only to modifiers inside the nP . The following configurations are the structures for the compounds in (10):

(12) a. $\hat{a}b+mive$ 'fruit juice'



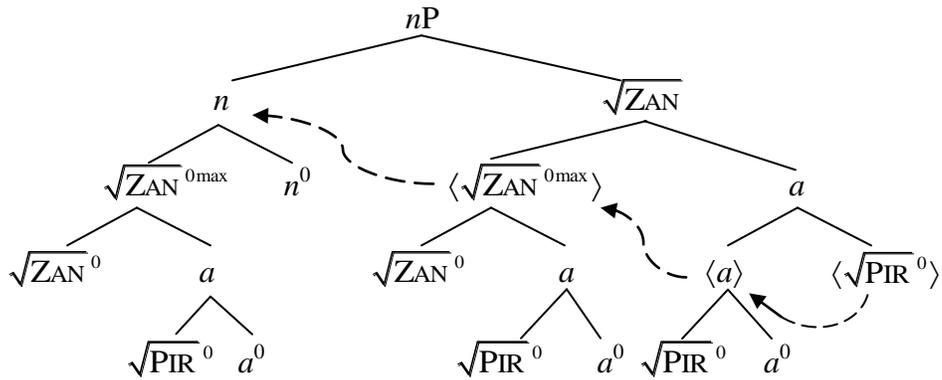
b. *šir+xošk* ‘dried milk’



As illustrated in (12), the modifiers of the Roots $\sqrt{\hat{A}B}^0$ and $\sqrt{\check{S}IR}^0$ are internally complex consisting of a Root, i.e. \sqrt{MIVE}^0 and $\sqrt{XOŠK}^0$, and a category-determining head, i.e. n^0 and a^0 . After being categorized, these modifiers incorporate into the nominal Root they modify. The resulting complex Roots head-move to their nominalizing functional projections and get categorized. What realizes as *âb-mive* ‘fruit juice’ and *šir-xošk* ‘dried milk’ at PF is thus the spell-out of these incorporated Roots embedded in the complex head nouns in (12)a&b.

The examples of inverse Ezafe compounds are derived pretty much the same way as other compounds. However, in PF derivation of inverse Ezafe compounds, the Root and the incorporated head modifier undergo reordering inside the M-Word n .

(13) *pir+zan* ‘old woman’

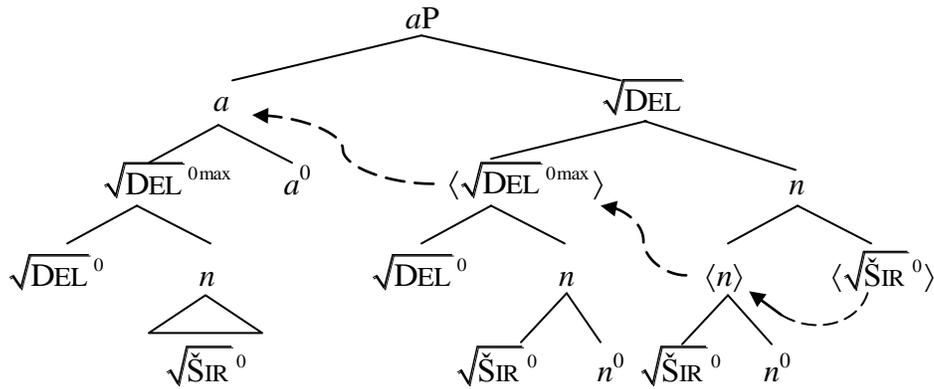


I assume that in the PF derivation of (13), the LD rule responsible for the reordering of the Root and the incorporated head modifier applies to the Concatenation statement $(\sqrt{\text{ZAN}}^0 \oplus a[\text{pir}])$ and yields $(a[\text{pir}] \oplus \sqrt{\text{ZAN}}^0)$. I will not pursue any further discussion of how LD applies M-Word-internally, as it is not relevant to the concern of this chapter⁶⁴.

In (11)a, we saw that combining the head modifier *šir* ‘lion’ with the noun *del* ‘heart’ results in an inverse adjectival compound *šir-del* ‘lion-hearted’. This shows that even in deriving adjectival compounds, head modifiers are base-generated as head-adjuncts to the uncategorized Root.

⁶⁴ Embick (2003) has found M-Word-internal LD as problematic for his theory of Local Dislocation. This type of LD operation does not bring about the structural consequences associated with adjunction. Unlike LD of M-Words in which the adjoining M-Word becomes a Subword, in M-Word-internal LD, the status of the elements involved is not affected, given that the operation adjoins a Subword to another Subword. I leave this issue for further exploration in the future. I also leave for future research the issue of why some compounds undergo reordering while others, such as those in (12), do not.

(14) *šir+del* ‘lion-hearted’



So far we have discussed compounds that involve combining two forms. I refer to the process of forming these compounds as *compounding*. Persian also has a process in which multiple forms are combined. I call this process *syntactic compounding*. Like compounding, syntactic compounding is a process in which head modifiers incorporate into the uncategorized Root prior to categorization. There is however three characteristics which distinguish these two types of compounding. First, in compounding, only one head modifier incorporates into the Root while in syntactic compounding one or more than one head modifiers undergo incorporation into the Root, as exemplified in (15). Second, in compounding, the combined form often does not have a compositional meaning while in syntactic compounding, the semantic contribution of the modifiers to the head noun is transparent. Third, in compounding, the combination of the head modifier with the head noun may produce compounds of different categories, e.g. nominal (*âb-mive* ‘fruit juice’) or adjectival compounds (*šir-del* ‘lion-hearted’). In syntactic compounding, however, the combination of the head noun with its modifier(s) generates only nouns.

(15) *Syntactic compounding*

Non-compounded phrase

- a. *pirhan-e âbi-e halqe-i-ye ċin+dâr*
dress-EZ blue-EZ loop-ADJ-EZ crease+have
'the blue creased loop dress'

Compound

- b. *pirhan âbi-e halqe-i-ye ċin+dâr*
c. *pirhan âbi halqe-i-ye ċin+dâr*
d. **pirhan âbi halqe-i ċin+dâr*

The comparison of the syntactic compounds in (15)b&c with the non-compounded version (15)a shows that in the former, the head and its modifier(s) are not linked with the Ezafe vowel *-e*. The ungrammaticality of (15)d indicates that syntactic compounding does not extend beyond two head modifiers.

Syntactic compounding is similar to compounding in that it does not allow phrasal modifiers to incorporate into the head noun.

- (16) *medâd-e PP[ru-ye miz] ⇒ *medâd ru-ye miz*
pencil-EZ on-EZ table
'the pencil on the table'

Also, phrasal modifiers cannot be incorporated into a syntactic compound consisting of the head noun and a head modifier.

- (17) medâd-e _a[meški]-ye _{PP}[ru-ye miz] ⇒ * *medâd meški ru-ye miz*
pencil-EZ black-EZ on-EZ table
‘the black pencil on the table’

Considering the facts presented so far, I suggest that both types of compounding in Persian apply within the same syntactic domain where head modifiers are head-adjoined to the uncategorized Root. Specifically, I propose that syntactic compounds are formed when an incorporated complex Root consisting of an uncategorized Root and at most two head modifiers is adjoined to the nominalizing *n* head via syntactic head movement. The nominalized complex Root, just like a simple Root, becomes part of the M-Word *n* and is thus expected to have word-level stress⁶⁵, which is in fact the case. The following configurations illustrate the procedures involved in deriving the syntactic compounds in (15)b and (15)c:

⁶⁵ See §4.5.1.1 for a discussion on the prosodic reflection of the difference between the head-adjunction structure forming the category-defining functional projection and the base-generated head-adjunction structure forming the Root projection.

b. *Compound*

* manzare+pošt+panjere-*hé*

This observation provides further support for assuming two domains for the two categorially distinct types of modifiers in the Ezafe construction. Specifically, if we assume, following Baker (1988), that syntactic compounding involves heads, the status of attributive modifiers as heads and not phrases follows. Let us consider the definiteness marking of a syntactic compound in a definite noun phrase in which both head and phrasal modifiers appear:

(21) a. *Non-compounded phrase*

ModP[*Mod*[*nP_{def}*[*n*[kafš-e *a*[meški-ye *a*[pâšneboland]]]]]-e *PP*[tu-ye vitrin]]]
shoe-EZ black-EZ high-heeled-EZ inside-EZ window-shop
'the high-heeled black shoe inside the window shop'

b. *Compound*

kafš+meški+pâšneboland-*é*-ye tu-ye vitrin

adjoined to the categorizing head via syntactic head-to-head movement. This analysis follows from our assumptions that attributive modifiers are heads and that they are base-generated as head-adjuncts to the Root. Thus, the impossibility of phrasal modifiers to enter into (syntactic) compounding implies the plausibility of assuming that they are generated outside the domain where head modifiers are generated.

Notice in (21) that prior to the application of syntactic compounding, both head and phrasal modifiers are connected to one another via the Ezafe vowel. This implies that, unlike the syntactic and PF operations we have discussed so far (e.g. licensing empty nouns and *-hâ* mobility), the application of Ezafe Insertion rule does not seem to be sensitive to the (inner vs. outer) domains of nominal modification. This is the issue that I will investigate in the next section.

5.2. Accounts of the presence of Ezafe

At the core of studies on the Ezafe construction lies the question about the motivation for the presence of the linking vowel *-e* between the head noun and its following modifiers. Before getting down to this, a quick look at the basic facts about the distribution of Ezafe is in order.

Ezafe is the unstressed vowel *-e* that connects

- the head noun to the following head/phrasal modifier;

(22) a. $nP[n[kif-e \ a[\check{c}arm-i]]]$ (N-EZ Mod_H)⁶⁷

bag-EZ leather-ADJ

‘leather bag’

b. $ModP[Mod[nP_{def}[n[kif]]-e \ PP[tu-ye \ vitrin]]]$ (N-EZ Mod_{Ph})

bag-EZ inside-EZ window shop

‘the bag inside the window shop’

- head modifiers to head/phrasal modifiers and phrasal modifiers to one another;

(23) a. Mod_H-EZ Mod_H

$nP[n[kif-e \ a[\check{c}arm-i-ye \ a[me\check{s}ki]]]]]$

bag-EZ leather-ADJ-EZ black

‘black leather bag’

b. Mod_H-EZ Mod_{Ph}

$ModP[Mod[nP_{def}[n[kif-e \ a[\check{c}arm-i]]]-ye \ PP[tu-ye \ vitrin]]]$

bag-EZ leather-ADJ-EZ inside-EZ window shop

‘the black leather bag inside the window shop’

⁶⁷ I use N for the head noun; Mod_H for head modifiers; Mod_{Ph} for phrasal modifiers; and XP for complements.

c. $\text{Mod}_{\text{Ph}}\text{-EZ Mod}_{\text{Ph}}$

$\text{ModP}[\text{Mod}[\text{nP}_{\text{def}}[\text{n}[\text{kif}]]\text{-e } \text{aP}[\text{por-e } \text{pul}]\text{-e } \text{PP}[\text{tu-ye } \text{mâšin}]]]$
 bag-EZ full-EZ money-EZ inside-EZ car
 ‘the bag full of money inside the car’

- and the head of phrasal modifiers to their complements.

(24) a. $\text{aP}(\text{A-EZ XP})$

$\text{mardom-e } \text{aP}[\text{a}[\text{âšeq-e } \text{nP}[\text{futbâl}]]]$
 people-EZ lover-EZ soccer
 ‘the soccer loving people’

b. $\text{PP}(\text{P-EZ XP})$

$\text{ketâb-e } \text{PP}[\text{P}[\text{tu-ye } \text{nP}_{\text{def}}[\text{ja'be}]]]$
 book-EZ inside-EZ box
 ‘the book inside the box’

Note in the above examples that Ezafe targets three categories: nouns, adjectives, and nominal prepositions. Through Ezafe Insertion, nouns are connected to their modifiers; adjectives are either linked to other adjectival head modifiers or to their complements; and nominal prepositions are connected to their complements. As for what triggers the occurrence of Ezafe, I will briefly give a critical survey of the main proposals in the literature and present my own proposal in §5.3.

5.2.1 Ezafe as a Case-marker (Samiian 1994; Larson & Yamakido 2005)

Samiian (1994) provides a unified account of the appearance of Ezafe on nouns, adjectives, and nominal prepositions. She suggests that, similar to English *of*, Ezafe in Persian is a dummy Case assigner that appears on categories that lack Case-assigning properties. Adopting Chomsky's (1970) syntactic feature system, she considers non-Case-assigning categories as including all [+N] categories (i.e. *n* & *a*). She further suggests that nominal prepositions are also among non-Case-assigners. Given that (true) prepositions [-V, -N] and verbs [+V, -N] share the feature [-N] and hence are Case-assigners, Samiian proposes that nominal prepositions are neutralized with respect to this feature. Drawing on the above assumptions, she proposes that Ezafe, as a dummy Case assigner, appears within phrases headed by non-Case-assigning categories, thus enabling them to Case-license their complements. She formulates a transformational rule which inserts the Ezafe vowel before every non-verbal phrasal category that occurs below the X' level (cf. (1)).

Larson and Yamakido (2005) build their analysis of Ezafe on Samiian's (1994) proposal that Ezafe is a Case-marker. With reference to the Larsonian DP structure (Larson 1991), according to which D like V selects thematic arguments, Larson and Yamakido propose that nominal modifiers in the Ezafe construction are generated as arguments of D in the postnominal position. As [+N] categories, these modifiers need to be Case-marked. Case-marking, however, occurs differently in Persian compared to English. In English, [+N] elements move up to receive Case from D while in Persian, Ezafe makes the Case available for them in the base position. This way Ezafe reveals the

underlying postnominal position of nominal modifiers, since they are Case-marked in situ.

Ghomeshi (1997) has discussed the problems with Samiian's (1994) analysis which I think can be extended to Larson and Yamakido's as well.

- Nominal modifiers in both accounts are assumed to be *complements*. Ghomeshi states that the notion of *complement* for nominal modifiers is neither semantically nor syntactically justifiable. Syntactically, the possibility of having multiple occurrences of modifiers in the Ezafe domain provides counter-argument for assuming that they are complements.
- Samiian's claim that all non-Case-assigning modifiers are Case-marked by Ezafe does not hold across the board. Ghomeshi refers to the lack of Ezafe on intensifiers occurring with adjectives, and adverbs within the sentence as counterevidence.
- The view of Ezafe as a Case-marker is particularly problematic for an analysis in which modifiers are assumed to be head-adjoined (cf. Ghomeshi's analysis and mine). Given that only maximal projections are Case-marked, it is not clear how the appearance of Ezafe on head-adjoined modifiers can be accounted for within Samiian's and Larson & Yamakido's approaches.

5.2.2 Ezafe insertion at PF (Ghomeshi 1997)

As discussed in §5.1.1.1, Ghomeshi (1997) argues that postnominal modifiers in the Ezafe domain are all bare heads, except the possessor phrase. From this she concludes that nouns in Persian are non-projecting. Given this assumption, she proposes that the

Persian NP consists of non-projecting [+N] heads (N, A, and nominal P) forming an X^0 -adjunction structure (cf. (2)). This adjunction structure is particularly identified by the appearance of the Ezafe vowel on each of the heads with which the structure is formed.

Ghomeshi (1997) proposes a PF insertion rule to account for the function of the Ezafe vowel.

(25) *Ezafe Insertion Rule*

Insert the vowel *-e* on a lexical X^0 head that bears the feature [+N] when it is followed by phonetically realized, non-affixal material within the same extended projection.

[Ghomeshi, 1997: 781.92]

According to (25), Ezafe is inserted at PF on a lexical [+N] head that is followed by an overt independent head or phrase⁶⁸ within the same extended projection⁶⁹. In other words, a lexical [+N] head that forms a constituent with another independent head or phrase needs Ezafe. Thus, Ghomeshi defines the function of Ezafe as identifying constituenthood.

With the Ezafe domain being an X^0 -adjunction structure, Ghomeshi's Ezafe Insertion Rule assigns the Ezafe vowel only to the non-projecting lexical [+N] heads. Given the fact that there are also phrasal modifiers e.g. *a*Ps and PPs in the Ezafe domain,

⁶⁸ Within Ghomeshi's approach, this is relevant for cases where, for example, the head noun is linked to the possessive noun phrase.

⁶⁹ Adopting Grimshaw's (1991) notion of *extended projection*, Ghomeshi (1997) considers the possessor phrase in the specifier of DP as an extended projection of NP on the grounds that N and D share the same categorial features.

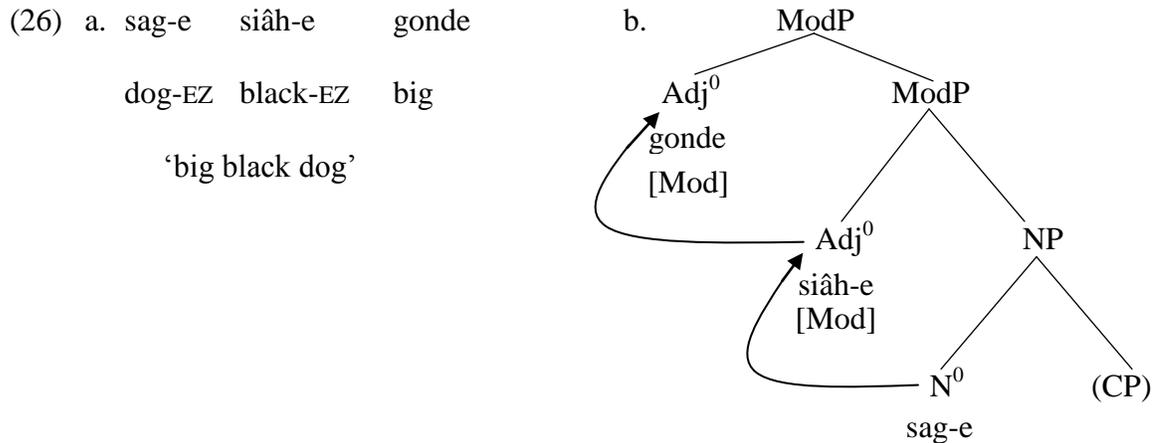
it is not clear how Ghomeshi's analysis can account for Ezafe linking maximal projections to one another.

5.2.3 Syntactic movement triggers Ezafe (Kahnemuyipour 2000, 2006)

Kahnemuyipour's (2000, 2006) analyses of the Ezafe construction are primarily built on Cinque's (1994) assumption that the cross-linguistic asymmetry concerning the relative order of nouns with respect to adjectives is the result of syntactic head raising of the noun to a functional head within the DP (see also Den Dikken & Singhapreecha 2004 and Den Dikken 2005). Relying on empirical evidence from Persian compounds, Kahnemuyipour claims that the dominance and productivity of the right-headed compounds in Persian should be taken as an indicative of the base-generation of modifiers in the prenominal position. Following Cinque, he assumes that adjectives (as well as other nominal modifiers) are base-generated in the functional projections between D and N; their postnominal surface position, then, is derived by movement of the noun. The manner in which the head noun moves and the structural status of modifiers (i.e. being heads or phrases) are two significant features that distinguish his two analyses.

Kahnemuyipour (2000) adopts Ghomeshi's (1997) views on the head status of modifiers in the Ezafe domain (cf. §5.1.1.1) and presents a head-movement analysis of the Ezafe construction. Accordingly, he assumes that modifiers are located in the heads of functional projections above NP. The modifiers have strong feature [Mod] and the functional projections are thus called Modifier Phrases (ModP). The noun, which also bears the feature [Mod], is required to move up and head-adjoin to the modifier for feature checking. Kahnemuyipour proposes that the movement of [Mod]-bearing

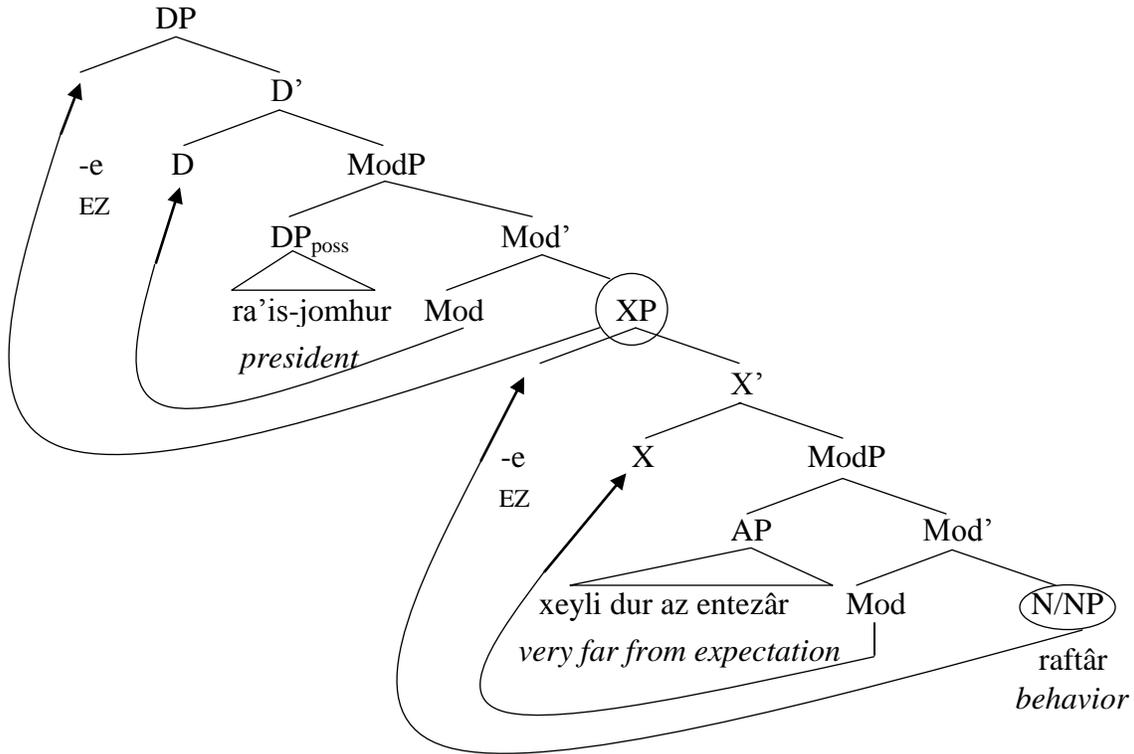
elements results in the appearance of the Ezafe vowel *-e*, which realizes the feature [Mod]. The analysis is shown in (26), taken from Kahnemuyipour (2006: 9.11):



Kahnemuyipour (2006) follows Samvelian (2006) in assuming that all the modifiers in the Ezafe domain appear maximally, i.e. with filled specifier and complement positions (cf. §5.1.1.2 above). Following Cinque (2005) among others⁷⁰, he adopts a roll-up movement analysis in which “there are additional functional projections (with null heads) mediating the noun and modifiers, with complements moving to the specifiers of these intermediate projections in a roll-up fashion” (p. 9). He particularly favors this approach because it “allows for the modifiers to be XPs occupying specifier positions rather than heads”. This type of movement analysis is illustrated in (27), taken from Kahnemuyipour (2006: 10.13):

⁷⁰ See Kahnemuyipour (2006) for the related references.

(27) raftâr-e [xeyli dur az entezâr]-e ra'is-jomhur
 behavior-EZ very far from expectation-EZ president
 'the president's totally unexpected behavior'



In this analysis, Kahnemuyipour proposes that the Ezafe vowel is the realization of the *overt* movement of Mod to the head of the XP above it or the *overt* movement of the complement to the specifier of XP. ModPs are introduced in this analysis simply as a convenient term for the locus of nominal modifiers. Kahnemuyipour leaves out any discussion as to the exact nature of these projections. This is contrary to Kahnemuyipour (2000) in which ModPs introduce modifiers as their functional heads bearing strong feature [Mod].

One problem with Kahnemuyipour’s analyses is that they do not capture the empirical facts discussed in Samiiian (1983) and Samvelian (2006). Specifically, his (2006) analysis does not reflect the fact that attributive modifiers in the Ezafe domain do not take complements. His (2000) analysis, on the other hand, does not address the fact that there can be *a*Ps and PPs in the Ezafe domain.

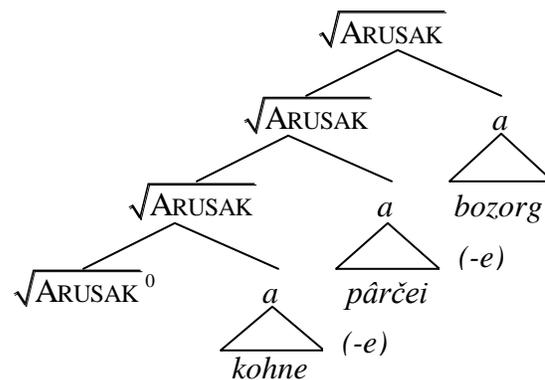
The other problem with Kahnemuyipour’s analyses arises from adopting Cinque’s assumption that nominal modifiers are generated in the functional projections above the NP. If Cinque’s assumption holds for Persian, one would expect a strict order of attributive adjectives in the Ezafe domain. This is however not borne out in Persian. As shown in (28), the adjectives *kohne* ‘old’, *bozorg* ‘big’, and *pârčei* ‘of cloth’ can be combined in various orders.

- (28) a. arusak-e kohne-ye pârče-i-ye bozorg
 doll-EZ old-EZ cloth-ADJ-EZ big
 ‘the old big cloth doll’
- b. arusak-e kohne-ye bozorg-e pârče-i
- c. arusak-e pârče-i-ye bozorg-e kohne
- d. arusak-e pârče-i-ye kohne-ye bozorg
- e. arusak-e bozorg-e kohne-ye pârče-i
- f. arusak-e bozorg-e pârče-i-ye kohne

In an informal survey of five native speakers of Persian that I conducted, almost all of the above possibilities were judged as acceptable. The ones the speakers disagreed on were random.

Given my proposed configuration in (7), the presence of multiple attributive adjectives in the inner (*nP*) domain of modification translates into multiple head-adjunction of those adjectives to the Root.

(29) Multiple attributive adjectives in the Ezafe domain



I suggest that the completely random order of attributive adjectives that we observed in (28) follows from their structural status as *adjuncts*. Hence, flexibility of the order of adjectives in the inner domain of modification provides further support for assuming that they are head-adjoined.

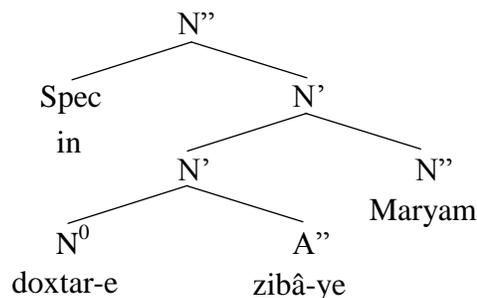
5.2.4 Ezafe as phrasal affix (Samvelian 2006)

Samvelian (2006) argues that Ezafe is a phrasal affix, morphologically realized at the lexical level and interpreted within syntax. Building on the lexicalist approach that she

adopts, “affixal elements must be handled morphologically and do not correspond to nodes in the syntactic representation” (p. 3). Relying on this assumption, she does not consider syntax to have any bearing on the restrictions that may exist on the Ezafe construction. She argues that different restrictions on the Ezafe construction emanate from the Ezafe’s morphological status and its interaction with other phrasal affixes that appear within the NP. Accordingly, the constraint she proposes on Ezafe is morphological by nature: phrasal affixes are barred from stacking.

Apart from Ezafe being treated as a phrasal affix, Samvelian also assumes, as noted earlier, that affixal elements do not generally correspond to nodes in the syntactic representation. Based on this premise, the Ezafe vowel along with other affixal elements such as indefinite determiner *-i*, pronominal enclitics, and the colloquial definite marker *-é* are handled morphologically and never head their projections. This leads Samvelian to oppose the DP-hypothesis adopted in the previous accounts and assume instead that NP is the ultimate projection of Ns. She proposes the structure in (30) for *in doxtar-e zibâ-ye Maryam* ‘this beautiful girl of Maryam’s (lit. this Maryam’s beautiful girl)’ as a representative of the structure for Persian NPs:

(30) *The structure of Persian NP* (Samvelian 2006: 45)



In (30), Samvelian has assumed a hierarchical binary-branching structure under N' for the Ezafe domain, in which all modifiers and the possessor NP appear. The nominal modifiers in the Ezafe domain are assumed to be all phrasal complements to the head noun. This is similar to the structure proposed by Samiiian in (1) except in the fact that Samvelian's proposed tree is not a flat structure.

I leave aside further discussion of Samvelian's analysis because of the differences that exist between the tenets of the Lexicalist framework and those of Distributed Morphology. Within DM, which I follow in this dissertation, word and sentence structure are handled both by a single generative engine, namely syntax. In the Lexicalist framework, however, morphology and syntax are two separate components of grammar with their own set of rules (see e.g. Halle & Marantz 1993 for further details).

5.3. Proposal: Ezafe insertion at the syntax-phonology interface

Having surveyed the previous accounts of Ezafe insertion, we have seen that the Ezafe vowel is viewed as being

- inserted presyntactically at the lexical level (Samvelian 2006);
- realized as a result of syntactic movement (Kahnemuyipour 2000, 2006);
- realized in the syntax as a Case-marker of D arguments (Larson & Yamakido 2005);
- inserted transformationally before every non-verbal phrasal category dominated by X' (Samiiian 1983); or
- inserted at PF to identify constituenthood (Ghomeshi 1997).

In my account, I adopt Ghomeshi's view on the Ezafe vowel as a linking element. According to her, the Ezafe vowel does not correspond to any lexical or semantic feature and, in fact, seems to have no semantic content at all. For this reason, she does not consider the Ezafe vowel as a *morpheme*. As such, it cannot appear as heading any sort of syntactic projection. Nor can this vowel attach presyntactically due to its sensitivity to syntactic environments. Instead, she proposes that Ezafe insertion takes place in the spell-out component of PF. She suggests that Ezafe insertion should be viewed as an example of PF rules "that concern syntactic constituents smaller than the sentence and which affect the phonological realization of these constituents", e.g. English stress rule or French liaison (p. 779). Ghomeshi, however, does not detail the mechanism involved in the insertion of the Ezafe vowel in PF. This is the task that I will undertake in this chapter.

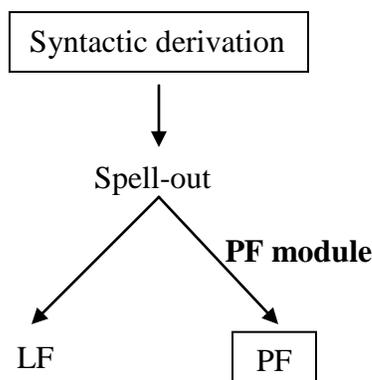
In this section, adopting Pak's (2008) model of syntax-phonology interface within DM, I will propose that the Ezafe vowel, as a linking element, is inserted in the process of linearization of the syntactic structure at PF. Under my proposal, Ezafe Insertion rule is a phonological rule that applies at the final stage of Linearization process. As a Late-Linearization rule, it inserts the Ezafe vowel onto an overt nominal head ([+N] M-Word) when it is followed by one or more nominal words. The Ezafe vowel is then spread rightward from the nominal head onto each and every nominal word that appears in the complement or functions as the modifier(s) of the head, stopping short of the last word that bears the phrasal stress. The nominal words that follow the nominal head must be members of categories *n* and *a*; words of categories other than these do not take the Ezafe

vowel. Before I lay out my proposal in detail, I will briefly review the relevant aspects of Pak’s (2008) model of syntax-phonology interface.

5.3.1 A theory of syntax-phonology interface (Pak 2008)

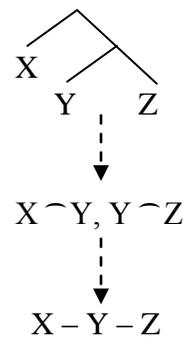
Pak (2008) proposes a model which (i) maintains a *direct* interface between phonology and syntax⁷¹, but that (ii) can accommodate cases of variability in syntax-phonology mapping. The fundamental assumption in Pak’s model is that syntactic structures are computed in phases and that phonological rules deal directly with the output of each phase. She adopts an articulated, derivational model of the syntax-phonology interface in which syntactic structures are converted into linear order and provided with phonological exponents by a series of PF operations.

(31) Architecture



The syntax deals with abstract, linearly unordered hierarchical structures, which are sent to the PF and LF components at spellout.

By the end of the PF branch, the structure has been fully linearized and ‘phonologized’, so that it can be input to the performance systems.



[Pak 2008: 41.54]

⁷¹ Pak’s (2008) model is an instantiation of a *direct-reference* model, which is based on the premise that “phonological rules apply directly to morphosyntactic structures” (p. 45)

Within this architecture, Pak (2008: 6) proposes that in the course of PF derivation, “phonological rules are interleaved with different kinds of *linearization procedures*,” such that linear order is established over objects of gradually increasing size—single words, then pairs of words ($X \text{ } Y$), then longer chains of words ($X\text{--}Y\text{--}Z$). Specifically, she introduces three types of linearization rules that are involved in this process (p. 98.28):

- (32) a. **Concatenation rules** apply within each spellout domain, as linear relations are established between pairs of M-words. As such, they are (i) highly local, unable to ‘see’ more than two M-words at a time; and (ii) unaffected by speech rate or other information that becomes available later in the derivation.
- b. **Chaining rules** apply later in the PF derivation, after an n -ary string of M-words has been created internal to each spellout domain.
- c. **Late-Linearization rules** apply when separate spellout domains are linearized with respect to one another. These rules use Chains as discrete chunks in the default case, but may merge Chains together (e.g. in fast speech) or split Chains apart (e.g. in slow or careful speech), depending on performance-related factors that come into play late in the derivation.

With respect to rule ordering, Pak’s model predicts that rules applying to smaller domains generally *precede* rules that apply to larger domains. Accordingly, Concatenation rules generally precede Chaining rules, and Chaining rules precede Late-Linearization rules.

Unlike many instantiations of Prosodic Hierarchy Theory (Nespor and Vogel 1986, Selkirk 1986, Inkelas and Zec 1995, among others), Pak's model allows some phonological rules to refer to particular aspects of the underlying morphosyntactic structure. Specifically, she refers to "Concatenation rules [that] can impose additional conditions on rule application based on specific morphosyntactic features and categories" (p. 73: ft. 6). As an example, she mentions Low Tone Deletion rule in Luganda which fails to apply in certain verb tenses (see her Chapter 5 for detailed discussion). Under this model, reference to morphosyntactic and categorial information is constrained by spellout domain boundaries.

Let us now turn to Ezafe insertion and see how it can be derived within Pak's model.

5.3.2 Sensitivity of Ezafe insertion to morphosyntactic structure

Ghomeshi (1997) notes that the implicit theme in Samiiian (1983) and Karimi and Brame (1986) is that the insertion of Ezafe is motivated by the morphosyntactic feature [+N] borne by lexical heads. She argues that nouns, adjectives, and nominal prepositions bear the feature [+N] and should thus appear with the Ezafe vowel. As implied in Ghomeshi's Ezafe Insertion Rule (cf. (25)), the mere presence of a [+N] lexical head does not trigger the insertion of the Ezafe vowel. Ezafe insertion rather occurs only when a [+N] head is followed by an overt modifier or complement. This explains why the Ezafe vowel does not appear on a bare noun (*ketâb(*-EZ)* 'book'). Thus, Ezafe appears on a [+N] head X when it is followed by a [+N] head Y, as schematized and exemplified in (33).

- (33) a. $X_{[+N]} - Y_{[+N]} \Rightarrow X-e Y$
 b. $x[\text{ketâb}] - e \ y[\text{sefid}]$
 book-EZ white
 ‘the white book’

Ghomeshi (1997) also explains that the Ezafe vowel does not appear on prenominal modifiers (e.g. demonstratives, quantifiers, superlatives, etc.) and on intensifiers preceding adjectives in the Ezafe domain because they are functional rather than lexical heads. Within the framework I have adopted in this thesis, this means that Ezafe appears only on the overt heads (M-Words) corresponding to the category-defining heads *n* and *a* (i.e. nouns and adjectives). Thus, the [+N] heads X and Y in (33)a correspond to *n* or *a*⁷².

What is more important for Ezafe insertion than the categorial identity of X and Y in (33) is the kind of relationship that holds between them. Considering the example in (33)b as well as those in (22) to (24), we see constituents following nominal heads *n* and *a* are either modifiers, possessors, or complements. I propose that the occurrence of Ezafe in these contexts is to establish relationship between nominal heads and their dependents. Here I use *dependent* as a cover term for modifiers, possessors, and complements of nominal heads in Persian noun phrases⁷³. The relevant generalization for Ezafe insertion is as follows:

⁷² As I will show later in my analysis of the Ezafe construction, nominal prepositions are nouns and thus head the noun phrase complement of a PP headed by a null locative preposition (cf. §5.3.3.2).

⁷³ Philip (2009) proposes a theory of linkers as dependent-markers. Under this theory, linkers in the complex noun phrase are particles that may be used to mark modifiers, possessors, and complements. I adopt this sense of the term dependent in this thesis, although I depart from Philip’s proposal that linkers are particles and thus present in the syntax.

(34) The Ezafe vowel *-e* is a linker that marks head-dependent relationship in phrases headed by *n* and *a*.

The significance of the generalization stated in (34) becomes particularly relevant when we consider contexts in which two nominal heads are not in head-dependent relationship and thus Ezafe insertion is not allowed. Consider the following example:

(35) x [ketâb](*-e) y [sefid] ast.
book(*-EZ) white is
'The book is white.'

In the example above, *ketâb* 'book' and *sefid* 'white' belong to two different phrases (*nP* and *aP*) occurring in distinct clausal positions: *ketâb* in subject position and *sefid* in predicative position. The lack of head-dependent relationship between these two nominal elements renders the presence of Ezafe ungrammatical.

Notice that in all examples of Ezafe constructions that we have considered so far, Ezafe appears on a nominal head or dependent that is followed by another dependent. For this reason, the last element in the Ezafe construction never takes Ezafe, though being in a head-dependent relationship.

Given the above explanation of the conditions on Ezafe insertion and the generalization in (34), the following rule can be formulated:

(36) *Ezafe Insertion Rule*

Insert the vowel *-e* on an overt [+N] M-Word X when followed by (an) overt [+N] M-Word(s) Y if

- a. X and Y are category-defining heads *n* or *a*; and
- b. Y is a dependent of X,

such that the vowel *-e* is spread rightward from X onto each instantiation of Y, stopping short of the last word.

The fact that Ezafe insertion on overt nominal heads is conditioned on the presence of dependents following the head implies that the Ezafe vowel is inserted at the level of the phrase, rather than the word (contra Samvelian 2006). Thus, the Ezafe Insertion Rule (EIR) applies to a phrase headed by an overt [+N] M-Word that is of category *n* or *a* and followed by other phonologically realized [+N] M-Words. The rule inserts the Ezafe vowel on the [+N] head and extends it rightward onto [+N] M-Words inside the modifiers and complements. Let us investigate in detail the implementation of EIR to examples of Persian noun phrases.

5.3.3 Towards establishing the Ezafe insertion domain

Recall our assumption from §5.1.2 that the domain of modification in Persian noun phrases is divided into two distinct subdomains. The *inner domain* (*nP*) is where head modifiers are generated, and the *outer domain* (*ModP/PossP*) is where phrasal modifiers occur. Licensing empty nouns, *-hâ* mobility, and compounding only occur within the

inner domain. As we have seen, however, (cf. §5.2) Ezafe appears within both subdomains; that is, Ezafe links both head and phrasal modifiers to the head noun.

In this section, I will show that EIR applies as a Late-Linearization rule within and across the inner and outer domains. To elaborate, within Pak’s model, I propose that EIR applies at the point of the PF derivation at which all spellout domains within a noun phrase are linearized. At this stage, I suggest, the inner and outer domains are merged into a single domain for the Late-Linearization rule of EIR. The rule picks out the head noun, inserts Ezafe on it, and extends the Ezafe vowel onto all following M-Words that meet condition (36)a&b, except the last one. This procedure is schematically shown below:

(37) *Ezafe insertion at Late-Linearization*

i. Input structure:

PRENOMINAL DOMAIN $PF_{NP}(inner\ domain)\ PF_{ModP/PossP}(outer\ domain)$

Internal structure: $(H_n\ Mod_H\ Mod_H)\ (Mod_{ph}\ Mod_{ph}\ Mod_{ph})$

ii. Chain-merging:

PRENOMINAL DOMAIN $(inner\ domain + outer\ domain = Ezafe\ insertion\ domain)$

Internal structure: $(H_n\ Mod_H\ Mod_H + Mod_{ph}\ Mod_{ph}\ Mod_{ph})$

iii. EIR:

It should be mentioned that phrasal modifiers (Mod_{ph}) that occur in the Ezafe insertion domain can themselves be comprised of a string of M-Words. Since these modifiers are headed by [+N] M-Words, an Ezafe insertion domain is formed phrase-internally in

which the head is linked to the [+N] M-Words in its complement via Ezafe (cf. §5.3.3.2 for an account of Ezafe insertion inside the phrasal modifiers).

To summarize, I will establish that the domain for the application of EIR is the whole modification domain in Persian noun phrases and Ezafe is a dependent-marker. I will also argue, following Matushansky (2008), that the prenominal position in Persian is the domain of quantifiers, and that the semantic distinction between quantification and modification explains the absence of Ezafe on demonstratives, quantifiers, superlatives, and numerals.

5.3.3.1 The *nP*

As we have established in the previous chapters, assuming a bottom-up structure building derivation, Persian noun phrases start out with the *nP* phase, where the head noun and its head modifiers are generated. Given Pak's model, the output of the *nP* phase will be subject to a series of PF operations that linearize and phonologize the M-Words contained in the spellout domain of the *nP* phase.

The term *spell-out domain* is standardly defined as referring to a phase head complement (Chomsky 1999 *et seq.*). Within Pak's model, however, phonological domains often map onto a phase head complement *and* the phase head, although in some cases, they line up with only a phase complement. This makes the spellout domain a vague concept in that it can include domains larger than a phase complement. Since in Persian noun phrases, the head noun in the *nP* phase must be spelled out with its following modifiers for the purposes of Ezafe insertion, I assume that the spellout domain

of the *nP* consists of the phase head and its complement⁷⁴. Given that there is no edge position in this phase, the spellout domain is in fact the whole phase.

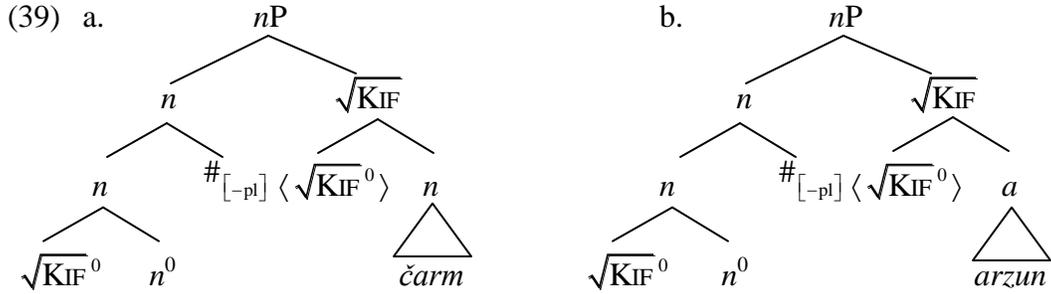
In the course of the PF linearization process, a number of intermediate representations are created, some of which constitute a domain for the application of certain morphological (e.g. Local Dislocation) or phonological rules (cf. Pak 2008 for examples). Within Pak's model, I propose that EIR, as a phonological rule, applies after these intermediate representations are clustered together at Chaining; hence, it is a Late-Linearization rule.

We begin with the PF derivation of a simple noun phrase in which the head noun is followed by a head modifier. Recall that in the *nP*, two types of head modifiers can follow the head noun: the attributive noun (38)a and the attributive adjective (38)b. The Ezafe vowel appears on the head noun only when these modifiers are present in the *nP*.

- | | | | | | |
|---------|--------|---------------|----|--------|-------------|
| (38) a. | kif-e | čarm | b. | kif-e | arzun |
| | bag-EZ | leather | | bag-EZ | cheap |
| | | 'leather bag' | | | 'cheap bag' |

Recall also that head modifiers of the noun appear in a base-generated head-adjunction structure, as shown in the structural representations of the above examples in (39).

⁷⁴ See Pak (2008: 106.ft. 12) for a similar case in Huave where the phase head is spelled out with its complement. She also refers to other phase-based proposals that predict that the phase head groups with its complement (see her Chapters 1 and 2 for a review of such proposals).



At PF, the linearization process gives us the statements of Concatenation derived from statements of structural left-adjacency (*-statements) between the M-Words in the above syntactic configurations⁷⁵. Chaining then strings together the M-Words contained in the nP spellout domain.

(40) PF: i. Concatenation⁷⁶

a. $n[kif] \frown n[\check{c}arm]$

b. $n[kif] \frown a[arzun]$

ii. Chaining

a. $n[kif]-n[\check{c}arm]$

b. $n[kif]-a[arzun]$

Given the EIR in (36), the Chains in (40)ii represent the morphosyntactic structure that meets the requirements for the application of EIR. The rule applies at Late-Linearization stage and inserts the Ezafe vowel *-e* onto the M-Word $n[kif]$, the head of the nP , thus

⁷⁵ See Chapters 1 and 4 for an explanation of steps involved in linearization of syntactic structures at PF.

⁷⁶ For reasons of space, I simply use the phonological exponents of M-Words instead of representing their internal structure. Thus, $n[kif]$ is the abbreviated representation of $(\sqrt{KIF} \oplus n)$.

connecting it to its following head modifiers $n[\check{c}arm]$ and $a[arzun]$. The phonologization of the output of EIR yields the surface phonetic representation or *Surface PF*⁷⁷.

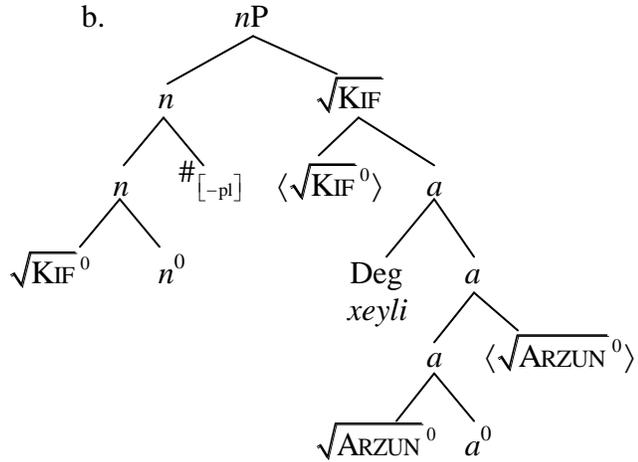
(41) Late-Linearization

- | | | |
|-----------------|-------------------------------|------------------------|
| i. EIR: | a. $n[kif]-e-n[\check{c}arm]$ | b. $n[kif]-e-a[arzun]$ |
| ii. Surface PF: | a. kif-e čarm | b. kif-e arzun |

Three remarks are in order with respect to the nature and mechanism of EIR as it applies in the nP domain. First, *EIR is not sensitive to linear adjacency*. Looking at the application of EIR in the PF derivation of the simple nPs above, it may seem that the Ezafe vowel is inserted on the head noun when it is linearly adjacent with its following modifier. But the head noun need not be adjacent to its modifier as in the case when intensifiers are adjoined to the attributive adjective in the inner domain of modification (nP). Consider the example in (42) and its PF derivation in (43).

⁷⁷ The term PF is used to refer to both the derivation along the PF branch and the surface phonetic representation that is produced at the end of the branch. Following Pak (2008: 26.ft. 1), I use the term *Surface PF* to specifically refer to the latter meaning.

(42) a. kif-e xeyli arzun
 bag-EZ very cheap
 ‘a very cheap bag’



(43) PF

- i. Concatenation: $n[\text{kif}] \frown \text{Deg}[\text{xeyli}]$
 $\text{Deg}[\text{xeyli}] \frown a[\text{arzun}]$
- ii. Chaining: $n[\text{kif}] - \text{Deg}[\text{xeyli}] - a[\text{arzun}]$
- iii. Late-Linearization
- iv. EIR: $n[\text{kif}] - e \text{ Deg}[\text{xeyli}] a[\text{arzun}]$
- v. Surface PF: kif-e xeyli arzun

If EIR were sensitive to linear adjacency, it would take $n[\text{kif}] \frown \text{Deg}[\text{xeyli}]$ in (43)i as the input and would fail to apply because the M-Word following $n[\text{kif}]$ is not of category n or a , and does not modify it. This is the incorrect result. Given that EIR can see past the immediately following M-Word $\text{Deg}[\text{xeyli}]$ ‘very’ to $a[\text{arzun}]$ ‘cheap’, it takes the entire nP as its domain, which in turn suggests that it is a Late-Linearization rule. This supports Pak’s (2008: 99-100) generalization that “[a]ny time a phonological rule requires reference to material spanning a domain larger than two immediately adjacent M-words,

[...] it applies after Chaining”. Accordingly, EIR applies to the output of Chaining in the manner schematized in (43)⁷⁸.

The second remark concerns Ezafe insertion in *nPs* containing multiple head modifiers. In such contexts, *the Ezafe vowel undergoes spreading* from the nominal head onto each of the following dependents ([+N] M-Words), stopping short of the last one. As can be seen in (44), the Ezafe vowel appears on the head noun and the two following head modifiers; the third modifier does not take Ezafe because it is the last dependent in the EIR domain. Ezafe thus links the head noun to a head modifier, and the three head modifiers to one another.

- (44) *nP*[*n*[kif-e *a*[čarm-i-ye *a*[meški-e *a*[gerun]]]]]]
 bag-EZ leather-ADJ-EZ black-EZ expensive
 ‘expensive black leather bag’

Again, this observation strengthens our proposal that EIR is a Late-Linearization rule. It takes an entire phrase—potentially containing more than two M-Words—and treats it as a *single object* for the purposes of Ezafe insertion. Given that in the *nP*, the domain of EIR is the entire phrase, the rule inserts Ezafe on the head noun and extends it rightward onto any number of head modifiers that appear in this domain. If the EIR were a Concatenation rule that could only see two M-Words at a time, the information that *meški* ‘black’ and *gerun* ‘expensive’ are modifiers of *kif* ‘bag’ would not be accessible.

⁷⁸ Taking the Concatenation statements in (43) as input, the Chaining operation ensures that each M-Word is pronounced once.

(45) Concatenation

- a. $n[\text{kif}] \frown a[\text{čarmi}]$
- b. $a[\text{čarmi}] \frown a[\text{meški}]$
- c. $a[\text{meški}] \frown a[\text{gerun}]$

Assuming that EIR is a Late-Linearization rule that applies to the entire nP after Chaining resolves the problem.

(46) PF

- i. Chaining: $n[\text{kif}]- a[\text{čarmi}]- a[\text{meški}]- a[\text{gerun}]$
- ii. Late-Linearization
EIR: $n[\text{kif}]-e- a[\text{čarmi}]-e- a[\text{meški}]-e- a[\text{gerun}]$
- iii. Surface PF: $\text{kif-e čarmi-e meški-e gerun}$

Thus, nPs with multiple head modifiers provide a sound argument in support of our claim that the EIR maps onto the entire spellout domain of the nP .

The third remark concerns the application of EIR in contexts in which the spellout domain of the nP is linked to the spellout domains of phrasal modifiers aP , PP , and Possessor nP at the Late-Linearization stage. In such cases, *EIR applies across spellout domains such that it merges the spellout domain of the nP with those of its following phrasal modifiers*. I will pursue this point in detail in the next subsection.

5.3.3.2 The phrasal modifiers in ModP/PossP

In addition to linking the head noun to its head modifiers in the *nP*, EIR as a Late-Linearization rule applies to connect (i) the head noun/the *nP* to its phrasal modifiers contained in the ModP/PossP (47); (ii) the transitive adjective to its complement (48); and (iii) the nominal preposition to its complement (49).

(47) a. *ModP*[*Mod*[*nP*_{def}[*n*[livân]]]-e *aP*[por-e âb]-e *PP*[*P*[ru-ye *nP*_{def}[miz]]]]]

glass-EZ full-EZ water-EZ on-EZ table

‘the full glass of water on the table’

b. *PossP*[*Poss*[*nP*_{def}[*n*[jib]]]-e *aP*[por-e pul]-e *nP*_{def}[*n*[Ali]]]]]

pocket-EZ full-EZ money-EZ Ali

‘Ali’s full pocket of money’

(48) *aP*[*a*[negarân-e *nP*[rokud-e eqtesâd-i]]] hast-im.

worried-EZ recession-EZ economy-ADJ is-1PL

‘We are worried about economic recession.’

(49) *PP*[*P*[tu-ye *nP*_{def}[ja’be]]]

inside-EZ box

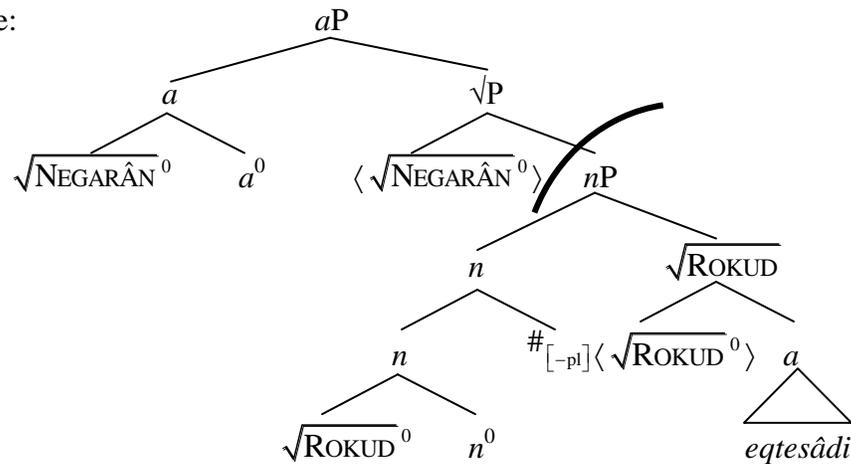
‘inside the box’

I will first discuss Ezafe insertion internal to the phrasal modifiers in (48) and (49), and then turn to how they are connected to the *nP* and to one another via the Ezafe vowel.

The *aP*

As can be seen in (48), the *nP* complement of the transitive adjective is introduced by the Ezafe vowel⁷⁹. Given our assumption that category-defining heads e.g. *n*, *a* are phase heads that trigger spellout (Embick 2008), the PF derivation of (48) involves two cycles. One is where the *nP* is spelled out and the other is where the material left under the root node (*aP*) is spelled out. I assume that the syntactic structure in (50) is the input to PF:

(50) Structure:



⁷⁹ Most transitive adjectives can either take prepositional complements or be connected to their complements via the Ezafe vowel. Transitive adjectives with Ezafe are used in colloquial Persian more often than the ones with prepositional complements.

- | | | | | |
|------------|------------|-----------|----------------|--------|
| (i) a. por | az | âb | b. por-e | âb |
| | full | of | | water |
| (ii) a. | alâqemand | be musiqi | b. alâqemand-e | musiqi |
| | interested | to music | | |

Since my focus in this thesis is on Ezafe insertion, I will not set aside transitive adjectives with PP complements.

At PF, Concatenation and Chaining apply internal to the *nP* and *aP* spellout domains, yielding the following statements:

(51) PF

- a. *nP* spellout domain (cycle I)
 - i. Concatenation: $n[\text{rokud}] \frown a[\text{eqtesâdi}]$
 - ii. Chaining: $n[\text{rokud}] - a[\text{eqtesâdi}]$
- b. *aP* spellout domain (cycle II)⁸⁰
 - i. \oplus -Concatenation: $\sqrt{\text{NEGARÂN}}^0 \oplus a^0$
 - ii. Chaining: $a[\text{negarân}]$

Since EIR is a Late-Linearization rule, its application in PF derivations involving two or more cycles is delayed until after the derivation is complete and all spellout Chains are linearized with respect to each other. The input to EIR is therefore the output of the Late-Linearization operation that strings together the separate spellout domains in a derivation. Accordingly, at the end of the PF derivation of the *aP* in (50), the spellout Chains $a[\text{negarân}]$ and $n[\text{rokud}] - a[\text{eqtesâdi}]$ are linearized. The result constitutes the domain in which EIR applies. The rule inserts Ezafe on the first [+N] M-Word in the string ($a[\text{negarân}]$) and extends it rightward onto the rest of the [+N] M-Words following $a[\text{negarân}]$. Given that the domain of the application of EIR is formed by clustering

⁸⁰ The *aP* spellout domain contains a single overt M-Word $a[\text{negarân}]$. Vocabulary Insertion and linearization in this context apply M-Word-internally where $\sqrt{\text{NEGARÂN}}^0$ is concatenated with the category-defining head $a(\sqrt{\text{NEGARÂN}}^0 \oplus a)$.

together two separate spellout domains, the two domains need to be merged into a single domain in order for EIR to establish head-dependent relationship between the first M-Word *a*[negarân], which is the head of the *a*P, and the relevant M-Words inside its complement, i.e. *n*[rokud] and *a*[eqtesâdi] (cf. (37) for a schematic representation of Chain-merging and the subsequent application of EIR). Within Pak’s (2008) model, Chain-merging is the hallmark property of Late-Linearization rules. The Late-Linearization process and application of EIR is illustrated below (parentheses mark domains):

- (52) i. Late-Linearization: (*a*[negarân])–(*n*[rokud]–*a*[eqtesâdi])
 ii. EIR: (*a*[negarân]–*e*–*n*[rokud]–*e*–*a*[eqtesâdi])
 iii. Surface PF: negarân–*e* rokud–*e* eqtesâdi

This derivation implies that morphosyntactic features are visible to phonological rules such as the EIR until the last stage of the PF derivation—when Late-Linearization rules apply. This is unsurprising in direct-reference models like Pak’s, given that “direct-reference theories [...] can in principle use any type of information that is available in the syntax to define phonological domains, including XP edges, head-complement relations, or branchingness...” (Pak 2008: 51).

Let us now turn to Ezafe insertion inside nominal PPs.

The nominal PP

The categorial status of what Ghomeshi (1997) terms as *nominal PPs* has been subject to controversy. Ghomeshi (1997), Karimi & Brame (1986), and Larson & Yamakido (2005) consider them as nouns; Samiian (1983) as prepositions; and Patcheva (2007) as Axial Parts, i.e. neither nouns nor prepositions. Prior to discussing Ezafe insertion inside this phrasal modifier, I will evaluate these views and present my own analysis as an alternative. I will propose that they are nominal locative expressions that appear as PossPs in argument positions and PPs in adjunct positions. I will show that as phrasal modifiers of a noun, they appear as locative PPs (PP_{LOC}).

Samiian (1983) makes a distinction between prepositions that appear with Ezafe and those that do not. Her P1 class of preposition includes prepositions like *dar* ‘in’, *bâ* ‘with’, *az* ‘from’, *bar* ‘on’, etc. that never appear with Ezafe. Her P2 class of preposition, which includes the majority of prepositions in Persian, is split between prepositions that obligatorily or optionally appear with Ezafe. Ghomeshi (1996) terms Samiian’s P1 class *true* prepositions. She calls P2 class of prepositions *nominal* prepositions because they show some properties of nouns; e.g. they appear with demonstratives (*in jelo* ‘this front’) and can take plural marking (*nazdik-â-ye zohr* near-PL-EZ noon ‘around noon’)⁸¹. Below is a list of (obligatory and optional) nominal prepositions (taken from Ghomeshi, 1997):

⁸¹ See also Pantcheva (2007: §3.2) for a thorough investigation of noun-like properties of nominal prepositions.

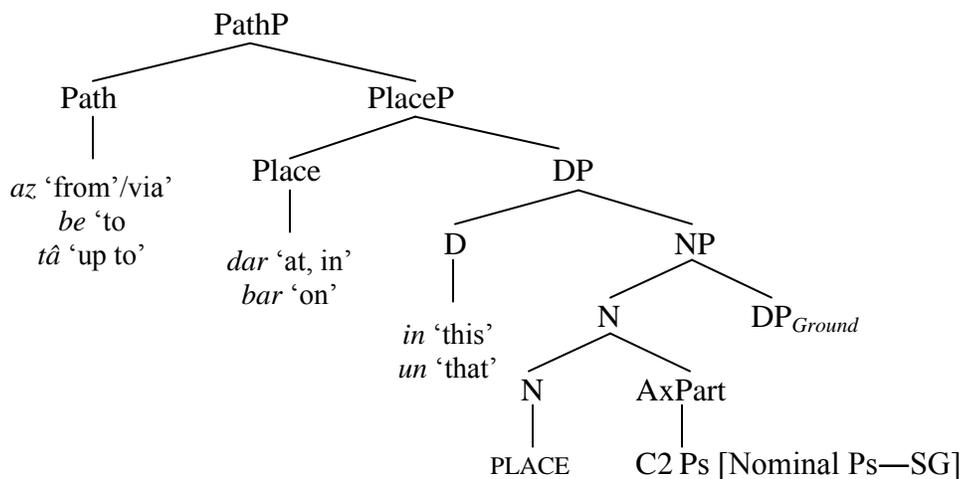
(53) *Nominal prepositions*

<u>P + -EZ</u> (Ezafe obligatory)		<u>P (+ -EZ)</u> (Ezafe optional)	
bedun-e	‘without’	tu(-ye)	‘inside’
beyn-e	‘between’	ru(-ye)	‘on’
bâvojud-e	‘in spite of’	jelo(-ye)	‘in front of’
zir-e	‘under’	pahlu(-ye)	‘next to’
sar-e	‘at’	bâlâ(-ye)	‘on top of’
pošt-e	‘behind’	kenâr(-e)	‘beside’
pain-e	‘below’	barâ(-ye)	‘for’
nazdik-e	‘close to/near’	*****	
mesl-e	‘like’	<u>P + -EZ</u> (Ezafe obligatory)	Cont’d
alâraqm-e	‘in spite of’	atrâf-e	‘around’
birun-e	‘outside’	piš-e	‘beside’
dâxel-e	‘inside’	dor-e	‘around’
xârej-e	‘outside’	alayh-e	‘against’ (in an argument)
miun-e	‘among’	lah-e	‘for’ (in an argument)
vasat-e	‘in between/the middle of’		

Contrary to Samiiian (1983) who argues against nominal prepositions being nouns, Ghomeshi—following Karimi & Brame (1986)—and Larson & Yamakido (2005) assume that they constitute a subclass of nouns. Thus, they consider nominal prepositions as [+N] categories along with nouns and adjectives.

Pantcheva (2007) follows Svenonius (2006) in assuming that noun-like locative elements, like most nominal prepositions in Persian, head Axial Parts⁸² (AxPart), a projection within an articulated PP structure, shown in (54). AxParts, according to Pantcheva, “do not refer to a certain part of an object but to a space specified with reference to that object” (p. 20). Pantcheva adopts the view held by Svenonius (2006) that AxParts constitute a separate syntactic category, distinct from nouns and prepositions. Under Pantcheva’s analysis, functional (true) prepositions select a DP containing a null noun PLACE modified by a nominal preposition, i.e. an AxPart.

(54) *Persian PPs* (Pantcheva 2007: 19.47)



The null noun PLACE in the above structure “defines a spatial relation slot which has to be satisfied by an AxPart” (Pantcheva 2007: 20). This suggests that the nominal properties of this class of prepositions come from the structural relationship between a nominal

⁸² Svenonius (2006) adopts the term *axial parts* from Jackendoff (1996). The *axial parts* of an object, according to Jackendoff, refer to its *top, bottom, front, back, sides, and ends*.

preposition (AxPart) and a null noun PLACE, where the AxPart modifies PLACE and hence inherits its nominal properties.

The nominal prepositions listed in (53) are not all instances of Axial Parts (cf. ft. 82). There are prepositions like *bedun-e* ‘without’, *barâ-ye* ‘for’, and *bâvojud-e* ‘in spite of’ that do not express spatial meanings. In addition, nominal prepositions are not always complements of true prepositions, contra Pantcheva (2007). As I will shortly discuss, nominal prepositions may appear in different categories based on the syntactic position in which they occur.

Introducing prepositions in Modern Persian, Sadeghi (1970) explains that what we call nominal prepositions here are in fact periphrastic in the sense that they are preceded by simple (true) prepositions. Thus, almost all of the nominal prepositions listed above can combine with one of the locative simple prepositions *dar* ‘in, at’, *bar* ‘on’, and *be* ‘to’ in the following order: $P_{\text{LOC}} + P_{\text{nom}}$ ⁸³, and do so appear in formal/literary Persian⁸⁴. However, Sadeghi acknowledges that due to the frequency of use, the simple preposition is commonly dropped.

⁸³ The exceptions are *bedun-e* ‘without’, *barâ(-ye)* ‘for’, *bâvojud-e* ‘in spite of’, and *alâraqm-e* ‘in spite of’. Perhaps the reason is that they already have a preposition in their morphological structure. The nominal preposition *bâvojud-e* can be broken down into a simple preposition *bâ* ‘with’ and the noun *vojud* ‘being, existence’; the form seems to be the result of incorporation. *Alâraqm-e* ‘in spite of’ is a loan word from Arabic. In this language, it consists of the preposition *alâ* meaning ‘on’ and the word *raqm* meaning ‘despite’. The problematic cases are *bedun-e* and *barâ(-ye)*. I presume they are also the result of the combination of a prepositional and a nominal element that have undergone lexicalization in the course of the development of Persian. For *bedun-e*, I can hypothesize that it consists of the locative preposition *be* ‘to’ and the Arabic loan word *dun* ‘without’; but finding out the etymology of *barâ* requires investigation, which I leave for future research. These exceptional cases will not be discussed in this section because the only expression among them that is allowed to appear as a part of a postnominal PP modifier is *bedun-e* ‘without’.

⁸⁴ Note that in formal Persian, the presence of Ezafe on nominal prepositions is always obligatory whether or not a simple preposition is spelled out before them, as shown in (55).

(55) a. ketâb-râ (bar) ru-ye miz gozâšt-am.
 book-OM (on) surface-EZ table put.PST-1SG
 ‘I put the book on the table.’

b. Sinâ dar-râ (be) ru-yam bâz=kard.
 Sina door-OM (to) face-CL.1SG open=did.3SG
 ‘Sina opened the door to me.’

If Sadeghi’s claim is correct that nominal prepositions are periphrastic, then simple locative prepositions ought to co-occur with nominal prepositions in *all* syntactic positions. This is, however, not borne out. Ghomeshi (1997) considers the occurrence of nominal prepositions in subject and direct object positions as evidence that they are nouns. As can be seen in the examples below, the apparent nominal PPs that occur in argument positions cannot co-occur with simple locative prepositions.

(56) a. (*dar) zir-e miz kasif-e.
 (*in) under-EZ table dirty-CL.is
 ‘Under the table is dirty.’

b. (*dar) zir-e miz-o tamiz=kard-am.
 (*in) under-EZ table clean=did-1SG
 ‘I cleaned under the table.’

I conclude from this observation that the constituents containing Ghomeshi's nominal prepositions are noun phrases in argument positions, not PPs. When these phrases occur in adjunct positions (cf. (55)) or as predicates in copular constructions (57), however, they are PPs because they can co-occur with simple locative prepositions. It is only in this environment that the expressions with locative meaning in (53) take up the spatial role—i.e. function as what Svenonius (2006) calls axial parts⁸⁵.

- (57) miz-am (dar) vasat-e otâq bud.
 desk-CL.1SG (in) middle-EZ room was
 'My desk was in the middle of the room.'

Thus, I assume that locative simple prepositions *dar* 'in, at', *bar* 'on', and *be* 'to' that can optionally occur before nominal locative expressions listed in (53), head the adjunct locative PP (PP_{LOC}). The head P_{LOC} selects a noun phrase headed by nominal locative expressions as its complement. In order to establish the constituent structure of the noun

⁸⁵ Liu (2008: 39) emphasizes the importance of syntactic structure in assigning a spatial role to expressions that are spatial in meaning. Here is an excerpt of what she has to say about this:

It is not the case that an expression which is spatial in meaning must correspond to a spatial role in a clause syntactically. For instance, while *Melbourne* is a place name, hence, a spatial expression semantically, in the sentence, 'Melbourne plays a big role in the economy and education of Australia', *Melbourne* does not take up the spatial role but a sort of agentive role, and functions as the subject of the sentence syntactically. It is only in sentences like 'We met each other in Melbourne' that *Melbourne* will take up the spatial role.

The interesting point about Liu's examples is that the spatial role is assigned to *Melbourne* only when it occurs in an adjunct position, as an adverbial, not when it appears in an argument position, as the subject of the sentence. This supports our conclusion about locative expressions in Persian: they take up the spatial role only when they function syntactically as adjuncts.

phrase complement of the head P_{LOC} , we need to explore the relationship between the nominal locative head (e.g. *vasat* ‘middle’) and its dependent (e.g. *otâq* ‘room’ in (57)).

I suggest that a nominal locative expression, as the head of a noun phrase, is a relational noun that refers to a part of an object, or to a specific location with reference to that object. Compare the following examples in which the noun phrase subject is headed by the nominal locative expressions *ru(-ye)* ‘on’ and *zir-e* ‘under’:

- (58) a. (*dar) ru(-ye) miz kasif-e.
 (*in) on(-EZ) table dirty-is
 ‘On the table is dirty.’ (The upper surface of the table is dirty.)
- b. (*dar) zir-e miz kasif-e.
 (*in) under-EZ table dirty-is
 ‘Under the table is dirty.’

The head-dependent relationship in the above relational noun phrases is similar to that in possessive constructions⁸⁶. Specifically, in both relational and possessive noun phrases, the head takes a full-fledged noun phrase as its argument.

⁸⁶ See Barker (2008) for a semantic account of the relationship between possessives and relational nouns.

(59) a. *Noun phrases with locative heads*

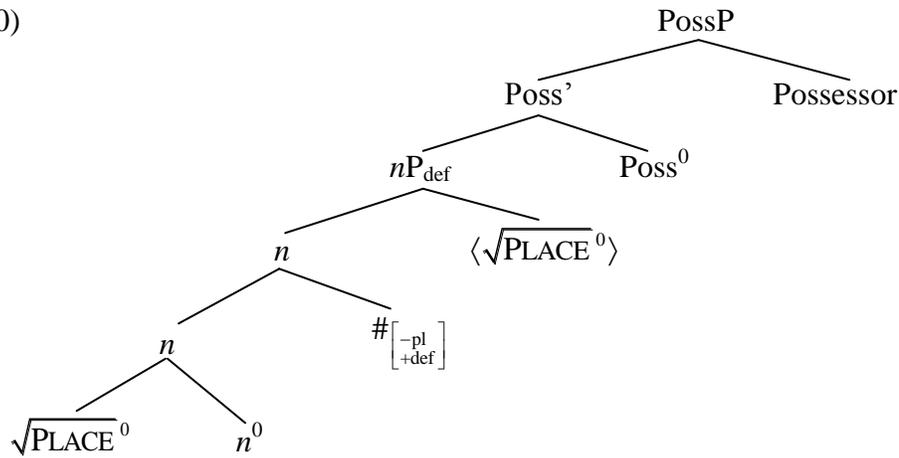
ru(-ye) in miz
 on(-EZ) this table
 ‘on this table’

b. *Possessive construction*

pâye-ye un miz
 leg-EZ that table
 ‘the leg of that table’

Given the argument above, it is reasonable to conclude that the relational noun phrases headed by nominal locative expressions are PossPs.

(60)



where $\sqrt{\text{PLACE}}$ = nominal locative expressions in (53)

As noted earlier, nominal locative expressions embedded in argument PossPs do not induce spatial reading. They take up the spatial role only when they are embedded in PP_{LOC} s that occur in adjunct or predicative positions. From this we conclude that the complement of the PP_{LOC} is a PossP. I propose that the spatial reading induced by these nominal locative expressions is the result of an Agree relation with the P_{LOC} head. The Agree relation is established by selection and c-command between the P_{LOC} head and the

nominal locative head of the nP_{def} embedded in the PossP. The syntactic configuration of the PP_{LOC} and the Agree relation is illustrated in (61)⁸⁷.

⁸⁷ Recall from examples (55) and (57) that the presence of an overt preposition in PP_{LOC} s is optional. The optionality of using simple (true) prepositions before nominal locative expressions in (53) is reminiscent of contexts in which nouns with locational semantics, e.g. lake, mountain, school, park, island, street, etc. take up the spatial role in adjunct (i) or predicative positions (ii) (see Ghomeshi 2008 for a discussion and analysis of bare locative nouns).

- (i) Sinâ raft (be) madrese ~ Sinâ (be) madrese raft.
 Sina went.3SG (to) school
 ‘Sina went to school.’
- (ii) Sinâ (dar) madrese ast.
 Sina (at) school is
 ‘Sina is at school.’

Like nominal locative expressions we have considered so far, nouns with locational semantics do not take up the spatial role when they occur in argument positions, as shown in (iii) and (iv).

- (iii) madrese naqš-e bozorg-i dar tarbiat-e bačče-hâ-mun dâre.
 school role-EZ big-IND in education-EZ child-PL-CL.1PL has
 ‘Schools have a major role in educating our children.’
- (iv) Sinâ madresa-ro dust=dâre.
 Sina school-OM friend-has
 ‘Sina likes school.’

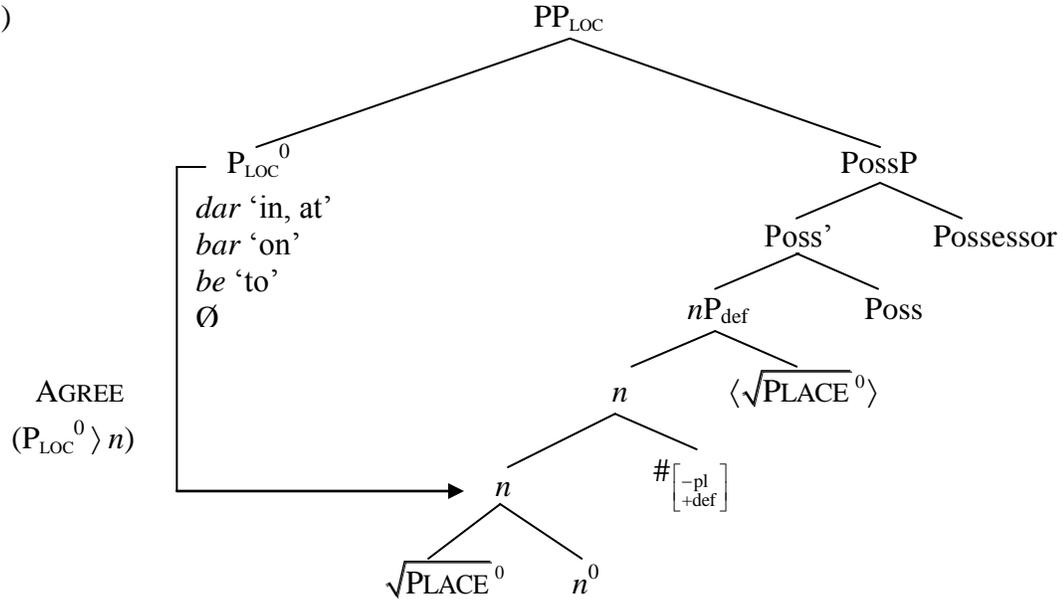
I propose that the Root $\sqrt{\text{PLACE}}$ in (60) and (61) also stand for nouns with locational semantics. However, they behave like other nouns in that they may appear in nPs , ModPs , or PossPs in argument positions. In adjunct positions, any of these categories containing nouns with locational semantics can be the complement of P_{LOC} .

That nominal locative expressions in (53) (Ghomeshi’s (1997) nominal prepositions) can stand alone in Persian is a fact that Ghomeshi (1997) and Larson & Yamakido (2005) mention to argue that they are nouns.

- (v) ketâb-â-ro bord-am tu.
 book-PL-OM took-1SG in
 ‘I took the books in.’

Within my approach, I use this fact to argue that these expressions and nouns with locational semantics belong to the same class. Since further discussion of this issue is tangential to the main concerns of this chapter, I leave it for future work.

(61)



Having proposed and discussed the phrasal categories and syntactic structures in which the nominal locative expressions in (53) occur, we now consider which one of the two possible phrasal categories is a phrasal modifier in the Persian noun phrase: the PossP or the PP_{LOC} . I suggest that only PP_{LOC} s can serve as phrasal modifiers because, given that phrasal modifiers are positioned in the specifier of ModP, they cannot be arguments. Moreover, that Ghomeshi's nominal prepositions are all construed as inducing spatial reading further supports my proposal that the phrasal modifier containing nominal locative expressions is a PP_{LOC} . Consider the following examples:

- (62) a. livân-e (*bar) ru-ye miz b. vilâ-ye (*dar) kenâr-e daryâ
glass-EZ (on) surface-EZ table cottage-EZ (in) beside-EZ sea
'the glass on the table' 'the cottage beside the sea'

Notice that unlike the examples of PP_{LocS} we have discussed so far, the nominal locative expressions in the above examples cannot be preceded by true prepositions. This presumably indicates that the head of PP_{LocS} has a null exponent when they appear as phrasal modifiers. PP_{LocS} in (62) are interpreted as reduced relative clauses. When they appear in their corresponding relative clauses, they are in predicative positions and true prepositions can occur with them.

(63) a. livân-i ke (bar) ru-ye miz as
 glass-REL that (on) surface-EZ table is
 ‘the glass that is on the table’

b. vilâ-i ke (dar) kenâr-e daryâ-st
 cottage-REL that (in) beside-EZ sea-CL.is
 ‘the cottage that is beside the sea’

What makes argument PossPs and adjunct PP_{LocS} similar to each other, however, is that in both of them nominal heads with locative meanings are connected to their dependents via the Ezafe vowel. In the PF derivation of both phrases, the presence of a nominal head followed by a [+N] M-Word dependent (i.e. the possessor) provides the environment for the application of EIR. As a review, the conditions on inserting Ezafe are repeated in (64):

(64) *Ezafe Insertion Rule*

Insert the vowel *-e* on an overt [+N] M-Word X when followed by (an) overt [+N] M-Word(s) Y if

a. X and Y are category-defining heads *n* or *a*; and

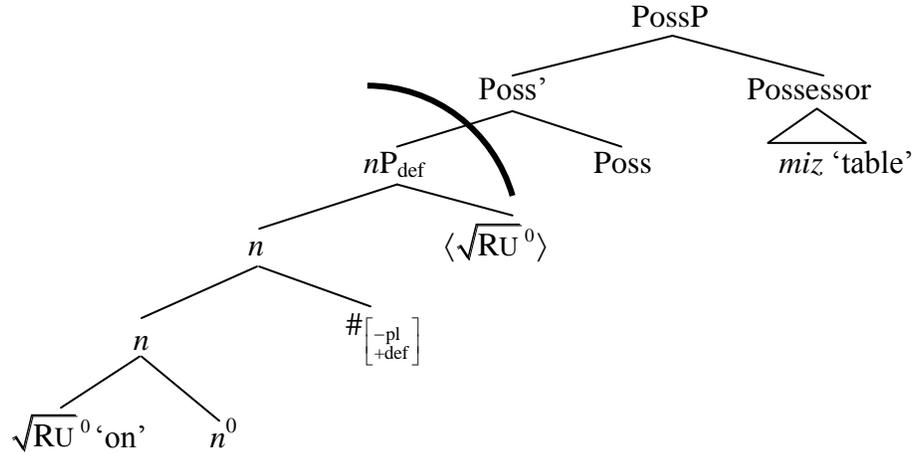
b. Y is a dependent of X,

such that the vowel *-e* is spread rightward from X onto each instantiation of Y, stopping short of the last word.

In deriving argument PossPs, EIR takes the PossP as its domain of application. The nominal locative expressions that head the embedded *nP* trigger the insertion of Ezafe on them when followed by [+N] M-Words inside the possessor *nP*. EIR will then spread the Ezafe vowel from the head onto each of the following dependents that meet condition (64)a. Recall that EIR, as a Late-Linearization rule, applies when PF has computed all spellout domains. In the case of adjunct PP_{LOC} , EIR applies when the spellout domain of the root node (PP_{LOC}) has been computed. The rule takes the PP_{LOC} as its domain. The nominal locative expression *ru* ‘on’ triggers the insertion of Ezafe. Ezafe is thus inserted on the nominal head and spread onto the following [+N] M-Words that meet the conditions of Ezafe Insertion Rule.

Let us now begin the analysis of Ezafe insertion in argument PossPs. The derivation of argument PossPs involves two cycles: (i) the *nP* phase and (ii) the root node PossP. Taking the PossP *ru-ye miz* ‘on the table’ in (58)a as an example, I assume the following syntactic structure is the input to PF:

(65)



PF computes linearization statements for the nP phase in the manner described in §5.3.3.1. The spellout domain of the nP_{def} consists of only one overt M-Word (n), which represents the head noun *ru* 'on'. Thus the output of PF derivation for the nP_{def} is $n[\text{ru}]$ at Chaining.

In the next cycle of the derivation, PF produces linearization statements for the root node PossP. The only overt M-Word in the spellout domain of this phrase is the head of the Possessor $n[\text{miz}]$. We now need to establish linear order between the spellout domains of the nP_{def} and the PossP. This is done at the Late-Linearization stage where separate spellout domains are chained together. The linearization statement produced at this stage will have the spellout domains of the nP_{def} and the PossP clustered together such that the M-Word $n[\text{ru}]$ in the nP_{def} is followed by the M-Word $n[\text{miz}]$ in the PossP. This characterizes the structural requirement for the application of EIR. When EIR applies, the two chained spellout domains are merged together. This Chain-merging forms the Ezafe insertion domain, where the Ezafe vowel is inserted between the two linearized [+N] M-Words that are in head-dependent relationship. The steps for the Late-Linearization of these two spellout domains are illustrated in (66):

(66) PF

- i. Late-Linearization: $(n[ru])-(n[miz])$
- ii. EIR: $(n[ru]-e-n[miz])$
- iii. Surface PF: ru-ye miz

As we saw in (59), possessor *n*Ps are full-fledged noun phrases that allow all pre- and postnominal modifiers to occur in them. Examples of argument PossPs with such noun phrases, particularly those with prenominal modifiers, provide further evidence in support of EIR as being sensitive not to the linear adjacency but to the morphosyntactic feature of the M-Words that follow nominal heads *n* and *a* and are in head-dependent relationship with them (see the discussion in §5.3.3.1 for a relevant case in (42)). Consider the following example of an argument PossP in which the possessor *n*P contains prenominal modifiers (parentheses mark the Ezafe domain here):

- (67) (zir-e in se miz)-o tamiz=kon-id.
under-EZ this three table-OM clean=do.IMP.2PL
'Clean under these three tables!'

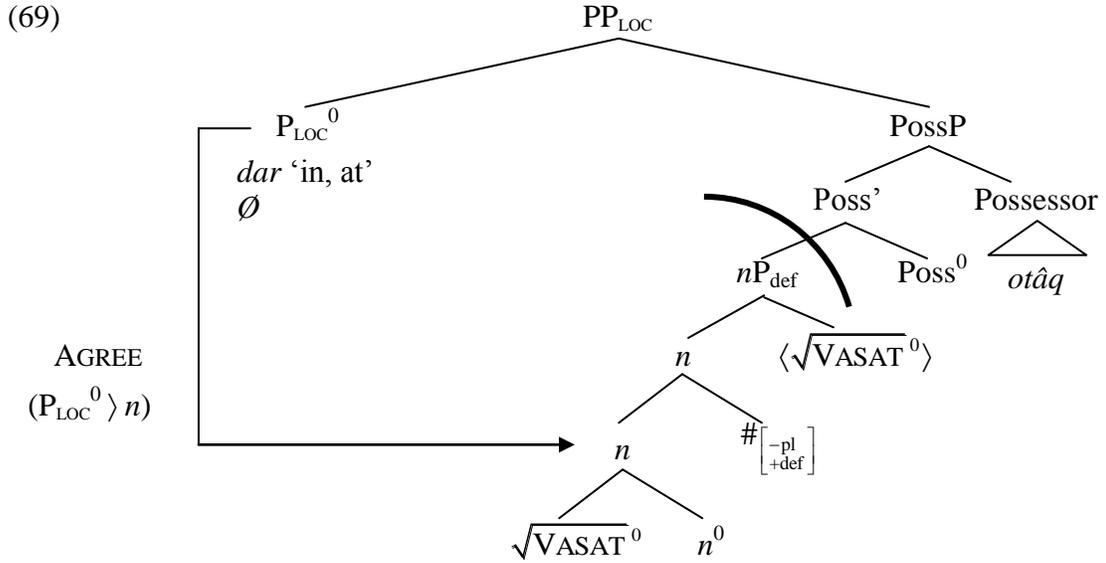
The PF derivation of the parenthesized argument PossP in (67) proceeds in the same manner described for (65). At the final stage of the derivation, Late-Linearization chains together the *n*P spellout domain containing the M-Word *n*[zir] and the PossP spellout domain containing the M-Words in the possessor *n*P, as shown in (68)i. Like intensifiers, prenominal modifiers are not [+N] M-Words of categories *n* and *a* and therefore do not

trigger Ezafe insertion on the head of the *nP zir* ‘under’. Nevertheless, Ezafe insertion on the head noun is licensed due to the presence of the head of the possessor *nP (miz* ‘table’) at the end of the Ezafe insertion domain. The visibility of this [+N] M-Word is ensured by Chain-merging that occurs as a consequence of applying EIR. Inserting Ezafe on the head noun *n[zir]*, EIR spreads it rightward in the Ezafe insertion domain onto only M-Words that satisfy the required categorial conditions.

(68) PF

- i. Late-Linearization: (*n[zir]*)-(Dem[in]-Card[se]-*n[miz]*)
- ii. EIR: (*n[zir]*-e-Dem[in]-Card[se]-*n[miz]*)
- iii. Surface PF: *zir-e in se miz*

Turning to the adjunct PP_{LocS} , the PF derivation involves two cycles, i.e. (i) the nP_{def} and (ii) the root node PP_{Loc} . Taking the adjunct PP_{Loc} (*dar vasat-e otâq* ‘in the middle of the room’ in (57) as an example, I suggest the following syntactic structure as the input to PF:



PF spells out the M-Word $n[\text{vasat}]$ ‘middle’ as the only content of the nP_{def} cycle. The head P_{LOC} gets spelled out with the PossP in the last cycle where the left-over content under the root node PP_{LOC} is spelled out. Note that prior to spell-out, the Agree relation between P_{LOC} and the nominal locative head of the nP_{def} $n[\text{vasat}]$ is established at syntax. At Late-Linearization, the nP_{def} and PP_{LOC} spellout domains are chained together. However, the M-Words in the PP_{LOC} spellout domain are chained to the nP_{def} spellout domain on two different sides. If the head P_{LOC} is phonologically realized, it is chained to the nP_{def} spellout domain on the left edge. The PossP, however, is chained to the nP_{def} spellout domain on the right edge as a dependent of the head noun *vasat*. This causes each of these constituents to form their own domains. The head P_{LOC} forms a separate domain from that of the nP_{def} because it is not a [+N] M-Word and thus cannot group together with the head noun in the Ezafe domain. The PossP also forms a separate domain because it is a domain where only phrasal modifiers of the noun can appear—recall that head modifiers can only appear in the nP (inner) domain. When EIR

applies, the nP_{def} spellout domain merges with the PossP (outer) domain. As noted earlier, this Chain-merging takes place uniformly in all such cases where a ModP or a PossP follows the nP . The purpose of the Chain-merging is to establish relationship between the head noun and its following phrasal modifiers, the Possessor in this case. I propose that the Ezafe vowel is the marker of this head-dependent relationship. The derivational process of the above adjunct PP_{LOC} at the Late-Linearization stage is schematically illustrated below:

(70) PF

- i. Late-Linearization: $(P_{\text{LOC}}[\text{dar}]-(n[\text{vasat}]-(n[\text{otâq}])))$
- ii. EIR: $(P_{\text{LOC}}[\text{dar}]-(n[\text{vasat}]-e-n[\text{otâq}])))$
- iii. Surface PF: dar vasat-e otâq

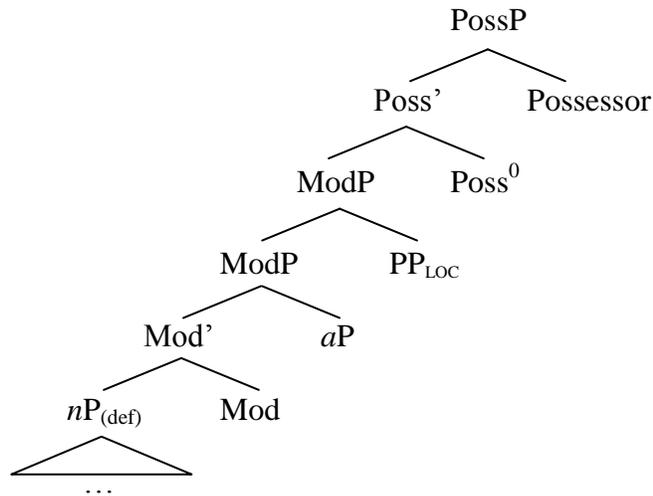
In this subsection, we argued that what Ghomeshi (1997) calls nominal prepositions are in fact nominal locative heads of embedded noun phrases in PossPs. We showed that such PossPs appear only in argument positions. In adjunct or predicative positions, however, the PossP appears as the complement of PP_{LOC} s. These nominal elements with locative meanings are therefore [+N] M-Words and should be connected with their following [+N] M-Words via the Ezafe vowel. Let us now consider how the Ezafe vowel is inserted between the head noun and its following phrasal modifiers— aP , PP_{LOC} , and possessive nP_{def} —in Persian noun phrases.

5.3.4. EIR in Persian noun phrases: the complete picture

So far we have established how EIR applies to constituents internal to the *nP* and those internal to the phrasal modifiers following it. In both cases EIR applies to a [+N] M-Word head and spreads the Ezafe vowel onto its following [+N] M-Word dependents except the last one. As such, we have introduced the Ezafe vowel as the marker of head-dependent relationship in Persian. In this section, we get down to explaining (i) how the Ezafe vowel is inserted between the *nP* and its following phrasal modifiers, and (ii) why Ezafe cannot appear between prenominal modifiers and the noun.

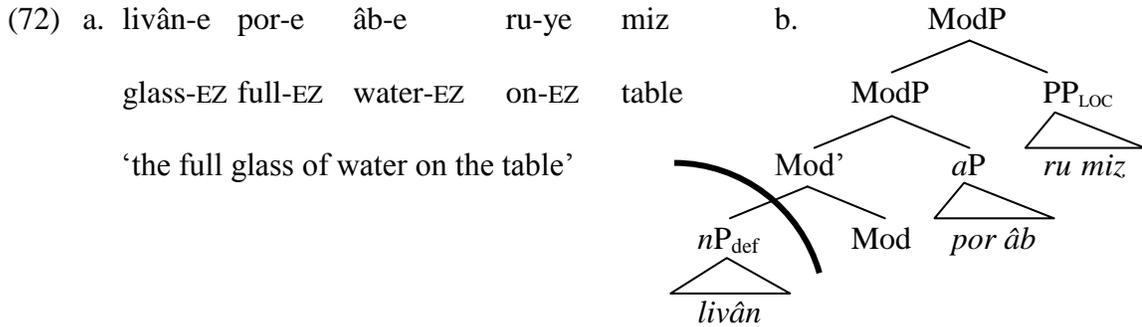
Recall from §5.1.2 that noun phrases with phrasal modifiers, excluding the possessor *nP_{def}*, appear as ModPs in the Ezafe domain and noun phrases with possessors occur as PossPs.

(71) *Phrasal modifiers in the Ezafe domain*



Recall that phrasal modifiers are base-generated in the specifiers of ModPs and PossPs. These projections define a domain for the phrasal modifiers (the outer domain)

distinct from that of the head modifiers inside the *nP* (the inner domain). For this reason, we assume that the *nP* is a phase and its content is spelled out separately. Let us begin with an example of a noun phrase which contains phrasal modifiers.

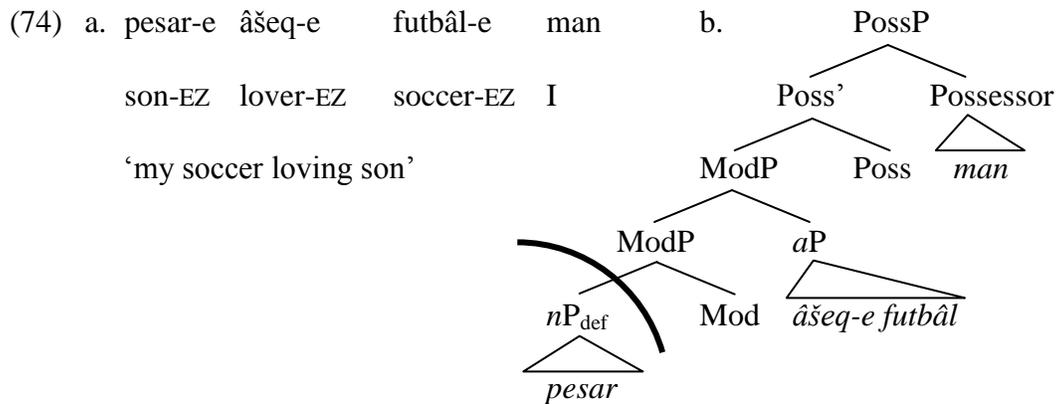


In the *nP* phase, the overt M-Word *n*[livân] constitutes the content of the spellout domain of the *nP*_{def}. In the final stage of the derivation, the phrasal modifiers in the specifier of the ModP are spelled out. At the Late-Linearization stage, the *nP*_{def} (inner) domain is chained together with the ModP (outer) domain. In order to establish head-dependent relationship between the head noun and its phrasal modifiers, EIR applies and merges the two domains into a single domain for Ezafe insertion. Thus, Ezafe is inserted on the head noun *livân* ‘glass’ and extended over the following dependents in the manner stipulated in Ezafe Insertion Rule (cf. (64)) and shown in (73):

(73) PF

- i. Late-Linearization: (*n*[livân])-(*a*[por]-*n*[âb]-*n*[ru]-*n*[miz])
- ii. EIR: (*n*[livân]-*e*-*a*[por]-*e*-*n*[âb]-*e*-*n*[ru]-*e*-*n*[miz])
- iii. Surface PF: livân-e por-e âb-e ru-ye miz

Now we consider an example of a PossP in which the *nP* is connected to its following phrasal modifiers via *Ezafe*. I should note that in PossPs, the phrasal modifier PP cannot co-occur with the Possessor⁸⁸.



⁸⁸ We established earlier in this chapter that the PP that appears in Persian noun phrases as a phrasal modifier is an adjunct locative PP (PP_{Loc}) in which the head takes a PossP as its complement. I suggest that PP_{Loc}s and PossPs in Persian are not allowed to co-occur in a noun phrase because the presence of two possessors in a row creates parsing problems. For instance, in

- (i) *ketâb-e ru-ye miz-e man*
 book-EZ on-EZ table-EZ I

miz and *man* are both possessive modifiers: *miz* is the possessor inside the PP_{Loc} *ru-ye miz* ‘on the table’ and *man* ‘I’ is the possessor of the whole noun phrase *ketâb-e ru-ye miz* ‘the book on the table’. Nevertheless, *man* ‘I’ is understood as being the possessor of *miz* ‘table’. Thus, the noun phrase in (i) is most naturally translated as *the book on my table* and not *my book on the table*.

The constraint on having two consecutive possessors in the Persian noun phrase can also be generalized to *aP*s containing PossP complements.

- (ii) *[*pesar-e aP[âšeq-e še’r-â-ye šâmlu]-ye man*
 boy-EZ lover-EZ poem-PL-EZ Shamlu-EZ I

As can be seen in (ii), the embedded possessor *šâmlu* inside the *aP*, and the noun phrase possessor *man* ‘I’ have occurred one after the other. The consecutive occurrence of the possessors in (ii) results in ungrammaticality because no plausible interpretation is available for this noun phrase.

The constraint banning two consecutive possessors provides an alternate explanation for Ghomeshi’s (1997) proposal that there is only one structural position for possessors, hence no Persian equivalent to the English *the enemy’s destruction of the city*.

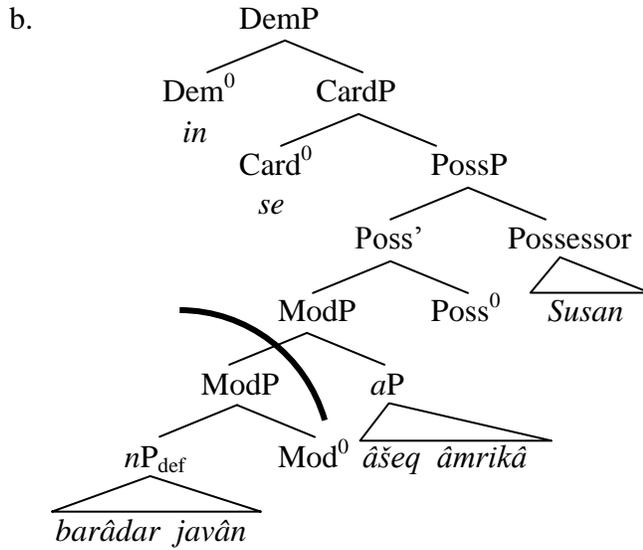
The same derivational procedure as (73) is involved in linking separate spellout domains in the PossP via the Ezafe vowel. I suffice to schematically illustrate the PF derivation in (75):

(75) PF

- i. Late-Linearization: $(n[\text{pesar}])-(a[\hat{\text{a}}\text{šeq}]-n[\text{futbâl}]-n[\text{man}])$
- ii. EIR: $(n[\text{pesar}]-e-a[\hat{\text{a}}\text{šeq}]-e-n[\text{futbâl}]-e-n[\text{man}])$
- iii. Surface PF: $\text{pesar-e } \hat{\text{a}}\text{šeq-e futbâl-e man}$

ModPs and PossPs in Persian noun phrases appear as complements to functional projections housing pronominal modifiers (e.g. demonstratives, quantifiers, superlatives, numerals). However, pronominal modifiers are not linked to the noun via the Ezafe vowel. Consider the following example and its syntactic representation:

- (76) a. in se barâdar-e javân-e âšeq-e âmrikâ-ye susan
 this three brother-EZ young-EZ lover-EZ America-EZ Susan
 ‘these three US-loving young brothers of Susan’s’



The immediate reason for the absence of Ezafe on prenominal modifiers is that they are not [+N] M-Words, given EIR. The other reason is suggested in Matushansky (2008: §9.5), where she accounts for why superlatives are the only Persian adjectives that appear prenominally. She proposes a theory under which superlatives behave like universal quantifiers that take the noun phrase they modify as their arguments. The argument, according to her, corresponds to the restrictor of the universal quantifier or the comparison set (see the original source for complete details). Based on this analysis, she generalizes that the prenominal position in Persian is reserved for quantifiers and that Ezafe marks the modification relation. She concludes that the absence of Ezafe on prenominal modifiers is explained by the semantic distinction between quantification and modification. I adopt her view on assuming that Persian noun phrases are at large divided into domains of quantification and modification and that Ezafe only appears in the domain of modification to connect the head noun to its following dependents.

Given EIR, the Ezafe vowel must be inserted on the *overt* head noun. The question that arises is how Ezafe insertion takes place in the Empty Noun Construction, where the head noun is not phonologically realized. We will pursue this issue in the next section.

5.4. Ezafe insertion in the ENC

Recall from Chapter 3 that the following assumptions were proposed to account for licensing the ENC in Persian:

(77) *Assumptions*

a. *Focus condition on the ENC in Persian*

A nominal subconstituent α can be left unpronounced in constituent β iff β has a salient antecedent δ that is mentioned or implied in the previous text or discourse, such that

- (iii) α is in identity-of-sense relationship with the corresponding subpart γ of δ ; and
- (iv) The remnant of β is non-identical to the corresponding part λ of δ .

b. *Licensing Condition on the empty noun in Persian*

The empty noun $\sqrt{\text{ROOT}_{[E]}}$ is licensed by an F-specified remnant modifier r_F in a phase-based manner, where the phase contains r_F .

Based on these assumptions, the licensing of the empty noun in Persian is due to the interaction between the properties of the E-feature specified on the nominal Root and those of the focus feature F specified on the remnant modifier (i.e. an attributive adjective). This interaction is carried out in a phase-based manner such that the empty noun is licensed by an F-specified modifier within the same cycle of derivation (the *nP* phase). In the absence of a remnant modifier in the *nP* domain, the empty noun is licensed by an inherently F-specified prenominal modifier in the next higher structure. In this case, the phase in which the empty noun is licensed will be the whole noun phrase. As discussed in Chapter 3, when all the required conditions for licensing the empty noun are met, the derivation is sent off to the interfaces for spellout. At PF, where the syntactic structure is linearized at Vocabulary Insertion, the E-feature specification on the nominal Root determines its null exponent. Since Vocabulary Insertion triggers Linearization (Embick & Noyer 2001), I assume it applies early in the PF derivation. Consequently, EIR, as a Late-Linearization rule, applies when the null exponent of the empty noun has already been assigned. Let us explore this further by considering the derivation of an *nP* with the empty noun and multiple head modifiers (78)B.

(78) A: Setâre çe jurâb-i mi-puš-e?

Setare what stocking-IND DUR-wear-3SG

‘What kind of stockings does Setare wear?’

(80) PF

- i. Concatenation: $n[\text{jurâb}] \frown a[\text{meški}]$
 $a[\text{meški}] \frown a[\text{nâzok}]$
 $a[\text{nâzok}] \frown a[\text{boland}]$
- ii. Chaining: $n[\text{jurâb}]-a[\text{meški}]-a[\text{nâzok}]-a[\text{boland}]$

At Late-Linearization, EIR takes the output of Chaining as its domain. The rule cannot insert the Ezafe vowel onto the head noun because it is not phonologically realized. This however does not affect the head-dependent relationship between the null head noun and its following modifiers. The Ezafe vowel is inserted on the following [+N] overt modifier and spread rightward onto other dependents except the last one in the modification domain. This implies that Ezafe insertion in the ENC is not only Vocabulary sensitive but also sensitive to the head-dependent relationship between the elements inside the modification domain. The Late-Linearization level of the PF derivation is shown in (81):

- (81) i. Late-Linearization: $n[\text{jurâb}]-a[\text{meški}]-a[\text{nâzok}]-a[\text{boland}]$
- ii. EIR: $n[\text{jurâb}]-a[\text{meški}]-e-a[\text{nâzok}]-e-a[\text{boland}]$
- iii. Surface PF: $\text{meški-e nâzok-e boland}$

EIR applies in a similar manner in cases in which the head modifiers are followed by phrasal modifiers. The only difference in such cases is that the derivation involves two cycles: the *nP* cycle and the cycle in which the root node is computed. Consider the following example:

- (82) Sinâ nP (~~kif~~ čarm-i-e meški)-e $ModP$ (pošt-e vitrin)-o
 Sina ~~bag~~ leather-ADJ-EZ black-EZ behind-EZ windowshop-OM
 xarid.
 bought.3SG
 ‘Sina bought the black leather [one] behind the window shop.’

The PF derivation of (82) yields the nP spellout domain and the $ModP$ spellout domain. At Late-Linearization stage, the Ezafe insertion (modification) domain is formed by merging the two separate spellout domains. As noted earlier, in the ENC, the Ezafe vowel is inserted on the overt modifier following the null noun and is extended onto the following dependents save the last one.

- (83) PF
- i. Late-Linearization: $(n[\del{kif}] - a[\check{c}armi] - a[meški]) - (n[pošt] - n[vitrin])$
 - ii. EIR: $(n[\del{kif}] - a[\check{c}armi] - e - a[meški] - e - n[pošt] - e - n[vitrin])$
 - iii. Surface PF: čarmi-e meški-e pošt-e vitrin

So far we have established how Ezafe is inserted in the ENCs with head and phrasal modifiers. In all the examples considered, the empty noun is licensed by an F-specified head modifier (i.e. attributive adjective) inside the nP phase. We have seen that in all such cases, the absence of an overt head noun results in Ezafe being inserted on the immediately following [+N] M-Word and spread rightward onto other following [+N] M-Words—representing head modifiers or elements inside phrasal modifiers—except the

last one. Let us now consider Ezafe insertion in the ENC_s with prenominal modifiers. Three possibilities are considered here: (i) when there is a prenominal and a postnominal head modifier (84); (ii) when there is a prenominal and postnominal head and phrasal modifiers (85); and (iii) when there is a prenominal and a postnominal phrasal modifier, as in (86).

(84) se-tâ kif meški-e sâde
 three-CLS ~~bag~~ black-EZ plain
 ‘three plain black [ones]’

(85) kučik-tar-in kif čarm-i-e tu maqâze
 small-COMP-SUP ~~bag~~ leather-ADJ-EZ in shop
 ‘the smallest leather [one] in the shop’

(86) do-tâ-ye ~~ketâb~~ Sinâ
 two-CLS-EZ ~~book~~ Sina
 ‘the two of Sina’

Recall from Chapter 3 (§3.3.3) that when in ENC_s like (84) and (85), the empty noun is licensed by a head modifier in the *n*P phase, the phase is complete and thus sent off to the interfaces for semantic and phonological interpretation. In such contexts, the presence of a prenominal modifier does not play any role in the process of licensing the empty noun. Accordingly, EIR applies in the PF derivation of (84) and (85) in the manner described

above for the ENC's in (78)B and (82). In the derivation of (84), two domains are formed at Late-Linearization: the domain of quantification followed by the *nP* spellout domain. EIR applies only to the *nP* spellout domain as it constitutes the domain of modification in this example.

- (87) (se-tâ) (kif meški sâde)
 three-CLS bag black plain
 ‘three plain black [ones]’
 → se-tâ meški-e sâde

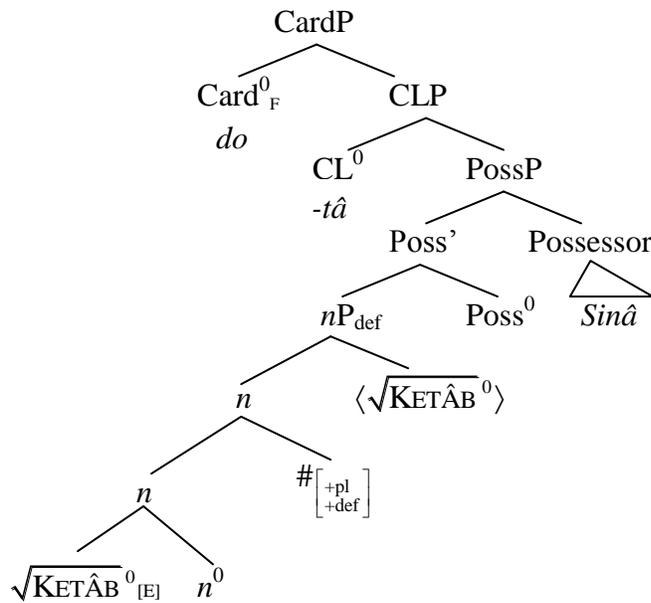
In (85), however, three domains are formed at Late-Linearization: the domain of quantification; the domain of head modifiers (*nP*); and the domain of phrasal modifiers (ModP). The application of EIR results in the merger of the head and phrasal modifier domains, thus forming the domain of modification or Ezafe insertion domain.

- (88) (kučik-tar-in) (kif čarm-i tu maqâze)
 small-COMP-SUP bag leather-ADJ in shop
 ‘the smallest leather [one] in the shop’
 → kučiktarin čarmi-e tu(-ye) maqâze

We have seen so far that the presence of a prenominal modifier in the ENC is insignificant in licensing the empty noun when there is an empty noun licenser inside the *nP* domain. It also does not affect the application of EIR. The prenominal modifiers in

both (84) and (85) are in the domain of quantification and thus do not bear the Ezafe vowel, as predicted. This is however not the case in (86) where the numeral+classifier has been linked to the possessor noun phrase *Sina* via Ezafe. In this context, the empty noun is licensed by the numeral+classifier because there is no licenser inside the *nP* domain. As I suggested in Chapter 3, the derivation of the ENC is sent to the interfaces after the empty noun is licensed. Since the empty noun is licensed by an element outside the *nP* domain (i.e. numeral+classifier), I assume that the syntactic derivation of (86), represented in (89), is sent off to the interfaces as a whole.

(89)



Given our assumptions in (77), the empty noun is identified through the interaction between the E-feature of the nominal Root *ketâb* ‘book’ and the inherent F-feature of the numeral+classifier *do-tâ* ‘two-CLS’. As a result of this interaction, I suggest, the numeral+classifier inherits the [+N] feature of the head noun. Under similar circumstances where the empty noun is licensed by other inherently F-specified

prenominal modifiers, I propose that the prenominal licenser would likewise inherit the [+N] feature from the head noun.

At PF, Linearization produces Concatenation statements for the overt M-Words in (89). In the first statement, the M-Word Card^0 is concatenated with the M-Word CL: $\text{Card}[\text{do}] \frown \text{CL}[-\text{t}\hat{a}]$. The default classifier $-\text{t}\hat{a}$ is an affix and is required to be attached to numerals. This triggers a string-vacuous Local Dislocation in which the two M-Words become one and $\text{CL}[-\text{t}\hat{a}]$ becomes a Sub-Word: $\text{Card}[\text{do}] \oplus \text{CL}[-\text{t}\hat{a}]$. The output of the LD is thus an M-Word that itself concatenates with the next overt M-Word $n[\text{Sin}\hat{a}]$: $(\text{Card}[\text{do}] \oplus \text{CL}[-\text{t}\hat{a}]) \frown n[\text{Sin}\hat{a}]$. When Concatenation is complete, the M-Words are clustered together at Chaining: $(\text{Card}[\text{do}] \oplus \text{CL}[-\text{t}\hat{a}]) - n[\text{Sin}\hat{a}]$. At Late-Linearization, the output of Chaining is the input to EIR. Given that the M-Word $\text{do-t}\hat{a}$ ‘two-CLS’ has inherited the [+N] feature of the empty noun as a result of the interaction between its F-feature and the E-feature of the empty noun, I suggest that it groups together with the possessor $\text{Sin}\hat{a}$ and forms the Ezafe insertion (modification) domain. As a result, EIR inserts the Ezafe vowel on the numeral+classifier $\text{do-t}\hat{a}$ and connects it to the possessor $\text{Sin}\hat{a}$: $\text{do-t}\hat{a}\text{-ye Sin}\hat{a}$.

5.4.1 *-h* \hat{a} mobility and Ezafe insertion

In Chapter 4 and the current chapter we have explored and elaborated on *-h* \hat{a} mobility and Ezafe insertion as two PF operations that apply in the linearization process. Within Pak’s (2008) model, the LD operation that handles *-h* \hat{a} mobility would be an *early*

morphophonological rule because it targets binary Concatenation statements⁸⁹ as its domain and operates in terms of linear adjacency. EIR, however, targets larger domains, as it applies either to a string of M-Words (Chain) inside a given spell-out domain or to chains of M-Words across spell-out domains. As such, its application occurs *late* in the derivation. Hence, it is a Late-Linearization rule and, as predicted in Pak’s model, follows the LD rule responsible for *-hâ* mobility.

Having established that LD and EIR belong to two distinct levels of PF derivation, we now take a look at an example of *-hâ* mobility to see how the ENC is derived at PF.

(90) Irân az fa’âl-tar-in kešvar-hâ-ye âsiâ-st.

Iran of active-COMP-SUP country-PL_{DEF}-EZ Asia-CL.is

‘Iran is among the most active countries of Asia.’

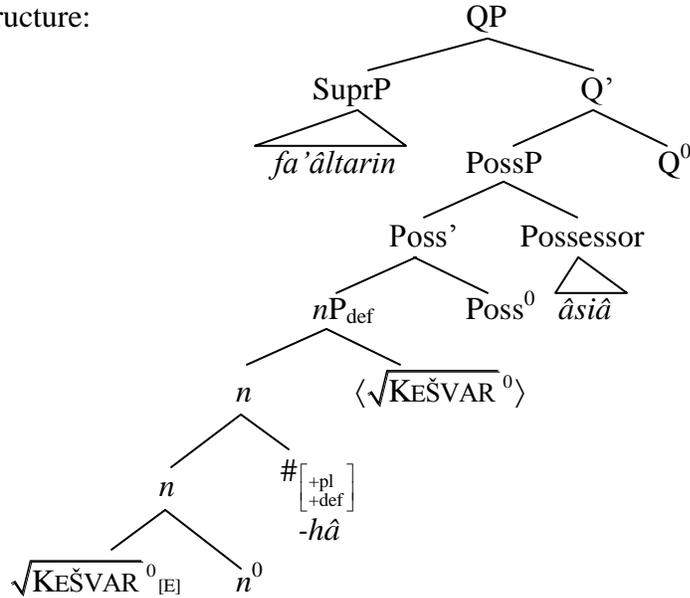
→ Irân az _{QP}[fa’âltarin]-***hâ-ye*** _{PossP}[_{nP}[kešvar<*-hâ*>]] _{nP}[âsiâ]]-st.

‘Iran is among the most active [ones] of Asia.’

I assume the following syntactic structure as the input to PF:

⁸⁹ Recall that Concatenation statements comprise of two M-Words, being generated within a given spell-out domain.

(91) Structure:



The derivation of (91) involves only one cycle because the empty noun is not licensed within the nP domain. As noted earlier in (89), in cases where the empty noun is not licensed inside the nP domain, the spellout of the nP is delayed until the empty noun is licensed in the higher structure. Accordingly, the whole noun phrase in (91) is spelled out together as one cycle. As such, the concatenation of the prenominal modifier *fa'altarin* and the stranded *-hâ* is made possible at PF Linearization. This provides the structural description for LD to apply and string-vacuously affix *-hâ* onto *fa'altarin*. Concatenation applies again and linearizes the output of LD with the possessor *âsiâ* 'Asia'. Chaining strings these M-Words together. As there is one spellout domain involved in this derivation, the output of Chaining will be the input to Late-Linearization. At Late-Linearization, two domains are formed: (i) the prenominal modifier *fa'altarin* that hosts the plural marker *-hâ* at this stage is part of the nP domain. As noted earlier, the interaction between the E-feature specified on the nominalized Root and the inherent F-feature of the prenominal licenser of the empty noun results in the prenominal modifier to

inherit the [+N] feature of the missing noun. (ii) The possessor *âsiâ* ‘Asia’ forms the domain of phrasal modifiers. The application of EIR at this point merges the two domains and forms the modification domain, where [+N] M-Words that modify the head noun are linked together via the Ezafe vowel. The steps of the PF derivation described above are schematized in (92):

(92) PF

- i. Concatenation: $\text{Supr}_F[\text{fa}'\hat{\text{a}}\text{ltarin}] \sim (\sqrt{\text{KEŠVAR}}^0_{[\text{E}]} \oplus n^0 \oplus -h\hat{\text{a}})$
- ii. LD: $\text{Supr}_F[\text{fa}'\hat{\text{a}}\text{ltarin}] \oplus -h\hat{\text{a}}$
- iii. Concatenation: $\text{Supr}_F[\text{fa}'\hat{\text{a}}\text{ltarin}] \oplus -h\hat{\text{a}} \sim n[\hat{\text{a}}\text{si}\hat{\text{a}}]$
- iv. Chaining: $\text{Supr}_F[\text{fa}'\hat{\text{a}}\text{ltarin}] \oplus -h\hat{\text{a}} - n[\hat{\text{a}}\text{si}\hat{\text{a}}]$
- v. Late-Linearization: $(\text{Supr}_F[\text{fa}'\hat{\text{a}}\text{ltarin}] \oplus -h\hat{\text{a}}) - (n[\hat{\text{a}}\text{si}\hat{\text{a}}])$
- vi. EIR: $(\text{Supr}_F[\text{fa}'\hat{\text{a}}\text{ltarin}] \oplus -h\hat{\text{a}} - e - n[\hat{\text{a}}\text{si}\hat{\text{a}}])$
- vii. Surface PF: $\text{fa}'\hat{\text{a}}\text{ltarin} - h\hat{\text{a}} - \text{ye} \quad \hat{\text{a}}\text{si}\hat{\text{a}}$

In this section, we discussed the application of EIR in the ENC, both in those with head and phrasal modifiers and those with pronominal modifiers. We then considered the interaction between EIR and the LD operation responsible for *-hâ* mobility in the ENC and showed that these rules apply at distinct levels of PF derivation and thus follow strict ordering.

5.5. Summary

In this chapter, I have presented a novel analysis of the Ezafe construction in Persian and discussed the consequences of this analysis for the ENC. From a broad perspective, my analysis addresses two main aspects of the Ezafe construction: (i) the categorial and structural status of postnominal modifiers; and (ii) Ezafe insertion. Building on my analyses of licensing empty nouns and *-hâ* mobility in Chapters 3 and 4, I have presented evidence from compounding and syntactic compounding in Persian which supports my proposal that postnominal modifiers are split in terms of both category and structural positions. Adopting Ghomeshi's (1997) view that Ezafe is a linking element inserted at PF, I have proposed that Ezafe insertion is determined at the interface of syntax and phonology. Within Pak's (2008) model of syntax-phonology interface, I have argued that Ezafe insertion rule (EIR) is a phonological rule that applies at Late-Linearization stage at PF. The rule inserts the Ezafe vowel onto an overt [+N] head and spreads it rightward onto the following [+N] M-Words that are modifiers/complements of the nominal head. In the final part of this chapter, I have argued that in the PF derivation of ENCs, EIR and the Local Dislocation (LD) operation responsible for *-hâ* mobility follow strict ordering such that LD applies early at Concatenation but EIR at Late-Linearization.

Chapter 6

Conclusions and Future Work

6.0. Introduction

In this thesis, we have concentrated on presenting an account of the empty noun in Persian, its licensing conditions, and its interaction with nominal modifiers, the plural marker, and the Ezafe linker. Proposals were made to address, among other related issues, the following fundamental research questions introduced in Chapter 1: (i) How does the interaction between having antecedents and the presence of a nominal modifier play a role in the licensing process of empty nouns in Persian? (ii) How does the non-pronunciation of the head noun affect the distribution of plural marking and what is its consequence for the syntax of the Persian noun phrase? (iii) What is the explanation for the behavior of the Ezafe vowel in the context of the Empty Noun Construction (ENC) and how does it contribute to our overall understanding of the nature and function of Ezafe in the Persian noun phrase? In this final chapter, I will briefly review some of these proposals and their implications, and make a few suggestions for future avenues of research.

6.1. Summary of proposals

With respect to the categorial status, the empty noun in Persian was argued to be a head.

Three reasons were suggested for this view:

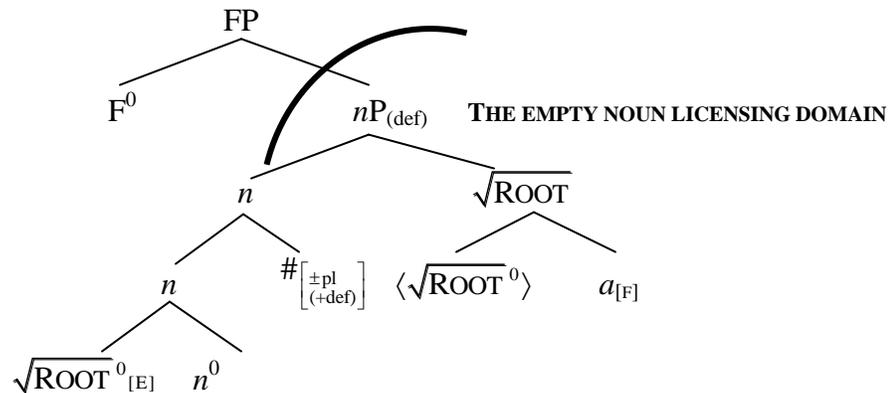
- (i) The Root of the empty noun and its attributive adjectival modifier(s) cannot take complements. For this reason, they are assumed, following Ghomeshi (1997), to be base-generated in a head-adjunction structure. As such, the noun and its attributive modifier(s) form the *nP* domain in which only heads can occur. I call this the *inner domain* of nominal modification.
- (ii) The empty noun allows for the standing of its attributive adjectival modifier(s).
- (iii) The non-pronunciation of the head noun along with its head modifier(s) is possible only if the modifier(s) undergo(es) incorporation into the nominal Root.

In this thesis, the analysis of the ENC in Persian has been proposed within a version of the PF-deletion approach which links ellipsis phenomenon to information-structural features, contrastive focus in particular (cf. §2.2). Within this approach, ellipsis/non-pronunciation applies at PF to redundant/given material under (i) semantic and syntactic identity with corresponding material in the antecedent clause; and (ii) contrastivity of the remnant with its antecedent. In my feature-based account of the ENC in Persian in Chapter 3, I have used Merchant's (2001) *E-feature* to specify the sister Root to the categorizing *n* head for non-pronunciation at syntax. This feature also ensures that the phonologically unexpressed material is in some sense identical with its corresponding antecedent at LF (cf. §3.6.1.1). Additionally, I have used the *F(ocus) feature* in the sense of Rooth (1992a, 1992b) and Schwarzschild (1999) to specify the remnant modifiers (i.e. the attributive adjectival modifier and certain pronominal modifiers) licensing empty nouns at syntax, and ensure that the contrastivity relationship between the remnant modifier and its counterpart in the antecedent clause is established at LF. Building on

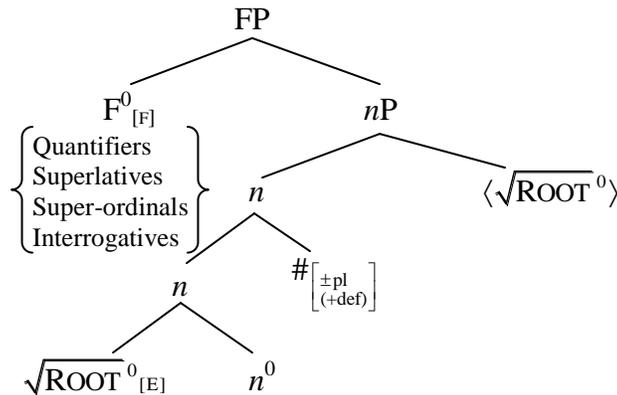
these assumptions, I proposed, along the lines of Gengel (2007), that the ENC in Persian is the outcome of the interaction between the E-feature and the F-feature.

The account of the interaction between the E- and F-features in the ENC was modeled within the framework of Distributed Morphology (DM) based on the notion of phases and cyclic spell-out. It was proposed that the licensing of the empty noun occurs as a result of the interaction between the information-structural features E and F within the same locality domain (cycle/phase). I have argued that the phase structure of the ENC is subject to variation depending on where in the derivation the licensing takes place. It was shown that in the ENC, *n* defines a locality boundary (phase) when the empty noun is licensed by an F-specified attributive adjectival modifier inside the *nP* domain. Whenever the empty noun is licensed in the domain of a higher head (namely, an inherently F-specified prenominal modifier), the *nP* and its higher constituent form a single phase. The phase structure of the ENC for these two possibilities of licensing the empty noun are shown schematically below (FP represents the projection that houses the prenominal licenser of the empty noun):

(1) a. *The ENC with the nP as a cycle/phase*



b. *The whole noun phrase as the empty noun licensing domain*



The above DM analysis of the ENC was shown to be adequate for accounting for the displacement of the plural marker in the ENC (cf. Chapter 4). The motivation for the movement of the plural marker onto its immediately adjacent empty noun licenser was suggested to be a morphophonological constraint requiring the plural marker to be spelled out on a phonologically realized head. It was argued that the constraint itself is motivated by the assumption that the plural marker in Persian is modificational rather than heading a functional projection. Adopting an articulated derivational model of PF (Embick & Noyer 2001, 2003; Pak 2008), I proposed that mobile plural marking in the ENC is handled by a PF operation called Local Dislocation. The rule applies early in the Vocabulary Insertion/Linearization process, adjoining the stranded plural marker to the empty noun licenser (i.e. either an attributive adjectival modifier in the *nP* domain or a prenominal head in the higher structure). In so doing, it applies in the same cycle in which the empty noun is licensed.

One important implication of the DM analysis of the empty noun and *-hâ* mobility is that it substantiates my proposal that postnominal modifiers in the Ezafe domain belong to two domains of modification: inner domain (*nP*), where the attributive (head)

modifiers (*n* and *a*) occur with the head noun, and outer domain (ModP/PossP), where phrasal modifiers (*a*P, PP, and *n*P_{poss}) appear. It was shown that the inability of phrasal modifiers to license the empty noun and host the stranded plural marker follows from this assumption.

The final issue addressed in this thesis was Ezafe insertion. Unlike licensing the empty noun and *-hâ* mobility which occur only in the inner domain, the application of Ezafe Insertion Rule (EIR) was shown as not being sensitive to the inner vs. outer domains of modification; it was argued that it applies within and across these domains. Adopting Pak's (2008) model of syntax-phonology interface, it was proposed that EIR applies at Late-Linearization, the point of the PF derivation at which all spell-out domains within a noun phrase are linearized. At this stage, the inner and outer domains are merged into a single domain for the Late-Linearization rule of EIR. We called the resulting domain the *modification domain*. The rule picks out the head noun, inserts Ezafe on it, and extends the Ezafe vowel onto all following [+N] M-Words except the last one. These [+N] M-Words must be categorially *n* or *a* and in modificational relationship with the head noun,. This procedure is schematically shown below:

(2) *Ezafe insertion at Late-Linearization*

i. Input structure:

PRENOMINAL DOMAIN PF_{np}(*i n n e r d o m a i n*) PF_{ModP/PossP}(*o u t e r d o m a i n*)

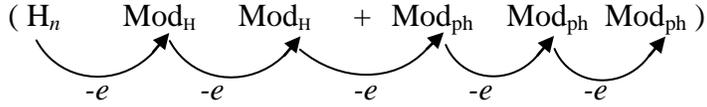
Internal structure: (H_n Mod_H Mod_H) (Mod_{ph} Mod_{ph} Mod_{ph})

ii. Chain-merging:

PRENOMINAL DOMAIN (*inner domain + outer domain = Ezafe insertion domain*)

Internal structure: (H_n Mod_H Mod_H + Mod_{ph} Mod_{ph} Mod_{ph})

iii. EIR:



It was discussed in detail that phrasal modifiers (Mod_{ph}) that occur in the Ezafe insertion/modification domain can themselves be comprised of a string of M-Words. Since these modifiers are headed by [+N] M-Words, an Ezafe insertion domain is formed phrase-internally in which the head is linked to the [+N] M-Words in its complement via Ezafe (cf. §5.3.3.2 for an account of Ezafe insertion inside the phrasal modifiers).

As for the absence of the Ezafe vowel on prenominal modifiers, I followed Matushansky (2008) in assuming that Ezafe only appears in the domain of modification; prenominal modifiers are in the domain of quantification and thus do not take this linker. It was also argued that Ezafe does not appear on prenominal modifiers because they are not [+N] M-Words of categories *n* and *a*. With respect to the appearance of Ezafe on the prenominal licenser of the empty noun, I proposed that the interaction between the E-feature specified on the nominalized Root and the inherent F-feature of the licensing modifier results in the modifier to inherit the [+N] feature of the empty noun. This feature inheritance causes widening of the *nP* domain to include the projection of the prenominal modifier licensing the empty noun.

In this thesis, I have presented EIR as a phonological rule whose application at Late-Linearization is sensitive to the (morpho)syntactic information of the constituents inside its domain. This provides further evidence in support of Pak's (2008) claim that

any type of syntactic information is visible to phonological rules at all levels of their applications in the PF derivation.

In the remaining of this chapter, we will discuss some potential areas for future research.

6.2. Future directions

In this section, we will consider two important potential areas of research that come out of this thesis, namely, the classification of adjectives that (dis)allow empty nouns and the internal structure of partitive noun phrases.

6.2.1 Classes of licensing adjectives in the ENC

We introduced attributive adjectives as the head modifier licensing empty nouns in all the examples of the ENC we used in this thesis. There are however adjectives in Persian that are head modifiers but do not license the empty noun. Consider the following example (non-licensing adjectives are in bold):

(3) [introducing some video clips]

... vali tosie=mi-kon-am didan-e in *kelip/kelip-e

but advice=DUR-do-1SG watching-EZ this clip/clip-EZ

jâleb / **zibâ** / **didan-i-(r)o** az=dast=na-d-in.

interesting/beautiful/spectacular-OM from=hand=NEG-give-2PL

‘...but I recommend you not to miss watching this interesting / beautiful /

spectacular [one=video clip].’

This observation calls for an investigation on the properties of the class of licensing adjectives in Persian ENC. According to Sleeman (1996), there is a strong cross-linguistic tendency for quantifiers to appear with empty nouns while the use of adjectives in ENC is subject to variation. Her findings show that some languages, e.g. English and Italian, are more restrictive and some, like most Germanic languages, are more permissive. She reports that, contrary to French and Sardinian, the licensing of empty nouns is generally impossible with adjectives of quality in languages like English, (some varieties of) Italian, and Spanish. These two groups of languages share the possibility of combining empty nouns with quantifiers and pronouns. Building on the above observation, Sleeman (1996: 151-152) suggests the following hierarchy of licensing elements in the ENC, with their position in the hierarchy related to the extent to which they are able to license empty nouns:

(4) *Hierarchy of licensing elements in the empty noun construction*

1. D-partitives (quantifiers, pronouns, superlatives, ordinals, ‘only’)
2. Color adjectives
3. ‘big’ - ‘small’
 ‘good’ - ‘bad’
 ‘old’ - ‘young, new’
4. All other adjectives of quality

In most of the languages considered in Sleeman’s study, empty nouns are only possible with elements of class one and sometimes class two. In French and Sardinian, empty nouns are also possible with adjectives of quality from classes three and (more marginally) four. While Sleeman states that she does not have an answer as to why this should be so, she suggests that there is a relation with the existence of the right-dislocated *de* + noun construction, which exists in both French and Sardinian and is not possible in Spanish and English.

(5) *French* la rouge, de voiture

Sardinian sa ruja de makkina

The red of car

‘the red car’

Accordingly, one can suggest that Persian can be grouped with languages like French and Sardinian in that the above possibility of right-dislocated nominal constructions also holds for Persian.

- (6) qermez-e in mâšîn
red-EZ this car
'this car in red (lit. red of this car)'

Although the findings of this thesis imply that Persian falls within the class of permissive kind of languages in allowing attributive adjectives to combine with empty nouns, further investigation is required to establish a more robust classification of adjectives that (dis)allow empty nouns. I leave this for future work.

6.2.2 Partitive constructions and empty nouns

One line of research to come out of this thesis is the area of partitive noun phrases. Partitivity, as a relation between a part or *subset* and a whole or *superset*, is expressed in two different constructions in Persian: (i) The subset is an inherently partitive element, viz. a quantifier, superlative, or numeral, that is connected to the partitive complement through Ezafe or the preposition *az* 'of, from' (7); (ii) The subset is an attributive adjective linked to a deictic noun phrase (superset) and the head noun is empty (8). I call the first type the *Q-partitive* and the second type the *N-partitive construction* (supersets are in curly brackets).

(7) *Q-partitive construction*

xeyli az / behtarin-e/ se-tâ az {bačče-hâ}
many of best-EZ three-CL of child-PL
'many/the best/three of the children'

(8) *N-partitive construction*

kət qermez-e {in kot}
eət red-EZ this coat
'this coat in red (lit. the red of this coat)'

Looking at the above examples of these two types of partitive constructions, the following questions arise:

- (i) Are these two types of partitive constructions generated as two distinct syntactic structures or they are derived from a single structure?
- (ii) Is the preposition *az* 'of, from' that connects the subset to the superset a linker like *Ezafe* or the head of a PP complement to the partitive elements in the subset?

Let us consider the first question. What distinguishes the two partitive constructions is that the N-partitive construction involves an empty noun while the Q-partitive construction does not. If this is correct, something has to be said about the different positions where the superset is generated in these two constructions. In the Q-partitive construction, the superset seems to be generated inside the *nP* domain while in the N-partitive construction, the superset is generated outside the *nP* domain.

There is also an alternative possibility of assuming that the Q-partitive construction involves empty nouns. As illustrated in the following naturally occurring examples of the Q-partitive construction, the noun representing the set is realized both in the subset and the superset. As a preliminary approximation, I take this as evidence that in Q-partitive constructions like in (7), there is an empty noun (partitive noun phrases are in []; subsets in (); and supersets in { }).

- (9) a. [(yek šâxe gol) az {gol-hâ}]-râ jodâ=kard-am va bâqi-râ
 one branch flower of flower-PL-OM separated=did-1SG and rest-OM
 be dast-e mâdar dâd-am.
 to hand-EZ mother gave-1SG
 ‘I separated a branch of flower from the flowers and handed the rest to the mother.’ <http://www.98ia.com/Forums-file-viewtopic-p-472328.html>

- b. ... va [(yek sarbâz) az {sarbâz-ân-e âšpazxâne}] va [(do sarbâz)
 and one soldier of soldier-PL-EZ kitchen and two soldier
 az {sarbâz-ân-e muzik}]-râ mi-koš-and.
 of soldier-PL-EZ music-OM DUR-kill-3PL
 ‘... and [they] kill one soldier from kitchen soldiers and two soldiers from music soldiers.’ <http://www.bultanneews.com/pages/?cid=7252>

Thus, it can be claimed that the two types of partitive constructions in Persian are derived from a single syntactic structure.

The second question concerns the role of the preposition *az* ‘of, from’ in partitive constructions. It appears that this preposition plays the same role as *Ezafe* in the context of partitive constructions: they both link subset to the superset. Specifically, I consider the preposition *az* in the context of partitive constructions a linker and thus a dummy element for the following reasons: (i) it does not have the semantics of a path preposition, and (ii) it does not head a PP, as it is dropped when the superset is pronominalized.

- | | | | | | | | |
|------|--------------------------|----|-------------|-------|--------------------------|---|------------------|
| (10) | se-tâ | az | ketâb-hâ-ye | hasan | <i>Pronominalization</i> | ⇒ | se-tâ-š |
| | three-CL | of | book-PL-EZ | Hasan | | | three-CLS-CL.3SG |
| | ‘three of Hasan’s books’ | | | | | | ‘three of them’ |

I leave further investigation on this issue and analysis of the syntactic structure of Persian partitive constructions for future work.

6.3. Concluding remarks

In this thesis, I have developed an analysis of the Empty Noun Construction in Persian on the basis of the PF-deletion approach to noun ellipsis. This approach is situated within the theory of Distributed Morphology based on the notions of phases and cyclic spell-out. It was shown that the analysis of the licensing of empty nouns in Persian has motivated several proposals on the internal structure of the noun phrase, plural marking, compounding, and *Ezafe* insertion, among others. These proposals were shown to be adequate both for capturing the empirical facts and generalizations emerging from the

data. Nevertheless, further investigation on other aspects of Persian language is needed to determine the plausibility of these proposals and suggestions.

References

- Abels, Klaus. 2003. Successive cyclicity, anti-locality, and adposition stranding. Ph.D. dissertation, University of Connecticut.
- Adger, David. 2003. *Core syntax: a Minimalist approach*. Oxford: Oxford University Press.
- Ahamdi-Givi, Hassan and Hassan Anvari. 1985. *Dastoor-e zabân-e Fârsi* [Persian language grammar]. Tehran: Fatemi Publishing.
- Alexiadou, Artemis. 2001. Adjective syntax and noun raising: word order asymmetries in the DP as the result of adjective distribution. *Studia Linguistica* 55.3: 217-248.
- Alexiadou, Artemis. 2004. Inflection class, gender and DP internal structure. In: Gereon Müller, Lutz Gunkel & Gisela Zifonun (eds.) *Explorations in Nominal Inflection*, 21-49. Berlin: Mouton de Gruyter.
- Alexiadou, Artemis and Kirsten Gengel. 2008. NP ellipsis without Focus movement: the role of Classifiers. Handout of a talk presented at the *Workshop on Interface-Based Approaches to Information Structure*, 13-15 September, University College London.
- Alexiadou, Artemis, Lilliane Haegeman, and Melita Stavrou. 2007. *Noun phrase in the generative perspective*. Berlin: Mouton de Gruyter.
- Baker, Mark. 1988. *Incorporation: a theory of grammatical function changing*. Chicago: University of Chicago Press.
- Barker, Chris. 2008. Possessives and relational nouns. Draft of an article for Maienborn, von Heusinger and Portner (eds.) *Semantics: An International Handbook by Natural Language Meaning*. Berlin: Mouton de Gruyter.

- Bateni, Mohammad Reza. 1969. *Towsif-e sâxtemân-e dasturi-ye zabâne fârsi* [Description of the grammatical structure of Farsi language]. Tehran, Iran: Amir Kabir Publishing.
- Beek, Janneke ter. 2008. *Restructuring and infinitival complements in Dutch*. Utrecht: LOT.
- Bouchard, Denis. 2002. *Adjectives, number and interfaces: why languages vary*. Oxford: Elsevier Science.
- Choi, Kwok Tim. 2006. Chinese SHI...DE focus constructions: an Optimality-Theoretic proposal. M.A. thesis, Simon Fraser University.
- Chomsky, Noam. 1970. Remarks on nominalization. In: Roderick A. Jacobs and Peter S. Rosenbaum (eds.) *Readings in English Transformational Grammar*, 184-221. Waltham, Mass.: Ginn.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, Massachusetts: MIT Press.
- Chomsky, Noam. 1999. Derivation by phase. *MIT Occasional Papers in Linguistics* 18.
- Chomsky, Noam. 2000. Minimalist inquiries: the framework. In: Roger Martin, David Michaels, and Juan Uriagereka (eds.) *Step by step: essays in honor of Howard Lasnik*, 89-155. Cambridge, Mass.: MIT Press.
- Cinque, Guglielmo. 1994. Evidence for partial N-movement in the Romance DP. *Italian Syntax and Universal Grammar*. Cambridge: Cambridge University Press, 287-309.
- Cinque, Guglielmo. 2005. Deriving Greenberg's Universal 20 and its exceptions. *Linguistic Inquiry* 36: 315-332.
- Corbett, Greville G. 2000. *Number*. Cambridge: Cambridge University Press.

- Corver, Norbert and Marjo van Koppen. 2005. Microvariation and ellipsis in the *wat voor*-construction. Handout of the talk presented at the workshop *Sounds of Silence: Empty Elements in Syntax and Phonology*, 19-22 October, Tilburg University, Netherlands.
- Den Dikken, Marcel. 2005. *Relators and linkers: The syntax of predication, predicate inversion and copulas*. LI Monograph 47. Cambridge: MIT Press.
- Den Dikken, Marcel. 2007. Phase Extension: Contours of a theory of the role of head movement in phrasal extraction. *Theoretical Linguistics* 33.1: 1-41.
- Den Dikken, Marcel and Pornsiri Singhapreecha. 2004. Complex noun phrases and linkers. *Syntax* 7.1: 1-54.
- Embick, David. 2003. Linearization and local dislocation: derivational mechanics and interactions. *Linguistic Analysis* 33: 303-336.
- Embick, David. 2007. Blocking effects and analytic/synthetic alternations. *Natural Language and Linguistic Theory* 25: 1-37.
- Embick, David. 2008. Localism versus globalism in morphology and phonology. Ms., University of Pennsylvania.
- Embick, David and Rolf Noyer. 2001. Movement operations after syntax. *Linguistic Inquiry* 32.4: 555-595.
- Embick, David and Rolf Noyer. 2006. Distributed Morphology and the Syntax/Morphology interface. In: G. Ramchand and C. Reiss (eds.) *The Oxford Handbook of Linguistic Interfaces*, 289-324. Oxford University Press.
- Embick, David and Alec Marantz. 2006. Architecture and blocking. Ms., University of Pennsylvania and MIT.

- Farshidvard, Khosro. 1969. *Dastoor-e emrooz* [Modern grammar]. Tehran: Vahidiyye Printing House.
- Fazlali, Fatemeh. 1994. Barrasi-ye hazf dar zabân-e Fârsi [A survey of ellipsis in Persian language]. Unpublished M.A. thesis, Allameh Tabataba'i University, Tehran, Iran.
- Gallego, Ángel and Juan Uriagereka. 2006. Sub-extraction from subjects. Paper presented at WCCFL 25 and LSRL 36.
- Gengel, Kirsten. 2007. Focus and ellipsis: a generative analysis of pseudogapping and other elliptical structures. Ph.D. dissertation, Institut für Linguistik/Anglistik der Universität Stuttgart.
- Ghaniabadi, Saeed. 2007. *-hâ* mobility and noun ellipsis. Ms., University of Manitoba.
- Ghaniabadi, Saeed. In preparation. Plural marking beyond count nouns. Ms., University of Manitoba.
- Gharib, Abdolazim; Bahar, Malekoshsho'ara; Foroozafar, Badiozaman; Homayi, Jalal; Yasami, Rashid. 1994. *Dastoor-e zaban-e Farsi (panj ostad)* [Persian language grammar (five professors)]. Tehran, Iran: Jahan-e Danesh Publishing.
- Ghomeshi, Jila. 1996. Projection and inflection: a study of Persian phrase structure. Ph.D. dissertation, University of Toronto.
- Ghomeshi, Jila. 1997. Non-projecting nouns and the Ezafe construction in Persian. *Natural Language and Linguistic Theory* 15.4: 729-788.
- Ghomeshi, Jila. 2003. Plural marking, indefiniteness, and the noun phrase. *Studia Linguistica* 57.2: 47-74.
- Ghomeshi, Jila. 2006. NP-anaphora and focused possessors in parallel architecture. *Linguistics* 44: 721-748.

- Ghomeshi, Jila. 2008. Markedness and bare nouns in Persian. In: Simin Karimi, Vida Samiian, and Donald Stilo (eds.) *Aspects of Iranian Linguistics*, 85-112. Cambridge Scholars Publishing.
- Giannakidou, Anastasia and Melita Stavrou. 1999. Nominalization and ellipsis in the Greek DP. *The Linguistic Review* 16.4: 97-132.
- Grimshaw, Jane. 1991. Extended projection. Ms., Brandeis University/Rutgers University.
- Grinder, John and Paul M. Postal. 1971. Missing antecedents. *Linguistic Inquiry* 2: 269-312.
- Gundel, Jeanette K. and Thorstein Fretheim. 2004. Topic and focus. In: Laurence R. Horn and Gregory Ward (eds.) *The Handbook of Pragmatics*, 175-196. Malden: Blackwell Publishing.
- Halle, Morris. 1997. Distributed Morphology: Impoverishment and Fission. In: Benjamin Bruening, Yoonjung Kang and Martha McGinnis (eds.) *MIT Working Papers in Linguistics 30: Papers at the Interface*, 425-449. Cambridge, Massachusetts: MITWPL.
- Halle, Morris and Alec Marantz. 1993. Distributed Morphology and the pieces of inflection. In: Hale, Kenneth and Samuel Jay Keyser (eds.) *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*. Cambridge, Massachusetts: MIT Press, 111-176.
- Halle, Morris and Alec Marantz. 1994. Some Key Features of Distributed Morphology. *MIT Working Papers in Linguistics: Papers on Phonology and Morphology* 21, 275-288.

- Harley, Heidi. 2005. *One-replacement, unaccusativity, acategorial roots and Bare Phrase Structure*. In: Slava Gorbachov and Andrew Nevins (eds.) *Harvard Working Papers in Linguistics*, Volume 11.
- Harley, Heidi. 2008. Compounding in Distributed Morphology. [lingBuzz/000602].
- Harley, Heidi and Rolf Noyer. 1999. State-of-the-art: Distributed Morphology. *Glott International* 4.4: 3-9.
- Hankamer, Jorg and Ivan Sag. 1976. Deep and surface anaphora. *Linguistic Inquiry* VII: 391-426.
- Huddleston, Rodney and Geoffrey Pullum. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.
- Inkelas, Sharon and Draga Zec. 1995. Syntax-phonology interface. In: John A. Goldsmith (ed.) *The handbook of phonological theory*. Cambridge: Blackwell, 535-549.
- Jackendoff, Ray. 1996. The architecture of the linguistic-spatial interface. In: Bloom, Paul, Mary A. Peterson, Lynn Nadel, and Merrill F. Garrett *Language and Space*, 1-30. Cambridge, MA: MIT Press.
- Kahnemuyipour, Arsalan. 2000. Persian Ezafe construction revisited: evidence for Modifier Phrase. In: Jensen, John T and Gerard van Herk (eds.) *Cahiers Linguistique d'Ottawa, proceedings of the 2000 annual conference of the Canadian Linguistic Association*, 173-185.
- Kahnemuyipour, Arsalan. 2006. Persian Ezafe construction: case, agreement or something else. In the *Proceedings of the 2nd Workshop on Persian Language and Computer*, 3-16. Tehran University.

- Kalbasi, Iran. 1992. *Saxt-e eshteghaghi-ye vazhe dær Farsi-ye emruz* [Word derivational structure in Modern Persian]. Tehran: Mo'assesse-ye Motale'at va Tahghihat-e Farhangi (Pazhuheshgah).
- Karimi, Simin and Michael Brame. 1986. A generalization concerning the EZAFE construction in Persian. Ms., University of Washington and University of Arizona; presented at the annual Western Conference of Linguistics, Canada.
- Kavousi Nejad, Soheila. 1992. *Hazf dar zaban-e Farsi* [Ellipsis in Persian language]. Unpublished M.A. thesis, Tehran University, Iran.
- Kavousi Nejad, Soheila. 1998. Hazf dar goruh-e esmi-ye zaban-e Farsi [Ellipsis in Persian noun phrases]. *Name-ye Farhangestan* 16.4: 109-127.
- Kester, Ellen-Petra. 1996. *The nature of adjectival inflection*. OTS, Utrecht.
- Khanlari, Parviz N. 1990. *Dastur-e zabân-e Fârsi [The grammar of Persian]*. Tehran: Toos.
- Kiss, Katalin É. 1998. Identificational focus versus information focus. *Language* 74.2: 245-273.
- Kornfeld, Laura M. and Andrés L. Saab. 2004. Nominal ellipsis and morphological structure in Spanish. In: Reineke Bok-Bennema; Bart Hollebrandse; Brigitte Kampers-Manhe; and Petra Sleeman (eds.) *Romance Languages and Linguistic Theory 2002*, 183-198.
- Lambrecht, Knud. 1994. *Information structure and sentence form: Topic, focus, and the mental representation of discourse referents*. Cambridge: Cambridge University Press.

- Larson, Richard and Naoko Takahashi. 2007. Order and interpretation in prenominal relative clauses. In: Meltem Kelepir and Belkiz Öztürk (eds.) *Proceedings of the 2nd Workshop on Altaic Formal Linguistics (WAFLL 2)*. MITWPL.
- Larson, Richard and Hiroko Yamakido. 2005. Ezafe and the position of nominal modifiers. Ms. Stony Brook University; presented at the LSA.
- Liu, Danqing. 2008. Syntax of space across Chinese dialects: conspiring and competing principles and factors. In: Xu, Dan (ed.) *Space in Languages of China: Cross-linguistic, Synchronic and Diachronic Perspectives*, 39-67. Springer Sciences + Business Media B.V.
- Lobeck, Anne. 1991. Phrase structure of ellipsis in English. In: Susan Rothstein (ed.) *Perspectives on Phrase Structure: Syntax and Semantics* 25, 81-103. San Diego: Academic Press.
- Lobeck, Anne. 1992. Licensing and identification of ellipted categories in English. In: Steve Berman and Arild Hestvik (eds.) *Proceedings of the Stuttgart Ellipsis Workshop* [Arbeitspapiere des Sonderforschungsbereichs 340, Bericht Nr. 29], IBM Germany, Heidelberg, 1-55..
- Lobeck, Anne. 1995. *Ellipsis: functional heads, licensing, and identification*. Oxford: Oxford University Press.
- Lyons, Christopher. 1999. *Definiteness*. Cambridge: Cambridge University Press.
- Mahootian, Shahrzad. 1997. *Persian*. London: Routledge.
- Marantz, Alec. 1996. "Cat" as a phrasal idiom: consequences of late insertion in Distributed Morphology. Ms., MIT, Cambridge, Massachusetts.

- Marantz, Alec. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own Lexicon. In: Alexis Dimitriadis, Laura Siegel, Clarissa Surek-Clark, and Alexander Williams (eds.) *University of Pennsylvania Working Papers in Linguistics 4.2: Proceedings of the 21st Annual Penn Linguistics Colloquium*, 201-225. Philadelphia: University of Pennsylvania.
- Marantz, Alec. 2001. Words and things. Ms., MIT.
- Marantz, Alec. 2007. Phases and words. In: Sook-Hee Choe (ed.) *Phases in the theory of grammar*, 191-222. Seoul: Dong In.
- Megerdooonian, Karine. 2008. Low-density language strategies for Persian and Armenian. Ms., The MITR Corporation.
- Matushansky, Ora. 2008. On the attributive nature of superlatives. *Syntax* 11.1: 26-90.
- Merchant, Jason. 2001. *The syntax of silence: Sluicing, islands, and the theory of ellipsis*. Oxford: Oxford University Press.
- Merchant, Jason. 2004. Fragments and ellipsis. *Linguistics and Philosophy* 27: 661-738.
- Munakata, Takashi. 2009. The division of C-I and the nature of the input, multiple transfer, and phases. In: Grohmann, Kleanthes K. (ed.) *InterPhases: phase-theoretic investigations of linguistic interfaces*, 48-81. Oxford University Press.
- Nespor, Marina and Irene Vogel. 1986. *Prosodic phonology*. Dordrecht: Foris.
- Nespor, Marina and Irene Vogel. 2007. *Prosodic phonology*. The Hague : Walter de Gruyter.
- Ntelitheos, Dimitrios. 2004. Syntax of elliptical and discontinuous nominals. M.A. thesis, University of California, Los Angeles.

- Ortmann, Albert. 2000. Where plural refuses to agree: feature unification and morphological economy. *Acta Linguistica Hungarica* 47: 249-288.
- Pak, Marjorie. 2008. The postsyntactic derivation and its phonological reflexes. Ph.D. dissertation, University of Pennsylvania.
- Panagiotidis, Phoevos. 2003. Empty nouns. *Natural Language & Linguistic Theory* 21: 381-432.
- Pantcheva, Marina. 2007. The place of PLACE in Persian. Ms., CASTL, University of Tromsø.
- Philip, Joy. 2009. Restrictions on the morphosyntactic marking of dependencies: what linkers tell us. Ms., University College London. Available from *lingBuzz*, [<http://ling.auf.net/lingBuzz/000878>], accessed 10/25/2009.
- Poot, Antonio González and Martha McGinnis. 2005. Local versus long-distance fission in Distributed Morphology. In: Claire Gurski (ed.) *Proceedings of the 2005 Canadian Linguistics Association Annual Conference*, 20 pages.
- Prince, Allan and Paul Smolensky. 1993. Optimality Theory: constraint interaction in generative grammar. Ms., New Brunswick: Rutgers University and Boulder: University of Colorado.
- Ross, John Robert. 1969. Guess who? In: R. Binnick, A. Davison, G. Green, and J. Morgan (eds.) *Papers from the 5th Regional Meeting of the Chicago Linguistic Society* 252-286. CLS, ChicagoIll.24.
- Rooth, Mats. 1992a. A theory of focus interpretation. *Natural Language Semantics* 1: 75-116.

- Rooth, Mats. 1992b. Ellipsis redundancy and reduction redundancy. In: Steve Berman and Arild Hestvik (eds.) *Proceeding of the Stuttgart Ellipsis Workshop, Report Series Sprachtheoretische Grundlagen fur die Computerlinguistik, no 29, SFB 340*. IBM Heidelberg.
- Sadat-Tehrani, Nima. 2007. The intonational grammar of Persian. Ph.D. dissertation, University of Manitoba.
- Sadeghi, Ali Ashraf. 1970. Horuf-e ezâfe dar Fârsi-ye Mo'âser [Prepositions in contemporary Persian]. *Journal of Tabriz College of Literature and Humanities* 95 and 96: 441-470.
- Samiian, Vida. 1983. Origins of phrasal categories in Persian, an X-bar analysis. Ph.D. dissertation, UCLA.
- Samiian, Vida. 1994. The Ezafe construction: some implications for the theory of X-bar syntax. In: Medi Marashi (ed.) *Persian Studies in North America*, 17-41. Bethesda, Maryland: Iranbooks.
- Samvelian, Pollet. 2006. When morphology does better than syntax: the Ezafe construction in Persian. Ms., Université de Paris 3 – Sorbonne Nouvelle.
- Schwarzschild, Roger. 1999. GIVENness, Avoid F and other constraints on the placement of focus. *Natural Language Semantics*, 7.2 : 141-177.
- Sedighi, Anousha. 2005. Subject-predicate agreement restrictions. Ph.D. dissertation, University of Ottawa.
- Selkirk, Elisabeth. 1986. On derived domains in sentence phonology. *Phonology Yearbook* 3: 371-405.

- Sigurðsson, Halldór Ármann. 1993. The structure of the Icelandic NP. *Studia Linguistica* 47.2: 177-197.
- Skinner, Tobin R. 2009. Investigation of downward movement. Ph.D. dissertation, McGill University.
- Sleeman, Petra. 1996. *Licensing empty nouns in French*. Onderzoeksinstituut voor Taal en Spraak, Utrecht.
- Svenonius, Peter. 2004. On the edge. In : David Adger, Cécile de Cat, and George Tsoulas (eds.) *Peripheries : Syntactic Edges and their Effects*, 259-287.
- Svenonius, Peter. 2006. The emergence of Axial Parts. In: Svenonius, Peter (ed.) *Nordlyd: Tromsø Working Papers in Linguistics* 33.1: 49-77. CASTL, Tromsø. Available at <http://www.ub.uit.no/munin/nordlyd/>.
- Szendrői, Kriszta. 2001. Focus and the syntax-phonology interface. Ph.D. dissertation, University College London.
- Williams, Edwin. 1997. Blocking and anaphora. *Linguistic Inquiry* 28: 577-628.
- Wiltschko, Martina. 2008. The syntax of non-inflectional plural marking. *Natural Language and Linguistic Theory* 26: 639-694.
- Winkler, Susanne. 2006. Ellipsis. In: Keith Brown (ed.) *The Encyclopedia of Language and Linguistics*, 2nd edition. Oxford: Elsevier.
- Zhu, Yao. 1997. The focus-marking function of *Shi* in Mandarin Chinese. Ph.D. dissertation, University of Minnesota.